BETTER QUEUE MANAGEMENT IN A BUSY PUBLIC HOSPITAL OF A DEVELOPING COUNTRY WITHOUT APPOINTMENT SYSTEM: AN APPLICATION USING DATA ENVELOPMENT ANALYSIS

Komal Aqeel Safdar

Doctor of Philosophy

ASTON UNIVERSITY

February 2016

ACKNOWLEDGEMENTS

First and foremost, I would like to thank the Almighty Allah for blessing me with the opportunity, strength and determination to complete the challenging task of PhD.

I am extremely grateful to my loving family, my Mom, Dad and brother, for allowing me to pursue a PhD away from home here in UK, and for their continuous support, encouragement and motivation through every step of PhD.

I am much obliged to Aston Business School for considering me for Graduate Teaching Assistant Scholarship (GTA), allowing me to undertake a PhD degree.

I am most grateful to my PhD supervisor Prof. Ali Emrouznejad and Associate supervisor Prof. Prasanta Kumar Dey, for accepting me as their PhD student, and for their continuous and useful guidance at every stage of my PhD.

I am very grateful to the Ethical committee of the designated hospital in Pakistan for permitting me to conduct research as well as all staff members for giving up their valuable time, useful comments and assistance.

I would also like to thank all my colleagues/academics/conference audience for their useful insights with regard to my research. I am deeply obliged to the anonymous experts on discussion boards and forums online for their worthwhile suggestions.

Lastly, my sincere thanks to the Research Degrees Programme (RDP) team for providing a comfortable research environment and for dealing with all academic and non-academic queries during my time here at Aston as a PhD researcher.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 5

CHAPTER 6

LIST OF APPENDICES

LIST OF ACRONYMS

A&E	Accident and Emergency
ABC	Activity-based Costing
ACU	Ambulatory Care Unit
ANN	Artificial Neural Network
AR	Assurance Region
BCC	Banker, Charnes and Cooper (1984)
CART	Classification and Regression Tree
CCR	Charnes, Cooper and Rhodes (1978)
CRS	Constant Returns to Scale
CTMC	Continuous Time Markov Chain
DEA	Data Envelopment Analysis
DHA	District Health Authority
DMU	Decision-making Unit
DRG	Diagnosis-related Group
DRS	Decreasing Returns to Scale

ED	Emergency Department
EI	Efficiency Index
ENT	Ear, Nose and Throat
FCFS	First-come-first-served
FDH	Free Disposal Hull
FHSA	Family Health Service Authority
GDMO	General Duty Medical Officer
GDP	Gross Domestic Product
GP	General Practitioner
ICU	Intensive Care Unit
IRS	Increasing Returns to Scale
IT	Information Technology
MPI	Malmquist Productivity Index
MR	Medical Registration
NHS	National Health Services
OECD	Organization for Economic Cooperation and Development
OPD	Outpatients' Department
OR	Operational Research
PI	Performance Indicator
PIM	Performance Improvement Management
SAS	Statistical Analysis System
SFA	Stochastic Frontier Analysis
SPSS	Statistical Package for the Social Sciences
VAH	Volunteering Agency Hospital
VBA	Visual Basic for Applications
VRS	Variable Returns to Scale
WHO	World Health Organization

CHAPTER 1 INTRODUCTION

Chapter Overview

The current chapter aims to identify various healthcare issues challenges faced by the developing world. These issues include a brief overview of the main health-related aspects. Additionally, a number of factors have been highlighted which limit the smooth running of health delivery systems within these countries. Furthermore, the health profile of Pakistan along with other related factors has been presented, demonstrating the deficient health situation and hardships endured by patients in a developing country. Several efficiency assessment measures of health systems are explored. It has been emphasized that queuing is an important criterion which represents the efficiency of health services and the level of patient care provided. Moreover, a few characteristics of health institutions and their effect on the queuing situation in developing countries have been identified. The chapter will conclude with a brief outline of the current research study.

The existence of any nation is a function of the survival of its citizens, which in turn depends on the provision of the health care facilities. Health, not only enables the people to lead a socially and economically productive life, but it also plays a significant role in the economic growth of a country (Ramanathan 2005; Adeleke *et al.,* 2009).

The health care service providers deal with multiple issues on a daily basis. Factors such as variability in patients' complaints, daily and weekly workload fluctuations and multiple objectives within and between facilities make it extremely challenging to understand the operational performance of a large hospital sub-system (Matta and Patterson, 2007). Given the increasing cost pressures, complexity of diseases and demand of quality and efficacy from highly aware and educated patients due to advances in technology and telecommunications, efficient and smooth provision of health services is becoming extremely important (Ramanathan 2005; Mehandiritta 2011). Despite these issues, there is a need of continuous assessment of the operational efficiency of hospitals. This allows the decision-makers to develop a better understanding of the management effectiveness (Chuang *et al.,* 2011), reflect on current practices and improve performance (Baril *et al.,* 2016), and provide valuable insights for enhancing the level of patient care.

1.1 Healthcare Issues in Developing Countries

The developing world accounts for more than three-quarters of the world's population. The gap between the developing and developed nations in terms of income and health expenditures, leads to a huge difference in the overall health infrastructure and health care delivery strategies adopted. (World Health Statistics, 2015; Gaimard 2014).

Figure 1-1a: Poor Sanitation in Developing Countries
(Source: Maynard 2014)

Figure 1-1b: Poor Condition of Hospitals in Developing Countries
(Source: Venkatasubramanian 2015)

The developing world faces numerous *health-related* issues (as shown in Figure 1-1a and 1-1b). The poor health indicators in these countries demonstrate the issues that they encounter. For instance, the infant mortality rate in low income countries is nearly 50%, as compared to high income countries which is around 5% to 15%. Similarly, the maternal mortality ratio in the developing countries is at least 250 (per 100,000 live births) as compared to around 20 to 50 for the developed nations. The life expectancy at birth for the developing nations is around 50 years, whereas for the developed nations, it is around 75 years. Furthermore, the incidence level of infectious diseases such as HIV/AIDS, Tuberculosis and Malaria is much higher in developing countries. Hence, the incidence level is extremely high in the low income countries, that is, more than 80, 11,000 and 240 (per 100,000 of population), as compared to only around 25, 210 and 20 for the developed countries for HIV/AIDS, Malaria and Tuberculosis respectively (World Health Statistics 2015; Global Health Observatory (GHO) Data, WHO). A comparison of developing and developed nations in terms of a few widely used 'health-related' indicators is given in Table 1-1.

Additionally, *other factors* such as lack of antenatal and postnatal care, immunization, nutritional deficiencies, clean drinking water and sanitation further increase the burden of health problems faced by the developing world (McCarthy *et al.,* 2012). The extent of difference in health indicators of developing nations, as compared to the developed nations, signifies the poor health of citizens living in these countries.

Table 1-1: Comparison of Significant Health-related Indicators in Developing and Developed World

(*Source: World Health Statistics Report 2015*)

The developing countries have an increased burden of diseases and at the same time, the health systems experience issues such as lack of financial means (Dhungel 2014), shortage of human resources, and lack of infrastructure; where these constraints hinder their capability to provide effective patient care (Peters *et al.,* 2009). The total expenditure on health as percentage of Gross Domestic Product (GDP) is only around 4% in low income countries, whereas in high income countries this percentage increases to nearly 10%. Also, the general government expenditure on hospital care (as % of total expenditure on health) is less than 40% for developing countries, as compared to nearly 60% for the developed nations (World Health Statistics, 2015; Gaimard 2014).

There is a huge gap between the demand of services and availability of resources to cater for this demand (Chandra 2015). The lack of resources and infrastructure hinders effective healthcare service, and is usually indicated by hospital beds (MBHSS, WHO 2010) and physicians (Jamison *et al.,* 2006).

Most developing nations have only 2 to 5 physicians (per 10,000 of population). The developed nations are considerably better with at least around 15 to 30 physicians. Similarly, the number of hospital beds varies from 10 to 20 (per 10,000 of population) in developing countries, whereas it is around 40 to 50 in the developed world. The number of hospitals is less than 1 (per 100,000 of population) in most under-developed countries (World Health Statistics 2015; Country Profile, WHO). In addition to the physicians and nurses, the developing countries face the shortage of other health personnel such as pharmacists, laboratory technicians, emergency medical personnel, public health specialists and administrative staff. The limited number of health personnel is more critical in the developing world, particularly Sub-Saharan Africa and South-Asia (Driouchi 2014).

Developing economies also suffer from the *inadequate allocation* of human resources, where on a macro level, it concerns with the disparity in the distribution of health care workers between regions, between rural and urban areas, and between public and private sectors (Driouchi 2014). On a micro level, this imbalance relates to the inadequate distribution of health personnel that prevails within one busy health institution. Poor planning, organisation, coordination and communication, also result in lack of flexibility in distribution of human and other resources; which are not synced with the high demand of services (Chandra 2015).

Although progress has been observed in the GDP growth rate and economic prosperity in developing countries in the past few years, however, increasing population is another fundamental issue which diminishes the overall positive effect on the economy (Ragan and Lipsey 2005; Gaimard 2014); and health service provision becomes insufficient and may lag (Pol and Thomas 2002). Other issues include limited access to health services especially by the rural population due to geographical distances and financial constraints (Geyndt 1995); and low education level resulting in poor understanding of healthcare information and lack of basic knowledge of preventive healthcare (Amzat and Razum 2014). Some additional problems are high level of poverty resulting in poor nutrition and high incidence of diseases (Amzat and Razum 2014), lack of training for health personnel, lack of information for finding bottlenecks and improving health provision, and lack of strategic planning to monitor the service provision (Sheikh *et al.,* 2015). Provision of effective health services is a global challenge, however, the health delivery issues are more acute in developing countries. Therefore, there is a dire need to investigate and address the numerous health care delivery problems, specifically from a developing country perspective, to facilitate patients.

1.2 Health Profile of Pakistan: A Representative of Developing Countries

The extent of variability between developed and developing world with regard to the level of healthcare services provided is considerable. Numerous factors, mainly limited resources and budgets, increased population and low literacy rates, account for poor health care condition of patients in developing nations, where *Pakistan* is one of them.

Pakistan is situated in the South-Asian region. It is the sixth most highly populated country in the world, with a population of nearly 180 million. More than half of the population resides in rural areas (Nishtar *et al.,* 2013) and is living below the poverty line (Cassum and Shah 2014). A simple map of Pakistan is shown in Figure 1-2, highlighting its four provinces and Kashmir, and its neighbouring countries.

Figure 1-2: Map of Pakistan

(Source: Pakistan in Pictures, 2010)

The *literacy rate* in Pakistan is quite low, only around 50%. This rate is much lower than almost 100% literacy level of the developed nations of the world (including UK, USA, Japan, Russia and others). Moreover, Pakistan's literacy rate is lower than that of some developing countries (such as India, Bangladesh, Mexico and others), which have a literacy rate of at least 60%. Also, the *GDP* per capita for Pakistan is only around $1200, which is lower than some developing countries (such as India, Indonesia, Nigeria and others) which is at least $1500. Additionally, the GDP per capita for most of the developed nations is not less than approximately $40,000 (including Japan, Germany, UK, USA and others) (WHO Statistics 2015). These characteristics create hurdles in facilitating the citizens even with primary healthcare services in a highly under-developed country like Pakistan (see Table 1-2 below).

lower than most developed countries which have total health expenditures of around 10% (while USA spends around 20%) (WHO Statistics 2015) (see Figure 1-3 below).

Figure 1-3: Health Expenditures of Pakistan as Compared to Other Countries

(Source: World Health Statistics 2015)

Furthermore, developing countries such as Pakistan experience dearth of resources. The two main measures which are commonly associated with the efficiency of healthcare delivery systems are number of *beds* and *physicians*. In Pakistan, the number of beds and physicians (per 10,000 of population) is 6 and 8 respectively. The statistics are almost similar for other developing nations. However, in the developing world (including Russia, Japan and Germany), on average, there are 100 beds and 35 physicians (WHO Statistics 2014 and 2015) (see Figure 1-4 below).

Figure 1-4: Number of Hospital Beds and Physicians of Pakistan as Compared to Other Countries

(Sources: World Health Statistics 2014 and 2015)

Therefore, Pakistan as well as other similar developing countries experience a major deficiency of financing, medical and non-medical personnel, medicines/drugs and functioning equipment (Fleba 2009). Additionally, the *demand* of health services by far exceeds the *availability* of resources; hence resulting in overloaded health systems. Consequently, it is extremely challenging for the administration to optimize the use of resources, and efficiently manage and organize such huge numbers of patients, while providing health services of acceptable quality to maximum population.

1.3 Queueing: An Important Efficiency Criterion

The pursuit of obtaining increased efficiency is a central objective within all health systems. Policy makers and administrators are keen to obtain consistency between the health expenditure spent and the preferences or expectations of the citizens residing in a country (Jacobs *et al.,* 2006). Therefore, the *efficiency* assessment of health institutions and the extent of usefulness of results (Hollingsworth 2012) has always been a matter of increased interest for government bodies and policy makers. It is a continuous process which targets at reducing the gap between the varied goals of health care organisations and high level of patient care. In this regard, *Data Envelopment Analysis (DEA)* is a popular efficiency assessment technique (O'Neill *et al.,* 2008). The results from DEA analysis are very useful, since they demonstrate not only the current status of the system under evaluation but also provides recommendations for improvement.

Generally, efficiency assessment using some specific indicators allows for the comparison of different ways of achieving the same goal in order to provide an optimal solution; by assessing the relationship between an outcome or output and the resources required to produce it (Rozner 2013). Efficiency indicators are particularly useful to assess the current availability and optimal utilization of human resources to provide health services; such as total number of staff employed, number of patients served per employee, average time taken by an employee to conduct a certain, number of trained personnel, and others (Rozner 2013; MBHSS: WHO 2010). Additionally, the capacity of some other resources to provide services can also be considered for efficiency assessment such as beds, equipment, medicines/basic supplies, laboratories and operation theatres, diagnostics, and others. *Scarcity*, along with *inappropriate allocation* of resources, is a matter of utmost concern for the health service providers because it hinders meeting the demands and needs of the population in an effective and efficient way (Hollingsworth and Peacock 2008; MBHSS: WHO 2010).

The healthcare costs are increasing globally with escalating public demand of high quality services (Jacobs *et al.,* 2006). Hence, efficiency assessment can also be conducted with regard to costs of health services. The efficiency indicators can include costs related to staff including salary and benefits per full-time equivalent or net patient revenue per full-time equivalent, cost per discharge for inpatient cost, cost per outpatient visit, costs related to supplies/drugs (Cleverly and Cameron 2007), administrative costs (OECD 2010), and others.

Additionally, some other indicators include inpatient/outpatient visits (MBHSS: WHO 2010), diagnosis error rates, order fill rates, stock wastage due to damage/expiration, and others (MBHSS: WHO 2010; Peters *et al.,* 2009). Furthermore, some efficiency indicators specifically relate to patient type. For instance, indicators for inpatient care include acute bed occupancy and turnover rates, number of admissions, number of discharges, average length of stay, and others (OECD 2010; MBHSS: WHO 2010). Criteria for outpatients include number of consultations per doctor and percentage of outpatient surgeries resulting in same-day discharge (Rozner 2013; OECD 2010); whereas factors for emergency department include number of patients seen within an hour given current staff availability and percentage of emergency visits with patients seen in less than fifteen minutes, and others (Rozner 2013).

The health institutions are in search of possibilities which can assist them in reducing costs with effective utilization of limited funds, while improving the quality of health services and increasing patient satisfaction; emphasizing on the crucial need of efficiency assessment of health institutions (Daultani *et al.,* 2016; Tan 2013).

Another extremely important criterion for efficiency measurement within the service industry, is the *waiting time* of customers. With regard to healthcare services, the queuing problem is immensely crucial due to the involvement of patients. Excessive waiting and service times can reduce customer demand, increase the cost of medical care, constitute a barrier to effective treatment (Biju and Naeema 2011; Yeboah and Thomas 2010; Silva and Serra 2008); and cause dissatisfaction among the patients (Oche and Adamu 2014) where they might leave the system without receiving the service. Due to increasing competition, healthcare organizations strive to provide fast and efficient health services to attract more patients. In this case, low patient waiting time is a significant factor which demonstrates improved efficiency (Mensah *et al.,* 2015), in addition to high patient throughput, a short length of stay at the clinic, and low clinic overtime, while maintaining adequate staff utilization rates and low physician idle times (Jun *et al.,* 1999). Hence, *queuing* is a key indicator of the level of patient care and smooth patient flow which requires continuous assessment.

1.4 Absence of Appointment Systems and Queuing: A Developing Country Perspective

Although *'Queuing'* is a global issue faced by health care systems all over the world, however, the problem is more severe in the *developing world* particularly in large *public hospitals* which cater to a high percentage of the population (as shown in Figure 1-5). The developing countries are not able to adequately meet the health care needs of the public due to extensive inefficiency within hospitals. Some of the crucial issues faced by healthcare delivery systems in these countries include high incidence of diseases, managing increasing demand, dearth of doctors/other health personnel, lack of medical equipment and aftercare facilities, inadequate distribution of resources, late arrival of doctors, and high population and low literacy rate (Mensah *et al.,* 2015; Bhattacharjee and Ray 2014; Zere *et al.,* 2001; Babes and Sarma 1991). As a consequence, the health systems are overcrowded with long queues and excessive waiting time for the patients.

Figure 1-5: Excessive Queuing in Developing Countries

(Sources: Velasquez 2014 (top left), Sinclair 2009 (top right), Bird 2012 (bottom left), Habib 2012 (bottom right))

The issue of waiting times and queue assessment has been widely researched in healthcare, however, maximum studies have been conducted in the developed world. Furthermore, very few studies have analysed the queuing systems in the developing world

(Afrane and Appah, 2014), with almost *negligible* studies in South-Asian region and no study in Pakistan. Hence, there is a dire need to assess and improve the existing queue systems specifically within large busy public hospitals of developing countries, given the current operational challenges, in order to reduce the excessive wait times of patients.

Pakistan as well as some other developing countries, lack *General Practitioner (GP)* referral systems (Afsar and Younus 2004; Naqvi *et al.,* 2009; Hashmi *et al.,* 2010), unlike developed nations. Almost all large public hospitals provide primary, secondary and tertiary care facilities to all patients. Even for a minor ailment, the patients have to visit these large unwieldy hospitals. Additionally, uncertainties due to randomness of arrival pattern of walk-in patients, multiple stages to be followed by patients from point of entry till exit, lack of coordination for optimal utilization of resources, and cumbersome administrative work and lay procedures; result in complex patient flow system and enhanced wait times, leading to increased patient inconvenience (Bhattacharjee and Ray 2014; Chandra 2015).

Healthcare institutions in the developed world typically have a dedicated Appointment System, which not only assists the hospital management to deal with patient flow and organize resources effectively, but also facilitates patients. In most developing countries, almost all health organizations are overcrowded and *do not* have a set *Appointment System* within the *outpatients'* department; creating hurdles in the efficient management and guidance for patients, resulting in long wait times (Wei *et al.,* 2015). Very few studies have considered *walk-in* patients, and even in these studies, the main emphasis was improving an already existing appointment system (Fetter and Thompson 1966; Rising *et al.,* 1973; Ashton *et al.,* 2005; Cayirli and Gunes 2014). There are negligible studies which have evaluated a queue system where *'all'* patients are walk-in. Hence, there is a necessity to evaluate the queue system specifically from a developing country perspective such as Pakistan in this case, given the absence of appointments.

In these countries, the appointment systems are not currently implementable due to a number of reasons. Low level of education (Amzat and Razum 2014) makes it extremely challenging to communicate with patients even regarding the most basic instructions. Therefore, scheduling appointments for patients and for them to remember their appointments is nearly impossible. Most busy public hospitals in large cities cater to many surrounding towns and villages involving long hours of travelling resulting in limited access (Geyndt 1995). Additionally, the public transport is not very reliable with variations in availability. Due to lack of proper roads in most of these villages/towns, people have to use multiple means of transport, which frequently lead to excessive delays. Therefore, there is a high probability that the patients will miss their appointments. This will lead to chaos as these patients might turn up late on the same day or the next day demanding to be examined at that time. It will be extremely challenging for the staff to adjust these patients with those who already have appointments. Due to lack of proper health information systems and communication (Nishtar *et al.,* 2013), there might be a delay in conveying information to patients. Post is not reliable, and might not reach people in villages/towns on time. Online appointment systems are not feasible as well, since most people do not have this facility. Besides, low literacy rates prevent people to keep track of information provided through different sources. Due to lack of independent GP referral systems (Kilic *et al.,* 2015), the public hospitals bear the responsibility of providing GP as well as specialist services (Wei *et al.,* 2015). It will be extremely challenging to manage appointments for both services for huge number of patients and to keep a track of them. If referred by the GP-equivalent doctors, the patients might be asked to come another day to see the specialist depending which appointment they get, which they might not remember. Due to issues such as work schedules, lack of transport and long travelling time, it is extremely inconvenient for patients to visit these hospitals very frequently. With a huge number of patients awaiting examination every day, they will be given very late appointments, not before a few months.

Considering absence of appointments and other crucial aspects, there is a need to address the challenge of handling *complex* queue systems within busy public hospitals of developing countries 'as they are'.

1.5 Motivation, Overview and Research Objectives of the Current Study

The poor healthcare situation in *developing countries*, such as Pakistan, makes the life of ill quite miserable. In addition to the 'health-related' challenges, there are a number of constraints which create hurdles in the smooth operations of health delivery systems. One of the most significant factors in this regard is the absence of appointment systems, which leads to uncertainty in the arrival pattern of patients along with excessive wait time and overcrowding. *Queuing* is also an extremely important criterion of efficiency assessment which represents the operational capability of large busy public hospitals, with excessive queuing causing inconvenience for the patients. The ultimate objective of the healthcare system of any developing country is to have enough hospitals and resources to cater for the huge inflow of patients. However, besides continuously striving to achieve this goal, there is a dire need to construct a framework which could guide the management of busy public hospitals in developing nations to *minimize* the excessive waiting by patients, in the absence of prior appointments, given the current meagre conditions.

The current study aims to assess and improve the efficiency of the *outpatients'* department in a large *public* hospital of a *developing* country, specifically with respect to *queuing* situation, in the *absence* of appointment systems. In order to achieve the desired objective, the specialized efficiency technique of DEA has been employed, presenting a novel application of this modelling tool and providing evidence of its effectiveness with regard to queue management. The idea is to propose a queue assessment model using DEA, which acts as a guideline for controlling excessive queuing efficiently, within busy public hospitals of developing countries, where all outpatients are walk-in. Additionally, the present study intends to develop a dynamic framework dedicated towards the implementation of the proposed model, which continuously monitors the queue situation, and guides the administration for taking necessary measures to minimize excessive wait times of patients.

The current study has been carried out within the outpatients' department of one of the largest and busiest public hospitals of Pakistan, a *representative* of struggling health care delivery systems in developing countries, with high waiting times of patients.

The main research objectives of the current study are:

- *Research Objective 1:* To identify various factors which lead to excessive queuing in a busy Public hospital of a developing country, where all outpatients are walk-in,
- *Research Objective 2:* To develop a Queuing-DEA model and assess its usefulness for evaluating the queue system in a busy public hospital of a developing country, where 'all' outpatients are walk-in,
- *Research Objective 3:* To develop a dynamic framework for practical implementation of the proposed model for continuous monitoring of the queue system in the absence of appointments, within a large public hospital of a developing country.

Chapter Summary

The present chapter highlighted the poor health care situation in developing countries mainly in terms of the factors constraining the smooth operations of health care systems, as well as briefly elaborating upon poor health-related indicators. Factors such as scarcity and inadequate allocation of resources, lack of financing and high level of population are some of the main issues experienced by developing countries. Additionally, the health profile and other related factors of Pakistan, a representative of developing countries, were highlighted in comparison to other developed and developing countries, including high level of population, low literacy, low health care expenditures and shortage of resources. Therefore, this evaluation reflected the poor health care situation and low level of patient care in a developing country, such as Pakistan. A number of aspects were identified which play an important role in demonstrating the efficiency of health care delivery systems, including high costs and limited budgets, and dearth and poor distribution of limited resources. Queuing is a significant measure which indicates the efficiency of operations in health organisations as well as the extent to which patients are facilitated. The public hospitals in developing countries are extremely overcrowded and the waiting times of patients are excessive. Absence of appointment systems greatly enhances the queuing problem in developing countries as it prevents the management to organize and plan pre-emptively. Additionally, non-existent GP referral systems, increased number of stages and cumbersome procedures and paperwork further burden the management. Therefore, the key focus of the current research study is to develop a model which facilitates busy public hospitals of developing countries, using Pakistan as an exemplar, to evaluate and improve the queuing system keeping in mind the present idiosyncrasies.

CHAPTER 2 LITERATURE REVIEW: QUEUE MANAGEMENT IN HEALTHCARE

Chapter Overview

The current chapter provides a review of literature on healthcare studies, with respect to queue management. It provides a detailed evaluation of various issues and queuing variables considered by various healthcare studies. Previous studies conducted in different departments of a hospital have been recognized. A contrast of studies in developed and developing countries has been carried out, elaborating on different issues highlighted in these studies in terms of queuing and scheduling. Moreover, a number of commonly used queuing parameters have been evaluated with emphasis of their utilization depending on the objective of the study. Additionally, a few operational techniques have been appraised which have been frequently employed to assess the queuing problem. Furthermore, a few major limitations of these techniques have been highlighted, which might hinder an effective queuing analysis. It has been stressed that the merits of an efficiency assessment technique such as Data Envelopment Analysis (DEA), make it more suitable for providing a comprehensive analysis of the queuing problem; especially in public hospitals of developing countries, where the queue system is highly disorganized.

All organizations which deliver services strive to get maximum output from minimum utilization of resources. They conceptualize and implement systems which are the most efficient (Fomundam and Hermann, 2007). This applies to the health care industry as well which is considered as one of the largest industries in the service sector. Health care administration is faced with the challenge of managing the trade-off between enhancing productivity while maintaining high quality of services (Wijewickrama and Takakuwa 2008).

Among other factors, queuing is considered to have a negative influence on consumer service perceptions (Bielen and Demoulin, 2007). Similarly, in case of healthcare, waiting time has become an important criterion to determine the service quality (Wijewickrama and Takakuwa 2008), efficiency of services (Jehu-Appiah *et al.,* 2014) and patient satisfaction (Patwardhan *et al.,* 2013). In developing countries, the queuing problem is more crucial within large public hospitals due to absence of appointment systems, and other health-care and operational challenges (as discussed in Section 1.2 and 1.5 above). The current study aims to analyse the queue system in a busy public hospital of a developing country, where all patients are walk-in with excessive wait times.

2.1 Patient Flow/Queuing Studies in Different Areas of the Hospitals

The patient flow system within a health centre is an extremely significant indicator of the effectiveness and efficiency of service delivery. A health institution comprises of many different sections mainly including inpatients, outpatients, emergency, diagnostics and pharmacy. There are numerous studies which emphasize on exploring issues related to the patient flow system in order to provide an optimal solution. Certainly, the sophisticated models utilized and various parameters incorporated depend on the main aim of the study and in some cases, on the particular section of the hospital where the study is being carried out. Some of the previous works specifically concentrate on queue management, while others had slightly different objectives but still required an evaluation of existing inflow of patients.

A number of studies have explored issues within an *inpatients'* department, which required an assessment of the patient flow. Proudlove *et al.,* (2007) provided an overview of the use of Operational Research (OR) techniques to assess and improve the inpatient flow. Many studies related to the inpatients' department are associated with *bed capacity* and utilization. Bruin *et al.,* (2010) and Harper and Shahani (2002) investigated the relationship of occupancy rate of beds with refused admission in an inpatient setting. Additionally, Bruin *et al.,* (2010) incorporated another factor, size of the hospital, to assess this relationship while Harper and Shahani (2002) argued that hourly/daily/monthly variations in length of stay are significant for capacity planning. Li *et al.,* (2009) had the objective of developing a decision support system which aids the management in the allocation of beds. Gorunescu *et al.,* (2002) developed a methodology which allowed for the estimation of bed occupancy and probability of lost demand, as well as allowing for optimal allocation of beds given the level of patient rejection. Seung-Chul *et al.,* (2000) explored various bed reservation strategies in order to minimize the number of cancelled surgeries by reserving beds for the use of elective-surgery patients. Kao and Tung (1981) emphasized on the use of forecast data to assess fluctuations in demand and patient load rather than historical data, for appropriate bed allocation in order to reduce overcrowding. In a study conducted in an Intensive Care Unit (ICU), Ridge *et al.,* (1998) explored the trade-off between the average occupancy level of beds and the transfer of patients due to lack of bed availability.

Apart from bed capacity planning, *nurse* staffing is another factor which has been studied in an inpatient setting. Griffiths *et al.,* (2005) developed a model to determine the required number of nurses per shift in an ICU such that the nursing staff costs are reduced. Biju and Naeema (2011) examined the issue of staff inadequacy in an inpatient setting and its relation to the queue sizes and waiting times. Yankovic and Green (2011) mainly investigated the interdependency between inpatient bed occupancy level and demand for nursing levels in order to assist hospital managers in making appropriate nurse staffing decisions which avoids delays in nursing care as well as beds._Kortbeek *et al.,* (2015) developed a method which examines the interaction between inpatient nurse staffing levels mainly using bed census predictions, along with other planning issues such as case mix, care size units, nurse-to-patient ratios and surgical block planning. George *et al.,* (1983) explored the waiting lists in a surgical department in terms of better allocation of resources including theatres, beds and availability, and time of surgeons. Goddard and Tavakoli (2008) evaluated improved prioritization techniques of inpatients on waiting lists for better management of these lists. Silvester *et al.,* (2004) emphasized on the issue of mismatch between demand and available capacity as one of the main factors for enhanced waiting lists, and suggested ways that require an in-depth understanding of the patient flow process and bottlenecks.

Several studies have also been carried out in *Accident and Emergency* (A&E) departments of the hospitals in relation to patient flow. Lane *et al.,* (2000) discusses the relationship of delays in A&E department with other factors such as demand patterns, resource availability, number of beds and other hospital processes. Brailsford *et al.,* (2004) examined different scenarios in order to assess the patient flow and identify bottlenecks in the system in an A&E department. McGuire (1994) conducted a study specifically related to the length of stay in an emergency department by testing different alternatives and selecting an optimal solution. Madelbaum *et al.,* (2012) explored the interface between emergency department and the internal wards. Mayhew and Smith (2008) and Laskowski *et al.,* (2009) evaluated the patient flow system within an A&E department. The former study emphasized on the completion times given the government target of 4 hours, while the latter investigated patient wait times and access. Ahmed and Alkhamis (2009) developed a decision support system which assisted in determining an appropriate number of medical personnel in an emergency department given budgetary constraints which maximizes patient throughput and reduces the wait time of patients. In another study, Song *et al.,* (2014) suggested that a 'dedicated' queuing system leads to a shorter length of stay for the patients as compared to if they were allocated under a pooled queuing system.

Researchers have identified that an ineffective and inefficient appointment system is one of the main issues for the long waiting time of patients in an *outpatients'* department. Therefore, various studies have been conducted which aim at improving the appointment scheduling in an outpatients' department. Cayirli and Veral (2003), and Lakshmi and Sivakumar (2013) provide a comprehensive survey on appointment scheduling problems in an outpatient setting.

The study by Bailey (1952) was one of the first studies to explore an appointment system for outpatients. In this particular study, an improved appointment system was recommended, by examining the average wait times of patients and doctors' idle time. Since then, numerous studies have been carried out which evaluated queuing and appointment scheduling problems in outpatients' departments considering various aspects. Fetter and Thompson (1966) and Klassen and Rohleder (1996) investigated the relationship between patient wait time and doctor's idle time in their studies as well. The former authors conducted a number of experiments to show the effect of different factors on the relationship between patients' waiting time and doctors' idle time. These factors included patient load, doctors' prompt arrival, appointment intervals and patient and doctor tardiness. The latter focused on comparing different appointment scheduling rules in order to propose an optimal rule which could minimize waiting time of patients and idle time of doctors. A few studies mainly explored the consultation time in relation to the wait times for assessing the queuing situation for outpatients. Aharonson-Daniel *et al.,* (1996) explored the excessive waiting by patients mainly in relation to the low consultation time in an outpatient setting. Different scenarios were tested and an appropriate appointment system was proposed to reduce the waiting times. Huarng and Lee (1996) conducted a study in the dermatology outpatients' department and assessed patients' queue as well as service utilization of staff members. They discussed recommendations for improvement not only in terms of consultation time but also by increasing physicians and number of sessions. However, Cayirli *et al.,* (2008) considered consultation time for new and returning patients as well as wait time and doctor's idle time and overtime in order to evaluate a number of appointment schedules for an outpatients' department.

Some studies particularly focused on identifying *multiple elements* which led to long queues and testing different appointment systems to identify an optimal system. Harper and Gamlin (2003) identified crucial factors which led to build-up of queues by evaluating the appointment schedules in an Ear, Nose and Throat (ENT) outpatients' department in a major hospital. Alternative appointment schedules were proposed and tested; and results were based on their effects on performance measures such as average wait time until the first service, percentage of patients who waited for more than thirty minutes for their first service and average time spent in the clinic. Wijewickrama and Takakuwa (2008) developed a number of appointment systems by incorporating different combinations of various appointment rules and patient sequences. These appointment systems were tested for punctual patients as well as for no-shows. Klassen and Yoogalingham (2009) developed a model of determining optimal rules for an outpatient appointment system. They suggested that this approach is more flexible and robust as it has the ability to incorporate multiple variables and performance measures, hence enhancing chances for practical implementation. Zhu *et al.,* (2012) evaluated various appointment systems to identify the factors causing long patients waiting time and clinic overtime in an outpatient setting. These factors included uneven distribution of appointment slots, late start of sessions, unused sessions when no patients are seen and irregular calling sequence of appointment patients for consultation. Bard *et al.,* (2014) proposed changes to the existing appointment schedule by investigating the effect of overtime, no-show rate and overbooking on congestion and increased length of stay.

In some studies related to outpatients, *additional issues* were highlighted. Lehaney *et al.,* (1999) emphasized that appointments in outpatients' department should be scheduled such that the unexpected non-attendance rate of patients is minimized. O'Keefe (1998) investigated the requirement of qualitative approaches to assess the operations of outpatient departments, and emphasized on the importance of implementable policies. Matta and Pattersen (2007) carried out a study in an oncology outpatients' department and emphasized on the need of stratification of certain performance measures such as days of the week, patient classes and multiple facility visits, in order to get a detailed overview of the patient flow. Yeboah and Thomas (2010) conducted a study in a Cancer outpatients' department and also proposed that the stratification of outpatients, that is, having separate clinics for different groups will reduce the waiting time of patients and improve the patient flow. Berg *et al.,* (2014) explored the issue of uncertainty in the completion times of procedures and patient attendance in an outpatient setting. They evaluated different appointment schedules to determine an optimal schedule mainly in terms of patient overbooking, no-show rates and sequencing of no-show patients, considering this uncertainty.

There are only a limited number of studies which particularly considered *Walk-in* patients in *outpatients'* departments. However, the Walk-in patients were analysed in the light of the appointment schedule system. Fetter and Thompson (1966) considered Walk-in patients when evaluating different appointment schedules, to identify the optimal schedule which could fit in walk-in patients. Rising *et al.,* (1973) investigated Walk-in patients in an outpatient setting, with respect to the appointment scheduling as well. They proposed a model where an estimate of walk-in and pre-booked patients was calculated; with the objective to decide for the number of appointments that should be given by day of the week and during one day. Huarng (2003) evaluated different appointment scheduling rules by considering the arrival rate of Walk-in patients in an outpatients' department, using two criteria. They were ratio of wait time to physician idle time per session and the closing time of the session. Wijewickrama and Takakuwa (2005) tested different appointment schedules by allotting different slots for walk-in patients, such as to reduce the average wait time of appointed patients and congestion. Cayirli *et al.,* (2008) considered the percentage of walk-in patients in addition to no-shows and assessed their effect on the wait time of appointment patients, and idle and overtime of doctors combined; in order to determine appropriate scheduling rules. It was suggested that overbooking and/or keeping some appointment slots empty can be used to cater for this type of variability. Isken *et al.,* (1999) proposed a general framework which could be used by modellers to prepare optimal appointment schedules and evaluate patient flow patterns, among others, in an outpatients' department. They concluded that the inclusion of walk-in patients along with no-shows and late arrivals should be considered when developing a model.

Ashton *et al.,* (2005) conducted a study in a Walk-in centre to assess the arrival pattern and flow of patients. The variation in the arrival pattern was considered as a significant factor for obstructing the smooth flow of patients. One of the suggestions proposed was to start an appointment system for patients with specific complaints, while maintaining a drop-in session for first-time patients. Cayirli and Gunes (2014) considered the effect of seasonal Walk-in patients on outpatient appointment scheduling. They investigated the impact of daily, weekly and monthly variation in walk-in patients on the adjustments required in the number and specific timings of appointments to be allocated daily. Recently, Mensah *et al.,* (2015) evaluated and compared the patient flow system in the outpatients' department of a public and private hospital considering walk-in patients only, disregarding appointment or emergency patients. They concluded that poor service provision is the main factor that leads to increased wait times, and this problem is more severe in public hospitals than private hospitals. They emphasized on improved supervision of staff and the services and the use of electronic based systems, to reduce the wait time of patients. However, in most public hospitals of developing countries like Pakistan, there are over-whelming and highly disorganized queues with lack of proper and reliable electronic systems. Hence, there is a need to develop a framework which continuously monitors the overcrowded and highly variable patient flow system as it is.

In addition to the main departments within a hospital, studies have been carried in *other areas* as well, with the objective of improving patient flow and resource utilization. Dansky and Miles (1997) and Santibanez *et al.,* (2009) conducted studies in an Ambulatory Care Unit (ACU). The former study explored the relationship between the wait time of patients and patient satisfaction with the ACU service, and highlighted the importance of handling and managing waiting times effectively. The latter study tested different scenarios including a number of factors regarding patient scheduling, resource allocation and operations, and highlighted that simultaneous implementation of changes regarding these aspects will lead to reduced wait times. Conducting a study in a transfusion centre, Angelis *et al.,* (2003) evaluated allocation of resources such as doctors, nurses and beds in a multiple service environment, where each patient follows a different path depending on the type of service(s) they require. Cote (1999) carried out a study in a local family clinic, and explored the relationship between the examining room capacity and the patient flow using a few performance measures including room utilization, likelihood of room(s) operating at full capacity, queue length and patient flow time. In a study carried out in a Paediatric department, Benneyan (1997) investigated the trade-off between the wait time of patients and distribution of resources such as examination rooms, availability of doctors and availability of support staff (receptionists, medical assistants and nurses) by evaluating different combinations scenarios that could reduce the wait time. Martin and Haugene (2003) conducted a study in a Geriatric department and explored different ways of improving patient flow and bed allocation which included better utilization of idle bed capacity in other wards and improving the routine procedures to reduce length of stay. Patwardhan *et al.,* (2013) compared convenient care centres and family practice physician offices with respect to waiting and consultation times, and identified that the former has relatively shorter wait times but longer consultation times.

In the field of healthcare, different studies have evaluated various factors when assessing the patient flow system, mainly resource allocation, in different departments of the hospital including inpatients, outpatients, emergency, or in other health institutions. Specifically, almost all studies in an outpatient setting focus on an already existing appointment system. Furthermore, there are negligible studies which explore the Walk-in patients specifically within an outpatients' department, and even these studies have been carried out in the developed countries. In previous works, improvements in the appointment scheduling were suggested, and it is believed that this will improve the flow of walk-in patients as well. None of the studies proposed a set framework which assesses a patient flow system where all patients are walk-in along given variation in the arrival pattern, high waiting times, and appointment systems are not implementable; which is the case in developing countries. There are a number of factors which make it highly unlikely to implement appointment systems in these countries; such as low education levels, poor understanding of standard procedures by patients, lack of effective information dissemination techniques, delays and variation in public transport, and lack of General Practitioner (GP) referral systems. Therefore, the present study aims to develop a queue assessment framework which evaluates the *outpatients'* inflow in a busy public hospital of a developing country, where there are *no* pre-booked appointments; and provides a set guideline to reduce long queues quickly (see Figure 2-1 below).

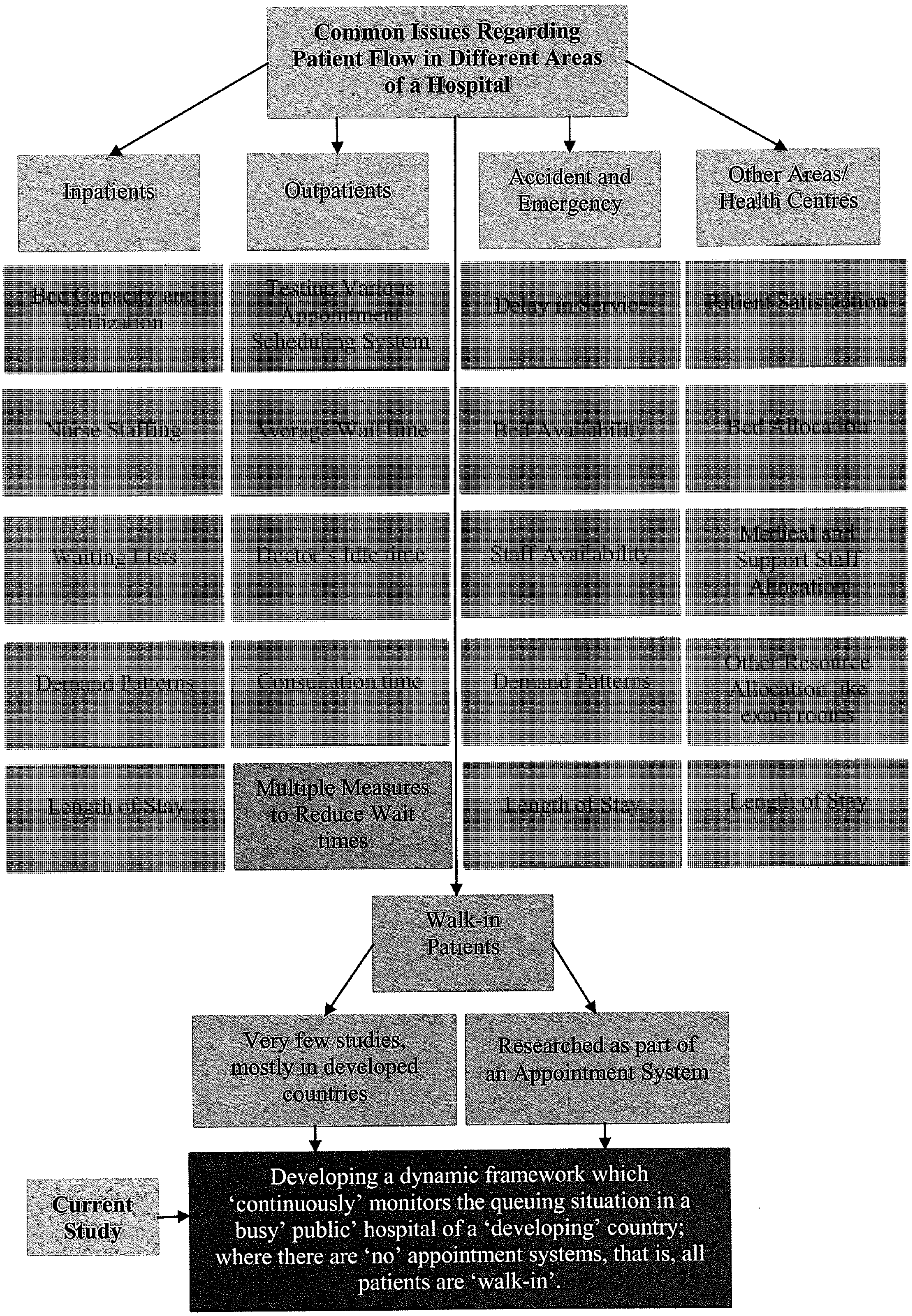


Figure 2-1: Summary of Most Common Issues Studied in Different Sections of a Hospital and Objective of the Current Study

2.2 Patient Flow/Queuing Studies in Developed Versus Developing Countries

Almost all health systems, developed and developing countries, face multiple challenges such as cost pressures, increased patient expectations for quality service, increased concerns for patient safety and limited resources (Harper and Shahani 2002; McDermott and Stock 2007). Waiting time is another crucial issue and different countries employ various methods to address this significant problem (Kozlowski and Worthington 2015). The developed countries have a dedicated appointment system in health institutions and almost all studies have focused on exploring factors affecting it. However, the health systems in developing countries face the challenge of assessing and improving the patient flow in the absence of appointment systems.

Considering *developed countries*, most of the previous works have been carried out in the USA and UK emphasizing upon different issues regarding patient flow. The *UK* National Health Services (NHS) has been investigated in a number of different studies in the past 50 years. The study conducted by Bailey (1952) was one of the first few which targeted the outpatients' department in the UK. This study highlighted the lack of guidance to monitor the patient flow in the outpatients' department in the UK. Therefore, special attention was given to exploring the appointment systems in terms of the patient wait time and consultation time. The study by Brahimi and Worthington (1991) investigated only those outpatients' clinics in the UK, which were responsible for checking the recovery of patients who have been to A&E department recently. In this study, the appointment system as well as number of repeat visits by patients was explored. Goddard and Tavakoli (2008) highlighted the increase in demand for NHS in the past few years, and identified the waiting lists for inpatients as a significant performance indicator of NHS. Harper and Shahani (2002) focused on bed capacity and utilization as significant performance measures for hospitals under NHS Trust in UK. Lane *et al.,* (2000) highlighted the subject of performance of A&E departments in UK hospitals and provided an account for reasons behind excessive delays in A&E for patients requiring admission. These factors included bed shortages, increased number of patients in A&E due to difficulties in gaining access through GP referral systems, discharge procedures of existing patients and the policy of maximum waiting which prevents admission of more severe cases. Lehaney *et al.,* (1999) elaborated that the waiting time of outpatients is one of the key performance indicators considered by NHS Performance Guide, and evaluated the long waiting time for patients once they have arrived at their appointment time in an outpatients' department. Proudlove *et al.,* (2007) provide a detailed overview of the various pathways and procedures followed by inpatients in NHS, UK along with various OR modelling techniques employed. They also elaborated on the government initiatives for improving the patient inflow, in terms of wait time of patients, bed capacity and in-depth understanding of the patient flow process. Mayhew and Smith (2008) aimed at assessing the A&E departments considering the completion time target of 98% of patients to be discharged within 4 months, as provided by the UK government. Allder *et al.,* (2010) emphasized on the importance of evaluating the functioning of the whole patient flow system for better long-term planning of bed capacity in UK NHS. They explored different factors leading to increased variation in bed availability and demand, and increased length of stay. Demir *et al.,* (2014) explored the increasing demand of neonatal services in the UK. They evaluated the performance of the neonatal system by assessing the length of stay at different pathways followed by the patients.

Similarly, a number of patient flow studies have been carried out in the *USA*. Cote (1999) observed the pathways followed by patients in a local family health clinic in USA, including nurse aid station, examination room, check-out and cashier. This patient-level data was used to assess mainly the examination room capacity in relation to some other performance measures. McDermott and Stock (2007) elaborated on the challenges faced by hospital management in USA, and evaluated the average length of stay in New York state hospitals in USA. Wright *et al.,* (2006) provided an outline of the nurse staffing issues reviewed by the government bodies and lawmakers in USA such as nurse-to-patient ratios, nurse staffing ratios and shortages, and overtime and excessive workload. The study is focused on developing a nurse scheduling system in a large acute hospital in USA, in the light of these issues considered by recent legislations in USA. Yankovic and Green (2011) also conducted a study regarding nurse staffing levels in one of the orthopaedic hospitals in USA. One of the objectives of this study was to identify the impact of certain factors on the nursing levels such as inpatient delay, unit size, nurse intensity and length of stay. The results were used to evaluate the trade-off between the nurse-to-patient ratio provided by the legislation of California and improvement in performance in terms of patient care. Kang *et al.,* (2014) emphasized on streamlining the admission process policy in emergency departments (ED) in USA. They evaluated the effect of various admission processes on the patient flow in ED in terms of reducing delays, overcrowding, length of stay and wait times of patients.

In addition to USA and UK, patient flow studies have been conducted in some *other developed countries* as well. Martin and Haugene (2003) highlighted the challenges faced by the health system in Norway, including increased life expectancy and high costs of medical equipment and medication. Some other governmental challenges were also mentioned such as decrease in taxes, flexibility for patients to choose any hospital and increasing demand of improved patient throughput while maintaining high level of patient care. This study aimed at exploring different ways to improve patient throughput, and reducing length of stay and wait times for patients in one of the largest hospitals of Norway. Zhang *et al.,* (2012) stressed upon the importance of long-term care for aging population in Canada. They focused on developing an approach which helped in evaluating the minimum required capacity planning for long-term care facilities and policies. In a study carried out in Japan, Wijewickarama and Takakuwa (2005) highlighted the increasing importance of reduced wait times. Multiple appointment systems were tested to identify optimal schedules which could lead to a huge reduction in the wait times of patients. Zhu *et al.,* (2012) conducted a study in one of the government hospitals of Singapore, and suggested recommendations for improvement in the existing appointment schedules to reduce patient wait time and clinic overtime without incurring extra workload. Kozlowski and Worthington (2015) explore the policy of maximum wait time guarantee within the Denmark health system. They evaluated the trade-off between reduced wait time and increased under-utilization of resources, particularly for public hospitals, in the light of this policy.

The healthcare issues are much more *severe* in resource-poor countries such as Pakistan. People in these countries lack even the most basic healthcare provision (Dhungel 2014) with long queues, and the management struggles with achieving a balance between providing quality healthcare and optimal utilization of resources, hence emphasizing on the crucial need of constant evaluation of patient flow system in developing nations.

In addition to developed countries, a few studies have been carried out in the *developing countries* as well. Babes and Sarma (1991) carried out a study in an outpatients' department of a public health centre in Algeria. The authors highlighted that the existing appointment systems lead to excessive wait times, and evaluated several factors which rendered these systems inefficient. These factors mainly included variation in the arrival pattern of patients, criteria for giving appointments and late start of consultation.

Aharonson-Daniel *et al.,* (1996) stressed on the issue of increasing demand with long wait times and dissatisfaction among patients in Hong Kong. Recommendations were provided by testing several scenarios including variations in the consultation time and the appointment systems to provide an optimal solution which can reduce the wait times. However, only a few health institutions in developing countries have an existing appointment system. Most of the large busy public hospitals have all patients as walk-in. In a study within the emergency department at a government hospital in Kuwait, Ahmed and Alkhamis (2009) evaluates the patient flow process and identifies different alternatives that lead to an optimal distribution of staff at various service points. Li *et al.,* (2009) highlighted the challenges faced by the public hospitals in China due to increasing development of private health institutions. They evaluated the bed allocation problem in one of the public hospitals in China, to allow for increased profits and chances of patient admission, to avoid patients turning to private hospitals. These factors included wait time being higher than average wait time, probability of all servers being busy and queue length. Huarng and Lee (1996) highlighted the healthcare issues faced by Taiwan including insurance policy for everyone, evaluation system for each hospital, increasing competition and increasing demand for quality services. They conducted a study in the small local hospital in Taiwan, in order to assess the waiting time and service utilization of staff and suggested recommendations for improvement, mainly in terms of consultation time.

Mital (2010) evaluated the number of doctors in an outpatients' department and number of beds in an inpatients' department within a medium-sized hospital in India using various queuing characteristics. Oche and Adamu (2014) investigated the general outpatients' department in a tertiary health institution in Nigeria and explored a few factors leading to excessive wait times of patients. They suggested that there should be an increase in the number of doctors and increased training for health workers in various analytical software packages to encourage implementation of appointment and triage systems to reduce the wait times. Although these suggestions are reasonable, but in almost all public hospitals of developing countries which are highly unorganised, provide multiple (primary, secondary and tertiary care) services, and all patients are walk-in; these recommendations are not currently implementable. In the hope of achieving an ideal streamlined patient flow system, no efforts are made to deal with the existing queue system. There is a crucial need to construct a framework which monitors the patient pathway as it is without suggesting any major changes which are not practical in the short-run.

Only a small number of studies conducted in the *developed countries* have considered the inflow of *walk-in* patients in an *outpatient* setting, and as part of a broad objective of assessing an appointment system. For instance, Rising *et al.,* (1973) conducted a study in health service in USA, and evaluated waiting times and doctor scheduling for walk-in patients in addition to appointment and second service patients. In a study in Japan, Wijewickrama and Takakuwa (2005) analysed different appointment schedules by reserving a certain slot for walk-in patients such that the congestion for appointed patients is minimized. Ashton *et al.,* (2005) assessed a NHS walk-in centre in the UK, and discussed recommendations for improving the flow of walk-in patients. These suggestions included triage of patients, creating an appointment schedule for all patients, promoting less busy times and providing information about services not provided. These suggestions might be plausible in developed countries. However, in developing nations, these recommendations are almost impossible to implement due to various demographic, social and cultural barriers (as discussed in Section 1.4); and requires some continuous intervention.

Considering *developing countries*, there have been negligible studies which have explicitly targeted *outpatients'* department of large public hospitals in developing countries where all patients are *walk-in*. Huarg and Lee (1996), Aharonson-Daniel *et al.,* (1996) and Mensah *et al.,* (2015) briefly discussed the option of implementing complete appointment systems in order to reduce the wait time of patients, in their studies in Taiwan, Hong Kong and Ghana respectively. In a recent study, Daultani *et al.,* (2016) evaluated the patient flow system within the outpatients' department of an ophthalmic hospital in India, considering number of patients served per day, service lead time, resource utilization along with the potential for cost savings.

To the best of our knowledge, apart from these studies, any other study cannot be found which explicitly evaluated a public health centre in a developing country where all patients are walk-in and appointment system is not a favourable option presently. Besides, due to increased variation in the arrival pattern of walk-in outpatients, the management cannot predict accurately the number of patients arriving in a day/week or maximum load; resulting in excessive waiting times for patients. This necessitates the formulation of a flexible framework which continuously evaluates the current walk-in patient flow system (see Figure 2-2 below).

