Identifying Technical Inefficiencies and Quality Concerns in Institutional Pharmacies of Private Hospital Groups using Data Envelopment Analysis

by

Nina Uys

Thesis presented in partial fulfilment of the requirements for the degree of Master of Engineering in the Faculty of Engineering at Stellenbosch University

Study Leader: Prof. J Bekker

Decembr 2015
Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

September 2015
Abstract

Private health care service providers continuously strive to find a balance in providing quality patient care while being cost-effective. This balance serves the interest of both the patient and the profit-driven organisations providing these services. Lower costs result in lower service fees, which is advantageous to organisation market share and patient medical care costs.

Institutional pharmaceutical services (i.e. those provided in a hospital) differ from other in-hospital medical specialities in that the hours of pharmacists and pharmacist’s assistants are not billed to individual patients, but are rather absorbed in the operational cost of a hospital. Improving the performance of the institutional pharmacy can thus directly affect a hospital’s bottom line.

The problem is that identifying performance improvement initiatives are difficult, as the factors affecting performance are non-commensurate. Case studies from the literature show that the physical environment, process design, inventory management, scheduling, and human resources management and well-being affect pharmacy performance. These factors are however not easily comparable or measurable when analysing performance.

Data Envelopment Analysis (DEA) is a frontier analysis technique used to measure the relative performance of Decision Making Units (DMUs) with common inputs and outputs. The primal and dual (and thus slack values) of the DEA linear programming problems provide insight into inefficiencies of a DMU compared to the rest of the DMUs in the set.

The aim of this study was to use DEA to identify technical inefficiencies and quality concerns in institutional pharmacies of private hospitals. This would enable pharmacy operational managers to identify underperforming pharmacies and to specify and garner financial support for performance enhancing interventions.

DEA was applied using data provided on the inputs and outputs of a private hospital group in South Africa. The measurable inputs used for the analysis included the employee hours per month, the percentage of aged stock, the number of call-outs for pharmacists per month and the number of reported incidents.
Outputs included the number of prescriptions filled for in-hospital use, discharged patients and retail customers per month.

Three DEA models, each with their primal and dual problems, were developed. Multiple models were developed to ensure that results were reasonable and consistent across the various models for verification and validation purposes. Two more models were developed to perform sensitivity analysis on model results.

The DMU results were related back to the case studies from the literature by interpreting the results of three example DMUs in the set. This gave context to the results and illustrated how to identify possible actionable plans for improvement initiatives.

As DEA only provides insight into how DMUs perform relative to each other, knowledge on how to improve the group of pharmacies continuously so as to remain competitive in a global context is also required. The literature on continuous improvement is presented, with case studies relating to the implementation of process improvement techniques and advanced pharmacy technologies. These studies are presented to be implemented in pharmacies already rated fully efficient through DEA, so as to continuously improve the standard for relative performance.
Private gesondheidsorg-dienste streef deurentyd na 'n balans tussen die verskaffing van gehalte-pasiëntesorg en hoe om koste-effektief te bly. So 'n balans bevoordeel die pasiënt se belange en dié van die winsgedrewe organisasies wat hierdie dienste lewer. Laer koster lei tot laer diensfooie, wat voordelig is vir organisatoriese markeandeel asook pasiënte se gesondheidsorg-koste.

Institusionele apteekdienste (d.i. dié wat in hospitale gelever word) verskil van hospitale se ander mediese spesialisdienste deurdat 'n pasiënt nie vir aptekers en hul assistente se dienstye betaal nie maar dat die hospitaal se operasionele koste hierdie uitgawes absorbeer. Beter werkverrigting in die institusionele apteek raak dus die hospitaalbegroting direk.

Die probleem is dat inisiatiewe vir beter werklewering moeilik uitkenbaar is want die faktore wat prestasie beïnvloed, is onvergelykbaar. Volgens gevallestudies uit die literatuur word apteekprestasie geraak deur die fisieke omgewing, prosesontwerp, voorradebestuur, skedulering, menslike hulpbronne en die algemene welstand van personeel. Vir prestasie-ontleding is dié faktore egter nie maklik vergelykbaar of meetbaar nie. Data Omvattings-Ontleding (DOO) [[Data Envelopment Analysis, DEA] is 'n voorpunt-ontledingstegniek waarmee Besluitnemingseenhede (BNE's) [Decision Making Units: DMU's] wat gemeenskaplike in- en uitsette het, se relatiewe prestasie gemeet word. Die DOO linieêre program se verwaarmste, tweeledige (en dus spelings-) waardes bied insig in die ondoeltreffendhede van 'n BNE teenoor die res van die BNE's in die stel voorbeelde.

Dié studie se mikpunt was om DOO te gebruik om tegniese ondoeltreffendhede asook die gehalte waaraar daar kommer bestaan in private hospitale se institusionele apteke te meet. Dit kan apteke se operasionele bestuurders in staat stel om onderpresterende apteke uit te ken en om, vir ingryping ter wille van beter prestasie, finansiële steun te spesifiseer en te bewillig.

DOO is toegepas deur die gebruik van data wat oor die in- en uitsette van 'n private hospitaalgroep in Suid-Afrika verskaf is. Die meetbare insette wat vir die ontleding gebruik is, het ingesluit die tyd wat die werknemers per maand
gewerk het, die persentasie verouderde voorraad, hoeveel keer elke apteker elke maand uitgeroep is en hoeveel voorvalle aangemeld is. Uitsette het ingesluit die getal voorskrifte wat maandeliks vir hospitaalgebruik, ontslag-pasiënte en kleinhandel-kliënte ingevul is.

Drie DOO-modelle is ontwikkel, elkeen met sy vernaamste én tweeledige probleme. Veelvuldige modelle is ontwikkel om te verseker dat resultate vir kontrole en bekragtiging dwarsoor die onderskeie modelle redelijk en konsekwent was. Om sensitiwiteitsontledings op modelresultate uit te voer, is nog twee modelle ontwikkel. Die DOO-modelle is met die gevallestudies uit die literatuur in verband gebring deur die vertolking van die resultate van drie uit die stel DOO-voorbeelde. Dit het die resultate in konteks geplaas en geïllustreer hoe werkbare planne vir verbeterings-inisiatiewe geïdentificeer kan word.

Omdat DOO net insig bied in hoe BNE’s in verhouding tot mekaar presteer, is kennis ook nodig oor hoe om die apteekgroep deurlopend te verbeter sodat dit gloaal mededingend bly. Die literatuur oor deurlopende verbetering is dus aangebied, saam met gevallestudies wat verband hou met die inwerkingstelling van tegnieke vir prosesverbetering en vir gevorderde apteektegnologieë. Dié studies is aangebied sodat dit in werking gestel kan word in apteke wat reeds deur DOO as ten volle doeltreffend geëvalueer is, om die vlak van relatiewe prestasie voortdurend te verbeter.
Acknowledgements

I would like to thank:

• **Prof James Bekker** for his guidance and patience.

• **Manda** and **Dirk** for their time and effort.

• All the pharmacists, pharmacist’s assistants, runners, interns and nurses for patiently answering all my questions, and for allowing me to observe them at work.

• **Lafras** for introducing me to Python, and then for teaching me the difference between 3 and 3.0.

• **Hans** for copy-editing this document.

• My other friends, family and colleagues, especially **Theo** for food from an oven, my parents for encouraging a continuous education, **Magdalena** for being a sterling example of how to finish a thesis and **Stefan** for his witticisms.

• My husband **Pieter**, who truly lives up to his name.
## Contents

Declaration i

Abstract ii

Opsomming iv

Acknowledgements vi

Nomenclature vii

1 Introduction 1
   1.1 Background of the Study .............................. 1
   1.2 The Research Aim and Objectives .................... 3
   1.3 Research Methodology ................................. 3
   1.4 Document Structure .................................. 4

2 Factors Affecting Pharmacy Performance 5
   2.1 Introduction ........................................... 5
   2.2 Physical Environment ................................ 7
       2.2.1 Facilities Design ................................. 7
       2.2.2 Workstation and Tool Design .................. 10
       2.2.3 Environmental Factors .......................... 11
   2.3 Process Planning ...................................... 12
       2.3.1 Task Allocation and Sequencing ................ 12
       2.3.2 Workload and Performance .................... 13
       2.3.3 Interruptions .................................. 14
   2.4 Inventory Management .................................. 15
   2.5 Scheduling ............................................ 16
   2.6 Human Resource Development, Management and Well-being 18
   2.7 Summary ................................................. 20
# Measuring Service Performance

## Introduction

## Efficiency Measuring Techniques

### Accounting Techniques

### Balanced Scorecard

### Frontier Efficiency Methods

## Data Envelopment Analysis

### Background

### DEA Models and Definitions

### Sensitivity Analysis of DEA Inputs and Results

### Considerations for DEA in Healthcare

### DEA Health Care Case Studies

## Summary

# Applying Data Envelopment Analysis to Institutional Pharmacies

## Introduction

## DEA Research Question

## Institutional Pharmacy Process Model

## DEA Inputs, Outputs and Attributes

## Input Data Analysis

## Data Envelopment Analysis Models

### DEA Model Development

### DEA Overall Results

## Sensitivity Analysis of Results

### Modelling Sensitivity Analysis

### Examples of DMU Results Interpretation

## Validation of Analysis

## Summary

# Continuous Improvement

## Continuous Improvement Methodologies

### Plan Do Study Act with Root Cause Analysis

### Lean Thinking

### Six Sigma

### Lean Six Sigma

### Theory of Constraints

## Pharmacy Technology Advancements

### Electronic Prescription Systems

### Automatic Counting Systems

### Robotics

# Stellenbosch University

https://scholar.sun.ac.za
List of Figures

2.1 Factors affecting pharmacy performance. . . . . . . . . . . . . . . . 7
2.2 Pharmacy shelving. . . . . . . . . . . . . . . . . . . . . . . . . . . 11
3.1 Frontier analysis methods used in health care studies. . . . . . . . 25
3.2 Simplified pharmacy DEA example: Inputs versus outputs. . . . 27
3.3 Simplified pharmacy DEA example: Changed efficiency frontier. . 27
3.4 Returns to scale frontiers. . . . . . . . . . . . . . . . . . . . . . . . 31
3.5 DEA radius of stability. . . . . . . . . . . . . . . . . . . . . . . . . 34
4.1 Night cupboard inventory control process. . . . . . . . . . . . . . 41
4.2 Measurable inputs and outputs of pharmacy processes for DEA . 44
4.3 DEA model source code flow diagram. . . . . . . . . . . . . . . . 45
4.4 CCR model study results. . . . . . . . . . . . . . . . . . . . . . . . 47
4.5 BCC model study results. . . . . . . . . . . . . . . . . . . . . . . . 48
4.6 Additive model study results. . . . . . . . . . . . . . . . . . . . . 48
4.7 DMU data analysis flow diagram. . . . . . . . . . . . . . . . . . . 55
List of Tables

3.1 Simplified pharmacy DEA example: Data. . . . . . . . . . . . . . . . 26

4.1 DMU j=27 results. . . . . . . . . . . . . . . . . . . . . . . . . . . 50
4.2 DMU j=19 results. . . . . . . . . . . . . . . . . . . . . . . . . . . 51
4.3 DMU j=43 results. . . . . . . . . . . . . . . . . . . . . . . . . . . 52
Nomenclature

List of Latin Symbols

\( m \)  Total number of inputs per DMU
\( n \)  Total number of DMUs
\( s \)  Total number of outputs per DMU
\( s_i^- \)  Input \( i \) slack value
\( S_i^- \)  Non-linear additive model input \( i \) slack value
\( s_r^+ \)  Output \( r \) slack value
\( S_r^+ \)  Non-linear additive model output \( r \) slack value
\( t \)  Non-linear additive model transformation variable
\( u \)  Output weight variable
\( v \)  Input weight variable
\( x_{io} \)  Input value \( i \) for DMU\(_o\)
\( y_{ro} \)  Output value \( r \) for DMU\(_o\)
List of Greek Symbols

\( \delta \)  
Radius of Stability

\( \epsilon \)  
Non-Archimedean element

\( \lambda \)  
Intensity variable

\( \Lambda \)  
Non-linear additive model intensity variable

\( \tau \)  
Linear additive model efficiency score

\( \theta \)  
Envelopment model efficiency score

List of Acronyms

BCC  
Banker, Charnes and Cooper

CCR  
Charnes, Cooper and Rhodes

CI  
Continuous Improvement

CRS  
Constant Returns to Scale

CSV  
Comma-Separated Values

DEA  
Data Envelopment Analysis

DES  
Discrete Event Simulation

DMU  
Decision Making Unit

DMAIC  
Define, Measure, Analyse, Improve, Control

DPMO  
Defects per Million Opportunities

ENT  
Ear, Nose and Throat
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDL</td>
<td>Essential Drug List</td>
</tr>
<tr>
<td>ESI</td>
<td>Emergency Severity Index</td>
</tr>
<tr>
<td>df</td>
<td>degrees of freedom</td>
</tr>
<tr>
<td>FIFO</td>
<td>First In First Out</td>
</tr>
<tr>
<td>FTEE</td>
<td>Full-Time Employee Equivalent</td>
</tr>
<tr>
<td>GPP</td>
<td>Good Pharmacy Practice</td>
</tr>
<tr>
<td>ICU</td>
<td>Intensive Care Unit</td>
</tr>
<tr>
<td>JIT</td>
<td>Just in Time</td>
</tr>
<tr>
<td>LP</td>
<td>Linear Programme</td>
</tr>
<tr>
<td>PDSA</td>
<td>Plan Do Study Act</td>
</tr>
<tr>
<td>RCA</td>
<td>Root Cause Analysis</td>
</tr>
<tr>
<td>SFA</td>
<td>Stochastic Frontier Analysis</td>
</tr>
<tr>
<td>TLS</td>
<td>Theory of Constraints, Lean and Six Sigma</td>
</tr>
<tr>
<td>TOC</td>
<td>Theory of Constraints</td>
</tr>
<tr>
<td>TQM</td>
<td>Total Quality Management</td>
</tr>
<tr>
<td>TTO</td>
<td>To Take Out</td>
</tr>
<tr>
<td>TPS</td>
<td>Toyota Production System</td>
</tr>
<tr>
<td>VMCA</td>
<td>Veteran Administration Medical Center</td>
</tr>
<tr>
<td>VRS</td>
<td>Variable Returns to Scale</td>
</tr>
<tr>
<td>WWU</td>
<td>Workload Weighted Unit</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Background of the Study

Joseph P. Newhouse, a noted health policy and management economist, states (2002) that:

“Despite the lack of a summary measure of its efficiency, many seem convinced that the (health care) industry’s performance falls short.”

In South Africa, the private health care industry is dominated by three large hospital groups, which cater mostly to middle to high income citizens with private medical scheme subscriptions. The Council of Medical Schemes reported in 2008 that in-hospital cost of private health care has increased by 8.1% annually over a period of seven years, thus more than is expected from inflation alone (Office of the Registrar of Medical Schemes, 2008). They attributed this as the most significant contributor to rising medical scheme costs. With the possibility of the introduction of private hospital health care cost regulations, the need to become more efficient is all the more a necessity for the current market stakeholders.

Health care service providers have to strive to find a balance though in providing quality patient care while being cost-effective. This balance serves the interest of both the patient and the profit-driven organisations providing these services. Lower costs result in lower service fees, which is advantageous to organisation market share and patient medical care costs.

Institutional pharmaceutical services (i.e. those provided in a hospital) differ from other in-hospital medical specialities in that the hours of pharmacists and pharmacist’s assistants are not billed directly to individual patients, but are rather absorbed in the operational cost of a hospital. Improving the performance of the institutional pharmacy can thus directly affect a hospital’s bottom line.
In South Africa, pharmaceutical services and facilities are regulated by the Good Pharmacy Practice (GPP) (2010) and the Medicine Act (2014). These standards and the act do not only demand safe and reliable service of its practitioners, but also a commitment to providing cost-effective services. The standards include those with clear compliance measures, such as the minimum work space for pharmacists dispensing areas and medication schedule classifications. However, it also includes requirements for the presence of error prevention measures and adequate working conditions, the presence of which are easily monitored but the efficacy of which are subjective. Any intervention to improve the cost-effectiveness of pharmaceutical services is at the very least subject to these restrictions.

The success of an improvement initiative on a system is measured by how much better the system performs its objective after implementation. We define the objective of an institutional pharmacy, i.e. a pharmacy in a hospital, to store and dispense medicine in a hospital in a safe and cost-effective manner while conforming to legislative requirements. All processes performed in the pharmacy support this objective.

Within this objective there is a prioritised list of goals. The highest priority is that the pharmacy must always comply to the GPP and other legislative standards — ranging from medicine storing temperatures to minimum pharmacy work space specifications. These specifications have clear compliance measures, either being correct or incorrect. The next priority addresses how well those specifications are met. Examples include the efficacy of error prevention processes during dispensing and accurate stock control to ensure the required medication does not age and is available when needed. The lowest priority is to ensure that the service is cost-effective. This is realised by exploiting the potential of all resources in the system.

The problem is that identifying performance improvement initiatives are difficult, as the factors affecting performance are non-commensurate. Factors range from subjective quality perceptions to optimal inventory levels to human resource scheduling — all important factors, but not easily comparable.

Operational pharmacy managers of hospital groups are however faced with this problem every day. They have to identify underperforming pharmacies in the group, determine root causes for poor performance and pinpoint improvement initiatives to increase performance.

Data Envelopment Analysis (DEA) is a technique to measure the relative performance of Decision Making Units (DMUs) with common inputs and outputs (Cooper et al., 2004). The use of it in private institutional pharmacies can be beneficial to identifying relative inefficiencies with high confidence levels. This can enable operational managers to specify and garner financial support for performance enhancing initiatives.
1.2 The Research Aim and Objectives

The problem, as detailed in the previous section, is that factors affecting the quality and effectiveness of pharmaceutical services are difficult to measure, compare and prioritise for improvement initiative purposes. The use of DEA can be beneficial to increasing the confidence levels that operational pharmacies require to specify and budget for improvement initiatives.

The research aim is formulated as follows:

Identify technical inefficiencies and quality concerns in the institutional pharmacies of South African private hospital groups via DEA for improvement initiatives.

The following objectives have been set to meet this research aim:

1. Determine the factors that influence the performance of institutional pharmacies through observation and literature study.

2. Map the institutional pharmacy inputs, outputs and processes through observation and interviews with health care professionals.

3. Apply and test DEA models and sensitivity analysis techniques in an institutional pharmacy case study.

4. Determine which continuous improvement methodologies can be used to improve pharmacies that already have 100% relative efficiency according to the DEA evaluation.

1.3 Research Methodology

An exploratory research design is followed, using both qualitative and quantitative methods. Qualitative data are gathered through site observations and interviews to obtain a general sense of the workings, elements and issues concerning institutional pharmacies. This will provide a basis for a systematic quantitative DEA case study, performed on one of the private South African hospital groups. Validation will be performed both quantitatively (through sensitivity analysis of results) and qualitatively (through interviews with industry stakeholders).

The qualitative observation and DEA case study are only performed in one of the South African private hospital groups. It is anticipated that this will not hamper the global applicability of applying DEA in South African private hospital pharmacies, as all of the groups work within the constraints of the GPP and thus have similar goals, constraints and required processes.
1.4 Document Structure

In Chapter 2, an overview of the literature regarding factors affecting pharmacy performance is presented. The factors are grouped into five subcategories, namely the physical environment, process, inventory management, scheduling, and human resource management and well-being.

The various efficiency measurement techniques of the service industry is discussed in Chapter 3, with a focus on DEA. The literature on different DEA models, sensitivity analysis of results and health care application case studies are discussed.

In Chapter 4, DEA is applied to the institutional pharmacies of a South African private hospital group. The pharmacy process maps are provided, detailing the required resources and outputs of these processes. Thereafter the various DEA models’ logic is shown. Sensitivity analysis is performed on the results of the models, and the results, with validation input of its applicability from industry stakeholders, are presented.

