Business Process Analysis of Yield Data Flow at a Newly Merged Pharmaceutical Company

by

Nan JIANG

Bachelor of Engineering in Aeronautical Engineering
Bachelor of Economics
Xiamen University, 2009

SUBMITTED TO THE DEPARTMENT OF MECHANICAL ENGINEERING IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ENGINEERING IN MANUFACTURING
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SEPTEMBER 2010

Signature of Author.....
Department of Mechanical Engineering
Aug 17, 2010

Certified by.....
Stanley B. Gershwin
Senior Research Scientist, Department of Mechanical Engineering
Thesis Supervisor

Accepted by.....
David E. Hardt
Ralph E. and Eloise F. Cross Professor of Mechanical Engineering
Chairman, Committee on Graduate Students

Business Process Analysis of Yield Data Flow at a Newly Merged Pharmaceutical Company

by

Nan JIANG

Submitted to the Department of Mechanical Engineering on August 17, 2010 in Partial Fulfillment of the Requirements for the Degree of Master of Engineering in Manufacturing

ABSTRACT

Tracking, monitoring, and documentation of the manufacturing performance are significant for pharmaceutical companies under the regulations of Food and Drug Administration (FDA). However, the current yield data are not consistent and the business procedures for yield data flow are not unified and optimized at the newly merged pharmaceutical company SJP Singapore. Therefore, a systematic analysis of the current yield data processing was performed in four facilities located at two campuses of this company.

Through this analysis, the current business procedures were visualized; the various yield concepts were clearly defined; the problems involved and their root causes were identified; the potential solutions were proposed; and a standard business process was developed taking into account the situation of each facility. As a result, the inconsistent yield data were harmonized and procedures in all four facilities were unified and standardized. This revised business process would enable each department to better fulfill its responsibilities and drive decisions on future improvements.

Keywords: business process, yield, data flow, pharmaceutical, API, postponement

Thesis Supervisor: Stanley B. Gershwin
Title: Senior Research Scientist of Mechanical Engineering

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my thesis advisor, Prof. Stanley B. Gershwin for his guidance and his keen insight into the problem. His probing questions inspired me to uncover hidden opportunities in this project.

Next, I would like to appreciate Dr. Brian Anthony and Ms. Jennifer Craig for keeping track of the project progress and giving constructive advice to the thesis writing.

I would also like to thank the internship company for sponsoring this project. Special thanks go to Mr. Paul D'Souza and Dr. Nicholas Almeida for their continuous support and encouragement throughout the period of internship. In addition, I would extend my gratitude to all the employees I have interacted with for their timely help.

My deepest thanks go to my family, for their love, encouragement and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Project Description

This thesis reports on yield data flow improvement at Company SJP Singapore. A systematic business process represented as a workflow chart, which defines the specific roles, responsibilities, and process standards of producing products was developed for this pharmaceutical company to continuously monitor and manage the production performance of all the in-line products measured by yield. To achieve this goal, the business process is expected to streamline the yield data flow and provide the visualization of yield status of the various products being manufactured in the company for the management team to decide on future improvements. Moreover, this process should align the data source, reduce or justify inter-department data inconsistencies and redundancies, and establish a standard data flow system.

To represent the different production nature of both pharmaceutical products and active pharmaceutical ingredient (API) products, and also to include the different data system environments at both south and west campuses, the current yield data flow at the two campuses of this company were tracked so as to identify the potential causes to the current inconsistencies in yield data. Four production lines were selected to be explored in sequence for detailed countermeasures, which were product "T", API "Eto", product "Z" and product "Na". Product "T" is a pharmaceutical product being manufactured at the south campus while API "Eto" is a chemical product being manufactured in the API facility at the south campus. Similarly, product "Z" and product “Na” are two pharmaceutical products of different types being manufactured at the west campus. These production lines were chosen because the processes involved are representative to cover more potential issues in the data flow. The products manufactured in the API facility at the west campus are generalized into a common production line; therefore no specific product needs to be selected.

1.2. Company Background

Located at Singapore Tuas Biomedical Park, SJP Singapore is a subsidiary wholly owned by a global pharmaceutical company Superior & Co., headquartered in USA. Superior is research-driven, and develops and manufactures a variety of products to improve human and animal health.

1.2.1. Characteristics of the pharmaceutical industry: FDA restricted and capital intensive

Since the pharmaceutical industry is strictly regulated by company regulatory bodies such as U.S. Food and Drug Administration (FDA), there are increasing needs of tracking, controlling and optimizing processes for companies who want to avoid FDA penalties. Furthermore, the operational efficiencies and plant performance can be greatly improved by process analytical technology and the concept of "Quality by Design". In furthering the objectives of "Quality by Design", FDA plays an instrumental role in ensuring the alignment of the many innovative applications, including the rapid corrective and preventive action (CAPA) that assures compliance of current good manufacturing practices (CGMP) maintaining process control and product quality [1].

Apart from the restrictions by FDA, it is the market characteristics of the pharmaceutical industry that drives the organizations to strive for quality excellence. Generally speaking, it takes a pharmaceutical company one billion US dollars to develop a new drug [2]. Therefore, any seemingly insignificant improvement would be of considerable value to a pharmaceutical company. Figure 1 is a graph demonstrating the feature of pharmaceutical industry as compared to Hi-Tech industry [2].

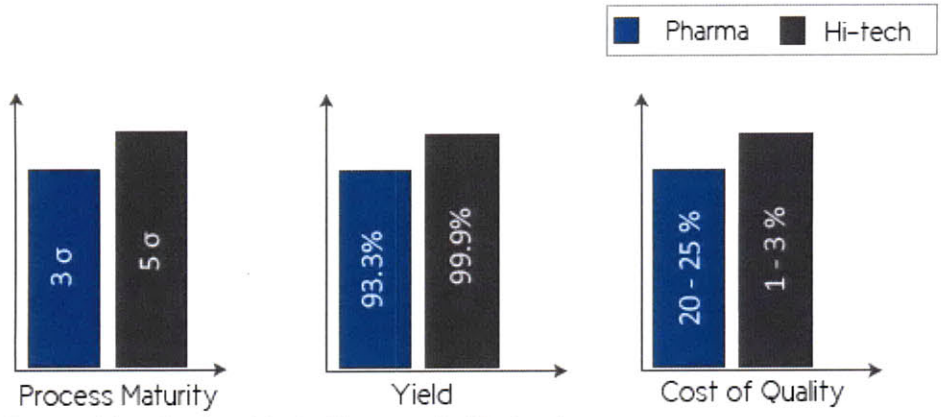

Figure 1: Comparison of Pharma and Hi-tech Industries [2]

1.2.2. Status in the industry: pharmaceutical industry market environment-more concerns on productivity.

Previously, being one of the largest pharmaceutical companies in the world both in terms of revenue and market capitalization, Superior has every reason to secure its position in this steadily growing pharmaceutical market. According to IMS Health, the global pharmaceutical market will grow 5-8% annually through the next five years and the patient demand will remain robust [3]. Since it becomes more difficult to develop new blockbuster products to seize the market, and the patents of the old blockbusters are expiring, it is crucial for the companies to maximize the productivity of in-line products and the operational efficiency.

1.2.3. Recent merger requires a common standard procedure

Superior merged with another company in 2009. As a global leader, Superior is able to supply a broader range of products after the merger and has been on the track of

completing integration actions to achieve the target of $3.5 billion annual savings in 2012 [4].

The manufacturing facility at SJP Singapore is comprised of two campuses. The south campus is the former SJP Singapore before the merger and the west campus refers to the manufacturing facility of the acquired company. Both campuses in Singapore are capable of the production of both API and pharmaceutical drugs, including the medicines in the cholesterol management, HIV treatment and respiratory areas.

However, it is also because the two campuses used to belong to two independent companies that they have entirely different organizations, management systems, work flow and databases, which make a standardized business process essential.

1.3. Project Motivation

Driven by the regulations of FDA, most pharmaceutical companies have the incentives of employing a plethora of management systems and databases to support tracking, monitoring, and documentation of the production performance including release quantities and yield percentages. In spite of this, the assumption that the summarized historical yield data is accurate and effective may not stand given the invisible and unclear data flow chain. Comparing to the solid system of Standard Operating Procedures (SOP) that covers every single step of the manufacturing processes and the data collection methods, the further yield data processing and reporting procedures are not well defined probably due to different organizational structures in different companies, or even in different branches of the same company.

Being one of the manufacturing subsidiaries of a global leading pharmaceutical company, SJP Singapore is concerned about the productivity which contributes to its survival in the competitive market. The funding of a new drug could potentially be

saved by 1-2 % of yield improvements in the pharmaceutical industry [2]. Hence, a transparent optimized yield data flow process is of great value to the company. Like most of the other companies, the previous improvements focused more on the production itself, and apart from the continuous improvement engineers, all of the other departments have been striving to achieve better performances as well, neglecting to explore information streaming. Therefore, not only is a business process that standardizes the whole yield data flow chain filling in the gap so as to support the data accuracy, reliability and promptness, but also providing the management teams an effective tool to make judgments and decisions so as to improve the system responsiveness. Moreover, if applicable, this business process can be universalized and applied to other manufacturing environments, for instance, production lines, pharmaceutical plants, or even other industries, which proves the external and long term interests of the business process development.

In the special case of Company Superior, the business process for yield data flow performs a significant function as part of the integration actions. Since each company has their own database and management systems, the yield data flow certainly will not automatically merge to form a uniform report. Therefore, it is necessary to establish a common business process so that the new company is all aligned when it comes to the yield information. As the yield data is reviewed by a series of departments from the technicians who operate the equipment to the Integrated Process Team (IPT) Leader, it is crucial that the data they obtain deliver the same information so that all these vital experts are working toward the same direction. In addition, the yield improvement projects typically require the collaboration of multiple departments, sharing a single consolidated data platform would help facilitate the cooperation and avoid potential errors, format mismatches, and communicational problems during data transfer.

Once a systematic business process is established to streamline the yield data flow in a pharmaceutical company, the regulatory compliance can be increased at a lower cost. And a visible yield data source and summarized yield metrics report would be a

more reliable and effective signal enabling the managers and production leaders to identify potential opportunities of improvements.

1.4. Goals

To meet the needs of SJP Singapore, the business process should be able to provide the solutions to the aforementioned problems. Considering the special case of SJP Singapore, a series of requirements are summarized as follows.

First, this business process should help ensure the yield data accuracy. Since the yield data diverge mainly at two nodes in the current flow chain, the raw data source and the calculation of the yield percentage along the chain, this business process should not only unify the raw data source and avoid errors from the root, but also reduce the data inconsistencies downstream along the flow. The latter further requires root cause identification.

Secondly, this business process should improve efficiency. In this business process, the departments involved in yield calculation should be minimized so as to reduce redundant work. Apart from the number of departments, the priority of the yield data access should be well defined in the data flow. For instance, to improve the promptness of responses, the department which is most closely related to the process operating should be among the first to obtain the yield information so as to fix the problem effectively. The feedback loop should be shortened if possible. The efficiency can be measured in terms of labor hours and labor costs.

Thirdly, this business process should justify the reliability of the yield data calculated and submitted to the managers. Unlike other yield and discard projects aiming at improving a certain production line, the current status of yield data flow is not visible and not even depicted in details. As it is more difficult to quantify and evaluate the

accuracy, efficiency and effectiveness of a data flow, this business process should enable the visualization of the yield data flow and further streamline the current flow.

Furthermore, the business process should be general. Developing a business process itself is an investment to be rewarded. To maximize the value of this process, it should be applicable to most of the in-line products in the company.

Finally, the business process structure should be able to continuously support the yield data flow in future operations. Particularly, this business process should be relatively robust to external changes and minimize the effect of heuristics and other subjective factors. For instance, the collection of data should not be dependent on any experienced engineers or better performing departments; instead, it should be a standard procedure system insensitive to people or performance.

1.5. Thesis Organization

This thesis is organized in ten chapters. Chapter 1 is the introduction of the whole project, including a brief description of the project, the background of Company SJP Singapore, the motivation for the company to launch this project, and the goals that the company expects to achieve. Chapter 2 is a more detailed problem statement. The current problems identified are stated in Section 2.1, while the contributions of this project in solving these problems are listed in Section 2.2. In Chapter 3, the relevant methodologies, and information about the data systems at SJP Singapore are provided with a theoretical background.

The business process development starts from Chapter 4, which is the methodology of how this project is performed to complete the procedure establishment. In Chapter 5, the various pharmaceutical and API processes and products are introduced and compared to justify the selection and generalization of the representative production lines to be studied in the following chapters. Subsequently, Chapter 6 and Chapter 7

provide a thorough analysis of the current business process of yield data flow at the south and west campuses, respectively. In these two chapters, the organizational structures of yield data processing at the two campuses are investigated; followed by specific definitions of the yield concepts, yield data flow tracking, customer analysis, problems and recommendations at each facility.

After the current business process and root causes of the problems are made visible by the four yield data flow diagrams and detailed yield calculations in batch sheets, a new business process is proposed in Chapter 8. In this chapter, the characteristics of the yield data flow in each facility are compared, followed by a gap analysis by summarizing the problems and setting requirements for the new business process. The new business process is developed based on these specific requirements.

Chapter 9 is the conclusion of this project. The future work is listed in Chapter 10.

2. PROBLEM STATEMENT

2.1. Current Problem Identification

The current problems observed are listed in this section. These problems can be categorized into five types, which are data inconsistencies, invisible data flow, data processing redundancies, data flow inefficiencies, and lacking of standard procedure.

2.1.1. Data Inconsistencies

1) Multiple Yield Definitions

The yield data of a same batch of the same product submitted by different departments are observed to be different. The gap in between the numbers is not significantly large. However, this phenomenon persists. It is attributed to the multiple definitions being used by different departments in different facilities. However, they are not clearly defined and differentiated in the various documents, reports, and even conversations. Therefore, not only do yields diverge from the root by definition, but probably also along the data transmission path by people's perception.

2) Multiple Data Sources for Yield Calculation

There are a series of documents and data systems recording the various yield data, including the shop floor batch sheets, batch release report, Data 3 System, SAP System, Distributed Control System (DCS), and Plant Information (PI) System. However, the raw data sources are not explicit, which renders the yield data unwarranted.

2.1.2. Invisible Data Flow

Other than the raw data sources, the whole yield data flow is invisible. The structure inside the yield data flow box is not visible to the outside, with the only observable information being the input from the shop floor and the output of yield metrics. The complexity of the processing of yield data among a plethora of departments is far beyond the imagination of a linear transmission. Moreover, since each department is only responsible for a small portion in this flow, a big picture of the whole data transmission process is yet to be discovered.

2.1.3. Data Processing Redundancies

1) Multiple Department Calculations

From the observation of the various outputs of yield data, there are at least three departments calculating or exporting the yield data, including the continuous improvement (CI) engineers, process engineers, and Finance. They collect the yield information from different data sources and use it for different purpose. Nevertheless, if they eventually derive the same yield data, there should be opportunity to reduce the redundancy; on the contrary, if they eventually derive different yield data, then the inconsistency need to be analyzed and either justified or eliminated.

2) Only One Type of Yield Is Related to Problem-Solving

Even if all the departments have good reasons to calculate the different yield data and set different targets accordingly, only one type of yield can be technically improved and only the engineers are capable of this improvement. Other departments such as Finance are using them as operation metrics; setting targets cannot help solve the problems unless it is connected to the production in some manners.

2.1.4. Data Flow Inefficiencies

The yield data, as other parameters depicting the manufacturing performance, is measured to capture the atypicals and drive necessary improvements. However, the current yield data flow is not problem-solving oriented, namely, clearly defined path for prompt response to yield problems is not complete and the departments in charge of data processing and problem solving are partially parallel.

2.1.5. No Standard Procedure Exists

1) Two Campuses Differences

As mentioned in the introduction, the SJP Singapore is a combination of the south campus and west campus, which were two independent companies a few months ago. They have different data systems and organizational structures, not to mention the yield data flow process. Since there is no standard procedure for yield data flow, the data among the various products of the same type may not be compared effectively.

2) Two Facilities Differences

Even before a further analysis, the yield data flow has already been noticed to be different due to the different manufacturing natures. Most of the API processes are close-coupled, which is not the case for the pharmaceutical processes. A close-coupled process means the materials can flow inside the pipelines and containers without exposing to human factors. This kind of process can be fully controlled by DCS and the data is therefore stored in PI without manual calculation. The pharmaceutical processes are modular designed to enable multipurpose production. Therefore, after manually transferred, the yield needs to be recorded at each stage, which is also manual calculation. These intrinsic characteristics explain why a common procedure does not exist yet.

2.2. Main Contributions

Considering the current problems analyzed above, the following objectives are achieved by the business process developed in this work:

1) To avoid confusion and unnecessary errors, all the yield concepts being used in SJP Singapore will be defined with formulae given if there is a specific calculation of it. The current users of each yield concept will be identified and summarized.

2) The yield data will be tracked from the raw data source to the top level recipients of this information. In addition, a data flow diagram will be plotted in each facility at the two campuses of the whole company to provide a transparent look into the black box. These diagrams are prerequisite of the further analysis, problem and opportunity identification, and improvements to make in the business process.

3) After the data flow diagrams are plotted, the data inconsistencies, the workflow redundancies, and inefficiencies should all be identified. Possible solutions will be discussed and recommended.

4) Moreover, a standard business process which can be applicable to all of the four facilities in this company will be developed based on the analysis.

3. THEORETICAL BACKGROUND

In this chapter, the basic methodologies used in this project are provided with theoretical background. Section 3.1 explains the principles of the DMAIC approach and it application in this project. Section 3.2 introduces the concept of SIPOC and clarifies its definition particularly in this thesis. Section 3.3 reviews the previous approaches to identify opportunities for business process improvements and to establish a new business process. The methods and logic for plotting a data flow diagram and the steps for performing gap analysis are briefly introduced in Section 3.4 and Section 3.5, respectively. Finally, the five different data management systems and their relations are described in Section 3.6.