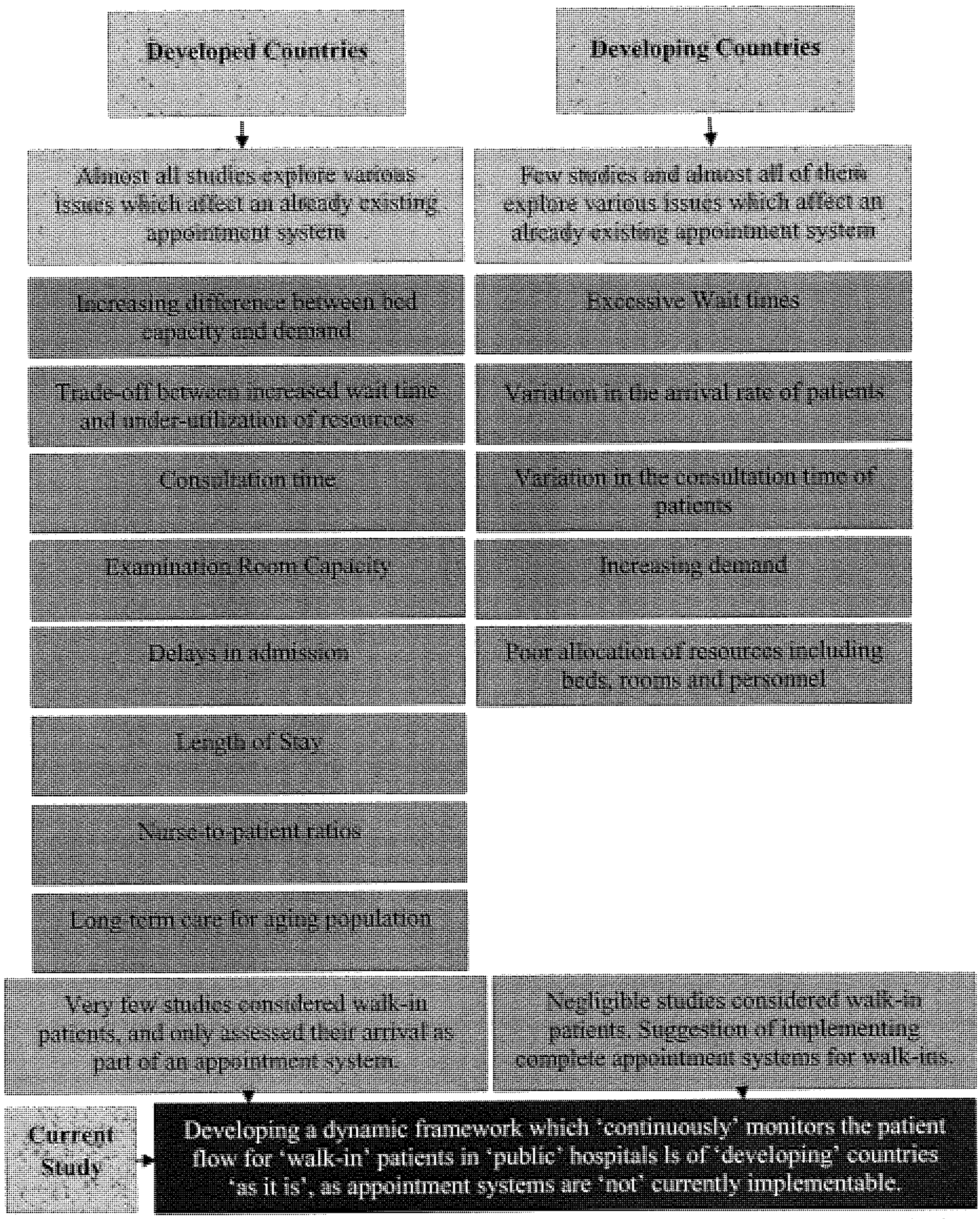


Figure 2-2: Summary of Most Common Issues Considered in Developed and Developing Countries and Objective of the Current Study

2.3 Parameters Used in Queue Management/Appointment Scheduling Studies

Different researchers have utilized a variety of *parameters* to assess the queue system depending on the nature of the research problem being investigated. There are some queuing factors which are used more often than the others.

The *waiting time* of patients is the most commonly used factor to assess a queue system. However, the waiting time has been defined and measured differently depending on the problem under consideration. Considering outpatients' departments, waiting time has been evaluated with respect to the existing appointment systems. In some studies, the waiting time is calculated as the difference between the consultation time and the time a patient arrived at the hospital (Adeleke *et al.,* 2009; Baril *et al.,* 2016). Additionally, Saremi *et al.,* (2015) also considered the costs associated with this wait time. In order to propose appropriate appointment schedules, they considered the effect of multiple sequential services required by different patient types on the costs of waiting time, given the availability of servers and service time at each of these stages. Some studies included the wait time as the difference between the time consultation started and the appointment time given to the patient (Babes and Sarma 1991; Harper and Gamlin 2003; Cayirli *et al.,* 2008). However, Fetter and Thompson (1966) considered waiting from arrival to appointment time and then from appointment time to first consultation.

In some cases, extensions to these wait times were included. O'Keefe (1998) and Zhu *et al.,* (2012) incorporated waiting time for both scenarios, early and late arrival for an appointment. Cayirli *et al.,* (2008) considered uniformity of waiting times across different appointment slots and also the number of patients seen within a pre-specified threshold (within 30 minutes of appointment time). Goddard and Tavakoli (2008) investigate the waiting time of patients on the waiting list for admission to the hospital considering three different scenarios; including patient's willingness to wait, access to list restricted for seriously ill patients and using clinical priority to determine how quickly a patient can be moved from the list to admission. In some previous works, the wait time for emergency patients has been considered. For instance, Lane *et al.,* (2000) considered the waiting time as the time a decision is made to admit a patient till the time the patient is actually admitted to the ward. They evaluated the bottlenecks in the admission process which was suggested as the main cause of delay, and hence increased wait times. While conducting a study on emergency Medical Services, Aboueljinane *et al.,* (2013) measured the patient wait time in terms of the 'response' time, that is, the time required for the rescue team to reach the patient location after an emergency call is received. They suggested that response time is a crucial factor especially if the level of emergency is high such as life-threatening emergencies. In almost all studies regarding outpatients, the wait time has been defined in relation to the appointment systems. However, nearly all public hospitals of developing countries have walk-in patients; therefore, the wait time will be defined with regard to the arrival behaviour of and services received by walk-in outpatients.

Another significant aspect of the queuing analysis is the *service time*. In most studies, consultation time with the doctor is considered as a measure of service time, which is the difference between the time it started and when it ended (Hill-Smith 1989; Brahimi and Worthington 1991; Harper and Gamlin 2003; Adeleke *et al.,* 2009). Liu and Liu (1998) set a specified consultation time (10 minutes) in order to assess its influence on the different appointment schedules. Zhu *et al.,* (2012) considered the unused session time and session utilization rate in addition to the actual consultation time, to improve the wait time and clinic overtime. Cote (1999) considered the waiting and service times for all patients who required initial or second consultations. Cayirli *et al.,* (2008) specifically compared the mean consultation time for new and return patients. Some studies have evaluated the service time at different stages. Giachetti *et al.,* (2005) consider waiting and service times for preliminary examination, second time examination and discharge, whereas Aharonson-Daniel *et al.,* (1996) consider waiting and service times for consultation with the doctor and pharmacy. Mankowska *et al.,* (2014) evaluated the issue of multiple services required by the patient and the level of tardiness, by assessing if these services are independent or inter-dependent, and can be carried out simultaneously or consecutively. Mandelbaum *et al.,* (2012) considered the start of patient 'service' in a ward before a patient arrives, including staff preparation, to improve the process of assignment of patients to different wards. Some studies have considered the service time as the total time in the system, that is, the difference between the times a patient enters the patient flow till exit (Lehaney *et al.,* 2009; Laskowski *et al.,* 2009). Huarng and Lee (1996) concluded that nearly all of the time is spent waiting as the consultation time is quite low. There is a need to assess the degree to which consultation/service time impacts the queue system in a developing country hospital.

The availability and utilization of *resources* has also been considered as a significant variable to assess a queue. Among different resources, bed capacity planning, and number of physicians and nurses are the most widely discussed issues. Some studies specifically evaluated the trade-off between *bed* capacity and demand. Lane *et al.,* (2000) investigated the impact of varying number of beds and increase in demand for beds on various performance measures including (average) wait time to consultation, time to admission, elective cancellations, daily hospital occupancy and daily utilization of doctors. Gorunescu *et al.,* (2002) explored the relationship of bed occupancy with the demand lost due to all beds being occupied, and assessed the trade-off between cost of empty beds and cost of delayed patients. Cochran and Roche (2008) investigated the issue of the level of bed demand, capacity and utilization for inpatients in multiple levels of care (Intensive Care Unit, intermediate, medical/surgical and obstetrics). Some other studies assessed the impact of factors related to bed capacity and planning, on a few performance measures. Harper and Shahani (2002) evaluated optimal number of beds by considering variation in different factors such as care units (ward, speciality or hospital as a whole), bed requirements (seasonal requirements/beds allotted on a part-time basis), demand patterns of patients (hourly/daily/monthly variations) and length of stay (using patient categorisation techniques, admission rules and deferral times). Li *et al.,* (2009) evaluated optimal profit in addition to the optimal number of beds required, and suggested that optimal profit can be achieved by balancing lost patients due to unavailability of beds with the number of idle beds. Mandelbaum *et al.,* (2012) considered the bed turnover rates in addition to idle bed availability and occupancy levels in order to evaluate an optimal level of bed capacity. Some studies have specifically explored the issue of transferred patients or refused admissions due to lack of bed capacity. Ridge *et al.,* (1998) investigated the trade-off between the bed occupancy level and bed numbers, and number of transfers due to lack of bed availability, for emergency and elective patients. Bruin *et al.,* (2010) evaluated the relationship between the size of a unit (wards/departments), the target occupancy rate for beds and refused admission in order to suggest an optimal allocation of beds. They mainly suggested merging small or less busy departments/wards to improve bed allocation.

Some studies have emphasized on the availability, demand and scheduling of *nurses* as part of resource utilization, and considered related variables. Griffiths *et al.,* (2005) investigated the scheduling and rostering of nurses by considering costs. They assessed the change in costs per shift by using varying number of nurses in each shift and the corresponding effect on the number of idle nurses in each shift. With the objective of evaluating nurse scheduling, Wright *et al.,* (2006) considered regular and overtime wages as well as variables related workload such as maximum and minimum number of shifts each week, weekends and overtime shifts and number of patients handled per nurse. Mandelbaum *et al.,* (2012) also investigated the issue of workload for nurses but they divided the load into two parts, load associated with treating hospitalized patients and load due to patient admissions/ discharges. Yankovic and Green (2011) identified the interdependency between bed occupancy level and demand for nursing by considering variables related to both resources. These parameters included probability of delay for a bed and for nursing care, average delay for a bed, average nurse utilization, and average number of patients waiting for a bed and nurse.

The availability and scheduling of *physicians* has been specifically considered in some studies in terms of resource utilization. Different variables have been considered in order to assess the availability of doctors. Huarng and Lee (1996) investigated the scheduling of doctors by incorporating factors such as average utilization rate, average number of patients seen, and maximum idle and busy time for doctors. Liu and Liu (1998) suggested an optimal scheduling system by considering the delay in the arrival of doctors. Klassen and Yoogalingham (2009) included cost parameters related to number of doctors such as that for physician overtime and idle time. A few studies specifically targeted physician rostering and related issues. Rising *et al.,* (1973) evaluated the physician scheduling in relation to allocating appointment slots to patients considering different types of patients including emergency, walk-in and second service patients. Lane *et al.,* (2000) explored the issue of variation in rostering of A&E doctors and its effect on the increase in backlog of patients waiting for consultation due to excessive load. Bruni and Detti (2014) specifically emphasized on workload balancing when evaluating the physician scheduling problem. Some studies have included variables which consider *other medical staff members* as well. Laskowski *et al.,* (2009) considered diagnostic staff available in addition to physicians and nurses. Mankowska *et al.,* (2014) proposed a heuristic to provide an optimal solution for an assignment problem in a home healthcare problem. The model includes factors such as distance to patients' home, time preference by the patients, and requirement of multiple services and if they need to be provided simultaneously.

In developing countries such as Pakistan, the situation is even worse since there are no appointment systems in large public hospitals, with shortage of doctors and other medical/non-medical staff. Hence, in addition to A&E, outpatients' departments as well as laboratory and pharmacy suffer from extreme overloading of patients leading to massive delays in treatment, and excessive wait times; which is addressed in the current study.

The *idle time* of doctors has been considered in some studies to propose better appointment schedules for patients. In this regard, Fetter and Thompson (1966) investigated the relationship between the wait time of patients and the idle time of doctors using various performance measures mainly patient and doctor arrival pattern (lateness or prompt), patient overload and different appointment intervals. Additionally, Klassen and Rohlder (1996) also explored the trade-off between the idle time of doctors and the percentage of urgent patients served. Cayirli *et al.,* (2008) assessed the combined effect of doctor's idle time and overtime on different factors such as patient wait time, type of patients (new/return) and no-shows. However, recently, Anderson *et al.,* (2015) evaluated the effect of overlapping appointment scheduling on the costs associated with doctor's idle time along with overtime and wait time of outpatients, not the other way round as in previous studies. However, in large public hospitals of developing countries, idle time of doctors is not a significant issue due to excessive overcrowding of patients. The doctors are occupied almost the entire time due to the huge number of patients awaiting consultation.

The *length of queue* has been specifically included as a significant queueing variable in some studies. Huarng and Lee (1996) investigated the average as well as maximum length of queue in an outpatients' department, and assessed their relationship with other queuing factors such as average wait time, average number of patients served and maximum idle and busy time for doctors. Cote (1999) considered the queue length for one particular examination room and evaluated its relationship with the arrival rate of patients. They concluded that although queue length increases as arrival rate escalates, but it is not overwhelming. Silvester *et al.,* (2004) emphasized on the building up of queues during different activities in a multi-stage facility. They concluded that some of these activities run smoothly but still show an 'artificial' increase in requirement of resource capacity, hence leading to wastage. They emphasized on an in-depth evaluation of the whole process in order to differentiate between bottlenecked and streamlined activities, so that resources can be allocated to activities which actually require them and avoid wastage. Au *et al.,* (2009) included length of queue as a measure of congestion in an emergency department particularly considering those patients who left without being seen due to excessive waits. They evaluated the queue length and wait times for two types of queues, including queue for treatment and queue for bed in the ward. Tang *et al.,* (2014) specifically considered routine and urgent patients in order to analyse and improve the current outpatient appointment scheduling, by assessing the queue length, wait time, and physician's idle time and overtime.

In addition to queue length, the *arrival rate* of patients also leads to extreme overloading in the queue system, and has been considered by a few studies to assess the queue situation. Wijewickrama and Takakuwa (2005) particularly considered the inter-arrival times of patients. Laskowski *et al.,* (2009) considered the arrival rate of heterogeneous patients. They specifically assessed the effect of serious patients on the common queue consisting of all patients in an emergency department. A few studies have considered collecting data at different times of the day and/or different days of the week in order to get a better insight into the variation in the arrival pattern of patients (Huarng and Lee 1996; Harper and Gamlin 2003; Zhu *et al.,* 2012).

In busy hospitals of developing nations, the arrival rate of patients is excessively high and variable due to non-existent appointment systems, which increases the queue length considerably. During busy hours, the arrival rate of patients can go up to at least 50 patients per hour, hence, the number of patients to be served is extremely high resulting in excessive wait times. However, during some hours of the day, the arrival rate is quite low. Therefore, there is a necessity to monitor and improve these over-whelming queues in a complex patient flow system in developing countries, considering the extreme variation in the arrival pattern of patients, which is the purpose of the current study.

Patient *no-shows* is another variable which has gained importance in literature in order to prepare optimal appointment schedules. Hassin and Mendel (2008) evaluated the effect of patient no-shows on the waiting time of patients with appointments and associated costs. Klassen and Yoogalingham (2009) considered the effect of different levels of no-shows on the combined cost of wait time of patients and doctor's idle time, and wait time of patients and overtime. They recommended using overbooking of appointments or variation in appointment slots in order to incorporate no-shows in appointment schedules. Feldman *et al.,* (2014) considered no-shows in addition to the cancellation behaviour of patients and patient preferences for appointment days, in order to prepare optimal appointment schedules. Samorani *et al.,* (2015) considered no-show predictions as well as daily show predictions to analyse the trade-off between low wait time and clinic overtime, in order to improve the appointment scheduling system for an outpatients' clinic. However, in developing countries where all patients are walk-in and appointment systems are non-existent, no-show behaviour is of little importance.

The *length of stay* in the hospital and the *discharge rate* of patients has been included by some studies as important criteria for optimal utilization of resources, mainly when considering inpatients and emergency patients. Dermott and Stock (2007) assess the impact of different structural variables on the average length of stay, such as capital spending, hospitals in urban and rural areas, high bed capacity and teaching versus non-teaching hospitals. Some other workforce management variables were also considered, including high salaries and high staffing levels. Griffiths *et al.,* (2005) evaluated length of stay with respect to different the sources of admission including other wards and emergency room within the hospital, higher dependency units, elective/emergency surgery and X-ray department. A few studies used the length of stay as a measure of bed capacity planning. Bruin *et al.,* (2010) evaluated the required number of beds in different clinical wards by investigating the variation in length of stay in these wards; along with other causes including prolonged stay of some patients and disproportionate resource consumption. Mandelbaum *et al.,* (2012) compared the average length of stay in four different wards, and assessed its relationship with bed occupancy as well as bed turnover rates. Considering *discharge rate*, Goddard and Tavakoli (2008) investigate the relationship between the discharge rate and number of patients on the waiting list to be admitted to the hospital, whereas Mandelbaum *et al.,* (2012) considered patient admissions and discharges to assess the amount of workload of medical staff in different wards.

Commonly used queuing variables employed by previous works either relate to improving appointment systems for outpatients' departments or optimal utilization of resources when considering inpatient and emergency services. However, in busy public hospitals of developing nations, all patients are walk-in and wait times are extremely high. Additionally, immense overloading of patients further add to the complexity of operating a queue system in an organized way. Therefore, these crucial aspects necessitates the need to first identify the queuing factors, followed by assessing the impact of these variables on the queue situation to identify bottlenecks which require an intervention. A summary of commonly used queuing variables is shown in Table 2-1.

Table 2-1: Commonly Used Queuing Variables in Queue Management Studies in Healthcare

Common Queuing variables	**Definition and Description**
Waiting time	*Generally, difference between when a patient arrived and when service started*
	-Difference between arrival time and the time consultation begins *(Adeleke et al., (2009); Saremi et al (2015); Baril et al.,(2016))*
	-Difference between arrival time and appointment time *(Fetter and Thompson (1966); Harper and Gamlin (2003))*
	-Difference between time consultation begins and appointment time (*Babes and Sarma (1991); Cayirli et al., (2008)*)
	-Waiting time for early and late arrival for appointment *(O'Keefe (1998); Zhu et al., (2012))*
	-Uniform waiting times over all appointment slots/waiting within a pre-specified threshold (*Cayirli et al.,(2008))*
	-Waiting time at the waiting list before being admitted for treatment (*Goddard and Tavakoli (2008)*)
	-Waiting from decision to admit until admission to the ward (*Lane et al., (2000)*)
	-Response time *(Aboueljinane et al., (2013))*
Service Time	*Generally, difference between the time when service started and when it ended*
	Difference between the time when consultation started and when it ended *(Liu and Liu (1998); Hill-Smith (1989); Brahimi and Worthington (1991); Adeleke et al., (2009); Harper and Gamlin (2003) ;Cayirli et al., (2008); Zhu et al., (2012); Mankowska et al.,(2014))*
	-Difference between the time a patient entered till when he/she exited the system/mean time in the system

	(Laskowski et al., (2009);Huarng and Lee (1996); Lehaney et al., (1999))
	-Service time started prior to a patient admitted the ward due to the procedures involved before a patient is admitted (*Mandelbaum et al., (2012)*)
Resource Utilization	*Optimal allocation and utilization of resources.*
Bed Capacity	-Relationship between bed occupancy and demand (*Lane et al., (2000); Gorunescu et al., (2002); Cochran and Roche (2008)*)
	-Optimal number of beds required (*Harper and Shahani (2002); Li et al., (2009)*)
	-Bed occupancy, bed turnover rates and idle bed capacity (*Mandelbaum et al., (2012)*)
	-Bed occupancy and refused admissions/transfers (*Ridge et al., (1998); Bruin et al., (2010)*)
Nurse Scheduling and Requirement	-Costs associated with variables related to nurse scheduling (*Griffiths et al., (2005)*)
	-Workload for nurses (*Wright et al., (2006); Mandelbaum et al.,(2012)*)
	-Relationship between demand for nurses and bed occupancy level (*Yankovic and Green (2011)*)
Physician Scheduling and Requirement	-Number of doctors required and related variables (*Huarng and Lee (1996); Liu and Liu (1998)*)
	-Costs associated with variables related to availability of doctors (*Klassen and Yoogalingham (2009)*)
	Specific issues regarding rostering of doctors (*Rising et al., (1973); Lane et al., (2000); Bruin and Detti (2014)*)
Other Medical staff members	-Other medical staff members and related variables (*Laskowski et al., (2009); Mankowska et al., (2014)*)
Length of Queue	*The length of the queue at a particular point in time*
	-Queue length and other related variables (*Huarng and Lee (1996); Cote (1999)*)
	-Queue length for multiple activities in a multi-stage facility (*Silvester et al., (2004)*)
	-Queue length for different types of patients (*Au et al., (2009); Tang et al., (2014)*)
Idle time of doctors	*The amount of time that the doctors are not attending to a patient*
	-Relationship between idle time of doctors and wait time of patients

	(Fetter and Thompson (1966); Klassen and Rohlder (1996)) -Idle time and overtime of doctors (*Cayirli et al., (2008)*) -Costs associated with idle time and related variables (*Anderson et al., (2015)*)
Arrival Rate of Patients	*The rate at which patients arrive to receive a service* -Interarrival times (*Wijewickrama and Takakuwa (2005)*) -Arrival rate of different types of patients (*Laskowski et al., (2009)*) -Variation in arrival rate due to hourly/daily and busy/not busy times (*Huarng and Lee (1996); Harper and Gamlin (2003); Zhu et al., (2012)*)
No-shows	*When patients do not arrive for their appointments* -No-shows and other related variables *(Hassin and Mendel (2008); Feldman et al., (2014))* -Costs associated with no-shows and related variables *(Klassen and Yoogalingham (2009))* -No-show and show predictions *(Samorani et al., (2015))*
Other Variables	
Length of Stay	*The difference in time from when a patient is admitted to the hospital till discharged* -Relationship between length of stay and structural variables (*Dermott and Stock (2007)*) -Length of stay and sources of admission (*Griffiths et al., (2008)*) -Relationship between length of stay and bed capacity and demand (*Bruin et al., (2010); Mandelbaum et al., (2012)*)
Discharge Rate	*The rate at which patients are discharged* -Relationship between discharge rate and wait time (*Goddard and Tavakoli (2008)*) Discharge rate and workload of staff (*Mandelbaum et al., (2012)*)

2.4 Different Modelling Techniques Used for Queuing Analysis

A number of different modelling techniques have been used by researchers to evaluate existing queue systems, by using a different set of parameters. Among various methodologies, Discrete-Event Simulation and Queuing Theory/Models are among the most commonly utilized modelling techniques for queuing analysis.

Discrete-event Simulation is an effective tool which is most commonly used to efficiently allocate resources to improve patient flow, while reducing costs (Jun *et al.,* 1999). Almost all studies regarding Simulation modelling have evaluated the patient flow system from the point of entry to exit, identified bottlenecks and provided recommendations for improvement. Most previous works have used Simulation to assess and improve an existing *appointment system* in *outpatients'* clinics. In most cases, a number of different appointment systems are proposed to select the optimal schedule. Klassen and Rohlder (1996) used simulation modelling to compare various appointment schedules in order to decide for the positioning of slots for urgent patients within the schedule (either at the start, middle or end). A number of factors were kept fixed in the simulation model including fixed interval for consultation time (10 minutes), fixed no-show level (at 5%) and fixed total session time (at 3.5h). Harper and Gamlin (2003) compared nine different appointment schedules using simulation modelling for urgent as well as other types of patients including regular, follow-up and ward discharge. Various schedules were assessed on the basis of different types of wait time including wait to first service, percentage of patients who were seen within 30 minutes of waiting and the average time in the system (clinic). Aharonson-Daniel *et al.,* (1996) considered a multi-server system, with five doctors in five different consultation rooms. Simulation modelling was used to assess the effect of varying arrival pattern of patients and service time on the wait times and queue lengths for each doctor; in order to prepare an appropriate schedule. Liu and Liu (1998) also modelled a multi-server system. In this case, Simulation was used to evaluate the number of appointments to be given in a block appointment schedule such that the wait time and doctor's idle time is reduced. A combination of different factors was used in the simulation model including number and arrival pattern of doctors, patient no-shows, and total number of appointment blocks and length of each block. Andersen *et al.,* (2015) assessed overlapping appointment systems where the simulation model tested various overlap periods to determine an optimal schedule. The simulation model also showed that largest overlap and reduction in costs for wait, doctor idle time and overtime, are achieved with high level and coefficient of variation of service time, and no-show rate. Zhu *et al.,* (2012) evaluated causes of uncertainty leading to long wait times and overtime of physicians, and incorporated changes in the simulation model accordingly, in order to observe the impact on these two factors. However, the absence of appointment systems in public hospitals of developing countries further increases the level of uncertainty in the patient flow system. Therefore, supervision of variable influx of patients in the health systems is a crucial issue, which is addressed by the current study.

A few studies have used *Simulation* modelling to incorporate *Walk-in* patients when assessing an *appointment scheduling* system. Fetter and Thompson used Simulation to identify the effect of variation in a few variables on the wait time of appointed and walk-in patients, and idle time for doctors. For instance, limiting early and late arrival of appointed patients to 5 minutes reduced the wait time of appointed patients allowing for more time for walk-in patients. However, a change in appointment intervals did not impact both wait times, but increased load of physicians increased both wait times without much decrease in idle time. Cayirli *et al.,* (2008) considered the percentage of walk-in as well as no-show patients as a measure of 'congestion' in the simulation model, when evaluating different appointment schedules. The simulation also included sequencing, consultation times and interval adjustment for appointments of different type of patients including new and return for both appointed and walk-in patients. In both studies, the walk-in patients have been considered as part of an appointment schedule. There is a dire need to formulate a model which evaluates the queue system where all patients are walk-in with no pre-booked appointments.

Some studies have utilized *Simulation* modelling for *resource utilization*. Cote (1999) and Benneyan *et al.,* (1994) used Discrete-event simulation to assess the impact of varying number of examination rooms on resource utilization and wait time of patients. Both studies concluded that increasing examination room capacity had little impact on patient wait times, whereas room utilization was affected to some extent. In a recent study carried out in an ophthalmic hospital, Daultani *et al.,* (2016) utilized Simulation to evaluate seven different scenarios, which considered different combinations of resources including two machines, a consultant and an optometrist. The study concluded that improvements are observed in the number of patients served, wait times and resource utilization, when considering different scenarios.

Some studies specifically target the *bed capacity* using *Simulation* modelling. In some studies bed allocation problem has been studied with respect to number of transfers and refused admission. Ridge *et al.,* (1998) and Harper and Shahani (2002) used simulation modelling to assess the impact of varying number of all and reserved beds for emergency patients on emergency and Elective patient transfers. The simulation model in former study highlighted the trade-off between the occupancy level of beds and the number of transfers. Also, reserved beds for emergency patients had a small decrease on emergency transfer but an increase in the percentage of planned transfers. The simulation model in the latter study illustrated the effect of lack of beds in multiple units including a ward, specialty bed pool and hospital as a whole. It was highlighted that refused emergency admissions will have to be transferred to another speciality leading to increased burden on other specialties, and refused elective admissions will increase patient deferrals and consequently elective patient waiting lists. Zhu *et al.,* (2012) used simulation modelling to evaluate the impact of future demand growth for emergency and elective admissions, on the requirement of beds in an Intensive care unit. Considering two different growth rates, the authors concluded that on average, one and two additional beds are required to meet 5% and 10% growth rate respectively.

Some studies have considered *staff* distribution using *Simulation.* Griffiths *et al.,* (2005) used simulation model to identify the required number of rostered and supplementary nurses in an Intensive care unit, and associated costs. The Simulation model considered variation in the arrival of patients from different sources (including elective and emergency surgery, emergency room, ward, another hospital, higher dependency unit and X-ray department), and variation in the time of admissions and length of stay of patients from these sources. The model showed that an increase in rostered nurses in line with increased patient demand will reduce additional costs for nursing staff. Gul and Guneri (2012) used Simulation modelling to determine the impact of varying number of nurses in addition to physicians and receptionists to assess the impact on their utilization, length of stay of patients, patient throughput (number of discharges in a day), and utilization of different units in the hospitals (emergency, monitor beds area, emergency response room and Resuscitation). Additionally, the obtained results were further analysed in terms of variation in patient demand patterns.

The issue of shortage and poor allocation of resources is more severe in large busy public hospitals of developing countries not only in emergency and inpatient departments but also in outpatients' department, due to absence of appointment systems. The bigger challenge is the lack of strategy which prevents the administration to conduct an in-depth evaluation of the patient flow system. Therefore, the current study aims to develop a model which allows for detailed assessment of a walk-in patient flow system.

Considering *Queuing models*, Brahimi and Worthington (1991) evaluated the *appointment schedule* in an *outpatients'* department for a single service. A number of factors were assessed including the arrival time of patients which was compared with their appointment time (either punctual, early or late), wait time and queue length. Using time-dependent queueing models, an optimal appointment schedule was proposed with more appointments at the beginning of the session with equal interval appointments later in the schedule, providing a better estimate of how the queue will proceed. In most public hospitals of developing countries, there are no appointment systems, which further magnifies the queue problem. Adeleke *et al.,* (2009) considered a single-channel queue system and evaluated the traffic intensity which shows the load of patients and average time in the system, in addition to total wait time and service time. The authors recommended implementing a multi-server queue system instead so that more patients can be examined at one time with more doctors, increasing the number of health centres, and increasing the number of paramedic officers who can conduct preliminary examination of patients and refer to physicians only if need be. These recommendations seem very generic and objective solutions are not provided to deal with extreme overload of patients. In developing countries, shortage of medical personnel and facilities are ongoing challenges. In addition to continuous efforts for increasing medical centres and personnel, there is a crucial need to model a queue system with patient overload and long waiting, as it is. Biju and Naeema (2011) also evaluated the issue of staff inadequacy using queuing models by considering various factors such as service and arrival rate, and service and wait times. They concluded that arrival rate is much higher than the service rate, leading to excessive queuing and highlighted the issue of lack of staff. The authors suggested that staff from other departments can be utilized to deal with patient load during peaks times, and more experienced staff members need to be employed for cash counters and for dispensing medicines at the pharmacy. Although the authors emphasize on the issue of staff inadequacy at peak times, however, these recommendations seem vague with lack of a practical solution. In developing countries, with excessive wait times and staff inadequacy, the need of a proper guideline is crucial in order to assess and improve the queue system.

Some studies have considered *additional issues* when utilizing *Queuing models*. Wang (2004) developed a priority queue model based on the issue of the risk of complications faced by the patient if treatment is delayed, considering these patients as urgent cases. The critical time for patients is considered as a significant variable in the queueing model. The model illustrates the difference in queue lengths and wait times with and without prioritization of urgent patients. The trade-off between the welfare of all patients in terms of reduced risk and increased queue length and wait time of less urgent patients is highlighted. Kowzlowski and Worthington (2015) use queuing models, specifically Continuous Time Markov Chain (CTMC) model to evaluate the maximum wait time policy. The model highlighted that increased probabilities of patient withdrawing and rescheduling and maximum wait time guarantee will reduce the wait times but will lead to an increased under-utilization of resources. This issue can be more severe for smaller specialties with small waiting lists.

Queuing theory and models have also been used to assess *resource allocation* problems. Briuin *et al.,* (2010) utilized the Erlang's loss model to assess the required number of beds for inpatients. The main factors considered were different levels of blocking due to refused admission, number of arrivals, average number of occupied beds, number of operational beds and average length of stay. Also, the effect of merging departments and mixing patient flows was observed on the required number of beds using 5% occupancy rate and average length of stay. Mayhew and Smith (2008) used queuing models in an A&E department, in order to assess the completion time of patients if shifted to another department from A&E. Varying completion targets from 1 to 12h, different percentages of patients discharged (90, 95, 98 and 99), and varying arrival and service rates were considered to observe the effect of truncating the long treatment path as compared with the initial model. It was concluded that the re-designed model has reduced completion times for a higher percentage of patients discharged. By evaluating the patient flow system as an Inverted V queuing model, Mandelbaum *et al.,* (2012) evaluated the process of assigning patients from emergency to different wards, considering bed availability and staff workload. Mittal (2010) used queuing models to investigate the effect of varying number of beds for three different wards on various queuing characteristics, including probability of zero patients in the system, waiting time for beds, probability of wait time exceeding a specific time and probability that all beds are busy. For developing countries, there is a shortage of resources including beds and staff, for inpatients. Additionally, there is increased workload, extreme overloading and limited staff in outpatients' department as well, since there are no appointment system. The current study aims to evaluate the queuing system for outpatients' department where all patients are walk-in.

Some *other techniques* have also been employed to assess the queue situation. System Dynamics has been used by some researchers as a way to evaluate queuing problem. Brailsford *et al.,* (2004) utilized the System Dynamics approach to assess the impact of changes in admission process on the level of bed occupancy. Based on model results, the authors showed that reduction in admissions of elderly patients (by referring them to other community services) and patients with respiratory problems, leads to a significant increase in the bed occupancy level; hence extending bed availability for increased planned GP referrals in the coming years. Lane *et al.,* (2000) used System Dynamics modelling to assess the effect of changing bed capacity and demand, separately and combined, on average occupancy level as well as other various performance measures including average time to A&E consultation, average percentage of elective cancellations and average daily A&E utilization. The model results show that reducing the number of beds only slightly improves bed occupancy but significantly increases elective cancellations. Also, reduction in A&E admissions leads to a slight increase in elective cancellations, however, waiting time of all patients reduces to a great extent. Smith and Roberts (2014) used System Dynamics to provide a detailed explanation of and identify factors which affect the patient flow process and the interaction among different departments. Some major factors identified through this model are delays in the rate of shifting patients from emergency department to other wards, delays in the rate of discharge and time of discharge, rate of transferring beds from one medical ward to another due to increased emergency admissions and bed occupancy rates of different wards. Additionally, the authors used sequential bifurcation technique to further break down the identified factors to highlight the most crucial elements affecting patient flow.

George *et al.,* (1983) formulated a Linear Programming model for a surgical department by considering resource usage and availability as well as diagnosis and urgency of patients. The model was used to evaluate the effect of increased bed capacity by considering the number of admissions depending on patient level of urgency and type (day case or inpatient), current bed availability, length of operation and current theatre space availability. Giachetti *et al.,* (2005) considered Open Access Scheduling to improve the appointment system at an outpatients' clinic to address problems such as long patient throughput, large backlog of appointments and high no-show rate. The factors analysed included variability in demand and capacity and its relationship to appointment backlog, and the effect of current scheduling system on the level of patient backlog and no-show rate.

Some authors used a *combination* of different *methodologies* to assess the queue situation. *Simulation* has been used in conjunction with *other* techniques. A few studies have used Simulation with Optimization to assess a queue problem. Ahmed and Alkhamis (2009) used simulation with optimization to provide a more comprehensive evaluation of different factors affecting the patient flow system in an emergency department. Varying budget levels and arrival rates of patients were used to evaluate the effect on patient throughput and staff allocation for different categories of patients including emergency, less severe and normal. Furthermore, varying wait times were used to assess the impact on budget and staff allocation. Klassen and Yoogalingham (2009) used Simulation with Optimization for an outpatients' department to evaluate the effect of plateau-dome appointment scheduling on different performance measures including wait time, day end time, idle time and overtime. Furthermore, the effect of varying no-show rates, total appointment session length and end of day time was assessed on this type of appointment system. Angelis *et al.,* (2003) used estimation of target function with Simulation and Optimization, to specifically evaluate the average time in a system for patients belonging to different categories, in a multiple heterogeneous service system. The target function where average time in system was a function of the arrival rate of patients and number of servers, was estimated by using the simulation model to identify the relationship among these variables. The estimated target function was then used in the optimization model to determine the average time in system with varying levels of constraints on budgets, arrival rate and number of servers. Baril *et al.,* (2016) utilized Discrete-event simulation to test and validate different scenarios proposed by employing Lean Management using a

Kaizen event; to reduce wait times of patients for a chemotherapy treatment in a haemotology/oncology clinic. This approach was utilized mainly to determine appropriate appointment schedules for two types of services including consulting with the doctor followed by start of treatment, and to achieve consistency between them. Zhang *et al.,* (2012) used Demographic and Survival analysis, in addition to Simulation and Optimization to develop a decision support system to determine an optimal level of bed capacity using different factors such as length of stay, initial conditions, population growth rate, per capita arrival rates and varying service level criterion. Lehaney *et al.,* (1999) used Soft systems methodology with Simulation to suggest an improved appointment scheduling in an outpatients' department. They emphasized that a number of steps need to be taken before constructing a simulation model including setting out the objectives, monitor and take control of actions needed, and knowing internal and external criteria for policy-making. Then, after the model has been constructed, a thorough follow-up need to be undertaken to assess its ability to be implemented. The current study has the objective of identifying different factors which lead to excessive queuing.

Queuing models have also been used in combination with *other* Operational Research techniques to assess a patient flow system. Chao *et al.,* (2003) used Queuing models with Optimization to evaluate the effect of customer 'switching' from one service to another based on better service, on demand and resource allocation. The rate of three different levels of switchable customers; no switching, partial and complete, is considered in addition to service level, aggregate arrival rate, service capacity and expected wait time. Li *et al.,* (2009) used Queuing theory in conjunction with Goal Programming for a bed allocation problem. The main objective was to assess the effect of number of beds and occupancy levels on the admission level of patients and daily profits for different departments. The goal programming model used the results from the queuing model as an input, considering the admission level and profit targets of each department. Silva and Serra (2008) use maximal covering allocation model with priority queuing to assign patients with two different priority levels to the closest facilities independent of their priority levels, such that maximum population is covered. In addition to the waiting time and service rate depending on the level of priority for each customer, the amount of time servers are busy and average delay experienced by a customer due to another customer in service are also considered. Laskowski *et al.,* (2009) used agent-based modelling with queuing models to explore the bottlenecks in the patient flow system in an emergency department. The issue of physician allocation and its effect on queuing length was highlighted for patients with high, moderate and low priority levels. Also, the effect of varying arrival rates on patient wait time and service time were emphasized upon. Although the authors did highlight the issues of high wait time and physician allocation, but an objective solution is lacking since the idea was to pinpoint issues which require further investigation. In developing countries, with high arrival rates and queue lengths of patients with no appointments, there is a crucial need of developing an analytical model, which not only pinpoints the areas of high wait times but provides recommendations for improvement. In a recent study, Saremi *et al.,* (2015) use Mathematical programming, Simulation and multi-objective Tabu search methods to suggest an optimal appointment schedule for patients with emphasis on minimizing wait times and completion time of facility. The Mathematical programming model is used to generate initial solution, which is further refined by Tabu search for optimal schedules. Furthermore, different schedules are evaluated using Simulation under added complexity due to variation in service rate, different patient types, resources constraints (doctors/nurses), and multiple services required by patients. Although these factors do make the queue system complex, however an existing appointment system facilitates in improving the patient flow system. In large public hospitals of developing countries, with no appointment systems, the patient flow system is not only more complex but there is an increased level of uncertainty with respect to patients' arrival rate at multiple stages with excessive wait time, and proper allocation of resources is challenging. Therefore, the current study aims at assessing a queue system by considering these factors.

A summary of various operational research techniques used to assess most common characteristics of queuing analysis is shown in Table 2-2.

Table 2-2: Summary of Different Operational Research Techniques for Common Patient Flow Factors

Most Common Techniques \ Common Patient Flow Factors	Appointment Scheduling	Wait time	Bed Capacity	Staff Allocation	Patient throughput	Arrival rate of patients	Level of Budget (cost/profit)
Simulation	✓	✓	✓	✓		✓	✓
Queuing Models	✓	✓	✓	✓	✓	✓	
Other Techniques							
System Dynamics			✓		✓		
Linear Programming			✓				
Open Access Scheduling	✓	✓					
Combination of Techniques							
Simulation and Optimization	✓	✓	✓	✓	✓	✓	✓
Simulation and Soft Systems Methodology	✓						
Queuing Models and Optimization				✓		✓	
Queuing Models and Goal Programming			✓				✓
Queuing Models and Agent-based Modelling				✓		✓	
Mathematical Programming, Simulation and others		✓			✓		

2.5 Some Limitations of Other Operational Research Techniques and the Use of Data Envelopment Analysis for the Current Study

Various OR modelling techniques have been used extensively in the literature to evaluate a queuing problem. However, they have a number of *limitations* which restrict their usefulness for the evaluation of patient flow system in large busy public hospitals of developing countries, where all patients are walk-in.

Modelling techniques like Simulation and Queuing theory, and statistical methods consider the *average* values of the queuing parameters included in the model. For instance, Angelis *et al.*, (2003) and Lehaney *et al.*, (1999) consider the average time in the system. Harper and Gamlin (2003) included the average wait time of patients in different sessions of the day (morning/afternoon) and the percentage of patients who waited within a certain time period. Wijewickrama and Takakuwa (2005) particularly evaluate the average process and delay time while Griffiths *et al.*, (2005) consider the average queue size. Zhu *et al.,* (2012) and Andersen *et al.,* (2015) use the mean service time to evaluate the patient flow system.

However, the average value for most commonly used queuing variables is not a suitable measure when evaluating a complex patient flow system as in developing countries. There is a high level of variability and uncertainty in the patient flow. Therefore, an average value does not represent the system accurately and hence lacks detailed information regarding the magnitude of queuing problem. Hence, for such disorganized queue systems, where arrival of patients is not fixed, there is a need to consider each and every value of all queuing variables, rather than average values in order to assess the uncertainty with greater accuracy.

Almost all modelling techniques employ *statistical distributions* to allow for a good fit to the actual data to increase the level of accuracy, as indicated by queue management literature. Huarng and Lee (1996) specify Exponential and Normal distributions for the service times in different departments, whereas Andersen *et al.,* (2015) used Uniform distribution. Mital (2010) assumed that service times as well as inter-arrival times follow Exponential distribution. Mandelbaum *et al.,* (2012) used Poisson distribution for the arrival pattern of patients and Exponential distribution for length of stay. Feldman *et al.,* (2014) used Poisson distribution for the number of appointment requests by patients in a day.

Although, *previously defined distributions* represent the actual system as accurately as possible, however they are only an 'approximation' of real-life problems and behaviour of patients (Cheema 2005) in a queue. Besides, developing mathematical and simulation models as well as validation of these models is extremely time-consuming, since they require an effort to imitate the actual system as accurately as possible (Benneyan *et al.,* 1994). For large public hospitals in developing countries, such as Pakistan, there exists increased variability and uncertainty in the patient flow, not to mention the unpredictable arrival pattern of patients. Furthermore, absence of appointment systems might obstruct an accurate representation and evaluation of the patient flow system. Therefore, these models with pre-specified probability distributions are not suitable since there are chances of getting misleading results.

The modelling techniques provide a detailed analysis of the current queue system and the bottlenecks as well as an opportunity for *'what-if'* analyses, where varying values of a parameter are utilized to assess the effect on other queue variables. For instance, to evaluate the effect on wait time and queue length, Aharonson-Daniel *et al.,* (1996) considered varying arrival rates and service times, whereas Cote (1999) considered varying number of examination rooms. Harper and Shahani (2002) used different number of beds to evaluate the impact on emergency and elective transfers. Gul and Guneri (2012) used varying number of nurses and physicians to determine the effect on length of stay and number of discharges.

However, in this case, there is a need to define different values prior to the analysis. Various models need to be run under different scenarios to determine the optimal level of the variable, which is not only time-consuming, but does not provide an objective solution quickly enough. There is a dire need of a methodology which simultaneously monitors the current queue situation and provides a definitive optimal target to reduce the wait time of patients.

The current study intends to evaluate the patient flow system for walk-in patients in a busy public hospital of a developing country. Considering the main aim of this study, the above mentioned limitations of other modelling techniques can severely constrain an in-depth and accurate evaluation of the queue system. *Data Envelopment Analysis (DEA)* is a popular efficiency assessment technique which has the ability to address these limitations, hence making it suitable to achieve the goal of the current study.

<u>Chapter Summary</u>

The present chapter provided a comprehensive review of literature in Queue management. Previous studies have been carried out in many different departments of a hospital. When considering outpatients, all studies focus on improving an already existing appointment system. Additionally, there are negligible studies which have evaluated the arrival pattern of walk-in patients in an outpatient setting, and only with respect to its effect on scheduled patients. None of the studies assessed a patient flow system where 'all' patients are walk-in. By conducting a review of studies carried out in various developed and developing countries, it can be concluded that though studies in developing countries have highlighted some queue-related issues and its consequences; however, none of the studies have analysed a queue system with large number of walk-in patients only. An evaluation of queuing variables was also conducted to identify the way they have been utilized in different studies. Among various queuing variables, waiting time and service time have been used in almost every study, however, they have been defined differently depending on the objective of the study. Some additional parameters include length of queue, variables associated with resources such as beds and physicians, arrival rate and some others. In case of outpatients', almost all previous works included these queuing variables aiming to improve an existing appointment system. However, in busy public hospitals of developing nations, all patients are walk-in with extreme overcrowding which creates hurdles in effective operation of a queuing system, leading to high wait times. Therefore, there is a necessity to identify queuing factors which affect the efficiency of the queue system, and to provide a framework for improving the queue system. Furthermore, an evaluation of some Operational Research techniques for assessing queue systems has been conducted, where Discrete-Event Simulation and Queuing Models are among the most frequently utilized. Additionally, a few limitations of these OR techniques have been elaborated. For instance, the use of averages and pre-defined statistical distributions lacks the detail required to effectively evaluate a complex queue system, which is highly uncertain and unpredictable, as in case of walk-in patients. Also, 'what-if' analysis is extremely time consuming and does not give straight-forward recommendations. Hence, the benefits of an efficiency technique, DEA, overcomes most of the limitations of other OR techniques making it extremely useful for the purpose of the current study, that is, evaluating the queue system in the absence of appointment system.

CHAPTER 3 LITERATURE REVIEW: DATA ENVELOPMENT ANALYSIS THEORETICAL FRAMEWORK AND HEALTHCARE APPLICATIONS

Chapter Overview

The current chapter mainly provides an exhaustive assessment of Data Envelopment Analysis (DEA) modelling as applied to healthcare. First of all, theoretical underpinnings of DEA modelling have been highlighted, with graphical and mathematical representations. Then various healthcare studies have been evaluated with respect to three different perspectives. The first part reviews the DEA studies which specifically compare the efficiency of hospitals or other health institutions. The second part compares the efficiency measurement studies conducted in developed and developing countries, by identifying the causes of inefficiency and recommendations provided. The last part evaluates those studies which have used DEA in combination with other analytical techniques, either to compare with DEA results or as pre or post-analysis. The next section identifies and evaluates commonly used health-related input and output variables in different studies. Furthermore, a few extended healthcare applications of DEA are demonstrated, emphasizing that DEA has the capability to provide suitable results for other health-related applications, including queuing problem, deviating from its traditional utilization.

Efficiency measurement represents the first step towards the evaluation of a coordinated health care system, and constitutes one of the basic means of audit for the rational distribution of human and economic resources (Ramanathan 2005). Assessing operational efficiency enables healthcare providers to better understand their management effectiveness (Chuang *et al.,* 2011), given increasing costs, resource constraints in terms of personnel and infrastructure (Aksezer 2016), and the challenge of providing high quality service to all patients. In this regard, efficiency analysis in terms of wait times needs to be addressed as it demonstrates the overall efficacy of health services and patient satisfaction (Chandra 2015).

Efficiency analysis can be considered as a broad framework consisting of several stages where developing a DEA model is part of it. These stages can act as a checklist for desired objectives, and can ensure that the results are reliable and manageable (Emrouznejad and De Witte 2010). Some studies provide guidelines specifically with regard to different aspects of DEA modelling (Osman et al 2014). Cook *et al.,* (2014) discuss some of the major elements of a DEA model, including model orientation, input/output variable selection and use of mixed or raw data; with the objective of providing clarification and guidance to researchers and practitioners alike, with regard to these issues. Brown (2006) suggests different approaches to deal with three specific issues within DEA modelling including outliers, measurement error in data and heterogeneity; and emphasize on differentiating among these three factors before conducting DEA analysis.

However, a few studies proposed a 'framework' which aimed at improving the process of performance assessment as a whole, by adopting a more strategic approach. In this regard, Emrouznejad and De Witte (2010) have proposed the 'Cooper' framework (see Figure 3-1 below). It specifically incorporates DEA modelling, particularly to assist in the process of performance evaluation of projects with large and complicated datasets (Emrouznejad and De Witte 2010). The proposed framework consists of six stages including Concepts and Objectives, On structuring data, Operational Models, A Performance Comparison Model, Evaluation, and Results and Deployment. The first two phases are related to developing an understanding of the problem under consideration and operational aspects of the units of analysis. The next two stages are concerned with the development of the appropriate DEA model for the concerned problem (Emrouznejad and De Witte 2010). The last two phases are related to analysing and interpreting the DEA results including preparation of documentation for non-DEA experts (Emrouznejad and De Witte 2010). Although particularly considered for DEA, the authors mention that this framework can be modified to incorporate parametric techniques as well, such as Free Disposal Hull (FDH) and Stochastic Frontier Analysis (SFA). Some other frameworks which can incorporate a DEA model have also been developed. For instance, Osman et al (2014) suggested a framework known as AIM-UP which consists of different stages including Understand, Prepare, Analyse, Implement and Monitor. In this case, Prepare and Analyse stages are associated with developing a DEA model, however, parametric techniques such as Stochastic Frontier Analysis (SFA), could also be used (Osman et al 2014). Azadeh *et al.,* (2009) developed a framework consisting of three methodological tools including Delphi, Voting Analytical Hierarchy Process (VAHP) and DEA. The authors emphasize on the benefits of using this framework to compare several IT/IS investment options to determine the optimal one.

Figure 3-1: DEA Model as Part of 'COOPER' Framework

(Source: Emrouznejad and De Witte (2010))

3.1 Theoretical Underpinnings of DEA Modelling

Data Envelopment Analysis is a multiple criteria decision-making technique, which is used to model the relationship between multiple inputs and outputs in order to assess the efficiency level of all units, known as decision-making units (DMUs). The basic concept of DEA comes from the *production frontier* approach as discussed by Farrell (1957) using the isoquant diagram, which displays the computation of efficiency scores by specifying a nonparametric piecewise linear frontier function. Based on this concept of efficiency measurement, a non-parametric mathematical programming framework, Data envelopment analysis, was developed by Charnes *et al.,* (1978) (Charnes, Cooper and Rhodes as CCR model). Later on, the traditional DEA model was extended by Banker *et al.,* (1984), most popularly known as the Banker, Charnes and Cooper (BCC) Model.

3.1.1 Graphical Representation of DEA Theory and Model

A DEA model allows for the construction of an *'efficient frontier'* formed by some DMUs which demonstrate best practices, among the entire set of DMUs. It then allocates the efficiency level to all units based on their distances from the efficient frontier. Units which lie on the frontier are said to be 100% efficient, while others which are away from the frontier are inefficient, with efficiency scores below 100%. DEA identifies a reference set for each inefficient DMU, which comprises of corresponding efficient DMUs which act as benchmarks for improvement. However, the advantage of DEA modelling is that apart from efficiency scores, it also provides additional information which can assist policy-makers in decision-making. For instance, DEA presents 'target' values for each input and output which show the exact amount required for each of these variables, in order to increase the efficiency level to 100% (Liu *et al.,* 2013a; Lee and Kim 2014). The provision of a target value is one of the most significant advantages of this modelling technique. It highlights the areas where inefficiency prevails as well as provides a guideline to all inefficient units for the degree of improvement improve the efficiency level. Hence, in case of large public hospitals in developing nations, where there is lack of strategy and proper guidance for assessing and improving patient flow in the absence of appointment systems, DEA is a highly appropriate assessment technique. As opposed to other statistical techniques which require pre-specified distributions, DEA is a non-parametric technique which does not require any a priori specification of the functional form or distribution of data, efficiency frontier and efficiency scores (Charnes *et al.,* 1994). Additionally, DEA produces results for each unit separately rather than providing an average across all data points. Therefore, these characteristics render DEA analysis as extremely suitable for problems where data is extremely variable and unpredictable, as in the case of walk-ins within a busy public hospital of a developing country, such as Pakistan.

Most common DEA models are either 'input-oriented' or 'output-oriented'. An input-oriented model is concerned with minimizing the quantity of inputs given that the same amount of outputs is produced. An output-oriented model focuses on maximizing the quantity of output produced subject to the condition that the amount of inputs remains unchanged (Charnes *et al.,* 1978; Banker *et al.,* 1984). A case of two-input and one-output model (Cooper *et al.,* 2007; Thanassoulis 2001) can be represented *graphically* as shown in Figure 3-2. There are four DMUs, A1, A2, A3 and A4, each producing the same amount of output, O1 while using different combinations of inputs, I1 and I2. This is an 'input-oriented' DEA model as the objective is to reduce the number of inputs given the same level of output. The line SS' is the estimated efficient frontier for the four DMUs.

Three of the DMUs, A1, A3 and A4 are 100% efficient as they are lying on the frontier. However, A2 is inefficient as it consumes more resources as compared to other DMUs. A2 can become 100% efficient if it operates at point M (at the efficiency frontier). Therefore, the efficiency of A2 at this point is OM/OA2. This type of efficiency is known as the 'technical' efficiency of a unit. A1 and A3 are the 'efficient peers' of A2, that is, the units which act as benchmarks for the inefficient unit A2. The DMUs representing the efficient peers are usually the ones which have the most closely related input/output combination to that of the particular inefficient unit.

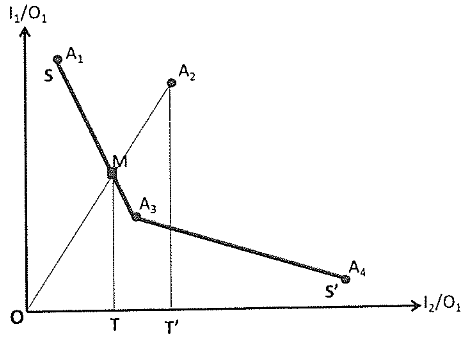


Figure 3-2: Measuring Efficiency Graphically for the Input-oriented DEA Model

However, when the number of inputs and outputs increase, the model becomes more complicated and is solved by constructing a linear programming model, as given in the next section.

3.1.2 Mathematical Formulation of DEA

In general terms, the efficiency is defined as the ratio of output to input. However, the DEA efficiency differs from an efficiency ratio since it has the ability to incorporate multiple inputs and outputs (Sherman and Zhu 2006). Assuming there are 'n' DMUs where each DMU uses 'm' inputs to produce 's' outputs, the efficiency score (which lies between 0 and 100% or 0 and 1) for any DMU_0, is mathematically defined as the ratio of the weighted sum of outputs ($\sum_{r=1}^{s} u_r y_{ro}$) to the weighted sum of inputs ($\sum_{i=1}^{m} v_i x_{io}$) for DMU_0 (as shown below):

$$0 \le \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \le 1$$

where:

y_{ro}: amount of output used by DMU_0

x_{io} : amount of input used by DMU_0

r: index for output generated by DMU_0 ($r = 1,\ldots,s$)

i: index for input utilized by DMU_0 ($i = 1,..,m$)

u_r: weight assigned by DEA to output r

v_i: weight assigned by DEA to input i

Hence considering the above efficiency score, Charnes *et al.,* (1978) developed a fractional programming model for measuring efficiency of each DMU in the dataset, known as the CCR Model (Cooper *et al.,* 2007), as shown in *Model 1a* below.

Model 1a (Fractional programming model)

$$\text{Maximize } \theta_o = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}}$$

subject to

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \le 1 \qquad j = 1,\ldots,n$$

$$u_r,\ v_i \ge 0 \qquad \forall\ r \text{ and } i$$

Given a dataset, the efficiency of each DMU is evaluated, hence 'n' optimization models will need to be run, one for each DMU_j. Consider DMU under estimation as DMU_0, in the above mathematical programming model (Model 1a), the objective is to obtain the output and input weights (u_r and v_i) that maximize the efficiency ratio θ_o for DMU_0 (DMU under evaluation). The basic idea behind weight allocation is that each unit might value inputs and outputs differently and therefore adopt different weights. Hence, each unit should be allowed to adopt the most favourable weights for inputs and outputs, in order to maximize its efficiency. The constraint shows that when this same set of input/output weight coefficients is applied to the remaining DMUs, the efficiency of 'each' DMU_j ($j = 1,\ldots,n$) should 'not' exceed 1. Hence, $\sum_{r=1}^{s} u_r y_{rj}$ is the weighted sum of outputs and $\sum_{i=1}^{m} v_i x_{ij}$ is the weighted sum of inputs, for 'jth' DMU ($j = 1,\ldots,n$). Additionally, non-negativity constraints apply to the whole fraction, that is, all outputs as well as inputs are assumed to be non-zero (Sherman and Zhu 2006; Cooper *et al.,* 2007; Ozcan 2008).