In Chapter 5 the literature on continuous improvement methodologies is discussed for use in pharmacies rated relatively fully efficient during DEA.

Finally, in Chapter 6 the various literature studies and evaluations are summarised. The research value is discussed, conclusions are drawn on how the research objectives were met and possible future work is discussed.
Chapter 2

Factors Affecting Pharmacy Performance

2.1 Introduction

The objective of an institutional pharmacy is defined as storing and dispensing medicine in a hospital in a safe and cost-effective manner while conforming to legislative requirements. A complex system including pharmacists, pharmacist’s assistants, runners (who deliver medication to wards), facilities, computer systems, inventory and management is integrated to achieve this goal.

Prescriptions, consisting of one or many medication line items, are ordered by patient attending physicians. These are captured on a patient’s physical file. The files are periodically collected during the day by the pharmacy runners, and delivered to the pharmacy.

In the pharmacy, the prescriptions are sorted by priority, with Intensive Care Unit (ICU) patients having the highest priority, followed by discharged patients (To Take Out (TTO) medication) and then patients from lesser critical care wards. In the pharmacy, the prescriptions are filled by pharmacists or their assistants, though the GPP (2010) dictates that the final checking of the prescription must be performed by a pharmacist.

Checks include that the correct medication and dosage has been dispensed and that the label correctly shows the medication name, storage and usage instructions. Pharmacists also take into account whether the prescribed medication can be taken by the specific patient, for example that the medication will not cause an allergic reaction or that the dosage meets patient age and ailment requisites. Any queries are raised with the prescribing physician to approve before applying changes.

The filled prescriptions are then delivered with the patient files to the various...
wards by the runners. In the case of TTO medications, patients collect from the pharmacy and receive counselling by a pharmacist on the use of prescribed medication.

Throughout the day physicians may also require “stat” medication, where pharmacists are requested to dispense medication immediately upon demand. Examples include when a patient has left the operating room and require immediate medicinal treatment, or in the case of an emergency.

Although all medications are ordered, received and managed by the pharmacy, select medications are stored throughout the hospital in wards, emergency carts and the night cupboard — an access controlled area for use by night nurses outside of pharmacy trading hours. The responsibility for inventory control, delivery and storage to these areas lie with the pharmacy manager.

Additional processes include generating various reports, such as monthly administration reports, semi-annual stock take reports and incident reporting. Pharmacist’s assistants generally order, receive and pack away medication, following a First In First Out (FIFO) system. According to the Medicine Act (2014) though, a pharmacist has to perform these inventory control processes for schedule 6 medication. Storage, handling and incident reporting of schedule 6 medication is highly regulated and time-consuming. The flow diagram of the business process under consideration in this thesis can be seen in Appendix A.

Optimising these procedures are beneficial to the hospital organisation, patients and the pharmacy itself. With the great number of resources that influence these processes, ample opportunity for improvement exists.

From observation and the literature, the factors affecting the performance of a pharmacy can be grouped into five subcategories, namely:

1. **Physical environment** — including facilities layout, workstation design and environmental factors such as illumination, ambient noise, temperature and ventilation.

2. **Process design** — such as the workload per employee, the number of distractions and interruptions experienced by pharmacists, as well as task assignment and sequencing.

3. **Inventory management** — considering reorder level policies and usage monitoring.

4. **Scheduling** — such as the number of personnel required and shift definition.

5. **Human resource management and well-being** — including management styles and employee job satisfaction and commitment.
These factors are discussed in the sections below (as per Figure 2.1). Each subsection details the relevant legal requirements, the case studies showing how each factor influences pharmacy performance and possible improvement interventions.

![Diagram of Pharmacy Performance](Figure 2.1: Factors affecting pharmacy performance.)

Other items not mentioned above may also be influential, such as the computer systems, tertiary education levels and administrative processes and forms. For the purpose of this study though these factors are disregarded as they are consistent in a hospital group that operates in one country. Only factors that differ from pharmacy to pharmacy were considered in this study.

The factors are subsequently discussed.

### 2.2 Physical Environment

#### 2.2.1 Facilities Design

Facilities analysis and design is concerned with improving material handling through the effective use of employees, equipment, space and energy. This includes increasing the economical use of available space, optimisation of flow of operations and minimising lead times while ensuring that the facility is flexible and adaptable. This all has to be done whilst being cognisant of legal requirements which, in the pharmaceutical industry, concerns the GPP (2010) and Medicine Act (2014). (Tomkins et al., 2003; McDowell and Huang, 2012)

In a pharmacy environment flexibility would indicate that all the possible hospital pharmaceutical requirements can be met by the pharmacy, such as temperature control for select medication and space to allow for dilution of medication as prescribed. Adaptability would indicate that the pharmacy is equally functional regardless of variation, such as calender cycles and peak times. It also indicates that
a pharmacy can be upgraded easily to allow for new technologies and equipment. The GPP (2010) has certain minimum requirements for institutional pharmacy facilities, such as:

- The design and layout of the pharmacy must allow for a logical flow in operations.
- The risk of cross-contamination and other errors must be minimised.
- Entrances, dispensing counters and patient consulting areas must be wheelchair accessible.
- The external appearance of the pharmacy must reflect the professional nature of the health care service as to inspire patient confidence.
- Certain signage and name-tag displays are required.
- The pharmacy area must have access control and adequate security.

Changes to a facility’s layout can be expensive and disruptive. In a health care environment where error prevention is prioritised, changes to procedures have to be well evaluated before implementation. High confidence levels are thus required to ensure that improvement initiatives will in fact increase performance. To this end, Discrete Event Simulation (DES) is often employed to test the expected improvements from various design alternatives.

Various case studies exist where simulation was used to evaluate not only alternative pharmacy layouts (McDowell and Huang, 2012; Lin et al., 1996), but also staffing levels (Reynolds et al., 2011; Al-Hawari et al., 2011; Rust et al., 2012; Hong et al., 2012), processes (Dean et al., 1999; Lin et al., 1996) and usage of automatic robotic dispensing systems (Reynolds et al., 2011). None of the results from the evaluations were however implemented in any of these case studies. There is thus still a gap in the literature for studies showing how to implement changes as evaluated by simulation studies whilst causing minimal disruption and evaluating the impact of these changes after implementation.

An example of a DES study is that of McDowell and Huang (2012). The authors used a weighted scoring system to evaluate various institutional pharmacy layouts considering factors such as feasibility, cost, patient and employee safety, robustness and others. Designs were developed through mapping activity flow charts, drawing communication and material relationship charts, defining space requirements and through interviews. These factors were then evaluated by simulating the various design options. The weighted scoring system did identify a clearly favourable design.

The problem with the simulation though is that the evaluated factors were not all mutually exclusive — for example the “feasibility” considered cost, but “cost”
was evaluated as a separate factor as well. This may have resulted in some factors carrying more weight than originally intended.

Both the McDowell and Huang (2012) and the Lin et al. (1996) case studies provided alternative layouts that align more with work patterns. Examples include placing fast moving items within easy reach of pharmacist workstations and providing quick and easy access for pharmacist to patient consulting areas. This reduces pharmacist travelling time, and the advantages regarding reducing lead time are clearly highlighted through simulation.

Certain facility design factors are however more difficult to model in simulation. In their evaluation on the impact of distraction and interruptions on prescription error rates, Flynn et al. (1999) found that the facility-related distractions did correlate with slightly increased error-rates. An example was a pharmacist continuously glancing up at a passer-by through the pharmacy window.

Other factors, such as perceived spaciousness, architectural design and other interior design factors also have been shown to have a positive psychological effect on health care employees which can improve their efficiency and decrease staff turnover (Mourshed and Zhao, 2012).

In their study on work-place stressors of Northern Ireland community pharmacies, McCann et al. (2009) found that the open workplace can be causal. Pharmacists felt pressured to fill a prescription as quickly as possible while a patient could see them at work. They reported that this discouraged thorough consideration and discussion amongst colleagues of possible pharmaceutical problems related to the prescription.

Mourshed and Zhao (2012) surveyed health care providers’ perception of the effects of facility design factors on environmental interaction (which included pharmacology, administration and management tiers). They found that the cleanliness and maintainability of a facility was deemed the most important factor by the workers themselves. The subjective factors relating to interior design was deemed less important than measurable factors such as proximity to wards, air quality, thermal comfort and noise levels.

Flynn et al. (2002) found that medication storage also influences dispensing errors. In a study of nearly 6 000 prescriptions in various American pharmacies, 91 errors and 74 near errors regarding content and labelling were found. A chi-squared correlation test revealed that two thirds of the content errors occurred in pharmacies where medication was packed tightly on the shelves.

In terms of measurable and predictable improvement initiatives, aligning the work space with work processes seems, from the literature and through observation, to be the most important aspect of facility design in pharmacies. This would require prioritising a layout based on the most frequent interactions with various elements. Examples include:
placing the most prescribed medication, label printers and packaging material as close as possible to pharmacists,

making patient consultation areas quickly and easily accessible and

being centrally located in the hospital for quick stock replenishment to decentralised medication distribution areas.

2.2.2 Workstation and Tool Design

The South African Occupational Health and Safety Act (1993) states that employers have a duty to inform employees of work place hazards. These include ergonomic factors, which from the literature and through observation of the pharmacy as work place, can include postural and repetitive motion injuries. These injuries have a gradual onset caused by repeated microtrauma resulting from excessive use of poorly fitted and designed equipment (Freivalds and Niebel, 2003).

The only work station demand the GPP (2010) has is that a surface of 900mm \(\times\) 1000mm of clean workspace must be provided for each pharmacist or registered person working in the pharmacy. No ergonomic design requirements are set.

The three main intervention types to reduce work place hazards resulting in musculoskeletal injury are:

1. redesigning the tool, job or work station so as to better fit the worker,

2. training workers in how to reduce the hazard and

3. only employing those whose physical capabilities exceed the physical job demand.

In 2001 the United States Occupational Safety and Health Administration introduced new ergonomic standards to reduce the specific risk of repetitive motion injuries for employees. Chi (2001) detailed how this affects the pharmacy industry where musculoskeletal injuries such as carpal tunnel syndrome are prevalent. These standards are set to prevent injury, absenteeism and employee turnover — all of which are harmful to employees but also to pharmacy operational cost. Advice from industry experts includes utilising ergonomic tools in the pharmacy, such as height-adjustable work benches, keyboards and monitors, as well as pharmacy specific equipment like pop-up vial dispensers and high density shelving (see Figure 2.2). They thus advise intervention relating to the first type discussed above.

In an Iranian study of 211 female pharmacists, Aminian et al. (2012) found that 87.7% of them had reported at least one musculoskeletal injury in the preceding year. In another American study, Fante et al. (2007) found that repetitive motion injuries had resulted in significant lost time injury days the preceding year. They analysed the postural behaviour of the pharmacy employees, and made recommendations
on several ergonomic interventions, such as a voice activated telephone system that could be used with ease while typing on a computer.

In another intervention of oncology pharmacists in Taipei, Chou et al. (2012) redesigned the traditional needle for drawing liquid from a vial for two types of oncology medications. The new device not only resulted in increased productivity, but also significantly reduced the hand muscle soreness complaints and fatigue symptoms of pharmacy employees.

From the few available studies and the views from experts in the field, redesigning the workstation and tools have been the prevalent solution to ergonomic hazards in the pharmacy. This stands to reason as the second intervention type regarding training is mostly related to using tools correctly to prevent injury, such as with heavy lifting and wearing protective gear. The third option regarding hiring exclusions would be unconstitutional in South Africa.

2.2.3 Environmental Factors

Objective environmental factor considerations in the GPP (2010) centre mostly around medication requirements, and not those of pharmacy workers (for example maintaining the storage temperature of medication). The only requirement for personnel is that “levels of heat, light, noise and ventilation must exert no adverse effect on personnel”.

Buchanan et al. (1991) showed in a controlled study that poor illumination had a direct effect on increased error rates in an outpatient pharmacy. The significance was present for each pharmacist in the pharmacy for three different illumination levels, regardless of age or visual acuity.

Flynn et al. (1996) discuss the effect of noise and loudness on pharmacists. In a study where intermittent and ambient noise levels were controlled and pharmacist error rates monitored, no linear correlation between the various variables were
found. The error rates did correlate with increasing loudness only until a certain point, when error rates started decreasing as loudness increased. The argument could be made that employees simply become “used to” noise. Flynn et al. (1996) cite seven studies where noise also increased performance in other industries, and 29 where it decreased performance. The error rates were also inconsistent across the various employees, indicating a personal threshold and tolerance to noise pollution.

No studies on noise reduction or ventilation interventions were observed in the literature. This could be because most countries have strict standards per job type regarding environmental factors, and adherence to these may have resulted in no need for intervention.

The Occupational Health and Safety Act (Department of Labour, 1993) do have strict requirements on specific workplace thermal regulation, ventilation, limiting noise pollution and illumination levels. These can be objectively measured and regulated in a pharmacy.

2.3 Process Planning

2.3.1 Task Allocation and Sequencing

The dispensing process, as defined by the GPP (2010), is divided into three phases, namely:

1. The pharmacist interprets, evaluates and assesses the prescription. This includes considering the pharmaceutical effects of the prescription for the patient, such as contra-indications and medication interactions.

2. Preparation and labelling, which include record keeping of the supply of medication, checking for accuracy and completeness and performing medication schedule specific administration activities.

3. The patient (in TTO cases) or the medical staff must be instructed on the medication use.

The GPP (2010) requirements strictly regulate what the process must deliver, but not necessarily how the deliverables must be achieved. For example, some hospitals have a bar code on the patient file that can be scanned to populate patient data automatically for record keeping, while other hospitals require manual data capturing.

One way to evaluate the efficiency of alternative processes is with simulation, as was discussed in the facility design section (2.2.1). In the same study where Lin et al. (1996) simulated different facility layout designs, various task allocations were also evaluated. Using fixed-interval work sampling, the analysts collected data over two months, three days a week every 1.5 minutes. In this way, they
estimated within a 90% confidence interval the time pharmacists and technicians (known as assistants in South Africa) spent on various tasks.

A new work assignment system was then designed that balanced the workload and standardised the assignment of tasks between the pharmacists and the technicians. The simulation showed that this assignment better utilised both pharmacists’ and technicians’ time. The new way, according to the simulation, would reduce patient waiting time without increasing the required personnel.

In a similar case, Ghandforoush (1993) used goal programming to schedule and assign tasks to pharmacists and technicians in a hospital that was experiencing increasing demand. They showed that the increasing demand could be met without additional employees through optimal time allocation throughout a typical day.

Another case that shows how performance can be tweaked by using varying equipment and process is that of Flynn et al. (2002) who studied the effectivity of prescription inspection systems. Some pharmacies used a bar-code verification system to perform final prescription inspection, while others relied on manual inspections. A chi-squared test revealed that the bar code system found 62% of errors, 37% more that the manual system.

The available studies show that even though the required prescription tasks are non-negotiable for legal reasons, there is room to improve the efficiency and error prevention of these processes by investigating how they are performed.

2.3.2 Workload and Performance

As with human musculoskeletal functions, cognitive functions have limitations as well. The field of cognitive ergonomics studies the workplace designs to address attention distribution, information to inform decision making and the usability of technology. The field also investigates the causes and repercussions of mental load, stress and errors. (Cañas et al., 2011)

In a literature review of 60 papers on dispensing errors in the United Kingdom, United States, Australia, Brazil and Spain, James et al. (2009) found that the most common contributory factors to dispensing errors were workload, interruptions and inadequate lighting.

The authors do however caution that the term “workload” was subjectively cited as reasons for error by pharmacy workers, and the various studies may not have comparable operational definitions of this term. Most pharmacies define workload as the number of prescriptions dispensed per hour, but variation exists in the definition of “prescription”—it is either defined as one patient’s order consisting of many prescribed items, or as every line item dispensed, regardless of the associated patient. Also, workload of the pharmacy may not be equal to the workload of the pharmacist, as one pharmacist may be more efficient than his or her peers.
Another literature study on factors affecting pharmacists’ performance by Schafheutle et al. (2011) (with some overlap of papers evaluated in the James et al. (2009) study) highlighted the ambiguity and lack of consensus on the relationship between workload and performance. Some studies, notably Szeinbach et al. (2007), Kistner et al. (1994) and Bond et al. (2002) found a positive correlation between dispensing errors and prescription volume per pharmacist. Others (Beso et al., 2005; James et al., 2008; Peterson et al., 1999; Roberts et al., 2002) discussed instances where hospital pharmacists cited high workload as contributing higher error rates during subjective performance self-evaluations. The problem though with self-evaluation is that evaluators may not disclose details that reflect poorly on their perceived competence, and these studies may not reflect truly objective error causes.

Another study of 30 hospital pharmacies by Wu (2000) found no correlation between workload and error rates. Schafheutle et al. (2011) also discuss a commentary piece by A.F. Grasha who surveyed 84 institutional pharmacies in the United States and concluded that higher error rates occurred when prescription volume was low. He argues that a decline in workload causes a decrease in pharmacists’ task engagement, resulting in more dispensing errors. This argument is aligned with theories of eustress (literally meaning “good stress”) that state that engaging and challenging work environments are predictors of work success. (Hargrove et al., 2013)

There is thus a lack of objective and robust research on whether higher workload causes increased dispensing errors. Most studies cite anecdotal evidence that does not confidently support or reject the hypothesis.

2.3.3 Interruptions

Studies have shown that interruptions in a pharmacy not only reduce productivity, but also increase error rates. (Burford et al., 2011; Flynn et al., 1999; Peterson et al., 1999; James et al., 2009; Schafheutle et al., 2011)

In the GPP (2010) there are no specified measures to prevent interruptions. Undue work place interruptions can have various causes, some facility design related as discussed previously (section 2.2.1), but others due to not planning and managing the response to inevitable work-related interruptions.

The Toyota Production System (TPS), a continuous improvement methodology that addresses process re-engineering, was used to improve a pharmacy in a medical centre servicing 500 inpatient orders per day and an ad hoc retail window for the public. Sobek and Jimmerson (2003) started by creating a value stream map, focussing on the medication order (prescription) filling process. Through observation they found that the pharmacy was in violation of the third TPS rule (discussed in detail in Chapter 5), that states that the pathway for a service has
to be well specified. During high volume times orders could be handled by one or two pharmacists, have varying filling sequences and be interrupted by phone calls. Changing this in a pilot study by assigning one pharmacist to work solely on fulfilling orders in a streamlined manner while another saw to phone calls and ad hoc queries, the pharmacy’s orders-in-system and order-to-delivery time decreased by 32%.

A further improvement initiative was aimed at reducing the number of phone calls. To do this, the pharmacists started communicating exceptions to orders more clearly. For example, certain medications require refrigeration in the ward and would thus not be included in a delivery package. By adding a bright sticker on the delivery note (“Refrigerated Meds”) nurses would first look in the ward refrigerator before calling the pharmacy. These and other initiatives in standardised customer communications and streamlined processes reduced the number of calls received per day by 40%.

Reducing interruptions, or just standardising the way in which they are addressed, have thus been shown to significantly improve efficiency and prevent errors in pharmacies.

2.4 Inventory Management

The GPP (2010) stipulates that, as a minimum, adequate stock levels of the Essential Drug List (EDL) (2012) of South Africa must be maintained. This list is published by the Department of Health and includes medications to treat the most common health problems diagnosed in South African secondary and tertiary hospitals.

The GPP (2010) further requires that standard operating procedures must exist to distribute pharmaceuticals to wards, departments theatres and other storage areas. The inventory control system must be able to indicate where inventory is being kept, to ensure adequate control of expired, obsolete or recalled medications. Ensuring that all storage areas have acceptable environmental conditions and that access control falls within the responsibility of the pharmacist manager.