3.1. DMAIC Approach

Define-Measure-Analyze-Improve-Control (DMAIC) is a problem-solving approach broadly used in business as a component of Lean Six Sigma methodology. The framework and implementation of this method were introduced explicitly and concisely by George, et al. [5]. Nunnally provided a detailed explanation of the application of DMAIC and Six Sigma in the pharmaceutical industry on controlling variation [6]. More emphasis was put on scoping in the define phase by Lynch, et al. [7]. SJP Singapore is applying this concept in diverse projects execution following the five basic steps:

1) Define the problem and launch the project, validate the scope and goals, collect voice of customer;
2) Measure the critical aspects of the current process, articulate the process through process mapping or value stream mapping, develop and execute a specific measurement plan, evaluate the performance of the process based on the measurement;

3) Analyze the inputs and outputs of the process, identify and investigate the root causes;

4) Improve the process based on the analysis, propose, prioritize and pilot the potential solutions;

5) Control the process after the implementation is launched, monitor and correct the deviation from targets, validate performance and document in standard operating procedures (SOP).

In this project, the first four phases have been completed with a business process established. The detailed steps are listed in Table 1. The implementation of this business process requires both administrative and operational arrangements which are out of the scope and beyond the time span of this project.

Table 1: DMAIC Process in This Project

Phase	Tasks
Define	Understand the current status, define the problem and validate goals.
Measure	• Define the different definitions of yield; • Select the representative product, track the yield data and plot the data flow diagram; • Identify the current and actual customers and perform gap analysis; • Compare the data input and output throughout the data flow.
Analysis	• Analyze the data inconsistencies and find the root cause; • Identify the redundancies and inefficiencies along the data flow and propose potential solutions;
Improve	Develop a standard business process to streamline the yield data flow.
Control	Not included in this thesis.

3.2. SIPOC

Supplier-Input-Process-Output-Customer (SIPOC) is a high level diagram of a process to help define the customers of the project so as to satisfy the needs. As Basu [8] pointed out, it can be applied during either the define phase or the measure phase with its impact upon the entire project life cycle. This methodology consists of five elements:

1) Supplier: the provider of the input to the process;
2) Input: whatever required for the process, including labor, materials or information as in this project;
3) Process: the transformation steps of the input to output;
4) Output: the concerned product of the process, which is the reported yield data in this project;
5) Customer: the receiver of the delivered product.

The definitions of customer are different in Voice of Customer (VOC) and SIPOC. The customer in VOC is the customer of the project who is influence by it; while the customer in SIPOC refers to the customer of the process who receives the product of the process. In the context of this project, the latter customer refers to the recipient of the yield data.

The concept of this method is adopted throughout the whole project. Instead of a traditional SIPOC Diagram plotted completely at the beginning of the project, the suppliers and customers are identified during the yield data tracking. The data flow diagrams demonstrate the yield data transmission process via the current suppliers and customers. Apart from the "S", "C" and "P" contained in the same diagrams, the input and output of different formats of yield data are analyzed in the data comparison sections.

3.3. Business Process

Business Process is a term that embodies the related activities or responsibilities of producing products, which can be service, information or goods for customers. It is often represented as a workflow chart, which defines the specific roles, responsibilities, process standards, etc.

Kemsley provided a series of reasons why business process should be improved and how opportunities can be discovered [9]. It was a general introduction of the business process in all industries, but is brief and to the point. Aalst and Hee eloquently pointed out the disadvantages of traditional process structuring that not enough attention was paid to process structure within the framework of traditional systems making the business process difficult to recognize and conceived incorrectly or incompletely [10]. From the perspective of information technology, they systematically discussed about the workflow management from the organizational structure to specific business process development and re-engineering. A variety of methods explained can be applied to the pharmaceutical industry as well. For instance, the tier process introduced later in this thesis is one of the hierarchical organizational structures described in this book [10].

3.4. Data Flow Diagram

A data flow diagram is a chart to visualize the processing of data. There are diversified sets of symbols, logic, and model organizations, which is elaborated in the paper by Li and Chen [11].

3.5. Gap Analysis

Gap analysis is a concept describing the approach to the optimized allocation by comparing the current status and the desired status extensively used in business process development, business direction and information technology. Jeston and Nelis introduced the steps of implementing gap analysis in business process management [12] while Bolstorff and Rosenbaum described the application of gap analysis in supply chain management [13].

3.6. DCS, PI, Data 3, SAP and JDE

Distributed Control System (DCS) is a system structure that allows the distributed controllers to control each component sub-system along the entire manufacturing system. The production processes at SJP Singapore are managed by DCS while DCS is constituted by network of sub-systems of Programmable Logic Controller (PLC). The data collected by DCS is stored in Plant Information (PI) System and ready for export [14].

Data 3 is used for material management at the south campus of SJP Singapore. All the production and transfer quantities are recorded in this system. Different from DCS, the data and material status in Data 3 is input manually. Moreover, the yield percentages are not documented in Data 3 as the quantities. At the west campus, SAP is the counterpart business management system of Data 3.

JDE is used as financial management software at the south campus to calculate yield variance and related costs.

4. METHODOLOGY

The methodology of DMAIC approach is adapted to the discovery nature of the business process development. The detailed steps of the investigation of the current yield data flow and the development of a standard business process is plotted in Figure 2. The dotted blocks demonstrate the distribution of contents in Chapter 5 to Chapter 10. As can be observed from the graph, the selection of representative production lines from each facility is discussed in Chapter 5, which is in the dotted block on top. The two dotted blocks below explain the detailed analysis of yield definitions and yield data flow at the two campuses respectively in Chapter 6 and Chapter 7. The upper right dotted block is the summary of the current status at the whole company by comparing the yield data flow and customers in each facility. The standard business process is established to satisfy the requirements derived from the analysis of the problems in all of the four facilities. These are the content in Chapter 8. Though this thesis is organized by different campuses, according to the definition of DMAIC, the upper half of the study in each facility until data flow tracking belongs to the "Define" and "Measure" phases, and the lower half and Chapter 8 all belong to the "Analyze" and "Improve" phases. The lower right is a conclusion of the whole project in Chapter 9. Finally the potential expansions of this business process and further exploration are listed in Chapter 10 as future work.

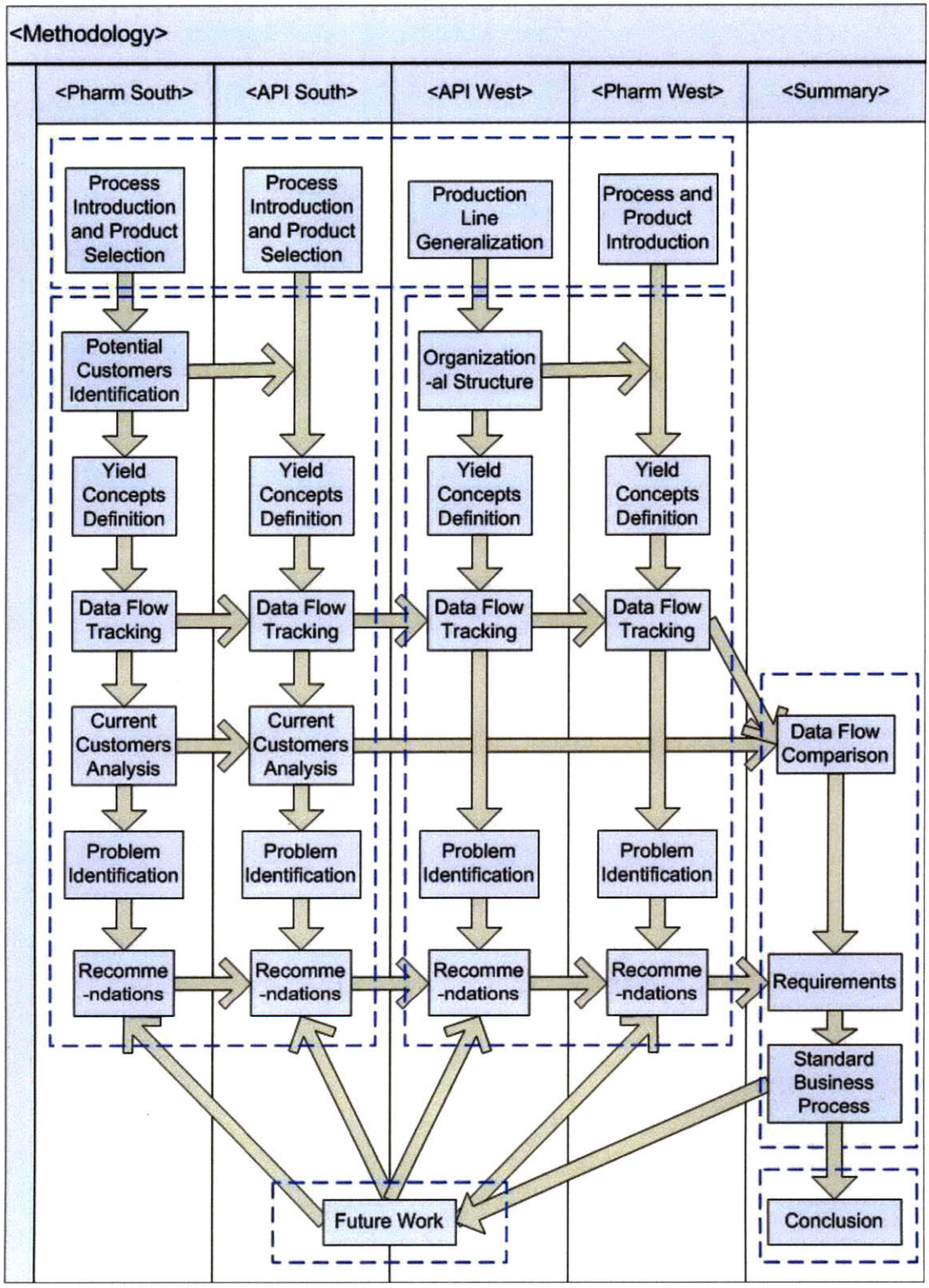

Figure 2: Business Process Development for Yield Data Flow Methodology

4.1. Define and Measure

The visible problems have been defined in the problem statement; however, the root causes for these problems are to be discovered during the process of plotting diagrams of the current data flow. To collect the voice of customer (VOC) of this project, the SIPOC method is partially incorporated by identifying the suppliers and customers of the yield data. However, unlike most of the other problems, the customer identification cannot be completed before the measure phase because the actual suppliers and customers are unknown before the data is tracked through the related departments.

The measurement phase of this project is therefore beyond pure data. Since there is a variety of products being manufactured in this company at two campuses, several representative products should be selected for further analysis. Considering the strong resemblance between the manufacturing processes and data flow process of each product in the same facility at the south campus, it is a reasonable and practical strategy to choose one sample production line which is representative to include more potential problems during the yield data transmission from each facility and then generalize to all products in that facility. At the west campus, on one hand, no specific product is selected from the API facility because the API production lines can be generalized into a common one; on the other hand, the two products of the pharmaceutical facility are all studied because they involve completely different manufacturing processes thus yield calculation might be different.

After the products have been selected, to prepare for the data tracking, the tier process, which is the organizational structure at the south campus, is investigated to understand the structure of the workflow and the potential departments involved in the yield data processing. Similarly, the organizational structure of the west campus is also introduced although the tier process has not been fully established at this campus. As shown in Figure 2, the various concepts of yield are then clearly defined so that

the objects to track are clarified. The yield concepts at the west campus are defined referring to the south campus.

Subsequently, by interviewing the engineers from the potential customer departments and verifying the specific format of the yield data, the current yield data flow diagrams are plotted.

4.2. Analyze and Improve

The current suppliers and customers, as well as the various purposes of using the yield data are identified and discussed after the data flow tracking. Moreover, by comparing the data and calculations acquired from different departments along the yield data flow chain, the root causes for data inconsistencies are investigated and discussed. During this process, the work redundancies and inefficiencies are also analyzed.

Targeting at the problems identified in each facility, potential solutions are proposed accordingly. These recommendations are also the requirements for the business process to be established to fill the gap between the current yield data flow and the desired yield data flow. Therefore, with a comparison of the current yield data flow in each facility and a summary of the requirements, a standard business process is established finally.

The potential expansion of this business process and the achievement of other opportunities would be the future work for this company.

5. PRODUCTION LINE SELECTION

This chapter describes the selection of representative production lines to be considered in this research. The processes and products being manufactured at the south campus are introduced and compared, and a representative production line is therefore selected from each of the pharmaceutical and API facility in Section 5.1 and 5.2 respectively. In Section 5.3, a generalized production line of the API facility at the west campus is described, followed by a supplementary introduction of pharmaceutical processes and product information obtained from the pharmaceutical facility at the west campus.

5.1. Product "T" at Pharm South

Since there is a variety of pharmaceutical products being manufactured at the south campus, a representative production line, product "T", was selected for further analysis. To justify the selection, the pharmaceutical processes are introduced in Section 5.1.1 and the pharmaceutical products being manufactured at the south campus are introduced in Section 5.1.2. Finally, one typical product is selected by process comparison in Section 5.1.3.

5.1.1. Pharmaceutical Processes Introduction

1) Dispensing and Charging:

In this process, raw materials are pre-weighed manually in the dispensing booth and then fed into a container through a charge chute using a charge hopper.

2) High Shear Granulation:

This process is a type of wet granulation in which powder particles adhere to form larger granules by adding liquid solution. Generally speaking, granulation is carried out for better flow and compression characteristics. After the wet granulation process, the wet granules are pumped to a fluidized bed for drying and the solvent is removed.

3) Roller Compaction:

This process is also called dry granulation in which multi-particle entities agglomerate without a liquid solution being added. Specifically, at the south campus of SJP Singapore, the powder is squeezed between two cantilevered rolls to form a ribbon. Subsequently, the ribbon is broken up by passing through a series of other related equipment. Roller compaction is often used due to its simpler process and lower costs, or when the product to be granulated is sensitive to moisture.

4) Lubrication:

This step is used to prevent the granules from sticking to the equipment during tableting by adding lubricants such as Magnesium Stearate.

5) Direct Compression:

This is a type of tableting process that is simpler and cheaper than wet granulation and dry granulation. In this process, a mixture of dry powders is filled into a tablet hopper and directly compressed into tablets.

6) Milling:

Instead of increasing particle size in the granulation process, the milling process aims at reducing and homogenizing the particle size and normalizing the distribution of particle sizes as well.

7) Blending:

The purpose of this process is to mix the APIs with various excipients. It is a critical process to ensure the content uniformity of the dosage units and to reduce the variability in the manufacturing processes.

8) Tablet Compression:

In this process, the granules are placed into tablet dies and punched to form tablets. Different from the direct compression which presses the dry ingredient powders, tablet compression here refers to the compression of granules.

9) Film Coating:

This process is required to hide the unpleasant taste of the tablets, to make large tablets smoother to swallow, and to control the dissolution of drugs in the gastrointestinal tract. For instance, some medicines may be affected by the time it takes to reach the small intestine where they are absorbed.

5.1.2. In-line Product Introduction

There are three products being manufactured in the pharmaceutical facility at the south campus of SJP Singapore, denoted as product "V", product "T", and product "I".

Product "V" is a medicine containing the combination of two APIs, namely API "Eze" and API "S" [15]. According to the different ratios of the two APIs, one batch of final products can be divided into four strengths. Nevertheless, since the manufacturing processes of these four strengths are identical, this factor would not be differentiated during the analysis of yield data flow. Product "V" has been the highest-volume and most profitable product manufactured for many years and will

still be in line for approximately another four years given the patent expiration date in 2014 [16].

Product "T" is a drug for treating dyslipidemia and primary hypercholesterolaemia with a combination of two APIs, which are API "L" and API "N". These two ingredients are mixed together by the tableting process called "Bi-layer Compression". Comparing to product "V", product "T" is younger and will be in line for an even longer period of time.

Product "I" is a new product manufactured for only one year after its approval in 2007 by FDA [17]. It is a medicine containing the API "R" used for the treatment of HIV.

5.1.3. Product Selection by Process Comparison

A comparison of the processes involved in the production of each product is drawn in Figure 3, and product "T" is selected as a typical production line for further analysis and data tracking. Not only because this production line is representative to cover more manufacturing processes thus more steps for yield calculation and to include more potential problems during the data flow, but also because product "T" is a younger product than the highest volume product "V".