However, in order to operationalize the above fractional programming model, it is algebraically converted to a linear programming model, using Charnes-Cooper transformation (1962) (Sherman and Zhu 2006) as shown in *Model 2a.*

Model 2a (Linear programming model)
Maximize $\theta_o = \sum_{r=1}^{s} u_r y_{ro}$ subject to $\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0 \qquad j = 1,\ldots,n$ $\sum_{i=1}^{m} v_i x_{io} = 1$ $u_r,\ v_i \ge 0 \qquad \forall\ r \text{ and } i$

This linear mathematical programming model can be reformulated as a dual problem, to obtain the DEA efficiency scores. In most instances, the dual program is used to solve linear program due to advantages in technical computation (Sherman and Zhu 2006).

Also, this method can provide more useful alternative interpretations, more detailed information about the benchmarks for particular DMUs and their weights, and a better insight on the relationship between the evaluator and the units being assessed (Bogetoft and Otto 2011; Ozcan 2008). The *dual* linear programming model for Model 2a is shown in *Model 3a* below:

Model 3a (Dual programming model)
Minimize θ_o subject to $\sum_{j=1}^{n} \lambda_j x_{ij} \le \theta x_{io}$ $\quad i = 1,\ldots,m$ $\sum_{j=1}^{n} \lambda_j y_{rj} \ge y_{ro}$ $\quad r = 1,\ldots,s$ $\lambda_j \ge 0$ $\quad j = 1,\ldots,n$

For the dual linear program displayed in Model 3a, the objective is to minimize the efficiency (since it is dual efficiency) of the concerned DMU given two sets of constraints (Ozcan 2008). Lambda (λ_j) shows the weights for each 'jth' DMU. The first constraint emphasizes that the weighted sum of inputs of all DMUs ($\sum_{j=1}^{n} \lambda_j x_{ij}$) should not less than or equal to the 'ith' input of DMU_o multiplied by its efficiency level (θx_{io}), where $i = 1,\ldots,m$. Similarly, the second constraint concerns outputs that is, the weighted sum of the outputs of all DMUs ($\sum_{j=1}^{n} \lambda_j y_{rj}$) should be greater than or equal to the 'rth' output of the DMU_o being evaluated (y_{ro}), where $r = 1,\ldots,s$ (Ramanathan 2003).

The CCR Model as explained in Model 2a, operates under the assumption of constant returns to scale (CRS). This traditional model was extended by Banker *et al.*, (1984) to work under the condition of variable returns to scale (VRS), by incorporating an additional constraint in Model 2a, $\sum_{j=1}^{n} \lambda_j = 1$ (for $j = 1,\ldots,n$). This new constraint enables the DEA model to determine if a firm is operating under variable returns to scale, which could be either CRS, increasing or decreasing returns to scale (IRS and DRS respectively), utilizing the concept of Sheppard's distance function as shown by Banker *et al.,* (1984) (Emrouznejad and Cabanda 2014).

The DEA analysis is further carried out by obtaining slacks (Cooper *et al.,* 2007; Ozcan 2008). In order to obtain the *slacks* from first-stage DEA analysis, a second stage linear programming model is required to be solved after solving Model 3a.

These slack values will then be used to determine the target values for each input or output (depending on the orientation of the model) for every inefficient DMU. For the slack model, a second-stage linear programming model is run after the dual programming model in Model 3a. This model is shown in *Model 4a* below:

Model 4a (Slack model)

Maximize $\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+$

subject to

$$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = \theta^* x_{io} \qquad i = 1,\ldots,m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = \theta^* y_{ro} \qquad r = 1,\ldots,s$$

$$\lambda_j \geq 0 \qquad j = 1,\ldots,n$$

The objective of this model is to maximize the slacks for inputs and outputs for all DMUs, as shown in Model 4a. For the above model, s_i^- and s_i^+ are the slacks for inputs and outputs, respectively (Ozcan 2008). The negative sign in the superscript on input slack indicated that it needs to be reduced, whereas the positive sign on output slacks show that outputs require an increase.

Furthermore, both models (Model 3a and 4a) under the DEA analysis can be reformulated to one model (Cooper *et al.,* 2007), as shown in *Model 5a* below:

Model 5a (Two-stage model combining dual and slack models)

$$\text{Minimize } \theta - \varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+)$$

subject to

$$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = \theta x_{io} \qquad i = 1,\ldots,m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - s_i^+ = y_{ro} \qquad r = 1,\ldots,s$$

$$\lambda_j \geq 0 \qquad j = 1,\ldots,n$$

As observed, the objective function consists of 'ε' which is known as 'non-Archimedean', representing a small number. Model 5a allows for the calculation of optimal efficiency scores (θ) from the dual model (Model 3a), and then determines and optimizes slack values to obtain the efficiency frontier for the set of DMUs (Ozcan 2008; Cooper *et al.*, 2007).

The *target values* for inputs and outputs can be calculated by utilizing slack values (Ozcan 2008). It should be noted that Models 3a, 4a and 5a represent an input-oriented DEA model. Therefore, the target values can be calculated as shown below:

Target Values for inputs and outputs (Input-oriented model)		
Inputs	$\hat{x}_{io} = \theta^* x_{io} - s_i^{-*}$	$i = 1,\ldots,m$
Outputs	$\hat{y}_{ro} = y_{ro} + s_i^{+*}$	$r = 1,\ldots,s$

The target values play a key role in providing a guideline to decision-makers, by demonstrating the degree to which the inefficient DMUs need to decrease the inputs (since input-oriented model) in order to become efficient, by keeping the outputs at the same level.

The above DEA model demonstrates the *mathematical* representation of an input-oriented DEA model. The *output-oriented* DEA model can be developed on a similar pattern. In this case, the objective is to 'maximize' the output level for each DMU, while keeping the level of inputs constant, hence, modifying the mathematical formulation of the DEA model accordingly. The development of an output-oriented DEA model (Cooper *et al.*, 2007; Ozcan 2008) is shown below.

Output-oriented DEA Models
Model 1b (Fractional programming model)
Minimize $\phi = \dfrac{\sum_{i=1}^{m} v_i x_{io}}{\sum_{r=1}^{s} u_r y_{rj}}$ Subject to $\dfrac{\sum_{i=1}^{m} v_i x_{io}}{\sum_{r=1}^{s} u_r y_{rj}} \geq 1$ $u_r,\ v_i \geq 0 \qquad \forall\ r$ and i
Model 2b (Linear programming model)
Minimize $\phi = \sum_{i=1}^{m} v_i x_{io}$ Subject to $\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} \geq 0 \qquad j = 1,\ldots,n$ $\sum_{r=1}^{s} u_r y_{ro} = 1$ $u_r,\ v_i \geq 0 \qquad \forall\ r$ and i
Model 3b (Dual programming model)
Maximize ϕ Subject to $\sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{io} \qquad i = 1,\ldots,m$ $\sum_{j=1}^{n} \lambda_j y_{rj} \geq \phi y_{ro} \qquad r = 1,\ldots,s$ $\lambda_j \geq 0 \qquad j = 1,\ldots,n$

Model 4b (Slack model)
Maximize $\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+$ Subject to $\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io}$ $\quad i = 1,\ldots,m$ $\sum_{j=1}^{n} \lambda_j y_{rj} - s_i^+ = \phi^* y_{ro}$ $\quad r = 1,\ldots,s$ $\lambda_j \geq 0$ $\quad j = 1,\ldots,n$
Model 5b (Two-stage model combining dual and slack models)
Maximize $\phi + \varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+)$ Subject to $\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io}$ $\quad i = 1,\ldots,m$ $\sum_{j=1}^{n} \lambda_j y_{rj} - s_i^+ = \phi^* y_{ro}$ $\quad r = 1,\ldots,s$ $\lambda_j \geq 0$ $\quad j = 1,\ldots,n$

Target Values Using Slacks (Output-oriented)		
Inputs	$\hat{x}_{io} = x_{io} - s_i^{-*}$	$i = 1,\ldots,m$
Outputs	$\hat{y}_{ro} = \phi^* y_{ro} + s_i^{+*}$	$r = 1,\ldots,s$

The fractional programming model (Model 1b) has been changed to 'minimizing' the ratio of weighted sum of 'input to output', followed by DEA linear programming model (Model 2b). Two main differences can be found in Models 3b, 4b and 5b, and the target values when compared with the input-oriented model. The output efficiency is represented by ϕ and the objective is to maximize this efficiency level of each DMU. Furthermore, the efficiency level ϕ appears in the second constraint which is associated with the outputs rather than the input-related constraint. Similarly, for the target values, the efficiency level will be placed in the target value for outputs and not inputs, hence, the required increase in the output level can be determined for each DMU whilst keeping the inputs at the same level.

3.2 DEA Applications in Healthcare

The DEA literature is dominated by real-world applications, while articulating developments in its own theory base, and linking this base with that of several other disciplines (Gattoufi *et al.,* 2004). Liu *et al.,* (2013a), Emrouznejad *et al.,* (2008) and Seiford (1996) provide an extensive review on the evolution of DEA over the years, including theoretical refinements and extensions along with various applications, and provide an exhaustive bibliography of the DEA literature.

Since the early 1980s, efficiency analysis has been used to measure and analyse the productive performance of health care services (Hollingsworth 2008). Data Envelopment Analysis (DEA) has proven to be an effective and versatile tool for health care efficiency measurement that is, quantifying efficiency and productivity (Harrison *et al.,* 2004; O'Neill *et al.,* 2008). Some studies have provided a comprehensive survey of various studies that have employed DEA to conduct efficiency analysis of health institutions including Liu *et al.,* (2013b), Hollingsworth (2008) and Worthington (2004). However, Pelone *et al.,* (2015) conducted a review of previous works which have employed DEA to specifically assess primary care efficiency, and evaluated these studies in terms of the extent to which the results and information generated is useful for decision-makers.

3.2.1 DEA Applications in Hospitals and/or Other Healthcare areas

Considering efficiency assessment within healthcare DEA literature, most of the studies had the objective of comparing the efficiency of *hospitals* or other health centres.

Chang *et al.,* (2004) and Sheikhzada *et al.,* (2012) compared *private and public hospitals.* The former study concluded that private hospitals (including district and regional) are more efficient than the public hospitals, whereas the latter study identified that public hospitals are more efficient than their private counterparts. Chang *et al.,* (2004) further validated the initial efficiency scores using statistical as well as DEA-based tests. Furthermore, some additional characteristics were considered to get a better insight for the difference in efficiency scores, including competition intensity, severity of illness, rural/urban location of hospitals and teaching/non-teaching hospitals. However, Sheikhzada *et al.,* (2012) specifically evaluated factors which can improve the efficiency of inefficient units. They recommended the transfer of excess medical and non-medical personnel from inefficient to efficient hospitals, and sending excess administrative staff to early retirement. In case of unused beds, it was recommended to either transfer them to efficient hospitals, sell them or give them on contract to private clinics. The current study specifically emphasizes on developing a framework for a public hospital in a developing country to evaluate the queue system, which can be implemented in similar public hospitals in other developing countries.

Comparing the efficiency of *public hospitals* only, Kirigia *et al.,* (2007) concluded that nearly 50% of the hospitals were efficient over a period of four years. The authors mainly recommended to increase the demand for health services (including domiciliary care, antenatal/postnatal care, immunization and others) by adopting health promotion strategies and providing financial incentives (such as reduced transport costs, wait and treatment time). However, reducing nursing hours was not considered as a viable option due to its importance in providing primary care. In a similar study where more than 50% of public hospitals were deemed inefficient, Applainaidu *et al.,* (2014) recommended that excess medical personnel and beds can be transferred to primary care facilities which experience shortage of resources. However, they did recognize that transfer of resources from one healthcare unit to another might not be simple due to the involvement of administrative or social issues. These objectives are extremely hard to achieve and implement in developing countries, at least not in the short-run. Hence, with regard to queuing problem in large public hospitals of developing countries, there is a need to assess and improve the system 'as it is', and DEA modelling is extremely appropriate for this purpose.

Hollingsworth and Parkin (1995) and Tsai and Molinero (2002) conducted efficiency analysis for *acute-care hospitals* using DEA. The former study specifically compared the efficiency results from both CRS and VRS DEA models, and emphasized that the VRS model provides more useful information for decisions regarding resource allocation. Furthermore, the reliability of efficient peers was evaluated. It was concluded those units are better benchmarks which appear as peers more frequently and are efficient under both CRS and VRS models. However, the latter study utilized a CRS model with further analysis using a multi-activity VRS model for different specialities within hospitals. The

CRS model indicated more than 60% units as efficient. Under VRS, only one hospital trust was efficient in all its sub-activities, while others consisted of activities which were a mixture of efficient, IRS and DRS. The authors emphasized that this analysis guide the decision of dividing operating expenditures among different specialties.

Considering *different types* of *hospitals*, Harrison *et al.,* (2004) compared the efficiency of federal hospitals while in a recent study, Harrison *et al.,* (2015) compared 'for-profit' hospitals only. The inefficiency results accompanied by slack analysis were utilized to discuss managerial, resource allocation and manpower implications. Harrison *et al.,* (2004) suggested that information regarding operating expenses should be taken into consideration when deciding for federal funding process and personnel should be rotated among different facilities as well as more training should be provided to improve overall efficiency. Harrison *et al.,* (2015) also indicated that reduction in operating expenses and excess personnel can lead to short-term benefits and improved efficiency.

In addition to the DEA analysis, Huang *et al.,* (1989) assessed the relationship of the efficiency level of rural health care programs, and the organisation form and population size. The statistical tests (Chi-square test) showed that organized group practices had the highest number of efficient programs whereas community health centres had the lowest. Considering population size, almost all programs serving small populations were efficient, regardless of the type of organisation form. Furthermore, the logistic regression was applied which confirmed these results. Ramirez-Valdivia *et al.,* (2011) conducted a DEA analysis for primary health centres in both types of municipalities, urban and rural. In this case, four different model specifications were considered (CRS and VRS for both municipalities) to identify consistency of efficiency scores. The results indicated that primary healthcare centres in urban municipalities are more efficient than the rural ones. The authors concluded that improvement in resource allocation and utilization will allow for *appropriate budget allocation* especially for units in rural areas with low income. They also recommended that inclusion of variables regarding managerial issues and quality of health services should be considered to provide a more comprehensive analysis. However, Salinez-Jimenez and Smith (1996) specifically considered quality indicators as outputs in the DEA model when assessing the efficiency of family health services authorities (FHSAs). Seven different quality indicators were considered which were related to the services provided by general medical practitioners and nurses, percentage of patients who received a specific service and the minimum standards satisfied by the practice premises. Additionally, two environmental variables were considered, namely standardized illness ratio and unemployment. A model was also run without environmental variables and it was observed that some of the efficient units were deemed inefficient, which showed the sensitivity of selection of these variables in disadvantaged areas. The relationship between the expenditures (input) and quality outcomes (output) was evaluated using both input and output oriented models.

Figure 3-2 provides a summary of existing DEA literature in healthcare along with the objective of the current study. Primarily, maximum healthcare DEA studies compare the efficiency of hospitals or other health institutions. Mainly, these studies consider resources as inputs (such as personnel, beds and expenses) with number of treated patients (mainly outpatient and inpatient visits) as outputs. Besides, almost all studies recommend transferring or reducing resources and improving access to health services to increase the efficiency. These recommendations are *nearly* impossible to implement in developing countries due to demographic, social and healthcare delivery challenges. The current study aims to provide evidence that DEA modelling is highly suitable and has the ability to provide useful and accurate results; to improve a queue system in a busy public hospital with high level of variability due to absence of appointment systems. To the best of our knowledge, this study is the first of its kind where DEA methodology has been applied in the context of a *queuing problem*. Additionally, this study presents a *novel* application of DEA since it has been applied within one hospital using patient level queuing data as opposed to a comparison of different hospitals as seen in earlier studies (see Figure 3-3 below).

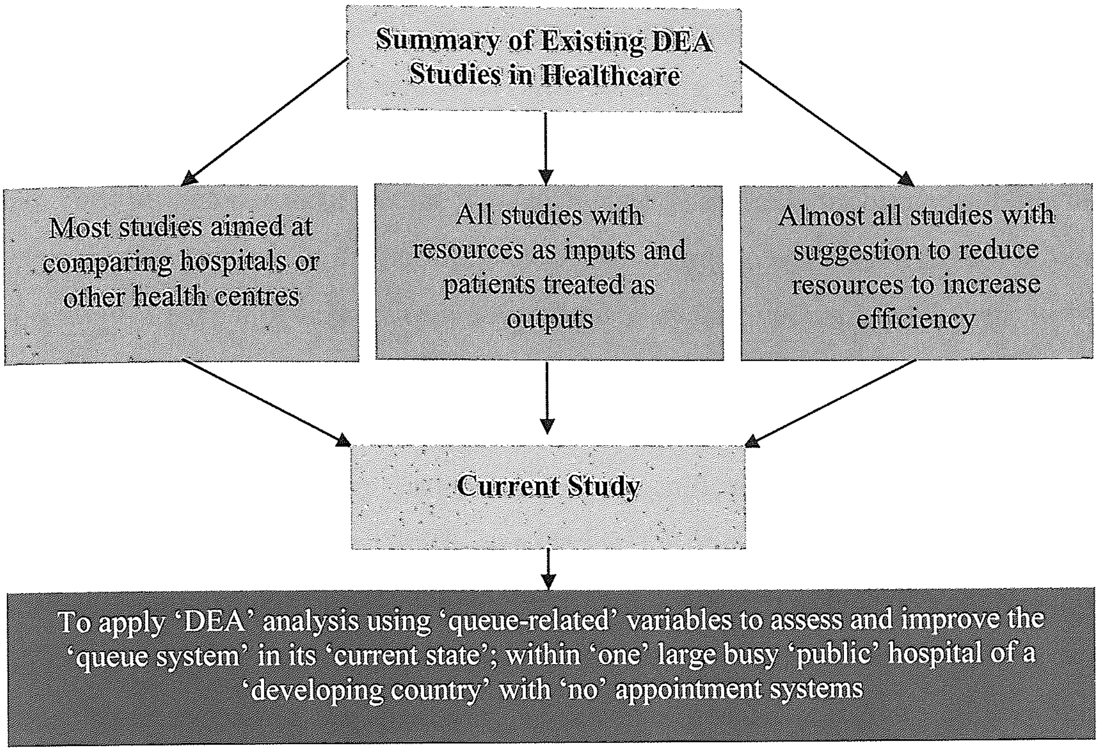


Figure 3-3: Summary of Most Common DEA Studies in Healthcare and the Objective of Current Study

3.2.2 DEA Research in Developed and Developing Countries

The healthcare DEA studies have been carried out in various different countries, including developed and developing countries. Spinks and Hollingsworth (2009) reviewed the usefulness of DEA efficiency results for policy implications within healthcare by conducting a cross-country comparison. Hadad *et al.,* (2013) evaluated the differences in the efficiency of health systems of Organization for Economic Cooperation and Development (OECD) countries, mainly in terms of type of DEA model selected and choice of input variables. Hollingsworth *et al.,* (1999) provided a review of application of DEA to hospitals and general health care for efficiency and productivity measurement in Europe and US. However, more research needs to be carried out to highlight the challenges and issues faced by developing nations.

Considering *developed countries*, many of these studies have been carried out in the UK and USA (as observed in queue management studies). With regard to *UK*, in addition to using DEA, some studies specifically evaluated the role of performance indicators (PIs) in providing information and assessing the efficiency of health services within the National Health Services (NHS). While assessing the performance of District Health

Authorities (DHAs) in terms of providing perinatal care, Thanassoulis *et al.,* (1996) concluded that PIs have the ability to provide easily understandable information to healthcare providers. Whereas DEA provides a better overall assessment of efficiency of units. They also emphasized that the two techniques when used jointly can lead to a more comprehensive analysis of the performance of units. Amado and Dyson (2008) provided a review of different efficiency techniques used to evaluate primary healthcare in NHS UK, including PIs, parametric techniques and DEA. The authors highlighted that PIs do not provide enough information on their own, and the parametric methods require prior definition of functional form which can be a significant drawback when assessing services, such as healthcare. When considering DEA, the authors concluded that in addition to outputs, 'outcomes' should also be included as part of the DEA analysis along with participation by various stakeholders, to get more useful and accurate results. Some studies have evaluated the efficiency of specific specialties within the healthcare institutions in UK. Tsai and Molinero (2002) conducted efficiency analysis to compare the performance of different specialties within UK hospital trusts, including medical, surgical, maternity, psychiatric and other specialities. The study also aimed at identifying the optimal allocation of budget among these specialties within inefficient units. Thanassoulis *et al.,* (2014) specifically assessed the efficiency and potential for cost savings among General Practitioner (GP) units within one specialty (geriatric and general medicine). The authors highlight the importance of GP service with regard to providing services to UK residents. The authors concluded that maximum possible savings can be achieved by decreasing the volume of drug prescriptions, followed by reduction in unit cost of prescriptions and inpatient unit costs.

Considering studies in *USA*, Nunamaker (1983) specifically compared the efficiency results from DEA and cost/patient day ratios, by assessing routine nursing services in non-profit hospitals in USA over a period of two years. The results obtained from the two methods were different as more than 50% of the units were inefficient in both years according to DEA, whereas none of the units were inefficient in the first year and less than 10% were inefficient in the second year for cost/patient day ratio. Additionally, the potential cost savings were much higher as indicated by DEA than cost/patient day model. In another study, Butler and Li (2005) employed DEA to compare the efficiency of rural hospitals in the US. The authors emphasized that information on returns to scale for the inefficient as well as the efficient units is crucial particularly with regard to optimal resource utilization. The authors concluded that more resources should be allocated appropriately to units which are inefficient as well as efficient under IRS, since both have potential to become fully efficient. Weng *et al.,* (2009) propose an extended DEA model that includes a new filtering benchmark process, named as Panel-based Benchmarking, to identify those hospitals which performed well, consistently, over a period of five years. The authors emphasized that this additional information is more useful to assist in decision-making regarding resource allocation (such as staff and equipment), and investments regarding training and process management for employees. In a recent study, Gholami *et al.,* (2015) discussed the increased pressure on hospitals, specifically in US, to provide services of increased quality as well as maintain efficiency given the current financial constraints. The main emphasis of this study was to evaluate the overall effect and usefulness of Information Technology (IT) investments on quality and operational efficiency of US hospitals, in order to influence decisions regarding allocation of funding. The results show that IT investments do lead to a higher service quality, but improves efficiency only up to a certain level.

Additionally, studies have been carried out in *other developed countries* as well. Blank and Van Hulst (2011) highlighted the importance and usefulness of corporate governance in the context of Dutch healthcare system, where the effect of various aspects of corporate governance on the cost efficiency scores was assessed. Among these factors, remuneration of supervisory board was significant under both CRS and VRS, remuneration of board and size of board were significant under CRS only, and external board members give a significant result under VRS only. Rouse *et al.,* (2011) provided an overview of the factors affecting the immunization activities in primary care practices in New Zealand. This study provided evidence of obtaining a better performance analysis if DEA and Activity-based costing (ABC) are used jointly. The authors identified that ABC can help in the selection as well as aggregation of suitable inputs for DEA, and DEA can provide a better insight on the information gathered from ABC. Furthermore, the authors also identified that practice efficiency is affected by additional administrative tasks such as GP involvement with immunisation advice, ordering vaccines and audit requirements. There is an increased level of administrative work and cumbersome procedures within public hospitals of some developing countries such as Pakistan, which further burdens scarce resources and leads to excessive wait times.

A few studies in developed countries included specific inputs/outputs to assess the efficiency of health institutions. Ouellette and Vierstrate (2004) reviewed a number of measures which were introduced to increase the efficiency in the Quebec health system in Canada, such as early physician and nursing personnel retirement program, hospital closures in the metropolitan region and reduction in hospital budgets. However, the authors argued that these measures did not lead to any increase in efficiency or cost savings, due to lack of proper management and incentives. Furthermore, quasi-fixed inputs are evaluated as part of DEA analysis. Four different DEA models are run, with and without quasi-fixed inputs, under both CRS and VRS. The authors highlighted the importance of incorporating fixed variables to provide more accurate assessment of inefficiencies with respect to inputs which are variable. Prior (2006) considered the level of infections as a quality indicator, and included it as an undesirable output when comparing Spanish small, medium and large hospitals over a period of three years. In this case, three different models were considered including one without this variable, and two other models where one introduced it as a weak disposable output and the other, a strong disposable input (different technological possibilities). The authors emphasized that quality attributes form a significant part of efficiency analysis, otherwise misspecification of models and unpredictable effects can distort results. In a recent study, Kittlesen *et al.,* (2015) conducted a cross-country comparison of acute-care public hospitals in four different developed countries, specifically Nordic countries including Sweden, Finland, Denmark and Norway. The authors highlighted that the health systems are similar with respect to public ownership and tax-based financing, however, there are differences in terms of administration of these hospitals and the incentives provided. The results showed that Finland had the highest efficiency and productivity among four countries considered. Additionally, the effect of some other variables on efficiency was evaluated. The results showed that efficiency has a positive and negative relationship with outpatient share and length of stay respectively, while no association was found with the status of being a university hospital or capital city hospital.

A few DEA studies have also been carried out in the *developing countries* in healthcare. Some of these studies have been carried out in *Africa*. Kirigia *et al.,* (2004) and Akazili *et al.,* (2008) compared the efficiency of health centres in Kenya and Ghana respectively. The studies provide an overview on the increasing need to improve the efficiency of health organizations in developing countries given the limited financial resources. The results from both studies showed that more than 50% of the health centres are inefficient, and highlighted the existence of excess inputs as identified by DEA analysis. The main recommendation was to transfer excess resources (beds and personnel) to more efficient facilities and utilize extra income to improve quality of services. However, decisions regarding increasing or decreasing resources, particularly human resources, seem impractical where developing countries are concerned due to various operational challenges and social barriers. One of the main objectives of this research study is to develop a framework which monitors the queue system in its existing state, avoiding such recommendations which might not be implementable in near future. Less than 15% of the hospitals were found efficient in a study by Zere *et al.,* (2001) in South Africa. The authors also carried out a second-stage analysis to evaluate the effect of a number of factors on the inefficiency level, including bed occupancy rate, average length of stay, outpatient visits as a proportion of inpatient days and location (three provinces were considered). The results showed that higher occupancy rate and outpatient visits lead to higher efficiency and are significant, whereas average length stay and location were not statistically significant. The authors emphasized that potential savings through efficiency gains are much higher than earnings through other financial sources such as revenues generated from user fees. Recently, Bwana *et al.,* (2015) carried out a study in Tanzania and compared private not-for-profit hospitals over a period of four years. The authors highlighted that the results from this study can provide useful information to concerned stakeholders including administrators, governing boards, owners and healthcare policy makers. The authors concluded that there is no improvement in the efficiency of hospitals over four years, with less than 30% of units being efficient overall, which shows the increase in inefficiency especially with respect to managing resources. They suggested that more resources should be allocated to units operating under IRS to further increase the output level and the efficiency level, and the government and hospital owners need to work in collaboration to improve the overall efficiency.

Considering the *developing countries in Europe*, Ersoy *et al.,* (1997) aimed at identifying the inputs and outputs which contribute to the inefficiency of acute general hospitals in Turkey. The authors highlighted the increase in health care expenditures in Turkey, and emphasized on increased importance of assessing the efficiency to justify health care costs. The results demonstrated that less than 10% of the hospitals are efficient only, and concluded that large number of personnel (specialist and primary care physicians) with double the amount of number of beds were being utilized by inefficient units. In a recent study, Mitrovic *et al.,* (2016) carried out a cross-country comparison to compare the efficiency of health system in Serbia with other countries. The authors highlighted the challenge faced by the Serbian health system to increase efficiency in order to justify increased government expenditure spending on healthcare, and rise in demand due to increase in ageing population. The DEA model with undesirable output was utilized since the output was mortality rate, with expenditures and human resources as inputs. Serbia had a difference of more than 30% with the most efficient units, and ranked 15th out of 21. The authors concluded that low level of efficiency might be the result of lack of reforms, deteriorating health status of population, poor funding and lack of long-term planning. They recommended further analysis to identify reasons for increased inefficiency.

A few DEA studies have been carried out in the developing nations of *Asia* as well. Al-Shammari (1999) and Shahhoseini *et al.,* (2011) assessed the efficiency of hospitals in Jordan and Iran respectively. The main emphasis of these studies is to provide empirical evidence and information regarding the current inefficiency within health services, to provide guidance for better utilization of resources. Al-Shammari (1999) highlighted the need to develop a system where information regarding the resource allocation is passed on regularly to policy makers and can be used to prepare broad guidelines for the management of these hospitals. Shahhoseini *et al.,* (2011) emphasized on the transfer of excess personnel to efficient ones, or re-consider using these inputs to increase the current level of output. Ramanathan (2005) compared the efficiency of public hospitals in Oman using DEA. The study provides an overview of the health system in Oman, and mentioned the different types of hospitals responsible for providing primary, secondary and tertiary care. The results indicate that nearly 50% of the hospitals are efficient. The author concluded that hospitals operating under IRS should be given preference in terms of increasing allocation of resources, investment and expansion. Dutta *et al.,* (2014) identified causes of inefficiency within government hospitals in India. They provide an elaborate overview of the demographic and health-related challenges faced by deprived areas in the context of India, including percentage of people living below poverty line, high population density, large rural population, shortage of resources, and disparity of health indicators among rural and urban areas. The average efficiency score for inefficient units was 60% and highlighted excess staff, however, the authors recommended close supervision but not reducing staff due to high requirement per bed. The second-stage analysis shows a requirement of reduction in average length of stay, better distribution of medicines, increasing outpatient days, and availability of all types of medical personnel (junior doctors, trainees, specialists and nurses). Ramanathan (2005) carried out a study in the Middle-Eastern part of Asia, whereas Dutta *et al.,* (2014) conducted efficiency analysis in the South-East region of Asia. However, only a very limited number of efficiency assessment studies have been conducted in Asia, with almost negligible in South-East Asia. To the best of our knowledge, the current study is the first one to conduct efficiency analysis in terms of queuing problem in Pakistan, and among the first few negligible studies in the South-east Asian region in general.

Numerous studies have been carried out in the developed world with regard to health system efficiency using DEA. Additionally, a few studies have been conducted in the context of developing countries as well, and most of these works have reported that at least half of the hospitals considered are inefficient. This illustrates the crucial challenges and operational inefficiencies experienced by the health delivery systems and the suffering of citizens within these countries. Almost all previous works in developing and developed countries have focused on increasing efficiency with respect to better resource allocation, mainly by reducing medical personnel and spending. However, there is still *lack* of *objective* solutions especially in *developing* countries, as those proposed in previous studies are unlikely to be implemented in the short-run. None of the previous works have evaluated a queue problem using efficiency analysis. Additionally, with regard to efficiency assessment in healthcare, there is a dearth of literature in developing countries, particularly the South-East Asian region where research is almost negligible. The current study aims at constructing a dynamic framework to assess and improve the existing queue system of a large public hospital specifically, in Pakistan, with the objective of generalizing results for other similar developing countries. Additionally, the research study also focuses on analysing the level of modification required in the DEA model for queue assessment (see Figure 3-4 below).

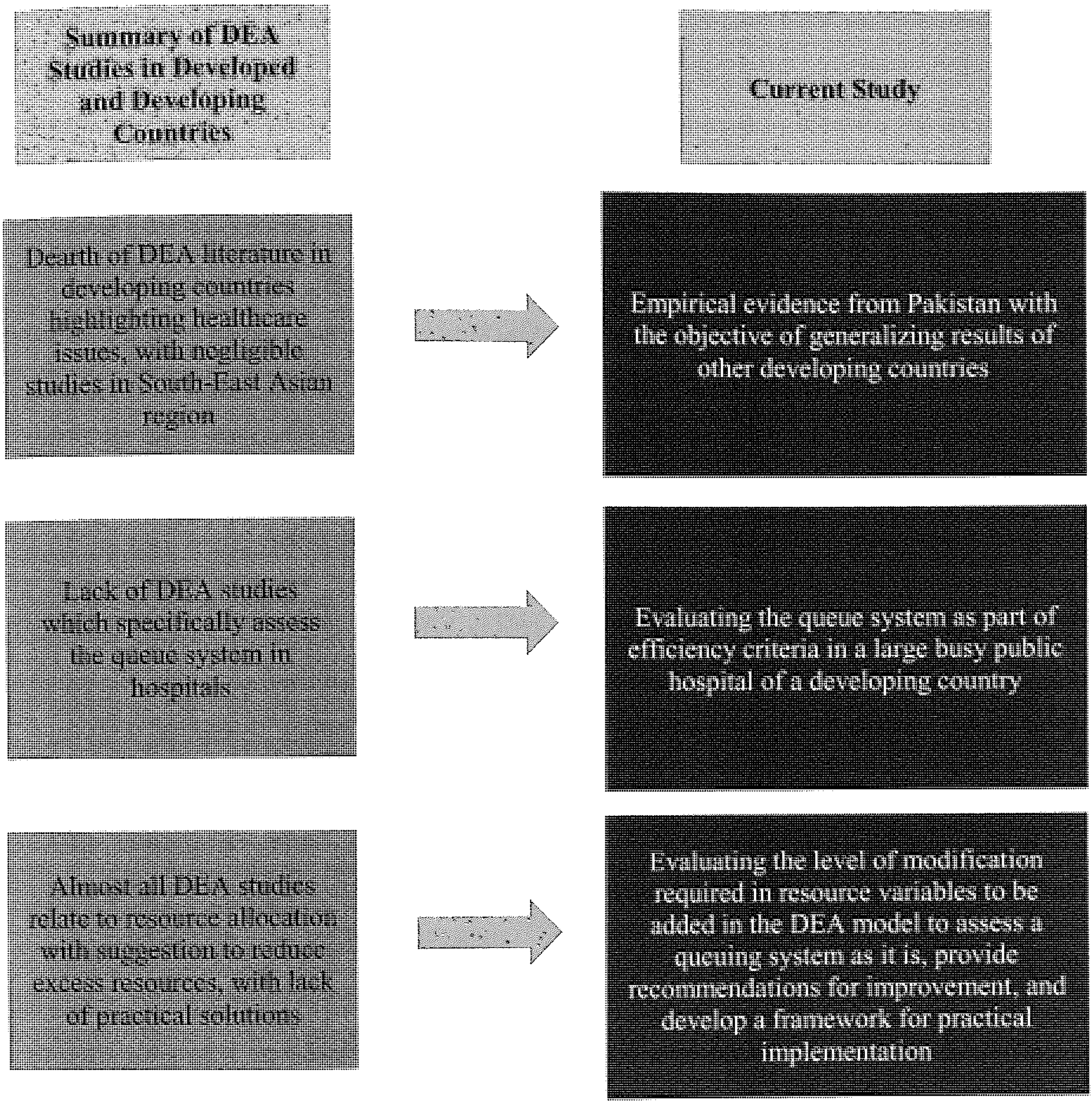


Figure 3-4: Summary of DEA Studies in Developed and Developing Countries

3.2.3 Use of DEA Modelling Combined with Other Techniques

In addition to the traditional DEA model, a number of healthcare studies have used either different types of DEA models or a combination of DEA with other techniques; to provide a more in-depth performance analysis of health institutions.

A few studies have utilized *different types* of DEA models. Chuang *et al.,* (2011) identified the effect on efficiency scores by contrasting different DEA models including the traditional model, DEA artificial neural network (ANN) model and DEA-assurance region (AR) model. The results showed DEA-ANN model to be superior to the other two, as it provided evidence of increased discriminatory power with a small number of efficient units. Also, it had the highest average efficiency score with the lowest standard deviation, leading to a more robust efficiency evaluation. Additionally, the study also utilized the classification and regression tree (CART) method to identify the exact level of inputs and outputs to improve efficiency, in terms of better resource allocation and reducing operating costs. Chilingerian and Sherman (1997) also employed cone-ratio DEA model, such that qualitative managerial information was incorporated to assess efficiency. The study particularly considered the practice styles of physicians as an additional variable. The results showed that at least 5% of the efficient units became inefficient when cone-ratio model was utilized, with nearly 50% decrease in the number of general medicine physicians. The authors concluded that practice style of this specialty requires close supervision to increase efficiency.

Flokou *et al* (2011) used *post-DEA* analytical techniques such as cluster analysis and Maverick index to further validate the results obtained from the first-stage DEA modelling. Maverick index was utilized to alert decision-makers regarding those units which have been performing consistently well. Cluster analysis allowed for setting more realistic targets for inefficient units through gradual benchmarking, rather than high unachievable targets in one step. Parkin and Hollingsworth (1997) discussed the issue of validity within DEA analytical results. Considering internal validity, the authors evaluated the efficiency scores by employing four different DEA models, with different combinations of inputs and outputs. The first model consisted of all inputs and outputs, whereas the second and third model included one output and input as an aggregated variable measure respectively, and the fourth model eliminated one of the input variables due to the possibility of high association with another input. The authors considered the last model as suitable, since it provided a certain level of discriminatory power, while maintaining the level of disaggregation among inputs and outputs.

A number of studies have considered *productivity assessment* in addition to efficiency evaluation, mainly to identify the actual source of progress or regress over a period of time. Hence, Malmquist Productivity Index (MPI) has been considered as part of the DEA analysis in many studies, mostly to assess the efficiency and productivity change over time. Kirigia *et al.,* (2008) evaluated the efficiency and productivity of public hospitals over a period of three years. The results showed that overall productivity increased by nearly 5%, on average. The authors concluded that this improvement was due to increase in efficiency rather than innovation (or technological change), since overall on average, efficiency increased by more than 10% while technological change regressed by nearly 7%. However, in a study by Ramanathan (2005), the results showed that the average MPI has declined over a period of two years by nearly 10%, with decrease in efficiency (nearly 4%) being less than the technical regress (nearly 7%). In a recent study, Ineveld *et al.,* (2015) evaluated the effect of health system reforms by assessing the change in efficiency and productivity over a period of five years. During the five-year period 2005-2010, slight overall improvements in efficiency and productivity were observed, except for the period 2009-2010. Hence, the increased overall efficiency in 2009 can be explained by technological improvement in efficient hospitals by 5%. Whereas the decline in overall efficiency in 2010 is a result of decrease in the efficiency level of inefficient hospitals, despite increased technological improvement by 1-2%. The authors concluded that the health care reform led to improved efficiency of small hospitals while the efficiency of large hospitals decreased.

Some studies have carried out a *second-stage analysis* specifically by utilizing regression models mainly to assess the effect of various variables on DEA inefficiency scores.In a study by Ramirez-Valdivia *et al.,* (2011), Biplot method was first employed to short-list the number of variables to be included in the Tobit regression model to assess the effect on inefficiency scores of primary health centres in rural and urban municipalities. Two of the variables including proportion of urban health centres and annual referral rate to specialists were significant for both rural and urban municipalities. Some additional significant variables of urban municipalities included proportion of PCs online, annual preventive medical exams and enrolled inhabitant less than 6 as well as population density. These variables were used to get a better insight of the factors leading to inefficiency in both urban and rural areas. In another study, Jehu-Appiah *et al.,* (2014) specifically considered ownership to assess the impact on efficiency scores of hospitals using a Tobit regression model. The different types of hospitals included government, mission, quasi-government and private. The results demonstrated that private for-profit hospitals had the highest (significant) impact on the level of inefficiency. The authors concluded that private health institutions need to increase their outputs by nearly three times to increase the efficiency level. They recommended that this objective might be achieved by creating additional demand for services. The authors also suggested that further analysis using decision-making and market dynamics can assist in developing reforms to assess and improve overall performance. Aksezer (2016) developed a two-stage model where DEA was carried out in the first stage, followed by optimal allocation model for resource allocation given budgetary constraints and to maximize network reliability.

Some *other techniques* have also been used in addition to DEA analysis, particularly with respect to validate the DEA results obtained. Chang *et al.,* (2004) employed statistical tests including Welch's two sample means and Wilcoxon two sample tests, to assess the differences in the efficiency scores for public and private hospitals. Additionally, DEA-based tests were utilized to validate the results obtained from statistical tests, which supported prior results that private hospitals are more efficient than the public, for both regional and district hospitals. Butler and Li (2005) highlighted the issue of identifying the returns to scale not only for efficient but also for inefficient units in order to assist in decisions regarding resource allocation. In order to validate the proposed methodology, discriminant analysis was utilized which proved that the extended DEA model is a valid approach for classifying inefficient units according to returns to scale. Weng *et al.,* (2009) utilized window analysis with different time periods, to assess temporal changes in DEA efficiency scores. The authors concluded that small and more windows reflect high discriminatory power since the number of benchmarks is low, providing a better insight and hence, more reliable benchmarks.

Some studies specifically *compared* DEA with other efficiency assessment techniques. Two of the most popular techniques include the parametric technique of Stochastic Frontier Analysis (SFA) and Ratio analysis. Kontodimopoulos *et al.,* (2011) utilized Tobit regression models to evaluate the difference in efficiency scores obtained from DEA and SFA, using a few environmental variables including ownership, location, years in operation and size of facility. The regression analysis indicated that private ownership had significant negative whereas location was not significant for both DEA and SFA efficiency scores. Operating for more than 12 years had a significant negative impact on DEA scores, with no effect on SFA scores. Increasing the size of facility had a significant but contradicting impact on SFA and DEA, which was negative and positive respectively. Linna and Harkenen (1999) compared the efficiency of acute-care hospitals using DEA and SFA to investigate factors affecting cost efficiency. The range of overall cost efficiency was narrower for SFA scores (between 84-86%) as compared to DEA scores (between 84-92%). Furthermore, SFA and DEA scores were regressed on a number of variables to evaluate the differences including degree of specialisation, use of modern technology, input allocations, quality control, scale of operations and patient transfers to other facilities. The results indicated that high level of specialisation and resident physicians led to high efficiency scores for both DEA and SFA. The authors concluded that the both DEA and SFA highlight the same determinants of efficiency level, even if some differences were observed in individual efficiency scores. In a recent study, Kittlesen *et al.,* (2015) compared the efficiency and productivity results from both techniques DEA and SFA, by conducting a cross-country comparison among four countries including Finland, Norway, Denmark and Sweden. For both techniques, the results indicated that hospitals in Finland are consistently most efficient and productive. The authors recommended including variables which truly reflect variation across hospitals rather than just factors at country level, or a comparison across various time periods to provide a comprehensive analysis.

Considering ratio analysis, Huang *et al.,* (1989) compared DEA rankings with the ranking provided by cost-related ratios, including average cost, average medical cost and self-sufficiency ratio. The authors concluded that the results obtained from both approaches are similar. Thanassoulis *et al.,* (1996) specifically highlighted the differences in results obtained from DEA and PIs, as well as the strengths and weaknesses of each technique. The authors evaluated the two techniques in terms of the reliability of efficiency measures and the target values obtained, and concluded that DEA provides results which reflect the overall performance of units, whereas PIs provide information on factor-specific performance. Furthermore, DEA provides guidance for the target values of inefficient units whereas PIs can provide some useful suggestions to set targets for efficient units. Hollingsworth and Parkin (1995) consider Efficiency index (EIs) in addition to DEA and PIs. They provide a discussion on all three assessment techniques particularly with respect to understanding the results and reliability of these results. The authors highlighted that PIs and EIs measure changes from year to year, rather than one point in time. Also, PIs consider numerous variables making it difficult to assess DMUs on a common criteria. Besides, some EIs require complex formulae and calculations. Hence, the authors concluded that in comparison, DEA has strong theoretical underpinnings and provides easily understandable efficiency results at one point in time with target values. The current study intends to use DEA modelling in order to provide results and guidelines for improvement using targets, such that these results are easily comprehensible by hospital administrators.

The DEA modelling technique has been used in *combination* with other methods in some studies, most commonly to validate DEA results, compare efficiency scores obtained from DEA and other techniques or to provide further analysis. Besides, extended DEA models have also been used to provide a more robust efficiency analysis. The objective of the current research is to develop a DEA model such that it accurately represents the *queuing system* within a busy public hospital in a developing country with no prior appointments, with the purpose of providing robust and useful results (see Figure 3-5 below).

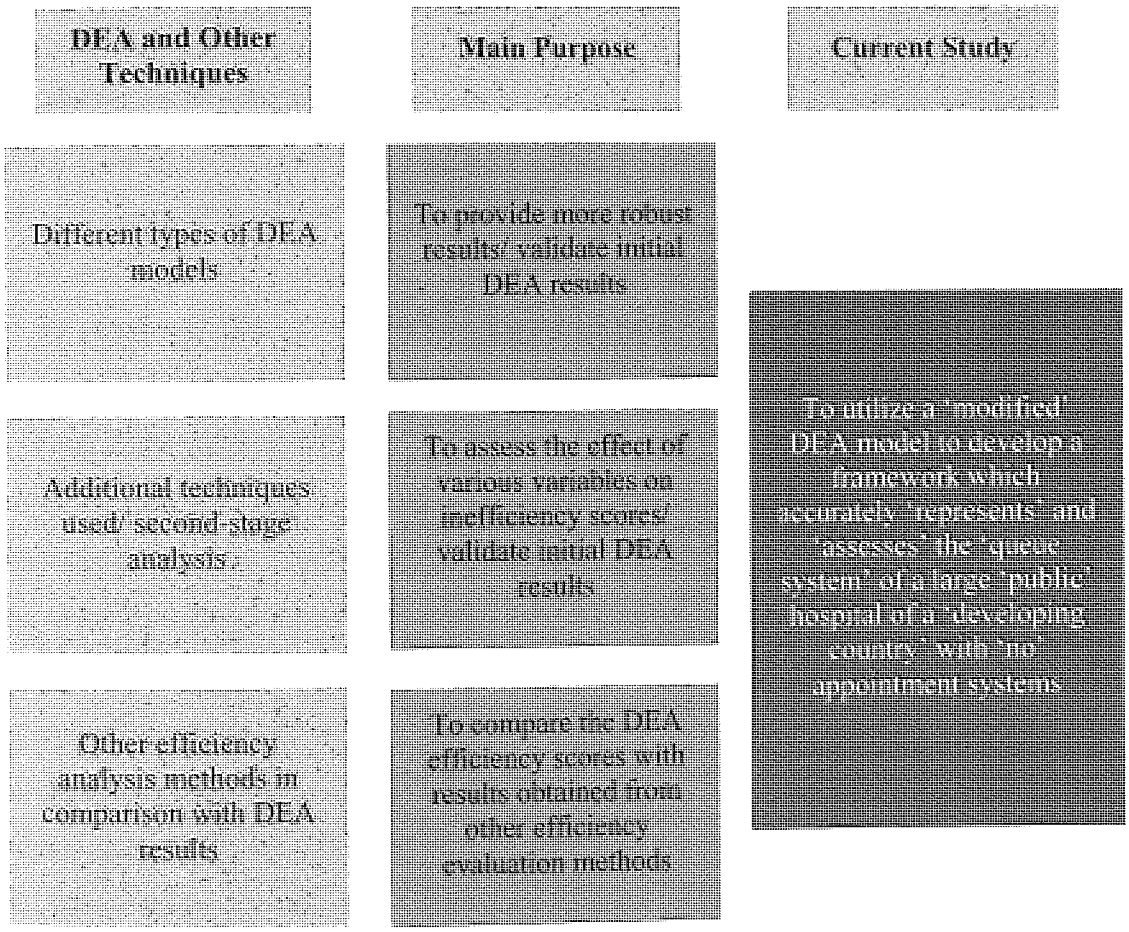


Figure 3-5: Summary of DEA in Combination with Other Techniques

3.3 Inputs and Outputs Used in DEA Healthcare Applications

Inclusion of inputs and outputs does not only affect the results but also the ability of a particular methodology to provide useful information (Hollingsworth and Parkin 1995). Selection of relevant inputs and outputs is one of the most important considerations in DEA modelling (Harrison *et al.,* 2004), where the actual selection is greatly dependent on the objective and expected results of the particular study. Some inputs and outputs, and their combinations, have been more frequently used in previous efficiency measurement studies in healthcare.

3.3.1 Outputs in DEA Healthcare Applications

Two of the most frequently used outputs in the healthcare literature are the outpatients' visits and inpatients' visits. Considering the former, a number of studies incorporated *outpatients' visits* including Chang *et al.,* (2004), Ramanathan (2005), Kirigia *et al.,* (2004), Akazili *et al.,* (2008), Prior (2006), Flokou *et al.,* (2011) and Bwana (2015). Zere *et al.,* (2001) considered outpatient visits as a proportion of inpatient days, and concluded that increasing outpatient activity in hospitals rather than constructing new clinics will lead to improved efficiency. When comparing the two outputs, outpatients' visits and inpatient admissions, Harrison *et al.,* (2004) concluded that the increase in outpatient visits was more than the inpatient admissions, due to increasing emphasis on ambulatory care in the past few years. However, overall shortage in outpatient visits was observed over a period of four years. Kirigia *et al.,* (2008) identified that the inefficient hospitals need to increase their outpatients' visits by at least nearly 50%. Tsai and Molinero (2002) considered outpatients' visits for different specialties as separate outputs including medical, surgical, psychiatric, maternity, and others. Puig-Junoy (2000) considered day-case patients and ambulatory visits as proxies to outpatient visits. The study utilized slack analysis to identify the level of outputs required. No major savings were observed in ambulatory visits, however, there is scope of improvement in the number of day-care services for inefficient hospitals. Chuang *et al.,* (2011) considered outpatient and emergency visits collectively, whereas, Sheikhzada *et al.,* (2012) included outpatient and emergency visits as two separate outputs. Butler and Li (2005) considered scheduled outpatient visits as well as unscheduled emergency room visits as two outputs, and evaluated the efficiency of hospitals using returns to scale analysis. However, none of the

DEA studies have considered unscheduled outpatient visits. The current study aims to assess the queue system of outpatients using DEA with random arrival of patients.

Considering the latter variable, a number of studies have included *inpatients' visits* in terms of number of inpatients (Ramanathan 2005; Lindlbauer *et al.,* 2016), inpatient days (Chuang *et al.,* 2011; Weng *et al.,* 2009; Bwana 2015) or inpatient admissions (Flokou *et al.,* 2011; Kirigia *et al.,* 2008; Harrison *et al.,* 2004; Dotoli *et al.,* 2015). Chang *et al.,* (2004) utilized inpatient days as a proxy to number of cases, due to limited data availability and the fee system which is determined according to per visit rather than per case. Furthermore, this single output variable consisted of three different types of patient days including general care, acute and intensive care, and chronic care. However, Puig-Junoy (2000) considered different types of inpatient days as separate output variables including acute and sub-acute patients in different specialties (medical, surgical, gynaecological, obstretics and paediatric), intensive care (neonatal and burn cases) and long-term care (psychiatric, long stay and tuberculosis). Slack analysis showed that there is scope for significant savings for intensive-care units, whereas no major savings were observed in acute-care and long-term care patient days. Tsai and Molinero (2002) considered inpatient days for five different categories of specialties including medical, surgical, psychiatric, maternity and all other specialties. However, Nunamaker (1983) formulated different DEA models considering separate as well as combined outputs for various specialties. For instance, total routine aged and paediatric days were considered jointly as well as separately in different models, along with routine maternity days and other routine days as the other two outputs in both models. Al-Shammari (1999) considered the 'patient days' with respect to inpatients only, where outpatient and emergency visits were not included in the DEA model.

Another significant output variable included in previous works is the number of *Surgeries* performed. Puig-Junoy (2000) identified from the DEA analysis that major potential savings can be generated from surgical interventions. Flokou *et al.,* (2011) considered inpatient treatment which included admissions as well as surgeries. The results were used specifically to identify those units which have been performing consistently well, in order to caution decision-makers regarding units with inconsistent performance. Dotoli *et al.,* (2015) considered total number of surgeries as an output, and used fuzzy logic framework to deal with uncertainty in input and output data. Butler and Li (2005) emphasized that surgeries is a significant variable to be included in the DEA model, as it utilizes different combinations of inputs including labour and capital. In this case, both inpatient and outpatient surgeries were included. However, Kose *et al.,* (2014) only considered surgical operations. Furthermore, a few studies have considered major and minor surgeries as two separate output variables (Al-Shammari 1999; Ramanathan 2005; Weng *et al.,* 2009). Some other variables have also been considered in previous works. For instance, a few studies used the total number of *Discharges* as an output variable. Dotoli *et al.,* (2015) and Puig-Junoy (2000) included total number of discharges as a single output. However, Blank and Van Hurst (2011) considered discharges from three different groups of patients as separate outputs, where these groups were generated according to the average length of stay and whether the patient was treated by a surgical specialty or not. When assessing medical and surgical services, Kose *et al.,* (2014) used number of discharged inpatients for the former while considered discharges as well as surgeries for the latter. A few studies have included *reverse outputs* in DEA models, where the objective is to reduce them. Lewis *et al.,* (2011) considered complication rate, whereas, Prior (2006) included infections (the level of nosocomial infections) as an 'undesirable' output. A few studies considered *antenatal visits* as an output. Kirigia *et al.,* (2008) considered antenatal visits as part of outpatient visits, whereas Kirigia *et al.,* (2004) and Akazili *et al.,* (2008) included them as two separate outputs, along with family planning visits and immunizations. However, Kirigia *et al.,* (2007) aggregated a number of services into a single output variable including pap-smear visits, family planning visits, maternal and child health visits, antenatal and postnatal visits, and number of children immunized. Mitrovic *et al.,* (2016) considered infant deaths and neonatal deaths (per 1000 live births) as two of the outputs, when conducting comparison of health system in different countries. Salinez-Jaminez and Smith (1996) considered proxies to *quality indicators* related to staff availability (GPs, nurses and front-line staff) and process variables (immunization, cervical smears etc), as separate outputs.

Considering some *additional output variables*, Huang *et al.,* (1989) considered visits to each of the staff members such as physician, nurse and health practitioner, as separate outputs as well as an additional output of total encounters consisting of all visits. Chuang *et al.,* (2011) considered time spent by a medical professional using expensive medical equipment as an output. Ramirez-Valdivia *et al.,* (2011) differentiated between annual number of medical visits and medical check-up visits as two separate outputs, as related to sickness and preventive care respectively. Rouse *et al.,* (2011) had a single output of number of vaccinations. Harrison *et al.,* (2004) considered number of services offered as a proxy output for capital assets. When evaluating the administrative department, Kawaguchi *et al.,* (2014) included medical income as an output variable. On the other hand, Kang *et al.,* (2014) assessed the efficiency of emergency departments only, and included two outputs of average length of stay and rate of leaving without treatment. Blank and Van Hurst (2011) included first-time visits as an output, that is, the number of patients treated by physicians without admission. The current study emphasizes on evaluating the queuing of patients including first-time and follow-up visits, but without prior appointments and random arrival pattern which results in excessive wait times.

A summary of most commonly used outputs in DEA healthcare studies is given in Table 3-1.