The choice on the variety of hospital medication inventory (called the pharmacy formulary) have various, and often conflicting, stakeholder requirements. Physicians have preferences on what medication they prescribe. These preferences are based on experience, favour a wide variety to best address individual patient needs, and are sometimes influenced by manufacturer sales representatives. Pharmacy managers on the other hand favour a limited formulary, to contain cost by benefiting from economies of scale. A smaller formulary also decreases handling, storage and administrative costs. This containment of medication cost is beneficial to the patient, but is also at odds with individualised care that physicians aim to
There are also unique medications that are very rarely prescribed and have a limited shelf-life, yet are critical to administer in a timely fashion when a patient is in need. An example is the anti-venom used to treat snake bites.

Pharmacy managers and inventory policy makers have various objectives with varying priorities when setting criteria for inventory management systems. Once these objectives have been identified, priorities have been set and constraints identified though, operations research techniques have been used to optimise inventory control policies.

Little and Coughlan (2008) built a constraint-based inventory policy model based on the knapsack problem. Constraints regarding space, delivery and criticality were considered to determine optimal stock levels. A service level objective was defined for consideration in the model. At the time of publication the authors were evaluating the model and system in a hospital. An issue that arose during implementation was a lack of quality data. The introduction of scanning technologies were being considered to improve the system’s accuracy.

In a different approach, Lapierre and Ruiz (2007) focused not on what to order, but rather the scheduling of orders. The objectives were to improve service levels and balance workload throughout the procurement cycle using a tabu search metaheuristic. This model was used to evaluate the logistics in a Montreal hospital. The authors note that the method has some drawbacks that still prevent proper implementation. They state that exact solving approaches may be required, or more thorough testing of the efficiency of the metaheuristic.

Kelle et al. (2012) used various models and methods to formulate an inventory policy in a hospital case. An (s,S) inventory model with a space and service level constraint was used to determine optimal reorder points and quantities to minimise emergency refill costs. They aim to improve the model to include multiple objectives that management have to consider.

No studies were available that discussed implementation results, but simulated results promised a reduction in inventory cost. All methods for demand forecasting do however require good historic data, and accurate data capturing methods are required when implementing inventory management systems.

### 2.5 Scheduling

The GPP (2010) defines the scope of practice for pharmacists, pharmacist’s assistants and pharmacy students. Duties range from the dispensing process discussed in section 2.3.1 to inventory management. Some duties may only be performed by students and assistants under the direct supervision of a pharmacist. The allocation of tasks have been discussed previously in section 2.3.1. The
scheduling of personnel throughout the day to ensure all tasks are performed is discussed here.

Regarding scheduling, the GPP (2010) only states that if an institutional pharmacy is not open 24 hours a day, a designated pharmacist must be available to call-out to supply pharmaceutical services in an emergency. These call-outs are however inconvenient for employees and require overtime payment. The minimisation thereof can be considered as an objective in inventory management problems (as discussed in section 2.4) by optimal distribution of stock in other storage areas.

Scheduling problems rather address issues such as meeting customer and organisational demands while considering flexible work time conditions, part-time work, balancing undesirable shifts and other employee preferences. (Ernst et al., 2004)

Scheduling problems were classified by De Causmaecker et al. (2004) to be focused on either permanence, fluctuation, mobility or project centred planning. Hospitals, with the exception of extreme emergency situations, is permanence centred. This means that a minimum employee coverage is required at all times, and employees thus work in shifts with cyclical schedules. Institutional pharmacies, even though their employees are not on site after hours but they are still scheduled for call-outs, thus conform to this requirement for permanence centred planning.

The other dimension of these problems is qualifications, as only certain employees can perform certain tasks as per legal requirements (discussed above).

When modelling this as an optimisation problem, the objective is to minimise labour cost. The constraints include minimum coverage and balancing undesirable shifts. The problem can also be multi-objective, where employee “happiness” is considered as an additional objective in terms of satisfying work time condition preferences.

In the literature, nurse-scheduling problems have been the most prevalent (see Bester et al. (2007) for various examples). Here an extra constraint exists regarding the distribution of nurses in various departments and wards, which is not applicable to the pharmacy.

Pharmacy specific applications of optimisation models and simulation have been limited in literature. Butt and Acar (2013) used simple linear programming problems to minimise the required number of technicians over a cycle and maximise pharmacists’ and technicians’ preferences for certain shifts. These models were then included in a simulation model to find optimal personnel schedules. Hong et al. (2012) also used discrete event simulation to determine optimal schedules at an outpatient pharmacy. Both studies showed that, by optimising scheduling, an increasing demand could be met without additional employees.

Mahaney et al. (2008) discuss the trend towards non-traditional work schedules
at institutional pharmacies. They argue that changing demographics, specifically the increase of working mothers who wish to work only part-time and increased employee focus on work-life balance, are encouraging employers to be more flexible in personnel scheduling.

They cite cases of medical centres in the United States where employees at the same pharmacy have varying shift lengths and intervals. Some technicians only work when demand is high for two hour shifts at a time, other pharmacists work long shifts for three weeks and then take a week off every month, and one night-shift employee works intermittently for a week at a time. Some hospitals even allow pharmacists to perform certain tasks at home, such as prescription evaluations. The hospitals in question did not use software or mathematical models to determine schedules, but rather a highly customisable employee self-scheduling system overseen by the pharmacy manager. In these cases, staff turnover is significantly less than in traditionally scheduled hospital environments.

The literature thus shows that depending on one’s objectives, e.g. minimising labour cost, maximising employee preferences or minimising employee turnover, various models and schedules can be used for optimisation.

2.6 Human Resource Development, Management and Well-being

The available literature on how human resource education, development and skill sets affect pharmacy performance tend to focuses on specific demographics and countries. Drawing conclusions from these studies that are globally relevant is thus difficult.

For example, there is a gap in the literature to address how varying educational levels correlate to pharmacist performance, especially in the South African context. Schafheutle et al. (2011) address this briefly in their literature study on pharmacist performance, but only address assessments made of “overseas” trained pharmacists (in this case outside of the USA and Canada). The conclusions they make regarding the relative under-performance of foreign employees can thus not be extrapolated to a South African context.

The South African Pharmacy Council does however set strict regulations regarding the accreditation of institutions that offer the Bachelor of Pharmacy degree as well as the professional registration requirements for pharmacists and pharmacist’ assistants. The GPP (2010) also states that “continuing professional development is a professional obligation” for pharmacy employees. The pharmacy manager is also required to set appraisal and development objectives, and employees must be involved by performing self-appraisal. The way in which managers lead and appraise though, is not regulated. There are no additional requirements for
assessing the socio-technical variables that may affect pharmacy employees.

Other studies in literature do however have a more universal application. Job dissatisfaction has been cited by pharmacists as increasing the risk of error rates (Szeinbach et al., 2007; Bond et al., 2002). There is however a gap in the literature that provides empirical evidence on how mental health and other socio-technical factors affect the performance of pharmacy employees, or how such an intervention can be approached (James et al., 2009; Szeinbach et al., 2007).

In a study of pharmacist (n=26) job satisfaction in relation to management perceptions, Ferguson et al. (2011) found that most interviewees were dissatisfied with pharmacy management. Reasons cited included the perception that management was disinterested and that there was a lack of recognition and support. The authors suggest that this may lead to increased employee turnover.

One field of study that examines job satisfaction and performance is the theory of self-determination. The theory states that the three great psychological needs are autonomy, relatedness and competence. (Deci and Ryan, 2000)

Autonomy in this case is defined as people’s perception that they have choices. It is argued that autonomy has great adaptive advantages, in that the more autonomous an individual is within a situation, the more she specifies, processes and hierarchises the response to the situation. A lack of autonomy thus means responses will be less regulated, the full capacity to respond less explored, and the generated solution less adaptive. (Deci and Ryan, 2000)

Autonomy has become synonymous with self-motivation. It has been shown that acting on one’s own volition and with a sense of choice, i.e. with intrinsic motivation, results in more willing and engaged participation in tasks than any external “carrot and stick” motivations, even if the tasks are considered menial and boring. In an organisational sense, this means that the more employees internalise (i.e. agree with and see the validity of) organisational rules, processes and procedures, the more these external factors become internal convictions, which result in employees acting more pro-actively and with volition. This is because satisfaction is gained from learning and performing the task itself. It is no longer an obligation executed apathetically for some external reward. (Trépanier et al., 2012; Deci and Ryan, 2000)

Relatedness is concerned with feeling valued and appreciated, whilst competence refers to having the capability and required resources to accomplish tasks. (Meyer and Gagné, 2008)

Leadership styles can either support or undermine the three basic psychological needs. One supportive style is that of transformational leadership, which is often seen in contrast to transactional leadership (also known as managerial leadership). This style is based on exchanges of reward and punishment and promotes monitoring and corrective behaviour. It is thus responsive in nature
to maintain organisational stability. In contrast, transformational leadership is pro-active, and helps with continuous organisational improvement. Whereas transactional leadership aims to motivate through control, transformational leadership addresses autonomous motivation. Where transactional leaders manage tasks, transformational leaders manage support. Where transactional leadership creates a fear of failure, transformational leadership encourages continuous learning and development. Employees’ basic needs for autonomy, relatedness and competence are thus expected to be met far better in transformational leadership environments. (Hetland et al., 2011)

Studies show that having the three needs met in an organisation results in better employee performance, commitment and job satisfaction. (Gagné et al., 2014; Kuvaas, 2009; Judge and Bono, 2001)

There is a gap in the literature though of where these types of interventions have been applied in the strictly regulated pharmacy environment.

2.7 Summary

The case studies discussed in this chapter have shown how factors from five categories potentially affect the pharmacy performance.

The physical environment was shown to affect pharmacy employees with regards to alignment of process and layout, ergonomic design of equipment and workspace and environmental factors such as lighting.

Regarding pharmacy processes, the importance of task allocation and sequencing was highlighted, as well as how responses to interruptions significantly affect pharmacy workers.

Inventory management was shown to have various conflicting stakeholder objectives, and various mathematical models exist to determine optimal formulary and inventory levels and reorder points.

The literature on scheduling of employees was discussed, and examples of models to minimise labour cost, balance undesirable shifts and maximise employee shift preferences were shown. Cases where the advantages of non-traditional self-scheduling have been highlighted were also discussed.

Lastly, studies on worker engagement and job satisfaction were discussed, and studies outside of the pharmacy realm that could possibly be considered for research were discussed.

In the next chapter, methods for measuring service performance for objective evaluation are discussed.
Chapter 3

Measuring Service Performance

3.1 Introduction

Effectiveness is the concern with doing the right thing, that which will create the most value for an organisation. Efficiency meanwhile is a measure of the loss or gain of a process (Chase et al., 2007), thus asking the question “are things done right?” — a relatively simple question in a production environment, but not as simple in the services industry. In his seminal paper, The Measurement of Productive Efficiency, Ferrell (1957) states why this measure is required:

“The problem of measuring the productive efficiency of an industry is important to both the economic theorist and the economic policy maker. If the theoretical arguments as to the relative efficiency of different economic systems are to be subjected to empirical testing, it is essential to be able to make some actual measurements of efficiency. Equally, if economic planning is to concern itself with particular industries, it is important to know how far a given industry can be expected to increase its output by simply increasing its efficiency, without absorbing further resources.”

These measurements require an understanding of the four components of efficiency, as defined by Shermann and Zhu (2006):

- Price efficiency refers to acquisition finances, like the cost of human capital and raw materials.
- Allocative efficiency looks at the optimal mix of alternative inputs, for example the cost of automation versus human resources.
Technical inefficiency addresses a lack of outputs with the given inputs, for example centralising and standardisation of services and processes.

Scale efficiency refers to optimal unit production volumes.

Productivity, the ratio of input to output, is often used interchangeably with efficiency, but refers in fact mostly to technical efficiency (Shermann and Zhu, 2006).

The problem with measuring the efficiency in a service organisation is that it is difficult to adequately develop standards for comparison of these elements of efficiency. In health care, the difficulty lies in the multitude of different outputs (Shermann and Zhu, 2006), the subjective dimensions of quality of service i.e. the evaluation of how well inputs are transformed into outputs) (Van Looy et al., 1998) and the difficulty in objectively weighing the importance of cost against the level of care (Cooper et al., 2004). Gupta and Boyd (2008) argue that institutional pharmacies especially provide such a wide range of services with varying degrees of intensity, not all of which are recorded (for example patient consultation), that collection of data for evaluation is difficult.

Thus, in health care, a service that is merely technically efficient or productive, but is not a quality service, fails in being a well delivered service. A measurement of overall performance is more valuable if it considers both efficiency and quality of service.

Several techniques have been developed to address the need for relative performance measures and optimal resource exploitation, as discussed in the sections below.

### 3.2 Efficiency Measuring Techniques

#### 3.2.1 Accounting Techniques

In environments where activities are repetitive and standardised, *standard cost systems* can be used to manage performance. To determine budgets or standard costs, the resource input price is predicted based on either historical cost or engineered standards. This cost is then compared to actual cost to determine the system performance and the effects of variation in volume and resource prices. Typically, this system is used in mass production industries. (Gowthorpe, 2005)

The problem with using this system in the services industry is threefold. Firstly, engineered standards are based on an in-depth understanding of a standardised process, using tools like time and motion studies to determine exact resource requirements (Shermann and Zhu, 2006). In a pharmacy, the possible prescription combinations, and thus motion combinations, are boundless, and such an estimation would be impracticable. Secondly, comparison to historical values may not be
insightful, as historical performance may be inefficient and of poor quality, leading to a false sense of good current performance (Shermann and Zhu, 2006). Lastly, prescribing a monetary value to quality of service is ethically murky in any health care environment.

Another technique is *ratio analysis*, where various output over input ratios are defined for comparison, for example housekeeping cost per bed-day or nurses per high-risk patients. The main concern with these ratios is that the extent of their interdependency is difficult to evaluate and the sheer number of possible ratios can result in contradictory conclusions (Shermann and Zhu, 2006) (Cooper et al., 2004). For example, a hospital may have high cost per patient, but also a high ratio of high to low resource intensive patients. The latter may result in the former, but may also not account for all of the higher patient cost. Also, ratios assume comparable units, implying constant returns to scale (Cooper et al., 2004). Comparison of ratios of different managerial units can thus lead to superficial corrective action. Ratios can however be used insightful in organisations with limited service types, singular inputs and measurable quality standards (Shermann and Zhu, 2006).

Accounting budgeting techniques include *zero base budgeting* and *programme budgeting*. These techniques use analysis of historic expense data and comparison with other organisational divisions or programmes to create new efficiency goals. This puts the onus on management to estimate and justify expenditures (compared to other better performing units) well in advance. The problem is that segregating units in an organisation and having them compete for financial resources may result in increased unit efficiency but not necessarily increased organisational efficiency. These techniques only work well where systems are independent, with mutually exclusive functions and clear goals. It also does not allow specifically for quality considerations. It has the advantage though of combating management complacency due to the excessive focus on addressing and justifying resource expenses. (Shermann and Zhu, 2006)

The use of *accounting performance measurement techniques* in health care have the advantage of focusing on minimising expenses and increasing returns on investments, which are especially required in private health care enterprises. As discussed before though, it is difficult to assign a monetary or weighted value to quality care. These techniques also focus on extreme analytical, and not overall system evaluation, thus not providing for the interdependent nature of successful health care services.

These accounting techniques traditionally used in assessing health care services, are thus insufficient in providing adequate insight into hospital performance.
3.2.2 Balanced Scorecard

One framework for performance measurement is the *balanced scorecard*, developed by Kaplan and Norton (1992) at Harvard Business School. The scorecard aims to enable managers to translate strategic objectives to a few linked measures from the customer, financial, internal business and innovation and learning perspectives. It does this by answering four questions:

1. “how do our customers see us?” (customer perspective),
2. “how do we regard our shareholders?” (financial perspective),
3. “what must we excel at?” (internal business perspective) and
4. “can we continue to improve and create value?” (innovation and learning perspective).

The balanced scorecard has received some critique, in that weighing and interdependency of different measures are subjective and that even though it shows performance measures, it does not do well in identifying resource inefficiencies. (Amado et al., 2012)

3.2.3 Frontier Efficiency Methods

Frontier analysis methods were developed to address the limitations of techniques like ratio analysis in addressing multi-input and output processes. In these techniques, a maximum production output possibility frontier of different input combinations is estimated empirically. The relative performance of a unit is determined based on its distance from this frontier. (Cooper et al., 2004)

Parametric (or econometric) frontier analysis requires functional form and distribution assumptions, as with Stochastic Frontier Analysis (SFA). The advantage of this approach is that it distinguishes between random fluctuations in inputs and inefficiencies in production. These distributional assumptions can however lead to specification bias. Non-parametric analysis requires fewer assumptions, but is then considered more sensitive to measurement errors, for example DEA. (Krüger, 2012)

In a literature review of 317 published papers on frontier analysis techniques in health care service from the past 30 years, Hollingsworth (2008) found that over 80% used non-parametric analysis. The breakdown of methods used can be seen in Figure 3.1. The author supposes that the small use of parametric methods could be due to the complexity of its application. As discussed in section 3.2.1, the service industry can be complex, in that the transformation of input to outputs can happen in many ways. There is a high risk of misspecification of the functional form of transforming inputs to outputs when using parametric
methods, which would greatly affect efficiency estimates. Another reason cited by Cooper et al. (2004) is due to the inherent nature and types of problems that SFA, the most popular parametric approach, addresses. SFA performs an overall optimisation across various observations to find inefficiency estimates, whereas DEA performs separate optimisations for each unit, thereby providing insight into the inefficiencies and quality performance of each individual unit. SFA and Stochastic Frontier Regression tools are thus often used to understand entire populations, whereas DEA is concerned with individual observations, which may be of more use in health care institutions.

Citing various Monte Carlo experiments evaluating various parametric and non-parametric performance analysis methods, Krüger (2012) concluded that there is no clear superior method in terms of accuracy and consistency. SFA and DEA do however fare better compared to other techniques such as Free Disposal Hull and Corrected Ordinary Least Squares. SFA does outperform DEA in cases with larger sample sizes (50, 100 and 200) and large measurement errors, whilst DEA is slightly superior with small sample sizes and small measurement errors.

In the application in this thesis, data are accurately captured over a large period of time, decreasing the likelihood of measurement errors. For these reasons, only DEA is discussed onwards.
3.3 Data Envelopment Analysis

3.3.1 Background

Data Envelopment Analysis (DEA) is a technique used to measure the relative performance of a set of Decision Making Units (DMUs) (Cooper et al., 2004). The term DMU is a generic and flexible term used in DEA to describe units that perform the same function of converting inputs to outputs. DEA was first developed by Charnes et al. (1978) as a method to evaluate the efficiency of activities in not-for-profit programmes. The original concept was to identify the best performers in a set of DMUs and that they would form the efficiency frontier that would create a best-practice benchmark for comparison of non-frontier units. It has since developed in theoretical foundation and application in the services and manufacturing industries (Cook and Seiford, 2009).

The basic principles of DEA can be illustrated with a simplified pharmacy example. Table 3.1 shows five hypothetical pharmacies (A to E), with the number of pharmacists employed by these and the number of line item prescriptions filled per week.

<table>
<thead>
<tr>
<th>Pharmacy</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmacists</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Line items filled per week</td>
<td>1300</td>
<td>1660</td>
<td>1900</td>
<td>1500</td>
<td>1250</td>
</tr>
</tbody>
</table>

Table 3.1: Simplified pharmacy DEA example: Data.

If we assume that an increase in the number of pharmacists results in an equivalent increase in the number of prescriptions filled, we can plot this data in relation to a linear efficiency frontier, as per Figure 3.2.

In this example, DMU C is 100% efficient relative to the other DMUs in the set. The efficiency frontier on which DMU C lies shows the best efficiency performance ratio, and is said to “envelop” the data.

The DMUs not on the efficiency frontier are relatively inefficient. The greater the distance to the frontier, the less efficient the DMU. So, in this example, DMU B is more efficient than DMU D.