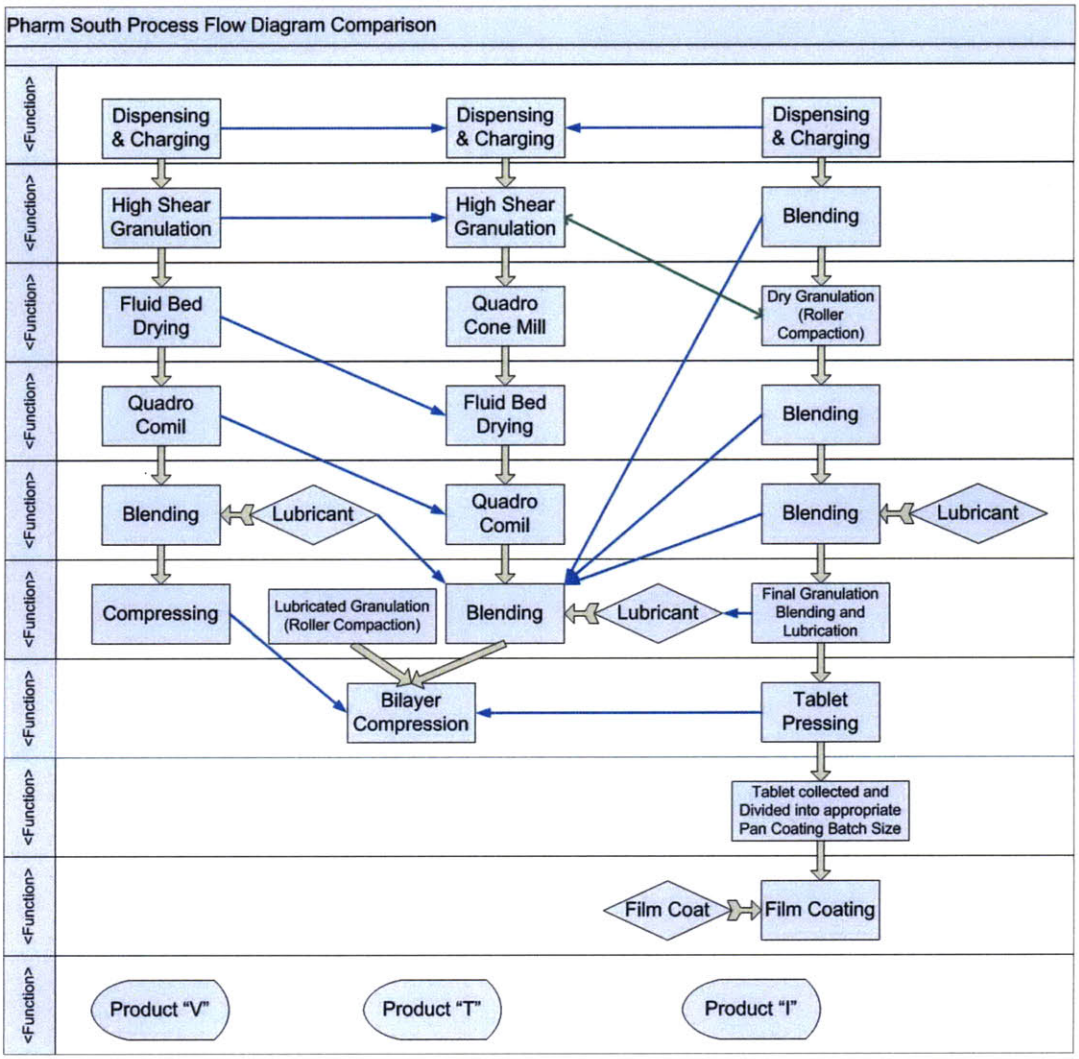

Figure 3: Pharm South Process Flow Diagram Comparison

As can be observed from the three process flow diagrams in Figure 3, all the processes involved in the production of product "V" are necessary steps for product "T", denoted by blue arrows. Most of the processes involved in the production of product "I" are covered by the product "T" production as well, except for roller compaction and film coating, denoted by the green double arrow. Product "I" is currently the only product at the south campus with roller compaction involved in its manufacturing processes. As explained previously, the dry granulation is considered as a more advanced technology than the wet granulation and consequently will be the future trend for granulation. However, in spite of the quality advantages, this process

might not significantly differ from the wet granulation in terms of yield data flow. Besides, the calculation of yield at the film coating stage is similar to the blending stage in terms of the material flow, the film coating is not necessary to be taken into account. Hence tracking product "T" will represent the current procedure in the whole facility.

5.2. API "Eto" at API South

There are mainly two API Products being manufactured at the south campus, API "Eto" and API "S". Given the characteristics of the production, the two intermediates of API "S" can be considered as independent production lines. Therefore, the selection of a typical production line should be justified. The API processes involved are described in Section 5.2.1 and the API products being manufactured at the south campus are introduced in Section 5.2.2. Finally, the processes involved in the production of each product are compared so as to select the representative API "Eto" in Section 5.2.3.

5.2.1. API Processes Introduction

1) Dissolution:

This is a process during which solid or liquid substances mix homogeneously in a solvent with the crystal lattice breaking down into individual molecules, ions, etc.

2) Informal Aging

The purpose of this process is to balance the cycle time like a buffer.

3) Extraction:

This is a method used to extract a substance from one liquid solvent to another based on relatively better solubility. The two solvents are immiscible liquids placed in one container.

4) Carbon Treatment:

This is a color reduction process that filters out the solid particles.

5) Distillation:

This is a unit process of separating substances based on different volatilities by boiling the liquid mixture.

6) Crystallization:

This process is to precipitate solid crystals from a solution.

7) Centrifugation:

It is a process to separate mixtures according to the densities by centrifugal force. Besides, it is also called as filtration.

8) Filter Drying:

This is a process with combination of filtration and drying completed in one container.

9) Drying:

This is simply to remove the moisture of the solvent, leaving the dry crystals in the container.

5.2.2. In-Line Product Introduction

In total two API products are being manufactured in the API facility at the south campus, which are API "Eto" and API "S". API "S" was mentioned in the introduction of product "V" as one of its ingredients. And API "Eto" is the ingredient of Product "A" for the treatment of chronic low back pain, acute pain, arthritis, gout, and ankylosing spondylitis.

Although there are only two products in this facility, there are four production lines involved in the manufacturing of them, of which two are for intermediates of API "S" and the other two are for final products.

5.2.3. Product Selection by Process Comparison

API “Eto” is selected for further analysis because the processes involved in its manufacturing not only represents the processes of most of the other products, but due to the semi-continuous characteristics, they also enables the application of DCS as well as the data transmission to PI system from where the yield data can be exported. Figure 4 is a comparison of the processes involved in the production of all the final products and intermediates. "EMSB" refers to API "Eto" Modified Salt Break, which is the complete manufacturing process for API "Eto", while "SC", "SP", "SMF" are the abbreviations of API "S" "Crude", API "S" "Pure", and API "S" "Milled Fine", respectively.

Although SMF is a final product, the process is not as typical as an API manufacturing process. It is more of a physical process because no chemical reaction takes place during this and the purpose of this step is just to reduce the particle size and blend the powders. To choose a production line that can represent more processes and to include more potential problems during the data flow, SMF is not desirable.

As can be observed from the parallel comparison, the SP process resembles the other two remaining processes, also without chemical reaction. It is a purification of the crude substances rather than an API manufacturing process. Hence SP is also a suboptimal choice.

Since the selection is simplified by the exclusion of two options, the comparison between EMSB and SC becomes rather obvious. The extraction, carbon treatment, and distillation processes evidently demonstrate that EMSB covers more possibilities for problem identification.

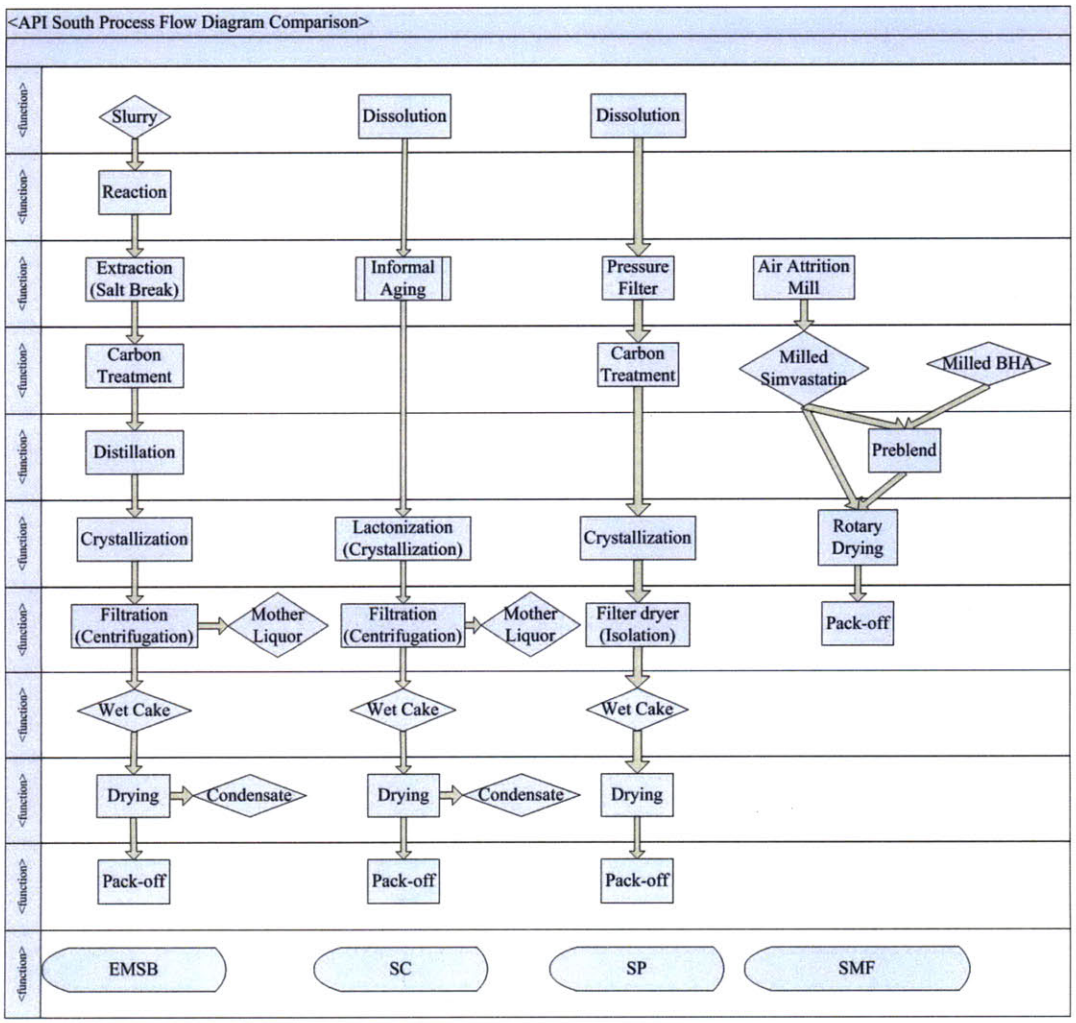

Figure 4: API South Process Flow Diagram Comparison

5.3. Production Line Generalization at API West

The various processes involved in the manufacturing of API products at the west campus can be generalized into one production line due to the similarities, and the generalized diagram is plotted in Figure 5 as compared to the processes in the API facility at the south campus.

As can be observed from Figure 5, most of the processes at the west campus are the same as in the south campus, while the only additional process milling is a physical process that does not introduce more yield calculation problem. And since the yield data flow in this facility is not controlled by DCS, it is not differentiated by the product manufacturing process.

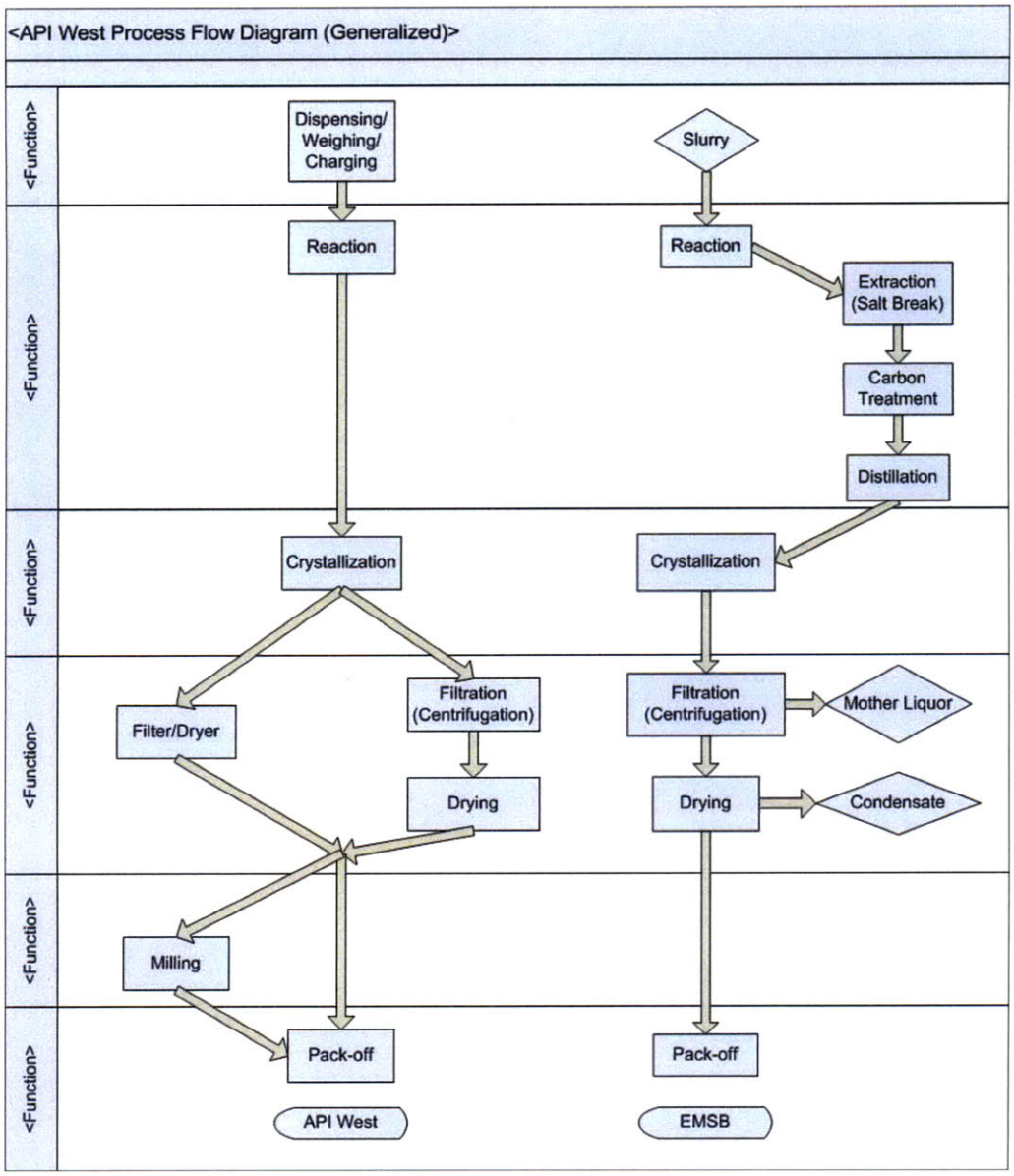

Figure 5: API West Process Flow Diagram (Generalized)

5.4. Product Introduction and Process Description at Pharm West

Due to the completely different product characteristics and time limit, the biotech facility at the west campus is not considered here, thus the pharmaceutical facility

only refers to the tablet facility. There are two products being manufactured in this facility at the west campus, product "Z" and product "Na". They are different types of pharmaceutical product. Product "Z" is tablets containing only one API "E", which is the same as one of the ingredients of product "V" manufactured in the pharmaceutical facility at the south campus. Product "Na" is a prescription nasal allergy spray involving only one step of manufacturing process [18].

The various processes for both products are listed in Figure 6.

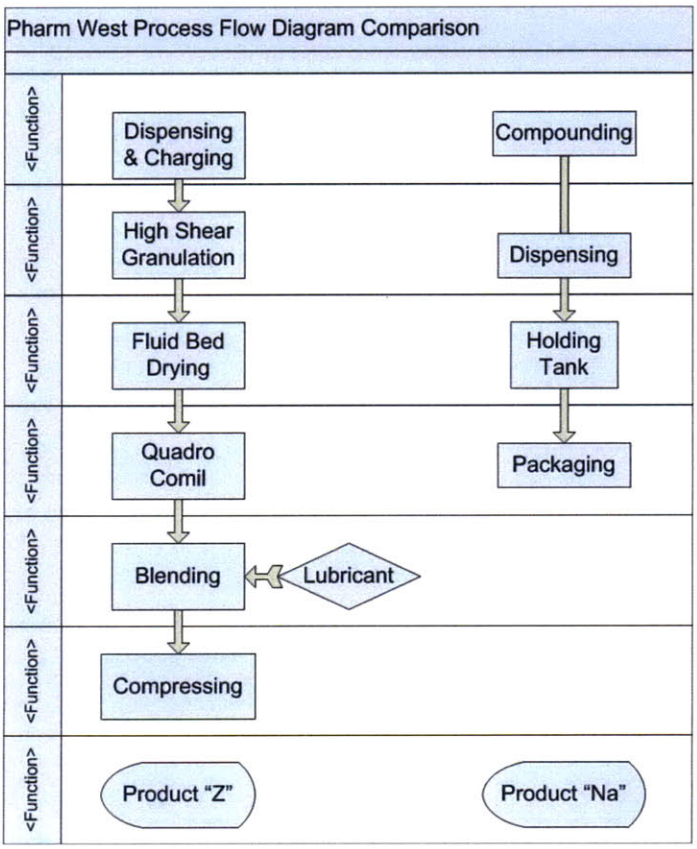

Figure 6: Pharm West Process Flow Diagram Comparison

As indicated by the diagram, these two products are totally not comparable. However, the processes involved in the production of Product "Z" are completely the same as the product "V" at the south campus, which are similar to the product "T". Therefore,

the analysis of the yield calculation of product "Z" will provide a comparison of the two campuses.

The only process for product "Na" requiring yield calculation is compounding, which basically means the mixture of all the ingredients in liquid. The involvement of yield calculation at this distinctive step will supplement the previous analysis toward pharmaceutical products at a relatively low cost of effort.

Consequently, the two products are all investigated in this facility.