Table 3-1: Most Common Outputs Used in Healthcare DEA Studies

Outputs	**Description of Most Commonly Used Outputs**
Outpatients' Visits	Number of Outpatients' Visits *(Chang et al., (2004); Ramanathan (2005); Harrison et al., (2004); Kirigia et al., (2004); Akazili et al (2008); Prior (2006); Kirigia et al., (2008); Flokou et al., (2011); Bwana et al., (2015))*
	Outpatients' Visits for different Specialties *(Tsai and Molinero (2002))*
	Outpatients' Visits with emergency Visits *(Butler and Li (2005); Chuang et al., (2011); Sheikhzada et al., (2012))*
	Outpatient Visits as Proportion of inpatient days *(Zere et al., (2001))*
Inpatients' Visits	Number of Inpatient days *(Al-Shammari (1999); Chang et al., (2004); Chuang et al., (2011); Weng et al., (2009); Bwana et al., (2015))*
	Number of Inpatients *(Ramanathan 2005; Lindlbauer et al., (2016))*
	Number of Inpatient days for different types of patients/Specialities *(Nunamaker (1983); Puig-Junoy (2000); Tsai and Molinero (2002))*
	Number of Inpatients' admissions

	(*Harrison et al., (2004); Kirigia et al., (2008); Flokou et al., (2011); Dotoli et al., (2015)*)
Surgeries	Total Number of Surgeries (*Puig-Junoy (2000); Flokou et al., (2011); Dotoli et al., (2015)*)
	Major and Minor Surgeries (*Al-Shammari (1999); Ramanathan (2005); Weng et al., (2009)*)
	Inpatient and Outpatient Surgeries (*Butler and Li (2005)*)
	surgical Operations only (*Kose et al., (2014)*)
Discharges	Total number of Discharges (*Puig-Junoy (2000); Dotoli et al., (2015)*)
	Number of Discharges for different types of patients (*Blank and Van Hurst (2011)*)
Some Other Outputs	Undesirable Outputs Level of Infections (*Prior (2006)*) Complication Rate (*Lewis et al., (2011)*)
	Antenatal Visits (*Kirigia et al., (2004); Akazili et al., (2008); Kirigia et al., (2008)*)
	Proxies to Quality Indicators (*Salinez-Jaminez and Smith (1996)*)
	First-time Visits (*Blank and Van Hurst (2011)*)
	Medical Income (*Kawaguchi et al., (2014)*)
	Average Length of Stay/Rate of leaving without treatment (*Kang et al., (2014)*)
	Number of Vaccinations (*Rouse et al., (2011)*)

3.3.2 Inputs in DEA Healthcare Applications

Among various inputs considered in DEA modelling in healthcare, number of *Beds* is one of the most frequently used variables as demonstrated in a number of studies including Sheikhzada *et al.,* (2012), Chuang *et al.,* (2011), Akazili *et al.,* (2008), Harrison *et al.,* (2004), Prior (2006), Weng *et al.,* (2009), Bwana *et al.,* (2015), Dotoli *et al.,* (2015) and Lindlbauer *et al.,* (2016). Some studies have specifically considered the number of beds as a proxy to capital assets (Flokou *et al.,* (2011); Butler and Li (2005); PuigJunoy (2000)). Zere *et al.,* (2001) compared three different groups of hospital with respect to size, and included number of beds as one of the inputs. For all three groups, the results showed that the number of beds are in excess and need to be reduced by at least 30%. However, when comparing public hospitals over a period of three years, Kirigia *et al.,* (2008) identified that the number of beds remained fairly constant over the period considered, with a required reduction of only 8%. The bed size has also been included as a significant variable in different models for comparison purposes. For instance, Chang *et al.,* (2004) included number of beds as an input variable when comparing public with private hospitals. When conducting DEA analysis of surgical and medical services, Kose *et al.,* (2014) included the number of beds as input in both models considering it as a significant variable for both types of services. Kawaguchi *et al.,* (2014) developed a 'network' DEA model where the objective was to assess the efficiency of two divisions of different hospitals, including administrative and medical services. These two divisions were assessed jointly by evaluating the transfer of resources from one to the other. In this case, the number of beds was included as a 'link' variable between these two divisions, such that it was considered as an output of the administrative process (generated through income) but an input for the provision of medical services (utilized to provide service).

The number of *Physicians* is another commonly used input as shown in various studies including Chuang *et al.,* (2011), Harrison *et al.,* (2004), Prior (2006), Weng *et al.,* (2009), Bwana (2015) and Dotoli *et al.,* (2015). When including physicians as an input, Puig-Junoy (2000) considered residents as part of the total number. Ramanathan (2005) considered different categories of physicians including administrators, specialists, consultants and dentists. When assessing the efficiency of private and public hospitals, Sheikhzada *et al.,* (2012) considered specialist and general physicians as two separate inputs. Kirigia *et al.,* (2007) included the total number of hours worked by the physicians rather than the actual number of physicians. Kose *et al.,* (2014) considered the number of physicians as an input in both models assessing the efficiency of surgical and medical services. The results demonstrated that number of physicians are in excess for both services.

In addition to physicians, previous works have also considered *other* medical and nonmedical *personnel* when evaluating the efficiency of healthcare institutions. In some cases, the number of physicians working are considered jointly with other staff members. Huang *et al.,* (1989), Chuang *et al.,* (2011) and Al-Shammari (1999) considered *medical* personnel only and included physicians, nurses and 'other' medical personnel as three separate inputs. However, rather than considering *Nurses* as a separate input, Ramanathan (2005) and Prior (2006) divided the personnel as two inputs including doctors and 'other' medical personnel where nurses were included in the latter category along with others. Weng *et al.,* (2009) and Dotoli *et al.,* (2015) also considered *non-medical* staff with medical staff as one variable. Mitrovic *et al.,* (2016) considered midwives and nurses as one input with physicians as a separate input. On the other hand, Sheikhzada *et al.,* (2012) included three inputs including clinical as well as non-clinical staff, with different combinations. The first included specialist doctors only, the second included general physicians, nurses and residents whereas the third included number of non-medical and support staff. Using a different combination, Puig-Junoy (2000) considered three inputs, where physicians and residents were considered jointly, with nurses and nonmedical staff as two additional inputs. Among these three inputs, maximum potential for improvement was observed in the number of nurses. Blank and Van Hurst (2011) included four different types of personnel as four separate inputs consisting of staff and administrative, nursing, paramedical and 'others.' Lindlbauer *et al.,* (2016) considered two separate inputs for non-medical staff including non-clinical staff and administrative staff. Akazili *et al.,* (2008) considered two general inputs including 'clinical' and 'non-clinical' staff members.

Kirigia *et al.,* (2008) did not consider other staff members, and included physicians and nurses only, and that as one input. Butler and Li (2005) and Bwana *et al.,* (2015) considered the total number of employees as one input related to all personnel, rather than dividing them into physicians or others. Kang *et al.,* (2014) used DEA to assess the efficiency of emergency departments in various hospitals. In this study, working hours for clinical and non-clinical staff were considered as two separate inputs rather than the actual number. Ouelette and Vierstrate (2004) considered variable as well as quasi-fixed inputs for DEA analysis. In this case, the number of hours worked by personnel excluding physicians was included as variable input, whereas number of full-time physicians was included as a quasi-fixed input. Kawaguchi *et al.,* (2014) analysed the administrative and medical services divisions separately for different hospitals. Therefore, the first division included number of non-medical staff (administrative and maintenance staff), whereas for the second division medical staff was considered, including doctors, nurses, assistant nurses and medical technicians, as four separate inputs. All DEA studies have included physicians and other medical/non-medical personnel as inputs to assess the efficiency of health-care institutions. The current study aims to evaluate the existing queue system, taking into consideration the present availability of personnel.

Some studies have utilized *other input variables* for DEA analysis. The total operating *costs* or expenditures have been utilized as input in various studies including Zere *et al.,* (2001), Butler and Li (2005), Chuang *et al.,* (2011), Harrison *et al.,* (2004), Kawaguchi *et al.,* (2014) and Mitrovic *et al.,* (2016). Tsai and Molinero (2002) considered a single input, total operating expenditures, with ten different outputs related to different specialties with regard to inpatients' and outpatients' visits. Nunamaker (1983) included total routine inpatient costs as input in different DEA models where different sample size and output variable specifications were considered. Kirigia *et al.,* (2008) specifically considered the expenditures in terms of pharmaceutical and non-pharmaceutical supplies, and concluded that these expenses need to be reduced by nearly 60%. Akazili *et al.,* (2008) also included the expenditures on drugs and supplies as an input. Lindlbauer *et al.,* (2016) included expenditures on supplies as a proxy to material resources. Ramirez-Valdivia *et al.,* (2011) considered three different types of costs as separate inputs including annual medical staff, annual general service and pharmacy. The authors recommended that annual medical costs should be further disaggregated with respect to different number of teams and number of members per team, to get a more detailed analysis. Rouse *et al.,* (2011) also considered two different types of costs including primary activities time and batch costs. However, the authors mentioned that these variables are aggregated, and provide very little guidance. Hence, in this case, ABC will provide an in-depth analysis on patterns of resource usage. Kawaguchi *et al.,* (2014) included the interest cost per year as input. Considering some other input variables, Butler and Li (2005) included the service complexity in terms of number of services offered as an input. Ouelette and Vierstrate (2004) considered the expenditure on furniture and equipment as an input whereas Blank and Van Hurst (2011) included the volume of material supplies as an input variable, including medical supplies, food and heating. While assessing the efficiency of emergency departments, Kose *et al.,* (2014) also included the bed occupancy rate as a significant variable. Generally, the DEA studies compare the efficiency of the hospitals as a whole rather than separate departments. The current study has the objective of extending DEA modelling for evaluating the queue system of a specific department, which is the outpatients' department.

The selection of variables in the DEA model depends on the objective of the study. Almost all studies compare the efficiency of hospitals using various inputs and outputs. The most commonly used outputs include the number of inpatient and outpatient visits, whereas number of beds and physicians are the frequently used inputs. The current study provides an *extension* of traditional DEA modelling by employing it for assessing the queuing problem by including relevant *queue-related variables*. Hence, the aim is to reduce the wait times by improving allocation of resources including personnel. A summary of various commonly used inputs in DEA healthcare studies is given in Table 3-2.

Table 3-2: Most Common Inputs Used in Healthcare DEA Studies

Inputs	Description of Most Commonly Used Inputs
Beds	Number of Beds *(Sheikhzada et al., (2012); Chuang et al., (2011); Akazili et al., (2008); Harrison et al., (2004); Prior (2006); Weng et al., (2009); Kirigia et al., (2008); Bwana et al., (2015); Dotoli et al., (2015); Lindlbauer et al., (2016))*
	Number of Beds as proxy to capital *(Flokou et al., (2011); Butler and Li (2005); Puig-Junoy (2000); Zere et al., (2001))*
	Number of Beds when evaluating different types of healthcare institutions/services *(Chang et al., (2004); Kose et al., (2014); Kawaguchi et al., (2014))*
Physicians	Number of Physicians *(Chaung et al., (2011); Harrison et al., (2004); Prior (2006); Weng et al., (2009); Kirigia et al., (2008); Ouelette and Vierstrate (2004); Bwana et al., (2015); Dotoli et al., (2015); Mitrovic et al., (2016))*

	Number of Physicians for different Specialities/ services *(Ramanathan (2005); Sheikhzada et al., (2012); Kose et al., (2014))*
	Number of working hours for Physicians (*Kirigia et al., (2007)*)
Nurses	Number of Nurses *(Huang et al., (1989); Al-Shammari (1999); Puig-Junoy (2000); Kirigia et al., (2008); Blank and Van Hurst (2011); Dotoli et al., (2015); Mitrovic et al., (2016))*
Other Medical/ Nonmedical Staff	Other Medical Staff *(Huang et al., (1989); Al-Shammari (1999); Ramanathan (2005); Prior (2006); Weng et al., (2009); Chuang et al., (2011); Sheikhzada et al., (2012); Kawaguchi et al., (2014); Dotoli et al., (2015))*
	Other Non-Medical Staff *(Puig-Junoy (2000); Akazili et al., (2008); Weng et al., (2009); Blank and Van Hurst (2011); Sheikhzada et al., (2012); Kawaguchi et al., (2014); Dotoli et al., (2015); Lindlbauer et al., (2016))*
	Number of working hours of Medical and Non-medical Staff *(Ouelette and Vierstrate (2004); Kang et al., (2014))*
Costs/Expenditures	Total Operating Costs *(Nunamaker (1983); Zere et al., (2001); Butler and Li (2005); Harrison et al., (2004); Tsai and Molinero (2002); Chuang et al., (2011); Kawaguchi et al., (2014); Mitrovic et al., (2016); Lindlbauer et al., (2016))*
	Expenditures on Pharmaceutical/drugs and non-Pharmaceutical/other supplies *(Akazili et al., (2008); Kirigia et al., (2008))*
	Expenditures on Furniture/equipment *(Ouelette and Vierstrate (2004))*
	Costs related to different services provided *(Ramirez-Valdivia et al., (2011))*
Some Other Inputs	Quantity of material supplies *(Blank and Van Hurst (2011))*
	Number of Services offered as proxy to capital assets *(Butler and Li (2005))*
	Primary activities and batch costs *(Rouse et al., (2011))*
	Interest cost per year *(Kawaguchi et al., (2014))*
	Bed occupancy rate *(Kose et al., (2014))*

3.4 Extended Applications of DEA Modelling in Healthcare and Extension to a Queuing Problem in the Current Study

In recent years, DEA modelling has moved away from its traditional and straightforward application in health, using specific inputs and outputs. A few studies have been conducted where DEA has been utilized to assess efficiency of 'units' other than healthcare organisations. These *extended applications* provide evidence that DEA theory has the capability to evaluate the efficiency of 'other' units of analysis, using suitable inputs and outputs, which deviates from its conventional implementation.

A few studies have evaluated the efficiency of hospital personnel including physicians and nurses. Chilingerian (1995) and Wagner *et al.,* (2003) assessed and compared the efficiency of physicians within one hospital. Therefore, the inputs and outputs considered were such that assessed the efficiency of physicians specifically. In the former study, total length of stay and total charges for ancillary services were the two inputs, whereas the number of discharges was the output. Discharges were divided into high severity and low severity cases as two separate outputs. Furthermore, two different models were run where surgeons and internists were considered separately in the first model, and the second model considered the case-mix of patients in addition to the type of physicians. The results showed significantly different efficiency and productivity results depending on the severity of patients, to aid in decision-making regarding the treatment of severe cases. A number of explanatory variables related to physicians' practice style and patient characteristics were also evaluated, where Health Maintenance Organisation (HMO) affiliation was significant in both models. The latter study carried out a stepwise procedure in DEA modelling to clearly evaluate the causes of inefficiency among physicians. This study conducted two separate DEA analyses to gain some additional insights, where the first model aggregated inpatient and outpatient measures, while the second model evaluated them separately. For the first model, aggregated as well as disaggregated inpatient and outpatient costs were considered as inputs, with quality indicators and severity of cases as outputs, added in a stepwise manner. The results showed that aggregation did not affect the efficiency results, however, some additional physicians became efficient as quality indicators and severity of patients were included. For the separate inpatient model, adding quality indicators only led to a slight change in efficiency results. However, for the outpatient model, disaggregated costs and adding quality indicators resulted in additional efficient units, but severity of cases did not affect the efficiency results. Osman *et al.,* (2011) assessed the efficiency of nurses in an intensive care unit, with respect to performance evaluation and appraisal, considering related input and output variables. The outputs included planning/organisation, nursing and technical practice performance, emergency work follow-up and problem-solving creativity. The inputs included job knowledge, work habits, teamwork and cooperation, interpersonal skills and communication. The results were used to identify the best practices among all in order to develop suitable career development plans for performance improvement of inefficient nurses.

Considering some *other DEA applications*, Rouse *et al.,* (2011) identified the efficiency of immunization activities in primary care practices. In this case, the performance was assessed using vaccination activities as an output, with two aggregated inputs, total time spent on primary activities and total batch level cost. Some other factors were considered as well, such as practice size, time spent on patients and urban population. The inefficiency is higher for those immunization activities which are carried out in small practice sizes and where the registered patients less 5 years old are in a greater number. However, large practices in more affluent areas spend longer on clients to discuss the benefits and risks of immunization activities, and hence use more resources leading to reduced efficiency. Vanelli *et al.,* (2013) used DEA to assess the efficiency of different colorectal cancer screening programmes. The regions considered were divided as large, medium and small according to admitted patients and number of found cancers, where these two variables were used as outputs. The number of colonoscopies was considered as a single input, which was the same for all three regions. The main goal was to evaluate the effectiveness of these programmes, in order to provide guidance for long-term planning and implementation. Gerard and Roderick (2003) evaluated the performance of haemodialysis satellite units (HDSUs), where patients treated per week was considered as the output with two trained nurses and dialysis machines as two inputs. The results showed that variation in resource utilization as well as difference in the number of patients treated and maximum slots available, led to inefficiency among HDSUs. The authors recommended that disaggregation of inputs, inclusion of quality measures as outputs and consideration of patient case-mix requiring dialysis, will lead to a more comprehensive analysis. Keshtkaran *et al.,* (2014) compared the efficiency of radiology units of hospitals. Two related output variables considered were number of admitted patients and number of photographic film. The inputs included the number of radiology devices, equipment maintenance and depreciation costs, number of expert personnel, number of technicians and the salaries of these staff members. The results showed that only 15% were efficient among all radiology units, and mainly expert personnel was in excess by nearly 40%. Another unique application of DEA is found in a study by Butt *et al.,* (2002). They carried out an efficiency assessment of stair-chairs, which are used by firefighters or emergency medical services staff members to transport a patient down the stairs. Six different stair-chair models were compared using DEA, under three different scenarios including leader facing backward, leader facing forward and leader facing according to manufacturer's recommendation. A single input of carrying task itself was considered with five different outputs including three-dimensional strength prediction program, lumbar motion monitor, risk of low back disorder, disc shear and compression. The sixth model outperformed all models, where the leader forward position was identified to have an advantage over leader backward position.

Jafari *et al.,* (2010) is the *only* paper to the best of our knowledge which specifically evaluated different queuing models using DEA. Hence, the queue-related inputs and outputs were considered. The inputs included inter-arrival rate, service rate, the average length queue, the average waiting time and busy period density. The output variables included number of servers and the average serviced customers. This study generally compared eight different queuing models in order to identify the waiting time and average length of queue. However, the current study is specifically concerned with providing a detailed evaluation of the queue system with no appointments in a large public hospital of a developing country; and providing some practical recommendations for improvement, where healthcare DEA studies are lacking.

In recent years, a few extended applications of DEA have been observed in the field of healthcare. These studies indicate the usefulness of DEA in assessing the efficiency of other units rather than the traditional application of performance evaluation of health institutions only. However, the number of studies highlighting such applications is very limited. The current study *extends DEA* modelling to assess the *queuing* situation within one hospital using patient level data with no appointments, in order to provide *evidence* of the effectiveness of DEA methodology specifically for improved queue management.

Chapter Summary

The current chapter mainly provides a review of theoretical aspects of DEA modelling and applications in healthcare. Firstly, the theoretical underpinnings of DEA have been specified, graphically and mathematically. The linear dual model of DEA is applied more frequently due to computational advantages. The slacks can be obtained by running another DEA model, and target values can be obtained from slacks. Most of the healthcare applications in DEA had the objective of comparing the efficiency of hospitals and other health centres. Besides, these DEA studies have been conducted in various developed and developing countries around the globe. However, only a limited number of studies have been conducted in Asia, with negligible studies in South-East Asian region. Almost all studies recommend transferring resources (beds/staff) from inefficient to efficient units, reducing excess resources, or complete shutdown of inefficient units. However, these recommendations seem highly improbable for implementation in developing countries which face social and demographic challenges in addition to constrained healthcare delivery. The current study aims at proposing a framework using DEA, which not only monitors a queue system but provides instant recommendations for improvement, which are suitable from an implementation perspective. Besides, some studies have used different types of DEA models or utilized DEA in conjunction with other statistical techniques, either to further validate (extended DEA models) or compare with (parametric methods such as SFA or PIs) initial DEA results, or to evaluate the factors causing inefficiencies (Regression Models such as Tobit Regression). Furthermore, an evaluation of the input and output variables used in healthcare DEA studies has been conducted. Most commonly used outputs in previous works are the outpatients' visits and inpatients' visits, whereas number of beds and physicians are the frequently used inputs. The current study aims at extending DEA modelling by developing a DEA-queuing model. Therefore, queue-related variables will be included in the DEA model. Additionally, considering the issue of staff shortage in public hospitals of developing countries, optimal utilization of personnel will be taken into consideration with the aim of reducing wait times. In recent years, it has been observed that DEA modelling has been utilized for efficiency assessment in healthcare, which is different from its traditional utilization for assessing the efficiency of health institutions only. These extended applications provide evidence that DEA has the ability to conduct efficiency analysis in other areas of healthcare, diverging from straight-forward applications. However, these studies are very few in number. The current study presents a novel application of DEA by applying for queue management for walk-in patients, highlighting the usefulness of this modelling technique other than its conventional usage.

CHAPTER 4 METHODOLOGY: RESEARCH DESIGN AND DEA MODEL DEVELOPMENT

Chapter Overview

The current chapter highlights the methodological underpinnings of the current study, divided into two mains segments of Research Design and a Data Envelopment Analysis (DEA) model for queue assessment (Queuing-DEA model).

The first-half of the chapter will elaborate upon the Research Design framework adopted by the current study. This section will commence with describing the philosophical underpinnings of the current study including epistemology, ontology and research approach. The next segment will evaluate the research strategy of case-study utilized in the current study as well as the specific case-study approach chosen. Furthermore, an evaluation of the specific duration of research is highlighted which is cross-sectional design. The next sub-section will elaborate upon the particular mixed methods approach adopted for the current study. Finally, a detailed explanation of the data collection techniques used to conduct empirical investigation will be provided.

The second-half of the chapter is dedicated towards the Queuing-DEA model development for the current study. This section will begin with developing an initial DEA model with respect to the queuing variables included and the corresponding data requirements. Next, the patient flow system at the case-study hospital of Pakistan is explained in detail, along with the identification of some additional characteristics prevailing within this queue system. Furthermore, an explanation of the actual data collection process at the designated hospital is provided. Based on the initial model and data requirements, patient flow system and the data gathering procedure, some crucial preliminary findings are evaluated. Additionally, some ethical issues have been identified which were taken into account whilst conducting the data collection process. Finally, considering the observations of the queue system and initial findings, a refined queuing-DEA model has been proposed. The various stages of the model have been evaluated including input/output selection, units of analysis; as well as the orientation of the model with mathematical representation of the developed model. It has been emphasized that the proposed DEA model has the capability of evaluating the queuing process under consideration.

4.1 Research Design Adopted for the Current Study

The philosophy and research design adopted for a research study has a major influence on the way research is carried out. It is of utmost importance that the philosophy, research design and strategies adopted are appropriate for the purpose of the research problem under consideration.

Saunders *et al.,* (2009) proposed the famous research 'onion' which depicts the different philosophical considerations when adopting a research strategy, all the way up to the choice of data collection and analysis techniques. This research 'onion' is shown in Figure 4-1.

Figure 4-1: The 'Research Onion' Representing Different Aspects of a Research Design

(Source: Saunders et al., (2009))

4.1.1 Philosophical Underpinnings

The research philosophy demonstrates the way in which a researcher views the world, the extent of contribution to knowledge and the process of developing this knowledge. The philosophical assumptions will underpin the research strategy and the methods chosen as part of that strategy (Saunders *et al.,* 2012).

There are two different types of *Ontological* positions, namely 'Subjectivism' and 'Objectivism'. Subjectivism, also known as social 'Constructivism' implies that social phenomena is created as a result of actions and perceptions of the social actors, and is in a constant state of revision due to changing social interactions. Therefore, this necessitates the need to conduct a detailed and in-depth investigation of a particular phenomenon or situation in order to fully understand it (Saunders *et al.,* 2012). Alternatively, Objectivism entails that a social entity has a reality external to the social actors within it. For instance, if the social entity is an 'organization', it has some fixed rules and regulations, adopts standard procedures and workforce has to fulfil tasks assigned to them. Therefore, an organization has a reality external to the people working within this set-up (Bryman and Bell 2015).

The current research study follows the ontological position of *'Objectivism'*. Like all health institutions, the public hospital where the current research is carried out is an 'organization' which has a fixed set of rules, regulations and procedures, independent of the 'people' working (medical and non-medical personnel) and visiting (patients). The staff and even the patients follow the standard procedures already set in place by the administration. Therefore, the current study aims to assess the queue situation keeping in mind the norms of the public hospital under investigation from an operations perspective, supporting objectivist approach; and not to understand the perceptions and behaviour of the staff and patients, as implied by the constructivist approach.

'Culture' is another aspect where the perception varies depending on the ontological position considered. Objectivism sees culture as a pre-determined norm whereas Subjectivism views culture as the 'result' of continuous social interactions (Saunders *et al.,* 2012). The current study specifically aims to investigate the queue system of a public hospital in a developing country. Therefore, a number of characteristics specifically associated with the health systems in developing countries are considered. The foremost

element is the absence of appointment systems, along with overcrowding and shortage of personnel. Furthermore, factors such high population, low literacy rate, poor transport facilities and lack of communication facilities such as internet or post, also create challenges for developing countries. All these aspects drive the 'culture' of the public hospitals, hence affecting the way health institutions are operated in developing countries, as inferred by the Objectivism, rather than 'culture' being the result of social behaviour within the hospitals as implied by Constructivism.

Following on to Ontology, *Epistemology* concerns the extent to which knowledge tends to be acceptable in a particular discipline (Bryman and Bell 2015). It regards the assumption about reality, that is, if the aim is to generate knowledge which is unbiased and generalizable or does it concern a specific time and place (Lee and Lings 2008). Epistemology can be divided into three different types including 'Positivism', 'Realism' and 'Interpretivism'. Positivism implies that a particular phenomenon is observable and allows for identifying consistency and causal relationships within the data, and create generalizations. Another aspect is that research is carried out in a value-free way, that is, researcher is external to the data collection process and does not alter the essence of the data to be collected. Furthermore, this approach emphasizes upon quantifiable observations related to a research problem, and carrying out quantitative analysis (Saunders *et al.,* 2012). Realism shares the same belief as positivism, which is, viewing the world from an objective perspective assuming that a particular phenomenon is observable and measurable. However, realists believe that there are certain aspects within a social reality which have an independent existence and cannot be measured objectively (Lee and Lings 2008). In contrast, the epistemological position of Interpretivism emphasizes on developing an in-depth understanding of the research subjects, and studying a phenomenon from their perspective. They are considered appropriate for the fields of business and management, marketing and human resource management where the idea is to indulge oneself in investigating human behaviour and interactions (Saunders *et al.,* 2012).

The current study has adopted a *Positivist* approach as far as epistemology is concerned. The emphasis is to identify the crucial queue variables which lead to excessive wait times of patients in a large busy public hospital of a developing country, where appointment systems are non-existent. Furthermore, the idea is to construct a generic framework for constant monitoring of the queue system, which can be utilized by other public hospitals in similar developing countries, for supervising the patient queues. The study is conducted in a 'value-free way', since the aim is to strictly observe and evaluate the queue process of outpatients as it is and suggest improvements accordingly, without altering the system to observe changes. Also, the study has emphasized on producing quantifiable data regarding the queuing situation to be utilized for DEA modelling, and generate suitable recommendations for improving the queue system as well as a dynamic framework for practical implementation. There are negligible studies which have evaluated the queue systems of walk-in patients in an outpatient setting. Therefore, it is extremely important to investigate the current functioning of the queue system in the absence of appointments, and the associated challenges faced by the administration in managing unorganized patient flows. Hence, data regarding the queue process has been gathered, by evaluating the significant queue factors and the conclusions are purely determined purely by the empirical results, as opposed to the Realist approach which assumes that some aspects may be non-quantifiable and not directly observable. Also, the current study aims to analyse the queuing process purely from an operations and efficiency perspective, not comprising of any subjective details regarding perceptions of the staff and patients, as encompassed by Interpretivism.

Following ontological and epistemological underpinnings, a particular research study could either have a *'Deductive'* or an *'Inductive'* approach. The 'Inductive' approach emphasizes that developing an in-depth understanding of the social phenomena, and the way people interpret their social world is important when conducting research. Therefore, rather than a rigid methodology as in case of deductive approach, a more flexible approach is adopted to allow for subjective explanations within the context. Hence, in this case, qualitative data is gathered mostly using a variety of methods with a small sample size (Saunders *et al.,* 2012). Alternatively, the 'Deductive' approach initiates with an existing theory regarding a social phenomenon, followed by developing a proposition(s) or hypothesis(s) based on these theoretical underpinnings, which are then empirically tested using suitable research strategies and data collection techniques (Bryman and Bell 2015). A highly structured methodology is used in this case to facilitate the observation and measurement of the concepts, often by adopting a quantitative design. Additionally, the deductive approach entails the process of 'reductionism' which ensures that the actual research problem is broken down in to simple elements to facilitate better understanding.

The generalization of results and conclusions is also considered as part of the deductive strategy, emphasizing on careful selection of variables, sample and sample size (Saunders *et al.,* 2012). A third type is the 'Abductive' approach which rather than moving from theory to data or data to theory, moves back and forth, thus combining deductive and inductive approaches. Here data is collected to explore a phenomenon with identifying themes to generate or modify a theory, which is then tested using additional data (Saunders *et al.,* 2012).

The current study adopted the *Deductive* approach mainly due to the nature of the research problem. The existing theoretical underpinnings and associated variables regarding queuing and DEA analysis have been highlighted. The current study emphasizes on the crucial issue of excessive queuing under a specific context, which is the non-existence of appointment systems in large public hospitals of developing countries. Some other characteristics affecting the health care delivery system in developing countries have also been considered. Additionally, it has been highlighted that DEA modelling has not been used to assess a queue system, therefore, providing an extension of DEA modelling.

Considering these aspects, *propositions* are framed which are then subjected to empirical analysis. The first is to identify the operational significance of existing queuing and DEA variables in the new environment of public hospitals without appointment systems. Following on to the first one, the second premise is to assess the usefulness of DEA modelling in evaluating a queuing problem. The third main proposition is to evaluate the crucial need of a dynamic framework with regard to practical implementation, which can constantly monitor the queue system.

In order to test these propositions, *quantitative data* has been gathered at the designated hospital of Pakistan, with the purpose of utilizing this dataset for DEA modelling and analysis. Hence, a structured methodology has been developed entirely dedicated to assessing the efficiency of the queue system from an operations perspective; as opposed to developing an understanding of the subjective elements within the context (such as patients' and staff perceptions and opinions) at any stage of the research, as indicated by the inductive and abductive approaches. 'Deductive' approach adopted by the current study is shown in Figure 4-2.

Aston University

Illustration has been removed for copyright restrictions

Figure 4-2: 'Deductive' Approach as Applied to the Current Study

(Source: Gill and Johnson (2002))

Therefore, considering the above mentioned ontological (Objectivism), epistemological (Positivism) and the theory/data orientation in the form of Deductive approach, the current study adopts a *Quantitative research design*; as opposed to the Qualitative research design which emphasizes on the generation of theory with interpretivist approach to understand the social behaviour encompassing a particular system (Bryman and Bell 2015).

4.1.2 Case Study Research Strategy

A research strategy can be defined as the methodological link between the philosophy adopted and the data collection and analysis methods used. The choice of research strategy mainly depends on the main aim and objectives of the research study, along with the extent to which it links to the philosophy undertaken and the existing knowledge about the research design adopted by similar studies. Access to potential data is another significant consideration as well as the availability of time and resources (Saunders *et al.,* 2012).

'Case study' is a very popular research strategy which is frequently adopted in business research. The 'basic' case study consists of a detailed analysis of a particular 'case' (or multiple cases), where the emphasis is to develop a rich understanding and an exhaustive investigation of a particular phenomenon within a specific context (Saunders *et al.,* 2012; Easterby-Smith *et al.,* 2012; Bryman and Bell 2015). With a case study research strategy, the objectives or research questions usually tend to answer 'how?' and 'why?' questions (Yin 2014), hence providing an opportunity to perform a detailed analysis of the subject matter. A case study approach can be used for theory testing, where the view is to assess if the theory can be applied to a real-life situation (De Vaus 2001), and to highlight the unique aspects of this case. Additionally, case studies consist of empirical inquiries focused on current events, rather than historical events. However, the environment or behaviours cannot be manipulated, as in an experiment. Due to these characteristics, case studies usually involve multiple data sources to develop an in-depth understanding of the problem (Yin 2014).

Within the case study strategy, a *'case'* can refer to an organization, a location, a person or an event (Bryman and Bell 2015), depending on the nature of the study. There are different types of cases including critical, unique, representative/typical or longitudinal. A 'typical' case can be defined as a case which is an exemplar to depict an everyday situation or norm (Bryman and Bell 2015), with the opportunity to analyse a phenomenon not observed by many before (Saunders *et al.,* 2012). The results generated from evaluation of this case can be implemented in other similar cases. A study could also have multiple cases, where usually the focus is to provide evidence if the results are similar across all (Saunders *et al.,* 2012).

The current study has adopted a *Case study* research strategy. The current study has selected the designated hospital in Pakistan as a *'typical'* case, which is one of the largest and busiest public hospitals of Pakistan, where all outpatients are 'walk-in'. Hence, it can be considered as a representative of large busy public hospitals in developing countries. The designated hospital provides primary and tertiary care facilities, and has a tremendous workload as nearly half a million patients are seen every year on average (641,447 in 2012), and at least 50-60 patients are examined every day on average in the busiest specialist outpatients' departments. The hospital has a bed capacity of 1,000 (1,300 in case of crisis). The total manpower, including medical and non-medical personnel is over 2,500. Specifically, there are nearly 90 specialists. This public hospital consists of over 30 different Specialties (Yearly Statistics Report 2013, Statistics Department, designated hospital). The number of patients and the workload indicate the excess overloading in the public hospital, and the staff numbers including the limited amount of specialists, demonstrate shortage of staff at this busy hospital. Additionally, this public hospital is situated in one of the largest cities in the country, and caters to a number of villages and towns in the suburbs as well. Therefore, considering these crucial characteristics, the index hospital selected is a 'representative' of most of the public hospitals in developing countries which have excessive queues in the outpatients' departments, and no appointment systems. The current study is interested in investigating 'how' a queue system in a hospital with no appointments operates and 'why' does a queue build up such that various factors specifically contributing to excessive queuing. Additionally, 'how' to improve queue by applying DEA modelling in this specific context.

Mainly, four major types of case study designs can be considered, including single or multiple case(s) with holistic (single unit of analysis) or embedded (including multiple units of analysis) design. When the study is related to the examination of a 'case' (such as an organization) as a whole, then this refers to a holistic approach. However, if the study is associated with the subunits within a single case, then this implies an embedded approach (Yin 2014). Similarly, these two approaches can be defined for the multiple cases, where the only difference is that the same approach is adopted in multiple cases rather than just one, to highlight the commonalities and differences among cases (Yin 2014).

The current study is specifically concerned with the operational evaluation of the 'queue system' within the designated hospital, for the outpatients, and not to investigate the overall functioning of the hospital. In order to assess the patient flow, wait times and other queueing data of 'patients' has been collected to evaluate the queue system within the designated hospital. Hence, patients are the 'subunits' or 'units of analysis' and the public hospital is the 'case.' Therefore, *single case* (designated hospital) with an *embedded design* (queuing data of patients) has been adopted for the current study. The Case Study approach used in the current study is shown in Figure 4-3.

Figure 4-3: Case Study Approach Used in the Current Study

(Source: Yin (2014))

4.1.3 Cross-sectional Research Design

Another consideration for conducting research, is whether to use 'cross-sectional' or 'longitudinal' research design. A cross-sectional design entails the collection of data at a single point in time. One aspect of cross-sectional design is dealing with 'variation' in data. This variation can exist with respect to people, organizations, nations or any other. In order to determine the extent of variation among units of analysis or cases, it is necessary to have a structured methodology which is standardized to measure variation to establish a consistent benchmark (Bryman and Bell 2015). Cross-sectional study uses quantitative techniques, however, they might use qualitative or mixed methods (Saunders *et al.*, 2012). In contrast, a longitudinal study allows for a study to be conducted over multiple time periods (Bryman and Bell 2015), therefore, providing an opportunity to study change or development over time (Saunders *et al.,* 2012). The selection of cross-sectional' or 'longitudinal' design depends on the objective of the study.

The current study has employed the *cross-sectional* design. The patient flow system has been evaluated using a case study approach at one point in time. The study is mainly quantitative in nature; however, a few interviews have been conducted only to further support the quantitative results. For the purpose of this study, cross sectional design is suitable as data for the 'existing' queue system was collected, and the results are used to provide guidelines for improvement. Therefore, data gathered over a short period was sufficient for the purpose of the study, and data over a prolonged time period was not required, as in longitudinal design. Furthermore, one of the objectives of the current study is to emphasize on the 'variation' in the arrival pattern of patients due to crucial factors such as excessive queuing and lack of proper management, since there are no appointments. Therefore, the current study aims to highlight this variation in order to assist in providing an accurate and a more comprehensive evaluation of the queue system, and appropriate recommendations for improvement. Additionally, due to constraints such as lack of time and resources and access to the hospital for multiple periods of time, cross-sectional design is preferred.

4.1.4 Mixed Methods Approach within Case Study Design

Due to the level of in-depth understanding which underpins a case study approach, usually an inductive approach is adopted with the collection of qualitative data. However, considering the main objectives of the current research study as well as the philosophical underpinnings adopted, the current study is mainly quantitative in nature. The case study research design can comprise of a quantitative design (Saunders *et al.,* 2012). If the dominant part of the research is quantitative, as for the current study, then the case study approach tends to have a deductive approach (Bryman and Bell 2015). In the current study, a deductive stance has been taken to assess the queue system in a public hospital (case) of a developing country, where the case study approach has been utilized within the quantitative framework. The current study strictly aims at assessing the efficiency of the queue system by employing DEA modelling using quantitative queuing data of patients, without indulging into the subjective elements such as behaviour or perceptions of patients and staff regarding queuing. Although the current study predominantly adopts a quantitative research design, however, interviews were used to gather some qualitative data, which is a form of mixed methods approach.

There are a number of factors that need to be taken into account when considering a mixed methods approach. The first consideration is the way quantitative and qualitative methods are joined, that is, if one is more dominant than the other. Another element is the level of integration between the two types of methods, specifically, whether the two are mixed at data collection, analysis or conclusion stage. The timing in mixed methods data collection is another element, that is, if both are conducted concurrently or sequentially (Onwuegbuzie Johnson and 2004; Creswell 2014).

Regarding the first consideration, Greene *et al.,* (1989) (as cited in Bryman 2006) provides five different ways of combining quantitative and qualitative data collection methods including triangulation, complementarity, development, initiation and expansion. The idea of 'complementarity' implies that one method merely 'complements' the other. Hence, one method further elaborates, enhances or clarifies the results gained from the other method (Bryman 2006; Johnson and Onwuegbuzie 2004). Similarly, all other ways have different roles to play in the data collection and analysis procedure. Considering the mixing of two methods at the data collection stage, the data, analysis and results from one strategy are used to frame the data collection procedure for the other (Creswell and Plano-Clark 2011). Finally, with regard to timing, there are different types of mixed methods designs including the convergent parallel, explanatory sequential, exploratory sequential, embedded, transformative and multiphase. The above two factors can also be considered when choosing the type of mixed methods approach. The 'Explanatory Sequential' design starts with the collection and analysis of quantitative data. This is then followed by the collection and analysis of qualitative data, which assists in further elaborating or clarifying the initial quantitative results. The quantitative research is the dominant strategy in this case. The mixing occurs at the data collection stage since the quantitative data is collected and analysed first, followed by the qualitative research. Also, the term sequential implies that the two types of techniques are conducted one after the other (Creswell 2014; Creswell and Plano-Clark 2011).

The current study aims at assessing the queue system of a public hospital in a developing country, from an operations perspective. Therefore, the major part of this research is quantitative in nature, as queuing data is utilized within DEA modelling to evaluate the queue system. However, some interviews have also been conducted to elaborate upon the quantitative results. Hence, in this case, interviews are 'complementing' the quantitative results as identified by the approach of *'Complementarity'*. Furthermore, the mix of two methods has occurred during the data collection stage, since the initial quantitative results have guided the data collection process of interviews, such as inquiring about the patient flow process and standard procedures. Although the quantitative data highlighted the severity of queuing issue and suggestions for improvement, however, the interviews provided further clarification of the results obtained from DEA analysis. Hence, the mixed methods approach of *'Explanatory Sequential'* has been adopted, where quantitative research is the key research strategy and was conducted first, followed by gathering qualitative data using interviews. This data and preliminary analysis further assisted in refining the DEA model for improved queue management. The 'Mixed Methods' design for the current study is shown in Figure 4-4.

Aston University

Content has been removed for copyright reasons

Source: 'Explanatory Sequential' Design by Creswell (2014)

Aston University

Content has been removed for copyright reasons

Figure 4-4: 'Explanatory Sequential' Approach for Mixed Methods as Used in the Current Study

4.1.5 Data Collection Techniques

Within a case study design, there are a number of methods which can be used to collect evidence. Some commonly used data collection techniques include documentation, archival records, interviews, direct observations, participant observation and physical artifacts (Yin 2014). The case study can incorporate any data collection method as long as it is practical (De Vaus 2001), and best serves the purpose of the research problem. In fact, often case study method can employ multiple methods of data collection (De Vaus 2001), as together they tend to provide a better understanding of the particular phenomenon under scrutiny (Runeson *et al.,* 2012).

The present study has used documentation for secondary data collection, and observation and interviews for primary data collection. Particularly, 'structured' observation has been used for quantitative data collection and interviews for the qualitative part to further support the quantitative results.

The information gained from documentation is considered to be relevant for every research study (Yin 2014). The documents can be in the form of letters, memoranda, email, written reports of events, progress reports, administrative documents or internal records (Yin 2014). Most commonly, documents are used to support and elaborate upon information gathered from other sources (Yin 2014). Furthermore, these documents provide an overall perspective of the operations of the organisation under investigation. The information gained from documents can assist in the selection of the particular case. For the current study, *documentation*, specifically, the yearly statistics report has been used which proved to be very useful in gathering general information about the structure of the hospital and its characteristics. The report provided aggregated as well as statistics separated in different categories such as, total number and 'types' of patients seen in a year (inpatients/emergency/outpatients), total manpower (medical/non-medical/support staff), different departments (outpatients'/pharmacy/administrative/wards), various Specialties and others. Therefore, some crucial issues were highlighted such as the absence of appointments, the busiest departments (specifically outpatients' departments), extent of overcrowding, shortage of staff and the cumbersome paperwork and procedures followed. This information highlighted that the designated hospital is a 'representative' of most public hospitals in developing countries, which will help in generating results, and a framework which might have the capability to be applied in other similar hospitals. A detailed overview of the data collected from the documents is discussed in Section 4.2.

Considering quantitative research design for the current study, *'Structured' observation*, also known as systematic observation, was used as the data collection technique. This method employs particular rules for observing and recording behaviours, known as the observation 'schedule', where the objective is to ensure that consistency is maintained when observing each participant's behaviour. Hence, it is possible to aggregate the behaviour of all respondents (Bryman and Bell 2015), in order to generate an overall conclusion if required. Observation, in general, could be either participant or non-participant. For the latter, the observer is not part of the social setting. In most cases, structured observers are non-participants as they strictly observe a particular phenomenon without being part of it. However, unstructured emphasizes on recording information in detail, and does not have a strict observation schedule or recording procedure. More frequently, participant observers use this strategy (Bryman and Bell 2015). For observation, the recording system should be easy to use (Bryman and Bell 2015), so that important observational aspects are not missed. Besides, observation can be done for short or long periods of time (Bryman and Bell 2015), including multiple short periods or few long periods. This also depends on the objective of the study, as well as time and resources available. Furthermore, structured observation is suitable for cross-sectional research.

For the present study, the Statistics department of the hospital and the clerical staff in the outpatients' departments (OPDs) confirmed that any data regarding the queuing of patients is not recorded by the hospital. Therefore, the current study considered 'structured' observation as the most suitable data collection technique for recording the waiting time and other queuing data of patients. In terms of methodology, one of the biggest strengths of the current study is that 'real-time' queuing data has been collected in order to utilize for DEA modelling. 'Structured' observation was suitable as focused data strictly regarding the queuing process, was collected within a cross-sectional design; without the requirement of recording any behavioural data (Easterby-Smith *et al.,* 2012). The data collected for each patient was same to allow for DEA analysis and comparison. Non-participant observation was undertaken, where the wait times of patients along with other aspects of the queue system were recorded from a distance. The recording method of wait times was kept very simple and quick to facilitate recording large amount of data in a short period of time, since the hospital was extremely overcrowded. The details of the actual data collection process through structured observation are discussed in Section 4.2.2.3.

'Interviews' were also used as one of the data collection techniques in the current study. Case study interviews tend to be flexible and guided, rather than structured and rigid (Yin 2014). In case of business research, the aim of an interview is to get maximum information from the respondent, regarding their own behaviour or that of others, or of an attitude, norm or belief (Bryman and Bell 2015). However, the interview also greatly depends on the nature of the research study and the objectives. For quantitative research, like for survey research, a structured interview is usually administered which is a standardized and consists of same questions for everyone. A semi-structured interview is more flexible, and the questions are more general or open as compared to a structured interview, and follow-up questions may be included as well (Bryman and Bell 2015). Some other types of interviews include unstructured, intensive, in-depth or qualitative, focused, focus group, oral history or life history. A 'focused' interview is one where the interviewer asks open questions but about a specific situation or phenomenon which is relevant to the research problem under consideration (Merton, Fiske and Kendall 1956; Bryman and Bell 2015). An effort should be made to record interview responses as accurately as possible. There are a number of factors which can distort interview results including probing (when the respondent is pushed for more information), prompting (when an interviewer suggests a possible answer to question to the respondent) and leading questions (Bryman and Bell 2015).

The current study aims at using *Interviews* to develop a better understanding of the patient flow process and the standard procedures, and further elaborate and support the quantitative data obtained from structured observation. A *'focused' interview* is more appropriate for the current study, since the interview questions will be directed specifically at the queuing system. The interviews (with open-ended questions) have been conducted with the clerical and administrative staff only, strictly focusing on understanding the patient flow process and set procedures and inquiring about the main factors which might lead to excessive queuing. The interviews did not incorporate any questions related to obtaining information about the perceptions or personal opinion of staff or patients, regarding the queue system.

In order to gather accurate information, the answers were recorded using *note-taking* at the same time. The interviewees did not allow for recording of their interviews by audio or video recorder. Additionally, an effort was made to ensure that the questions are strictly related to the queuing process, and were structured such that they are not probing, prompting or leading in any way. However, in some cases, the respondent was directed back to the subject of the question if the answer seemed incomplete, in order to maintain consistency and to ensure that required response is obtained. The details of the actual data collection process via interviews are discussed in Section 4.2.2.3. The complete Research Design adopted by the current study is shown in Figure 4-5.

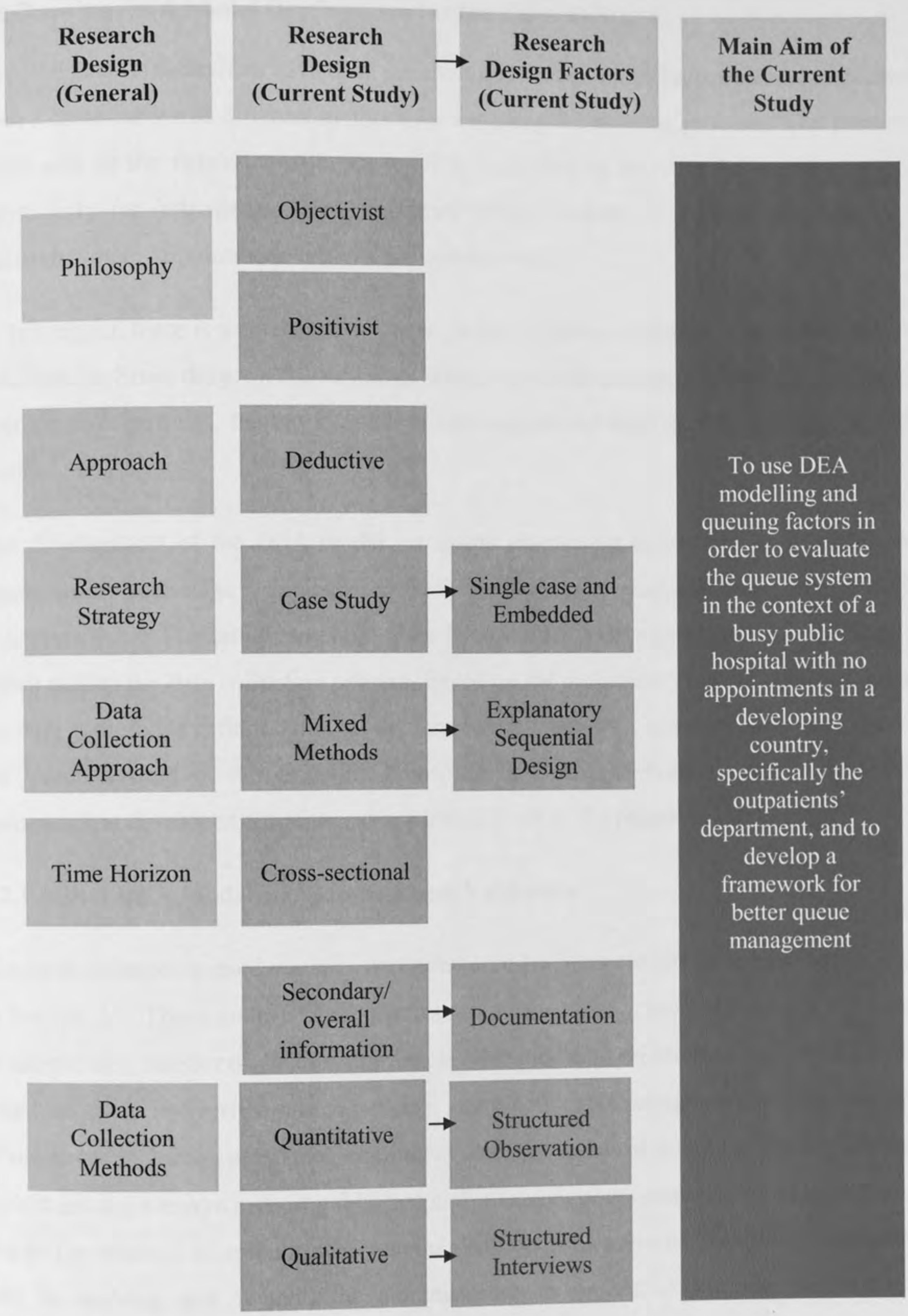


Figure 4-5: Research Design Adopted for the Current Study

4.2 Queuing-DEA Model Development for the Current Study

Over the years, researchers have used different OR modelling techniques and sometimes even a combination of different methods for assessing the queuing problem. The present study will be the first attempt to apply DEA modelling to evaluate the queue system, particularly for outpatients, in a large busy public hospital of a developing country Pakistan, where appointment systems are non-existent.

In this regard, there is a need to select *appropriate* variables in the DEA model to obtain valid results. Since the present study aims to investigate the queuing system and excessive wait times of patients, therefore, queue-related inputs and outputs were selected in the model.

The development of the DEA model for queue assessment consists of eight different stages, where each stage is explained in the following sections and a summary is provided in Section 4.2.6. The initial structure of the DEA model is developed first and validated, which guided the data collection process. Based on the preliminary results obtained from the data gathered, a refined DEA model was constructed, which is then used to evaluate the queue system at the designated hospital in Pakistan. Furthermore, a dynamic framework is developed for practical implementation of the proposed model.

4.2.1 Initial DEA Model Development and Validation

The most commonly used *queuing variables* used by other studies have been identified in Section 2.3. These mainly include wait time, service time (consultation time in terms of healthcare), number of doctors, number of other medical/non-medical personnel, beds, length of queue and arrival rate of patients; along with overloading and poor distribution of resources including personnel, equipment and other medical facilities. These queuing factors are more severe in most public hospitals of developing countries, such as Pakistan, due to the absence of appointment systems. Without a measure of 'how' many patients will be arriving and 'when', the administration is unable to pre-plan effectively. Consequently, the wait times of outpatients are excessive, and require a deliberate intervention. Similarly, some commonly used *input and output* variables in DEA were also identified in Section 3.3, with respect to healthcare. Some of these inputs include beds, physicians, nurses, other medical/non-medical staff and costs. Whereas some of the major outputs consist of outpatients' visits, inpatients' visits, major and minor surgeries, and discharges. The input and output values are aggregated such that they assess the efficiency of the whole hospital. However, none of these studies have included any queuing variables. The current study aims to provide evidence of the usefulness of DEA modelling in assessing a queue system; hence extending the application of this modelling technique. Therefore, the basic structure of the DEA model for queue assessment was developed considering these most frequently used queuing variables in queue management studies, and input/output variables within DEA studies, mainly including waiting time, length of queue and number of personnel.

Furthermore, in order to investigate the operational importance of these queuing factors within the initial DEA model, as well as other factors, a few observations and interviews were conducted in the busiest outpatients' department at the designated hospital in Pakistan. These observations/interviews were utilized to gather some preliminary information mainly regarding the arrival pattern of patients and the overall queue system such as the different stages followed by patients from entry till exit and the waiting time involved at each stage.

Some of these variables are related to inpatients such as number of beds, length of stay and discharge rate, however, the current study is specifically associated with assessing the queuing of walk-in outpatients. The initial observations and interviews at the designated hospital mainly demonstrated long wait times, overloading, unpredictable arrival behaviour of patients due to walk-in patients and lack of availability of personnel, as major factors. Hence, based on the preliminary observations/interviews, a second-stage DEA model was developed (please refer to Safdar *et al.,* 2016 for details). In this case, the *waiting time* of patients has been included as an input. Additionally, the *length of queue* can be considered as an appropriate indicator of the overloading in public hospitals, due to lack of appointment system, and has been included as an input as well. Some other queuing factors were also considered, for instance, consultation time was observed to be low, no-show behaviour is of little importance since there are no appointment systems, or idle time of doctors which is not a major issue due to high rate of overcrowding of walk-ins, keeping the doctors busy the entire session. Furthermore, the availability of personnel (doctors/pharmacists) was considered, highlighting the crucial factor of shortage of staff. In previous studies, the number of personnel (doctors/nurses/other personnel) has been included as an input (as identified from Section 3.3), where the objective is to reduce it in order to improve efficiency. However, the main goal of the current study is to evaluate and reduce the waiting time of patients, considering the current availability of doctors at a particular time when a patient arrived. Hence, the number of personnel is included in the initial DEA model as an output.

The preliminary observations highlighted some major aspects of the walk-in queue system with regard to the wait times, variable arrival pattern of patients and available personnel. Therefore, these initial observations were utilized to 'simulate' a dataset of 200 patients, representing the patient flow at the outpatients' department. The simulated data was subjected to DEA analysis for the main purpose of validation of the initial DEA model developed above. The results showed the effectiveness of the DEA model for assessing a queue problem. The comparison among the actual and target values of wait time and length of queue of each unit (patient) were useful in pinpointing the long wait times and/or length of queue given the current availability of doctors, and the required reduction (given by target values), highlighting the severity of the queuing problem. The validation of the initial DEA model proved that the information gathered is very useful since it alerts the administrators of high wait times and changing queue situation, so that appropriate measures can be undertaken for minimizing these wait times.

4.2.2 Data Collection Process

The development and validation of the second-stage DEA model as explained in Section 4.2.1 above, guided in initiating 'focused' data gathering. The data collection process consisted of observing the patient flow system including different patient pathways, recording real-time queuing data using structured observation, and few focused interviews to further elaborate on the functioning of the queue system. Additionally, some ethical considerations have been highlighted.

4.2.2.1 The Patient Flow System at a Busy Public Hospital of Pakistan

The *patient flow system* at the designated hospital in Pakistan was observed to identify the major wait times of outpatients where there are no appointment systems. Based on preliminary observations, it was observed that there are a number of different pathways that the outpatients need to follow, from entry till exit. When a patient first arrives at the hospital, he/she needs to be examined by a General Duty Medical Officer (GDMO). As mentioned in Section 1.4, there is a lack of independent General Practitioner (GP) referral clinics in some developing countries. Hence, unlike health systems in developed nations, the patients are either self-referred or referred by doctors who practice within large public hospitals. Therefore, these public hospitals are responsible for providing primary care in addition to secondary and tertiary care facilities to patients. The patient proceeds to the GDMO clinic reception, and collects a token where the token number is allocated on a first-come-first-serve (FCFS) basis. Next, the patient waits to be examined by a GDMO. Depending on the advice given by the GDMO based on the medical complaint of the patient, there are four different paths that a patient can take. One of the possibilities is that the patient is admitted to the hospital, and is categorized as an inpatient. Since the present study deals with the outpatients only, the admitted patients will not be considered for any further analysis, as they have a different procedure.