For these inefficient DMUs to become more efficient, they either have to:

- reduce the number of pharmacists (i.e. the input) used to fill prescriptions while keeping the line items filled per week constant,
- increase the number of line items filled per pharmacists (i.e. the output) using the same number of pharmacists, or
- reduce both the number of pharmacists and increase the number of line items filled.
If a DMU implements such performance interventions successfully, it is possible that the efficiency frontier will move, and that DMU C loses its efficient classification. So, for example, if DMU E increase the number of line items filled per week to 1 400 with the same number of pharmacists, the efficiency frontier would change as per Figure 3.3.

3.3.2 DEA Models and Definitions

As discussed previously, frontier methods like DEA were developed to address the deficiencies in using the index number approach for efficiency measurement,
i.e. the need to assign weights to input and output measures prior to evaluation. This is achieved by evaluating relative, and not absolute, efficiency. Cooper et al. (2004) offer the following definition for relative efficiency when using DEA:

“A DMU is to be rated as fully (100%) efficient on the basis of available evidence if and only if the performance of other DMUs does not show that some of its inputs or outputs can be improved without worsening some of its other inputs or outputs.”

The Charnes et al. (1978) model measures the efficiency of a DMU as the maximum ratio of the multiple weighted outputs to multiple weighted inputs subject to the similar ratio of every DMU being less than or equal to one. This can be modelled as

\[
\begin{align*}
\text{maximize} & \quad h_0(u, v) = \frac{\sum_{r} u_r y_{ro}}{\sum_{i} v_i x_{io}} \\
\text{subject to} & \quad \frac{\sum_{r} u_r y_{rj}}{\sum_{i} v_i x_{ij}} \leq 1 \quad j = 1, ..., n \\
& \quad u_r, v_i \geq 0 \text{ for all } i \text{ and } r
\end{align*}
\]

where \(y_{ro}\) and \(x_{io}\) in the objective function (3.1) are the respective output and input values for the DMU_0 under evaluation. The intensity variables, or weights, \(u\) and \(v\) are determined during the DEA evaluation for each input \(i\) and output \(r\). The constraint in (3.2) shows that the ratio of output to input of each other DMU \(j\) in the set must be equal to or less than unity, as no DMU can be more than 100% efficient.

DEA also utilises the duality of linear programming, i.e. the ability to look at a problem from either a minimisation or maximisation point of view. So, if \(\theta^*\) is defined as the efficiency score of a DMU, and \(s_i^{-}\) and \(s_i^{+}\) as the slack variables, Cooper et al. (2004) defines the following:

“The performance of DMU_0 is fully (100%) efficient if and only if both (i) \(\theta^* = 1\) and (ii) all slacks \(s_i^{-} = s_i^{+} = 0\).”

In the same way:

“The performance of DMU_0 is weakly efficient if and only if both (i) \(\theta^* = 1\) and (ii) all slacks \(s_i^{-} \neq 0\) and/or \(s_i^{+} \neq 0\) for some \(i\) and \(r\) in some alternate optima.”
The Charnes et al. (1978) DEA models, also called the CCR models after the authors, are broadly classified into either input or output orientated models. If an organisation wishes to achieve the same output level with the more efficient use of fewer inputs, a minimisation input orientated model is used. In contrast, an output orientated model determines the potential maximum production (or output) given fixed inputs.

The CCR input-orientated envelopment model is defined by Cooper et al. (2004) as

\[
\text{minimise} \quad \theta - \epsilon \left( \sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+ \right)
\]
\[
\text{subject to} \quad \sum_{j=1}^{n} x_{ij} \lambda_j + s_i^- = \theta x_{io} \quad i = 1, ..., m
\]
\[
\sum_{j=1}^{n} y_{rj} \lambda_j - s_r^+ = y_{ro} \quad r = 1, ..., s
\]
\[
\lambda_j \geq 0 \quad j = 1, ..., n.
\]

(3.3)

The associated input-orientated multiplier model is defined as

\[
\text{maximise} \quad z = \sum_{r=1}^{s} u_r y_{ro}
\]
\[
\text{subject to} \quad \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0
\]
\[
\sum_{i=1}^{m} v_i x_{io} = 1
\]
\[
u_i, v_i \geq \epsilon > 0.
\]

(3.4)

The CCR output-orientated envelopment model is defined as

\[
\text{maximise} \quad \phi - \epsilon \left( \sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+ \right)
\]
\[
\text{subject to} \quad \sum_{j=1}^{n} x_{ij} \lambda_j + s_i^- = x_{io} \quad i = 1, ..., m
\]
\[
\sum_{j=1}^{n} y_{rj} \lambda_j - s_r^+ = \phi y_{ro} \quad r = 1, ..., s
\]
\[
\lambda_j \geq 0 \quad j = 1, ..., n.
\]
The associated output-orientated multiplier model is defined as

\[
\begin{align*}
\text{minimise} \quad q &= \sum_{i=1}^{m} v_i x_{io} \\
\text{subject to} \quad \sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} &\geq 0 \\
\sum_{r=1}^{s} u_r y_{ro} &= 1 \\
u_r, v_i &\geq \epsilon > 0.
\end{align*}
\]

In the envelopment models above, \( \lambda_j \) is the intensity vector that gives an indication of how far \( DMU_j \) is from the efficiency frontier. The variable \( \epsilon \) is a non-Archimedean element (defined as infinitesimal or smaller than any real number). The weight variables, \( u_r \) and \( v_i \), show how much that specific output or input contributed to the efficiency score of the DMU (an efficient input or output will contribute more to the efficiency score). Also, \( n \) is the total number of DMUs, \( s \) the number of outputs and \( m \) the number of inputs per DMU.

These models are known as Constant Returns to Scale (CRS) models (Cooper et al., 2004). It works under the assumption that if the inputs are increased or decreased by a factor, the outputs will reflect the same proportional change — as was discussed in the example in section 3.3.1. Various Variable Returns to Scale (VRS) models have been developed to evaluate economies and diseconomies of scale, most notably from Banker et al. (1984), Thrall and Banker (1992) and Färe and Grosskopf (1994).

Figure 3.4 illustrates the difference between CRS and VRS models.

Ray \( OBC \) shows an efficiency frontier. If a CRS model was used, DMUs on ray \( AB \) and \( CD \) would be considered inefficient and those on \( BC \) as 100% efficient. However, in an VRS model, ray \( AB \) would show increasing returns to scale, and \( CD \) decreasing returns to scale.

Banker et al. (1984) built on this with the BCC model, which assumes VRS, as
Figure 3.4: Returns to scale frontiers.

\[
\begin{align*}
\text{minimise} & \quad \sum_{j=1}^{n} \theta_j + \epsilon \left( \sum_{i=1}^{m} \hat{s}_i^- + \sum_{r=1}^{s} \hat{s}_r^+ \right) \\
\text{subject to} & \quad \sum_{j=1}^{n} x_{ij} \lambda_j + s_i^- = \theta_i x_{io} \quad i = 1, \ldots, m \\
& \quad \sum_{j=1}^{n} y_{rj} \lambda_j - s_r^+ = y_{ro} \quad r = 1, \ldots, s \\
& \quad \sum_{j=1}^{n} \lambda_j = 1 \quad j = 1, \ldots, n \quad (3.5) \\
\text{with} & \quad 0 \leq \hat{\theta}_j, \hat{s}_i^-, \hat{s}_r^+ \forall i, j, r.
\end{align*}
\]

The added constraint (3.5) is a convexity constraint to allow for VRS. Banker et al. (1984) note that if the same method is employed with the output-orientated CCR model (3.4), the returns to scale may change. They suggest evaluating the results of both models before augmenting any inputs or outputs.

The multiplier of the BCC model is

31
maximise \[ z = \sum_{r=1}^{s} u_r y_{ro} - u_0 \] (3.6)

subject to \[ \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} - u_0 \leq 0 \]
\[ \sum_{i=1}^{m} v_i x_{io} = 1 \]
\[ u_r, v_i \geq \epsilon, u_0 \text{ free in sign.} \]

Thrall and Banker (1992) proposed that once the (3.6) model has been solved to obtain \( u_o \), the following can be said:

• if \( u_o < 0 \) at \((\hat{x}_o, \hat{y}_o)\) for all alternate optima then increasing returns to scale prevail, so an increase in input values will result in a proportionally larger increase in output values,

• if \( u_o > 0 \) at \((\hat{x}_o, \hat{y}_o)\) for all alternate optima then decreasing returns to scale prevail, so an increase in input values will result in proportionally smaller increased output values,

• if \( u_o = 0 \) at \((\hat{x}_o, \hat{y}_o)\) for at least one optimal solution then constant returns to scale prevail, so an increase in inputs will result in an equivalent increase in outputs.

One DEA model that does not distinguish between input and output orientation is the additive model (Cooper et al., 2004). This is because the objective function maximises outputs and minimises inputs simultaneously, as can be seen in (3.7).

maximise \[ \sum_{r=1}^{s} s_r^+ + \sum_{i=1}^{m} s_i^- \] (3.7)

subject to \[ \sum_{j=1}^{n} y_{rj} \lambda_j - s_r^+ = y_{ro} \quad r = 1, ..., s \]
\[ \sum_{j=1}^{n} x_{ij} \lambda_j + s_i^- = x_{io} \quad i = 1, ..., m \]
\[ \sum_{j=1}^{n} \lambda_j = 1 \quad j = 1, ..., n \]

with \[ 0 \leq \lambda_j, s_i^-, s_r^+ \quad \forall \ i, j, r. \]

This model however provides no measure for efficiency and summing slacks cannot be done where inputs and outputs have non-commensurate units. One
solution is proposed by Tone (2001) to create a slacks-based measure model, by replacing the objective function in (3.7) with

$$\text{maximise } \rho = \frac{1 - (1/m) \sum_{i=1}^{m} s_i^- / x_{io}}{1 - (1/s) \sum_{r=1}^{s} s_r^+ / y_{ro}}.$$ 

This substitution however changes the objective function to be non-linear. Through transformation of (3.7), Tone (2001) define the equivalent linear problem to be

$$\text{maximise } \tau = t - \frac{1}{m} \sum_{i=1}^{m} S_i^-$$
subject to
$$t + \frac{1}{s} \sum_{r=1}^{s} y_{rj} S_r^+ = 1$$
$$X \Lambda + S^- = tx_o$$
$$Y \Lambda + S^+ = ty_o$$
with
$$0 \leq \Lambda_j, S^-, S^+; 0 < t.$$

The optimal solution of this Linear Programme (LP) for a DMU ($\tau^*, t^*, \Lambda^*, S_i^-^*, S_r^+^*$ ) of the original additive DEA model is defined by

$$\rho^* = \tau^*, \lambda^* = \Lambda^*/t^*, s_{i}^-^* = S_{i}^-^*/t^*, s_{r}^+^* = S_{r}^+^*/t^*.$$ 

The associated multiplier model is

$$\text{maximise } u y_o - v x_o$$
subject to
$$u Y - v X \leq 0$$
$$\frac{1}{m} [1/x_o] \leq v$$
$$\frac{1 - v x_o + u y_o}{m} [1/y_o] \leq u.$$ 

Other notable DEA models as discussed by Cook and Seiford (2009) include:

- multiple level network models that evaluate activities in a process, for example where decisions in one activity affect the next or where multiple
machines are available for production,

- multiplier restriction level models where bounds are placed on the inputs and outputs and
- the Malmquist index that evaluates the change in productivity over time.

### 3.3.3 Sensitivity Analysis of DEA Inputs and Results

Criticisms have been raised on the use of DEA because it does not allow for random fluctuations in input or output data, for example when an epidemic occurs. This may result in deviations from the efficiency frontier being interpreted as inefficiencies or the frontier moving direction during analysis. (Newhouse, 1994)

One consideration is that the DEA methodology is sensitive to degrees of freedom (df). The number of inputs \(m\) and outputs \(s\) decreases df and the number of DMUs \(n\) increases df. To choose \(n\), Cooper et al. (2004) suggest the rule of thumb

\[ n \geq \max\{m \times s, 3(m + s)\}. \]

Sensitivity analysis methodologies discussed further here assume that this condition has been met.

Charnes et al. (1992) introduced a metric approach for DEA sensitivity analysis that evaluates the stability of a DMU, i.e., how far it is from changing status from “inefficient” to “efficient”. In Figure 3.5, this stability region for node F is shown. F is considered stable as it has a large radius of stability.

![Figure 3.5: DEA radius of stability. (Charnes et al., 1992)](https://scholar.sun.ac.za)

This distance can be modelled by modifying the additive model to reflect an “improvement” in both inputs and outputs, through
maximise \( \delta \)

subject to \[
\sum_{j=1}^{n} y_{rj} \lambda_j - s_r^+ - \delta d_r^+ = y_{ro} \quad r = 1, \ldots, s
\]

\[
\sum_{j=1}^{n} x_{ij} \lambda_j + s_i^- + \delta d_i^- = x_{io} \quad i = 1, \ldots, m
\]

\[
\sum_{j=1}^{n} \lambda_j = 1 \quad j = 1, \ldots, n
\]

with \( 0 \leq \lambda_j, \delta, s_i^-, s_r^+ \forall i, j, r. \) (3.8)

If all inputs and outputs are considered equally important, \( d_r \) and \( d_i \) is set equal to one. In this case, from (3.8), a DMU will remain inefficient for output data ranging from \( y_{ro} \) to \( y_{ro} + \delta^* \) and input data ranging from \( x_{io} \) to \( x_{io} - \delta^* \) where \( \delta^* \) is the maximum radius of stability for the DMU.

Evaluating when an efficient DMU will become inefficient is done in a similar way, only \( \delta \) is minimised and the inputs and outputs of the DMU are “worsened” by (Cooper et al., 2004)

minimise \( \delta \)

subject to \[
\sum_{j=1, j \neq 0}^{n} y_{rj} \lambda_j - s_r^+ - \delta d_r^+ = y_{ro} \quad r = 1, \ldots, s \quad (3.9)
\]

\[
\sum_{j=1, j \neq 0}^{n} x_{ij} \lambda_j + s_i^- - \delta d_i^- = x_{io} \quad i = 1, \ldots, m \quad (3.10)
\]

\[
\sum_{j=1, j \neq 0}^{n} \lambda_j = 1 \quad j = 1, \ldots, n \quad (3.11)
\]

with \( 0 \leq \lambda_j, \delta, s_i^-, s_r^+ \forall i, j, r. \)

The exclusion of the efficient DMU under consideration in (3.9), (3.10) and (3.11) is required, otherwise the solution will be equal to \( \delta = 0 \) and \( \lambda_o = 1 \). This would indicate that the efficient DMU has a zero radius of stability.

3.3.4 Considerations for DEA in Healthcare

Throughout the literature on DEA in health care applications, various considerations are noted for analysts.

Cooper et al. (2004) note that with respect to DEA application in health care, one should consider the severity of medical attention required as a quality measure.
In a pharmacy for example, some prescriptions, like “stat” medication and intensive care unit (ICU) patient prescriptions require quicker action than normal ward prescriptions. Case mix and complexity is thus an important consideration. Also, to evaluate the performance of a unit the outcome of services provided must be measured, both in terms of error rate and in response to severity. A unit should only be considered fully efficient if it produces constant quality outcomes.

Zere et al. (2001) notes that the health status of a patient is difficult to measure, and is also not a reliable outcome of an efficient and successful health care system. The successful and quality completion of a process or service in health care is thus the preferred measurement. (Zere et al., 2001)

In a critique on applying DEA in health care, Newhouse (1994) raised the concern that DEA assumes no random error or fluctuation in data, critical inputs such as human resources are often omitted in studies and case mix measures are not considered in studies. The first concern can be addressed using the sensitivity analysis techniques described in section 3.3.3. To address the second concern, guidance can be taken from Cooper et al. (2004) who note that health care incurs high labour cost, and that it is sensible to distinguish between employees of different pay grades as input when performing DEA. The third concern can be addressed in the pharmacy environment by looking at prescription line items, and not prescriptions as a whole, as output of the DEA study.

### 3.3.5 DEA Health Care Case Studies

Sexton et al. (1989) used a Charnes, Cooper and Rhodes (CCR) DEA model to evaluate the inefficiencies of Veteran Administration Medical Centers (VMCAs) in the United States. They argue that since medical treatment to veterans is free of charge, the total benefit received (or aggregated output) cannot be valued and so traditional accounting measures cannot be used. As inputs to the model, the evaluators used Full-Time Employee Equivalent (FTEE) units, which is equal to 2087 labour hours per year. The FTEEs of nurses, physicians, part-time physicians, residents and health technicians were evaluated, as well as drug and supplies cost and equipment cost. As output, the evaluators used Workload Weighted Units (WWUs), a unit assigned to each diagnosis-related group according to its labour requirements.

The WWUs of the surgical, medical, psychiatric, nursing-home, intermediary care and outpatients groups were evaluated. It was argued that these outputs are dependent on demand and equal across all DMUs, and sources of inefficiencies lie in reducing inputs, not increasing outputs. Of the 157 VAMCs evaluated, 107 were found to be on the efficiency frontier. The evaluators grouped the DMUs on the efficiency frontier to set input targets for similar DMUs not on the frontier. They found that VMCAs with university affiliations were more likely to be inefficient,
but concluded that the labour required for training and research had unquantifiable intrinsic value and the affiliations are not undesirable.

DEA has also been used to determine the CRS and VRS efficiencies of South African public hospitals in Gauteng (Kibambe and Koch, 2007) and the Western, Eastern and Northern Cape (Zere et al., 2001). In both studies the analysts noted a lack of data as their primary concern in using the method, and that care should be taken in interpretation of results in these studies. For example, Kibambe and Koch (2007) states that participation in their study was voluntary, and that only those hospitals that were responsive and thorough in data request were analysed. This could lead to biased results, as efficiently run hospitals will tend to be better able and willing to access their own data. Zere et al. (2001) found that in some hospitals the beds were underutilised (thus decreasing returns of scale) even though in the South African public sector the population to hospital bed ratio is relatively high. The results of both these studies thus did not offer conclusive results for immediate implementation, but can aid in decision making on the reallocation of resources on a national level and on the national health care strategy.

In the pharmaceutical services industry, studies have been performed mainly in Sweden, where the retail pharmaceutical services have been performed by one public monopoly since 1971. The act that governs this corporation stated that they should “provide an adequate supply of drugs while maintaining drug cost at the lowest possible level” and to maintain a “good standard of service” (Löthgren and Tambour, 1999).

Färe et al. (1995) used a Malmquist index to evaluate productivity and quality changes of these pharmacies over one year by measuring three inputs, seven outputs and three attributes. The inputs were the total pharmacists and technical staff service hours as well as an expense index that included cleaning cost, depreciation of equipment, energy and material cost. The outputs included the number of outpatient prescription packages, dose package deliveries, drug deliveries to health care units, special articles for the handicapped, special food for the handicapped, over the counter goods and information on drugs. The service level attributes to evaluate quality included queuing time, trading hours and timeous delivery of prescriptions.

Even though these attributes require more resources while maintaining the same outputs, a better “standard of service” is achieved which is in line with the corporation’s mandate. The inclusion also allows for fair evaluation of various types of pharmacies. The authors found that including the quality attributes significantly affected the measured change in overall performance over the year.

Färe and Grosskopf (1996) also developed a network DEA model that was modified by Löthgren and Tambour (1999) to evaluate customer satisfaction in 23 of the Swedish pharmacies. They argue that since pharmacists’ time is spent
both in the traditional production activities of filling a prescription and with customer orientated activities like consultation, the effect of this time spent should be evaluated.

The network DEA model allows for production and consumption activities to be presented as two nodes in a network and determines optimal attribute levels for both. Inputs considered were cost factors and labour hours, while evaluated outputs included the number of outpatient prescriptions, over-the-counter prescriptions and other expeditions.

They also developed the method to include quality attributes, including trading hours and number of prescriptions filled per day. The customer satisfactions quality factors and impact were measured in a customer satisfaction barometer questionnaire and included availability of the service, the pharmacy premises, service on prescription drugs, service on prescription free drugs and queue service. Results showed that inclusion of customer satisfaction in evaluations lowered the efficiency rankings of the DMUs.