6. CURRENT YIELD DATA FLOW ANALYSIS AT THE SOUTH CAMPUS

In this chapter, the various manufacturing processes and the organizational structure of the current yield data processing at the south campus are given a detailed description. Section 6.1 introduces the tier process at the south campus and identifies the potential customers of the yield data. The various current yield concepts in the pharmaceutical facility at the south campus (pharm south) are listed and defined clearly in Section 6.2. Section 6.3 tracks the yield data of the selected product "T" through each single step and department and discusses the current customers and their responsibilities at pharm south. The problems with the yield data calculation and processing in this facility are identified in Section 6.4 while Section 6.5 proposes potential solutions to these problems. Similarly, the yield concepts in the API facility at the south campus (API south) are defined in Section 6.6. Section 6.7 tracks the yield data of the selected API "Eto" through the departments and analyzes the current customers. The problems in this facility are identified in Section 6.8 with the potential solutions provided in Section 6.9.

6.1. Potential Customers Identification from Tier Process

Following the DMAIC Approach, the customers in this project are defined as whoever collects, calculates, proceeds, uses, or reports the yield data since all might benefit from this project by saving labor hours or adding value to the current activities. To track the data through the potential customers, the organizational structure at the south campus is introduced in Section 6.1.1, followed by the potential customer identification in Section 6.1.2.

6.1.1. Tier Process

The tier process is both a cultural mind set and a set of standard methods to visualize the current problems, operational performance, and accountabilities, it also enables quick identification and resolution of root causes [19].

In this thesis, the tier process is considered more as a data review process than a human resources allocation. It describes the hierarchy of the information flow.

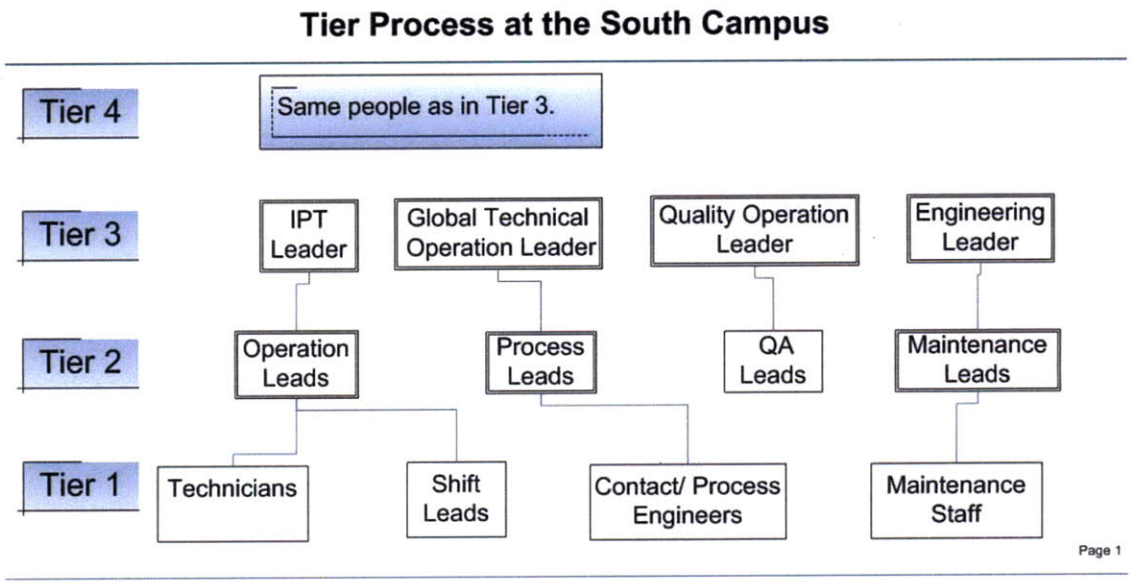

Figure 7: Tier Process at the South Campus

As shown in Figure 7, tier 1 is the base level of the tier process. The data is collected on a shift basis and can be considered as approximate real-time update. Currently at the south campus of SJP Singapore, this tier level consists of the technicians who operate the equipment, the shift leads who are in charge of shift operations, the contact engineers and process engineers who work directly on the manufacturing processes, and the maintenance staff who are responsible for mechanical issues. They review daily shop metrics and communicate issues and countermeasures. The data span of tier 1 is eight to twelve hours depending on the shift frequency.

Tier 2 is a level higher than tier 1 in the tier process. The data is obtained on a daily basis during normal office hours. At the south campus, tier 2 generally consists of the supervisors of the people in tier 1. As depicted in the diagram by the grey lines, technicians and shift leads report to the operation leads on the manufacturing operations; contact and process engineers report to the process leads on the process performance; maintenance staff report to the maintenance leads; and particularly, also involved in this daily level are quality assessment (QA) leads, who are in charge of a series of document review activities related to quality, such as batch sheet review, batch release, and annual product review (APR). The data span of tier 2 is twenty-four hours and the response actions are taken within the office hours roughly from 8:00 am to 4:00 pm.

Tier 3 is a level higher than tier 2 in the tier process. The data is reviewed weekly or daily as well but less frequently than tier 2. Instead of constantly being processed and monitored during the entire office hours, tier 3 data such as the trends in metrics are checked once in a day, at 10:30 am for instance. Similarly to tier 2, tier 3 consists of the supervisors of the people in tier 2. The operation leads report to the IPT Leader, who is managing the manufacturing of all the products. At the south campus, specifically, both the pharmaceutical facility and the API facility are currently under the supervision of the same IPT Leader. Accordingly, the process leads of both facilities report to the GTO Leader on the process performance and improvements; the QA leads report to the Quality Operation (QO) Leader on the quality related activities; and the maintenance leads report to the Engineering Leader on the maintenance progresses. The data span of tier 3 is one day or the past one week and the response time is usually longer than tier 2 since these leaders prioritize continuous improvement efforts and settle higher-level problems.

Except for the top level tier 5, which comprised of the monthly reviews by the plant manager, the highest tier level closely related to the data flow would be tier 4. It is undertaken by the same group of people as in tier 3; whereas, the data is tabulated and summarized into metrics both weekly and monthly depending on which department it

serves to. The response to the abnormalities found in this tier level would be investigated on a project basis.

6.1.2. Potential Customers

It can be assumed that all the departments except for the maintenance branch in the tier process are concerned about yield data, because yield percentage is an index to assess the performance of both the operation and the processes. In addition, two more departments closely related to the yield data processing are continuous improvement (CI) engineers and Finance. CI engineers generate metrics of yield and other parameters for the IPT Leader to charter improvements while Finance calculates the yield variance and the costs of not meeting the yield standards.

Theoretically, the capacity planning department should also be considered as a customer due to the impact of yield on production capacity. However, the yield deviation is too small in the pharmaceutical industry for the capacity planning to take into account. Hence the capacity planning can be assumed to be unaffected by the yield data flow.

6.2. Yield Definitions at Pharm South

Yield is a parameter used to measure the performance and capacity of production in almost all industries. However, since it is defined in multiple ways even in the same pharmaceutical company between different facilities, it may cause confusion, miscommunication, and data inconsistency thus losing its effectiveness in demonstrating the real production performance. For instance, yield may refer to the absolute yield quantity in the unit of kilogram, tablet or mole; it may also refer to the fractional yield of the output to the input, or the actual output to the theoretical output. Furthermore, there are different ways of measuring it, such as release yield and

accountable yield. To avoid these confusions, all the yield terminologies currently being used at SJP Singapore are listed and defined as follows.

1) Release Yield

The release yield is the actual release weight of a process such as granulation and blending. It is measured in kilograms. At the south campus, it is only used on the shop floor in the pharmaceutical facility and is documented in batch sheets by technicians. In this thesis, this concept is denoted as y_{Re} for the whole production line and the initial of the process name is used before the "Re" to distinguish the various stages. For instance, y_{GRe} refers to the release yield from granulation.

2) Theoretical Yield

The theoretical yield is the ideal case of the output weight of one or a series of processes. In the pharmaceutical facility, since there is no chemical reaction, the theoretical yield equals to the input weight expressed in kilograms. For a single stage, this input weight equals to the sum of the upstream release weight and the material weight added in this stage. For the whole production line comprising a series of processes, this input weight equals to the total input weight along all the processes. To be noted here is that, the input weight refers to the net inputs. The liquid added in the processes and removed completely in the end is not considered as the input weight. The theoretical yield is denoted as y_{TH}, and similar to the notation of release yield, the initial of the process name is used to distinguish the process, such as y_{GTH}.

3) % Release Yield

At the south campus, this concept is used in the pharmaceutical facility only and is recorded in shop floor batch sheets and CI engineers' spreadsheets. It is defined as the yield percentage of the release yield over the theoretical yield of that single stage and denoted as Y_{Re}. The weights are all in kilograms. For instance, the blending process is

performed after granulation, the theoretical yield or the input weight of the blending process is the sum of the release yield from granulation and the weight of material added. Therefore,

$$y_{BTH} = y_{GRe} + x_B \quad (1)$$

$$Y_{BRe} = \frac{y_{BRe}}{y_{BTH}} = \frac{y_{BRe}}{y_{GRe}+x_B} \times 100\% \quad (2)$$

In which, x_B denotes the additional materials added in the blending stage, which are the lubricants particularly in this thesis.

4) Accountable Yield

Let l_S and l_R denote the sampled weight and reject weight respectively. These two types of losses are added to the release yield to calculate the accountable yield, y_{Ac}. For instance, the accountable yield for blending is derived in Equation (3).

$$y_{BAc} = y_{BRe} + l_{BS} + l_{BR} \quad (3)$$

At the south campus, this term only appears in shop floor batch sheets in the pharmaceutical facility documented by technicians. In this thesis, this concept is denoted as y_{Ac} for the whole production line and the notation for single processes follow the same rule explained previously.

5) % Accountable Yield

The % accountable yield is defined as the yield percentage of the accountable yield over the theoretical yield and denoted as Y_{Ac}. The weights are in kilograms. Still take the blending process as an example,

$$Y_{BAc} = \frac{y_{BAc}}{y_{BTH}} = \frac{y_{BRe}+l_{BS}+l_{BR}}{y_{GRe}+x_B} \times 100\% \quad (4)$$

At the south campus, this concept is used in shop floor batch sheets and only the % accountable yield for the whole production line Y_{Ac} is recorded in the CI engineers' spreadsheets.

6) Batch Release Yield

The batch release yield is exactly the same value as the total % accountable yield Y_{Ac} in batch sheets. However, the name of this value is changed to batch release yield in the CI engineers' metrics. At the south campus, this number is not only used in shop floor batch sheets and CI engineers' spreadsheets, but also submitted to the IPT Leader and Global Technical Operation (GTO) Leader for decision making. Therefore, the calculation of this value is of high importance. Since the calculation of this concept requires the data of all the processes of a particular product and involves potential problems during the derivation, it is discussed in details in Section 6.4.1 after data tracking.

7) Standard Yield/Standard Lot Size

The standard yield used by Finance is also called standard lot size in Data 3 system. There is no specific formula for this concept because it is determined based on theoretical yield data and historical trends jointly by GTO and IPT Leaders. This number is given in the unit of thousand tablets for pharmaceutical products at the south campus.

8) (Financial) Yield

The actual yield that Finance refers to is named financial yield in this thesis because it is exported from the JDE system which extracts the data from Data 3 system. Theoretically, the value for this concept is the number of tablets transformed from the

total release yield of a batch y_{Re} using the weight per tablet. Therefore, the financial yield is expressed in the unit of thousand tablets.

Since all the yield data mentioned in the previous definitions are collected before the Quality Assessment (QA) review, the release yield refers to the actual output quantities but not necessarily the quantities approved to be released. However, the release quantities used by Finance come from the Data 3 which also processes the batch release information provided by QA. Hence, whether they are using the approved quantities or exactly the same quantities that the technicians input to the Data 3 system depends on the interfaces between Data 3 and JDE. Since most of the batches can be released eventually and the current staffs are mostly end users, this information is not available and this term is currently not distinguished yet.

9) Yield Variance

The yield variance is defined as the difference between the financial yield and the standard yield measured in thousand tablets or dollars. It is real-time calculated in the JDE system and exported monthly by Finance when they close the transactions in Data 3 system every second last business day. The dollar form yield variance is derived by multiplying the unit standard cost. This data is included in the monthly management report that Finance feedback to all the senior managers including IPT Leader, GTO Leader and Quality Operation (QO) Leader.

6.3. Data Flow Tracking at Pharm South

Based on the previous customer identification, the yield data of product "T" is tracked in the pharmaceutical facility of the south campus through all the departments from data collection to final report, which is defined as the "Process" in SIPOC. A Yield Data Flow Diagram is plotted in Section 6.3.1 with a detailed description of the role of each department. In addition, the data format at each stage is provided with a clear

explanation in this section. Through this process, the current customers are confirmed and discussed in Section 6.3.2.

6.3.1. Data Flow Diagram

A Yield Data Flow Diagram is plotted in Figure 8 to depict the current yield data flow chain.

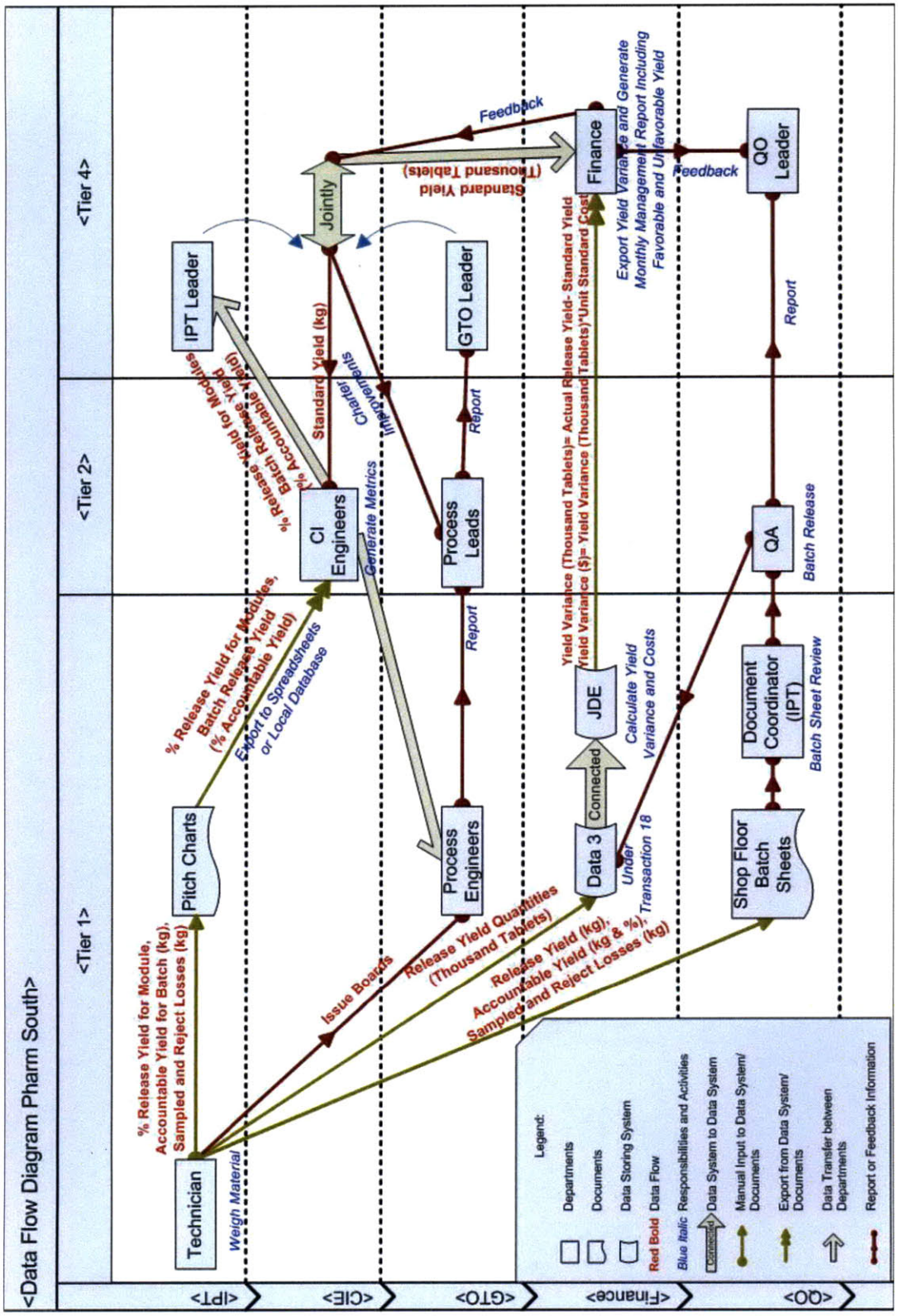

Figure 8: Data Flow Diagram Pharm South

In Figure 8, the columns divide the data flow process into different tier levels while the rows categorize the activities by department. Starting from the top left of the graph, the raw materials, intermediates, and final products are weighed by the IPT technicians, recorded manually in the shop floor batch sheets, and input to the Data 3 system under Transaction 18, which specifically refers to a work order receipt to stock. Comparing to the release yield quantities recorded in thousand tablets in the Data 3 system, batch sheets document the whole calculation of the release yield, accountable yield, % release yield and % accountable yield. Meanwhile, a pitch chart is drawn to show the stepwise increment of the material weight. The % release yield of each module, accountable yield for the whole batch, which is called the batch release yield here, and the sampled and reject losses are included in the pitch chart. This job is also performed by the technicians.

Take the selected Product "T" as an example, the calculation of the yield in the batch sheet are given in Table 2. Among the various processes, only the high shear granulation and blending modules of the manufacturing of API "L", the charging and blending modules of API "N" and the bi-layer compression of these two ingredients are involved in yield calculation. The capital letter "Y" is used to denote the yield percentage while the small letter "y" is used to denote the yield quantities. x_B denotes the additional materials added in the blending stage, which are the lubricants particularly for Product "T". x_C denotes the API "N" added to the granules containing API "L" during the compressing stage.