The other three options are related to *outpatients*. The first possibility is that the GDMO advises the patients for any necessary tests/investigations. In this case, patients proceed to the laboratory (Radiology/Pathology lab) for tests. Most commonly, these tests consist of blood tests, urine tests or chest X-rays. The patients can get the results either on the same day or some other day, if they are not routine tests. If the patients get the results on the same day, they consult with the GDMO again in light of these test results. Once they get a prescription, they proceed to the pharmacy. For this pathway, the patients wait at three different stages, including first GDMO consultation, laboratory, second GDMO consultation (regarding test results) and pharmacy. However, patients requiring second same-day consultation with GDMO are usually seen before other patients. Mostly though, the GDMO recommends them to come again the next day. Both situations cause inconvenience for patients, since same-day second consultations lead to extra waiting hours and other day consultations are not suitable for patients who have travelled from other towns/villages. If patients do not get the results on the same day, they exit the system. There is also the possibility that some special/non-routine tests require some prerequisite (like an empty stomach) and are not conducted on the same day. In this case as well, the patients exit the system. The second possibility is that the patients do not require any tests and further consultation. Therefore, the patients may proceed to the pharmacy in order to collect medicines prescribed by the GDMO. In this case, the patients have to wait at two stages, consultation with the GDMO and at the pharmacy. The third possibility is that the GDMO refers the patient to a specialist. Same-day consultations will require the patients to proceed to the specific specialist OPD depending on the type of medical complaint, for example, medical, ENT, skin, eye, child, orthopaedics, urology and rehabilitation/physiotherapy. At the specialist OPD, patients receive a token number and are served on a FCFS basis. However, since all Specialties do not have regular sessions every day, if a particular Specialty does not have a session that day, the patients are advised to come another day, and they exit the system.

Considering that the patients are examined by the specialist on the same day, there are two further possibilities. Either the patients do not need any further investigations and continue to the next stage which is pharmacy; or the specialist recommends some specific tests to be performed. In the latter case, there are several possibilities again, similar to the GDMO clinic. Firstly, the tests are not conducted on the same day and they exit the system. Secondly, the tests are conducted but they were unable to get the results, and the patients exit the system. Thirdly, they do get the results on the same day, and wait for a second consultation with the specialist, and then proceed to the pharmacy. The patients requiring second consultation are usually preferred to be seen by the specialist as soon as possible. Lastly, if they are unable to have a second consultation with the specialist on the same day (as recommended by the specialist or due to some other reasons), they exit the system. Therefore, these different scenarios lead the patients waiting at different stages including:

GDMO, specialist and pharmacy (if no further tests are required)
GDMO and specialist (if there are special tests which require some pre-conditions and cannot be conducted on the same day)
GDMO, specialist, laboratory, specialist, pharmacy (if they do get the test results on the same day)
GDMO, specialist and laboratory (if they do not get the test results on the same day/they do get results but second specialist consultation is on another day)

Considering various pathways taken by patients, the four major areas of waiting for patients include:

Consultation with GDMO (WAIT 1)
Pharmacy (WAIT 2)
Tests/Investigation (WAIT 3)
Consultation with the specialist (WAIT 4)

However, in addition to 'first-time' patients, there are many 'follow-up' patients as well. These patients do not necessarily require consultation with a GDMO. For instance, if the specialist advised the patient for a follow-up, then next time, patients can proceed to the specialist OPD directly. Therefore, an additional pathway, with regard to follow-up patients is:

Specialist and Pharmacy

Similarly, if the GDMO asked for the follow-up visit, the patient will proceed to the GDMO clinic. However, this time the GDMO may refer the patient for some further tests/specialist, if need be. In this case, the patients will follow the same stages as explained above. The patient flow system at the designated hospital is demonstrated in Figure 4-6.

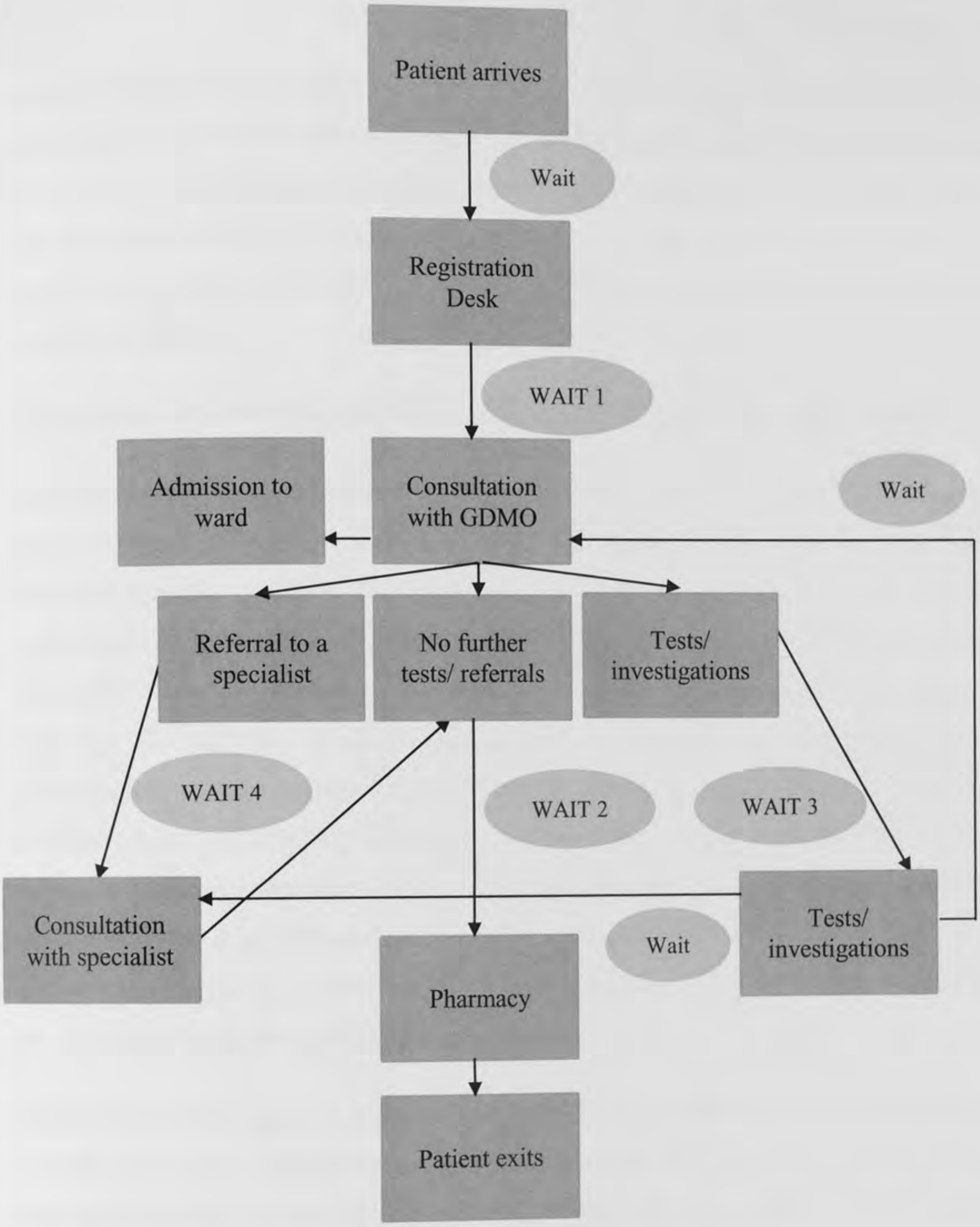


Figure 4-6: Patient Flow System at a Busy Public Hospital in Pakistan

It has also been observed that the different stages in the patient pathway are not dependent on each other, specifically in terms of queuing. Various departments such as GDMO clinic, specialist OPDs, laboratory and pharmacy, operate independently and have a separate queuing system. For instance, at the pharmacy, patients arrive from many different departments of the hospital including GDMO clinic, different specialist OPDs, and other independent centres such as Cardiology, Dental and Psychiatric. Also, there are patients who only came to collect their medicines, although they received consultation on the previous day. Similarly, for the GDMO clinic and specialists OPDs, there are different types of patients including first-time, regular, follow-up as advised, requiring the checking of lab reports or simply for renewal of prescriptions. Therefore, all departments have a 'mix' of patients and/or coming from different departments of the hospital. Hence, the queuing system of each department was observed '*independently*', with the objective that these mutually exclusive evaluations will lead to an overall improvement of the queuing situation.

4.2.2.2 Characteristics of the Data Collected at the Index Hospital in Pakistan

Detailed queuing data was collected at the designated hospital in Pakistan over a *period* of 7 weeks (23rd April to 28th May 2014) at GDMO clinic, radiology/pathology department, a few specialist OPDs and pharmacy. The queuing data was collected for a total sample of 741 patients. Due to incomplete information, a total of 37 data forms were discarded. Therefore, the data of 704 patients was further considered for DEA analysis. The data was collected during the working days (Monday-Friday) and working hours (7:30am-4pm). With regard to specialist OPDs, initially, the data was collected in a number of OPDs including medical, eye/ENT, skin, child, rehabilitation, urology, orthopaedic and gynaecology/obstetrics. However, later on, 'focused' data collection was conducted in the three busiest departments including GDMO clinic, in the three busiest departments including GDMO clinic, busiest specialist OPD and pharmacy, where excessive queuing prevailed (see details in Section 4.2.3)

The pharmacy has *separate* counters for men (serving, retired and senior citizens) and women. Similarly, GDMO clinic and a few specialist OPDs (medical, gynae) operate separately for males and family (women and children under 12 years) as well. This unique system prevails in most of the Islamic developing nations and some other developing nations as well. Overall, initial observations showed that excessive queues prevailed at male OPDs and male counters (serving/retired/senior citizens), therefore, more focused data was gathered for these departments only, later on.

It was observed that the dataset of all three busy departments consisted of some *outliers* with regard to the wait times of patients. Considering the busy *specialist OPD*, some patients were examined only within a few minutes after their arrival, and had a wait time of less than 1 hour. This is possible due to a number of reasons including urgent/emergency cases, or patients only requiring a renewal of medicine prescriptions to be signed off by the specialist. Some rare cases can be VIPs, patients known to the specialist, or if the people accompanying patients have to return to essential duties during working hours, where these patients are triaged by the receptionists. Additionally, there were some patients who waited for nearly 4 hours. Considering other observations, any wait time beyond 4 hours may have other unusual reasons such as attending other specialist OPDs or going to laboratory for investigations, leaving the hospital to attend to personal commitments and return back later or taking an extended tea/lunch break, though after reporting at the Reception of this OPD. In all these instances, the patients left the concerned OPD without reporting at the reception and returned back after undertaking these activities. Hence, the busy specialist OPD had a total of 182 observations, among which 22 data points had wait times that fall into the outliers' range (below 1h and above 4h), therefore DEA analysis was carried out for the remaining 160 observations for the busiest specialist OPD.

Among the three busiest counters at the *pharmacy* including serving, retired and senior citizens; the serving personnel counter had the maximum overload with excessive waits. The current study conducted DEA analysis for the 'serving' personnel counter at the pharmacy, although the same model can be applied to other busy counters as well. Overall, it was observed that the patient trickle in slowly in the mornings at the pharmacy counters. Hence there are no queues and the wait times are minimal. Usually these patients are the ones who may have obtained the prescription the day before but did not collect the medicines at that time due to unavailability of medicines or they had to leave to attend to some other urgent tasks. Therefore, they reported early the next day. These patients had a wait time of less than 30 minutes. On the contrary, it was also observed that some patients had a wait time of over 3 hours, for similar causes as that for specialist OPD. Therefore, from a sample of 68 observations for serving counter at the pharmacy, DEA results were obtained for 52 observations, hence 16 data points lied within the outliers' range (below 30 minutes and above 3h).

With regard to the *GDMO clinic*, it was observed that during early mornings, the inflow of walk-in patients is gradual. Hence, there were minimal queues with less wait times. These patients had a wait time of less than 30 minutes. On the other hand, some patients waited in excess of 3 hours. The reasons observed for these extreme values were similar to those identified for the specialist OPD. Hence, 20 observations from a total of 134 data points for the GDMO clinic had wait times falling within the omitted range (below 30 minutes and above 3h), where further analysis was conducted for the remaining 114 observations.

Hence, for each set of observations (for the three departments), the data points consisting of extreme values were *disregarded* before conducting any further analysis, for the sake of homogeneity of data as well as to improve the accuracy of results. A summary of the data utilized for the current study is shown in Table 4-1.

Table 4-1: Summary of Data Used for the Current Study

Data for Three Busiest Departments at the Designated Hospital	Total Data	Outliers (Excluded Data)	Data Considered for DEA Analysis
Busy Specialist OPD	182	22	160
Pharmacy (Serving Counter)	68	16	52
GDMO Clinic	134	20	114

The data collection procedure was extremely tedious due to a huge influx of patients, mostly arriving at the same time, hence leading to difficulty in keeping a track of 'all' patients arriving in a sequence. Therefore, due to the excessive load of patients, the data was collected for every 'third' or 'fourth' patient that arrived in queue in any department. Due to the huge number of patients, this technique was considered appropriate as an effort was made to collect data for a large sample size.

Furthermore, in order to observe the variability in the arrival pattern and load of patients, data was collected during *different* hours of the day and *different* days of the week. For instance, the medical, skin and gynaecology OPD sessions were conducted three times a week (Monday, Wednesday and Friday) whereas orthopaedics, eye, ENT and urology OPDs had sessions two times a week (Tuesday and Thursday). Therefore, observations were made during those days when these OPDs had sessions. Similarly, Swas operational all five days of the week, and observations were made accordingly. However, all busy departments experienced huge inflow of patients all week round, therefore, a combined analysis was conducted to cater for all days.

4.2.2.3 Data Collection for the Queue System Using Observations/Interviews

Based on the initial DEA model, data requirements, the availability of data and the various pathways followed by outpatients, the data collection process was initiated, using structured observation and focused interviews (as explained in Section 4.1.5). According to variables identified in the initial DEA model, data has been recorded regarding the wait time, length of queue and number of doctors, as well as consultation time.

Almost all large public hospitals in developing countries do not collect specific data regarding the 'queuing' of patients. Additionally, there are no definitive strategies in place, which solely monitor the queue system within these large public hospitals with no appointment systems. Therefore, in order to observe the exact waiting times and queuing data for each patient, a *time sheet* was prepared which was used to collect data via *observation*, in each busy department. The personal information on the time sheet data was recorded strictly to identify the data of each patient when conducting analysis later on. The other information (name, patient number (MR number) and gender) is already recorded by the receptionist when a patient arrived, and was copied on the time sheet. There was no interaction with the patients to collect this data. Similarly, a brief description of the medical complaint was also recorded by the receptionist as requested by the doctors to keep a record of different types of patients examined on a particular day. The waiting time data was collected in such a manner that once a patient arrived at the reception, the time was noted and the time sheet was given to the patient. The patients were advised to keep the time sheet safe and carry on with their following activities in the usual way, but when they were leaving that department, the time sheet should be returned to the reception. When a patient returned a time sheet, the time was noted again. During the busy hours of the day, some patients did not return the time sheets probably either because they misunderstood, forgot or simply lost the form. Although a record was kept of the number of time sheets given and returned, however, the unreturned/incomplete time sheets were discarded and not used for further analysis.

Similarly, when a patient arrived at the reception (GDMO clinic/specialist OPD/pharmacy), the number of patients already waiting to be examined by the doctor was recorded, to determine the length of a queue at the time a particular patient arrived. This information was noted on a separate sheet.

The number of personnel (GDMOs/specialists) varied throughout the working hours of the day. However, this information was already recorded by the receptionist under the category of 'Doctors Attending'. Therefore, in this case, when a new patient arrived, the number of doctors working at that particular time was copied down on a separate sheet. However, the process was different at the pharmacy. The main pharmacy of the hospital has different counters including serving, women, retired and senior citizens, where each category had two counters (two pharmacists) which was fixed.

The time sheet for observations of queuing data is shown in Figure 4-7 below.

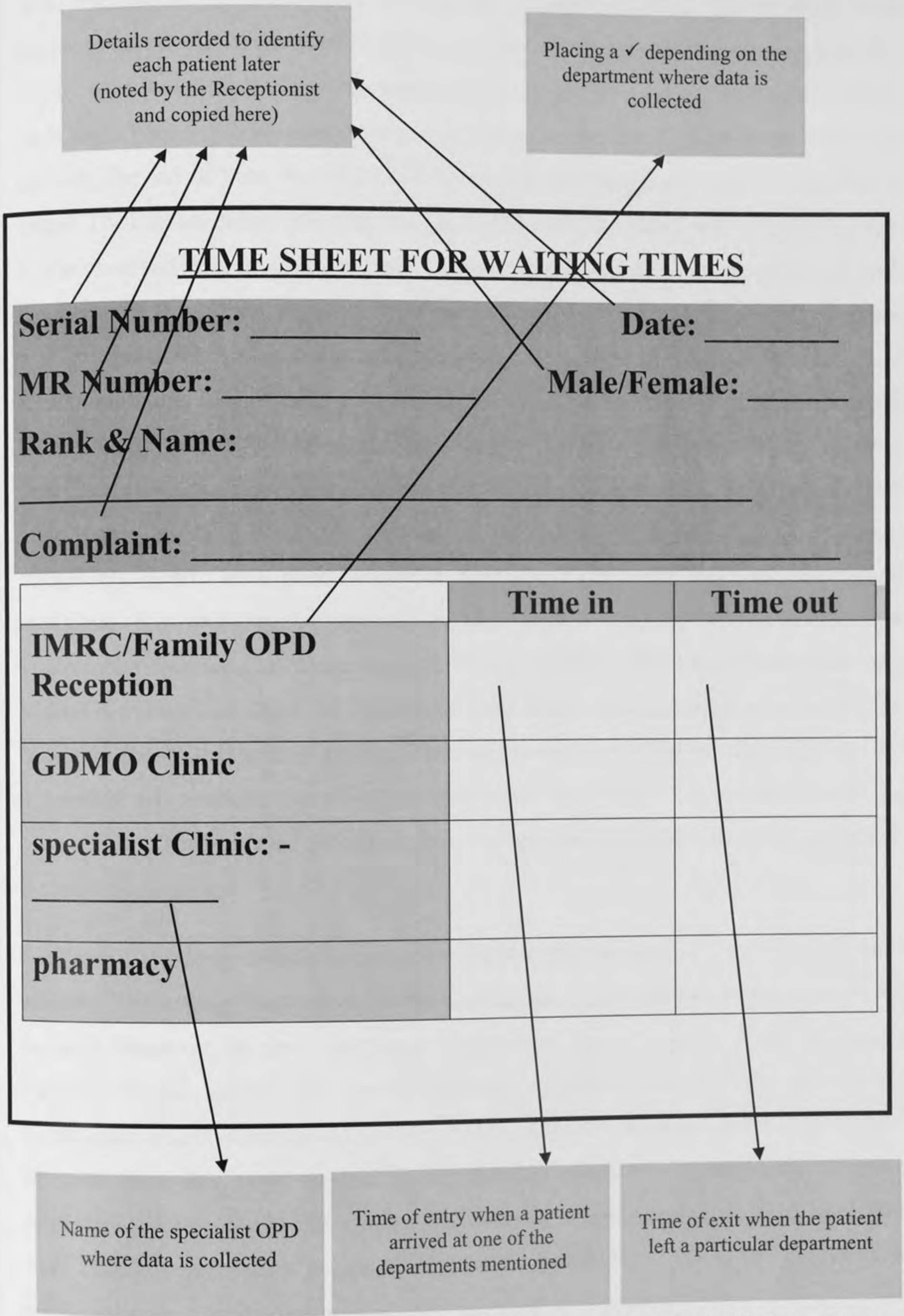


Figure 4-7: Time Sheet Used for Observation of Waiting Time of Patients in All Departments

With reference to the *consultation time*, due to the long queue of patients, the doctors are unable to spend a lot of time on one patient, as they have a huge number waiting, and they had to make sure that all outpatients were seen on the same day (though the number varied each day). Nearly 20-30 observations were made in the consultation room. For these patients, the arrival time was noted, the time when the consultation started and when it ended. The doctors did not allow for any more observations, since it was a private session. It was observed that, on average, the consultation time was only 5-6 minutes, even with the specialist. Therefore, for other observations conducted later on, the consultation time was extrapolated based on the actual observations. A number of factors were considered for extrapolation. One of the factors was the age of the patient. The children usually took a little longer than the adults as observed since the doctor needed to carefully examine them, and also since the children are not very cooperative. Similarly, the elderly patients took longer as well, as in some cases, they were unable to explain their medical complaints clearly or it took longer for doctor to explain the diagnosis and use of medicines. However, regular patients (diabetes, hepatitis injection or hypertension) or viral/regular diseases (flu, fever, diarrhoea or sore throat) took extremely less time, only around 4 minutes as observed. Second or third follow-ups (checking reports or post-operative check-ups) took on average 6 minutes as observed. Patients who only came for a renewal of medicine prescriptions took even less, only 2-3 minutes based on observations. The first time patients, however, depending on what their diseases, took 6-8 minutes on average.

Similarly, some observations were made at the pharmacy to obtain the service time when allowed. The average time taken for the pharmacists to give medicines was around 4-5 minutes. However, the time varied depending on the type of counter. There was slight variation in the service time at the pharmacy depending on the type of counter. Sometimes, it took more time to provide medicines to elderly women and senior citizens as more time was spent explaining the different medicines. Additionally, if more medicines are required by the elderly, occasionally it took longer to collect medicines from different sections of pharmacy. Therefore, in both these cases, the provision of medicines took 7-9 minutes on average as observed.

In addition to the observations as part of the data collection, *focused interviews* were also conducted mainly to further complement the quantitative results obtained from observation regarding the severity of the queuing issue. A total of 28 focused interviews have been conducted. The respondents included doctors including GDMOs and specialists, Pharmacists as well as Head of pharmacy, and personal assistants/nursing staff in GDMO clinic, specialist OPDs and pharmacy. Additionally some interviews were also conducted with the senior administrative staff including the Deputy Chief Executive as well as, quality assurance officer and statistics officer.

As with observing the queuing process, the initial round of *interviews* (as discussed in Section 4.2.1) were focused on gathering information about the actual patient flow, such as, the different pathways followed by patients and if the load of patients and wait time varies in different departments. These preliminary interviews assisted in gathering information about the patient influx, and provided a basis for more focused interviews conducted later, which particularly emphasized on identifying key elements which result in long queues and ways of to improve the queue system. During focused interviews, respondents were asked if the load of patients varies during the day or different days during the week, and the associated challenges. Also, in their expert opinion, what are the factors or bottlenecks which lead to long wait times of patients; and if there are certain activities which can be either re-organized or removed to streamline the queue system?. Additionally, the respondents were asked that in their opinion, what strategies can be adopted in order to improve the queuing situation which are feasible.

During interviews conducted with nursing staff, extra information was also gathered with regard to different activities which take place during the consultation. For instance, if any pre-examination data gathered from the patient (patient history or other demographic details), if this information is collected before or during the consultation, how much time does it take, if this information is recorded by the doctor or another staff member, and if medicine prescriptions are provided by the doctor during consultation or another staff member after consultation.

Therefore, these brief focused interviews with the staff/administrators led to a better understanding of the queuing process, and further assisted in quantitative analysis.

The interview *questionnaire* is displayed in Figure 4-8.

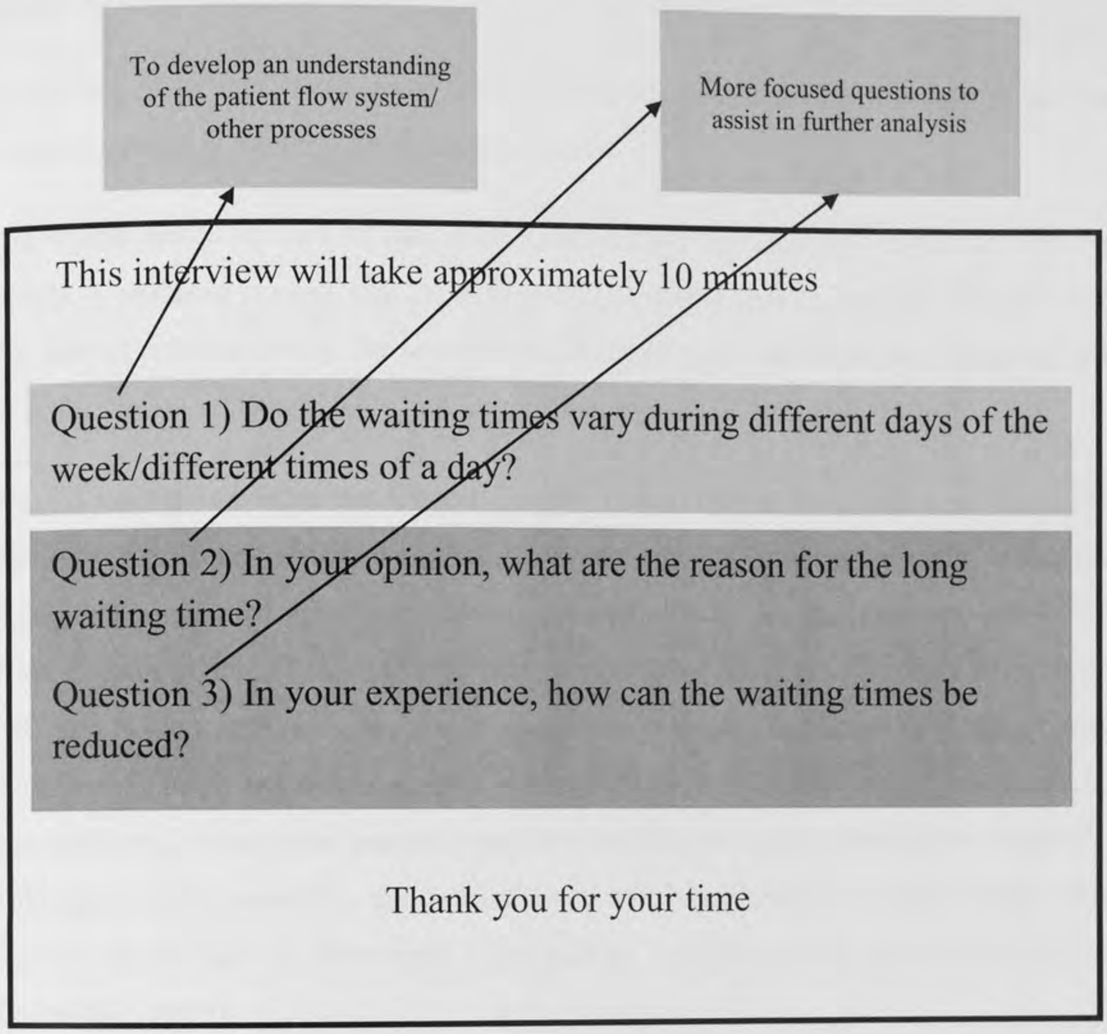


Figure 4-8: Focused Interview Questionnaire for Staff/Administrators

Therefore, the *initial DEA model* helped in establishing data requirements. The preliminary data collection (by observation and interviews) helped in developing an understanding the operation of the patient flow system. Consequently, a more 'focused' data collection procedure was adopted which targeted the busiest departments of the hospital. Hence, the results obtained helped in refining the DEA model for a comprehensive evaluation of the queuing process.

4.2.2.4 Ethical Considerations

A number of ethical issues have been considered specifically with regard to the data collection process adopted for the current study.

During the entire process of data collection, analysis, generating results and reporting results, it has been ensured that the case study hospital is kept *anonymous*. Therefore, for the sake of confidentiality, the selected health organization in Pakistan has been referred to as the 'index', 'designated', 'public' or the 'case-study' hospital.

Formal *permission* from the Ethical Review Committee of the designated hospital of Pakistan has been obtained before commencing the process of data collection. Additionally, Ethical approval has also been acquired from the Ethics Committee of Aston Business School, Aston University, Birmingham, UK (see Appendix 1). An effort has been made to adhere to the ethical guidelines proposed by Research Ethics of Aston Business School (Aston Business School, Research Ethics, Issues to Consider). Furthermore, it has been ensured the concerned body at the designated hospital in Pakistan is fully aware that the data collected will be utilized in a research study being carried out in the UK (European Commission, Data Protection and Privacy Ethical Guidelines, 2009).

The use of the *data collection technique* of structured observation was strictly targeted at gathering data regarding the wait times of patients in different departments, for the outpatients. Therefore, the arrival time of the patient was noted when he/she reported at the Reception of each department, and the time when the patient left the department. As explained in Section 4.2.2.3, the arrival time of patient was noted and the time sheet was given to the patient with instructions to return the time sheet when they left the department. The time sheet was handed over to the patients by the receptionist who explained all instructions. Therefore, any verbal interaction with the patients was not required. The time sheet had some additional data such as name, patient number, age, gender and medical complaint. This information was noted down by the receptionist for all patients. Hence, this information was copied down on the time sheet, without engaging with the patients. The main reason for noting down the demographic data of patients was to differentiate the data of one patient from another for the purpose of data entry and analysis, later. Some data characteristics such as age, gender and medical complaints, were also used to identify different types of patients specifically to elaborate on the queuing situation. It has been strictly ensured that the *anonymity of patients* has been preserved during all stages such as data collection, data entry, analysis and when presenting results. For this purpose, all patients have been referred as Patient 1, Patient 2, Patient 3 and so on.

Some structured interviews were also conducted with the administrative, nursing and other staff members. These interviews strictly targeted at obtaining any additional information about the patient flow system and functioning of different OPDs, and the standard procedures involved which may affect the queuing problem. Questions directly related to personal opinions or sensitive issues were completely avoided since they were not required for the purpose of this study. For the sake of anonymity, information regarding name and position was only gathered to identify one interview from another, and was not used in data entry and analysis later. In almost all developing countries, such as Pakistan, the interviewees refuse to give an interview if asked for a signature or recording interview with a tape-recorder, even if assured of *confidentiality*. Therefore, the interview was recorded using note-taking, as well as the name and position of each interviewee with date was written only to identify different interviews. Furthermore, in order to avoid causing any obstruction in the usual work routine of the interviewees, it was ensured that the interviews are conducted at less busy times or lunch/tea breaks, and do not exceed fifteen minutes.

Some *secondary data* was collected from the Statistics department of the designated hospital. Information regarding the basic structure and different departments of the hospital was gathered, along with associated statistics related to total number of patients, workload and staff. The data obtained from any annual reports/power-point presentations was kept *anonymized* preventing any disclosure of the designated hospital.

In order to collect data from interviews, *informed consent* was obtained from the respondents by providing a consent form accompanied by an information sheet (see Appendices 2a and 2b). The information sheet consisted of an explanation of the problem being investigated, reasons of investigating this problem, motivation behind collecting data from the designated hospital, and how this study could benefit the hospital and the level of patient care in general. Additionally, it also included type of data being collected, how data will be recorded, and how anonymity of respondents and the institution will be ensured. Work email address was also provided if the respondents wish to obtain any further information. Formal verbal consent was also obtained from the head of department (of each department where data was collected) for structured observation of wait times. When providing time sheets, the patients were also briefly explained by the receptionists about the purpose of the study. During initial stages of data collection, in order to familiarize with the operations of each department, some volunteers were requested to assist in the data collection process, and informed consent was obtained (see Appendices 3a and 3b). All respondents and volunteers had the right to withdraw from the research study if they desired.

Another important consideration with regard to ethics is the *storage* and *disposal* of data gathered. All raw data, analysis and results gathered is locked (hard-copy) and password protected (electronic copy). In order to prevent loss of data, back-up has been set up using Dropbox which is password protected as well. Only the academic supervisors are able to access the data collected, if required. The consent form prepared for the respondents clearly mentions about the access to data by the supervisors only (The Research Ethics Guidebook, Data Storage and Data Security). After the submission of PhD thesis and Viva, and any journal publication related to this research study has been completed, the data collected (electronic) will be disposed of, which is usually after a period of five years (Aston Business School, Research Ethics, Issues to Consider).

Throughout the course of the PhD programme, efforts have been made to *present* the current research including basic objectives, contribution and results. Oral presentations have been given at various international conferences, as well as poster competitions and seminars within Aston University in order to get constructive feedback. Additionally, a journal paper related to the initial model developed in the current study has been recently published. An effort will be made to continually write papers for journal publication, with regard to the current research study.

The data gathered and the results obtained will be shared with the designated hospital in Pakistan, if desired by the Head of the institution and/or other administrative staff at any time, especially if required with regard to the implementation of the results and recommendations proposed. For this purpose, if demanded, a short report can be prepared which will summarize the analysis and results obtained from the current study, which is easily understandable, avoiding the technical aspects of mathematical modelling.

Therefore, some crucial characteristics of the data collection process have been highlighted which demonstrate the ethical considerations taken into account with regard to the current study (see Figure 4-9 below).

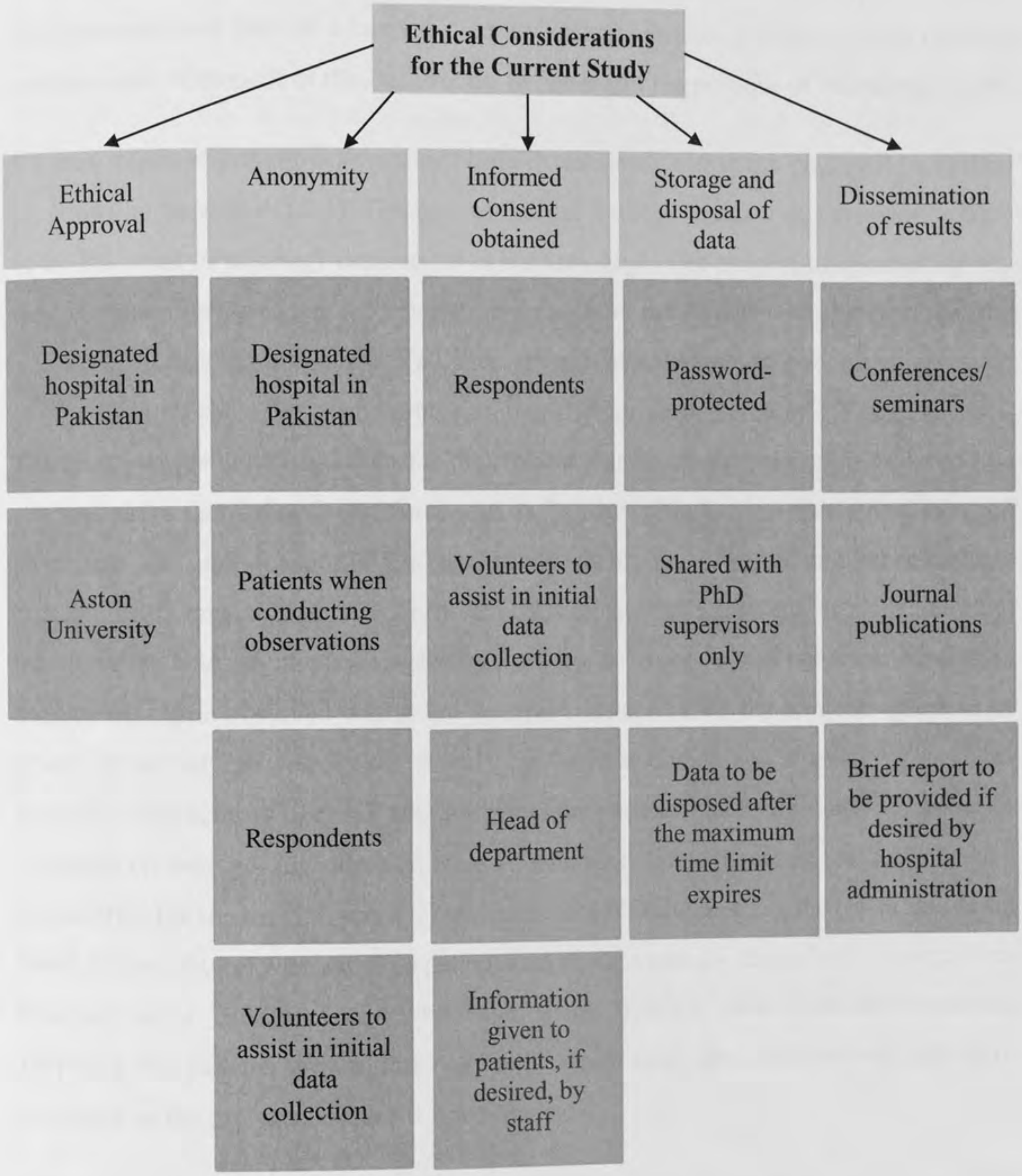


Figure 4-9: Summary of the Ethical Issues Considered by the Current Study

4.2.3 Preliminary Findings Obtained from Data Collection Process

The initial framework of the queuing-DEA model and the corresponding real-time data obtained from observations, and interviews resulted in some useful findings. These interesting findings emphasize on the prevailing inefficiency in the queue system of walk-in outpatients, and provide a basis for a more comprehensive evaluation of the queuing problem and refinement of the DEA model to better suit the purpose of the current study.

Various stages of *patient waiting* have been investigated within the patient flow system (as shown in Section 4.2.2.1). The data and initial findings indicate that the waiting time at the Radiology/Pathology department is not very high. The process of conducting any tests is streamlined and has little variation throughout the day. It was observed that the wait time, on average is one hour for different types of tests including common tests such as blood, urine, chest X-ray and ultrasound, or the rare tests including CT scan or MRI. Therefore, it was concluded that this department was comparatively organized and had less wait times than other departments such as GDMO clinic, busiest specialist OPDs and pharmacy. The main reasons in this regard are probably that a limited number of patients requiring tests/investigations or since all OPDs do not have sessions on the same day, therefore, the load of patients is distributed among different days of the week. Similarly, another wait time observed was at the Reception desk of each department, where brief details of the patients were taken quickly by the receptionist, and a token number was provided. This activity does not take long, and the patients had to wait only for about 2-3 minutes on average. Additionally, the wait time for patients requiring second same-day consultation (as shown in Figure 4-7) with either the GDMO or specialist (re-examination based on test results) was low. The patients are given a note by the doctor advising them to come for a follow-up check-up after going through some tests/investigations. Therefore, the patients showed this note to the receptionist upon their re-visit, and were examined by the physician within a few minutes.

Furthermore, as mentioned in Section 4.2.2.2, some departments had *separate areas* of examination for males and family (women and children under 12 years of age). It was observed that the waiting times were lower (one to one and a half hour on average) for family general OPD (GDMO clinic for women and children under 12 years), specialist OPDs for women (medical and gynaecology/obstetrics) and the women's counter at the pharmacy. Also, among various specialist OPDs observed, one of the OPDs was the busiest among all, and had extreme overloading and excessive wait times. Also, based on observations, the busiest counters at the pharmacy are those of serving, retired and senior citizens (males). Generally speaking, serving personnel mostly collected medicines for their families or parents as well. Besides, since all of them are working, many of them arrived at specific times such as during lunch break or later in the afternoon. Therefore, a huge load of patients builds up during these times, leading to excessive waits. With regard to retired/senior citizens, most of these patients were regular or follow-up patients and had to come for a check-up more often than other patients. Besides, some of them only came for the renewal of medical prescriptions and hence go to the pharmacy to collect their medicines. Therefore, it can be concluded that wait times at Radiology/Pathology department, reception of each department, second same-day consultations, and patient flow for Family OPD (women's counter at pharmacy) can be ignored. Hence, emphasis and a more comprehensive analysis is required to evaluate the excessive queuing at *three stages* including GDMO clinic, the busiest specialist OPDs and pharmacy (counters for serving, retired and senior citizens).

Due to absence of appointment systems and the resulting variation in the arrival pattern of patients, *overloading* was observed to be a critical element at the designated hospital. For instance, considering the busiest specialist OPD, it was observed that the load of patients is high during early morning as compared to the later part of the day. A large number of patients arrive at around 7am, since the OPD session starts at about 8:30am. On one of the days, it was observed that at least 50 patients were waiting at about 7:30am. On another instance, nearly 60 patients were waiting at around 10am. In most cases, if the arrival rate is about 50 patients per hour, the patients who arrive between 7-7:30am are not examined before 10am. This demonstrates the overcrowding in the OPD. The earlier the patients arrive, the lower is the token number, and they are among the first ones to be examined when the OPD session starts. For the GDMO clinic, on one of the days, it was observed that maximum number of patients arrived between 8am and 9am. Additionally, even during this hour, a large number of patients arrived in the first quarter of the hour (8-8:15am). Also, later in the day, about 25 patients were waiting at around 11:15am. At this rate, the 25th patient was examined between 1-1:30pm on average. However, at the pharmacy, it was observed that the crowd started building up at around 10am. It is less busy during early mornings, since they collect medicines in the end, after they have visited other stages (such as consultation with the GDMO/specialist). At around

10am though, the queue builds up just within a few minutes and a drastic increase in the number of patients was observed. In the early part of the afternoon, at around 12:30pm, it was observed that there were still about 40 patients waiting, and at this rate, the 40th patient was served at around 3pm. The pharmacy has less patients in the morning, as at that time only those patients arrive who are here just to collect medicines, either because the medicines were unavailable on the day of consultation or they did not have time. Hence, the main conclusion drawn from these observations is that the arrival pattern of patients varies from department to department. Also, the load of patients changes drastically just within few minutes during the busiest hours of the day. Therefore, there is a crucial need to 'measure' the level of overload before the queue builds up excessively.

The observations at the designated hospital also show that the consultation time did not last longer than 6 minutes on average, with GDMO and even with the specialist (as mentioned in Section 4.2.3). One of the reasons for low consultation as observed is that most of the patients require a regular check-up or follow-up or simply a renewal of medicine prescriptions, which takes less time and can be dealt by the doctors quickly enough. Only patients visiting for the first-time, requiring a detailed examination (children or nature of medical complaint) or having multiple complaints (mostly aged patients), require more time. However, such patients form a small percentage among the huge number of patients, and even in this case, the doctors do not spend more than 8-9 minutes on average as observed. Also, if the patients are seriously ill, the doctors advise them to proceed to the emergency department and get admitted. Additionally, the overcrowding of patients prevents the doctors from spending a long time on one patient, in order to ensure that all outpatients in a day are examined on the same day. Similarly, with regard to pharmacy, the average service time taken to provide medicines was observed as 4-5 minutes (as mentioned in Section 4.2.3). However, it was observed that it took slightly longer to serve elderly patients (men and women) if they require a large number of medicines or some additional explanation. Furthermore, there are some medicines which require special permission from the Head of pharmacy. In this case, the service time exceeds slightly at about 7-9 minutes. Therefore, it can be concluded that the low consultation/service time is a significant finding, and needs to be incorporated in the DEA model for precise representation of the existing patient flow system, and a more comprehensive evaluation of the queuing problem.

A significant finding from the observations of the existing queue system is the *inadequate scheduling* of *personnel* (GDMOs/specialists/pharmacists), which is more critical than the actual 'number'. For instance, the maximum load of patients is in the morning (7-8:30am) at GDMO clinic and one of the specialist OPDs, with at least 50 patients waiting at one time during the busiest times, however, it was observed that only 2 to 3 doctors are working when the session starts at 08:30am. The availability of doctors varies over the day. It was observed that around 10:30am and later at around 11am, the number of available doctors increases to four and five respectively. However, by this time, a long queue of patients is already built up, with most patients waiting since early morning (7-7:30am). Therefore, it was observed that the right number of doctors are not available at the busiest times of the day in all departments, leading to excessive wait times. The observations and brief interviews with the administrative staff/doctors demonstrated that in addition to conducting OPD sessions, the doctors have to carry out other tasks, including academic activities, training sessions, ward rounds and administrative work or paperwork related to ward rounds and OPDs. Almost all these activities are conducted during early morning, and the doctors arrive at the OPDs later on in the day. Since the OPDs experience extreme overloading during early mornings, by the time doctors arrive, there is an excessively long queue of patients which will require an additional number of hours. Hence, increasing the availability of doctors at the later part of morning seemed pointless. In case of the pharmacy though, there are only two counters operating 'all day' long for each category of patients (serving, retired, senior and women), irrespective of the variation in the load of patients throughout the day. For the pharmacy, the queue starts building up excessively between 10am-12pm generally, at the busy counters (serving, retired and senior) leading to long wait times for patients. Due to the absence of appointment systems, it is crucial to identify 'how' many doctors/pharmacists are required and 'when' so that the queue can be controlled quickly, before it builds up exponentially since any action at this point will not be very rewarding.

The initial findings obtained from observing the patient flow system of the public hospital in Pakistan are significant since they highlight the operational challenges in the queuing process, when appointment systems are non-existent. Furthermore, the queuing-DEA model can be further modified in the light of these *preliminary findings*, in order to provide a better assessment of the queuing system (see Figure 4-10 below).

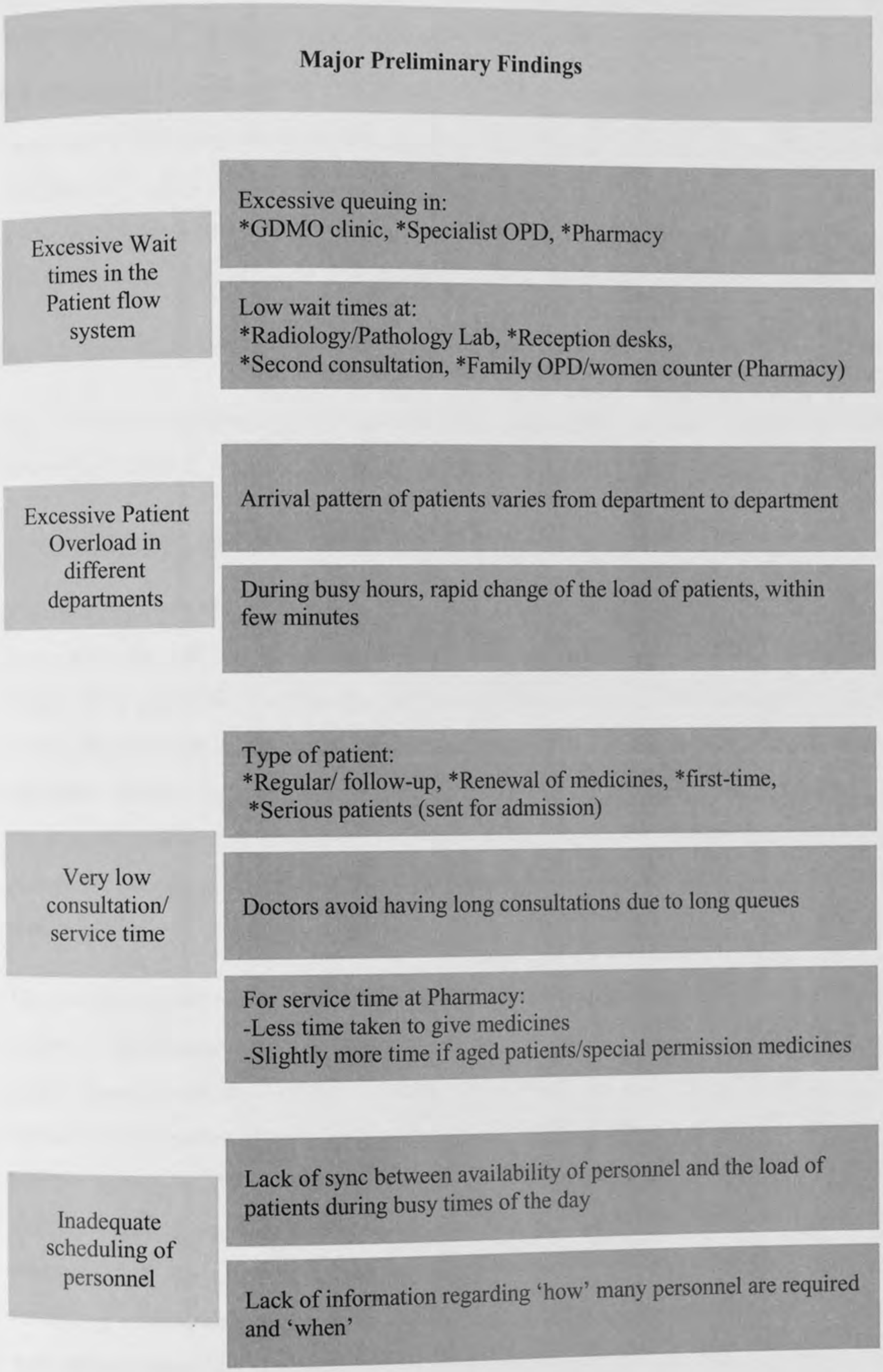


Figure 4-10: Summary of Preliminary Findings from Observing the Patient Flow System

4.2.4 'Refined' DEA Model for Assessment of the Queue System

One of the challenges in DEA modelling is the selection of appropriate input and output variables to obtain useful and valid solutions. The primary structure of the DEA model, real-time data and preliminary findings, are utilized to further *refine* the DEA model to better suit the purpose of the current study; which is evaluating the queuing system in the absence of appointment systems of a large public hospital in a developing country.

4.2.4.1 Characteristics of Refined Queuing-DEA Model

The different elements of the queuing-DEA model proposed are explained in the following sections.

Assessment of the Queuing System Using Patient Level Data in the DEA model

The objective of the current study is to evaluate the queue system, hence, *patients* are included in the DEA model as *'units of analysis.'* The utilization of patient level data is crucial as it allows to pinpoint the extreme wait times experienced by patients and the factors affecting these wait times. The data is noted for each individual patient for all four variables added in the DEA model, including consultation time, number of doctors/pharmacists, waiting time and length of queue. Hence, the queuing data for each patient will then be included in the DEA model, to assess the overall efficiency of the queue system.

Most of the queue management techniques utilize average values to assess the queue system as demonstrated by previous works (as discussed in Section 2.5). However, in public hospitals of developing countries, the arrival pattern of patients is extremely variable. The number of patients is not fixed due to the absence of appointment systems and the queue builds up quickly. Therefore, it is essential that the rapidly changing queuing situation is taken into consideration in order to provide a more comprehensive evaluation of the queuing situation. The efficiency assessment technique of DEA modelling is extremely useful in this regard. By comparing the queuing characteristics of each patient separately, DEA allows for the efficiency evaluation of the queue system at different times of the day in the busiest departments; in order to identify the critical issues affecting the queue situation and provide recommendations for improvement.

The current study will be among the first studies which employ patients as 'units of analysis', and assesses a queue system using DEA and that within 'one' public hospital. This analytical approach provide evidence that DEA can be extended in other applications such as a queue system using patient level data; rather than its more traditional usage of comparing the efficiency of 'hospitals' using aggregated values of inputs/outputs.

Input 1: Wait times of patients

As observed from initial findings, the absence of appointment systems is a significant factor which leads to variation in the arrival pattern of patients and lack of proper management of patients, further increasing the wait times. Therefore, there is a need to assess the efficiency of the queue system with regard to the *wait times* of patients which might vary at different times of the day. The current study aims to evaluate the queue situation using wait time as an input in the DEA model. This analysis will assist in identifying high wait times experienced by patients and the factors which affect them, so that appropriate measures can be undertaken to improve the existing queue situation.

Almost all prior queuing studies are associated with improving an existing appointment system. Therefore, the waiting times in these studies are defined with respect to the appointment times allocated to patients. The goal of the present study is to use DEA modelling to assess and improve the queue system of a busy public hospital in a developing country where appointment systems do not exist, and the patient arrival pattern is not consistent. Therefore, the waiting times are defined differently as compared to previous works since all patients are walk-in. In this case, the wait time is represented as the difference between the times a patient reports at the reception and when the patient enters the consultation room in case of GDMO clinic/specialist OPD, or when the patient arrives at the service counter at the pharmacy. Also, these wait times will be recorded for each individual patient to cater for the variability in the arrival pattern, as 'average' wait time might not be a suitable measure as utilized by previous studies (see Section 2.3 for discussion on queuing parameters). The preliminary observations and findings indicated that high wait times prevail in specific departments including GDMO clinic, the busiest specialist OPD and pharmacy. Therefore, for the current study, the DEA model is constructed for these busy departments.

Input 2: Length of Queue

The initial observations and findings show that due to absence of appointment systems, the overload situation changes drastically in busy departments and a massive queue builds up just within few minutes. Hence the unpredictability in the arrival times of patients with large number of patients arriving at the 'same' time during certain busy times, leads to painfully long wait times for all patients. Although the influx of patients cannot be controlled, however there is a need to measure the level of 'overload' of patients frequently, in order to highlight the extent of overcrowding within the system. This analysis can then be used to assist in deriving appropriate strategies to counter this overload, if possible. Hence, *'length of queue'* is the other input included in the DEA model as an 'indicator' of overload in the queue system.

Similar to wait times, the length of queue has been considered by various research studies, but with respect to improving an already existing appointment system, where 'average' length of queue is utilized to represent a queue system with prior appointments (see Section 2.3 for discussion on queuing parameters). As with the wait time, due to the extreme variability in the arrival pattern of patients, the average is not a suitable representative of the long queue prevailing in busy public hospitals of developing countries. Therefore, the current study aims at utilizing the DEA modelling where the length of queue will be noted each time a patient arrives. In this regard, the length of queue is defined as the total number of patients already waiting at the time a walk-in patient joins the queue. Hence, if more patients are waiting at the time of arrival of a particular patient, the longer will be the wait time of the current patient and subsequent patients. Therefore, the current study intends to utilize the length of queue when assessing the busiest departments at the designated hospital.

Output 1: Non-discretionary Output: Consultation time (GDMO/Specialist)/ Service time (Pharmacy)

For the current study, consultation time with the GDMO/specialist and the service time at the pharmacy are considered. The preliminary findings indicated that the consultation/service time at the designated hospital is quite low, specifically 6 minutes and 5 minutes on average for consultation with GDMO/specialist and service at the pharmacy, respectively. This result highlights that although this variable greatly influences the queuing situation as indicated by previous studies (see Section 2.3), it has less weightage in case of public hospitals of developing countries. With no appointment systems in place, the influx of patients is huge and unpredictable, therefore, the doctors examine patients quickly due to a long queue of patients waiting. Additionally, in most cases, the type of patients such as regular/follow-up require a quick consultation, where the serious patients requiring immediate care are transferred to the emergency department instantly.

Depending on the diagnosis and task involved, the patients are seen in a logical fashion, minimizing the chances of missing important and critical diagnoses. Also, the consultation time determines the level of quality of healthcare service provided and patient satisfaction. Therefore, it is perhaps in the best interest not to actively influence further reduction in doctor-patient interaction time. The consultation time is already quite low and if reduced any further, might increase the chances of errors. Similarly, the service time at the pharmacy is just adequate for 'safe' dispensing of medicines and perhaps should not be curtailed any further. On the other hand, efforts should be made, not to allow the consultation/service time to increase either, since additional time spent on one patient will lead to rapidly increasing wait times of subsequent patients. Therefore, for the purpose of this study and considering the observations and preliminary findings, the consultation/service time should not be allowed to either increase or decrease. However, there is a crucial need to utilize this variable to observe its effect on other queuing factors as well as the overall efficiency of the queue system, irrespective of whether it causes a major or minor impact on results. Hence, the *consultation/service time* has been included in the Queuing-DEA model as a *'non-discretionary'* output.

In DEA modelling, a 'non-discretionary' variable is defined as an exogenously fixed variable, which is beyond the control of decision-making units (DMUs). For instance, with regard to inputs, weather conditions for farming or air-force wings, demographic or socio-economic factors when comparing services within various countries (population, education, income and others), or even additional data (number of competitors, intensity of competition, unemployment rate) when evaluating organisations (banks, fast-foods and others). With reference to outputs, some examples can be checking cash transactions in a bank which is directly related to the gratis service, or army recruiting which depends on personal career options (Cooper *et al.,* 2011; Bank and Morey 1986; Golany and Roll

1993). The target values obtained for exogenous variables may not provide useful conclusions since they cannot be influenced by the management. However, the idea is to deduce the efficiency score and the expected increase (decrease) in the level of outputs (inputs) considering the discretionary variables, while keeping the uncontrollable variables fixed (Banker and Morey 1986).