### 3.4 Summary

In this chapter the traditional accounting efficiency measurement techniques were discussed, and their shortcomings when measuring health service performance were highlighted. Frontier efficiency methods were found to be better suited to the complex health care services industry, with the non-parametric DEA method being the most predominant technique seen in health care literature.

The case studies discussed illustrate how DEA has been used to measure pharmacy performance, whilst considering both efficiency and quality attributes. The South African case studies do however show that where there is a lack of data, the results from DEA studies have to be interpreted with caution. In these cases one must be mindful of the context within which the DEA is performed when considering implications of the results, such as the educational context of the VAMC studies. The Swedish case studies gave insight into the typical input and output factors used in pharmacy applications of DEA, especially when also considering quality of service, and not efficiency alone. The reasoning behind choosing input orientated models is also noted.

In the next chapter, the application of DEA in the institutional pharmacies of a South African private hospital group is detailed. The considerations from the literature regarding chosen input in health care and sensitivity analysis as well as the discussed case studies are included.
Chapter 4

Applying Data Envelopment Analysis to Institutional Pharmacies

4.1 Introduction

Cooper et al. (2004) define an eight step procedure to apply DEA in a healthcare setting. These steps are:

1. Identify the managerial or research question.
2. Develop a conceptual model of the health care process under consideration.
3. Identify factors influencing the health care map.
4. Select factors to include in a study based on available data.
5. Analyse input data using statistical methods.
6. Run various DEA models.
7. Perform sensitivity analysis on DEA results.
8. Get input from health care providers on results for further validation.

These steps will form the sections of this chapter where the DEA case study is discussed.
4.2 DEA Research Question

This thesis research aim is to use DEA in South African private health care institutional pharmacies to identify technical inefficiencies and quality concerns, to enable operational managers to specify and budget for improvement initiatives.

DEA is used to find initially under-performing pharmacies, but also continuously (perhaps semi-annually) to monitor improvements and intervention efficacy.

To do this, it would improve the feasibility of using DEA as evaluation tool if it could be performed as a desktop study. If a pharmacy manager needed to collect data in a consistent manner of all the pharmacies in his region, it could take months, if not years, to gather significant data through observation. It is thus more feasible to use data captured electronically with the pharmacy management system only.

The other factor that would increase the feasibility of DEA adoption is software choice. Software with high licensing fees would be impractical if open-source software could perform the DEA.

4.3 Institutional Pharmacy Process Model

Institutional pharmacies provide dispensing services to the hospital wards, patients who have been discharged and require TTO medication, and walk-in members of the public (i.e. the retail section). The pharmacies are also responsible for storing and inventory control of medication throughout the hospital. Select fast-moving inventory is stored in the various wards and also the night cupboard — a strict access-controlled storage facility that hospital staff uses outside of pharmacy operating hours.

Data on the responsibilities of institutional pharmacies regarding these functions were garnered by observation and interviews during site visits. Process models were drawn from observation at one pharmacy, and then verified and altered based on observations at other pharmacies to ensure globally relevant process maps. With the exception of tasks with strict legal requirements, not all tasks are performed by the same roles in all hospitals. For example, the GPP (2010) states that a pharmacist’s assistant may not handle schedule 6 medication unless under direct supervision of a pharmacist. Medication billing may however be performed by a pharmacist, pharmacist’s assistant or an administrative assistant, depending on the hospital size and resource availability. The process maps thus do not show specific process enablers, but rather the possible process role players.

An example of one of these process maps can be seen in Figure 4.1. This is the process by which a pharmacist replenishes the stock of the night cupboard that was dispensed outside of the pharmacy operating hours. This process led to a better
understanding of the stock control duties of pharmacists, origins of incident reports (when a stock take reveals that some medication is unaccounted for) and patient billing. Through the mapping of all the relevant pharmacy processes a complete understanding was gained of the factors influencing pharmaceutical services. The balance of the process maps can be seen in Appendix A.

### 4.4 DEA Inputs, Outputs and Attributes

From the defined pharmacy goal, literature guidelines and the process maps, a list of the ideal input, output and attribute factors was developed for inclusion in the DEA evaluation.
DEA is performed with the assumption that there is a correlation between process inputs and outputs. It is however possible also to include quality attributes into the evaluation, as discussed by Löthgren and Tambour (1999) and Färe et al. (1995). DEA also tries to show DMUs in the “best light”, as it will assign proportionally higher weights to the best inputs and outputs of a DMU, resulting in the best possible efficiency score (Cooper et al., 2004). Factors contributing to inefficiencies are thus given a lower weight during DEA.

In the first iteration of this DEA study a quality attribute considering the number of incidents in the pharmacy was included. An incident is where a medication dispensing error was reported, such as the incorrect dosage or even the incorrect medication being dispensed. It was found however that the inclusion of this as an input (i.e. something that has to be minimised even though not directly correlated to output quantities) resulted in a disproportionally high weighting being assigned to this input where a DMU had zero incidents. This resulted in the other factors, those measuring productivity, being assigned weights approaching zero. It was thus difficult to gain insight into the DMU efficiency overall while this quality factor was included.

In Chapter 1 it was discussed that the pharmacy should have a prioritised list of goals. The highest priority is to adhere to measurable legal specifications. Next, pharmacies must prioritise improving adherence to more subjective regulations, such as the efficacy of error prevention measures. Finally, cost-effectiveness of pharmacy operations must be considered.

If a pharmacy is experiencing incidents, a clear priority must be to evaluate error prevention measures before evaluating efficiency. However, if a pharmacy has no incidents, a decision maker requires insight into what factors are affecting efficiency. For this reason, the number of incident reports was excluded from the DEA, but is still included in the overall DMU evaluation process, as will be discussed in section 4.9.

The following data outputs were chosen for inclusion in the DEA evaluation:

- The number of in-hospital prescriptions (including that of the ICU) filled per month.
- The number of TTO prescriptions filled per month.
- The number of retail prescriptions filled per month.

The following data inputs were chosen for inclusion in the DEA evaluation:

- The number of pharmacist hours per month — this is the number of full time pharmacists in service times 177 work hours per month.
• The number of pharmacist’s assistant hours per month — this is the number of full time pharmacist’s assistants in service times 177 work hours per month.

• The number of intern hours per month — these are pharmacology students who can perform limited (but necessary) tasks in the pharmacy as part of in-hospital training.

• The percentage aged stock — this is the percentage of stock that has been on the pharmacy shelf for longer than three months, and is put up for sale at a reduced price to other pharmacies in the area to see if it can be used before it expires.

• The number of call-outs per month — these events occur outside of pharmacy operating hours when a patient requires medication not available from the ward or night cupboard stock. In these cases a pharmacist has to go into the hospital to fill the prescription from the pharmacy stock.

The output factors are non-discretionary variables, i.e. they cannot be varied at the discretion of a manager, whereas the input factors are discretionary in this case study.

At first it was thought that it would be more ideal to include as input the time spent to fill a prescription type. This however proved infeasible for two reasons.

Firstly, the patient file does not have a time stamp for when it is received by the pharmacy or when it leaves. A study using that as input data would require significant field data collection, whereas the chosen input, namely employee hours per month, is readily available and more feasible for a desktop study.

Secondly, the pharmacies use various prescription filling processes, making process times non-commensurate. For example, in some pharmacies, a pharmacist takes a prescription and performs all the dispensing activities sequentially. In other pharmacies, a “production line” type process is followed. Here a pharmacist takes all the incoming prescriptions and performs the evaluation step only for all of them while an assistant picks the medication. The pharmacist then performs the final inspection, labels the medication and readies the prescriptions for delivery.

The chosen process inputs and outputs that are used for the DEA are thus all measurable and readily available from the hospital management system. The pharmacy processes, which are the result of integrating the factors described in Chapter 2, are modelled as per Figure 4.2 for the purposes of the DEA.

4.5 Input Data Analysis

Data from 46 pharmacies in one private hospital group was used in the DEA. These pharmacies are located throughout South Africa. The pharmacies vary in size,
Pharmacy processes
(Integration of physical environment, process design, inventory management, scheduling, and human resource management and well-being)

Pharmacist’s assistant hours
(input variable\textsubscript{2})

Pharmacist hours
(input variable\textsubscript{1})

Intern hours
(input variable\textsubscript{3})

Aged stock percentage
(input variable\textsubscript{4})

Number of call-outs
(input variable\textsubscript{5})

Number of in-hospital prescriptions
(output variable\textsubscript{1})

Number of TTO prescriptions
(output variable\textsubscript{2})

Number of retail prescriptions
(output variable\textsubscript{3})

Figure 4.2: Measurable inputs and outputs of pharmacy processes for DEA.

with the smallest employing only one pharmacist and the largest 7 pharmacists and 6 pharmacist’s assistants.

Input data analysis was limited to a visual inspection of the data, where the completeness and basic sensibility of the data were evaluated. The average of four months of data was taken to evaluate the DEA evaluation methodology. The effect of possible variation was evaluated during the sensitivity analysis, discussed in section 4.7.

The effect of zeroes in data had an effect on one of the DEA models. How this was accommodated is discussed in section 4.6.

4.6 Data Envelopment Analysis Models

4.6.1 DEA Model Development

To ensure robustness of results, three different models were selected to perform the DEA, namely a CCR Input Orientated model, a BBC Input Oriented model, and an Additive Slacks Based Measure model. The envelopment and multiplier of each model were programmed. The flow diagram depicting the programming logic as applied in each model can be seen in Figure 4.3.

The models were programmed using the CVXOPT library (Andersen et al., 2014), a free software convex optimisation package for use in Python interpreters. LPs are special cases of the general convex optimisation problem (for more see Boyd and Vandenberghe (2004)), and is solved in this case using an interior points method (for more see Andersen et al. (2012)).
Define global variables

Read data from data source

Convert data to matrix form for optimisation

All DMUs evaluated?

Define optimisation variables

Define optimisation constraints

Define objective function

Perform optimisation

Prepare results for data output file

Write results to data file

stop

Figure 4.3: DEA model source code flow diagram.
The CVXOPT library in Python was selected due to its ease of use over some popular alternatives, namely Microsoft® Excel® (with Solver Add-in) and MathWorks® MatLAB®. Each DEA model requires solving a LP for each DMU with a unique set of constraints — in this case 6 DEA models for each of the 46 DMUs resulting in 276 unique LPs. Automation of creating these, using Python, is simpler than implementation would have been using Microsoft® Excel® with Solver Add-in. Python was used instead of MathWorks® MatLAB because it is open source, and health care institutions wishing to use DEA analysis based on the example in this thesis would not be required to pay licensing fees to duplicate the models. Python 2.7 was used as the CVXOPT library is only supported in this version.

Data was read from Comma-Separated Values (CSV) files and using a decimal dot and comma separated format. CSV was used because of its ease of use when working with large data sets. Most pharmacy management systems also have the option to export data in CSV format.

The DEA models were verified using the following tests:

- Results from the envelopment and multiplier forms of each model were compared to ensure that they were equal.
- Hypothetical data illustrating extreme conditions were used to ensure that models behaved as expected, for example DMUs with high inputs and low outputs resulted in low efficiency scores, and vice versa.
- DMUs that were rated as 100% efficient in the CCR models were compared to the BBC model results, where they should also be rated 100% efficient.
- Data from models in literature were used as input for comparison to ensure correct model results.

The source code for the three different models, with both the envelopment and multiplier model for each, can be seen in Appendix B.

The additive slacks based measure models are sensitive to zero data points, as division by zero would then occur. For example, as discussed in Chapter 3, the objective function of the envelopment model is

\[
\text{maximise } \tau = t - \frac{1}{m} \sum_{i=1}^{m} \frac{S_i^-}{x_{30}}
\]

and the first constraint is
\[ t + \frac{1}{s} \sum_{r=1}^{s} y_{rj} \frac{S^+}{y_{ro}} = 1. \]

It can be seen that if \( y_{ro} \) or \( x_{io} \) are equal to zero, division by zero would occur. Tone (2001) proves that in these cases, the inverse term can be deleted for consideration in the constraint or objective function by setting it equal to zero. To model this, the inverse for each input and output data point was first defined before applying DEA. For points equal to zero, the inverse was defined as equal to zero. This can be seen in the source code of both the envelopment and multiplier slacks based measure additive models in Appendix B.5 and B.6.

4.6.2 DEA Overall Results

Figures 4.4, 4.5 and 4.6 show the overall efficiency scores of the DMUs for the three different models.

![Efficiency Score Chart](chart.png)

**Figure 4.4: CCR model study results.**

The majority of the DMUs have an efficiency score of one in all the cases. Cooper et al. (2004) refers to DEA evaluations with such results as having evaluated a complex production technology. This means that there are many slightly different ways of providing the service. No one input and output mix is thus optimal for all DMUs.

In the case of the institutional pharmacy, this holds true. As discussed previously, the dispensing process can be performed in many ways. Individuals can take responsibility for a prescription and perform all dispensing processes for it, or a production line dispensing process can be followed. Also, the tasks that can be as-
signed to the various employee types, namely pharmacists, pharmacist’s assistants and interns, are not mutually exclusive. Although only a pharmacist can evaluate and perform final prescription inspections and handle schedule 6 medication, any employee can pick other medications, man the ad hoc retail window, perform and receive general orders and perform stock takes. Various employee type mixes and task allocation practices can thus be equally efficient.

The CCR model assumes CRS, as opposed to the Banker, Charnes and Cooper (BCC) model that assumes VRS. Therefore the efficiency ratings of CCR DMUs (in Figure 4.4) are slightly different from those evaluated with the BCC model.
(in Figure 4.5). This is because in a pharmacy VRS prevails, as the number of prescriptions filled will not necessarily increase proportionally if there are more employees.

This was confirmed by the DEA when the condition for constant returns to scale (as defined in Chapter 3) was not met. The conditions for strict decreasing or strict increasing returns to scale were also not met.

The input orientated CCR and BCC models consider how much input can be reduced while fixing outputs. In contrast to this, the slacks based measure models are neither input or output orientated, and are “less forgiving” of output shortfall. So, as defined in Chapter 2, a DMU can be rated 100% efficient, even though the slacks are not equal to zero. Such units are weakly efficient. The slacks based measure model does not allow a 100% efficiency score for a DMU if there is any shortfall of inputs or outputs (i.e. the slacks are not zero). This is why the results of the slacks based measure DEA model has lower efficiency scores.

There was however consistency in the results of the various models. Several DMUs were rated fully efficient as per the definition in Chapter 3, namely that the efficiency rating is 100% and all slacks were equal to zero in all models.

Various insights can be garnered from the various models. In the next section, the results of three DMUs are discussed in detail as examples of how these insights can be obtained.

4.7 Sensitivity Analysis of Results

4.7.1 Modelling Sensitivity Analysis

The source code for the models for sensitivity analysis of efficient and inefficient DMUs are included in Appendix B.7 and B.8.

The sensitivity analysis technique where a radius of stability is found for each DMU was used. As discussed in section 3.3.3 the number of inputs \((m)\) and outputs \((s)\) decreases df and the number of DMUs \((n)\) increases df.

The df condition is that

\[ n \geq \max\{m \ast s, 3(m + s)\}. \]

In this analysis, this is true as

\[ 46 \geq \max\{5 \ast 3, 3(5 + 3)\}. \]

4.7.2 Examples of DMU Results Interpretation

To illustrate how the data from the DEA can be interpreted, results from three DMUs from the analysis are discussed in this section.
All the models showed that DMU 27 was one of the most efficient in the set. It achieved a 100% efficiency rating and slack values of zero for all inputs and outputs with all three envelopment models (see Table 4.1). As per the definition of DEA efficiency provided in Chapter 3, DMU 27 is thus fully efficient.

<table>
<thead>
<tr>
<th>Theta</th>
<th>BCC Model</th>
<th>CCR Model</th>
<th>Additive Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_{i1}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i2}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i3}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i4}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i5}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i6}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i7}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i8}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i9}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i10}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i11}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i12}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i13}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i14}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i15}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i16}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i17}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i18}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i19}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i20}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i21}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i22}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i23}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i24}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i25}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i26}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i27}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i28}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i29}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i30}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i31}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i32}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i33}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i34}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i35}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i36}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i37}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i38}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i39}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i40}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i41}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i42}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i43}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i44}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i45}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i46}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i47}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i48}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i49}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$s_{i50}$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 4.1: DMU j=27 results.

The sensitivity analysis did show that DMU 27 has the largest radius of stability for the efficient DMUs. This indicates that the inputs can increase, or outputs decrease, significantly without a reclassification from efficient to inefficient. Because this radius has a large finite value (in this case 3 840) the DMU is classified as stable. This means that even if the input data were to vary in future evaluations, the DMU would still be classified as efficient.

Continuous improvement methodologies, as discussed in the next chapter, can thus be deployed to ensure that DMU 27 remains competitively adaptive.

The result of the second DMU for detailed discussion (number 19) is shown in Table 4.2. Even though the DMU was not rated 100% efficient, it achieved high efficiency scores from both the input oriented models. As discussed, the slacks based additive model is neither input or output oriented, and is “less forgiving” of shortcomings in output data, and the efficiency score is thus lower.

The fourth input variable, namely aged stock percentage, contributed most to the high efficiency rating of the DMU in all three models. This can be seen from the large weighting value assigned to it ($v_4$) and the zero slack values ($s_{i4}$). This
indicates that compared to the others, this pharmacy has a good and balanced inventory managements policy (as discussed in Chapter 2).

The employee hours however contributed to decreases in the efficiency rating. This is evident from the zero weighting assigned to the input variables \(v_1\), \(v_2\) and \(v_3\) and the presence of slack values \(s_{i1}\), \(s_{i2}\) and \(s_{i3}\). It is also evident due to the difference in efficiency rating between the input orientated models and the additive slacks based measure model. The additive model is penalising the DMU since others produce significantly more output with the same input.

This indicates that this pharmacy, in relation to others, requires more employee hours to fill prescriptions. Evaluation of the pharmacy will need to determine if this is due to poor scheduling (i.e. the employees are scheduled to be on duty but do not have enough work to do) or due to inefficiencies. If the former is not observed, benchmarking with other pharmacies on process design and the physical pharmacy environment (as discussed in Chapter 2) is required.

The sensitivity analysis shows that the radius of stability is high (740), meaning that the DMU is considered stable in its inefficient classification.

The results of the third DMU under consideration (number 43) can be seen in Table 4.3.