Table 2: Yield Calculation in Batch Sheets, Pharm South

<table>
<tr><th colspan="4">Product "T"</th></tr>
<tr><th colspan="2">API "L"</th><th colspan="2">API "N"</th></tr>
<tr><td>Granulation</td><td>$y_{GTH} = y_{Input}$
$Y_{GRe} = \frac{y_{GRe}}{y_{GTH}} \times 100\%$
$y_{GAc} = y_{GRe} + l_{GS} + l_{GR}$
$Y_{GAc} = \frac{y_{GAc}}{y_{GTH}} \times 100\%$</td><td>Charging</td><td>$y_{NCTH} = y_{Input}$
$Y_{NCAc} = \frac{y_{NCRe} + l_{NCS} + l_{NCR}}{y_{NCTH}} \times 100\%$</td></tr>
<tr><td>Blending</td><td>$y_{BTH} = y_{GRe} + x_B$
$Y_{BRe} = \frac{y_{BRe}}{y_{BTH}} \times 100\%$
$y_{BAc} = y_{BRe} + l_{BS} + l_{BR}$
$Y_{BAc} = \frac{y_{BAc}}{y_{BTH}} \times 100\%$</td><td>Blending</td><td>$y_{NBTH} = y_{NCRe}$
$Y_{NBAc} = \frac{y_{NBRe} + l_{NBS} + l_{NBR}}{y_{NBTH}} \times 100\%$</td></tr>
<tr><td colspan="4">Compressing</td></tr>
<tr><td colspan="4">$y_{CTH} = y_{BRe} + x_C$
$Y_{CRe} = \frac{y_{CRe}}{y_{CTH}} \times 100\%$
$y_{CAc} = y_{CRe} + l_{CS} + l_{CR}$
$Y_{CAc} = \frac{y_{CAc}}{y_{CTH}} \times 100\%$</td></tr>
<tr><td colspan="4">Batch Yield</td></tr>
<tr><td colspan="4">y_{TH} = Number of Tablets × Theoretical Tablet Weight
$y_{Re} = y_{CRe}$
$Y_{Ac} = \frac{y_{Re} + l_{CS} + l_{CR} + l_{BS} + l_{BR} + l_{GS} + l_{GR}}{y_{TH}} \times 100\%$
Y_{Ac}="Batch Release Yield" in pitch charts and metrics by CI engineers</td></tr>
</table>

Apart from the processes to manufacture Product "T", this method of calculation can also be applied to film coating. Since film coating is a process similar to blending in terms of material flow and yield calculation, it can be inferred that the analysis of

Product "T" can be generalized to all pharmaceutical products at the south campus including Product "I".

On the production site, the pitch charts are collected by the CI engineers once a day, tabulated into spreadsheets, and input to their internal database which are subsequently updated once a day. Specifically, they record the % release yield for the process modules and the % accountable yield for the whole batch, which is recorded as batch release yield from this point on. The metrics are then generated by CI engineers by comparing the data in the spreadsheets to the standard yield determined jointly by IPT Leader and GTO Leader. The metrics are submitted to IPT Leader weekly and monthly on the tier 4 level.

The GTO Leader obtains the information mainly from the process engineers while the process engineers acquire the firsthand information from the issue board edited by the technicians on tier 1 basis. They also share the yield metrics generated by CI engineers and report to the GTO Leader. The two leaders charter yield improvement projects together for which either CI engineers or process leads and process engineers are the current practitioners to improve the yield. It depends on the nature of the problem that who would be capable of solving the problems.

The standard yields are not only used by CI engineers, they are also provided to Finance. This stream of the yield data originates from the Data 3 system, which is a system used at the south campus to process all the material flow. Since the Data 3 system is connected to JDE system, which enables Finance to export the processed yield variance from JDE, the standard yields are also input to JDE annually. The formula for the yield variance calculations are built in JDE as indicated by Equation (5) and (6).

$$\text{Yield Variance (Thousand Tablets)} = \text{Actual Release Yield (Financial Yield)} - \text{Standard Yield} \quad (5)$$

Yield Variance ($)

= Yield Variance (Thousand Tablets)*Unit Standard Cost (6)

These data belong to tier 1 level because they can be accessed anytime. However, Finance will only export the yield variance monthly to summarize the favorable and unfavorable yields in the monthly management report and feed back to all the senior managements including GTO, IPT and QO.

After the batch sheets are filled, according to Current Good Manufacturing Practice, they are reviewed by multiple departments. Beside the double-check on the shop floor, the batch sheets are first sent to the document coordinator and then passed on to QA officers for further review. If the batch is approved by QA to be released, the QO and Quality Control (QC) status in Data 3 system will be changed simultaneously. In this case, the production information and the product release information are both stored in Data 3 system.

6.3.2. Current Customers Analysis

Judging from the actual yield data flow in the pharmaceutical facility at the south campus, the current customers of the yield data are technicians, CI engineers, IPT Leader, process engineers, process leads, GTO Leader, Finance, QA and QO Leader. The document coordinator is not considered as customer here because his responsibility is checking the format of documenting instead of the yield data. Comparing to the tier process, the GTO Leader is supplemented to the customer list because of their role in determining standard yield and chartering yield improvement projects. The process engineers work under GTO Leader and have expertise on the various pharmaceutical processes.

A good business process can either satisfy the maximum number of actual customers or minimize the number of actual customers to satisfy. Hence it is essential to identify the actual required customers in the business process development. To avoid

confusion, the customers of the project defined and identified in the previous sections include both the suppliers and customers of the yield data in light of SIPOC.

Suppliers

The current suppliers match the actual suppliers.

All of the manual inputs are completed by technicians under IPT, including the material weight filled in the batch sheets, pitch chart, issue board and transaction 18 in Data 3 system. This is reasonable because they are working on the shop floor and are able to record the firsthand data in details promptly. Further parties taking over these would have to communicate with them, which is inefficient and prone to mistakes due to more steps required and more departments involved.

QA also have access to Data 3, they only change the QO and QC status instead of any specific numbers. However, as explained in the definition of financial yield, whether it is an input depends on whether this information is synchronized with JDE system. If it is, and if JDE is smart enough to filter the yield data that technician input under transaction 18 by the product release information provided by QA, then this QO and QC status changes are "Input" in SIPOC, thus QA is regarded as supplier. On the contrary, if JDE is only linked to transaction 18 in Data 3, or the algorithm of JDE only involves the value under transaction 18, then it means whatever QA does has no impact on the whole yield data flow chain.

Customers

All of the other departments are both suppliers and customers because not only do they receive the yield data but also feed back into the data flow. Since the technicians are only responsible to collect, calculate and record data, not to summarize and analyze the data to make it meaningful, the roles of CI engineers and Finance are rendered vital. However, they are not exposed to the production as much as they are

exposed to the yield data. CI engineers need the yield data to generate the metrics demonstrating the production performance. For this purpose, the yield data should be able to reveal the manufacturing capability regardless of the costs, which explains why the data are collected even before the product release. In contrast, Finance cares about the actual release yield regardless of the reasons for material losses. This misalignment of the purposes for the yield data used by these two departments leads to independent data sources, different calculation methods, and consequently inconsistent yield metrics.

In spite of sharing the yield metrics generated by CI engineers, the process engineers are more involved in the information flow. They acquire the atypical yield information from issue boards and solve the process related problems. Instead of looking at the yield trends, what they need is a conclusion indicating which batch and which process module is out of specification. Specific yield data analysis is not their job.

With the data collected, processed, and summarized, it is the IPT and GTO Leaders' job to make decisions on setting yield standards, prioritizing problem-solving and driving improvements. Moreover, they are also responsible to identify who are capable of solving the problem in terms of the nature of root causes. For instance, CI engineers are able to reduce the variations by standardizing the operation steps, while process engineers can minimize the carry-over materials of each batch in the equipment to reduce the losses. They close up the far-end loop. To meet their needs, the format of the yield data submitted to them should be precise, concise, and problem-emphasized to drive instant response.

Beyond the product release responsibility of QA, the QO Leader is also in charge of quality analysis. Since yield is also a factor indicating whether the current production satisfies the customer need of quality, they receive the feedback information about yield from Finance. Other quality testing activities are not related to yield data,

therefore they are not considered in this thesis, which makes the QO Leader a pure customer.

6.4. Problems Identification at Pharm South

The problems with the current yield data flow are made visible by the previous analysis. There are generally three types of problems. Section 6.4.1 analyzes the various root causes for data inconsistencies. Section 6.4.2 identifies the major redundancies in the current workflow. In Section 6.4.3, the inefficiencies are briefly discussed.

6.4.1. Root Causes for Data Inconsistencies

1) Accountable Yield Overestimation

By adding back the sampled and reject losses to calculate the accountable yield as indicated in Table 2, there is a hidden assumption that these losses could have yield 100% out of the following processes, which is apparently not possible. Therefore, the current accountable yield is overestimated. A simplified mathematical model is constructed as follows to illustrate this problem.

Consider a production line of two process modules in Figure 9, process 1 and process 2. Assume the % accountable yield and the sampled and reject losses of both processes are constant regardless of the input weight.

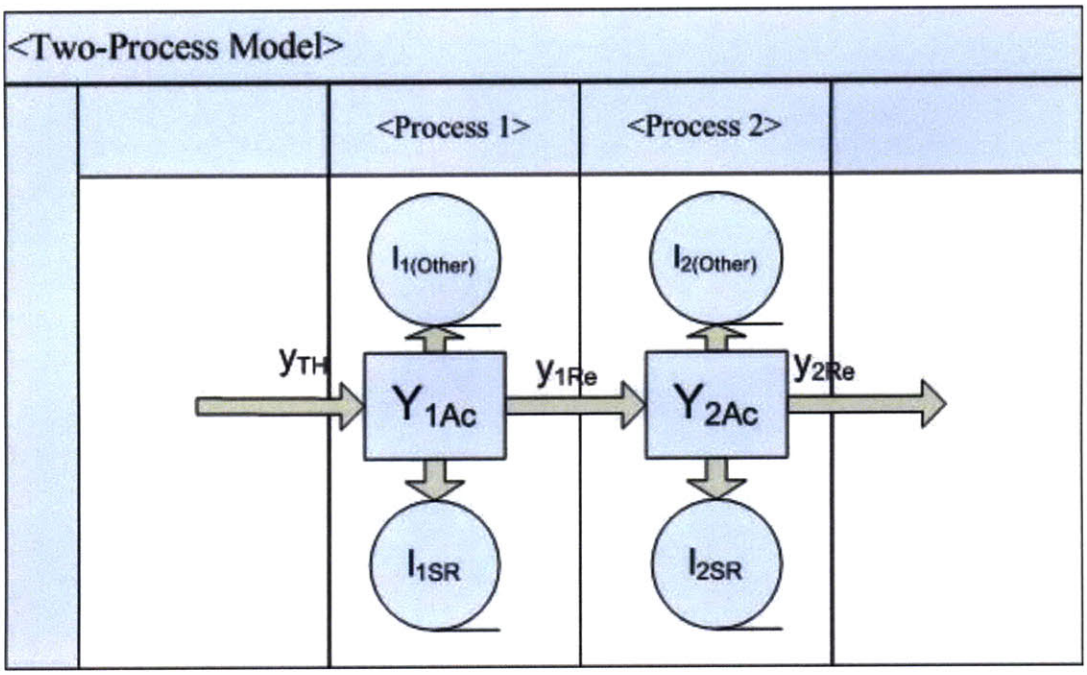

Figure 9: Two-Process Model

Assume the two processes are independent and the yield of them are not affected by each other, theoretically, the accountable yield for the whole batch can be calculated as follows:

$$Theoretical\ \%\ Accountable\ Yield = Y_{1Ac} \cdot Y_{2Ac} \tag{7}$$

$$Theoretical\ Accountable\ Yield = y_{TH} \cdot Y_{1Ac} \cdot Y_{2Ac} \tag{8}$$

In terms of material flow, the theoretical yield is the same as the input weight, which equals the sum of release yield, sampled and reject losses, and other unaccountable losses.

$$y_{TH} = y_{1Re} + l_{1SR} + l_{1(Other)} = y_{TH} \cdot Y_{1Ac} + l_{1(Other)} \tag{9}$$

$$y_{1Re} = y_{2Re} + l_{2SR} + l_{2(Other)} = y_{1Re} \cdot Y_{2Ac} + l_{2(Other)} \tag{10}$$

The current way of calculating accountable yield is adding all the sampled and reject losses back to the release yield. Using Equation (9) and Equation (10), the current

calculation of accountable yield can be transformed into a comparable form to the theoretical accountable yield as in Equation (11).

$$\begin{aligned} &Current\ Accountable\ Yield \\ &= y_{2Re} + l_{2SR} + l_{1SR} \\ &= (y_{TH} \cdot Y_{1Ac} - l_{1SR}) \cdot Y_{2Ac} + l_{1SR} \\ &= y_{TH} \cdot Y_{1Ac} \cdot Y_{2Ac} + l_{1SR} \cdot (1 - Y_{2Ac}) \end{aligned} \quad (11)$$

Apparently, this value is greater than the theoretical value, which means that currently the accountable yield is being overestimated. A more reasonable solution according to this simplified model is to correct the sampled and reject losses by the accountable yield of the downstream process, as in Equation (12).

$$\begin{aligned} &Suggested\ Accountable\ Yield \\ &= y_{2Re} + l_{2SR} + l_{1SR} \cdot Y_{2Ac} \\ &= y_{TH} \cdot Y_{1Ac} \cdot Y_{2Ac} \end{aligned} \quad (12)$$

The simplified model explained how the current accountable yield is overestimated. Unfortunately, in the real production, the assumptions of this model are challenged. The accountable yield cannot be a constant value. Particularly, it is affected by how much is taken out of the upstream process for sampling and how much is put into this step. The manufacturing yield capability varies as the input weight changes. Moreover, the reject loss is uncertain and the amount for sampling is not a fixed weight, or a fixed percentage of the weight. Therefore, there is no fixed value to correct the upstream losses. In other word, it is impossible to judge what would happen if the upstream losses did not occur.

2) Inconsistent Names and One Name with Inconsistent Meanings

Although all the yield data are well documented in batch sheets, hardly anyone will retrieve them from the archive after the QA release. Therefore, when "yield" is

mentioned, it refers to the data that are included in the CI engineers' metrics, and the data exported by Finance. Since what Finance exports from JDE is already processed yield variance, it is crucial that the data that CI engineers provide are precise and unambiguous.

However, the names of the yield are confusing. The metrics include the % release yield for each process module, whereas the batch release yield is actually the % accountable yield in the batch sheet. Finance also name the quantities in Data 3 yield, nevertheless, those numbers are normally 0.1-0.2% away from the batch release yield in the yield metrics. Because financial yield, accountable yield, release yield and batch release yield are all called yield and are not distinguished except in batch sheets, the root cause for the data inconsistencies are made invisible.

3) Standard Yield Derived from Accountable Yield but Referred by Financial Yield

The IPT Leader and GTO Leader determine the yield standards once a year based on the batch release yield provided by CI engineers, which is also the accountable yield in the batch sheet. However, these data are referred as standards of the financial yield in the JDE system. Considering the difference between the financial yield and release yield and the difference between release yield and accountable yield, this mismatch may result in a set of ineffective yield standards.

4) Multiple Systems

The multiple systems being used is another cause for the data inconsistencies. As explained in the definition of financial yield, whether the product release information is synchronized with JDE system is unknown. If it is, and if JDE is smart enough to filter the yield data that technician input under transaction 18 by the product release information provided by QA, then the financial yield can be different from the release yield. On the contrary, if JDE is only linked to transaction 18 in Data 3, or the

algorithm of JDE only involves the value under transaction 18, the financial yield should be equal to the release yield but not necessarily the batch release yield.

6.4.2. Workflow Redundancies

There are two data systems in the current yield data flow, whereas the metrics still comes from CI engineers' manual input. Moreover, the existence of the data systems causes more problems instead of facilitating and harmonizing the data flow.

In addition to the manual operation, repetitive work also contributes to redundancy problem. Both CI engineers and Finance are responsible for comparing the actual yield to the yield standards and report to the senior managements, despite the different purposes of calculating yield, it is not reasonable for the IPT and GTO Leader to receive two sets of yield data especially when these two sets of data are inconsistent because they only determine one set of yield standards.

6.4.3. Inefficiencies

In the pharmaceutical facility of the south campus, process engineers are part of tier 1 process in terms of their responsibility on process improvement. However, the yield data that they can access are the tier 4 metrics, either weekly or monthly. This information is not sufficient or convenient to support and drive daily improvements. Consequently, the process engineers are not closely involved as the way they should be.

6.5. Recommendations at Pharm South

Targeting at the problems identified and summarized in Section 6.4, the potential solutions are proposed in this section.

6.5.1. Yield Concepts Harmonization

As discussed in the previous section, the current way of calculating the accountable yield is suboptimal and there are few if any method to correct this calculation to precisely evaluate the real production capability. Considering that all the other facilities at both campuses in this company are using release yield to monitor the production performance, the release yield should be adopted in the pharmaceutical facility as well to generate metrics. This substitution is recommended because it is helpful to unify the yield data throughout the whole company.

If the release yield is adopted, there will not be inconsistent names problem, and particularly, the batch release yield will be the release yield of the whole batch instead of the accountable yield. In addition, the difference between the metrics and the financial yield will also be reduced by eliminating the deviation of accountable yield from the release yield, leaving the unreleased batches as the only difference between the financial yield and the metrics.

To avoid confusion, the financial yield should always be referred to by its full name until the release yield and the financial yield are made completely identical.

6.5.2. Yield Divergence Postponement

The difference between the release yield and the financial yield is the root cause for the problem 3) and 4) discussed in Section 6.4.1.