The concept of *non-discretionary* variables has been applied in a few studies, mainly to provide a more robust analysis. Considering healthcare, Afonso and Aubyn (2006) assessed the health services in various Organization for Economic Cooperation and Development (OECD) countries using a two-stage method. The first stage consisted of a DEA analysis, whereas the second-stage included Tobit regression, and single and double bootstrap procedure. The efficiency scores in the first stage were regressed against some explanatory variables in the second stage analysis including Gross Domestic Product (GDP), education, smoking habits and obesity. The authors concluded that these variables have a 'significant' effect such that higher level of GDP and education led to better health services, whereas, smoking habits and obesity worsen the health performance. Some other non-discretionary variables (such as income, health expenditures, population, eating habits and others) were also added but none of these variables were statistically significant. In another healthcare application, Herwartz and Schrumann (2014) compared the efficiency of hospitals before and after the introduction of DRG-based (Diagnosis-related groups) financing system in Germany, included number of beds as a non-discretionary input. Additionally, other explanatory variables were included to further explain the efficiency scores including 'profit' or 'non-profit' private hospitals, market share, occupancy rate, mortality rate, hospital budget, people aged above 65 years and others. In this case, both parametric and non-parametric techniques of DEA and Stochastic Frontier Analysis (SFA) respectively, were used to assess the effect of non-discretionary variables. Both analytical approaches indicated a decrease in the efficiency level of hospitals. Mitropoulos *et al.,* (2013) used a combination of location allocation models and DEA in order to evaluate the performance of health centres. In this case, treatment population (number of inhabitants assigned to facilities) was used as a non-discretionary input. The authors claimed that although this variable is not under the control of management, however, it was considered to have a huge influence on the performance of health centre since services at a short distance had more chance of being utilized. Therefore, the variable and the two techniques used jointly, helped in identifying suitable targets for the operation of health centres, hence improving efficiency.

Considering other applications of DEA modelling with non-discretionary inputs, Cordero-Ferrera *et al.,* (2010) conducted a comparison of schools. Although eleven variables were considered as non-discretionary inputs, however, using principal component analysis, they were combined to three main variables. These factors were categorized as the student's family and socio-economic background, abilities and attitude. A four-stage model was used consisting of DEA, Tobit regression and bootstrap procedure. However, in the second stage, instead of efficiency scores, the slack values of the two outputs were regressed on the explanatory variables. Using slack predictions, new input and output levels were calculated and the DEA model was re-run to obtain new efficiency scores and more realistic targets. In a recent study, Kyriacou *et al.,* (2015) evaluated the efficiency of various developed and developing countries over three decades in terms of redistribution of funds, using DEA and panel truncated OLS regression analysis. The second stage analysis was utilized to evaluate the impact of non-discretionary variables related to economic development, government quality and demographic features. The authors concluded that rich countries redistributing income through fiscal policies, had higher efficiency, whereas, higher asset inequalities reduced redistributive efficiency for some nations. Furthermore, aged cohort had more weightage than young cohort showing that pubic pensions and health expenditures have a stronger impact on redistributive efficiency than education spending. In a study by Gomes and Lins (2008), a non-discretionary output was considered rather than input. Among the three outputs included, population was treated as exogenously fixed whereas total energy consumption and GDP were considered as discretionary outputs. The study used a zero-sum gains DEA model to consider the scenario of CO_2 emission reallocation among various countries, where this variable was considered as an undesirable output modelled as an input. Similarly, Saen (2009) conducted a study to evaluate supplier efficiency and ranking, and considered supplier variety as a non-discretionary output, along with R&D expenditures as a discretionary output and cost as input.

There are very few studies which have utilized the concept of non-discretionary variables under DEA modelling, and almost all of them have included exogenously fixed inputs. There are negligible studies which considered non-discretionary outputs. Furthermore, most of these studies have considered a second-stage analysis using regression analysis to evaluate the effect of these variables. Additionally, very few applications can be found in the field of healthcare. The selection of non-discretionary variables immensely depends on the objective of the study. Although the variables cannot be influenced by the DMUs, however, they assist in providing better explanation of the efficiency scores. For the current study, the consultation/service time variable will be considered along with other queuing factors to evaluate the efficiency of the queuing system, but without causing any change in it. Therefore, the consultation/service time is considered as a 'non-discretionary' output in the proposed queuing-DEA model.

Output 2: Number of Personnel (GDMOs/Specialists/Pharmacists)

As indicated by the initial findings, the inadequate scheduling of personnel is another significant factor which further worsens the queue situation. The lack of sync between the availability of staff and excessive queuing is a crucial issue which needs to be addressed. In an ideal situation, more doctors can be employed and number of clinics enhanced. But with limited staff, along with excessive queuing, it is essential to determine the *'optimal'* utilization of resources, using certain criteria. Additionally, due to absence of Appointment Systems, the inflow of patients cannot be controlled since the number of patients arriving in a day is *not* 'fixed', hence creating hurdles in organizing the walk-in patients effectively. The key is to know the exact number of personnel required. This crucial information will allow for the doctors/pharmacists to be rotated 'quickly' from less urgent duties to busy areas, if possible, when wait times become excessive. Therefore, the high level of wait times and overload in the system (length of queue) can be used to spot the discrepancy in the doctor/pharmacist availability at that particular time. With the right scheduling at the right time, the patient care will improve in terms of reduced waiting time and provision of prompt service, leading to increased patient satisfaction.

The issue of excessive queuing and inappropriate scheduling has a major bearing on the development and orientation of the DEA model proposed. In this case, the objective is to *'increase'* the availability of doctors given excessive queuing, hence resulting in a reduction of wait times for subsequent patients. In almost all queue management and DEA studies, the number of personnel is one of the most widely used variable where the objective is to 'minimize' the number of staff in order to improve the efficiency of a health institution, with emphasis on optimal utilization of resources. However, in the current study, the objective is to reduce the excessive waiting times of patients. The administrators require an 'interactive' and 'flexible' framework which is responsive to the varying wait times and load of patients at any given time of working hours within busy departments, so that the manpower can be optimally utilized leading to the minimization of the queue. Therefore, from the perspective of the current study, the wait time and length of queue are added as *'inputs'* in order to determine the 'required' number of doctors/pharmacists; hence the *number of personnel* are included as an *'output'* in the DEA model.

Almost all previous DEA works (as discussed in Section 3.2.1) have the objective of comparing the efficiency of a number of hospitals, where the number of doctors represents the total number of doctors working at each hospital in the sample. However, the goal of the current study is to identify the efficiency of the queue system at different times of a day in the outpatients' department within 'one' busy public hospital of a developing country, using patient level data. The availability of personnel varies from department to department. For instance, the number of doctors working at the GDMO clinic/specialist OPD varies throughout the working hours of the day, whereas, the number of pharmacists at the pharmacy are fixed at each counter. Therefore, in this case, the number of doctors (for GDMO clinic and the specialist OPD) and pharmacists (for the pharmacy) was recorded when each patient arrived. This information provides an opportunity to evaluate the current availability of personnel and to determine the 'required' number to reduce the wait time of subsequent patients.

In developing countries, where there are no pre-booked appointments and no definite control on the variable influx of patients; optimal utilization of available staff in a quick dynamic fashion is vital, to reduce the excessive waiting of patients rapidly enough. With the aid of DEA modelling, the varying wait times and length of queue will be utilized to identify the *'required'* number of doctors to improve the efficiency of the queue system. This essential information will allow the management to ensure that more personnel are available at busiest times whereas they can be re-assigned to other tasks when it is not busy to avoid wastage of resources.

4.2.4.2 Summary of the DEA Model Proposed

The current study aims at evaluating the queue system of a large busy public hospital in a developing country and propose a framework for improving the queue system, using DEA modelling. The preliminary observations and *findings* indicate that the *inappropriate scheduling* of personnel is a major cause of long wait times of patients in busy public hospitals of developing countries. The right number of personnel (doctors/pharmacists) is not available at the right time hence leading to excessive queuing. Currently, there are no definitive strategies in place which can determine the exact requirement of personnel. Therefore, it is critical to have a set framework which could help liaison the variable arrival pattern of patients due to absence of prior appointments, and the availability of personnel. Henceforth, the *Queuing-DEA model* is developed such that it assesses the patient flow system and assists in preparing an appropriate strategy with respect to the scheduling of doctors/pharmacists; to deal with the overcrowding and minimize wait times. With this perspective, the number of doctors/pharmacists has been included as an output. The two significant determinants of the queue system namely wait time and length of queue are included in the DEA model as inputs; to determine the *'target'* number of doctors/pharmacists and diminish the queue instantly. Therefore, an 'output-oriented' DEA model has been run to suit the purpose of the study. Additionally, the consultation/service time has been added in the model as a 'non-discretionary' output. Being already quite low, this variable cannot be minimized any further and it cannot be allowed to increase either as this will further increase the wait times of patients. Hence, the consultation time will be included to observe its effect on other queuing variables and the overall efficiency of the queuing system at a particular time, however, without allowing it to change. Furthermore, *patient level data* has utilized, such that the value of each queuing variable included in the DEA model has been recorded for each patient to determine the efficiency of the queuing system at different times. The input and output variables included in the proposed DEA models for the three busiest departments are shown in Figure 4-11 below.

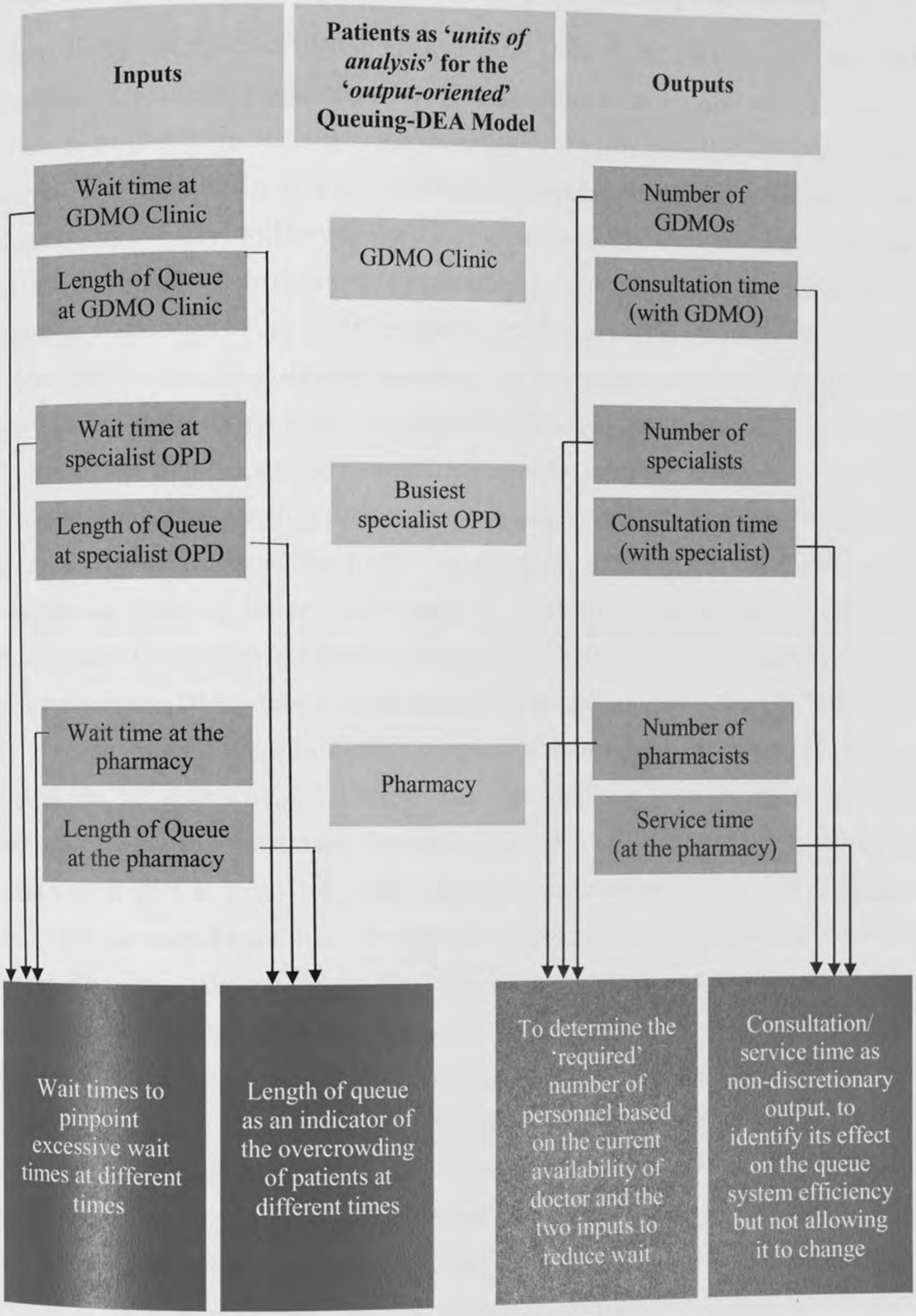


Figure 4-11: Proposed DEA Models for the Current Study

4.2.4.3 Mathematical Representation of the Proposed Queuing-DEA Model

The mathematical representation of the traditional DEA model has been discussed and shown in Section 3.1.2. For the current study, the traditional DEA model will be modified to suit the purpose of the study, which is construction of a DEA model to evaluate a queue system. Due to inclusion of a non-discretionary variable, the DEA model with non-discretionary variables will be employed for the current study. Over the years, a number of DEA models have been developed by researchers which incorporate non-discretionary variables. Banker and Morey (1986) propose a modification in the constraints of the DEA model to include exogenously fixed variables. The constraints for discretionary and non-discretionary variables are separated, where the efficiency component is included only for constraints with discretionary variables. Therefore, separate models are used which incorporate either the non-discretionary inputs or outputs depending on the orientation of the model. However, Golany and Roll (1993) developed a DEA model which allowed for simultaneous handling of non-discretionary inputs as well as outputs within the same model. Also, they proposed a model which deals with partial non-discretionary factors, which shows that DMUs have limited control over the factors and can only be influenced to a certain level. Ruggeiro (1998) provides a comprehensive review of various approaches to deal with fixed exogenous variables, and also compares multiple techniques using simulated data. The author highlights one-stage as well as two-stage models to handle exogenous variables. The two-stage models incorporate regression analysis in the second stage to test the effect of non-discretionary inputs on the efficiency scores. The author also provides a discussion on the strengths and weaknesses of each model. The Ray (1991) model uses regression which requires a-priori specification of the functional form whereas Ruggeiro's (1996) model is constructed such that if the number of non-discretionary inputs increases, efficiency will be over-estimated. The author proposes a three-stage model where the efficiency scores are re-calculated by running a third-stage linear program. It is concluded that this model overcomes the weaknesses of the models proposed by Ray (1991) and Ruggeiro (1996). Muniz *et al.,* (2006) also provide a detailed evaluation and comparison of techniques with non-discretionary variables. However, this review extends the one provided by Ruggeiro (1998) by including two additional models developed by Muniz (2002), and Yang and Paradi (2003). The authors conclude that although Ruggeiro (1998), Muniz (2002) and Yang and Paradi (2003) perform better than the Banker and Morey (1986) models, however they have drawbacks. The Ruggeiro (1998) model has the assumption of a priori functional form for the regression, Muniz (2002) considers DEA only but as number of variables increase the results might become inaccurate and Yang and Paradi (2003) model has a tendency of overestimating efficiency. Therefore, the authors suggest that each method has its own advantages and disadvantages, and recommend using alternative models to test for robustness.

The Banker and Morey (1986) model is the most frequently used model in the literature (Muniz *et al.,* 2006). The main benefit of this model is that non-discretionary and discretionary variables can be introduced simultaneously in the same DEA model, unlike two-stage models. In these models, discretionary variables are included in the DEA model followed by regression analysis to identify the effect of non-discretionary variables on the efficiency scores (Ishizaka and Nemery 2013; Saen 2009). Additionally, the model has computational advantages as compared to other sophisticated models. Although Banker and Morey (1986) model can include either a non-discretionary input or output, but for the purpose of this study, this model seems appropriate, since consultation time is the only non-discretionary output being considered.

Hence, using the modified version as proposed by Banker and Morey (1986) for an output-oriented model, the *Queuing-DEA model* developed for the current study is as follows:

Queuing-DEA Model For the Current Study		
Maximize $\phi + \varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{rEO_D}^{s} s_r^+)$		Output-oriented DEA model
subject to		
$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io}$	$i = 1,\ldots,m$	Two inputs $[i = 1, 2]$: Wait times and length of queue
$\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = \phi y_{ro}$	$r \in O_D$	One discretionary output $[O_D=1]$: Number of personnel
$n\sum_{j=1} \lambda_j y_{rj} - s_r^+ = y_{ro}$	$r \in O_N$	One non-discretionary output $[O_N=1]$: Consultation time
$\lambda_j \geq 0$	$j = 1,\ldots,n$	

Therefore, the above model shows an *output-oriented* Queuing-DEA model with a *non-discretionary output*. In this case, the objective is to 'maximize' the output given the number of inputs, without causing any change in the non-discretionary output. For the current study, there are two inputs namely wait times and length of queue, and number of personnel and consultation time is the discretionary and non-discretionary output respectively. There are two main differences from the traditional DEA model. Firstly, the constraints are different. The constraint for the non-discretionary output is separate which does not consist of the efficiency variable ϕ since efficiency level will not be affected by this variable as it is kept 'fixed' in the model. The efficiency level ϕ only appears in the constraint of discretionary output which is number of personnel. Secondly, the slack values of non-discretionary output will not be included in the objective function, that is, only the non-zero slack values of the discretionary variable will be added with the efficiency scores in the objective function. Furthermore, the target value for the non-discretionary output will not include the efficiency variable ϕ and hence, will remain unchanged (Banker and Morey 1986; Cooper *et al.,* 2011). Only the target value of number of personnel will be observed and evaluated, in order to identify the required number of personnel (doctors/pharmacists).

The proposed queuing-DEA model as represented above, will be utilized to fulfil the objective of determining optimal scheduling of personnel; in a rapidly changing queue situation, where all outpatients are walk-in.

4.2.5 Analysis: DEA Results and Construction of Dynamic Framework for Implementation

The proposed Queuing-DEA model as explained in Section 4.2.4 has been utilized to assess the queue system of three departments at the designated hospital, including GDMO clinic, busiest specialist OPD and pharmacy. The DEA analysis was carried out in Performance Improvement Management (PIM-DEA: Version 3.2)) software (PIMsoft 2014).Furthermore, for practical implementation of the proposed model, a dynamic framework has been developed in Excel (Excel 2013). The dynamic framework has been validated using a simulated dataset ensuring its smooth glitch-free running (Details of construction and validation aspects of dynamic framework are given in Section 5.2.2.).

4.2.6 Summary of Different Stages for DEA Model Development

The development of the Queuing-DEA model proposed in the current study consisted of several stages. The first stage involved constructing an initial DEA for queue assessment. In order to identify the extent to which these queuing factors affect a queue system without appointments, a few observations and interviews were conducted within the outpatients department at the designated hospital in Pakistan, which was the second stage. These preliminary observations/interviews mainly highlighted the variable arrival behaviour of patients along with overcrowding, long wait times and poor allocation of staff with high wait times. Therefore, the second-stage DEA model consisted of wait time and length of queue as inputs, with number of doctors as an output. The third stage consisted of validating the second-stage DEA model. Based on results obtained from the validated DEA model, 'focused' data collection plan was initiated, which was the fourth stage. The 'focused' data collection consisted of first observing the multiple patient pathways for outpatients; followed by observations regarding queuing variables (including wait times, length of queue and number of personnel, as well as consultation time), and interviews to develop a better understanding of the quantitative results and operational aspects of the queue system. The fifth stage consisted of identifying major findings based on real-time queuing data gathered through observations and interviews, where the most crucial finding was the inadequate allocation of personnel. Additionally, the consultation time was found to be extremely low, as well as high patient arrival rate and high wait times. The sixth stage comprised of developing a 'refined' DEA model based on these preliminary findings, such that the Queuing-DEA model is output-oriented. The model determines the exact requirement number of personnel to reduce queuing hence including current number of personnel as an output, with wait time and length of queue as the two inputs; while keeping the consultation time as a non-discretionary output (details provided in Section 4.2.6.3). The proposed Queuing-DEA model is then applied to the three busiest departments including GDMO clinic, busiest specialist OPD and pharmacy, hence, the seventh stage was to obtain DEA analytical results for these three departments using PIM-DEA software. The eighth and last stage was associated with developing a dynamic framework in Excel for practical implementation of the proposed model, for constant monitoring of the queue system of outpatients in the absence of appointments (see Figure 4-12 below).

Stage 1 → Initial DEA Model

Stage 2 → Few Preliminary Observations/Interviews and Second-stage DEA Model

Stage 3 → Validation of Second-stage DEA Model

Stage 4 → Focused Data Collection (Patient Flow System, Observations/Interviews and Ethical Considerations)

Stage 5 → Preliminary Findings Obtained from Focused Data Collection

Stage 6 → Refined DEA Model for Queue Assessment Obtained from Preliminary Findings

Stage 7 → Analytical Results Obtained by Applying the Proposed Model at the Three Departments at the Designated Hospital

Stage 8 → Construction of a Dynamic Framework for Practical Implementation of the Proposed Model

Figure 4-12: Summary of Different Stages for DEA Model Development

Chapter Summary

The current chapter elaborates upon the methodological features adopted in the current study to conduct research, consisting of two main parts which includes the Research Design and the DEA Model Development. Considering the former, the philosophical underpinnings of the current study are highlighted in terms of ontology, epistemology and the approach undertaken. The current study follows the ontological and epistemological position of Objectivism and Positivism respectively, with a Deductive approach. The present study utilizes the theoretical underpinnings and factors of a queue problem to develop a DEA analytical model in order to specifically assess a queue system in a large public hospital of a developing country in the absence of appointments for outpatients. Therefore, an empirical investigation has been conducted using quantifiable data to derive results, which will in turn extend the application of DEA modelling in the context of a queuing problem. Furthermore, the case study research has been adopted for the current study, considering the designated hospital in Pakistan as a 'typical' case, as this hospital is an exemplar which depicts the queue situation of almost all public hospitals in developing countries with issues such as shortage of staff, increased workload for existing staff and patient overload. Furthermore, among different case-study designs, 'Single case with an Embedded design' has been utilized, since the objective is to assess the queue system using patients as units of analysis within the designated hospital representing multiple sub-units with a single case.

Additionally, a cross-sectional research design has been adopted since the queueing data has been collected in one point in time which is suitable for the purpose of the study. Although a mixed method design has been adopted, however, the quantitative analysis is the dominant one. Hence, the qualitative results are used to 'complement' and further explain the quantitative results. Therefore, among various mixed method designs, the Explanatory Sequential design has been considered where the quantitative research is the main research strategy. It is followed by collecting qualitative data to further support the quantitative results, and assisted in refining the DEA model developed for the current study. Finally, considering data collection techniques, secondary data was collected using documentation which mainly included the annual statistical reports which provided information about the overall structure of the organization, patient overload and staff shortage. With regard to primary data, the quantitative data regarding the pathway followed by patients and wait times were recorded using non-participant structured observation. Additionally, interviews with staff members were conducted to further complement the results obtained from observation, and develop a better understanding of the patient flow and the standard procedures at the designated hospital.

Considering the latter part of the present chapter, the initial DEA model was developed and validated, which provided a basis for 'focused' data collection at the designated hospital in Pakistan. The 'focused' data collection process firstly included observation of the patient flow system at the designated hospital, in order to identify the different stages that a patient follows from entry till exit, the different types of patients (such as first-time or regular) and the specific departments with excessive waits. Based on the initial DEA model and the identified characteristics of the patient flow system, the data collection process consisted of a time sheet for structured observation of wait time, length of queue, consultation time and current availability of personnel (doctors/personnel) for all patients. Interviews with staff members comprised of questions regarding the time-consuming procedures and factors leading to queuing as well as any suggestions for improvement to reduce wait times. Some ethical issues have also been considered when conducting data collection. Firstly, ethical approval for data collection has been obtained from the designated hospital in Pakistan and Aston University. Some other ethical considerations include maintaining the anonymity of the designated hospital, all respondents and volunteers involved in data collection process, as well as obtaining informed consent from them and the head of departments. Considering storage and disposal of data, all data gathered is password-protected and has only been shared with academic supervisors. Regarding dissemination of data and results, the derived results and propositions can be presented at various platforms including conferences, seminars and journal publications. Furthermore, the results and recommendations will be shared with the administration of the index hospital, if desired, especially when considering the implementation of obtained results and propositions.

The initial framework of the queuing-DEA model and real-time data led to some useful findings. The most significant finding was the inadequate scheduling of doctors. Also, the consultation/service time was found to be extremely low. Some other findings include excessive waits at GDMO, specialist and pharmacy departments only, extreme overload of patients and high variability in the arrival pattern of patients. These preliminary findings highlighted the operational challenges in the absence of appointments, and were used to develop a refined DEA model. Considering the most crucial finding of inappropriate staff scheduling, the Queuing-DEA model was refined such that it determines the 'required' number of doctors, before the queue builds up excessively. For this purpose, the wait times and length of queue were included as inputs, where the latter is an indicator of the level of overcrowding. Additionally, consultation time already being quite low, will be included as a non-discretionary output, where it will added to observe its effect on the overall efficiency of the queue system without changing it. There are very few studies in healthcare using a DEA model with non-discretionary variables, and among those, there are negligible studies which have considered non-discretionary 'outputs'. Different DEA models have been proposed over the years, however, the current study aims to utilize the Banker and Morey (1986) model to include non-discretionary output for an output-oriented DEA model. The model has an additional constraint for the 'fixed' output, which does 'not' include the efficiency level component as in the case of discretionary output. Also, the slacks included in the objective function are only for discretionary outputs. The target values for the non-discretionary output remains unchanged. Patient level data has been utilized for the DEA model proposed where the queuing data will be noted for each patient 'separately', to provide a more comprehensive analysis of the changing queuing situation. Therefore, this refined DEA model will be further utilized to determine the optimal level of personnel to improve the queuing situation.

CHAPTER 5 RESULTS AND IMPLEMENTATION

Chapter Overview

The current chapter is divided into two main sections of DEA analysis and results, and the practical implementation of the proposed model. The chapter concludes with highlighting some additional issues and recommendations for improving the overall queue system.

The first-half of the chapter will elaborate upon the DEA analysis and results obtained from applying the proposed Queuing-DEA model in the three busiest departments of the designated hospital. The segment will commence with highlighting the considerations taken into account before conducting further analysis. The next three sub-sections will provide an in-depth evaluation of the DEA analytical results for the three departments including the busy specialist OPD, GDMO clinic and pharmacy. The section will conclude with providing a summary and usefulness of the proposed DEA model and the results obtained, with respect to evaluation of a queue system within large public hospitals of developing countries with no appointments.

The second-half of the chapter will be dedicated towards developing a dynamic framework for practical implementation of the proposed DEA model and results obtained. The first section will emphasize on elaborating upon the different operational features of the dynamic framework for queue assessment. The next section will provide an explanation of the technical aspects regarding the development of the Queuing-DEA model within the dynamic framework. The segment will conclude with demonstrating the usefulness of the flexible dynamic framework for continuous assessment of the queue system with unpredictable patient arrivals and no appointments within public hospitals of developing countries.

Lastly, the chapter will conclude with pinpointing some additional issues which were observed or highlighted by staff members during interviews, which can be taken into consideration when developing any long-term plans for better queue management.

Based on preliminary observations and results, the General Duty Medical Officer (GDMO) clinic, pharmacy and the busiest specialist outpatients' department (OPD) had extremely high wait times. Therefore, the Queuing-DEA model developed in Section 4.2.6.2, has been applied within the designated hospital in Pakistan; to pinpoint the excessive wait times and extreme variability in the arrival pattern of patients, due to absence of appointment systems. Hence, the main objective is to identify the exact requirement of personnel, by evaluating the wait times and length of queue, and including consultation/service time as a 'non-discretionary' output. This information is extremely significant since it acts as a guideline for the management to determine an appropriate scheduling of personnel at frequent intervals to control the queue 'quickly', in order to cater for imminent long queues and overload.

5.1 DEA Analysis and Results for the Busiest Departments at the Index Hospital in Pakistan

In order to conduct a detailed analysis of the excessive queuing of walk-in patients, the proposed Queuing-DEA model has been applied to the three busiest departments in the designated public hospital of Pakistan. The extent of elaboration of DEA analysis and results for all three departments has been considered in the current study. As per previous DEA studies, the analysis has been most frequently subjected to identifying the fully efficient (100%) and inefficient (less than 100%) units. Depending on the objective of the studies, this differentiation was considered adequate in order to determine the level of improvement required in the 'inefficient' units. However, for the current study, it must be emphasized that due to variability in the queuing characteristics and the changing queuing situation within few minutes, mere differentiation between efficient and inefficient units is not sufficient. The foremost objective of the present study is to determine the 'required' number of doctors for nearly all observations to provide a more comprehensive evaluation, and facilitate understanding of the operation of the queue system with no appointments. In this regard, the efficiency value provides supplementary information to further support the results obtained for the target number of personnel, as well as indicating the overall efficiency of the queue system at the time 'each' patient arrived.

5.1.1 DEA Analysis and Results: 'Required' Number of specialists at the Busy Specialist OPD

Among different specialist OPDs, one of the OPDs was observed to have excessive overloading and wait times of patients. Some patients knowing that there is mostly a long wait time to see the specialist, arrive as soon as the doors open even though the start time of the OPD session is at least an hour later. Hence the queue is always long and waiting times are unimaginably high which is further compounded by less number of specialists, with even fewer at the start of the clinic. The DEA model proposed in Section 4.2.4 has been utilized to determine the 'required' number of specialists when analysing the queue system at the *busy specialist OPD* as shown in Table 5-1.

Table 5-1: Inputs and Outputs for the Queuing-DEA Model for the Busy Specialist OPD

Inputs	**Outputs**
Wait time for specialist OPD	Number of specialists
Length of Queue for specialist OPD	Consultation time when examined by the specialist **(Non-discretionary Output)**

The *availability* of doctors varied in the specialist OPD during the working hours due to other tasks that need to be performed by the specialists, in addition to conducting OPD sessions. As verified by the interviews with the specialist OPDs administrators, these tasks involve academic activities such as training sessions, ward rounds and other administrative responsibilities or paperwork. Therefore, depending on the required amount of time spent on other activities, they arrive at the OPD at varying times. The dataset shows that the *availability* of doctors varies from 2 to 5. From a total of 160 observations, the current availability was 2 was 42 observations, 3 for 24 observations, 4 for 90 observations and 5 for only 4 observations (see Figure 5-1 below); demonstrating that the availability varied throughout the day/week.

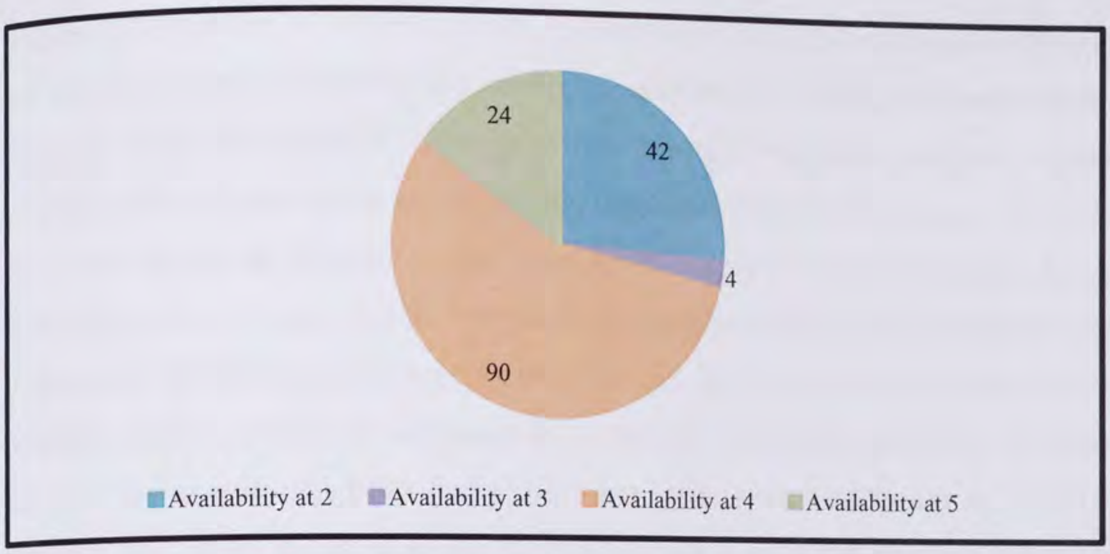


Figure 5-1: Number of Observations at Different Current Availability of Specialists in the Busy Specialist OPD with a Total of 160 Observations

From the total sample size of 160 observations, 7 are *100% efficient.* All of these observations had the availability and target number of doctors as 4 (since the efficiency is 100%). For three of these observations, the wait time was 1h, the length of queue was between 12 and 18, and the consultation time varied from 2 to 5 minutes. For the other 3 observations, the wait time was 1.5h. In this case, although the length of queue varied from 13 to 14, however the consultation time was higher, at 7 to 8 minutes. The last observation was obscure as compared to the others in the sense that the wait time and length of queue were quite high as compared to other observations (2.5h and 38 respectively), and the consultation time was quite high as well (13 minutes). Additionally, this observation is the efficient peer to only one inefficient data point in the whole sample. Therefore, it can be concluded that this observation, although being 100% efficient, might be not truly representative of the efficient queue situation, as observed from the other results. The two observations (P065 and P140) which were peers to maximum number of observations showed that with 4 doctors, the wait time should be between 1h and 1.5h and queue length should vary from 12 to 16, keeping the consultation time unchanged. Therefore, according to the DEA model, this combination of queuing variables represents an 'ideal' situation. Of course, there is a chance that due to various factors, this might not be achieved at all times, however, this target acts as a guideline which the department administrators can aim for, in order to reduce excessive queuing.

Although the availability of *4 doctors* represents a fully efficient queue system, however, as the queuing characteristics worsen, the 'required' number of doctors increases drastically. Therefore, a further analysis of observations with lower levels of efficiency and varying current availability of doctors is conducted. Considering the current number of doctors at 4, observation P157 (with an efficiency of 76%) had a wait time of nearly 2h (107 minutes) with queue length 16, and consultation time of 7 minutes. Therefore, due to increased wait time the 'target' number of doctors increased to 5, given that the consultation time is not altered. For another observation P173, the length of queue increased to 20 although wait time was still around 2h (126 minutes). In this case, the required number of doctors increased to 6, and the efficiency was 62%. In another instance, the observation P178 had an increase in the wait time to around 2.5h (134 minutes), the queue length increased slightly as compared to the previous observation which was 22, leading to an increase in the required number of doctors to 7 with efficiency level of below 60% (58%). For P107, the required number of doctors went further up to 8 with efficiency level 49%, since the queue length increased to 28 although the wait time remained at 2.5h (154 minutes). Some observations portrayed that the required number of doctors is even higher than 8, given extremely high values of either wait time or length of queue or both. For instance, P034 had high wait time and length of queue of 3h (173 minutes) and 41 respectively, therefore, the required number of doctors increased to 9 (consultation time was 8 minutes). For P172, the length of queue was low at 29, but the wait time was observed as around 3.5h (205 minutes) hence, the required number of doctors increased to 10. Similarly, for P042, the wait time has increased to an excessive value of 4h (239 minutes) with a length of queue of 34, leading to a further increase in the target of 11 doctors. For P006 (34% efficiency level), the wait time was 3.5h (209 minutes), with a very high length of queue of above 50 (55). Therefore, considering the extremely large number of patients waiting to be served at the time this patient arrived, the required number of doctors increased to 12. In case of P007, both queue variables had extreme values with wait time at nearly 4h (233 minutes) and length of queue of 58, therefore, leading to an extremely high requirement of 13 doctors.

The results were also evaluated when *2 doctors* were available. For P079, the wait time was around 1.5h (104 minutes) with length of queue of 16. Although the wait time is low but there are many patients waiting in the queue before this patient, leading to the required number of doctors of 5. Compared to P079, the values of P008 had high values for both inputs, that is, wait time was around 2h (128 minutes) and length of queue was 32. Hence, the model showed that 6 doctors are required. For P053, although the length of queue was

22, however, the wait time increased to nearly 2.5h (144 minutes), therefore maintaining the consultation time of 5 minutes, the required number of doctors increased to 7. For P012, however, although the wait time was maintained at about 2.5h (157 minutes) but the length of queue was 39, therefore, 8 doctors are required as shown by the model. For P025, both wait time and length of queue have high values of about 3h (167 minutes) and 42 respectively, leading to a further increase in the requirement of doctors of 9. The target number of doctors increased to 10, when the wait time observed for P143 was nearly a high 3.5h (222 minutes), even though length of queue was comparatively lower at 29.

Having an availability of *doctors at 3*, for P132 and P134, there was only a slight difference in the length of queue (30 and 29 respectively), however, the wait time of P134 was 3h (180 minutes) and was even higher for P132 at nearly 3.5h. Hence the number of required doctors was 9 and 10 for P134 and P132 respectively. For P015, the wait time was around 3.5h (198 minutes), however, the length of queue was extremely high of 50, pushing the requirement of doctors to 11. The consultation time for all three cases was 6-7 minutes which remains unaltered.

A maximum availability of *5 doctors* was observed at the busy specialist OPD. However, even with such a high availability, at most instances the wait time was observed as 3h or above. This highlights the crucial issue of the incorrect 'timing' of bringing in more doctors. At this time the queue has 'already' reached to an excessive level, therefore, even a high availability is ineffective at the moment and will take another few hours to bring down the queue. Hence, the model shows an extremely high requirement of doctors to diminish this queue. For P130 and P041, the wait time is around 3h for both (167 and 171 minutes respectively), however, the length of queue is 24 and 43 respectively. Therefore, the required number of doctors is 8 and 9 respectively. For P085, the length of queue is low at 29, however the wait time is nearly 3.5h (211 minutes) with consultation time of 6 minutes, therefore, the target further increased to 10. Having the same wait time of almost 3.5h (211 minutes) but a high queue length of 53 patients, the required number of doctors increased to 11 when observing P035. For P096, the wait time is even higher at 4h (229 minutes) with length of queue at 45 hence, increasing the required number of doctors to 12.

Therefore, it can be *concluded* that with current availability of doctors as 2 to 3, on average, a change in the wait time and queue length of 13 to 24 minutes and/or 10 to 20 respectively; led to one additional increase in the requirement of doctors. However, when the current availability is 4 to 5, on average, the target increased by one for change in the wait time and queue length of 14 to 40 minutes and/or 8 to 24 respectively. Furthermore, it was also observed that overall, extremely high values of wait time and length of queue, or even one of these factors, resulted in a massive increase in the requirement of doctors, irrespective of the current availability of doctors. Hence, overall, wait time of nearly 3 hours and above or with length of queue of around 30 and above, or both simultaneously; lead to a drastic increase in the target number of doctors between 9 and 13. Although the required number of doctors as shown might not be practically suitable, however, it does highlight the criticality of the queuing situation. Therefore, there is a necessity of an effective intervention which can control the queue quickly, before it builds up exponentially resulting in such high target values of doctors.

Figure 5-2 shows the values of the wait time, length of queue and the corresponding 'required' number of specialists, for 'each' of the 160 observations, where values for P065 and P140 (100% efficient units) are represented in red circles. For example, considering P065, the values of required number of specialists, wait time and length of queue are, 4, 73 minutes and 16 respectively, whereas values for any other data point say P041 are 9, 171 minutes and 43 respectively. Hence, Figure 5-2 demonstrates a highly variable pattern followed by the wait times and length of queue, and corresponding 'required' number of specialists as obtained from the proposed model. Table 5-2 illustrates a 'summary' of DEA results from all 160 data points in the busy specialist OPD, which shows the rapidly increasing required number of specialists, with increase in wait time and length of queue. For example, on average, when the wait time was between 1h and 1.5h (60-90 minutes) and the length of queue varied between 14 and 28, the corresponding required number of specialists is 4 to 5. However, when the wait time has a high value of between 3.5h and 4h (210-240 minutes) and length of queue is 31 to 58, the corresponding requirement is between 10 and 13. Hence, Figure 5-2 and Table 5-2 highlight that crucial information with regard to 'exact' requirement of specialists with rapidly changing queue situation, can guide the administrators to take action early with regard to adequate staff allocation.

The analytical results carried out in PIM-DEA are given in Appendix 4, along with the complete set of results of the busy specialist OPD in Appendix 5.

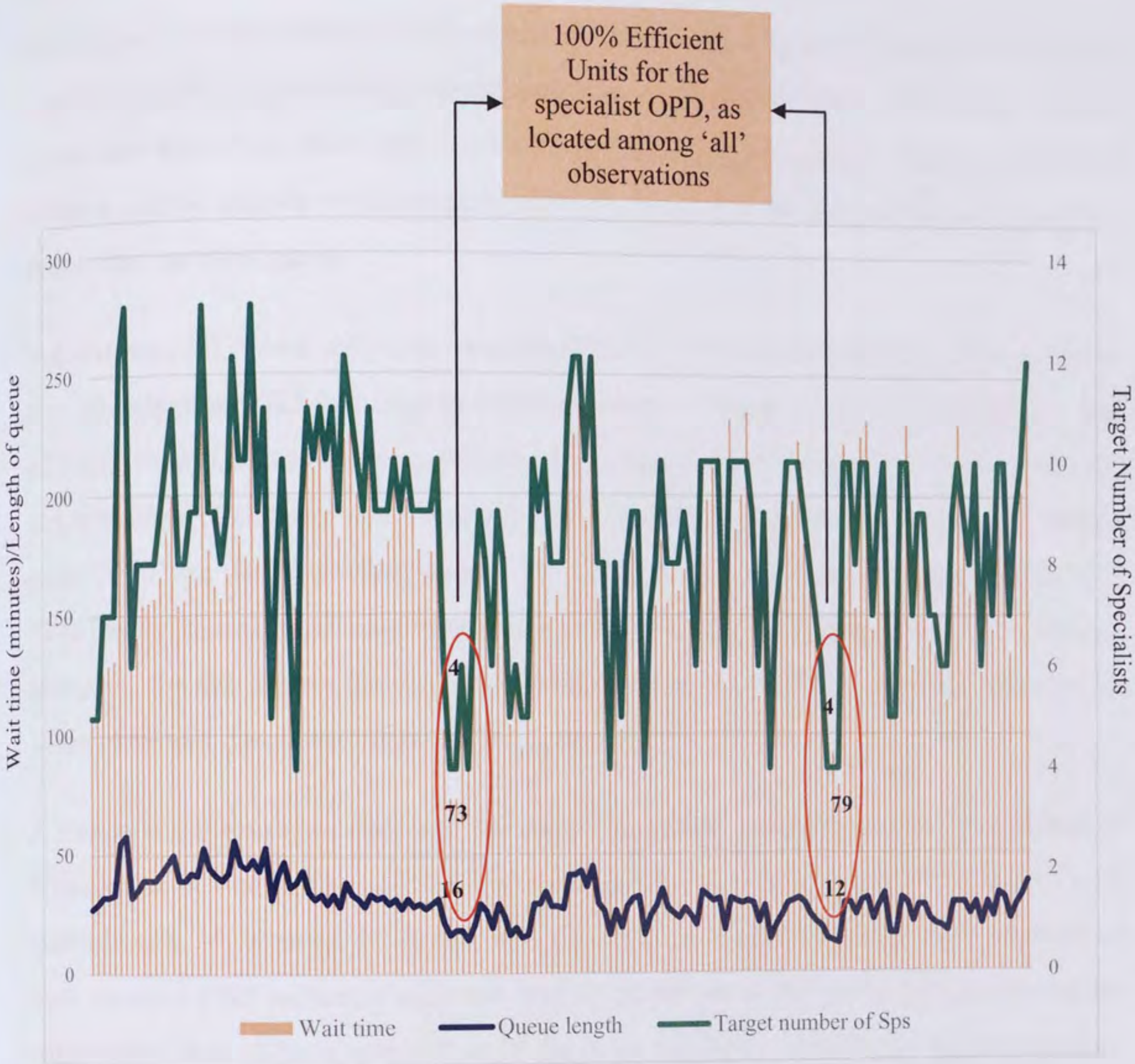


Figure 5-2: Variable Pattern Followed by the Two Inputs and the 'Target' Number of Specialists

Table 5-2: Summary of 'Required' Number of Specialists (on average) in Specialist OPD Given the 'Wait time' and 'Length of Queue' Categories

Wait time Category	Length of Queue Category	Required Number of Specialists (Specialist OPD)
60-90 mins	14 to 28	4 to 5
90-120 mins	16 to 29	5 to 6
120-150 mins	20 to 33	5 to 8
150-180 mins	23 to 44	8 to 10
180-210 mins	28 to 47	9 to 10
210-240 mins	31 to 58	10 to 13

Additionally, in few cases, it was observed that the varying requirement of doctors is sensitive to just a *slight* change in the wait time and length of queue. This might be due to the fact that when these input variables are already quite high, a slight increase will create a drastic impact on the required number of doctors since it would take longer to reduce the massive queue.

For instance, P024 had wait time was around 2.5h (157 minutes) and the length of queue was 39, whereas P023 has slightly different values of input variables (162 minutes and 42 respectively). Both observations had current availability of doctors as 2. However, the target number of doctors was 8 and 9 for P024 and P023 respectively. At another sample point P070, the wait time was around 3h (181 minutes) with queue length 29. The unit P145 has the same queue length with only a slight increase in the wait time (192 minutes). However, for the former the required number of doctors was 9 but for the latter was 10 where both had current availability of doctors at 2.

A similar trend was also observed for some data points when the current availability of doctors was 4. For instance, P106 had waiting time of around 2.5h (142 minutes) with queue length 24, whereas P100 has wait time 2.5h as well with just a slight increase of three minutes (145 minutes) with two additional people in the queue (26). However the required number of doctors was 7 and 8 for P106 and P100 respectively. In another case, the wait time for P054 and P108 is around 3h with only a slight difference (173 and 181 minutes respectively) and a minor change in length of queue which is 32 and 35 respectively. However, the requirement for the former is 9 and the latter is 10. Similarly for P046, the wait time is nearly 4h (226 minutes) and length of queue 33, with the requirement of doctors as 11. Although only a minor difference was observed for P181 with regard to input variables (239 minutes and 38 respectively); however, the target number of doctors was 12.

Considering the current availability of 5 doctors, for two observations P171 and P061, the wait times were around 2.5h (159 and 166 minutes respectively) with queue length 24 and 27, however the required number of doctors is 8 and 9 respectively. For P121, P052 and P048, the wait time is around 4h (232, 228 and 223 minutes respectively) with queue lengths 31, 33 and 38, however, the requirement of doctors is 10, 11 and 12 respectively.

These results demonstrate that only a *slight change* in input variables sometimes led to an *increase* in the required number of doctors at 8 or above. This not only highlights the extreme variability in the arrival pattern of patients, but also displays the importance of having an adequate number of doctors at the right time. For all current availability of doctors, on average, a change of 8-13 minutes of wait time and 2-5 queue length led to an additional increase in the requirement of doctors. Therefore, it is crucial that a sufficient number of doctors are available when queue starts building up, before it reaches a drastic level, showing an unfeasible target value of doctors.

5.1.2 DEA Analysis and Results: 'Required' Number of Pharmacists at the Pharmacy

As mentioned in Section 4.2.2, the main pharmacy consists of patients arriving from different departments of the hospital is operational on all working days. The pharmacy is usually the last stage in the patient flow system, as the patients visit other departments and collect medicines just before exiting the system. The DEA model proposed in Section 4.2.4 has been utilized to determine the 'required' number of pharmacists when analysing the queue system at the *serving counter of the pharmacy* as shown in Table 5-3.

Table 5-3: Inputs and Outputs for the Queuing-DEA Model for the Serving Counter at Pharmacy

Inputs	**Outputs**
Wait time for serving Counter at pharmacy	Number of pharmacists
Length of Queue for serving Counter at pharmacy	Service time when served by the pharmacist at the counter **(Non-discretionary Output)**

Unlike the specialist OPD and GDMO clinic, the pharmacy has a *fixed* availability of pharmacists at all times for all counters. In case of the serving personnel counter, there are 2 pharmacists available during the day.

The DEA analysis showed only one *100% efficient* unit among the observations which is P007. The wait time is around 0.5h and length of queue is 4, with the availability of pharmacists at 2 as being sufficient and service time to be maintained at 8 minutes. Therefore, this unit can act as a benchmark for other observations.

Considering the inefficient units, for P065 and P022, the wait is around 1h (54 minutes and 63 minutes respectively, but the queue length is 8 and 12 respectively. Hence, the required number of pharmacists is 3 and 4 for P065 and P022 respectively. For P002 and P042 (with 40% and 33% efficiency respectively), the wait time has increased to around 1.5h (82 and 101 minutes respectively); with only a marginal increase in queue length of P002 at 14 and a little higher for P042 at 16. The requirement of pharmacists was 5 and 6 for P002 and P042 respectively. For data point P020 and P016, the wait time is around 2h (112 minutes and 125 minutes respectively), but the length of queue of P016 (23) is higher than P020 (20). The requirement of pharmacists is 7 and 8 for P020 and P016 respectively. For P019, an increase was observed in both inputs, with wait time around 2.5h (149 minutes) and queue length 26, hence the target value of pharmacists is 9. For P053, it was observed that the queue length is unchanged at 26, however, the wait time has further increased to a very high value of 3h (168 minutes). Whereas for data point P015, the wait time is the same which is around 3h (166 minutes) but queue length is a high value of 34. The requirement of doctors for both P053 and P015 is 10.

Therefore, it can be *concluded* from the DEA analysis of the serving counter at pharmacy that one additional increase in the required number of pharmacists resulted from an increase in wait time between 11 and 24 minutes and queue length of 3 to 4, as shown by maximum data points. Hence, it is important that necessary measures are taken sooner than later to prevent excessive queuing and unrealistic targets for the requirement of pharmacists.

Figure 5-3 shows the values of the wait time, length of queue and the corresponding 'required' number of Pharmacists, for 'each' of the 52 observations of serving counter at the pharmacy, where the value for 100% efficient unit, P007, is represented in a red circle. Considering P007, the values of required number of specialists, wait time and length of queue are, 2, 33 minutes and 4 respectively, whereas values for any other data point say P041 are 7, 109 and 19 respectively. Hence, Figure 5-3 demonstrates a highly variable pattern followed by the wait times and length of queue, and corresponding 'required' number of Pharmacists as obtained from the proposed model. Table 5-4 illustrates a 'summary' of DEA results for all 52 data points from pharmacy serving counter, which shows increase in the required number of Pharmacists, corresponding to increasing wait time and length of queue. For example, on average, when the wait time was between 0.5h and 1h (30-60 minutes) and the length of queue was between 4 and 8, the corresponding target number of Pharmacists is 2 to 3. However, the required number increased to a high 9 to 10, when the wait time increased to 2.5h to 3h (150-180 minutes) with length of queue varying between 23 and 34. Hence, like busy specialist OPD, Figure 5-3 and Table 5-4 highlight the issue of rapidly changing queue situation so that administrators can control queue pre-emptively using information provided by the proposed model of the required staff availability.

The analytical results carried out in PIM-DEA are given in Appendix 6, along with the complete set of results for the serving counter at the pharmacy in Appendix 7.

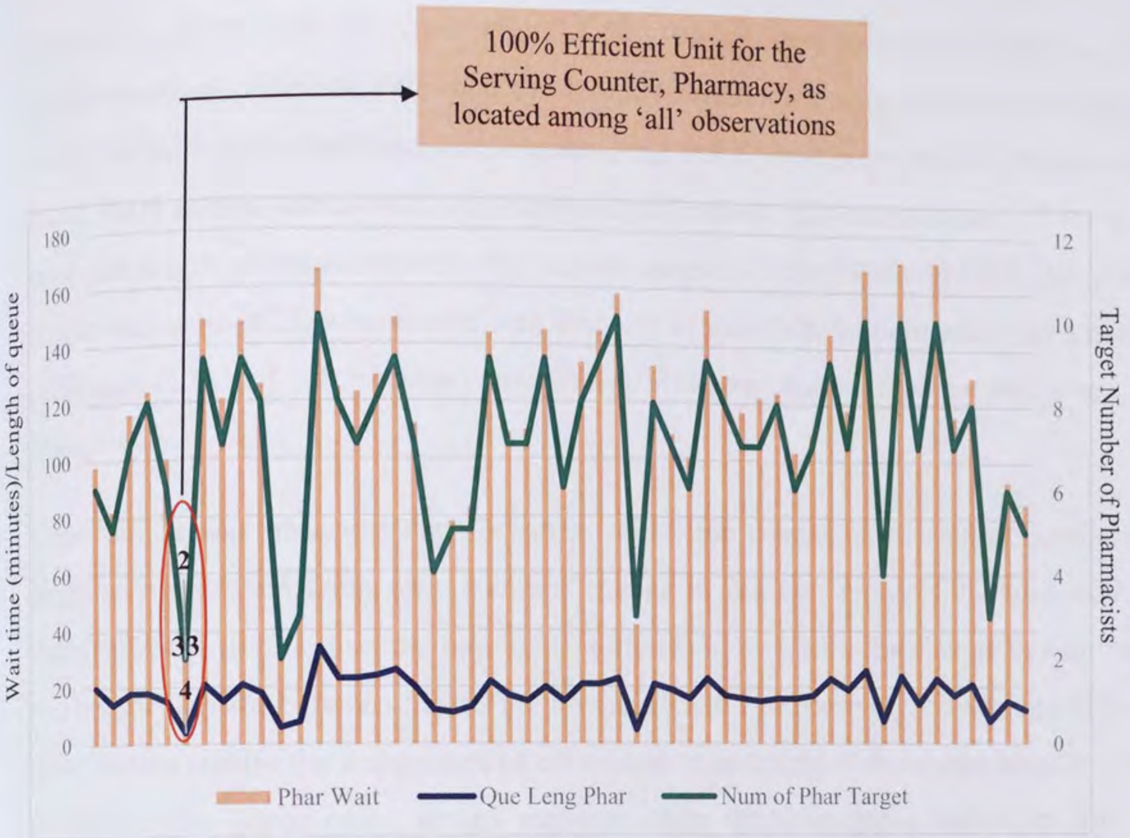


Figure 5-3: Variable Pattern Followed by the Two Inputs and the 'Target' Number of Pharmacists

Table 5-4: Summary of 'Required' Number of pharmacists (on average) at Serving Counter, Pharmacy Given the 'Wait time' and 'Length of Queue' Categories

Wait time Category	Length of Queue Category	Required Number of pharmacists (Serving Counter, Pharmacy)
30-60 mins	4 to 8	2 to 3
60-90 mins	11 to 14	4 to 5
90-120 mins	14 to 20	6 to 7
120-150 mins	15 to 24	7 to 9
150-180 mins	23 to 34	9 to 10

Further DEA Analysis of Few Observations at the Pharmacy

However, similar to the busy specialist OPD, the findings of the serving counter of the pharmacy also had a few findings where the required number of pharmacists changed with only a *slight* increase in wait time or queue length or both. For instance, comparing P037 with the 100% efficient unit P007 (with wait time of 33 minutes and queue length at 4), the wait time and queue length are only slightly increased (42 minutes and 5 respectively). However, the required number pharmacists for P037 is 3. Similarly, when considering P049 with wait time around 1.5h (103 minutes) and queue length 16, the requirement is 6. Whereas, P027 had a wait time (109 minutes) only a few minutes higher than P049 with just a slight increase in queue length of 17, but the target is 7. On the other hand, P004 shows values with only marginal differences (124 minutes and 18 for wait time and length of queue respectively), but the target is 8. For P043 and P035, the queue length was same at 23, whereas the wait time was around 2.5h for both with only a minor difference (153 and 158 minutes respectively). However, the requirement was 9 and 10 respectively.