The DMU scored very low efficiency ratings from all models. As with DMU 19, the employee hours have virtually zero assigned weightings \(v_1\), \(v_2\) and \(v_3\). The slacks however show that the pharmacist’s assistant and intern hours have excess

<table>
<thead>
<tr>
<th></th>
<th>BCC Model</th>
<th>CCR Model</th>
<th>Additive Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theta</td>
<td>0.978</td>
<td>0.97</td>
<td>0.639</td>
</tr>
<tr>
<td>(s_{i1})</td>
<td>18.3</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>(s_{i2})</td>
<td>338</td>
<td>336</td>
<td>346</td>
</tr>
<tr>
<td>(s_{i3})</td>
<td>75.2</td>
<td>70.9</td>
<td>73.1</td>
</tr>
<tr>
<td>(s_{i4})</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>(s_{i5})</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>(s_{o1})</td>
<td>1930</td>
<td>1960</td>
<td>2410</td>
</tr>
<tr>
<td>(s_{o2})</td>
<td>519</td>
<td>532</td>
<td>578</td>
</tr>
<tr>
<td>(s_{o3})</td>
<td>0.000</td>
<td>0.000</td>
<td>220</td>
</tr>
<tr>
<td>(v_1)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>(v_2)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>(v_3)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>(v_4)</td>
<td>36.4</td>
<td>43.2</td>
<td>29.3</td>
</tr>
<tr>
<td>(v_5)</td>
<td>0.339</td>
<td>0.176</td>
<td>0.083</td>
</tr>
<tr>
<td>(u_1)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>(u_2)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>(u_3)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 4.2: DMU j=19 results.
<table>
<thead>
<tr>
<th></th>
<th>BCC Model</th>
<th>CCR Model</th>
<th>Additive Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Theta )</td>
<td>0.642</td>
<td>0.613</td>
<td>0.298</td>
</tr>
<tr>
<td>( s_{i1} )</td>
<td>0.000</td>
<td>0.000</td>
<td>111</td>
</tr>
<tr>
<td>( s_{i2} )</td>
<td>138</td>
<td>125</td>
<td>487</td>
</tr>
<tr>
<td>( s_{i3} )</td>
<td>219</td>
<td>163</td>
<td>354</td>
</tr>
<tr>
<td>( s_{i4} )</td>
<td>2.55</td>
<td>1.12</td>
<td>8</td>
</tr>
<tr>
<td>( s_{i5} )</td>
<td>0.000</td>
<td>0.000</td>
<td>1.01</td>
</tr>
<tr>
<td>( s_{o1} )</td>
<td>1880</td>
<td>0.000</td>
<td>4440</td>
</tr>
<tr>
<td>( s_{o2} )</td>
<td>12.5</td>
<td>0.000</td>
<td>11.4</td>
</tr>
<tr>
<td>( s_{o3} )</td>
<td>0.000</td>
<td>0.000</td>
<td>1980</td>
</tr>
<tr>
<td>( Z )</td>
<td>0.642</td>
<td>0.613</td>
<td>0.298</td>
</tr>
<tr>
<td>( v_{1} )</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( v_{2} )</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( v_{3} )</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( v_{4} )</td>
<td>0.000</td>
<td>0.000</td>
<td>0.025</td>
</tr>
<tr>
<td>( v_{5} )</td>
<td>0.151</td>
<td>0.151</td>
<td>0.048</td>
</tr>
<tr>
<td>( u_{1} )</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( u_{2} )</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( u_{3} )</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 4.3: DMU \( j=43 \) results.

slack values \( s_{i2} \) and \( s_{i3} \) while the pharmacist hours have zero slack values \( s_{i1} \) for the input orientated models. This could indicate that there is an imbalance in the assignment of tasks. It would be advisable to evaluate the process design of the system first to identify inefficiencies.

The aged stock percentage was also assigned almost zero weightings \( v_{4} \) and has relatively high slack values \( s_{i4} \). Evaluation of inventory management policies is thus required.

The pharmacy does however have very few call-outs as can be seen from the high weighting assigned to \( v_{5} \). This could mean that a wide variety of stock is distributed amongst the wards resulting in fewer call-outs, even though this means too much excess stock is kept on hand which then ages. A more balanced inventory policy could increase efficiency.

This DMU has the largest radius of stability (a value of 1 500) of the inefficient DMUs. It is thus unlikely to be reclassified unless significant efficiency improvement initiatives are implemented.

4.8 Validation of Analysis

As discussed in Chapter 3, the current method for measuring performance, namely ratio analysis, is insufficient for gaining applicable insight into pharmacy inefficiencies and quality concerns. This is because it assumes that constant returns of
scales can only be used with comparable units and sheer number of possible ratios can lead to contradicting data analysis. The operational pharmacy manager for the case study was specifically concerned that the current ratio analysis makes no distinction between hospitals of various sizes.

A method for overcoming this shortcoming in current performance measurement is needed. DEA is a better tool for analysis than ratio analysis, as it can compare DMUs of various sizes, uses readily available data and provides clear and analytically sound direction into inefficiencies that must be addressed.

An interview was also conducted with the programme coordinator of the Pharmaceutical Sciences School of the University of KwaZulu Natal on the undergraduate pharmacy degree content. Pharmacy and pharmacology students are prepared at undergraduate level for a myriad of different workplaces, including various pharmacy types, manufacturing, sales, policy making, research and development.

The strict regulations surrounding course accreditation focus mainly on the medicinal aspect of pharmacology. One course is presented on pharmacy management, with various pharmacy settings (private, public, retail, mobile and wholesale) covered. Practical experience of the institutional pharmacy as work environment is acquired through internships for some students.

The onus of gaining knowledge and insight into efficiently working in and managing institutional pharmacies specifically thus lies with the employer, and not through undergraduate course work. Also, as per the Occupational Health and Safety Act (1993), the onus lies with the employer to make employees aware of specific workplace hazards, as discussed in Chapter 2. The employer must provide training and focus on providing insight to help employees to improve efficiency and overcome workplace ergonomic hazards.

It is thus not realistic to expect employees to have the intrinsic knowledge to improve pharmacy operations. The pharmacy has many parallel processes, and bottlenecks are difficult to identify without the proper analytical know-how and with limited experience regarding management as a science. Knowledge of DEA and the associated scientific management literature can thus aid the employees (who eventually get promoted to management) to become more aware of possible pharmacy operational improvements.

### 4.9 Summary

The overall approach applied for identifying technical inefficiencies and quality concerns in South African private hospital pharmacies using DEA for improvement initiatives can be viewed in Figure 4.7.

First, relevant data of the DMU set is collected. The input data includes phar-
macist, pharmacist’s assistant and intern hours per month, aged stock percentage and the number of call-outs per month. The output data includes the number of prescriptions filled for in-hospital, TTO and retail use. The quality attribute data includes the number of reported incidents per month.

Second, analysis of the quality attribute data is performed. Investigation is initially required at DMUs experiencing significant incidents, because efficient error-prevention procedures are prioritised over efficiency. Considerations for such investigations can include, as discussed in Chapter 2:

- Redesigning the prescription evaluation and final inspection procedures.
- Evaluation of possible distractions in the pharmacy due to the pharmacy layout.
- Standardising procedures for reacting to interruptions (such as telephone calls and requests for “stat” medication).
- Evaluation of environmental factors, such as lighting.

Improvement initiatives can then be targeted towards the cost-effectiveness of pharmacy services. For this, the third step is to perform DEA using the various models (including sensitivity analysis).

When interpreting DEA results, the efficiency score is the first consideration. An efficiency score of 100% for all models and zero slacks indicates that the DMU is fully efficient. This can be confirmed by the radius of stability determined through sensitivity analysis that shows if it is stable in its efficiency classification. Continuous improvement thinking can be introduced in these cases to enable the group as a whole to remain competitive (as is discussed in the next chapter).

For weakly efficient and inefficient DMUs, high slack values and low weighting variables give an indication of pharmacy process inputs that can be improved. The performance inefficiencies and improvement initiatives discussed in Chapter 2 are related to these factors.

Inefficient use of employee hours (i.e. the first three input variables) can have three causes. Firstly, the employees may be scheduled inefficiently (as discussed in section 2.5). Observation can show if employees are not busy for the full working day, and interventions to improve scheduling can solve this. Secondly, employees may be working inefficiently due to poor facility design (see section 2.2.1) and poor process design (see section 2.3). Thirdly, employees may not be motivated, as discussed in section 2.6.

Inefficiencies caused by the fourth input variable, namely the percentage of aged stock, can be addressed by better inventory policies and management, as discussed in section 2.4.
Collect DMU data

Incidents ≥ 1? Yes → Evaluate error prevention interventions from literature (Chapter 2)

No → Perform DEA (Chapter 4)

Rated fully DEA efficient? No → Evaluate efficiency improvement interventions from literature (Chapter 2)

Yes → Evaluate intervention possibilities from CI literature (Chapter 5)

Figure 4.7: DMU data analysis flow diagram.
Inefficiencies caused by the fifth variable, namely the number of call-outs, can be addressed either by more efficient scheduling (if there is consistency in the time and frequency of call-outs) or better distributed stock management.

The DMU data analysis process can be followed iteratively to evaluate the efficacy of improvement initiatives.

In the next chapter, the continuous improvement methodologies that can be used to improve the DMUs proactively are discussed.
Chapter 5

Continuous Improvement

The DEA results showed the relative efficiency of a group of pharmacies. Those that were rated inefficient can benefit from the case study considerations discussed in Chapter 2.

The ultimate purpose of improvement is however not to have a group of pharmacies all rated 100% relatively efficient in relation to each other. The private health care industry remains competitive, and overall continuous improvement is required to advance and maintain a dominant market position.

For this reason, proactive performance improvement is discussed in this chapter. The two main focus areas are that of continuous process and quality improvement methodologies and advanced pharmacy technologies.

5.1 Continuous Improvement Methodologies

Quality management has evolved from being reactive and inspection-orientated to an ideology of Total Quality Management (TQM), defined as “managing the entire organisation so that it excels on all dimensions of products and services that are important to the customer”. This concept of managing quality mostly evolved from the work of Edward Deming, Philip Crosby, Kaoru Ishikawa and Joseph Juran, who defined quality as “fitness for use” or conformance to requirements. The goal of TQM is the careful design of the service or product in question, and gearing the entire organisation to consistently produce this design. (Chase et al., 2007)

The core principles of TQM include a focus on customers, process, empowerment and involvement of employees, education, commitment by top management, continuous improvement, synergy between customers and suppliers and a systematic, scientific approach. (Dahlgaard-Park, 2011)
The term TQM however has been less prevalent in literature in the last decade as Continuous Improvement (CI) concepts (or kaizen) such as six sigma and lean processes have become more prominent. In a literature study of TQM, Dahlgaard-Park (2011) concluded that this is due to TQM not having unique precepts. CI techniques share many of the values of TQM, but with more definite focus areas. The techniques have however also been criticised for not adequately addressing the behavioural dimension of their implementation. As CI is often a deliberate change in behaviour and culture that evolves over time, it cannot be expected to effect significant change in the short term or in a once-off implementation (Bessant et al., 2001).

Adoption of these techniques in an organisation can produce exceptional results in the long run (as will be seen in the next sections). Continuous incremental improvements can be desirable where an organisation perceives no glaring quality impediments in its products or services. In the following sections these techniques and case studies are discussed.

5.1.1 Plan Do Study Act with Root Cause Analysis

The Plan Do Study Act (PDSA) cycle (also known as the Shewart Cycle or Deming Wheel/Cycle) is one of the fundamental continuous improvement tools. This iterative cycle starts with researching and choosing a desired future state and the actions to achieve it (Plan), then implementation (Do), followed by reflection and evaluation of results (Study) and finally adjusting in response to the results to achieve standardisation or the desired effect (Act). (Orsini, 2013)

The focus of PDSA lies in the planning phase, where root cause and risk analysis of the system in question is performed. For this, one of the seven quality tools of Ishikawa (1976) can be used, namely:

- *Cause-and-effect (or fishbone) diagrams* are used to determine and group common causal factors of a quality characteristic.

- *Check sheets* are used to obtain data in such a way as to allow for automated analysis.

- *Histograms* are used in data analysis to observe the frequency of data in a defined interval.

- *Control charts* are used to observe the impact of a changing factor in a process over time.

- *Pareto charts* are used to determine the most prevalent common causal factor of a quality characteristic.
• Scatter diagrams are a visual representation of the relationship between two kinds of data.

• Sampling is the selection of a subset of a population to gain knowledge of the entire population.

Other Root Cause Analysis (RCA) tools include:

• Gap analysis is used to study the difference in performance and requirements for system goals (Gitlow et al., 2005).

• Matrix diagrams are used to organise data of two factors into subsets for comparison (Gitlow et al., 2005).

• Why-why diagrams relentlessly ask “why?” to a problem and the subsequent answers until a root cause is identified (Nelsen, 2003).

• Failure Modes, Effects and Criticality Analysis (FMECA) is a prospective tool used to evaluate potential failure modes and their criticality in a system to determine the likelihood, frequency and detection possibility of a failure (Nguyen et al., 2013).

• A Program Decision Process Chart is a prospective tool used to identify possible problems arising from a plan to be executed and to develop contingency plans (Gitlow et al., 2005).

These tools are used to eliminate inefficiencies in many industry sectors, from manufacturing to services. A kaizen initiative using root cause analysis tools was successfully implemented at the Mitchell’s Plain Community Health Centre in the Western Cape to address the overcrowded triage area (Isaacs and Hellenberg, 2009). The implementation team used Pareto analysis to group patients entering the triage area and found that those requiring reissuing of prescriptions contributed the most to overcrowding. A fishbone diagram was then drawn up to identify reasons for this issue, and surveys were then completed to quantify the contributing factors. Finally, the Why-Why method was used to determine countermeasures for the identified factors. The initiative resulted in a decrease from 46 to 22 reissues per day within three months after implementation.

5.1.2 Lean Thinking

Lean Thinking is a management philosophy intent on achieving a streamlined production line by exposing bottlenecks and eliminating waste. The term “lean” was first coined in a Massachusetts Institute of Technology study of Japanese manufacturing plants (especially that of Toyota) (Sobek and Jimmerson, 2003).
Due to their intertwining principles and philosophies regarding waste, Lean Manufacturing and the Toyota Production System (TPS) have become synonymous. (Chase et al., 2007)

Fujio Cho, the honorary chairman of the Toyota Motoring Company, defines waste as “anything other than the minimum amount of equipment, materials, parts and workers (working time) which are absolutely essential to production”. Types of waste thus include overproduction, waiting time, transportation, inventory, processing, motion and product defects. Elements to address such waste include implementing Just in Time (JIT) production systems, uniform plant loading, clustering and *kanban* “pull” systems (Chase et al., 2007).

Duplication of Toyota's success with this production system was however rare in other companies in early literature. In their review of the TPS, Spear and Bowen (1999) argue that this was due to a paradox where the Toyota manufacturing plants seem excessively planned and rigid, but are in fact flexible and continuously improved. They identified four underlying rules to the philosophies that would make it easier to implement elsewhere. Firstly, the content, sequence, timing and deliverable of all work is detailed to reduce variation and deviation from specifications. Secondly, every worker connection is standardised and correct. Workers have supplier-customer relationships with each other, with clear expectations of the form and quantity of goods or services to be exchanged. Thirdly, every service or product travels along a well specified path, unless there is deliberate redesign of a process. Finally, continuous improvement is enabled through the scientific method. The hypothesis logic is presented and tested before implementation. Any redesign is done at the lowest possible level of the organisation under the guidance of a teacher.

Based on the TPS and other Lean Thinking initiatives, Womack and Jones (2003) developed five principles of Lean Production. These include defining value from the customers' perspective, mapping the value stream from problem solving through information management to transformation tasks, ensuring continuous flow of value-adding services or goods, using pull systems to inform actions and pursuing continuous improvement.

TPS and Lean Thinking have been implemented increasingly outside of the manufacturing industry. In Chapter 2 the case of Sobek and Jimmerson (2003) was discussed where the standardising of reactions to interruptions in pharmacies significantly improved lead times and reduce telephone calls. In a literature review of 30 topical articles published in peer reviewed journals, Poksinska (2010) showed that Lean was implemented in various health care case studies. Characteristics of Lean for these included (in order of number of occurrences) process improvement, continuous flow, value mapping, waste elimination, team work, lead time reduction and *kaizen*. The seven types of waste have also been translated to the healthcare
environment to illustrate Lean implementation relevance and value (Machado and Leitner, 2010) (Radnor et al., 2012). Examples include the staff movement to obtain patient information files (motion waste), waiting for treatment, diagnosis and discharge (waiting time waste), ageing stock and supplies (inventory waste) and duplication of information (processing waste).

Lean Thinking has also been used in facility planning of the emergency department of a 561-bed university hospital in Illinois in the United States. The designers found that as with clustering of products requiring similar processes in manufacturing improve efficiency, the same could be achieved with patient processes. The emergency department was sectioned in three parts, namely Emergency Severity Index (ESI) 1-2, ESI 3-5 and trauma. ESI 3-5 accounts for 78% of the patient volume. By optimising the layout (like proximity to the nurses’ station and supply closet) for this section, total nurse walking time (and thus motion waste) was reduced by 71% compared to the previous layout. (Nicholas, 2012)

There are however challenges in implementing Lean Thinking in healthcare. Acceptance by healthcare practitioners is not always forthcoming, as the history of Lean in automotive manufacturing can seem unrelated to the personal and life threatening nature of healthcare. Educators require non-manufacturing examples and case studies to convince health care workers of its applicability in their industry.

Another challenge is the definition of customer value in health care. A customer can be either the patient, a family member, a referring health care specialist, a medical aid or the government, all with varying perspectives of what “value” is in health services. Another challenge is the strict hierarchical nature of healthcare, as the decision makers (physicians) on top are usually not the ones experiencing inefficiencies. Lastly, health care consists of many interdependent units. There is thus a risk that positive changes in one unit could negatively effect another. Overall system efficiencies thus need to be taken into account. (Poksinska, 2010)

5.1.3 Six Sigma

The Six Sigma Black Belt Handbook (McCarty et al., 2005) describes Six Sigma as a metric, methodology and overall management system. As a metric, Six Sigma is used to measure quality on a scale that shows Defects per Million Opportunities (DPMO), with the target of Six Sigma being 3.4 DPMO. In this way, various processes can be compared to each other. As a problem solving and continuous improvement methodology, Six Sigma uses the metric and root cause analysis tools like Define, Measure, Analyse, Improve, Control (DMAIC) to reduce process variation and errors. As a management system (inclusive of the metric and the methodology), Six Sigma focuses on understanding customer requirements and orientating processes to meet them, reducing process variation through rigorous data analysis and encouraging quick and sustainable continuous improvement.
Initially introduced by Motorola, who won the 1988 Malcolm Baldrige National Quality Award in the United States after developing and implementing it, there have been many successful Six Sigma case studies in various sectors (Drohomeretski et al., 2014). For example, the United States National Association for Healthcare Quality (Carboneau et al., 2010) used the Six Sigma DMAIC approach to increase hand hygiene compliance at a public hospital in New Mexico. It was identified as a problem that infection caused more deaths than AIDS, breast cancer and car accidents combined in the United States.

The implementation team used DMAIC to first determine (D) behavioural and physical barriers to hand hygiene, and then to measure (M) the current level of adherence to the relevant policies. In the analysis (A) phase, the team identified education, a culture of non-compliance and the availability of working hands-free sinks as the root causes of the problem. By addressing these root causes in the improve (I) phase and putting in an acceptable threshold monitoring plan in the control (C) phase, the hospital managed to reduce hospital acquired infections by 51% in 12 months. The net savings of the initiative amounted to approximately US$276,000 in 2012 due to a decrease in preventable, prolonged, paid-for hospital stays.

In a survey on Six Sigma in the services industry globally, Chakraborty and Chuan (2013) found that “top management commitment and involvement” is a main critical success factor for implementation. As with Lean Thinking, education and training are vital to success. Other factors include overcoming resistance to change and increasing customer focus.

### 5.1.4 Lean Six Sigma

Six Sigma initiatives focus heavily on quality, often neglecting the need for efficiency. Lean processes however focus on efficiency, the elimination of waste and process simplification, often neglecting rigorous data analysis and effectiveness (Antony, 2011). The two philosophies have proven to be complementary, forming a Lean Six Sigma business strategy that has been accepted in many sectors (Drohomeretski et al., 2014; Atmaca and Girenes, 2011; Salah et al., 2010).

Lean Six Sigma also uses the Six Sigma DMAIC tool, but during the Analysis phase lean tools like kanban, push/pull and JIT are used to eliminate waste as well as improve quality. (Snee, 2010)

### 5.1.5 Theory of Constraints

Theory of Constraints (TOC) is a managerial philosophy aimed at effecting continuous improvement in an organisation by focusing on the constraints affecting the system. Eliyah Goldratt, the pioneer of TOC, defined a constraint as “anything
that limits a system from achieving higher performance versus its goal” (Chase et al., 2007). The focus of TOC is to systematically identify, exploit and elevate the constraints in a system to achieve higher system performance through continuous improvement. Rahman (1998) cites two underlying principles of TOC from Goldratt’s theory and business novels: firstly, that every system must have at least one constraint and secondly that the existence of constraints indicates that system intervention for improvement is possible.

Goldratt (1990) argued that a system is built around a purpose, and that before attempting an improvement intervention on the system, every system element must be evaluated in terms of its impact on this purpose. He defined five steps to improve a system:

1. Identify the constraints and rank their effect on the system purpose.
2. Exploit the limiting constraints to maximise throughput.
3. Subordinate the non-constraints to the limiting constraints.
4. Elevate the limiting constraint by adding capacity to it.
5. Reiterate the steps to achieve continuous improvement.