On one hand, if the release yield is adopted for the batches in the metrics, the standard yield will be determined upon the release yield and referred by the financial yield. On the other hand, if the definition of the release yield is identical with the definition of the financial yield, no matter which transaction in Data 3 system is connected to the JDE system, there should not be any problem with multiple data systems.

To meet the need of IPT and GTO Leaders, the format of the yield data submitted to them should be precise, concise, and problem-emphasized to drive instant response. However, the yield trends information submitted by CI engineers and Finance are not only repetitive, but also inconsistent. The root cause for this repetitive calculation is also the difference between the release yield and the financial yield.

Considering all the problems caused by the same reason, the release yield and the financial yield should be unified. However, since it is significant for Finance to know that whether the product is released while it does not affect the production capability that CI engineers are concerned about, it is difficult to ignore the production release information. Therefore, to achieve one version of truth by merging the two parallel streams of data flow, and to reduce the work redundancies, the divergence of the yield data should be postponed.

This idea is derived from the postponement strategy used in supply chain management to minimize the risk by delaying further investment until the demand information is available. Here the proposed postponement refers to the adoption of release yield by both Finance and CI engineers until the product is confirmed not to be released. Namely, these two departments share the same database, either Data 3 or CI engineers' internal database to monitor the release yield. This requires the correction of the interfaces between JDE and Data 3 system to insure that only the release yield under transaction 18 is connected so that CI engineers can also export the yield data from Data 3 system. An alternative way is that technicians input the batch sheet yield data directly to the CI engineers' database instead of inputting to Data 3 system. If the product is not released, the data flow should be reprocessed to eliminate the data from the financial yield, meanwhile, this information should be fed back to the senior managements for further investigation.

In this case, this strategy not only reduces the repetitive work and times of manual operations, but also gives prominence to the abnormal batches. Furthermore, the IPT,

GTO, QO Leaders will receive one set of harmonized yield data and a notification or report of the unreleased batch.

Finally, by this strategy, all of the yield data are stored in the same data system and updated daily by technicians; hence process engineers can access the database anytime to review the related process yields so as to fix the problems responsively.

6.6. Yield Definitions at API South

Because of the different nature of the API production comparing to pharmaceutical production, some of the yield concepts are defined differently. The terminologies for API products are explained as follows.

1) (Operational) Yield

Since most of the API processes, especially the processes for API "Eto" at the south campus are semi-continuous, which is also called close-coupled, the whole production line is controlled by DCS with the yield data stored in the PI system. The yields of individual processes are not manually recorded in the batch sheet as in the pharmaceutical facility, instead, only the actual net charged material weight and the net product weight are weighed in DCS and recorded manually in the batch sheet. The API production involves chemical reaction; therefore the yield is defined in terms of moles, as in Equation (13). The term "operational" is added to distinguish it from the other yield concepts in this thesis. This concept is denoted as Y_{Op}.

$$Y_{Op} = \frac{Net\ Product\ Weight \times Assay\ of\ Output}{Actual\ Net\ Charged\ Weight \times Assay\ of\ Input} \times \frac{Molecular\ Weight\ of\ Input}{Molecular\ Weight\ of\ Output} \times 100\% \quad (13)$$

In which, assay of output and assay of input refers to the percentage of the major ingredients in the net product weight and the actual net charged weight, respectively. They are factors of purity. The unit of molecular weight is kg/mol. Since only the net product weight and actual net charged weight are not fixed, the result of the constant part is calculated for each product and built in DCS. In this thesis, this constant is named correction factor, as indicated in Equation (14).

$$Correction\ Factor = \frac{Assay\ of\ Output}{Assay\ of\ Input} \times \frac{Molecular\ Weight\ of\ Input}{Molecular\ Weight\ of\ Output} \tag{14}$$

The operational yield is stored in PI and shared by everyone on the production site.

2) Theoretical Yield

The theoretical yield is the ideal net product weight. It is denoted as y_{TH} and calculated as in Equation (15). When $Y_{Op} = 1$,

$$y_{TH} = Actual\ Net\ Charged\ Weight \div Correction\ Factor$$
$$= Actual\ Net\ Charged\ Weight \times \frac{Assay\ of\ Input}{Assay\ of\ Output} \times \frac{Molecular\ Weight\ of\ Output}{Molecular\ Weight\ of\ Input} \tag{15}$$

The unit is kilogram.

The theoretical yield is only considered by the process engineers.

3) Standard Yield/Standard Lot Size

The standard yield used by Finance, also called standard lot size in Data 3 system is determined by GTO and IPT jointly based on theoretical data and historical trends similarly to the process in the pharmaceutical facility. However, in API, this number is given in kilograms.

4) (Financial) Yield

The financial yield is obtained exactly the same way as it is obtained for the pharmaceutical products at the south campus. Therefore, it is defined similarly and involves the same problem related to the QA release. If the unreleased product can be assumed to be a rare case considering the high-quality environment, the financial yield should be equal to the net product weight. Therefore, the financial yield is expressed in kilograms.

5) Yield Variance

The yield variance is defined as the difference between the financial yield and the standard yield measured in kilograms or dollars. Apart from the unit difference, everything follows the same procedures as for pharmaceutical products.

6.7. Data Flow Tracking at API South

The tier process in the API facility is the same as the pharmaceutical facility provided that the tier 3 and tier 4 processes are performed by the same set of people; and although the departments are independent for different facilities at the tier 1 and tier 2 levels, they are attached to the same senior managements. In addition, the other potential customers can initially be assumed to be the same and examined in the data tracking. Therefore, based on the experience gained in the analysis of the pharmaceutical facility at the south campus, the yield data of API "Eto" is tracked in the API facility of the south campus. Similarly to the investigation conducted in the pharmaceutical facility, a Yield Data Flow Diagram is plotted in Section 6.7.1 with a detailed description of the role of each department. The data format at each stage is also explained in this section. The current customers are confirmed in Section 6.7.2.

6.7.1. Data Flow Diagram

The Yield Data Flow Diagram is plotted in Figure 10 to depict the current yield data flow chain. This yield data flow is applicable to most of the API products as API "Eto" when the processes are close-coupled thus can be controlled by DCS.

Figure 10 follows the same denotation convention as in Figure 8. Starting from the top left of the graph, the actual net charged material weight and the net product weight are weighed manually by the technicians, input to DCS, and recorded in shop floor batch sheets and the Data 3 system under transaction 18. Although there are also batch sheets in the API facility, they are not used for generating yield metrics. Instead, since everything is dynamically captured in the DCS and stored in PI system, the yield data can be exported from the PI system whenever it is needed by anyone who has authorized access to PI. Particularly, the operational yield formula is embedded in the DCS system, thus the data exported from PI are operational yield expressed in percentage.

Different from the pharmaceutical facility, the yield data is exported everyday by shift leads and passed on to the CI engineers in spreadsheets. The metrics are generated by the CI engineers based on the data in the spreadsheets and the standard yield determined jointly by IPT Leader and GTO Leader, and then submitted to the IPT Leader both weekly and monthly on the tier 4 level.

The process engineers obtain the information from both the issue board edited by the technicians on tier 1 basis and the batch release report exported from the PI system. They report to process leads.

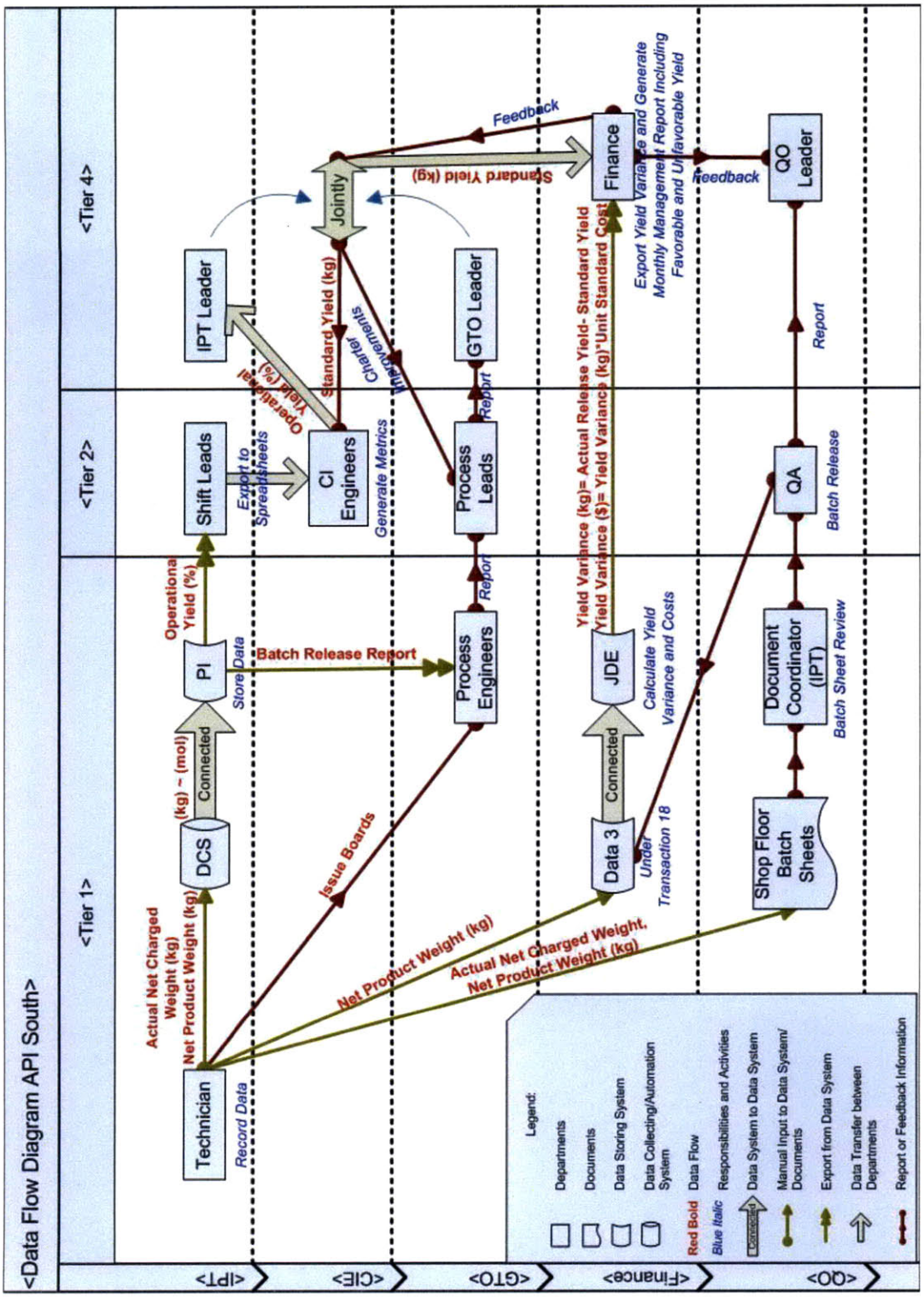

Figure 10: Data Flow Diagram API South

The rest of the data flow is the same as in the pharmaceutical facility except the unit of the data are kilogram, percentage, and dollars instead of thousand tablets. For instance, the formula of yield variance becomes:

$$\text{Yield Variance (kg)} = \text{Actual Release Yield (Financial Yield)} - \text{Standard Yield} \quad (16)$$

$$\text{Yield Variance (\$)} = \text{Yield Variance (kg)} * \text{Unit Standard Cost} \quad (17)$$

And the data stored in Data 3 system are in kilograms.

6.7.2. Current Customers Analysis

Judging from the actual yield data flow in the API facility at the south campus, the current customers of the yield data are technicians, shift leads, CI engineers, IPT Leader, process engineers, process leads, GTO Leader, Finance, QA and QO Leader. Apart from the shift leads, all the other departments are performing the same function as in the pharmaceutical facility. In the API facility, the shift leads share part of CI engineers' responsibility for summarizing the data. Therefore, they are both the supplier and the customer of the yield data because not only do they export the yield data but also submit it.

6.8. Problems Identification at API South

The problems with the current yield data flow in the API facility are similar to the flow in the pharmaceutical facility. However, process engineers are more closely involved in the real-time yield data enabled by DCS and PI system, which makes the data flow more efficient. In addition, the DCS and PI system help reduce the manual operation. Instead of transferring the yield data from a paper document to spreadsheets, the yield data can be exported from the data system, which saves a lot

of labor hours and avoids potential human errors. Since the only difference between the operational yield and the financial yield comes from the unreleased product, the production release information should not affect the yield standard.

What DCS records is equivalent to the % release yield in the pharmaceutical facility, all the problems about accountable yield including the overestimation and name with inconsistent meanings does not exist, leaving the problems related to the difference between operational yield and financial yield unsolved, which are briefly listed below:

1) Since whether the product release information is synchronized with JDE system is unknown, the data in JDE may be different from the operational yield if the product is not released.

2) Both CI engineers and Finance compare the actual yield to the yield standards by different methods and report the equivalent information to the senior managements, which is repetitive.

6.9. Recommendations at API South

All the problems left can be resolved by the postponement concept proposed in Section 6.5.2. The postponement can be achieved if both Finance and CI engineers use the same database, either Data 3 or PI to monitor the operational yield. If Data 3 is employed, the correction of the interfaces is required between JDE and Data 3 system to insure that only the operational yield under transaction 18 is connected so that CI engineers can also export the yield data from Data 3 system. An alternative way is that Finance export the data from PI system instead. If the product is not released, the data flow should be reprocessed to eliminate the data from the financial yield, meanwhile, this information should be fed back to the senior managements for further investigation.

In this case, this strategy reduces the repetitive work, unifies the data system, and gives prominence to the abnormal batches.

7. CURRENT YIELD DATA FLOW ANALYSIS AT THE WEST CAMPUS

The west campus was previously an independent company; therefore, the production lines, the organizational structure and the yield data flow process are completely different from those of the south campus. Since the information at the west campus is subjected to changes due to the internal restructuring after the merger, this thesis focuses less on the manufacturing processes and the organizational structure of yield data processing at the west campus. The information obtained from the west campus supplements the detailed description of the current yield data flow at the south campus. Section 7.1 describes the organizational structure at the west campus, namely the tier process. Section 7.2 defines the yield concepts in the API facility at the west campus. In Section 7.3, the yield data flow is plotted to demonstrate the different way of processing the yield data at API west while the problems and recommendations are discussed in Section 7.4. Similarly, the yield definitions and the yield data flow at pharm west are explained in Section 7.5 while the problems and recommendations are discussed in Section 7.6.

7.1. Tier Process at the West Campus

There are only two tier levels at the west campus, because the complete tier process establishment is still in progress. The current organizational structure is depicted in Figure 11.

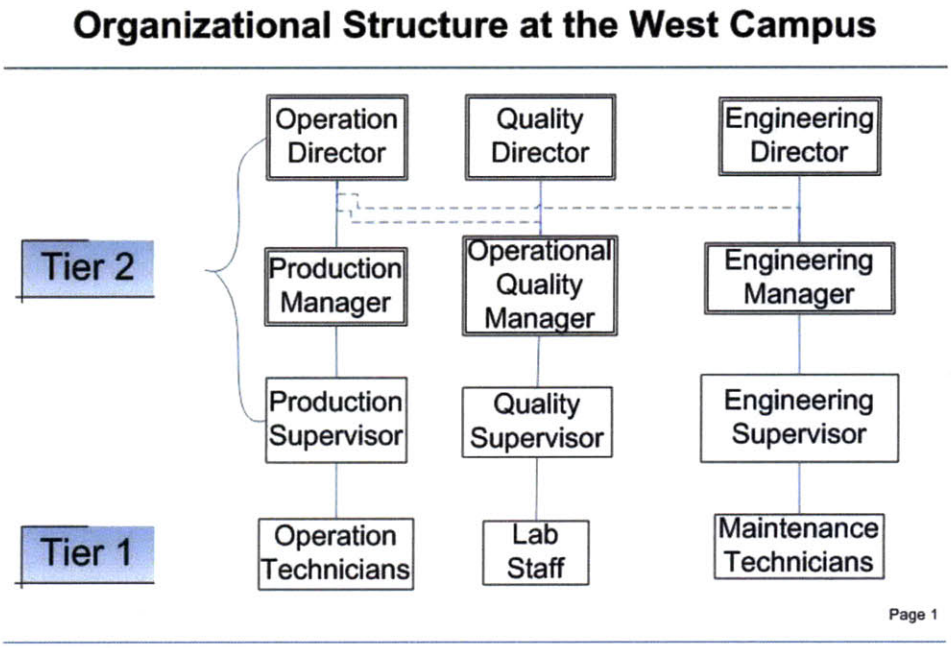

Figure 11: Organizational Structure at the West Campus

As indicated in Figure 11, each department consists of four levels and their relations are all direct reports except that the operational quality manager and the engineering manager also need to report to the operation director.

7.2. Yield Definitions at API West

The yield definitions in the API facility at the west campus include the yield definitions of both pharmaceutical and API facilities at the south campus.

1) % Actual Yield

The concept is exactly the same as the % release yield defined in the pharmaceutical facility at the south campus. However, the numerator of the % actual yield is called net discharged weight; and the input weight is named net charged weight here.

2) Actual Yield

The actual yield is the net discharged weight, which is actually the same as the concept of release yield in the pharmaceutical facility at the south campus.

3) (% Accountable) Yield

There is no specific name for this concept defined in batch sheets; instead, it is simply called "yield". To differentiate the yield concepts in this thesis, it is named as % accountable yield, because similar to % actual yield, this concept is exactly the same meaning as the % accountable yield defined in the pharmaceutical facility at the south campus. The formula is given below in Equation (18).

$$\%\ Accountable\ Yield = \frac{Actual\ Yield + Sampled\ and\ Reject\ Losses}{Theoretical\ Yield} \times 100\% \qquad (18)$$

4) Theoretical Yield

The theoretical yield is the ideal net discharged weight. It is denoted as y_{TH} as in the API facility at the south campus, and calculated as in Equation (19).

$$y_{TH} = \text{Net Charged Weight} \div \text{Correction Factor}$$
$$= \text{Net Charged Weight} \times \frac{\text{Assay of Input}}{\text{Assay of Output}} \times \frac{\text{Molecular Weight of Output}}{\text{Molecular Weight of Input}} \qquad (19)$$

The unit is kilogram.