Therefore, it was observed that for some values the change in the target number of pharmacists resulted from only a minor change in the two inputs simultaneously or independently. In this case, the requirement increased by 1 with the change in wait time and length of queue between 5 and 9 minutes and 0 and 1 respectively, on average. Hence, these results signify the importance of continuous monitoring of the queue situation and controlling the queue early, as any measure taken after the queue builds up will be pointless.

5.1.3 DEA Analysis and Results: 'Required' Number of GDMOs at the GDMO Clinic

The GDMO clinic is one of the most important departments of the hospital. As identified in the patient flow system shown in Figure 4-7, all first-time patients 'must' be referred from GDMOs before they can proceed to any departments such as specialist OPDs or laboratory (for investigations). Therefore, the GDMO clinic is very busy with wait times of up to 3 hours.

The DEA model proposed in Section 4.2.4 has been utilized to determine the 'required' number of GDMOs when analysing the queue system at the *GDMO Clinic* as shown in Table 5-5.

Table 5-5: Inputs and Outputs for the Queuing-DEA Model for the GDMO Clinic

Inputs	**Outputs**
Wait time for GDMO Clinic	Number of GDMOs
Length of Queue for GDMO Clinic	Consultation time to be examined by the pharmacist at the counter **(Non-discretionary Output)**

The *availability* of doctors at the GDMO clinic was less variable as compared to the specialist OPD, as they have a minimum of 2 and a maximum of 3 doctors, every day of the week. Among a total of 114 observations, the current availability was 2 for 65 observations, whereas for the remaining 49 observations, the current availability was 3 (as shown in Figure 5-4 below); demonstrating that the availability of doctors varied throughout the day/week.

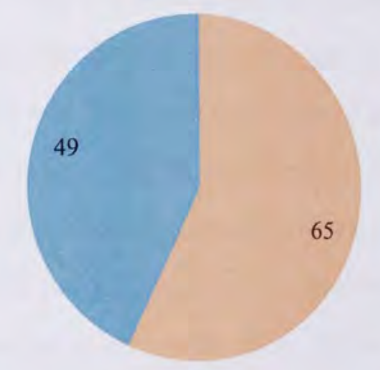


Figure 5-4: Number of Observations at Current Availability of GDMOs at the GDMO Clinic with a Total of 114 Observations

Considering the efficient units, there were a total of 4 observations which had a *100%* efficiency level. All four observations have a queue length of 4 to 5 patients, and a wait time of around 0.5h, with the availability of doctors as either 2 or 3. Three of the observations require a consultation time of 8-10 minutes while one unit which has a consultation time of 5 minutes. However, an investigation of the results of efficient peers indicates that P064 and P125 are peers to maximum observations in the dataset, 52 and 107 observations respectively. For these two observations, the availability of doctors is 3 with wait time of around 0.5h, queue length of around 5 and consultation time of around 5-8 minutes (which remains unchanged). Therefore, considering this supplementary information, these observations act as a benchmark. As mentioned in case for specialist OPD, it might not be possible to achieve this 'target' all the time, but it does provide a guideline.

With regard to the inefficient units, the 'required' number of doctors varies given the current availability of either 2 or 3 in the GDMO clinic, depending on the varying queuing characteristics. The two values having an efficiency level of 93% had very similar characteristics as to those of 100%, which proves that the efficiency level of above 90% is almost the same as 100%. Considering the current availability of *doctors at 3*, for P099 with efficiency level 80%, the wait is around 40 minutes, with a slight increase in the queue length which is 6. Therefore, the requirement of doctors has increased to 4. For P081, although the queue length had a small increment (to 7) but the wait time increased to nearly 1h (58 minutes), hence the requirement of doctors is 5. In case of P088 (52% efficiency), the wait time although still being around 1h is greater (65 minutes) with little increase in queue length (at 10), led to a target of 6 doctors. The wait time for P036 is 1.5h (85 minutes) with slight increase in queue length to 12, increasing the requirement to 7. Although the units P073 and P080 had wait time of around 1.5h (95 and 102 minutes respectively), however, the queue length was slightly higher at 13 and 15, hence the required number of doctors was 8 and 9 respectively. For P089, P031, and P129, the wait time was around 2h (116, 124 and 133 minutes respectively) and queue length was 16, 18 and 21 respectively. However, even with these minor differences among the units regarding the two inputs, the requirement of doctors was different which was 10, 11 and 12 for P089, P031 and P129 respectively (all had a consultation time of 3 minutes). For P093, however, a sharp increase was observed in the wait time and length of queue, to around 2.5h (153 minutes) and 26 respectively. This change led to a requirement of 14, that is, an increment of 2 doctors. For P079 and P022, the wait time was around 3h (171 minutes and 178 minutes respectively), however, queue length of P022 was slightly higher than P079 (29 and 26 respectively). Hence, the target for P079 and P022 was 15 and 16 respectively.

Similar trend was observed when the availability of *doctors was 2*. For P105, the wait time is around 40 minutes, with length of queue at 6, hence increasing the requirement to 3 whilst keeping the consultation time fixed at 7 minutes. For P113 with efficiency level 49%, the wait time is closer to 1h (46 minutes) and length of queue is 7 increasing the requirement to 4. For P020 and P059, the wait time was around 1h (55 and 73 minutes respectively), however, the queue length was 9 and 10 respectively. However, even this slight difference in the inputs for the two data points, led to a different requirement of doctors which was 5 and 6 respectively. In case of P106, the wait time was around 1.5h (80 minutes), with queue length only marginally higher at 11, but the requirement increased to 7. For P076 with an extremely low efficiency of 24%, the wait time was also around 1.5h (93 minutes) but the queue length was at 16, increasing the target to 8. For

P055 and P027, the queue length was same at 15, however, the wait time was around 1.5h (104 minutes) and 2h (110 minutes) respectively. Hence the target number of doctors was 9 and 10 respectively. For P083 and P009, the wait time was around 2h (123 and 133 minutes respectively), with length of queue at 19 and 21 respectively. The data point P083 had lower input values than P009, even though slightly. Hence, the target number of doctors was a high 11 and 12 respectively. For P075 with an efficiency level of below 20% (15%), the wait time had an even higher value of around 2.5h (151 minutes), with a sharp increase in queue length at 29, and therefore, the target number of doctors has further increased to 13.

Overall, the DEA analysis for the GDMO clinic demonstrated that on average, a change in the wait time of 7-13 minutes with a variation of 1 to 3 people in the queue length, either simultaneously or independently, led to one additional number of required doctors. Therefore, it can be concluded that for the GDMO clinic, the variation in the availability of doctors is affected with a *small* increment in either one or both inputs. In case of busy specialist OPD and pharmacy, this was the case for only a few observations. However, for the GDMO clinic, nearly all observations show that the required number of doctors is sensitive to even a slight change in the wait time or queue length or both. This differentiation between GDMO clinic and the other two departments might be due to the variation in the arrival pattern of patients. Although all three departments experience immense overload of patients due to absence of prior appointments, however, the busy specialist OPD and pharmacy experience more patients arriving *'at one time'* leading to a queue build up. Hence, the wait time and length of queue spikes up just within few minutes. However, in GDMO clinic, although the number of patients arriving is high but the arrival pattern is more *spread out*. With the arrival of each new patient, the increase in the queue length and wait time is not as drastic as for the other two departments. Hence, the change in the requirement of doctors is observed for *less variation* with respect to GDMO clinic when compared with specialist OPD and pharmacy. However, in both cases, the issue remains that it is extremely significant to monitor a queue at more frequent intervals to avoid either situation.

Figure 5-5 shows the values of the wait time, length of queue and the corresponding 'required' number of GDMOs, for 'each' of the 114 observations, where values for P064 and P125 (100% efficient units) are represented in red circles. For example, considering

P064, the values of required number of specialists, wait time and length of queue are, 3, 35 minutes and 4 respectively, whereas values for say P041 which are 8, 95 minutes and 16 respectively. Hence, Figure 5-5 demonstrates a highly variable pattern followed by the wait times and length of queue, and corresponding 'required' number of specialists as obtained from the proposed model. Table 5-6 illustrates a 'summary' of DEA results from all 114 data points in the GDMO clinic, which shows the rapidly increasing required number of specialists, when wait time and length of queue increases. For example, on average, when the wait time was between 0.5h and 1h (30-60 minutes) with length of queue between 5 and 8, the corresponding required number of specialists is 3 to 5. However, when the wait time increased to 2.5h to 3h (150-180 minutes) and length of queue varies from 26 to 29, the corresponding requirement is a high 13 to 16. Hence, although Figure 5-5 and Table 5-6 show high variability among the three variables though increase in the required number of GDMOs is observed for a lower level of variation in the wait time and length of queue, as compared to the busy specialist OPD and pharmacy. However, these results signify the need to control the queue early, to prevent building up of queue and a high requirement (sometimes to an unfeasible number) of GDMOs.

The analytical results carried out in PIM-DEA are given in Appendix 8, along with the complete set of results for the GDMO clinic in Appendix 9.

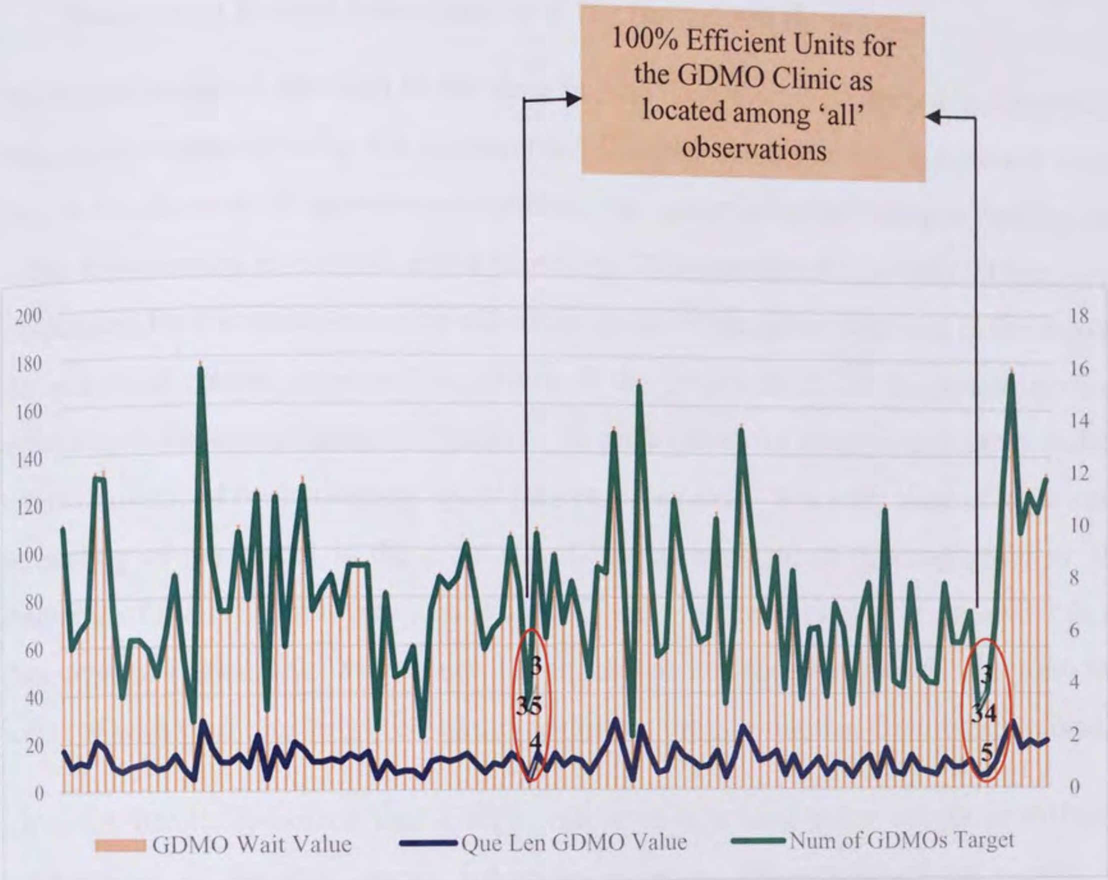


Figure 5-5: Variable Pattern Followed by the Two Inputs and the 'Target' Number of GDMOs

Table 5-6: Summary of 'Required' Number of GDMOs (on average) at the GDMO Clinic Given the 'Wait time' and 'Length of Queue' Categories

Wait time Category	Length of Queue Category	Required Number of GDMOs (GDMO Clinic)
30-60 mins	5 to 8	3 to 5
60-90 mins	9 to 13	5 to 8
90-120 mins	14 to 17	8 to 10
120-150 mins	18 to 21	11 to 12
150-180 mins	26 to 29	13 to 16

5.1.4 Summary and Usefulness of DEA Analysis/Results for Assessing Queuing Situation in Busiest Departments of the Designated Hospital

The excessive queue situation in the three busiest departments of the public hospital in Pakistan is evaluated using the Queuing-DEA model constructed in the current study. Due to the absence of appointment systems, the queue situation changes rapidly, just within few minutes, as patients arrive randomly. This unpredictability makes it extremely challenging for the management to effectively monitor the queue situation. In this regard, the analytical results depicted the ability of the proposed model to capture extreme variability in the arrival pattern of patients, by pinpointing the extreme wait times and the extent of overload by evaluating 'each' patient *'separately.'* Since the issue of inadequate scheduling of personnel is the most significant determinant of queuing problem; the emphasis of the DEA analysis was to identify the 'required' number of personnel in all departments. Hence, the '*target'* number of doctors has been determined using the two inputs of wait time and length of queue, given that the consultation time remains fixed.

The DEA results indicated that a high wait time or a long queue length or different combinations of the two inputs, led to an increased requirement of the number of personnel. In some cases, the DEA results demonstrated an *unfeasible* target number of personnel when the queue reached an extremely high level, irrespective of the current availability of personnel. Hence, it can be concluded that any measure undertaken at this point will be ineffective and it will take a considerable amount of time to bring down the queue. Therefore, the current Queuing-DEA framework plays a significant role as it can alert the management about the *'exact'* number of personnel required at a specific time to control the queue *pre-emptively*. This will not only ensure that the long wait times of subsequent patients are minimized by taking action *early*, but will also lead an *optimal* utilization of personnel. In this way, when queue is excessive, personnel can be briefly withdrawn from other tasks to cater to long queues. However, when the queue dies down, the management will know instantly and the staff members can undertake other pending activities.

Furthermore, in a few cases, the DEA results also indicated that the requirement of personnel changed with just a *slight variation* in the wait time and length of queue. This demonstrates the varying arrival pattern of patients and the quick queue build-up just within few minutes, resulting in one additional requirement of personnel. Hence, in this regard, the proposed model provides an opportunity to observe the queue at frequent intervals and capture rapid changes, in order to immediately determine the 'exact' requirement of personnel.

A summary of the usefulness for the results obtained from Queuing-DEA model proposed in the current study is given in Figure 5-6.

Usefulness of DEA Analysis for Assessing a Queue System

Provides a guideline for the 'exact' number of personnel required at a specific time to control the queue pre-emptively and reduce excessive queueing, and for optimal utilization of personnel

Can assess the queue system at multiple time intervals to determine the requirement of personnel

Evaluates 'each' patient separately to capture variability in the arrival pattern of patients and drastic changes in the queue within few minutes

Depicts the major as well as minor changes in the wait time and queue length, and different combinations of both, resulting in a change in the target values of personnel

Figure 5-6: Summary of the Usefulness of DEA Analysis and Results Obtained when Assessing the Queue System

5.2 Towards Practical Implementation of the Queuing-DEA Model and Results by Constructing a Dynamic Framework

The preliminary investigation of the queue system, in terms of observation and few complementary interviews, demonstrated that the inadequate scheduling of doctors is the most crucial issue which leads to long queues and wait times for patients. At present, there are *no* existing effective strategies which aim at better queue management within the busy departments. This further signifies the dire need for an effective measure which controls the queue as soon as it starts building up. The challenge is to determine the *'exact'* requirement of doctors at a specific time, and 'when' they are needed, for optimal staff utilization and minimization of wait times of patients. The analytical results in Section 5.1 illustrate the 'required' number of personnel (doctors/pharmacists) at each of the busiest departments. It was also observed that the target number of personnel increased substantially to an unfeasible number, once the wait time and level of overload spiked. Therefore, the DEA results lead to the overall conclusion that there is a need to detect rapid changes in the queue situation to avoid excessive waits and for optimum staff utilization.

Given the useful DEA results, there is a need to develop a *dynamic framework* dedicated towards the *practical implementation* of the Queuing-DEA model; such that it can be applied in every busy department of the designated hospital and other similar hospitals even in other developing countries. The aim for constructing the dynamic framework is to continuously supervise the changing queue situation and ensuring adequate staff availability.

5.2.1 Characteristics of the Dynamic Framework to Assess the Queue System

The *dynamic framework*, consisting of the proposed Queuing-DEA model, has been developed in Microsoft Excel. The primary objective is to ensure that the administration is kept updated with the changing queue situation. Hence, immediate action can be taken by increasing the required number of doctors. The framework has been developed in Excel as it is commonly available and staff are well-versed with its functionality. Also, its simple and easy-to-understand features would require minimal effort on part of the users, hence allowing for instant implementation. The dynamic framework has been designed such that it evaluates the queue system each time a patient arrives, and displays the required number of doctors along with the efficiency of the system 'at that moment', with minimum 'setting-up' issues. A screenshot of the dynamic framework as developed for the specialist OPD is shown in Figure 5-7, and can be developed for other departments on a similar pattern. Column A requires the patient number and name. Due to overcrowding in the OPD, the name of the patients will help in quickly identifying the patient in the database along with patient number speeding up the process. The next four columns, B, C, D and E represent the queuing information that needs to be entered manually by the receptionists. As soon as the patient reports at the reception, information will be entered in Column A, B and D, including the patient name and number, time of arrival and the current availability of doctors. The receptionists already have the information about the number of doctors working currently, since they sign in when they arrive at the Reception. At this instant, the token number will be given to the patient. The same pattern will be followed for all other patients who arrive. The information in Columns C and E will be added when the patients report back at the Reception, before leaving this specific department. This information consists of the time of departure or the time out (Column C), along with the consultation time (Column E). The consultation time will be noted by the doctor or the trainees assisting senior doctors. The 'time of entry' in the examination room will be noted when a patient enters and 'time of exit' will be noted when the patient leaves. This information can be noted on a time sheet and given to the patient, who can then show this time sheet to the receptionist who can enter the consultation time in Column E of the spreadsheet. As soon as 'time in' is entered in Column B when a patient arrives, the length of queue will appear in Column G; whereas upon entering 'time out' in Column C, the value of wait time will appear in Column F. The calculations for the two variables are automated and do not need to be entered manually, providing real time data.

With this information, the necessary data for the DEA model is available including the inputs (wait time and length of queue in Columns F and G respectively), and the outputs (current number of specialists and Consultation time in Columns D and E respectively). As soon as the receptionist enters the consultation time, which is the last piece of information required, the Queuing-DEA model as developed within the framework, will run 'automatically', and will re-run each time the data of a patient is entered. For instance, after data has been entered for Patient 10, the DEA model will re-run. Similarly, as soon as the data is entered for Patient 11, the DEA model will run again, and so on and so forth. After a DEA model runs, values will start appearing in Columns P and BO where Column

P represents the 'required' number of doctors, and Column BO gives the efficiency level of the system at that instant. The information displayed by Column P is extremely significant as it provides a basis for decision-making regarding increasing the number of doctors if required. This decision can be influenced by a threshold pre-specified by the administrators. For instance, the threshold for required number of doctors can be set at 3, where the receptionist can notify the Head of Department (HOD) if the value exceeds 3, who can then re-assign the doctors to the OPD. If the 'required' availability is equal to the 'current' availability of doctors, it shows that more doctors are not required as shown by the DEA model, and no action is required at this stage. However, since the DEA model gives a comparative analysis where the efficiency level and targets are provided relative to the other units, it might also be possible that target number of doctors is the same as current number, although the wait time is high. Therefore, in this case, more doctors are required since the wait time of all patients is high. This might happen rarely, however, it is important to observe the wait times independently as well. A threshold for an 'acceptable' level of wait times can be specified after consensus with the administrators. With each run of the DEA model, the efficiency level of the system will be displayed in Column BO.

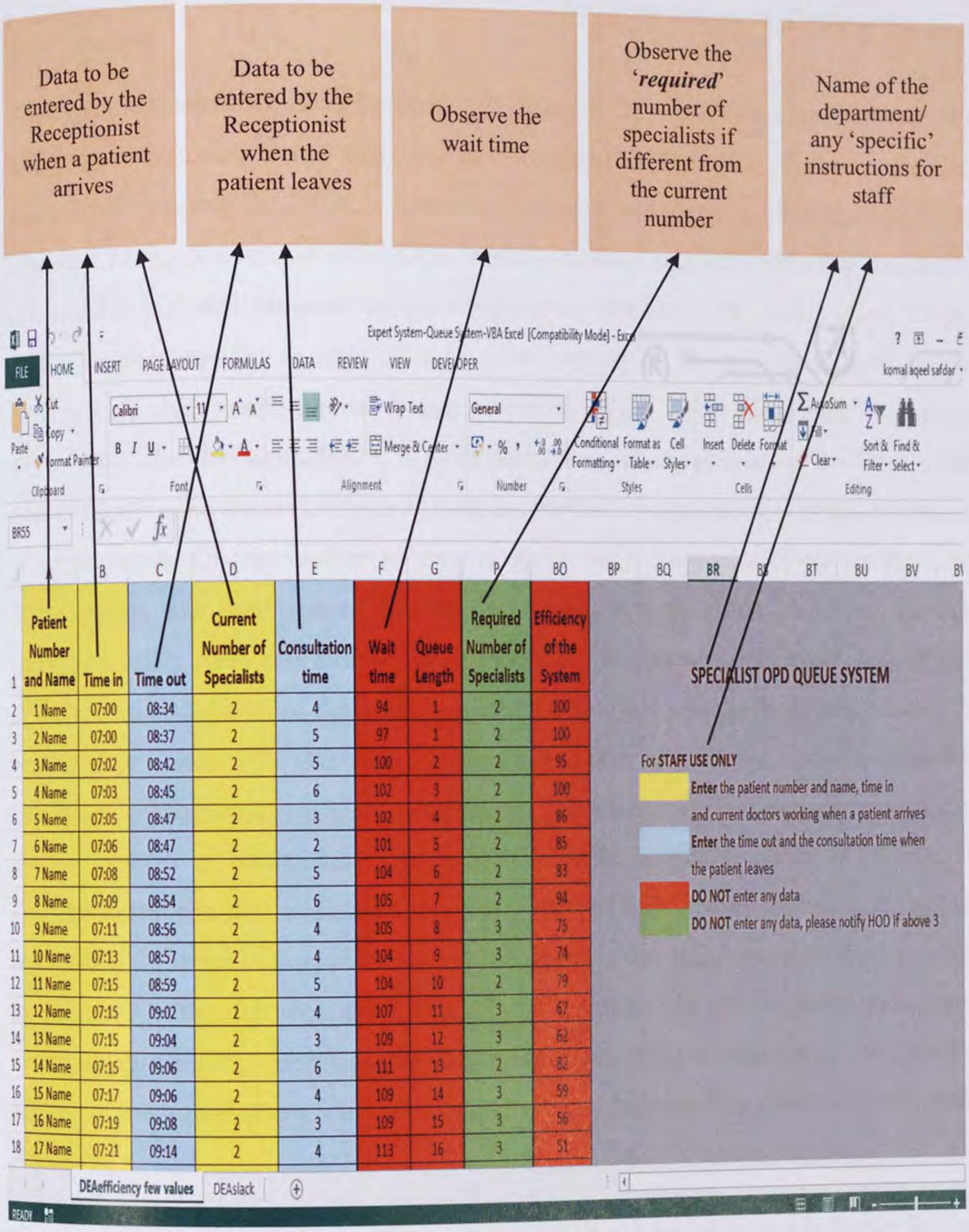


Figure 5-7: Proposed Dynamic Framework for Queue Assessment

Therefore, by evaluating the current queue situation as soon as the first few patients start arriving, the *requirement* of doctors can be determined at that stage. Hence, appropriate action can be taken earlier on, by re-assigning the number of doctors as suggested by the model. This information is crucial as it allows the administration to identify the 'required' number of personnel, and manage the queue *dynamically* and efficiently. Therefore, the wait times of subsequent patients can be minimized, which is the ultimate benefit of this useful and 'easy-to-use' dynamic system.

5.2.2 Technical Aspects of the Dynamic Framework for Assessing the Queue System

The dynamic framework using the proposed Queuing-DEA model has been developed in *Microsoft Excel* using Visual Basic for Applications (VBA) coding (the complete VBA codes for developing the dynamic framework in Excel, are given in Appendix 10). In this case, the VBA code for the entire DEA model generated was run, including the output-oriented DEA model followed by the slack model. Based on the slack values, 'target' values for the required number of doctors are calculated. Since both of these models require a 'Solver' programme to run, therefore, they were developed in separate worksheets. The first worksheet (DEAefficiency few values) consists of the DEA model whereas the second sheet (DEAslack) includes the slack model, along with the calculation of target values for the number of doctors. However, for the sake of convenience, the target values are then copied to the first worksheet, to avoid switching between worksheets after every run, simplifying the system. Also, the columns which strictly refer to the technical aspects of the DEA modelling and VBA programming are 'hidden' as displayed in Figure 5-7. The code has been developed for a total of 50 units (patients), but can be increased if required. Additionally, the VBA code has been developed such that the process is 'automated'. Hence, as soon as the consultation time is entered, the DEA model in the first worksheet will run, followed by the slack model in the second worksheet; resulting in the calculation of the target values which are then finally copied to Column P of the first worksheet. Therefore, the entire code will run in the background automatically, and the receptionists/other users need not be involved with the technical aspects of the model. They are only required to enter minimal data manually in Columns A to E.

Considering the *first worksheet* 'DEAefficiency few values' as shown in Figure 5-8, the inputs and outputs of the model are displayed in Columns D to E and F to G respectively. Zhu (2009) develops a DEA model in Excel using VBA, along with running a slack model. Therefore, based on the formulae and VBA codes suggested by Zhu (2009) for an input-oriented model, the codes were developed for the Queuing-DEA model as proposed by the current study. The DEA model has been set up in columns J to O. The formulae for the constraints (four constraints as there is a total of four variables) are developed in cells J4 to M7, where J4, K4, L4 and M4 represent the first constraint related to the wait time in the specialist clinic. Similarly, the three rows below represent other constraints.

The cell N2 represents the unit under evaluation, in this case the patient for which the DEA will be run. As demonstrated in Section 3.1.2, the DEA model runs for all units in the sample separately to display the efficiency of a particular unit relative to other units in the sample. Therefore, every time a new patient arrives, the entire DEA model will re-run 'for all units'. Therefore, at the end of each run, the number 50 will appear in N2 since this is the last unit in the sample. The cell O3 represents the efficiency level of the concerned unit. In this case as well, the efficiency of the 50^{th} unit will display when the DEA model will run for all 50 units each time. Additionally, the DEA model has already been set up for a total of '50' units initially even if some of these values will display 0, since there are no patients to show. Therefore, when a new patient arrives, the values will start appearing for the current patient and the patients before. But the data for the following cells will display the value zero. Furthermore, the values of weights (identified as lambdas in the DEA model) will appear in Columns Q to BN. When running the DEA model for each patient, the lambda values for all 50 units will be displayed in these columns. For instance, the lambda values for all units when running the DEA model for Unit 1 will appear in rows Q2 to BN2 and so on, till Unit 50 which will be displayed in Q51 to BN51. Column H shows the lambda values for the particular DMU under assessment and will show the values of the 50^{th} unit as well. Therefore, the efficiency levels are then displayed in Column I for all units, which are then converted into percentage form in Column BO. For instance, for unit 3, the efficiency level is greater than 1 by approximately 5%. Therefore, this means that the inefficient unit needs to increase the output by 5% to obtain 100% efficiency, hence the efficiency level is 95%.

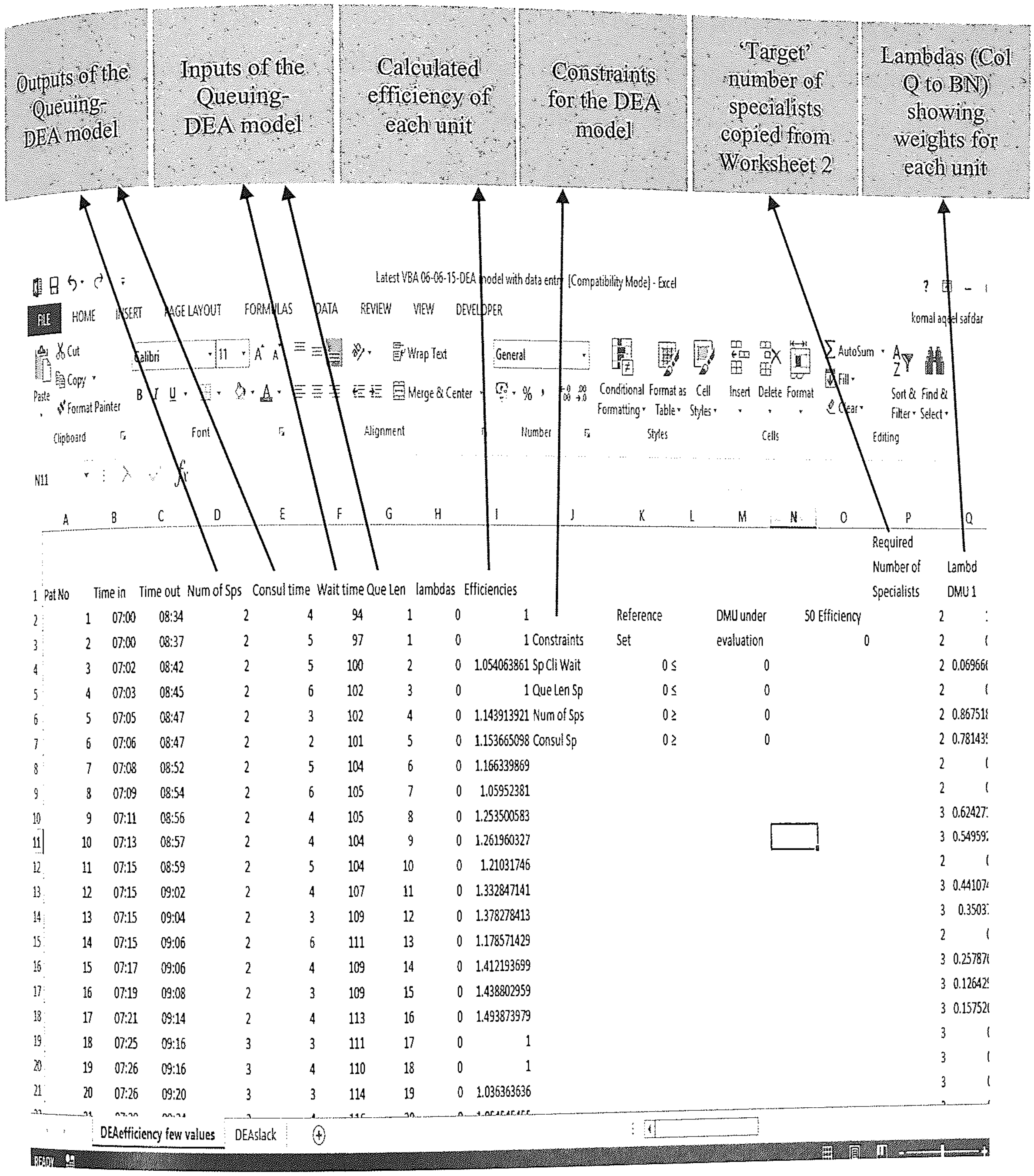


Figure 5-8: Worksheet 'DEAefficiency few values' of the Dynamic Framework Displaying the Queuing-DEA Model Developed in Excel

In the *second worksheet* 'DEAslack' as illustrated in Figure 5-9, the slack model is displayed which is run 'after' the DEA model is run. Columns A to I, and cell O3 which shows the efficiency level, are copied from first to the second worksheet since the constraints of the slack model use the input/output data as well, avoiding the need to refer back to the first worksheet. The constraints are set up from J4 to M7, with unit under evaluation in cell N2. The objective function of the slack model is the sum of slacks which is displayed in cell P8. The slacks for each input and output variable are the decision variables and will also appear in the constraints. The slack variables are set-up in cells P4, P5, P6 and P7. These cells will be the 'changing cells' since the values will change every time the DEA model will run for each unit. After the slack model is run, the slacks are displayed in Columns R to U for each of the 50 units. For instance, cells R2, S2, T2 and U2 will represent the slacks for all four input and output variables for Unit 1, and so on, till R51, S51, T51 and U51 for Unit 50.

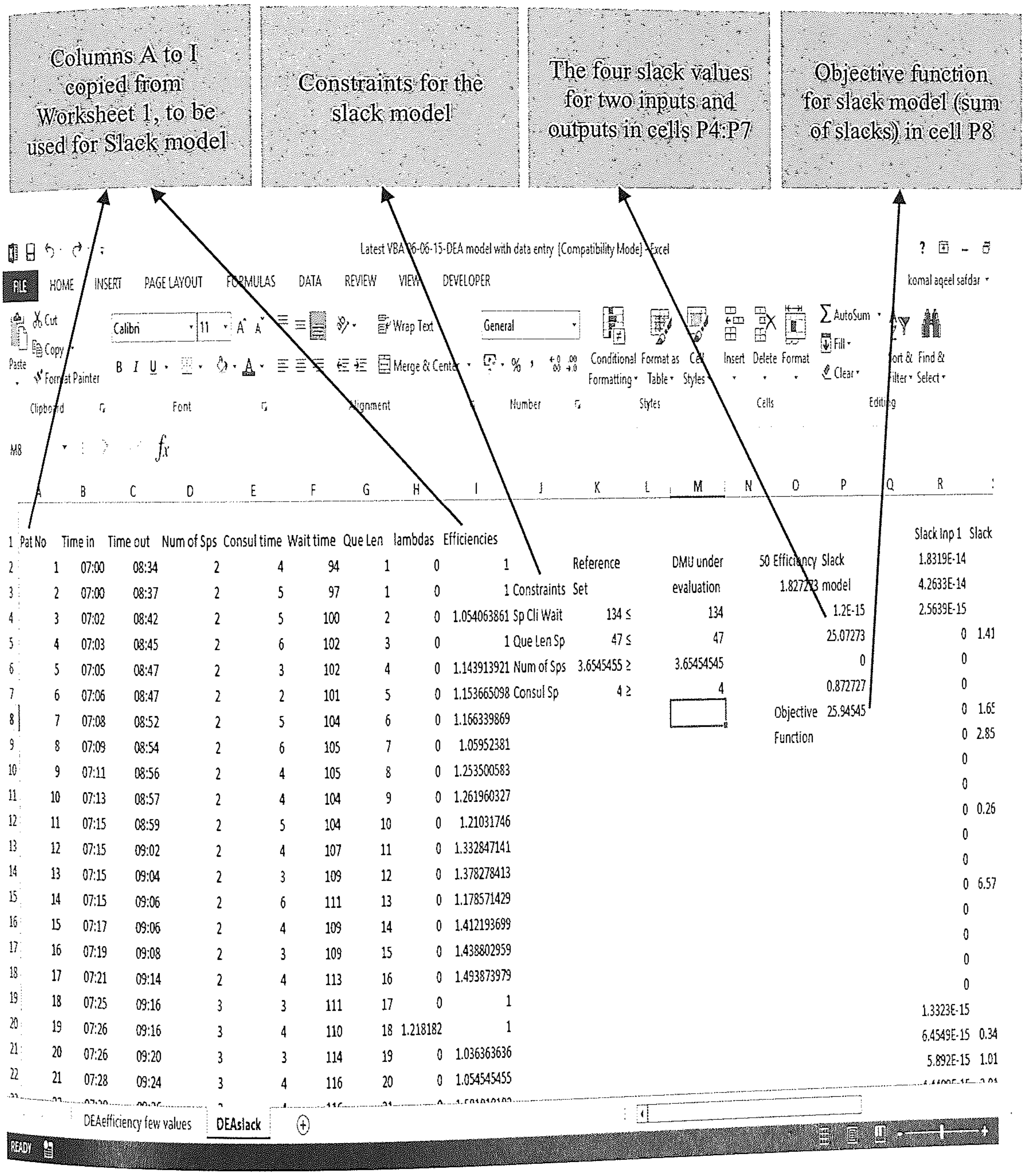
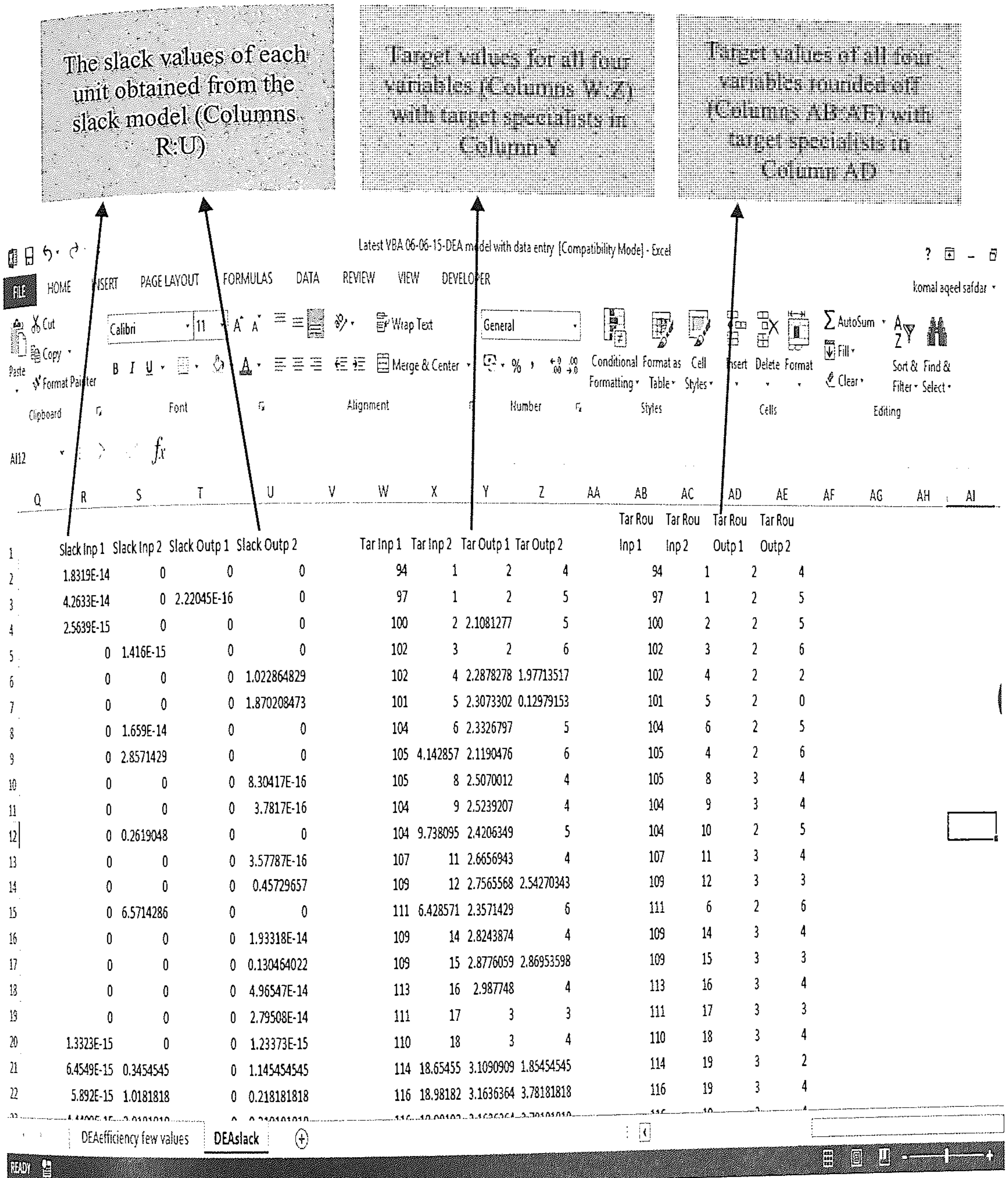


Figure 5-9: Worksheet 'DEAslack' of the Dynamic Framework Displaying the Slack Model Developed Using VBA Coding in Excel

Using the slack values and the original input/output values, the *target values* of each input and output are calculated and shown in Columns W to Z, as shown in Figure 5-10. Additionally, the target values are then rounded and displayed in Columns AB to AE. In this case, Column AD represents the target values of the number of specialists. This column is copied to Column P of the first worksheet. Also, the VBA code is written such that after all models are run and the target values calculated and copied, the active worksheet is first one. This is to ensure that the users do not have to manually shift to the first worksheet as this is where the data needs to be entered manually. Although the DEA model has been developed for 50 patients, but it can be extended to any number of patients such as the maximum load experienced in any department, which can be up to 300 patients. The following rows which do not have any values, will have a value zero, since there are no patients and they are empty. However, as each patient's data will be entered, the required doctors and efficiency will keep updating *'automatically'*.



	R	S	T	U	W	X	Y	Z	AB	AC	AD	AE
1	Slack Inp 1	Slack Inp 2	Slack Outp 1	Slack Outp 2	Tar Inp 1	Tar Inp 2	Tar Outp 1	Tar Outp 2	Tar Rou Inp 1	Tar Rou Inp 2	Tar Rou Outp 1	Tar Rou Outp 2
2	1.8319E-14	0	0	0	94	1	2	4	94	1	2	4
3	4.2633E-14	0	2.22045E-16	0	97	1	2	5	97	1	2	5
4	2.5639E-15	0	0	0	100	2	2.1081277	5	100	2	2	5
5	0	1.416E-15	0	0	102	3	2	6	102	3	2	6
6	0	0	0	1.022864829	102	4	2.2878278	1.97713517	102	4	2	2
7	0	0	0	1.870208473	101	5	2.3073302	0.12979153	101	5	2	0
8	0	1.659E-14	0	0	104	6	2.3326797	5	104	6	2	5
9	0	2.8571429	0	0	105	4.142857	2.1190476	6	105	4	2	6
10	0	0	0	8.30417E-16	105	8	2.5070012	4	105	8	3	4
11	0	0	0	3.7817E-16	104	9	2.5239207	4	104	9	3	4
12	0	0.2619048	0	0	104	9.738095	2.4206349	5	104	10	2	5
13	0	0	0	3.57787E-16	107	11	2.6656943	4	107	11	3	4
14	0	0	0	0.45729657	109	12	2.7565568	2.54270343	109	12	3	3
15	0	6.5714286	0	0	111	6.428571	2.3571429	6	111	6	2	6
16	0	0	0	1.93318E-14	109	14	2.8243874	4	109	14	3	4
17	0	0	0	0.130464022	109	15	2.8776059	2.86953598	109	15	3	3
18	0	0	0	4.96547E-14	113	16	2.987748	4	113	16	3	4
19	0	0	0	2.79508E-14	111	17	3	3	111	17	3	3
20	1.3323E-15	0	0	1.23373E-15	110	18	3	4	110	18	3	4
21	6.4549E-15	0.3454545	0	1.145454545	114	18.65455	3.1090909	1.85454545	114	19	3	2
22	5.892E-15	1.0181818	0	0.218181818	116	18.98182	3.1636364	3.78181818	116	19	3	4

Figure 5-10: Worksheet 'DEAslack' Displaying Slack Values and Target Values for all Inputs and Outputs for the Queuing-DEA Model

Hence, the dynamic framework consisting of the Queuing-DEA model was developed within *Microsoft Excel* spreadsheet, to assess the queue continuously by automating the DEA model, for implementation. The dynamic framework was validated by developing a simulated dataset in Excel to observe if the desired results are achieved. Varying values of wait time, length of queue and current availability of doctors, were utilized; where change in the efficiency level and required number of doctors was observed, which is the main purpose of the framework. The major advantage of this DEA model is that it keeps on updating continuously as soon as a new patient arrives, and previous patients leave. The VBA macros have been developed in a way such that it requires less work for the receptionists. As soon as they enter the time out and consultation time of a patient that has just left, the DEA model will re-run automatically, and show the 'latest' efficiency levels and requirement of doctors at this stage.

5.2.3 Summary and Practical Usefulness of the Dynamic Framework for Queue Assessment

The DEA model as proposed in Section 4.2.6.2 has been applied to the busiest departments at the designated hospital in Pakistan, in order to evaluate the excessive queue situation. The main conclusions drawn from the DEA results are that the wait times and load of patients are excessive, the load of patients changes drastically just within few minutes due to absence of appointment systems, and the inappropriate scheduling of doctors with unfeasible target number of doctors in case of excessive queuing. Hence, all these factors create hurdles in the efficient management of patients, and hence lead to excessive wait times.

In this regard, the proposed *dynamic framework* allows for the practical implementation of the Queuing-DEA model due to the provision of useful results, and to *continuously* monitor the queue situation with the arrival of each patient. Hence, the administration will be able to cater for the *quickly* changing queue situation due to walk-in patients. Since there are no effective strategies in place to control the queuing problem, this framework provides an *'exact'* requirement of doctors at a specific stage. In this way, the administrators can re-assign the doctors from other tasks to where they are needed the most. This action taken earlier on, will greatly minimize the wait times of subsequent patients, and ensure optimal staff utilization as more doctors are present when required and less when the OPD is not crowded. Even if the administrators are not able to fully follow the suggestions as given by the model all the time, but they have a guideline which allows them to take immediate action most of the times. Additionally, the dynamic framework provides some flexibility with regard to the 'timing' of the proposed action taken, as the queue can be monitored either continuously or at less frequent intervals. For instance, the administration may prefer to refer to the framework more frequently during early mornings say at a fifteen-minute interval from 7-9am, and less frequent intervals say every one hour from 9am-12pm; and take action accordingly. Furthermore, if desired by the administrators, a 'threshold' can be set either using the required number of doctors or wait time, to fix the 'exact' time when action regarding re-allocation of personnel is undertaken. For example, the administrators may decide to re-assign personnel only when requirement 3 or above or if the wait time exceeds 2h. This scenario will maintain the flexibility which is main advantage of the dynamic framework, however at the same time, will provide a set time-frame for administrators to re-allocate staff where the 'exact' number will be provided by the framework.

The unpredictability in the arrival behaviour of patients due to absence of appointments, as well as rapidly changing queue situation, necessitates the use of a dynamic framework which shows the 'moment-to-moment' change in the queue system and allows the administration to take action depending on the queue situation. Therefore, the dynamic nature of this framework makes it extremely beneficial for a busy public hospital of a developing country such as Pakistan, where all patients are walk-in and inflow of patients cannot be controlled. Therefore, in this regard, a 'fixed' schedule for staff availability will be less beneficial. The variation in the queue during every hour of a day and during different days of the week requires a dynamic framework as proposed in the current study, which provides recommendations for the required number of personnel 'as the queue situation changes'. Consequently, the wait times of patients can be minimized, and administrators can ensure greater availability when queue is excessive at the OPD, and re-assigning personnel for other activities when the queue is less. However, if the administrators are willing to construct a 'fixed' schedule for the availability of personnel beforehand, large amount of data need to be collected over longer periods to identify a pattern in patient arrivals which can then be used to prepare a schedule for a day/week. Nevertheless, a 'dynamic' framework which allows the administrators to take action as and when required proves to be more effective for a constantly changing complex queue system.

Furthermore, the dynamic framework has been developed in *Microsoft Excel* due to its 'easy-to-use' functionality. Excel does not require internet or a license (short-term which expires) like other sophisticated software packages available these days. Although, these software packages may be better suited for the purpose, however, they are not feasible with regard to *implementation* in *resource-poor* public hospitals in *developing* countries, with lack of technical skills among the staff members. The use of Excel requires minimal training for the staff members, since almost all staff members are familiar with its application, and it is available on most of the desktop computers at Reception desks.

Therefore, the emphasis is to ensure smooth and quick implementation of the dynamic framework, with easily understandable structure, as opposed to suitability and sophistication. The DEA model will run in the background without displaying any technical functions to avoid confusion. For the sake of convenience, the dynamic framework has been developed as such that the entire DEA model calculations have been 'automated' using appropriate VBA codes This is to ensure that minimal work is required by the receptionists/other users. They are only required to enter the 'time in' and the 'number of specialists' when a patient arrives, and 'time out' and 'consultation time' when a patient leaves. At this instant, the entire DEA programme will re-run automatically, displaying the required number of doctors in another column on the same spreadsheet to avoid switching between worksheets. This will ensure *quick* response to the changing queuing situation.

Therefore, the *dynamic* framework dedicated towards the practical *implementation* is extremely beneficial with regard to continuous assessment of the rapidly changing queue situation to ensure timely action in terms of bringing in more doctors; as well as for optimal utilization of personnel. Additionally, the proposed framework is generic such that it can be applied in any department of any busy public hospital with no appointment systems even in other developing countries.

A summary of the usefulness of the dynamic framework is given in Figure 5-11.

Usefulness of the Proposed Dynamic Framework

Continuous monitoring of the 'quickly' changing queuing situation with random patient arrival pattern

Identifies the 'exact' requirement of personnel such that immediate action can be taken to minimize the queuing of subsequent patients and optimal utilization of personnel

'Feasible' and 'easy-to-implement' with automated running of DEA modelling requiring minimal training and work for the users (receptionists)

'Generic' and 'flexible' framework which can be applied in any department any busy public hospital with no appointment systems in other developing countries

Figure 5-11: Summary of the Usefulness of the Dynamic Framework Developed to Assess the Queue System

5.3 Additional Issues and Recommendations for Assessment of the Queue System

A few *additional issues* influencing the overload and wait times of patients, were also observed or pointed out by the staff members. These factors only have a minor effect on the overload and wait times of patients, and it might not be possible to take any permanent action instantly. However, they can be taken into consideration for any future measures with respect to queue management.

As mentioned in Section 4.2.6.1, it was observed in the busy specialist OPD that patients, although being aware of the *start* of the OPD session (at 8:30am) arrived early morning, between 7 and 7:30am. One of the main reasons for this early arrival is that patients are of the opinion that the sooner they arrive, the earlier they will be examined as soon as the OPD session starts. Therefore, they would prefer to wait before the start time obtaining an early token number, and are among the first ones to be seen when the OPD session starts. Hence, it allows them to return back to other commitments without wasting the whole day. Furthermore, as mentioned in Section 4.1.2, the designated hospital is a *large* public hospital which caters to many towns and villages in the suburbs of the city. Patients who have travelled a long way, prefer to come early and be examined as soon as possible to have sufficient time to get back. The staff members also commented that the designated hospital has some highly specialized units (oncology, neurology, specialized urology and orthopaedics), therefore, more patients prefer to visit this hospital. Although the influx of patients cannot be controlled, due to absence of appointment systems, but the queue of patients can be better managed by taking immediate action with respect to optimum availability of doctors, as proposed via the dynamic framework.

Additionally, it was also observed in some *better* organized specialist OPDs, the sessions start at 08:00 *sharp*, with 2 to 3 doctors. In this way, maximum number of patients who arrived earlier in the morning were examined just within 40-45 minutes after their arrival, avoiding the building up of early morning rush. Although, it must also be considered that these OPDs are half as crowded as the busiest OPDs. It was observed that the OPD sessions within some busy departments always started with a delay of 15 to 20 minutes due to insufficient number of doctors, who might be involved with other non-OPD essential duties. Therefore, starting the OPD punctually, will reduce load of the patients. Although, it is possible that the patients will start arriving even earlier if the OPD session starts early, just to be among the first few to be examined.

Another observation made was the *difference* in the *load* of patients in early mornings as compared to later in the day, as mentioned in Section 4.2.6.1. It was observed that many patients arrive just within few minutes of each other during the mornings (between 7am and 8:30am), leading to a drastic increase in the overload. However, the pharmacy has an overload after 10am, as people usually collect their medicines once they complete all other activities. Therefore, the varying load of patients should be taken into consideration when defining strategies to control the load of patients. Also, usually Mondays and Fridays are the busiest days of the week (beginning and end of the week respectively). The flexible dynamic framework can be used to monitor the queue either at more or less frequent intervals, depending on the varying load of patients during different times of the day or different days of the week. Even a threshold for queue monitoring can be set depending on the level of overload within different departments, such as early versus late mornings and days of the week.

During interviews, some specialists commented that a large number of patients are *referred* unnecessarily to the specialist OPDs increasing the load of patients, although they can be examined by the GDMOs, requiring primary care only. They were of the opinion that the referral systems need to improve. However, the GDMOs remarked that they refer the patients to specialists even if they are a bit unsure about the diagnosis. Also, they stated that in some cases, patients do not 'trust' the advice given by GDMOs since the patients consider them inexperienced, and hence request for a referral to specialist. Therefore, for the sake of patient satisfaction, they refer them to a specialist. The administration might need to improve the referral systems within the hospital, or propose an initiative for long-term planning of having independent GDMO clinics somewhat similar to GP clinics in developed nations. Hence, the patients with only minor ailments, will be 'filtered' before they arrive at the hospital, which might lead to a reduction in the load of patients.

For two departments, it was observed that a *small* pharmacy is situated near the reception desk, which was responsible for providing most frequently used medicines. If any medicine was not available, only then the patients were advised to proceed to the main pharmacy. Although this is a long-term plan and might not be implementable immediately, however, the hospital administration can consider opening a small pharmacy within the busiest departments to provide most frequently used medicines. The data for the most commonly used medicines can be obtained from the main pharmacy. This will not only reduce the load of patients at the main pharmacy, but will be convenient for the patients as well, since they can obtain medicines from the same department where they were examined and leave sooner, and avoid waiting at the main pharmacy.

Another consideration is the *'type'* of patient including first-time, follow-up, regular and patients who just require a renewal of medicines. The medicine prescriptions are given for up to 3 months, and need to be signed off by the concerned doctor for any further renewal. The doctors commented that a separate counter for patients requiring a renewal only, can reduce the queue at the concerned department and wait times for these patients, and allow the doctors to spend more time on patients who require actual consultation. Additionally, it is also recommended that if this counter is situated at the pharmacy, it will be more beneficial for the patients. They can proceed directly to the pharmacy to obtain a new medical prescription and to collect medicines. Also, it was observed at the busiest specialist OPDs that the maximum number of patients are 'follow-up' or 'regular'. Therefore, the queues for first-time and follow-up/regular patients can be separated, as it might reduce the wait times for both groups of patients. However, this might not be practical in a developing country public hospital, since the patients will be confused with regard to their queue, or might intentionally join the queue which has less patients. An additional staff member directing patients to their respective queues may be of help in this scenario.

In order to improve the queue situation, some *other* recommendations were given by the staff members. One of the doctors suggested to hold regular seminars given by experienced doctors, to increase awareness among patients regarding certain diseases. The patients can be informed about the symptoms, diagnosis and prevention of some common diseases such as hepatitis, malaria, chest infection, malnutrition in children and others. This might reduce the overload of patients, since some patients only visit to ask for further clarification or are over-concerned about a minor disease. A few junior doctors suggested a change in assigning the duty shift patterns. They commented that sometimes they have early morning duties although they had a night shift on the previous day. This might be one of the reasons for late start of some OPD sessions. Additionally, other activities such as paperwork/administrative work adds to the burden of work. These activities may be undertaken by other staff members. However, these issues are debatable as the administrators have to make the best use of available personnel. Furthermore, it was observed that due to lack of proper guidance and low level of education, the patients are unable to comprehend the advice/diagnosis given by the doctors, instructions regarding medicines or tests/investigations to be conducted or referral to specific specialist OPD as advised by the GDMO. Hence, sometimes the patients visit the doctors multiple times to clarify certain matters. This not only prevents a smooth flow of patients, but also leads to over-crowding. Also, the doctors are unable to spend too much time on one patient, since they have a large number of patients to examine. Therefore, separate information desks might be set-up at different parts of the hospital to assist patients regarding different issues.