Goldratt developed the “thinking process” for TOC implementation teams, described simply as identifying what to change, what to change it to and how to cause the change. He encouraged the use of cause-effect diagrams and current reality trees when identifying constraints to ensure that the underlying problem of a system, and not symptoms of the problem, were addressed. This notion is reminiscent of the Pareto Principle, which in the operations management field is the assumption that a few root causes result in a significant number of perceived problems (Polito et al., 2006).

A notable case study by Womack and Flowers (1999) on the application of TOC in the health care environment is that of the 366th Medical Group, a United States Air Force unit medical facility in Idaho. The facility was struggling to meet the target waiting time of seven days for patients wishing to make routine appointments, even though the target waiting time for acute appointments (24 hours) and health maintenance exams (four weeks) were consistently met. The TOC steps were followed, and the implementation team identified the root cause of the constraint as being a rigid scheduling system. By modifying the schedule templates to be dynamic in response to variation in appointment type demand, the team exploited the constraint and subordinated other appointment types to the limiting type.

The new improvement iteration identified the limiting constraint as the medical technicians. This seemed counter-intuitive, as the health care providers were the
most expensive human resource providers. The providers however reported that much of their time was consumed in tracking down technicians to perform various tasks. By hiring more technicians to match the number of practitioners (and thus elevating the constraint) the facility consistently met its target waiting times.

The TOC initiative was expanded to other facilities, and the United States Department of Defence generated $1.6 million additional revenue at a cost of less than $200 000.

Other successful implementations of TOC and the thinking process span across various industries. Mabin et al. (2001) used the thinking process in a bank merger situation in New Zealand to facilitate change management. The thinking process tools were also used by Polito et al. (2006) to reduce expendable and non-expendable in-flight inventory for Best Airlines in the United States. An implementation of the five TOC steps in the Hitachi Tool Engineering manufacturing plant in Japan led to such great improvements in lead times, quality and inventory management, that the TOC principles have been adopted in all sections of the organisation (Umble et al., 2006). There are many more examples — as Kim et al. (2008) note, 57 papers on the successful application of TOC and the thinking process were published in reviewed journals and conference proceedings between 1999 and 2006.

Even with the many instances of successful practical implementations of TOC and the related thinking process in literature, it is still regarded with scepticism in academic spheres. In their review of the evolution of TOC, Watson et al. (2007) noted that even with the increased practical implementation and acceptance as a viable management philosophy, TOC has not yet been accepted as a scientific theory with sound theoretical foundations.

Validation of the technique is complicated owing to there being no method to show how well the TOC method was implemented in a case study. Subsequently, it is hard to distinguish between cases where TOC would never have been a successful methodology for improvement, and cases where it was merely poorly applied and yielded no success.

This is illustrated by the case study published by Lubitsh et al. (2005) where TOC was implemented in three National Health Service (NHS) departments in the United Kingdom. The expectation was that TOC initiatives in the Eyes, Neurosurgery and Ear, Nose and Throat (ENT) departments would result in increases efficiency and output. At the end of the two-year study, the ENT and Eyes departments showed improvements in several success factors, such as number of outpatients seen within 30 minutes, number of appointment cancellations and finished consulting episodes. The same improvements were however not achieved in the neurosurgery department. The authors discusses this in the paper, noting that it could be because the department is larger and involves more complex processes and that the implementation was less structured. It could thus be that
the department would never achieve success with TOC, or that it was just not
implemented well.

In a study of TOC and thinking process peer reviewed literature, Kim et al.
(2008) revealed four research areas that are not widely addressed. Firstly, the
literature does not address the successes and failures of TOC in comparison to
other business improvement processes such as Lean and Six Sigma. Secondly, TOC
has only been used to amend perceived existing problems, and not to problems in
a positive sense, such as stretched goals and changes in strategic thinking. Thirdly,
no literature documents the empirical nature of TOC. Finally, it is questioned in
the literature if combinatorial business improvement methodologies cannot show
better results than blindly following only the TOC steps. The final two conclusions
were also made by Gupta and Boyd (2008) in their study of TOC as a unifying
theory for operations management.

Pirasteh and Farah (2006) wrote of a global electronics contract manufacturer
with customers in the medical, aerospace, defence, telecommunication and storage
computing industries. To meet growing demand while reducing cost and increasing
quality and value, the organisation sought a process improvement methodology to
implement in their 21 plants. It was suggested by consultants that they implement
a combinatorial solution, including the best principles of Theory of Constraints,
Lean and Six Sigma (TLS) in one initiative. To do a comparative study, trial were
run in all the plants. Six sigma was implemented in 11 plants, the lean process
at four plants and the new TLS method at six plants. The effectiveness of the
methodologies were researched in a double-blind format to reduce potential biases,
so all plant personnel and methodology trainers were unaware of the comparative
nature of the study. In the TLS plants, the improvement initiative started with
identifying the high level organisational problems using the five TOC steps.

Next, the first four steps of the Lean process as well as the 5S method was
followed to create waste-less value stream. A pull system was then implemented,
and finally a Six Sigma process was followed to eliminate process deviations. After
a two year study, it was found that the combinatorial TLS method accounted for
89% of the resulting cost saving from the entire improvement initiative.

The literature cites other non-comparative study examples where TOC in
combination with other improvement technologies yielded successful results (Ehie
and Sheu, 2005; Jin et al., 2009; Creasy and Ramey, 2013). TOC has also been
combined with linear programming and heuristics in product-mix problems to
achieve better results (Fredendall and Lea, 1997; Mabin and Gibson, 1998; Mabin
et al., 2001; Wang et al., 2009).

It can be seen that even though there are many TOC success stories, without a
sound theoretical foundation, TOC cannot guarantee significant improvement for a
system. The combination of TOC with other process improvement methodologies
to better guarantee success is however gaining more interest in literature and academic spheres. The Goldratt Institute even presents workshops on integrating TOC with Lean and Six Sigma (AGI-Goldratt Institute, 2013).

5.2 Pharmacy Technology Advancements

Technologies to enable the automation of dispensing activities have advanced significantly in recent years, and have been implemented in various pharmacy settings. Introduction of these various technologies in a hospital group can improve performance, as is discussed below.

5.2.1 Electronic Prescription Systems

From observation of institutional pharmacies, one of the most prominent issues for pharmacies is the availability of the patient file. Physicians make annotations on it and use it for prescribing items, nurses use it for annotations and to follow care and medication instruction and pharmacists require it to fulfil prescriptions.

Runners collect the files intermittently during the day, usually based on typical physician ward visits schedules. They return the file with the prescription as groups of ward prescriptions are filled. The problem is that medication is often required outside of these pre-defined ward visits, and cause many interruptions in the pharmacy for “stat” prescriptions.

A physician writes out prescriptions on the file that pharmacists then transcribe on the pharmacy system for labelling and billing purposes. This transcription process has been shown to cause preventable errors, mostly due to poor handwriting. (Peterson et al., 1999; James et al., 2009)

An electronic prescription system can avoid these errors. It enables authorised users (i.e. physicians) to type in the prescription, which is then immediately available to the institutional pharmacy. These systems have been shown also to include errors where physicians press the wrong key, but these errors are outnumbered by previously experienced errors and often prevented when pharmacists review the prescription. (Thomsen, 2004)

5.2.2 Automatic Counting Systems

Typically pharmacies use a pill tray and spatula to count pills manually. These trays allow the pharmacist to tip out a number of pills from a large container, count out the required number for dispensing, and funnel these into repackaging materials.

In their study on causes of dispensing errors in the United States, James et al. (2009) found that incorrect dispensing quantity was one of the five most
prevalent unpreventable errors. It was also the most prevented error, meaning that the error was corrected during final inspection before release. The United Kingdom experienced similar quantity related error rates.

Automated counting systems have been shown to reduce these types of errors and increase productivity in hospital pharmacies. (James et al., 2009; Thomsen, 2004)

There is however a concern in the industry that pill residue in these systems may cause harmful cross-contamination. Pill residue on trays is visible, and cleaning them is simple. Cleaning best-practices and technologies utilising vacuums could address this issue for automatic systems. (Sipkoff, 2006a,b)

5.2.3 Robotics

Robotics are increasingly being used to count, bottle and label fast-moving items in various pharmacies, including institutional pharmacies. The purpose of these is to increase pharmacy productivity, whilst freeing the pharmacists from menial labour and allowing them to focus on prescription evaluation and patient counselling. Technologies mostly utilise a robotic arm that can select appropriate packaging (including vials for liquid medication) and perform bar code scanning to select and label medication. (Coblio, 2011; Lin et al., 2007)

There is a concern that automation systems, if implemented without proper consideration for pharmacy needs, may not be cost-effective or decrease error rates. Technological intervention can still not prevent interruptions and divided attention or relieve mental workload of pharmacists. Systems also have limitations on the number of medications they can dispense, and are mostly integrated with manual prescription-filling processes. (Holden et al., 2010)

Oswald and Caldwell (2007) studied the error rates of a tertiary care university hospital before and after implementing an automated pharmacy carousel system. They found that the system did reduce the error rates of the line item medication dispensed by the automatic system, but not the error rates of the manually filled prescription line items.

Lin et al. (2007) showed that the introduction of a robotic system in a hospital reduced prescription-filling time, but not inspection time. The “saved” time also did not translate to increased patient consultation time as is intended with these technology interventions. Similar results were observed by Angelo et al. (2005) who found that patient satisfaction and counselling rates with pharmacy services were unaffected by automation.
5.3 Summary

Continuous improvement methodologies and technology interventions require careful evaluation and consideration before implementation.

Process improvement methodologies are not necessarily empirical in nature, and the way they are approached and implemented play a large role in the success of improvement interventions.

Technological interventions require careful evaluation to ensure that actual performance obstacles are being removed through this implementation. Obstacles to improving errors and productivity, such as interruptions and mental workload, are not necessarily removed with technological upgrades.

Most importantly, both technological and continuous improvement methodologies require whole system changes. It is thus not just a matter of intervention at the institutional pharmacy, but changing the entire hospital operational activities.
Chapter 6

Conclusion

6.1 Project Summary

The research aim for this project was to identify technical inefficiencies and quality concerns in the institutional pharmacies of South African private hospital groups using DEA for improvement initiatives.

To achieve this, the research objectives were set as follows:

1. Determine the factors that influence the performance of institutional pharmacies through observation and literature study.

2. Map the institutional pharmacy inputs, outputs and processes through observation and interviews with health care professionals.

3. Apply and test DEA models and sensitivity analysis techniques in an institutional pharmacy case study.

4. Determine which continuous improvement methodologies can be used to improve pharmacies that already have 100% relative efficiency according to the DEA evaluation.

In Chapter 2, the first objective was addressed. Factors were identified from literature that can potentially affect pharmacy performance. These factors were classified into five categories. The physical environment was shown to affect pharmacy employees with regards to alignment of process and layout, ergonomic design of equipment and workspace and environmental factors such as lighting. With pharmacy process design, the importance of task allocation and sequencing was highlighted, as well as how responses to interruptions significantly affect pharmacy workers. Inventory management was shown to have various conflicting stakeholder objectives, and various mathematical models exist to determine optimal
formulary and inventory levels and reorder points. The literature on scheduling of employees was discussed, and examples of models to minimise labour cost, balance undesirable shifts and maximise employee shift preferences were shown. Cases where the advantages of non-traditional self-scheduling have been highlighted were also discussed. Lastly, studies on worker engagement and job satisfaction were discussed, and studies outside of the pharmacy realm that could possibly be considered for research were discussed.

In Chapter 3, various performance measurement techniques were discussed. It was highlighted that DEA was a superior analytical tool for cases where insight is required in a set of similar DMUs compared to traditional accounting techniques such as ratio analysis. The case studies where DEA had been applied in healthcare and pharmacy settings were discussed, and issues and notes for application highlighted.

DEA was then applied in a case study as discussed in Chapter 4 where the second and third objectives were met. It was found that it was impractical to include the only identified quality factor, namely the number of incident reports, with the current DEA models. This was because it resulted in high weightings being assigned to that factor during analysis, so that insight into efficiency shortcomings was unattainable. For this reason, it was excluded from the DEA, but included in the overall process for DMU analysis.

The examples of interpretation of the DEA results showed that the efficiency scores, weightings and slacks provided insight into inefficiencies in the institutional pharmacies. These inefficiencies can be related to the performance factors discussed in Chapter 2. This is especially valuable to operational pharmacy managers who need to firstly identify underperforming pharmacies in their regions, and also need guidance on improvement initiatives to implement.

The ultimate purpose of improvement is however not to have a group of pharmacies all rated 100% relatively efficient in relation to each other. The private health care industry remains competitive, and overall continuous improvement is required to advance and maintain a dominant market position. So, finally, in Chapter 5, the Continuous Improvement methodologies that can be introduced by the DMUs rated fully efficient were discussed to address objective 4.

The first focus area in this chapter was that of continuous process and quality improvement methodologies, such as the Toyota Production System, Lean Thinking, Six Sigma and Theory of Constraints. The case studies where these methodologies were implemented were discussed. Advanced pharmacy technologies were then discussed for consideration in CI.

It was thus shown that DEA can be used to great effect to identify technical inefficiencies in institutional pharmacies in private hospitals in South Africa. It was also shown though that quality metrics are not yet as readily available as
efficiency measurements, and the single metric does not aid insight for operational managers when included in DEA. The complete proposed method for using DEA in performance evaluations was illustrated in Figure 4.7.

6.2 Limitations of Current Study and Future Research Recommendations

As discussed above, even though DEA was used successfully in this study to gain insight into efficiency shortcomings, the inclusion of quality metrics in DEA can provide even more insight into DMU performance. Possible future research can include how to acquire more quality metrics for inclusion in the DMU analysis, and investigating if and how applying DEA would occur. An important note in such studies would be to ask how to compare and prioritise efficiency and quality metrics in a single evaluation where health care services are concerned.

One limitation of the current study was that the assumption was made that all pharmacy employees have the potential to be equally efficient. There have however been studies showing that factors such as age, relevant experience and even gender can influence pharmacists’ performance. Future studies can possibly take this into consideration by evaluating if a pharmacy’s success or failure is not heavily dependent on a specific outlier employee.

Other expansions on the current research could be to include a Malmquist Index (as discussed in section 3.3.5) to evaluate DMU efficiency over time as discussed in Chapter 2. This would give decision makers important insight into the efficacy of improvement initiatives.

Finally, continuous improvement initiatives (both process and technology improvements) require whole hospital adaptation of new technological systems and ways of working. In these cases, it might become difficult to draw a clear system boundary around the pharmacy for analytical purposes. For example, introduction of electronic prescription systems would require the prescribing physicians to be included in the pharmacy boundary, as knowledge of their use of the system would become important. Continuous improvement methodologies have also shown that fluid interaction between sub-systems, such as the ward nurses and the pharmacists, lead to better overall system improvement. In these cases, the DEA models as proposed in this thesis would no longer be relevant. Consideration for entire hospital DEA studies may have to be considered.

6.3 Research Value

The research value in this thesis lies in the unique combination of the literature studies with the application of DEA in institutional pharmacies. Most DEA studies
only provide the DEA results, without context of how to apply these findings in the evaluated system.

The collection, categorisation and discussion of the case studies for both the inefficiencies and continuous improvement of pharmacies are also unique and can provide insight for pharmacy managers.

A final thought on improvement initiatives: In a paper on quality management in the Western World presented at the Institute of Sciences in Osaka in 1989, Edward Deming said (Orsini, 2013):

"It is not enough for everyone to do his best. Everyone is already doing his best. Efforts, to be effective, must go in the right direction."
Bibliography


Dean, B., van Ackere, A., Gallivan, S., and Barber, N. When should pharmacists visit their wards? An application of simulation to planning hospital pharmacy services. *Health care management science*, 2:35–42, 1999.


Appendix A

Pharmacy Process Flow Diagrams

A.1 Process Credits
A.2 Process Ward Requisition Books

Start

- Pick ward requisition book
- Bill medication to patients
- Pick used medication from shelves
- Count out required quantities
- Return medication to shelf
- Note in schedule book
- Pack medication
- Make notations in ward schedule book
- Pack medication in ward cupboard
- Return to pharmacy

More requisition books?

- Yes
- No

Stop
A.3 Perform Rounds Process

1. Start
2. Collect deliverables
3. Walk to ward
4. Deliver any filled prescriptions
5. Receive signed delivery slips
6. Pick up receivables (new prescriptions/books)
7. All wards visited?
   - No: Repeat steps 3-6
   - Yes: Return to pharmacy
8. Process Credits
10. Sort prescriptions
11. Process Prescriptions
12. Stop
A.4 Dispensing Process

Start

Pick new prescription

Check prescription sanity

Have query? Yes

Phone prescribing doctor

Have query? No

Make notation of any changes

Capture patient data

Print labels, instructions and delivery sticker

Fetch label

Fetch packing medium

Any solutions/dilutions required? Yes

Make solutions/dilutions

Any solutions/dilutions required? No
Pick medication from shelves
Count out required quantities
Return medication to shelf

Any schedule 3 or 6 medication?  

| Yes | Pick medication from shelves  
|     | Count out required quantities  
|     | Return medication to shelf  
|     | Note in schedule book  

| No | Make notations on prescription document  
|    | Label  

Repeat prescription?  

| Yes | Make copy of prescription  
| No  | Pack medication, delivery sticker and prescription doc in bag  

Pack in ward container

More prescription documents?  

| Yes | Stop  
| No  | Stop  

Stop
A.5 Stock Ordering Process

Start

Check pharmacist's notes

Check system for required orders

Order available on market? Yes → Order available stock

No → Change balance of order

Order available from preferred supplier? Yes → Order available stock

No → Change balance of order

Order available from other supplier? Yes → Order available stock

No → Change balance of order

Add to out of stock list

Stop
A.6 Stock Receiving Process

Start

Receive order and invoice

Check for damages and actual goods against invoice

Order correct with no damages?