5) Target Yield

The target yield is determined annually by the operation director based on the historical yield data of the actual yield and the theoretical knowledge. It is used by Finance to calculate costs regarding yield monthly.

7.3. Data Flow Tracking at API West

Not only are the production lines similar to each other and can be generalized to a common one, the yield data flow is not differentiated by product in this facility either. A Data Flow Diagram is plotted in Figure 12. To facilitate the comparison of the data flow of all of the four facilities in this company, the diagrams all follow the same denotation convention.

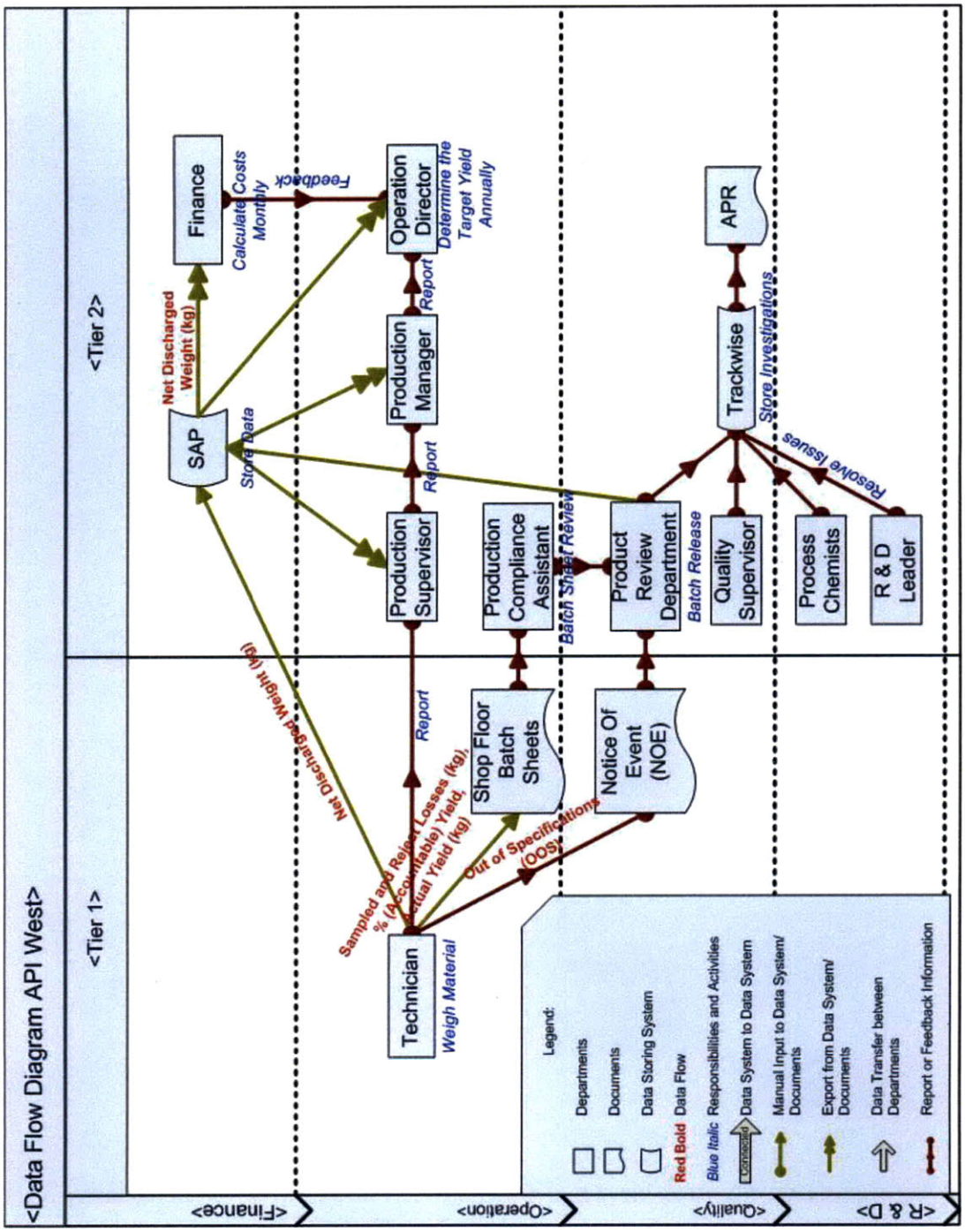

Figure 12: Data Flow Diagram API West

In the API facility at the west campus, SAP is the central data system that records not only all the material flow but also the yield trend. The unit of yield in SAP is kilogram.

Same as all the other facilities, the raw data come from the shop floor and are documented by the technicians in batch sheets and the SAP system. If there is any yield problem, they will raise a notice of event (NOE) to the product review department, and then product review, quality supervisors, process chemists, and the R& D Leader will solve the problem together. Apart from the NOE, their information also comes from the batch sheets reviewed by the production compliance assistant and the product review department. The whole process of problem solving is stored in Trackwise, which is a quality management software. Furthermore, the annual product review (APR) will also capture the information stored in Trackwise. Along the information line from the technician to the operation director, there is no data transfer involved because of the centralized database SAP. In addition, since Finance can export data from SAP instead and feed back to the operation director, the application of SAP also enables the elimination of JDE from the data flow.

Note that only the net charged products and the net discharged products are weighed and recorded in batch sheets. Also included in batch sheets are the theoretical yield, actual yield, and % accountable yield calculation.

$$Theoretical\ Yield = Net\ Charged\ Weight \times \frac{1}{Correction\ Factor} \tag{20}$$

$$Actual\ Yield = Net\ Discharged\ Weight \tag{21}$$

$$\%\ Accountable\ Yield = \frac{Actual\ Yield + Sampled\ and\ Reject\ Losses}{Theoretical\ Yield} \times 100\% \tag{22}$$

The target yield is determined annually by the operation director and provided to Finance to monitor the yield variance and related costs. Here the yield variance refers to the difference between the actual yield and the target yield. Finance will feed back

to the production department every month of the favorable and unfavorable batch information.

7.4. Problems and Recommendations at API West

In the API facility at the west campus, all the departments are considering the actual yield. Therefore, there is no problem caused by the inconsistency of definitions. In addition, comparing to the south campus where two to four data systems are involved in the yield data flow, the west campus has only one centralized data system, SAP. Trackwise documents the process of quality improvement, and is not considered as a data system. Hence data system does not cause any problem either.

Comparing to the data flow processes at the south campus, it is observed that GTO is not involved in the yield data flow in this facility, because they are mainly in charge of new product validation and transport at the west campus; however, it might be better to involve GTO in the yield data flow considering their ability in process improvement. This involvement is not only the access to the central data system, but also a standard procedure to drive the integration of their manufacturing process knowledge to the business process flow.

As mentioned previously, there are only two tier levels established at the west campus. This might not be enough to avoid ambiguous responsibility and miscommunication. SAP helps save labor hours in the transmission of the yield data; however, simply storing the yield data in a centralized system does not necessarily strengthen the effectiveness of the information. Without processing the data level by level in the tier process, there will not be visible conclusion of the yield problems to drive immediate response from the upper management. The identification of problems and prioritization of projects is actually achieved along the yield information flow, thus two tier levels can be interpreted as that the yield data is only processed once, which is not a sufficient organizational structure for problem solving. Therefore, the tier

process at the west campus should be further developed to standardize the responsibility in the yield data flow.

Another reason for further developing the tier process is that the most frequent review of the yield data is monthly in the API facility at the west campus by Finance while yield is reviewed and updated every day either by CI engineers or shift leads at the south campus. Considering the high costs of the pharmaceutical industry, it is reasonable to introduce some mechanism to capture the atypical on a more frequent basis.

7.5. Yield Definitions and Data Flow Tracking at Pharm West

The organizational structure in the pharmaceutical facility at the west campus is generally the same as in the API facility at the same campus. Nevertheless, the products manufactured in this facility are totally different.

The actual yield, the % actual yield, the accounted (accountable) yield and the % accounted yield in the batch sheet of this facility are all the same as the definitions in the API facility at the west campus. The only difference is that for compression process, the unit is tablet instead of kilogram. The target yield is still determined based on historical yield data. However, the theoretical yield for each process module is defined as the input weight as in the pharmaceutical facility at the south campus.

Since the yield data flow through the departments is the same for product "Z" and product "Na", it is tracked and plotted in Figure 13. As can be observed from the intersecting lines, the data and information flow demonstrate the cross function activities in this facility. As in all the other facilities, the raw data are collected by technicians. They record all the yield data including actual yield and % accountable yield in batch sheets and input to the SAP system. If there is any out-of-specifications

(OOS), they will report to the production supervisor. The rest of this flow to operation director in the operation department are all information flow, which means they do not transfer the yield data itself, but report the phenomena and problems. The batch sheets are reviewed by the production supervisor and production manager and then forwarded to the quality supervisor where the batches are released. If a batch is approved for release, the quality supervisor will change the status of this item in the SAP system.

Annually, the yield data is summarized into the APR by technicians. The APR is utilized by GTO, the production manager, and the quality manager to make improvements accordingly. In addition, the NOE is also captured by GTO who is responsible for technical issues.

The yield is reviewed every month by the operation director, the production manager, and the operation excellence leader in the Monthly Line Performance Review Meeting. The operation excellence leader exports the yield data from SAP and summarized into spreadsheets manually to show the yield trends to the operation director and the production manager during this meeting. And the latter two respond to the problems and charter improvements. The production manager is also responsible for setting the yield target annually for financial need.

With the target yield provided by the production manager, Finance exports the actual yield from SAP to calculate the yield variance to justify whether the yield for each batch is favorable or not, and feed back to the production manager and the operation director during the Monthly Finance Review Meeting.

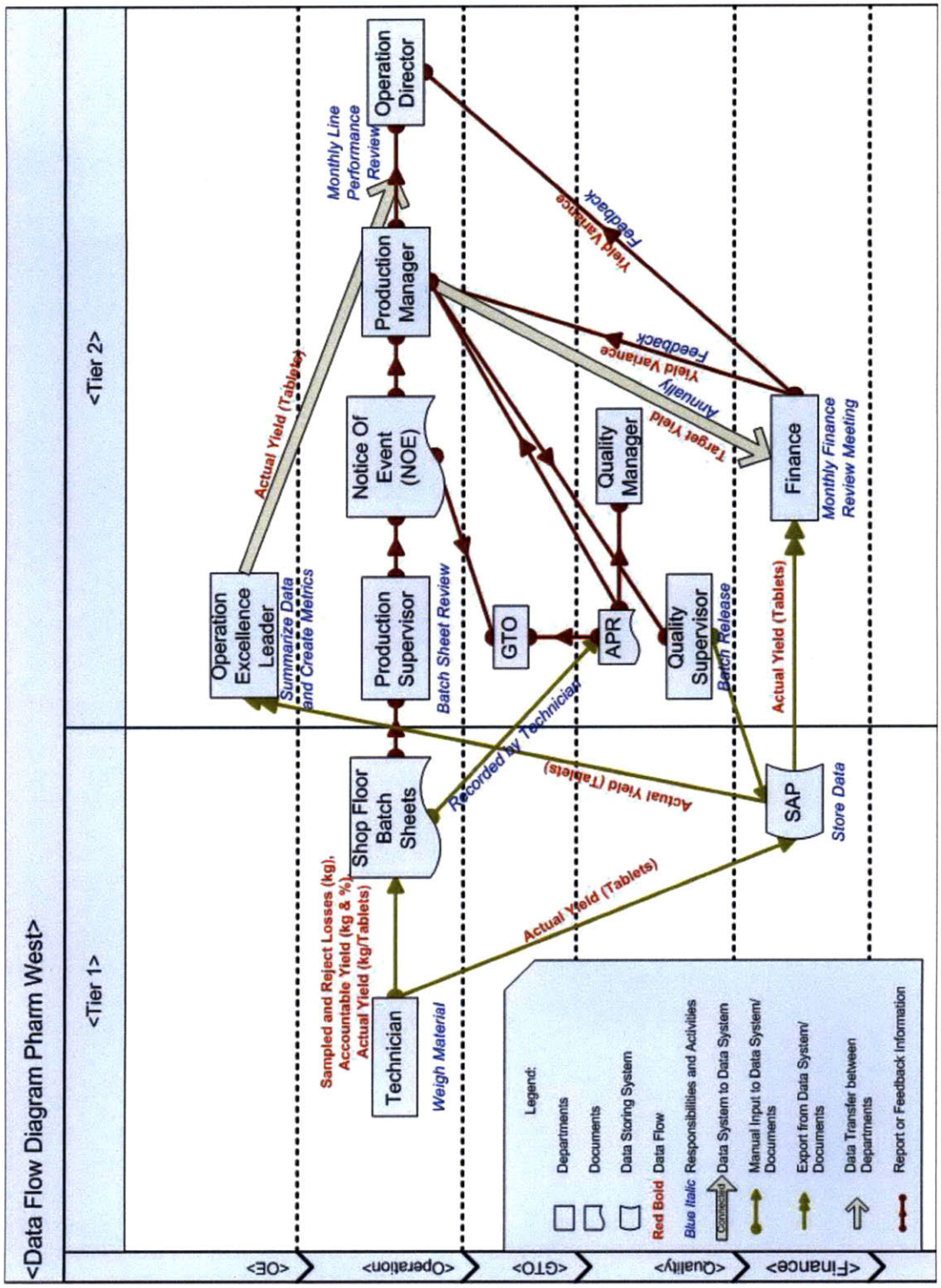

Figure 13: Data Flow Diagram Pharm West

While the data flow through the departments is the same for both products, the specific yield calculations are not necessarily identical. Moreover, the yield calculations for product "Z" can be compared with those of product "T" to show the inconsistencies between the two pharmaceutical facilities. The specific calculations are given in Table 3 and Table 4, for product "Z" and product "Na" respectively.

Table 3: Product "Z" Yield Calculation in Batch Sheets, Pharm West

Product "Z"	
Processes	Formula
Granulation	Total Theoretical Yield: $y_{GTH} = y_{Charged}$ Actual Granulation Yield: $y_{GActual}$ Assignable Granulation Loss: l_{GAc} Total Accounted Granulation: $y_{GAc} = y_{GActual} + l_{GAc}$ % Actual Yield: $Y_{GActual} = \frac{y_{GActual}}{y_{GTH}} \times 100\%$
Lubricant Blend	Actual Blended Yield: $y_{BActual}$ Assignable Blending Loss: l_{BAc} Total Accounted Blend Yield: $y_{BAc} = y_{BActual} + l_{BAc}$ Adjusted Theoretical Blend Yield: $y_{BTH} = y_{GActual} + x_B$ % Total Accounted Blended Yield: $Y_{BAc} = \frac{y_{BAc}}{y_{BTH}} \times 100\%$ In which x_B is the weight of the lubricants.
Compression	Total Net Weight of Tablets: $y_{CActual}$ Tablet Rejects+ Recovered Granulation+ Quality Control Samples: l_{CAc} Total Accounted Tablet Yield: $y_{CAc} = y_{CActual} + l_{CAc}$ % Total Accounted Yield: $Y_{CAc} = \frac{y_{CAc}}{y_{BActual}} \times 100\%$
Yield recorded in SAP	Total Number of Tablets Compressed $= \frac{y_{CActual}(kg) \times 1000000}{\text{Average Tablet Weight (mg)}}$

Table 4: Product "Na" Yield Calculation in Batch Sheets, Pharm West

Product "Na"	
Processes	Formula
Compounding	Net Weight of Compounding Tank: $y_{Charged}$ Net Weight of Holding Tank: y_{Actual} Process Rejects+ Samples: l_{Ac} Accounted Yield: $y_{Ac} = y_{Actual} + l_{Ac}$ % Actual Yield: $Y_{Actual} = \frac{y_{Actual}}{y_{Charged}} \times 100\%$ % Accounted Yield: $Y_{Ac} = \frac{y_{Ac}}{y_{Charged}} \times 100\%$
Yield Recorded in SAP	y_{Actual} (kg)

Comparing the product "Z" with product "T", it can be noticed that the yield percentage for the whole batch is not calculated. What finally recorded in SAP is the actual number of tablets, which is equivalent to the actual yield in the API facility at the west campus, the net product weight in the API facility at the south campus, and the release yield in the pharmaceutical facility at the south campus. Despite the different names, the calculation of yields for product "Z" and product "T" are identical. The cause of the only inconsistency is that the batch release yield, which is actually accountable yield for the batch, is used in the metrics of product "T" while the actual yield is recorded in the SAP system as the main data source.

The principle of calculating the yield for product "Z" and product "Na" are the same as indicated in the formulae. The reason why they are recorded in different formats and units in SAP is that product "Na" is liquid.

7.6. Problems and Recommendations at Pharm West

Since the actual yield is considered in all the departments, there is no problem caused by inconsistent yield definitions. In addition, there should not be any problem caused by multiple data systems because only SAP is involved in the yield data flow in this facility.

Although GTO is in charge of solving the problems captured in NOE and APR, they are not closely involved in the yield data flow. Considering their expertise in the various processes, the yield data would be of great value to identify the problems of a process. On the other hand, it would be too late for GTO to trace back the historical yield data if the yield problem has already occurred.