Yes

Pack away goods

No

Call supplier

Receive reference number

Pack away damaged/incomplete goods

Label package with order and reference number

Await supplier pickup

Receive invoice in system

Stop
A.7 Aged Stock Process

1. Start
2. Receive system notification of stock older than 6 months
3. Place stock on internal hospital group market
4. Wait for expiration date
5. Check if stock is still on shelf:
   - Yes: Collect stock from shelf
   - No: Write off stock in systems, Dispense of stock according to relevant regulation
6. Stop
Appendix B

DEA Models’ Python Source Code

B.1 CCR Multiplier Model

```python
__author__ = 'Nina Uys'

import numpy as np
import csv
from cvxopt import matrix
from cvxopt.modeling import op, variable

inputs = 6  # Defines number of inputs
outputs = 3  # Defines number of outputs
DMUs = 46  # Defines number of DMUs

# Reads in DMU input data from CSV file
x = np.genfromtxt('InputDataCorrected.csv', dtype=float, delimiter=',')
# Ensure correct matrix size - 6 inputs with 46 DMUs
x = np.reshape(x, (inputs, DMUs))
# Transposes matrix x for LP
x = matrix(x.T)

# Repeat for output data
y = np.genfromtxt('OutputData.csv', dtype=float, delimiter=',', )
y = np.reshape(y, (outputs, DMUs))
y = matrix(y.T)

# Define the output weight variable u and input weight variable v
u = variable(outputs)
v = variable(inputs)

# Define headings for DEA output file
columnheadings = ['DMU', 'Z']
for headings in range(1, 4):
```

B-1
columnheadings.append("u" + str(headings))
for headings in range(1,7):
    columnheadings.append("v" + str(headings))
# Write to CSV file
c = csv.writer(open("CCRInputMMResults.csv", "wb"), dialect='excel', delimiter=',')
c.writerow(columnheadings)

# DEA will be performed for each individual DMUm (entire for-loop per DMU)
for m in range(0, DMUs):
    # Define constraints
    ineq = (((y * u) - (x * v)) <= 0)
    eq = (x[m, :] * v == 1)
    c1 = (u >= 0.00000000000001)
    c2 = (v >= 0.00000000000001)
    # defines "lp" problem, objective function is multiplied by −1 as 'op' minimises
    lp = op(-1*(y[m, :] * u), [ineq, eq, c1, c2])
    # solve for defined linear problem "lp"
    lp.solve()

# prepare data for CSV output
#allows for correct variable definition by stripping non-number elements
z = lp.objective.value()
tempvariable = float(str(z).strip('\n').strip('[').strip(']'))
tempvariable = tempvariable * (-1)
# prepare output data for CSV files
data = [m, tempvariable]
for eachu in u:
    data.append(float(str(eachu.value()).strip('[').strip(']')))
for eachv in v:
    data.append(float(str(eachv.value()).strip('[').strip(']')))
# write to CSV
c.writerow(data)
B.2 CCR Envelopment Model

```python
__author__ = 'Nina Uys'
import numpy as np
import csv
from cvxopt import matrix
from cvxopt.modeling import op, variable, sum

x = matrix(np.genfromtxt('InputDataCorrected.csv', dtype=float, delimiter=','))
y = matrix(np.genfromtxt('OutputData.csv', dtype=float, delimiter=','))

inputs = 6  # Defines number of inputs
outputs = 3  # Defines number of outputs
DMUs = 46  # Defines number of DMUs

# Defines column headers for CSV file
columnheadings = ["DMU", "Theta"]
for headings in range(1,47):
    columnheadings.append("Lambda" + str(headings))
for headings in range(1,7):
    columnheadings.append("s−" + str(headings))
for headings in range(1,4):
    columnheadings.append("s+" + str(headings))
# Writes column headings to CSV file
with open("CCRInputEMResults.csv", "wb") as f:
    writer = csv.writer(f, dialect='excel', delimiter=',')
    writer.writerow(columnheadings)

# perform entire for loop for each DMU
for m in range(0,DMUs):
    # stage 1
    # define lambda array
    l = variable(DMUs)
    # define efficiency score theta variable
    t = variable()
    # define constraints
    eq1 = x*l - t*x[:,m] <= 0
    eq2 = y*l - y[:,m] >= 0
    eq3 = l >= 0
    # define stage 1 LP1, including objective function and constraints
    lp1 = op(t, [eq1, eq2, eq3])
    lp1.solve()
    tstar = lp1.objective.value()
    # stage 2
    # define slack variables
    si = variable(inputs)
```

B-3
so = variable(outputs)

# define stage 2 constraints
eq4 = (x*l - tstar*x[:,m] + si == 0)
eq5 = (y*l - y[:,m] - so == 0)
eq6 = (si >= 0)
eq7 = (so >= 0)

# define stage 2 LP2, including objective function and constraints (maximise slacks)
lp2 = op(-1*(sum(si) + sum(so)), [eq3, eq4, eq5, eq6, eq7])
lp2.solve()
maxs = lp2.objective.value()

# prepare data for CSV output
tempvariable = float(str(tstar).strip('
').strip('[').strip(']'))
data = [m, tempvariable]

# convert lambda, si and so to floats
for eachlambda in l:
data.append(float(str(eachlambda.value()).strip().strip('[]')))
for slack in si:
data.append(float(str(slack.value()).strip().strip('[]')))
for slack in so:
data.append(float(str(slack.value()).strip().strip('[]')))

# write to CSV
c.writerow(data)
B.3 BCC Multiplier Model

```python
__author__ = 'NinaUys'
import numpy as np
import csv
from cvxopt import matrix
from cvxopt.modeling import op, variable

inputs = 6  # Defines number of inputs
outputs = 3  # Defines number of outputs
DMUs = 46  # Defines number of DMUs

# Read in DMU input data from CSV file
x = np.genfromtxt('InputDataCorrected.csv', dtype=float, delimiter=',')
# Ensure correct matrix size - 6 inputs with 46 DMUs
x = np.reshape(x, (inputs, DMUs))
# Transposes matrix x for LP
x = matrix(x.T)

# Repeat for outputs
y = np.genfromtxt('OutputData.csv', dtype=float, delimiter=',')
y = np.reshape(y, (outputs, DMUs))
y = matrix(y.T)

# Define the intensity variables
u = variable(outputs)
v = variable(inputs)

# DMU under evaluation
u_0 = variable()

# Define headings for DEA output file
columnheadings = ['DMU', 'Z', 'U_0']
for headings in range(1, 4):
columnheadings.append('u' + str(headings))
for headings in range(1, 7):
columnheadings.append('v' + str(headings))

# Write to CSV file
c = csv.writer(open('BCCInputMMResults.csv', 'wb'), dialect='excel', delimiter=',')
c.writerow(columnheadings)

# DEA will be performed for each individual DMU m
for m in range(0, 46):
    ineq = ((y * u) - (x * v)) - u_0 <= 0
    eq = (x[m, :] * v == 1)
c1 = (u >= 0.0000001)
c2 = (v >= 0.0000001)
    # define objective function and constraints, multiplied by -1 as
```
'op' minimises
lp = op(-1*(y[m, :] * u - u_0), [ineq, eq, c1, c2])

# solve objective function
lp.solve()
z = lp.objective.value()
tempvariable2 = u_0.value

print u_0

# allows for correct variable definition by stripping non-number characters
tempvariable = float(str(z).strip(\n).strip([' '] .strip ('))
tempvariable = tempvariable * (-1) # minimisation value

tempvariable2 =
    float(str(tempvariable2).strip().strip(\n).strip([' '] .strip ('))

# prepare output data for CSV files
data = [m, tempvariable, u_0]
for eachu in u:
    data.append(float(str(eachu.value()).strip().strip([' '] .strip (')))
for eachv in v:
    data.append(float(str(eachv.value()).strip().strip([' '] .strip (')))

# write to CSV
c.writerow(data)
B.4 BCC Envelopment Model

```python
__author__ = 'NinaUys'
import numpy as np
import csv
from cvxopt import matrix
from cvxopt.modeling import op, variable, sum

x = matrix(np.genfromtxt('InputDataCorrected.csv', dtype=float,
delimiter=','))
y = matrix(np.genfromtxt('OutputData.csv', dtype=float,
delimiter=','))

inputs = 6  # Defines number of inputs
outputs = 3  # Defines number of outputs
DMUs = 46  # Defines number of DMUs

# Defines column headers for CSV file
columnheadings = ['DMU', 'Theta']
for headings in range(1,47):
    columnheadings.append("Lambda" + str(headings))
for headings in range(1,7):
    columnheadings.append("s−" + str(headings))
for headings in range(1,4):
    columnheadings.append("s+" + str(headings))
# Writes column headings to CSV file
c = csv.writer(open("BCCInputEMResults.csv", "wb"), dialect='excel', delimiter=',')
c.writerow(columnheadings)

for m in range(0,DMUs):
    # define lambda array
    l = variable(DMUs)
    # define theta value
    t = variable()
    eq1 = x*l - t*x[:,:] <= 0
    eq2 = y*l - y[:,:] >= 0
    # Additional BCC condition
    eq3 = l >= 0
    eq4 = sum(1) == 1
    # define stage 1 - determine theta
    lp1 = op(t, [eq1, eq2, eq3, eq4])
    # solve stage 1
    lp1.solve()
    # start stage 2 to optimise slacks
    # tstar is the calculated theta value from stage1
    tstar = lp1.objective.value()
    # define slack variables
    si = variable(inputs)
```

so = variable(outputs)

# define constraints
eq5 = (x*l - tstar*x[:,m] + si == 0)

eq6 = (y*l - y[:,m] - so == 0)

eq7 = (si >= 0)

eq8 = (so >= 0)

# define objective function to solve stage 2 (maximise slacks)
lp2 = op(-1*(sum(si) + sum(so)), [eq3, eq4, eq5, eq6, eq7, eq8])

# solve stage 2
lp2.solve()

maxs = lp2.objective.value()

# prepare data for CSV output

tempvariable =
    float(str(tstar).strip('n').strip('[').strip('])'))

data = [m, tempvariable]

# convert lambda, si and so to floats
for eachlambda in l:
    data.append(float(str(eachlambda.value()).strip().strip('[[ ]])))

for slack in si:
    data.append(float(str(slack.value()).strip().strip('[[ ]))))

for slack in so:
    data.append(float(str(slack.value()).strip().strip('[[ ]))))

# write to CSV
c.writerow(data)
B.5 Slacks Based Measure Multiplier Model

```python
__author__ = 'NinaUys'
import numpy as np
import csv
from cvxopt import matrix
from cvxopt.modeling import op, variable

# Import data from CSV files
x = (np.genfromtxt('InputDataCorrected.csv', dtype=float, delimiter=','))
y = (np.genfromtxt('OutputData.csv', dtype=float, delimiter=','))

# define new matrix, with arbitrary values 1–276
newx = matrix(range(276), (6,46), 'd')
# go through each value in x to prevent dividing with 0
for i in range(0,46):
    for j in range(0,6):
        if x[j, i] == 0.0:
            # for every zero value in x, make newx value 0 too
            newx[j, i] = 0.0
        else:
            # for all non-zero values define newx value as inverse of x value
            newx[j, i] = (1.0 / x[j, i])

# repeat for outputs
newy = matrix(range(138), (3,46), 'd')
for i in range(0,46):
    for j in range(0,3):
        if y[j, i] == 0.0:
            newy[j, i] = 0.0
        else:
            newy[j, i] = (1.0 / y[j, i])

# Transpose matrices for LP
x = matrix(x.T)
y = matrix(y.T)

# Redefine as real value – Python requirements for division
m = 6.0
s = 3.0

# Defines column headers for CSV file
columnheadings = ['"DMU", "Z"
for headings in range(1,4):
    columnheadings.append("u" + str(headings))
for headings in range(1,7):
    columnheadings.append("v" + str(headings))
```
# Write to CSV file

c = csv.writer(open("AddInputMMResults.csv", "wb"), dialect='excel', delimiter=',')
c.writerow(columnheadings)

# Writes column headings to CSV file

for n in range(0, 46):
    # define intensity variables
    u = variable(3)
    v = variable(6)
    # define constraints
    eq1 = ((y * u) - (x * v) <= 0)
    eq2 = (v - 1 / m * (newx[:, n])) >= 0
    eq3 = (u - (1 - x[n, :] * v + y[n, :] * u) / s * (newy[:, n])) >= 0
    # define objective function
    lp = op((x[n, :] * v) - (y[n, :] * u + 1), [eq1, eq2, eq3])
    # solve objective function
    lp.solve()
    # Change to maximisation answer
    z = -1 * lp.objective.value()

    # prepare output data for CSV files
    tempvariable = float(str(z).strip(\n').strip(['').strip(']'))
data = [n, tempvariable]
for eachu in u:
    data.append(float(str(eachu.value()).strip(['\').strip(']')))
for eachv in v:
    data.append(float(str(eachv.value()).strip(['\').strip(']')))

# write to CSV

c.writerow(data)
B.6 Slacks Based Measure Envelopment Model

```python
# Import numpy and csv
import numpy as np
import csv
from cvxopt import matrix
from cvxopt.modeling import op, variable

# Import data from CSV files
x = matrix(np.genfromtxt('InputDataCorrected.csv', dtype=float, delimiter=','))
y = matrix(np.genfromtxt('OutputData.csv', dtype=float, delimiter=','))

inputs = 6  # Defines number of inputs
outputs = 3  # Defines number of outputs
DMUs = 46  # Defines number of DMUs

# Define new matrix, with arbitrary values 1–276
newx = matrix(range(276), (inputs, DMUs), 'd')
# go through each value in x to prevent dividing with 0
for i in range(0, DMUs):
    for j in range(0, inputs):
        # for every zero value in x, make newx value 0 too
        if x[j, i] == 0.0:
            newx[j, i] = 0.0
        else:
            # for all non-zero values define newx value as inverse of x value
            newx[j, i] = (1.0 / x[j, i])

# Repeat for outputs
newy = matrix(range(138), (outputs, DMUs), 'd')
for i in range(0, DMUs):
    for j in range(0, outputs):
        if y[j, i] == 0.0:
            newy[j, i] = 0.0
        else:
            newy[j, i] = (1.0 / y[j, i])

# Redefine as real value - Python requirements for division
inputsR = 6.0
outputsR = 3.0

# Defines column headers for CSV file
columnheadings = ['DMU', 'Theta']
for headings in range(1, 47):
    columnheadings.append('Lambda' + str(headings))
for headings in range(1, 7):
```

B-11
columnheadings.append("s−" + str(headings))

for headings in range(1,4):
    columnheadings.append("s+" + str(headings))

for headings in range(1,4):
    columnheadings.append("s+" + str(headings))

# Writes column headings to CSV file

c = csv.writer(open("AddInputEMResults.csv", "wb"), dialect='excel', delimiter=',')
c.writerow(columnheadings)

for n in range(0,DMUs): # Optimisation performed individually for all DMUs in the set
    Si = variable() # initialise variables
    So = variable(inputs) # variables converted for LP, will transform back at end
    L = variable(DMUs)
    
    # constraints

    eq1 = t + 1/outputsR * (So[0] * newy[0, n] + So[1] * newy[1, n] + So[2] * newy[2, n]) - 1 == 0
    eq2 = x * L + Si - t*x[:, n] == 0
    eq3 = y * L - So - t*y[:, n] == 0
    eq4 = L >= 0
    eq5 = Si >= 0
    eq6 = So >= 0
    eq7 = t >= 0.00000000000001 # cannot be zero, function does not allow >
    
    # define lp as the objective function

             * newx[4, n] + Si[5] * newx[5, n]), [eq1, eq2, eq3, eq4, eq5, eq6, eq7])

    lp.solve()

tau = lp.objective.value() # tau is the optimised objective function value

# convert back to original DEA non-linear model values

l = L
l.value = l.value / t.value
so = So
so.value = so.value / t.value
si = Si
si.value = si.value / t.value
t2 = t.value

# ready output to write to CSV
# convert tau to float value

tempvariable = float(str(tau).strip('\n').strip('[').strip(']'))

# include DMU # and tau in output csv

data = [n, tempvariable]

# convert lambda, si and so to floats
for eachlambda in l:
    data.append(float(str(eachlambda.value()).strip().strip('[]')))
for slack in si:
    data.append(float(str(slack.value()).strip().strip('[]')))
for slack in so:
    data.append(float(str(slack.value()).strip().strip('[]')))
# write to CSV
c.writerow(data)
B.7 BCC Sensitivity Analysis Model for Efficient DMUs

```python
__author__ = 'Nina'
import numpy as np
import csv
from cvxopt import matrix, solvers
from cvxopt.modeling import op, dot, variable, sum

# Reads in DMU input data from CSV file
data = np.genfromtxt('InputDataCorrected.csv', dtype=float, delimiter='
')
x = data
# sum the values for each column
x_DMU = np.sum(x, axis=0)
x_DMU = np.reshape(x_DMU, (1, 46))
x_DMU = matrix(x_DMU)

# Reads in DMU output data from CSV file
data = np.genfromtxt('OutputData.csv', dtype=float, delimiter='
')
y = data
# sum the values for each column
y_DMU = np.sum(y, axis=0)
y_DMU = np.reshape(y_DMU, (1, 46))
y_DMU = matrix(y_DMU)

# define intensity variable
lambdaj = variable(46)
slackoutput = variable(3)
slackinput = variable(6)

# define radius of stability variable
rho1 = variable(1)

# Defines column headers for CSV file
columnheadings = ["DMU", "Rho"]
for headings in range(1,47):
    columnheadings.append("Lambda" + str(headings))
for headings in range(1,7):
    columnheadings.append("s−" + str(headings))
for headings in range(1,4):
    columnheadings.append("s+" + str(headings))

# Writes column headings to CSV file
c = csv.writer(open("SensitivityAnalysisEff.csv", "wb"),
               dialect='excel', delimiter=',
')
c.writerow(columnheadings)

for m in range(0, 46):
    # define constraints
    eq1 = (x_DMU * lambdaj - x_DMU[m] * lambdaj[m] + slackinput - rho1
           - x_DMU[m] == 0)
```

B-14
eq2 = (y[i]DMU * lambdaj - y[i]DMU[m] * lambdaj[m] - slackoutput + 
        rho1 - y[i]DMU[m] == 0)
eq3 = (\sum(lambdaj) - lambdaj[m] == 1)
c1 = (lambdaj >= 0)
c2 = (slackinput >0)
c3 = (slackoutput >0)
c4 = (rho1 >0)
# define LP, objective function and constraints
lp = op(rho1, [eq1, eq2, eq3, c1, c2, c3, c4])
# solve LP
lp.solve()

z = lp.objective.value()
tempvariable = float(str(z).strip('n').strip('[').strip(']'))
# prepare data for CSV output
data = [m, tempvariable]
# convert lambda, si and so to floats
for eachlambda in lambdaj:
    data.append(float(str(eachlambda.value()).strip().strip('[')))
for slack in slackinput:
    data.append(float(str(slack.value()).strip().strip('[')))
for slack in slackoutput:
    data.append(float(str(slack.value()).strip().strip('[')))
# write to CSV
c.writerow(data)
B.8 BCC Sensitivity Analysis Model for Inefficient DMUs

```python
__author__ = 'Nina'
import numpy as np
import csv
from cvxopt import matrix, solvers
from cvxopt.modeling import op, dot, variable, sum

# Reads in DMU input data from CSV file
data = np.genfromtxt('InputDataCorrected.csv', dtype=float, delimiter=' , ')
x = data
# sum the values for each column
x_DMU = np.sum(x, axis=0)
x_DMU = np.reshape(x_DMU, (1, 46))
x_DMU = matrix(x_DMU)

# Reads in DMU output data from CSV file
data = np.genfromtxt('OutputData.csv', dtype=float, delimiter=' , ')
y = data
# sum the values for each column
y_DMU = np.sum(y, axis=0)
y_DMU = np.reshape(y_DMU, (1, 46))
y_DMU = matrix(y_DMU)

# define intensity variable
lambdaj = variable(46)
slackoutput = variable(3)
slackinput = variable(6)

# define radius of stability variable
rho1 = variable(1)

# Defines column headers for CSV file
columnheadings = ['DMU', 'Rho']
for headings in range(1, 47):
    columnheadings.append("Lambda" + str(headings))
for headings in range(1, 7):
    columnheadings.append("s-" + str(headings))
for headings in range(1, 4):
    columnheadings.append("s+")

# Writes column headings to CSV file
with open("SensitivityAnalysisInEff.csv", "wb") as c:
    c.writerow(columnheadings)

for m in range(0, 46):
    # define constraints
```

eq1 = (x_{DMU} \cdot \lambda_j + \text{slack input} + \rho_1 - x_{\text{DMU}[m]} == 0)
eq2 = (y_{\text{DMU}} \cdot \lambda_j - \text{slack output} - \rho_1 - y_{\text{DMU}[m]} == 0)
eq3 = (\sum(\lambda_j) == 1)
c1 = (\lambda_j >= 0)
c2 = (\text{slack input} >= 0)
c3 = (\text{slack output} >= 0)
c4 = (\rho_1 >= 0)
# define lp, includes objective function and constraints
# -1 because cvxopt minimises function, whereas this is a maximisation problem
lp = op(-1*\rho_1, [eq1, eq2, eq3, c1, c2, c3, c4])
# solve lp
lp.solve()

z = lp.objective.value()
tempvariable = float(str(z).strip(’\n’).strip(’[’).strip(’]’))
tempvariable = tempvariable * -1
# prepare data for CSV output
data = [m, tempvariable]
# convert lambda, si and so to floats
for eachlambda in lambda:
    data.append(float(str(eachlambda.value()).strip().strip(’[’))))
for slack in slack input:
    data.append(float(str(slack.value()).strip().strip(’[’))))
for slack in slack output:
    data.append(float(str(slack.value()).strip().strip(’[’))))
# write to CSV
c.writerow(data)