As in the API facility at the west campus, there are only two tier levels, the complete 5-tier process has not been established yet, which renders the yield data not fully utilized. For instance, the summarized spreadsheets are only used during the Monthly Line Performance Review Meeting. Neither GTO nor Quality is exposed to this processed data. Moreover, the most frequent response to the yield data is monthly, which might not be the optimized frequency. A daily or weekly review of the yield trends should be performed by someone who is familiar with the production processes and are capable of practicing improvements.

8. DISCUSSION

After defining the yield concepts and explaining the various organizational structures, the yield data flow was analyzed, the problems were identified and the potential solutions were proposed in all of the four facilities in this company. Therefore, a comparison of the yield data flow at the four facilities is made in Section 8.1. In Section 8.2, the problems for each facility are summarized; the recommendations for each facility are also the requirements for the new business process. Finally, a new business process is established in Section 8.3.

8.1. Yield Data Flow Comparison

The yield data flow in each facility is compared in Table 5. As summarized in the table, though the materials at all of the four are weighed manually, the calculation in the API facility at the south campus is done by DCS while it is done by technicians in the other facilities. This difference in measurement contributes to different processing time and different responsibilities for CI engineers and shift leads at different facilities. In the pharmaceutical facility at the south campus, it is the CI engineers who collect the yield data and generate the yield metrics. However, in the API facility, shift leads undertake the job of exporting yield data from PI for CI engineers to generate yield metrics, which means the same process is completed by two departments, which also means the data is collected on a more frequent tier 2 daily basis. At the west campus, the Operation Excellence (OE) Leader is in charge of summarizing the yield data from SAP in the pharmaceutical facility while it is not summarized into spreadsheets in the API facility.

In spite of different names used at the two campuses, the actual yield is an equivalent concept to the release yield. And all of the four facilities are considering the actual

(release) yield except the pharmaceutical facility at the south campus, which means the IPT Leader is looking at different data for the two facilities at the south campus.

By further comparison of the two campuses, the most obvious difference is the data systems. Two data systems are involved in the calculation of the financial yield at the south campus, which causes potential problems. The centralized SAP system eliminates the inconsistency caused by multiple systems.

Other differences between the two campuses include the more specifically defined tier process and more intense involvement of GTO at the south campus. In addition, though more frequently reviewed, the yield information is not captured in Annual Product Review (APR) at the south campus. By comparing the batch sheets at both campuses, it can be noticed that at the west campus only the yield of the process modules are calculated. In contrast, the yield of the whole batch is also calculated at the south campus beside the yield of the process modules.

	Pharmaceutical	API
South	• Manually weigh the materials and calculate yield • CI engineer summarize data • Accountable yield in metrics • Data 3 and JDE applied • 4 tier level involved • GTO involved • Data daily and weekly reviewed • Yield not captured in APR • Batch yield calculated	• Manually weigh but automatically calculate yield • Shift lead engineer summarize data • Release yield in metrics • Data 3 and JDE applied • 4 tier level involved • GTO involved • Data daily and weekly reviewed • Yield not captured in APR • Batch yield calculated
West	• Manually weigh the materials and calculate yield • OE summarize data • Actual yield considered • SAP applied • 2 tier level involved • GTO involved • Data monthly reviewed • Yield captured in APR • Only process module yield calculated	• Manually weight the materials and calculate yield • Data not summarized • Actual yield considered • SAP applied • 2 tier level involved • GTO not involved • Data monthly reviewed • Yield captured in APR • Only process module yield calculated

Table 5: Yield Data Flow Comparison

8.2. Problems and Requirements for the New Business Process

The various problems identified previously in all of the four facilities are summarized in Table 6. These are the results of performing gap analysis. The current business processes were compared to the desired process and the problems, that is to say the gaps were identified. The method of proposing solutions toward these problems follows the theory of gap analysis, because the recommendations proposed to each facility are also the requirements for the new business process to be established in the next section.

Facility\ Analysis	Problems (Gaps)	Recommendations
Pharm South	• Accountable yield overestimation • Inconsistent name and meanings • Mismatch of yield standards • Multiple systems	• Use release yield for metrics to harmonize concepts
	• Repetitive work • Inefficient information to drive improvements	• Centralize database to achieve yield divergence postponement
API South	• Inconsistency of operational and financial yield • Repetitive work	• Centralize database to achieve yield divergence postponement
API West	• GTO not effectively involved • Only two tier levels • Monthly review of yield data not enough	• Involve GTO in tier process • Specify the a more detailed tier process • Standardize the daily and weekly review
Pharm West	• GTO not effectively involved • Only two tier levels • Monthly review of yield data not enough	• Involve GTO in tier process • Specify the a more detailed tier process • Standardize the daily and weekly review

Table 6: Gap Analysis

8.3. Standard Business Process

A standard business process has been proposed to improve the communication of yield data based on the analysis of this company which is intended to improve the communication of yield data. Considering all the problems, requirements and actual required customers discussed previously, the proposed Data Flow Diagram is plotted in Figure 14. The implementation of the proposed business process requires further development of the tier process at the west campus.

The improved business process begins as technicians weigh the materials. The batch sheets should be completed and should include all the detailed information of the actual/release yield, the accountable yield, and sampled and reject losses. The technicians are also responsible for inputting the actual/release yield to a central database, be it SAP, Data 3 or any other software with integrated functions of both material and financial management. In addition, if there are any out-of-specifications found in the measurement, technicians should report to the upper management by editing issue boards or raising NOE. The batch sheets are reviewed by a designated person on the production site, be it the production supervisor or the document coordinator; the batch sheets are then passed on to the quality department. If the batch is approved for release, the batch release status will be changed in the central database.

The yield data stored in the central database are exported by both GTO and CI engineers or their counterpart Operation Excellence (OE) Leader. Here the process engineers are considered under GTO. CI/OE are the major users of the central database because they are in charge of generating the yield metrics as well as calculating the yield variance and providing the metrics and yield variance to GTO, Finance, Quality, and IPT Leader who can also be called as Operation Director. To calculate the yield variance, they also need the target/standard yield determined by the production site, typically the IPT Leader (Operation Director). GTO should utilize

both the raw data directly from the central database and the summarized yield metrics from CI to analyze a specific process in real time and to monitor the yield trends.

The Issue Board and NOE should first be submitted to the IPT Leader. After the IPT Leader identifies the area to which the problem belongs, they will cooperate with those related departments to resolve the problems. This whole process will be recorded in some software like Trackwise and retrieved every year to be captured in APR.

Theoretically, a centralized data system should be able to process both the yield information and the product release information. However, if there is any inconsistency caused by mismatch of these two sources of information, the actual/release yield information should be captured first in the metrics with the unreleased batches reprocessed. The reprocess procedure consists of the correction the financial yield, the submission of a report including the particulars of the unreleased batches to from IPT Leader or Operation Director and the decision on engineering improvement from IPT Leader or Operation Director. In this case, the concept of postponement is applied in this revised business process. In further improvement of this business process, the algorithm of the data systems and the interfaces between the material management system and the financial management system can be further investigated to isolate the production release information from the yield data in metrics.

Comparing to the current business processes, the revised business process helps save labor hours by employing a centralized data system. The customers are prioritized in the data flow according to their responsibilities and activities by utilizing a complete tier process. The closer their job is to the manufacturing processes, the more frequently they are able to access the yield data. For instance, the technicians are the closest to the production and thus deal with the raw data on a tier 1 shift basis. On the contrary, the IPT Leader/Operation Director reviews the summarized yield metrics on a tier 4 weekly or monthly basis. The most significant improvement is that the various

sets of yield data are harmonized to the release/actual yield. Furthermore, this unified process is a standard process that can be applied to most of the in-line products at both campuses and in both types of facilities. Finally, since the revised business process is developed based on the analysis of currently most representative products, it can be applied to future products of similar type without modification. Even if they are of completely different types with distinct yield definitions, measurements and calculations, the revised business process can also be generalized with a few modifications.

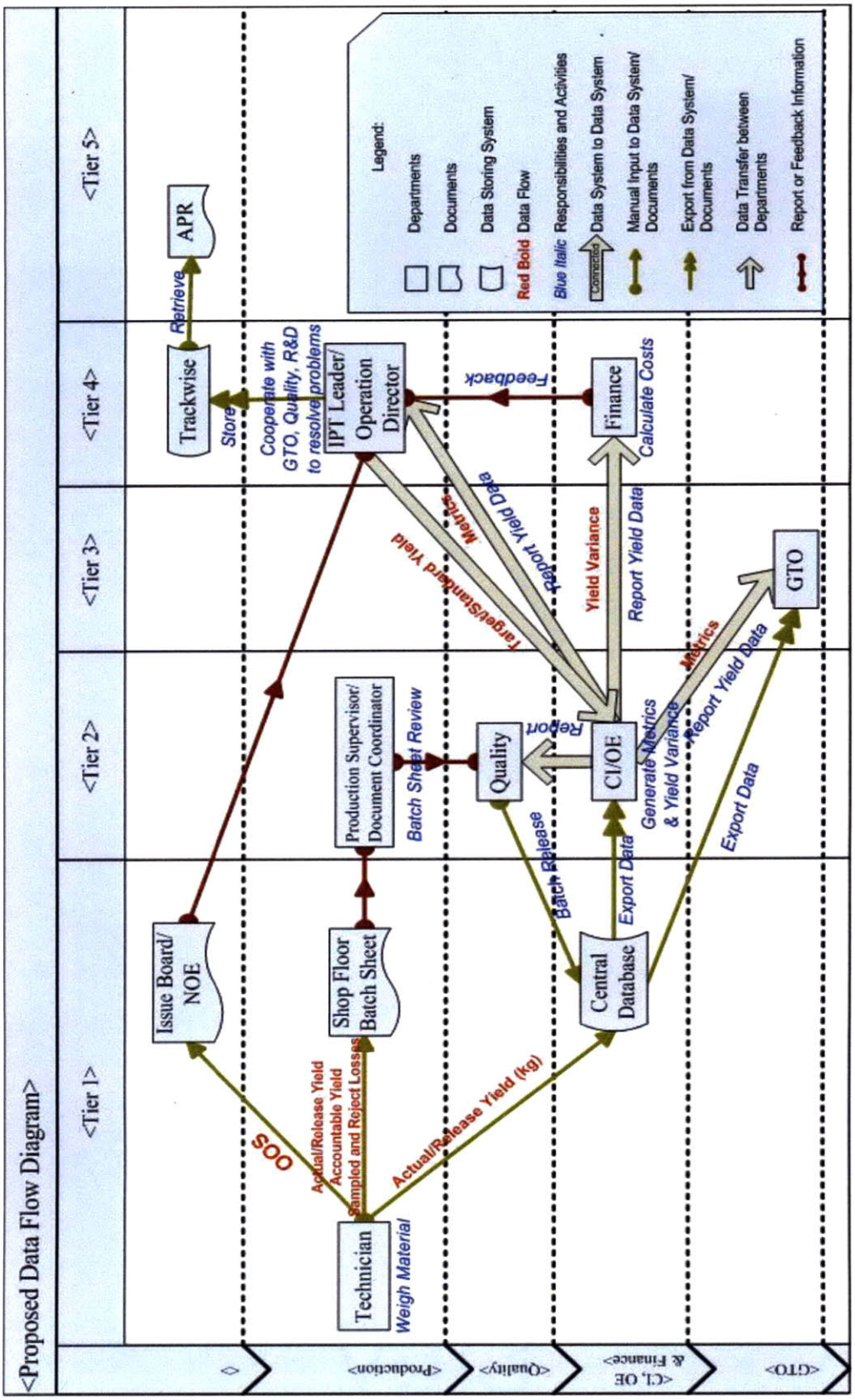

Figure 14: Proposed Data Flow Diagram

9. CONCLUSION

The current yield data flow at SJP Singapore expose a series of problems including:

1) Data inconsistencies
2) Invisible data flow
3) Workflow redundancies
4) Data flow inefficiencies

Through the study of the current business process from the raw data collection in the batch sheets to the data flow through all relevant customers by taking the representative production lines as examples, the yield concepts at each facility of this company were defined and clarified and the data flow was eventually visualized in four data flow diagrams. Meanwhile, the root causes for the problems mentioned previously were identified and summarized as listed.

1) The data inconsistencies are caused by accountable yield overestimation, inconsistent names and meanings, mismatch of yield standards, multiple data systems and inconsistency between operational and financial yield.

2) The workflow redundancies and data flow inefficiencies are caused by incomplete tier process and inconsistent data used at different departments.

Recommendations were proposed to each facility to solve their individual problems while all these recommendations added up are also the requirements to fill the gap to the desired procedure. A new business process was then developed and proposed which is intended to solve the problems from the root causes thus satisfying the needs of the customers identified. Moreover, it is a unified standard business process that can be applied to both pharmaceutical and API facilities at both the south and west campuses.

As a result, all the goals stated previously are achieved, namely:

1) This revised business process can help ensure the yield data accuracy by unification of different yield concepts.

2) This business process can improve efficiency by postponing the data divergence, suggesting a centralized data system and prioritizing the data access by customer responsibility.

3) This business process is visible so that the data reliability can be justified.

4) This business process is a general structure applicable to different products and a standard process applicable to different campuses.

10. FUTURE WORK

Considering the steadily growing pharmaceutical market and the leading position of this pharmaceutical company, diverse new products are expected to be developed and manufactured. The revised standard business process can be further updated taking into account the characteristics of the new products.

The paper document may be eliminated by further research on utilizing the current automation systems such as DCS and PI to store the yield data so that the batch information can be retrieved anytime by multiple departments instead of being kept in archive and losing its value.

Beside yield, other data such as discards and costs can also be investigated and benefit from an improved business process for the data flow, because a good business process is not only able to maximize the operational efficiency, but also enhance data accuracy. Moreover, a standard process assures compliance of CGMP and thus further promotes quality excellence.

The minimization of departments involved in a certain task such as yield processing and minimization of the labor hours required in each department can substantially improve the performance thus bringing considerable value to the company. Therefore, beside the general tier process, the operational structure should be examined and further optimized in terms of different tasks.

The investigation methodology employed in this project can be applied to other analyses. The steps of selecting representative production lines and generalizing manufacturing processes saved approximately 70% of the time for data tracking. In addition, the data flow diagrams provided the visibility of the current status and the root causes of the problems. Finally, the new business process was established by gap

analysis. Therefore, these critical approaches and tools can be applied to future study for continuous improvement.

APPENDIX – GLOSSARY

API: Active Pharmaceutical Ingredient

API South: the API facility at the south campus

API West: the API facility at the west campus

APR: Annual Product Review

CI: Continuous Improvement

Data 3: It is used for material management at the south campus of SJP Singapore. All the production and transfer quantities are recorded in this system.

DCS: Distributed Control System

DMAIC: Define-Measure-Analyze-Improve-Control

GTO: Global Technical Operation

IPT: Integrated Process Team

JDE: It is financial management software at the south campus to calculate yield variance and related costs.

NOE: Notice of Event

OOS: Out-of-specification

Pharm South: the pharmaceutical facility at the south campus

Pharm West: the pharmaceutical facility at the west campus

PI: Plant Information

SAP: It is the counterpart business management system of Data 3 at the west campus of SJP Singapore.

SIPOC: Supplier-Input-Process-Output-Customer

Trackwise: It is quality management software.

REFERENCE

[1] FDA, 2009, "Pharmaceutical cGMPS for the 21st Century - A Risk-Based Approach: Second Progress Report and Implementation Plan," http://www.fda.gov

[2] Wipro-SAS, 2009, "Manufacturing Quality Analysis for Pharmaceutical Industry," www.sas.com

[3] IMS, 2010, "IMS Forecasts Global Pharmaceutical Market Growth of 5-8% Annually through 2014; Maintains Expectations of 4-6% Growth in 2010," http://www.imshealth.com

[4] SJP, 2010, Internal Material.

[5] George, M. L., Rowlands, D., Price, M., and Maxey, J., 2005, *The Lean Six Sigma Pocket Toolbook,* McGraw-Hill, New York, Chap. 1.

[6] Nunnally, B. K., 2007, *Six Sigma in the Pharmaceutical Industry: Understanding, Reducing, and Controlling Variation in Pharmaceuticals and Biologics,* CRC Press.

[7] Lynch, D. P., Bertolino, S., and Cloutier, E., 2003, *How to Scope DMAIC Projects,* Quality Progress, pp. 37-41.

[8] Basu, R., 2009, *Implementing Six Sigma and Lean: A Practical Guide to Tools and Techniques,* Butterworth-Heinemann.

[9] Kemsley, S., 2007, "Business Process Discovery," http://www.tibco.com

[10] van der Aalst, W., and van Hee, K., 2004, *Workflow Management: Models, Methods, and Systems,* MIT Press, Cambridge, MA.

[11] Li, Q., and Chen, Y. L., 2009, *Modeling and Analysis of Enterprise and Information Systems.* Higher Education Press, Springer, Beijing, Berlin, Heidelberg, New York.

[12] Jeston, J., and Nelis, J., 2006, *Business Process Management: Practical Guidelines to Successful Implementations,* Butterworth-Heinemann, Oxford.

[13] Bolstorff, P., and Rosenbaum, R., 2007, *Supply Chain Excellence: A Handbook for Dramatic Improvement Using the SCOR Model,*AMACOM, New York.

[14] CSE Global, http://www.cse-global.com/pdf/PIS_I.pdf

[15] Superior & Co., 2010, "Patient Information about Vxxx," Internal Material.

[16] Monson, K., and Schoenstadt, A., 2010, http://cholesterol.emedtv.com/vxxx/vxxx.html

[17] FDA, 2008, http://www.fda.gov

[18] Superior & Co., Inc., 2010, http://www.naxxx.com

[19] SJP Singapore, 2009, "Tier Process," Internal Material.