These additional concerns (as shown in Figure 5-12 below) might affect the overload and wait times of patients to some extent. Some of these issues are addressed by the dynamic framework, however, immediate action might *not* be possible for some recommendations but they can be taken into account when developing *long-term* strategies for improving the overall patient flow system.

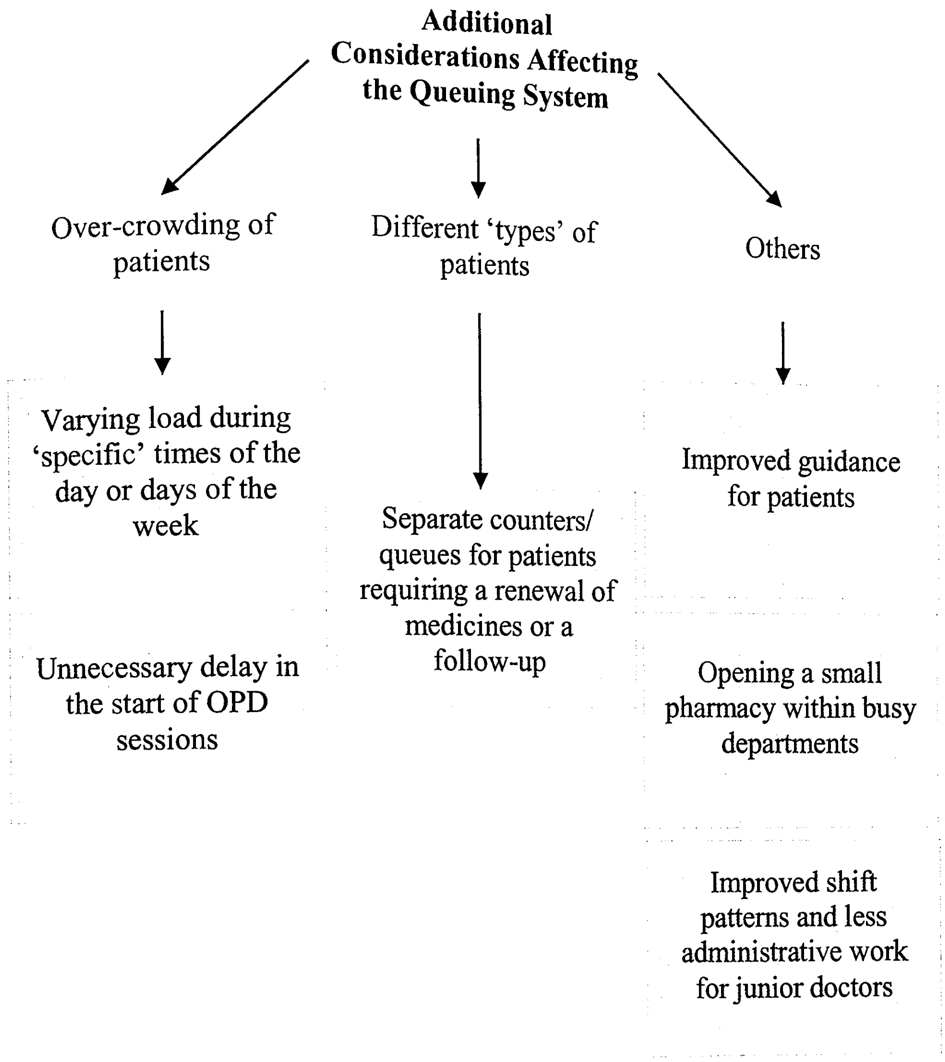


Figure 5-12: Summary of Additional Considerations for Assessing the Queue System

Chapter Summary

The current chapter consists of two main sections including the DEA analysis and results when evaluating the queue system within the designated hospital, and proposition for the practical implementation of the model. Considering the former section, the Queuing-DEA model proposed in Section 4.5.3 has been applied to the three busiest departments of the designated hospital in Pakistan. These departments include the busiest specialist OPD, the pharmacy and the GDMO clinic. The main aim was to determine the target personnel, given the excessive wait times and length of queue coupled with variability in the arrival pattern of patients; keeping the consultation time 'fixed'. Hence, this significant information will provide a guideline to the administration for preparing an adequate scheduling of personnel, syncing the availability of personnel with the overload of patients.

The preliminary analysis demonstrated some outliers in the data set for each department. These data points were disregarded when conducting further analysis to maintain homogeneity within the results and improve the accuracy. Additionally, an in-depth analysis was conducted for both efficient and inefficient units to provide a more comprehensive analysis regarding the required number of doctors at various points in the queue system. The availability of doctors was different for all three departments. The availability of doctors varied between 2 to 5 for the specialist OPD and between 2 and 3 for GDMO clinic, however, it was fixed at 2 for the pharmacy. The 'ideal' situation represented by the efficient units for the three departments included an availability of doctors of 4 and 3 for the specialist OPD and GDMO clinic respectively, and 2 for the pharmacy. The corresponding wait times and length of queue were 1.5h and 12 to 18, 0.5h and 5, and 0.5h and 4 for the specialist OPD, GDMO clinic and pharmacy respectively. One additional increase in the requirement of personnel was observed when the wait time and length of queue changed from 13 to 24 minutes and 10 to 20, 11 and 24 minutes and 3 to 4 for the specialist OPD and pharmacy respectively. However, in case of the GDMO clinic, it was observed that a slight change in the two inputs of 7 to 13 minutes and 1 to 3, led to one additional increase in the required number of doctors. Additionally, it was also observed that the target number of doctors increased drastically to an unfeasible level in some instances with very high wait times and length of queue.

The analytical results provide evidence that the proposed Queuing-DEA model has the capability to capture the rapidly changing queuing situation due to absence of appointment systems, and to provide an 'exact' requirement of personnel. This crucial information will assist in controlling the queue quickly, before it builds up excessively demonstrating an unfeasible requirement of personnel. Hence, this will result in minimization of the wait times of subsequent patients and lead to optimal utilization of personnel.

The latter section is dedicated towards the practical implementation of the Queuing-DEA model, based on the useful analytical results provided previously. In this regard a dynamic framework has been developed which allows for constant supervision of the queue system, and provides the 'exact' requirement of doctors so that timely action can be taken to control the queue 'as it happens'. With regard to the dynamic framework, different sections of the spreadsheet represent different aspects of the Queuing-DEA model. The first five columns require the data to be entered manually by the receptionists/other users, which will provide the basis for running the Queuing-DEA model. In this case, the patient name and number, time of arrival of the patient and the current availability of doctors will be entered when the patient arrives. Additionally, when the patient leaves, the time of departure and consultation time will be entered, where the consultation time is recorded by the doctors or their assistants. As soon as this information is added, the wait time and length of queue will be calculated 'automatically', and are not required to be entered manually. Hence the DEA model developed within the framework will run, displaying the 'required' number of personnel. Immediate action can be taken when the required number is greater than the current availability, or based on consensus with the administrators, a 'threshold' can be set.

The Queuing-DEA model has been developed within the proposed dynamic framework in Microsoft Excel using VBA coding. In this case, the VBA code was run first for the DEA model, and then for the slack model. Both models have been developed in separate worksheets. The slack values for all input and output variables were then used to calculate the 'target' values for all variables, including the number of the doctors. However, for the sake of convenience, the target values are then copied to the first worksheet as this is the primary section. In this way the users do not need to switch between worksheets every time the model runs, for the sake of convenience. The DEA model has been set-up for a total of 50 patients at the moment, where the model will re-run every time a new patient arrives and the data is entered. The VBA code for the entire DEA model has been 'automated'. As soon as the information is entered by the users, the model will re-run 'automatically' displaying the 'required' number of doctors for each unit, without any manual input from the users, simplifying the functioning of the dynamic framework.

Therefore, this dynamic framework is the first step towards the practical implementation of a queue assessment model to continuously monitor the rapidly changing queuing situation due to absence of appointments. With updated information regarding the 'required' number of doctors at that time, the administrators can quickly re-assign personnel from other tasks to where they are needed the most; in order to reduce the wait times of following patients and to allow for optimal staff allocation as synced with variable arrival pattern of walk-in patients. The model has been developed in Excel and has been simplified for the sake of users, to ensure smooth and quick implementation;

and requires minimal training for the users. Furthermore, the DEA model has been developed such that it runs automatically, in the background, requiring minimum effort from the users. Moreover, the proposed dynamic framework is generic and can be developed for other public hospitals in other developing countries, with no appointment systems and variable patient overload.

Some additional issues affecting the wait times and overload of patients, were also observed or suggested by the respondents. These factors include greater overload during early mornings in some OPDs as compared to later in the day, varying load on different days of the week where Mondays and Fridays are usually the busiest; and delay in the start of OPD sessions on a regular basis where the unavailability of doctors might be one of the factors. In this case, the dynamic framework proves to be very useful since it allows for the flexibility to assess the queue situation at more or less frequent intervals, depending on the load. Another issue identified is the different types of patients arriving in a day. Since some patients only come for renewal of medicines, it might be possible to have a separate desk or counter for these patients at the pharmacy so that they can proceed directly to the pharmacy. Additionally, it might be considered to have separate queues for first-time and follow-up patients. However, this might lead to increased mismanagement as patients might not be aware of this differentiation or join less busy queues intentionally. Some other recommendations which can be taken into account to address the overloading may include opening a small pharmacy within each department to provide more frequently used medicines, and awareness seminars for most common ailments to prevent patients from turning up just for clarification. Other suggestions can be improved guidance for the patients by other staff members to provide further explanation for treatment, and improved shift patterns and less administrative work for junior doctors. Although, immediate action might not be possible for most of these recommendations, however, they can be considered when preparing any long-term strategies for improving the patient flow system.

CHAPTER 6 DISCUSSION AND CONCLUSION

Chapter Overview

The current chapter primarily highlights the importance of the research context, proposed model and analytical results, and practical implications of the current study.

The first section of the chapter gives a detailed discussion of the main aspects of the current study by providing a review for each of the three research objectives. The discussion will emphasize on the significant aspects of the current study, with respect to each of the three research objectives.

The next section emphasizes on the contribution of the current study, which is also further divided into three categories including contribution within the research context, theoretical contribution and practical contribution.

The following section elaborates upon the limitations of the current study, and the extent to which these considerations posed challenges in conducting the current research.

Lastly, the chapter will conclude with exploring some future research avenues based on the modelling and analytical results from the current research study.

6.1 Discussion

The current study addresses the significant issue of excessive queuing in the outpatients' department of a busy public hospital of a developing country, where appointment systems do not exist; using Pakistan as an exemplar. Data Envelopment Analysis (DEA) modelling has been employed to assess and improve the complex queue system, extending its application in the context of queue management.

6.1.1 Research Objective 1: <u>*To identify various factors which lead to excessive queuing in a busy Public hospital of a developing country, where all Outpatients are walk-in*</u>

A number of *major factors* have been identified as crucial for the assessment of queue system of walk-in patients, within a large public hospital of a developing country, Pakistan Due to the absence of appointment systems, there is increased unpredictability in the arrival pattern of patients with extreme overload as the number of patients is not fixed; resulting in excessive wait times. High overcrowding and wait times were observed for walk-in patients at the three busiest departments of the designated hospital including General Duty Medical Officer (GDMO) clinic, specialist outpatients' department (OPD) and the pharmacy. The preliminary findings indicated that the inadequate scheduling of personnel is the most critical issue leading to excessive wait times. Additionally, another interesting finding of the current study is the extremely low consultation, which contradicts most of the previous works. Furthermore, the consultation time does not vary for different types of patients.

Very few studies have considered *walk-in* patients in an *outpatient* setting. Additionally, these studies have evaluated the walk-in patients with respect to improving an already existing appointment system. The current study, however, explicitly targeted a large public hospital of a developing country, where 'all' outpatients are walk-in, and analysed the associated operational aspects of a queue system, in the absence of appointment system. Zhu *et al.*, (2012) considered some uncertainties with the objective of improving the appointment system; including patients arriving earlier than appointment, late start of clinic by nearly 30 minutes and uneven distribution of appointment slots. Fetter and Thompson (1966) considered the change in the appointment interval from 15 to 20 minutes using different patient loads. For patient loads of 70% (26 and 21 for 15 and 20 minute interval respectively) and 93% (31 and 26 for each interval respectively), the authors concluded that only minor differences were found in the patient wait time.

Aharonson-Daniel *et al.,* (1996) simulated an optimal appointment system where 50 patients were allowed within every 30-minute interval, for a total of 3.5 hours (total 50 patients). The results showed the mean consultation time of about 2 minutes while the waiting time was over 50 minutes. Cayirli and Gunes (2014) evaluated the effect of seasonal walk-ins to determine appropriate appointment rules by considering the probability of walk-ins at 20% and 40%, representing low and high walk-in demand respectively. With appointments, the number of patients can be fixed and the overload can be controlled. However, the current study showed that when all patients are walk-in, the overload and wait times are excessive, as the influx of patients cannot be controlled. For instance, as observed in the busiest specialist OPD, the load can increase to nearly 50 patients 'at one time' with wait times of up to a high value 4 hours, signifying the severity of the queuing situation.

Due to the *absence* of appointment systems, the current study demonstrated *high wait times* as well as extreme variability in the wait times experienced by the patients. Huarng and Lee (1996) showed that the wait time was 30.59 minutes, which decreased further to 19.9 minutes when an additional OPD session was added in a dermatology department. A study by Harper and Gamlin (2003) showed that the average wait time is 26 minutes with standard deviation 25.4 minutes, and the wait time varied between 18 to 34 minutes with variation in the standard deviation between 15.5 to 41.1 minutes. Adeleke *et al.,* (2009) displayed that the average time in the system is 51 minutes which is a total of waiting and service time. Baril *et al.,* (2016) also showed that the total lead time including wait and service time, was around an hour (74.76 minutes) where the maximum time (93%) was spent waiting, which was 69.31 minutes. As compared to these studies, the wait times observed in the current study are exceedingly high. For the busy specialist OPD, maximum patients had wait times between 2.5h and 3.5h, which is nearly double or three times the wait time of most previous works. In a recent study, Oche and Adamu (2014) demonstrated that 61% of patients waited between 1.5h and 3h to see the doctor. Although closer, but the current study displays slightly higher wait times, on average. Additionally, the results show a much larger range of variation such as between 0.5h and 3h for GDMO clinic and pharmacy, and between 1h and 4h for the busy specialist OPD. The current study addresses the high wait time and variation by means of controlling the queue using a continuous monitoring tool.

In addition to the wait time, *extreme overload* was also observed in the busiest departments with high variation in the arrival pattern of patients. Cote (1999) considered 6 patients per hour as 'busy' arrival rate. By changing the load of patients from 60 to 70% (23 to 26 patients for the appointment system), Fetter and Thompson (1966) considered the effect on the wait times of appointive and non-appointive patients which was 69 and 75 minutes per day respectively. Aharonson-Daniel *et al.,* (1996) considered different consultation rooms and observed that for one room, the queue reaches a peak of about 35 patients in less than two hours, indicating that all patients will not be served within the four hour session length. Laskowski *et al.,* (2009) when evaluating an emergency department concluded that with two doctors, the patient queue is increasing but with four doctors the queue is zero with staff being underutilized. For an optimal appointment schedule, Tang *et al.,* (2014) evaluated the allocation of slots for urgent patients such that there is a maximum of 16 patients, with routine patients varying between 1 and 15. The arrival rate of patients at the designated hospital is higher and more variable as compared to previous works, and changes just within few minutes. The findings for the busy specialist OPD indicated that usually early mornings are the busiest, with the load of patients of nearly 50-60 patients gathering within as less as an hour. These patients were not examined before 2.5h to 3h after arrival, and more patients arrive by that time. Hence, increased load of patients and uncertainty in the arrival behaviour, makes it extremely challenging for the management to differentiate between busy and slack periods. The current study proposes a dynamic framework to best handle the uncertainties efficiently, for better queue management.

Low consultation time was observed for the current study, which was similar for all types of patients. This is contradictory to most previous works regarding appointment systems, where the consultation time is quite high and the objective is to minimize it. Fetter and Thompson (1966) observed varying consultation times for appointive and non-appointive patients at 12.6 and 9.8 minutes respectively. When allocating appointment slots for 10 patients in four clinics, Cayirli *et al.,* (2008) considered mean consultation time as fixed for returning patients at 15 minutes, and varied between 20-30 minutes for new patients. When evaluating the appointment schedule for an oncology department, Santibanez *et al.,* (2009) observed that the average consultation lasts around 1h and is highly variable for new patients whereas for follow-up patients, consultation time is 15 minutes with less variability. Bard *et al.,* (2014) concluded that if nine patients are to be examined according to the appointment schedule, a consultation time of 35 minutes is quite high and should be reduced to 20-25 minutes per patient. For the current study, the findings indicate that the average consultation time is only 6 minutes on average for the specialist OPD and GDMO clinic, with only 4-5 minutes service time at the pharmacy. Additionally, in almost all cases, the consultation time did not vary according to different types of patients including first-time and regular/follow-up and renewal of medicines. Only a few rare cases, such as patients requiring a detailed examination (children) or patients with multiple complaints (mostly aged patients), had a slightly higher consultation between 8-12 minutes. Even for the pharmacy, the service time increased only a little to about 7-9 minutes in some cases, for instance, patients which required a large number of medicines or additional explanation or medicines requiring special permission. The consultation time observed in the studies by Aharonson-Daniel *et al.,* (1996) and Oche and Adamu (2014) was 3.3 minutes and 7 minutes respectively, which is closer to the average consultation time of 6 minutes for the current study. However, for the former study, the corresponding average wait time obtained was 75 minutes, which is much lower than the current study as maximum patients had wait times between 2.5h and 3.5h. Considering the latter study, the consultation varied between 1 and 25 minutes, whereas the current study has a range of variation with half the width, between 2 and 12 minutes.

The *inappropriate scheduling* of personnel, has been identified as a major factor leading to excessive queuing. Liu and Liu (1998) utilized different number of doctors (2, 3 and 5) and also considered the varying arrival times of these doctors with arrival interval within 0 and 6, to assess the optimal appointment schedule. Considering staff scheduling, Rising *et al.,* (1973) concluded that five physicians should work in the first two hours and seven in the last six hours, for identifying the best appointment schedule during an eight-hour day session. The final changes showed that although, the overall wait time remained the same, however, the mean wait time for walk-in patients decreased from 38 to 28 minutes and increased from 12 to 26 minutes for appointment patients due to increase in the number of appointments. When evaluating an emergency department, Lane *et al.,* (2000) concluded that although more doctors are working from 7am till early afternoon (seven doctors) but the availability is not enough to deal with the huge number of emergency patients. Also, early afternoons are less busy but there is a rush of patients in the evenings. Paul and Lin (2012) showed that addition of an extra physician in the main emergency department during busy hours which is between 11am and midnight, improved the daily patient throughput and average length of stay. In another study, Bruni and Detti (2014) identified a number of factors that need to be considered when preparing a staff roster. These factors included allocating shifts for different days (weekdays, weekends, nights and public holidays) and equal allocation of shifts among the staff. The current study has demonstrated that although the availability of doctors varies throughout the day (such as for specialist OPD and GDMO clinic), however, it is not synced with the excessive queuing situation at a certain time. Also, with the rapidly changing queuing situation due to walk-in patients, a fixed schedule is not considered appropriate. Hence, the current study developed a dynamic framework which illustrates the exact requirement of personnel and updates continuously with changing queue situation; allowing for an optimal and flexible staff scheduling.

For the current study, *real-time* queuing data has been collected for each patient in the sample for the busiest departments of the designated hospital in Pakistan. Rising *et al.,* (1973) utilized the arrival pattern from previous years to estimate the number of walk-in patients for each weekday in the next year, which was then used to determine the best allocation of appointments. However, since all outpatients are walk-in within the designated hospital, therefore, real-time queuing data has been gathered for all weekdays, in order to capture the unpredictable arrival behaviour of walk-in patients. Zhu *et al.,* (2012) used a survey to collect data of patients with regard to improving an appointment system. Information was recorded regarding the type of patient and detailed timings at every stage (appointment, arrival and consultation time), and some staff remarks were also added. The current study used a time sheet mainly to record the wait times where the data was gathered separately for each department. Although similar to the method adopted by Zhu *et al.,* (2012), however, the time sheet used for observations within the current study was simplified to allow for 'quick' data collection mainly regarding wait times of patients, and cover a large sample as the number of patients were excessive.

Considering different patient categories, Harper and Gamlin (2003) collected data for different types of patients including new, returning, diary and urgent. The consultation time was recorded as different for each type of patients, varying between 1 and 35 minutes. Babes and Sarma (1991) noted that the consultation time was high for first visits and low for repeat visits, and consultation was much quicker when queue was long. Rising *et al.,* (1973) developed a priority queue where patients were seen in a sequence with first priority to emergency or those returning from labs, with next priority to appointment patients and lastly walk-in patients served on a first-come-first-served (FCFS) basis. Huarng and Lee (1996) identified that patients had low service times at some stages such as cash desk and pharmacy which was 1.10 and 1.18 minutes respectively. For the current study, the observations were made for *different types* of patients including first-time, follow-up and renewal of medicines. However, all patients joined a single queue served on a FCFS basis, and the consultation time was similar for all types of patients. Patients were triaged, and had low wait times and service times mainly if they are emergency cases or require a same-day second consultation, and at some stages such as reception desks and laboratory.

The *non-existent* appointment systems and the associated crucial factors identified prevent the smooth and efficient operation of a queue system within a typical busy public hospital of a developing country, using Pakistan as an exemplar; resulting in excessive queues and wait times.

6.1.2 Research Objective 2: *To develop a Queuing-DEA model and assess its usefulness for evaluating the queue system in a busy public hospital of a developing country, where 'all' Outpatients are walk-in*

Considering the operational characteristics identified and preliminary findings regarding the queue system within the designated hospital in Pakistan, DEA was employed to develop a queue assessment model. The proposed model is developed with the view of providing an in-depth analysis of the queue system for walk-in outpatients in a busy public hospital of a developing country.

Over the years, different Operational Research (OR) modelling techniques have been used to assess the queue problem. However, these techniques have some limitations which make them unfavourable in the context of the current study; which is to assess the queue system in the absence of appointment systems with high wait times and variable arrival pattern of patients. These *limitations* mainly include the use of 'average' values and pre-specified statistical distributions to represent a queue system, as used by almost all queue management studies.

For instance, to represent inter-arrival times of different types of non-appointed patients, Wijewickrama and Takakuwa (2005) utilized Exponential distribution with parameters, 1.5, 6.15 and 3 for walk-ins, patients for tests only and new patients respectively. Also, by changing the rate of transfer from 2% to 10%, it was observed that the average wait time for walk-ins and patients for tests decreased with little increase for appointed patients. Daultani *et al.,* (2016) used Erlang and uniform distributions to represent the service time by consultant and two different types of machines. Rising *et al.,* (1973) compared the mean wait time as well as mean consultation time of appointed and non-appointed patients over two years, and concluded that they are similar. Liu and Liu (1998) used different distributions for the service time including Uniform, Exponential and Weibull. Additionally, varying arrival rate of physicians was considered using Uniform distribution. Griffiths *et al.,* (2005) considered negative exponential distribution for inter-arrival times and Pearson VI for length of stay of patient referrals to intensive care unit from different sources, including emergency surgery, emergency room and ward. Mandelbaum (2012) evaluated the process of transferring patients from emergency department to multiple internal wards, and considered Poisson and exponential distribution for arrival rate and length of stay respectively. Mital (2010) considered two types of average wait times including waiting for a short while and waiting within a 'specific' time interval. Additionally, the change in variation over a period of four years was determined for average bed occupancy, average length of stay and the arrival rate of patients within inpatient wards. For determining an optimal overlap appointment schedule, Anderson *et al.,* (2015) utilized different distributions for service time including Uniform, Exponential, Erlang and Normal. One of the main conclusions drawn was that for all service time distributions, overlapped saves more than 40% of total cost, when cost ratio is higher than 8 and no-show rate is above 30%. In developing countries, most public hospitals do not have a set appointment system, and experience extreme overloading with high variability in the arrival rate of patients. Therefore, this necessitates the need of an assessment tool, which provides more detailed analysis of the queuing situation by evaluating each patient separately. Hence, the current study evaluated the queue system at the designated hospital in Pakistan by utilizing the popular efficiency assessment technique of Data Envelopment Analysis (DEA).

Considering DEA in healthcare, a *limited* number of studies have highlighted and applied DEA modelling to assess other healthcare management aspects. Wagner *et al.,* (2003) and Osman *et al.,* (2011) compared the efficiency of physicians and nurses respectively. The former study assessed the efficiency of physicians with and without aggregation of inpatient and outpatient measures, however, the efficiency results with both aspects only differed slightly. The latter study evaluated the efficiency of nurses in an intensive care unit, specifically with regard to performance evaluation and appraisal. Vanelli *et al.,* (2013) assessed colorectal cancer screening programmes to assess their effectiveness with respect to long-term planning. In a recent study, Keshtkaran *et al.,* (2014) specifically compared the efficiency of radiology units of different hospitals, showing that DEA can be utilized for comparing specific units or departments as well. These studies indicate the usefulness of DEA modelling in other areas of healthcare, moving away from the more traditional application of efficiency assessment of hospitals. However, such studies are very few in number. The current study extends DEA modelling in the context of queue management in developing countries in the absence of appointment systems, considering patients as units of analysis. Therefore, by analysing the queuing data for 'each' patient, the present study demonstrated the effectiveness of DEA modelling for assessing a queue problem, diverging from the conventional applications.

The *preliminary observations and findings* highlighted a number of factors, which assisted in refining the proposed Queuing-DEA with regard to efficiency assessment of the queue system. The most significant finding is the inadequate scheduling of personnel. Given high wait times and overload due to absence of appointment systems, the crucial issue is to identify 'how' many personnel are required at a specific time. Hence, the number of personnel representing the current availability when a patient arrives, is included in the model as an output. In almost all previous DEA studies, the objective is to minimize the number of personnel to improve efficiency and has been considered as an input to the model. Huarng and McLaughlin (1989) and Chaung *et al.,* (2011) considered physicians as an input with nurses and other health professionals as additional inputs. The latter study also utilized the Classification and Regression Tree (CART) analysis to further analyse the required improvement in resources. For instance, one scenario suggested that in order to improve the efficiency of hospitals, the number of physicians and health professionals should be greater than 50 and less than 190 respectively. Harrison *et al.,* (2004) considered full-time employees as a single input to compare the efficiency of federal hospitals in two years, and showed that the average 'excess' in employees only changed slightly from 27 to 37. Sheikhzada *et al.,* (2012) considered two inputs for staff where the first was the sum of general physicians, nurses and personnel having a medical degree, whereas the second was the sum of medical team having a diploma of 24 years, non-medical and support staff. However, in a recent study for dealing with uncertainty in inputs and outputs using a Cross-efficiency Fuzzy DEA model, Dotoli *et al.,* (2015) considered three inputs including doctors, nurses and sum of other employees and administrative staff. However, Lindlbauer *et al.,* (2016) considered five separate staff-related inputs including physicians, nurses, administrative staff, other clinical staff and other non-clinical staff. However, the current study has added personnel as an 'output' to the DEA model for determining the required number of personnel to improve a queue situation in an outpatient setting.

Furthermore, the consultation time has been included in the model as a 'non-discretionary' output. As observed in the current study, the consultation/service time is already quite low (6 minutes for specialist OPD and GDMO clinic, and 5 minutes at pharmacy), and should not be curtailed any further as it might increase the chances of errors in critical diagnoses, and safe dispensing of medicines in case of pharmacy. Hence, it can be concluded that this variable has less weightage when analysing the queue problem in developing countries. However, it has been considered as a significant variable in other queue management studies; but mostly with the objective of minimizing it. Only a limited number of DEA studies have considered 'non-discretionary' variables, with negligible in the field of healthcare and even these studies have considered fixed inputs. Mostly socio-economic or demographic factors are considered fixed, where a second-stage analysis is conducted to assess their impact on efficiency. Afonso and Aubyn (2006) considered the effect of Gross Domestic Product (GDP), education, smoking habits and obesity on health services provided in different countries. Herwartz and Schrumann (2014) utilized some explanatory variables including profit or non-profit private hospitals, mortality rate, people aged above 65 years and some others. Additionally, Kyriacou *et al.,* (2015) also considered variables related to government quality to evaluate the redistribution of funds in various countries. The current study has evaluated the queue system of a public hospital in a developing country, where the consultation time is added as a 'non-discretionary' variable; with a view of keeping it in constant in the DEA analysis.

Due to the absence of appointment systems, the initial findings of the current study show excessive wait time and overload with rapid queue build-up. Hence, the wait time and length of queue have been considered as inputs to the model, where the latter has been utilized as an indicator of overload in the queue system of the outpatients' department.

Using *patients as 'units of analysis'* in the DEA model, the current study determined the required number of doctors using corresponding wait time and length of queue values, whilst keeping consultation time unchanged.

Most of the previous DEA studies have conducted an efficiency comparison of '*hospitals*'. The current study extended the application of DEA modelling specifically for the assessment of a queue system, using patient level data. Considering previous works, Ramanathan (2005) and Flokou *et al.,* (2011) concluded that 50% and 37% of the hospitals were efficient respectively. Kang *et al.,* (2014) assessed the efficiency of emergency departments of different hospitals by dividing into six groups based on patient volume, with fully efficient units in each group being 61.54%, 68.97%, 20.55%, 30.28%, 21.12% and 39.22%, with patient volumes 100, 80-100, 60-80, 40-60, 20-40 and 1-20 (all in thousands); displaying a significant difference between groups of large and small patient volumes. In another study, Akazili *et al.,* (2008) concluded that 65% of the health centres (58 out of 89) are inefficient, and further identified that 24% had an efficiency score of less than 50%. The current study, however, showed that nearly 80% of the observations in each department had an efficiency level below 55%, including specialist OPD (79%), pharmacy (89%) and GDMO clinic (83%). These results demonstrate a highly inefficient queue system with long wait times for patients.

Bwana (2015) compared the efficiency of Volunteering Agency Hospitals (VAHs) over a period of four years, and identified that the lowest efficiency score in 2009 till 2012 is 31.6%, 28.8%, 23.4% and 24.3% respectively. For the current study, a few observations had an efficiency score of even lower than 20% with the lowest score being 19% for specialist OPD and pharmacy and 15% for GDMO clinic. As observed, the wait times experienced by these patients have been a high 3h to 4h for the specialist OPD and between 2.5h to 3h for the GDMO clinic and pharmacy.

Kirigia *et al.,* (2008) assessed the efficiency of public hospitals over a period of three years, and concluded that from a total of 28 hospitals, 61%, 57%, and 64% were inefficient in 2000 to 2002 respectively. Both studies included number of physicians as an input, with nurses and clinical staff as the other staff input in Kirigia *et al.,* (2008) and Akazili *et al.,* (2008) respectively. These studies suggested that excess staff can be transferred to other hospitals with shortage of staff. However, such extreme measures are difficult to operationalize, particularly in developing countries. Therefore, the current study proposed a queue assessment framework which monitors the queue situation 'as it is', and provides recommendations to optimally utilize the available staff through quick and dynamic scheduling to minimize excessive wait times.

The current study mainly aimed at identifying the crucial information of *requirement* of *personnel* in a queue system without appointments, utilizing DEA modelling. The analytical results indicated high variability in the wait times, length of queue and the corresponding requirement of doctors within the busiest departments of the designated hospital. Additionally, the results showed that extremely high wait times or length of queue or a combination of both led to an unfeasible high target of personnel. Additionally, the target was high irrespective of the current availability of personnel which varied between 2 and 5 and 2 and 3 for the busy specialist OPD and GDMO clinic, but was fixed at 2 for the pharmacy. For the busy specialist OPD, nearly 40% of observations showed a high requirement of doctors of 9 to 13 when the corresponding wait time was between 3h and 4h, and queue length was between 28 and 58. Similarly, for some observations at the pharmacy and GDMO clinic, with wait time between 2.5h and 3h, the target personnel increased to between 9 and 10 and 13 and 16 respectively. The queue length varied from 23 to 34 for the pharmacy and from 26 to 29 for the GDMO clinic. Although these high targets might be unachievable, however, they provide evidence of the severity of the queuing problem and demonstrates the need to control queue pre-emptively. The Queuing-DEA model developed in the current study is a monitoring tool which acts as a guideline for the management to identify the 'exact' requirement of personnel at a particular time.

Furthermore, the results for nearly all observations at the GDMO clinic and a few data points in the specialist OPD and pharmacy showed, that a slight change in the wait time and length of queue resulted in one additional increase in the number of personnel. These results demonstrate the *unpredictable* arrival pattern of patients due to absence of appointments, which lead to drastic changes in the load of patients within few minutes. The proposed DEA model has the ability to assess the queue at frequent intervals to capture the rapid changes, and provide corresponding requirement of doctors.

Factors such as absence of appointment systems with random arrival pattern and excessive wait times of patients, along with lack of queue management strategies; necessitates optimal utilization of available staff. Equipped with information for 'required' personnel as displayed by the proposed model, the management can take immediate action ensuring *minimal* wait times of subsequent patients and *appropriate* scheduling of personnel. Additionally, the proposed Queuing-DEA model is flexible and *generic* and can be applied to other public hospitals in similar developing countries, where appointment systems are absent and excessive queuing prevails.

6.1.3 Research Objective 3: *To develop a dynamic framework for practical implementation of the proposed model for continuous monitoring of the queue system in the absence of appointments, within a large public hospital of a developing country*

The preliminary investigation of the queue system demonstrated that the inappropriate scheduling of the personnel is a major issue leading to excessive wait times and long queues. Currently, there are no definitive strategies which lead to better queue management for busy outpatients' departments within the designated hospital. The proposed Queuing-DEA model has the capability to determine required number of personnel, for dynamic and efficient scheduling of personnel which can be varied depending on the severe queue situation. Based on the significant usefulness of the proposed queue assessment model, the current study has developed a dynamic framework for practical implementation of the Queuing-DEA model; such that the rapidly changing queue situation can be constantly observed.

The proposed *dynamic framework* has been developed in *Excel*, using Visual Basic for Applications (VBA) programming. The dynamic framework portrays all the necessary information required to assess the queue, including wait time, length of queue, consultation time and the required number of personnel. The key benefit of the framework is that it displays the 'required' number of personnel 'at that moment'. Utilizing the information that the *required* availability has exceeded the current availability, the management can take quick action by re-assigning the personnel from less urgent jobs to the OPD sessions. Hence, by taking action early, the wait times of subsequent patients can be reduced and queue can be controlled before it reaches a heightened level as observed from the DEA analytical results.

Very few studies have considered the implementation of the derived analytical results. Among them, some have specifically utilized Excel for improving a queue system with regard to implementation. However, almost all of them are associated with appointment systems. None of the studies have specifically developed a framework for implementation for assessing a queue system with no appointment systems, whilst taking into account the operational challenges experienced by public hospitals in developing countries. In a study by Wijewickrama and Takakuwa (2008), multiple appointment schedules using different patient sequences were developed in Excel whereas an animated simulation model representing these schedules, was developed in Arena. Harper and Gamlin (2003) utilized Excel for preliminary analysis of patient flow systems in nine different clinics, with more detailed analysis using statistical packages such as Statistical Package for the Social Sciences (SPSS) and Statgraphics. The obtained data was then used to develop a Simulation model. However, when evaluating the sequencing and scheduling of patients for two types of patients including routine and same-day, Chen and Robinson (2014) recommended using Excel for post-analysis with regard to implementation of modelling results. Bruin *et al.,* (2010) specifically aimed at developing a decision support system in Excel using VBA for implementation purposes based on Erlang loss queue model; where historical data was directly obtained from the hospital information system regarding average length of stay, admissions and number of beds. The authors considered this system as user-friendly and technologically accessible. For the current study, the dynamic framework has been developed in Excel for implementation of the proposed model. The framework proves to be extremely useful for an outpatients' department where all patients are 'walk-in'; given highly variable arrival pattern of patients, excessive overload and long wait times of patients.

Zhang (2012) developed a decision support system using a simulation-optimization model for optimal capacity planning. In this case, Excel consisted of information such as arrival and length of stay distributions and wait times, which was added to Arena for simulation modelling, and resulting optimal solutions were recorded in Excel. The authors mentioned that this decision support system has been implemented in an organization where Arena was already available, and training for Arena and Statistical Analysis System (SAS) software packages was provided to the employees. Considering simulation models, Lehaney *et al.,* (1999) remarked that the hospital management required a transparent and easy-to-use model which is less costly, and represents the system accurately. The authors developed a simulation model to test different schedules for patient bookings using Simul8 which was considered suitable, avoiding any sophisticated software packages. Anderson *et al.,* (2015) utilized simulation models mainly to evaluate the effect of various overlap periods on the costs of patient no-show rate, wait time, idle time and overtime. The authors considered the issue of implementation and concluded that the simulation model, when identifying different scenarios, provide a basis for application in real-life. In resource-poor public hospitals of developing countries, advanced statistical packages are highly unlikely to be utilized due to sophisticated functionalities, even though they may be better suited for the purpose. For the current study, the dynamic framework has been developed in Excel with technical functions operating 'automatically' in the background, requiring little input from users with minimal training. Additionally, Excel is user-friendly with 'easy-to-understand' features, operates without internet, is currently available at the reception desks of outpatients' departments and most users are familiar with its application. Hence, Excel is a viable option encouraging smooth and quick implementation.

The proposed dynamic framework has been designed with the sole objective of practical *implementation* of the Queuing-DEA model for *better* queue management. The framework has been designed such that it continuously monitors 'real-time' queuing data of walk-in outpatients within a busy public hospital of a developing country, and updates 'automatically' every time a patient exits the system. The dynamic framework provides crucial information regarding the current required number of personnel. Hence, the management can take immediate action by increasing the availability of personnel, to control the queue pre-emptively, minimizing the wait times of subsequent patients and for optimum staff utilization. Even if the administrators are not able to fully follow the suggested recommendations all the time, but it will pinpoint when the queue is excessive. Additionally, the dynamic framework also provides flexibility with regard to the monitoring intervals. For instance, the administrators can set a threshold for the required personnel after which action will be taken, or might allow for more frequent monitoring during busy periods. Furthermore, the proposed dynamic framework has generic characteristics and can be considered for implementation in other similarpublic hospitals where appointment system does not exist, even in other developing countries.

A summary of the research objectives of the current study and their main characteristics are shown in Figure 6-1.

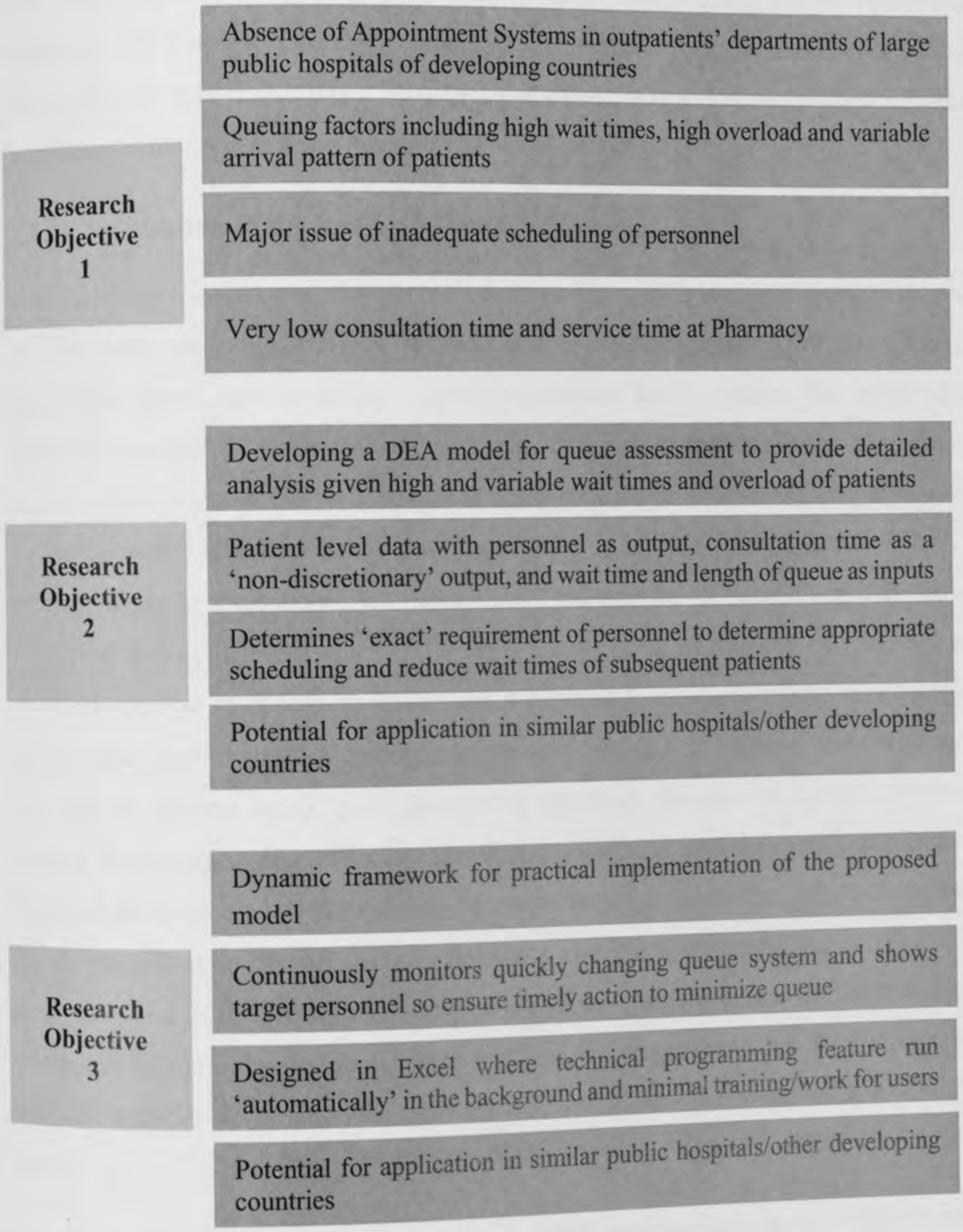


Figure 6-1: Research Objectives and Associated Components for the Current Study

6.2 Contribution of the Current Study

The current study evaluated the queue system of the outpatients' department within a busy public hospital of a developing country, in the absence of appointment systems. The proposed queue assessment model was developed using DEA, with the objective of providing recommendations for improving a queue system given that all patients are 'walk-in'. The study was conducted at a 'typical' large public hospital in Pakistan, which exemplifies excessive queuing for patients, and the prevailing challenges which constrain healthcare management from providing efficient services.

6.2.1 Contribution of the Research Context

The health care systems in developing countries face multiple challenges which create hurdles in the smooth operations of health delivery systems, making the life of ill quite miserable. These *aspects* mainly include overloaded health systems due to increased demand, shortage of staff, lack of resources, deficient budget and poor allocation of limited resources. Another crucial factor is the absence of appointment systems within the outpatients' departments of public hospitals. Additionally the excessive load of patients leads to excessive wait times. Since the number of patients and the arrival behaviour during different times of the day is not fixed, the management is unable to pre-plan and organize resources pre-emptively. Additionally, the implementation of complete appointment systems is not currently implementable, due to multiple economic, social and cultural barriers faced by the developing countries. The main challenges which can prevent the operation of an effective appointment system include low levels of education, large public hospitals accommodating patients who reside in surrounding towns/villages highly dependent on unreliable public transport, and lack of communication services such as internet and post. Hence, it is highly probable that patients do not understand, forget or are late for their appointments which will lead to further mismanagement. Also, given the huge number of patients, the appointments will need to be scheduled a few months ahead.

The ideal situation will be to have enough hospitals and resources to deal with excessive load of patients. However, this is a continuous long-term objective that all developing countries strive to achieve. It is imperative that the *complex* queue system for walk-in outpatients within large public hospitals is monitored *'as it is'*, with a specific guideline for improvement. The current study aimed at developing a dynamic queuing model which enables the management to continuously assess and improve the queue system for walk-in outpatients.

With regard to *outpatients*, almost all queue management studies have evaluated an already existing appointment system. Only a handful of these studies have considered *walk-ins* in the context of an outpatient setting, and those which have, only assessed the effect of walk-ins on determining the optimal appointment schedules. Additionally, there is dearth of literature in developing countries, and negligible in the South-east Asian region. The current study has been carried out in Pakistan, which is a representative of the poor healthcare conditions in developing countries. The present study specifically targeted the evaluation of a queue system with no prior appointments in an outpatients' department of a busy public hospital; and has identified the operational characteristics of different queuing variables particularly within this context.

6.2.2 Theoretical Contribution

Although, in recent years, extended healthcare applications of DEA modelling have been found but they are very limited in number and none of them have employed DEA in the context of queue management. The current study presents a novel application of DEA modelling by utilizing it for the evaluation of a queue system with no prior appointments, from a developing country perspective. Additionally, real-time queuing data of 'each' patient was evaluated using DEA, hence providing a more comprehensive analysis of the queuing situation for the outpatients' department at the designated hospital. Hence, the current study provides evidence that DEA modelling has the capability to perform efficiency analysis in other aspects of healthcare management, deviating from its more traditional usage for the efficiency comparison of health institutions.

The most *crucial* finding from the current study is the *inadequate scheduling* of personnel, therefore, the number of personnel (doctors/pharmacists) has been included as an output when developing the Queuing-DEA model. In a conventional DEA study of comparing the efficiency of hospitals, personnel has always been included as an input with the objective of minimizing it in order to improve efficiency. However, the current study is associated with improving the queuing system with insufficient availability of personnel which is inconsistent with the queue situation. Hence, in this case, the doctors/pharmacists have been included as an *output*. The proposed Queuing-DEA model provides an opportunity to pinpoint the extreme wait times and overload in the system, and determine the 'exact' corresponding target number of personnel. In developing countries, with absence of appointments, variable arrival pattern and rapid queue build-up, this additional knowledge proves to be extremely significant for better queue management. The proposed model demonstrating target personnel, provides explicit guidance to the management to take prompt action and re-allocate personnel in a dynamic fashion, synchronized with the excessive queue situation.

Additionally, *consultation/service* time was added as a '*non-discretionary*' output in the proposed model. Only a limited number of healthcare studies have included 'non-discretionary' variables, and they have considered fixed inputs, with negligible studies which consisted of a fixed output. Besides, mostly second stage analysis or some other technique is utilized to assess the effect of fixed variables. The findings indicated that the consultation time is already quite low, therefore, it could not be allowed to decrease any further and due to excessive queuing, it cannot be increased either. Hence, it was included as a fixed output. The proposed DEA model acknowledged the consultation/service time experienced by each patient, but kept it constant when determining the required number of personnel using wait times and length of queue. Hence the extended model provided an opportunity to assess the queue system taking into account the specific queuing characteristics defining the system. DEA analysis provides a comprehensive analysis of each unit of analysis considered with instant recommendations, overcoming the limitations of some other OR techniques, which makes it extremely suitable for the purpose of the study.

The extended DEA model designed in the current study has the potential for exploitation in similar public hospitals, even in other developing countries, given that there are no set appointment systems for outpatients with high and variable load of patients; which is the case in most developing countries.

6.2.3 Practical Contribution

The patient flow systems at the three busiest departments for outpatients at a busy public hospital in a developing country, Pakistan, were evaluated by applying the proposed Queuing-DEA model. These departments included the GDMO clinic, a busy specialist OPD and the pharmacy. The '*required*' number of personnel have been identified at each of three departments, given excessive wait times and extreme variability in the arrival pattern of patients due to the absence of appointment systems.

The analytical results from all three departments indicated that high wait times and length of queue cause the target personnel to increase *greatly*. In some cases, the DEA results even demonstrated an unfeasible target, irrespective of the current availability. This indicates that the queue builds up just within *few* minutes, and once it reaches an excessively high level, any intervention will not generate immediate results and it will take an additional number of hours to bring down the queue. Hence, these results highlight the criticality of the queue situation in busy public hospitals of developing countries, where all outpatients are walk-in. In this scenario, information regarding the 'exact' required number of personnel, will assist the management to respond *early,* and control the queue well before it spikes to a level where any action taken will be futile.

In addition to the benefits of the proposed model, *smooth implementation* is an equally crucial issue, but is extremely challenging when it comes to healthcare delivery systems in developing countries which face numerous challenges. Considering previous works in queue management and DEA, very few studies have actually considered the issue of implementation. Most of the recommendations proposed by previous studies require long-term planning, and might not be implementable in the short-run. In view of the usefulness of DEA results obtained in the current study and the overall increasing importance of implementation, a dynamic framework has been proposed which can be considered as a basis for the practical implementation of the Queuing-DEA model.

The dynamic framework has been developed such that it accommodates real-time queuing data, with the aim of *continuous monitoring* of the queue system and detect the rapidly changing queue situation. Furthermore, the dynamic framework has been designed in a way that it updates 'automatically' with the arrival of each new patient, displaying the *latest* required number of personnel. Equipped with this information, the management can re-assign the personnel (doctors/pharmacists) from less urgent jobs to the outpatients' department where they are needed the most. Consequently, the excessive wait times of subsequent patients and the load will be minimized which are major issues faced by public hospitals with no appointment systems in developing countries. Conversely, when the queue dies down, the personnel can be switched to other pending jobs. Hence, with continuous monitoring, the queue can be controlled instantly with *optimal* utilization of personnel.

The dynamic framework has been constructed in *MS Excel*. Although advanced software packages may be more efficient, however, Excel seemed more feasible with regard to implementation in resource-poor public hospitals of developing countries, due to its easy-to-understand functionality. Additionally, particular attention has been given to design the dynamic framework such that it requires *minimal* training for the users. The users are only required to enter the basic queuing data for each patient, and the system will run 'automatically' in the background, displaying the required number of personnel at that time. The emphasis has been to ensure smooth and *quick* implementation with easily understandable functional characteristics, with minimum effort for the staff.

Also, the proposed dynamic framework is generic and can be applied to similar public hospitals in other developing countries, to ensure constant supervision of the queue system at frequent intervals to determine any changes due to absence of appointments; and determine the required number of personnel leading to optimal utilization.

6.2.4 Summary of the Contribution of Current Study

The current study assessed the efficiency of an outpatients' department of a public hospital in a developing country, Pakistan, with respect to the queuing problem in the absence of appointment systems. The present study demonstrates the *extension* of DEA modelling with respect to evaluating a queue system using patient level data, and highlights the usefulness of the proposed model, hence diverging from the more traditional applications in healthcare. The analytical results obtained from applying the proposed model to the three busiest departments of the designated hospital, indicated high and extremely variable wait times and length of queue, with high target number of personnel. The essential information regarding the specific *requirement* of personnel, will allow the management to control the queue pre-emptively, and reduce the excessive wait times of subsequent patients. Additionally, a *dynamic framework* has been designed for smooth *implementation* of the proposed model. The framework enables *continuous monitoring* of the changing queue situation and allows the management to take immediate action with regard to *dynamic scheduling* of personnel. The dynamic framework has been designed in Excel with the technical aspects to operate automatically, to ensure straight-forward usage and minimal work for the users. The proposed Queuing-DEA model and the dynamic framework can be applied to similar public hospitals in other developing countries, for monitoring a queue system with non-existent appointment systems (see Figure 6-2 below).

Contribution of Current Study	
Research Context	Evaluated the queue system of an outpatients' department in a busy public hospitals of a developing country, in the absence of appointment systems
	Identified the operation of queuing factors specifically for a queue system with no appointments in developing countries
Theoretical	Extension of DEA Modelling in the context of a queue problem
	Including Personnel as output to determine required number of personnel to minimize excessive queuing, and consultation/service time as 'non-discretionary' output to keep it constant
	Potential for the proposed Queuing-DEA model to be applied in similar public hospitals of other developing countries, with no appointment systems for outpatients
Practical	Crucial information for management regarding 'exact' requirement of personnel to take early action and control queue quickly, minimizing wait times and for optimal utilization of staff
	Dynamic framework for implementation for constant monitoring of queue displaying required personnel using real-time queuing data, with straight-forward and automated technical features
	Generic framework which can be applied in similar resource-poor public hospitals, with no appointments in outpatients' department, even in other developing countries

Figure 6-2: Contribution of the Current Study

6.3 Limitations of the Current Study

The current study has a few limitations with regard to the methodological characteristics. The present study has utilized DEA modelling to assess the queue system of a large public hospital of a developing country. With numerous advantages of DEA, the modelling technique has some limitations with respect to the current study.

The DEA model can be sensitive to outlier data and the associated information is difficult to interpret (Charnes *et al.,* 1994). For the current study, the preliminary observations highlighted some extreme values in the datasets, which were disregarded when conducting further DEA analysis, in order to improve accuracy of results. Another aspect of DEA modelling is that the efficiency score is 'relative' (Hollingsworth and Peacock 2008; Rao 2013; Garcia-Marquez and Lev, 2015), that is, the efficiency score is determined in comparison with other units. With regard to the current study, this might lead to difficulty in interpreting results, specifically for non-DEA audience such as hospital administrators, hence, better elaboration of results is required. Additionally, another common concern with DEA is the selection of variables and whether to include them as inputs or outputs (Osman *et al.,* 2014), and choosing relevant variables when number of units is small (Ozcan 2008). The current study strictly emphasized on the queuing problem within the public hospital, where appointment systems do not exist, an effort has been made to choose relevant queuing variables which provide a comprehensive analysis, to suit the purpose of the study.

With regard to data availability, due to lack of access to data and time constraints, data has been collected at a single busy public hospital in Pakistan, however, an effort has been made to provide an in-depth evaluation of the queue system within the designated hospital. Although a few observations have been conducted in different specialist OPDs, however, due to difficulty in the data collection process, detailed data was collected at one busy specialist OPD (along with GDMO clinic and pharmacy).

Although the current study has some *shortcomings*, however, it must be emphasized that these limitations did not significantly constrain the modelling and data collection process in order to fulfil the research objectives.

6.4 Future Research

The analytical framework of the current study provides a basis for better queue management in public hospitals of developing countries, where appointment systems do not exist.

The proposed Queuing-DEA model has the potential to be applied in similar public hospitals of other developing countries where appointment systems do not exist. Hence, a research study can be conducted in similar public hospitals in Pakistan as well as other developing countries; to provide a basis for comparison with the regard to evaluation of the queue system in each hospital by employing the proposed model.

One of the main objectives of the current study has been to construct a dynamic framework dedicated towards the practical implementation of the proposed Queuing-DEA model. Therefore, the next step is to strive towards actual implementation of the dynamic framework in the designated hospital by consulting with the administrative team. The issue of excessive queuing within the outpatients' department and the recommendation for improvement in terms of the proposed model and dynamic framework for implementation, has been briefly discussed with the administrators. Some staff members acknowledged the severity of queuing problem and were open to discussion regarding improvement strategies, while some were sceptical about the recommendations. However, an effort will be made to hold detailed negotiations with the administrative team elaborating upon the prevailing queuing problem and the potential of the proposed model and dynamic framework for quick implementation as a continuous monitoring tool for the queuing situation.

Additionally, *later on*, an integrated queue management system can be developed which not only monitors the queue system within each department, but also uploads information from all departments to the main data management system in the Statistics department. The Statistics department can keep a track of the patient flow in each department instantly assisting them to conduct collective analysis of the queue management within the whole hospital.

These *future research* avenues can be explored, based on the proposed model, and results generated by the current study.

Chapter Summary

The current chapter highlights the critical elements of the current research study, mainly by providing a discussion regarding each research objective and emphasizing on the contribution of the current study.

The first section provides a comprehensive discussion with regard to the three research objectives. The first main objective of the current study was to identify the operational importance of queuing factors to assess the queue system for outpatients, in the absence of appointment systems. Very few studies have evaluated walk-in patients, and only with respect to improving existing appointment systems. The current study evaluates the queue system where 'all' patients are walk-in. Due to absence of appointments, the arrival rate of patients is extremely variable, with high overload, resulting in excessive wait times of patients. Furthermore, the preliminary findings indicated that the inadequate scheduling of personnel is a critical issue. Additionally, the consultation time was extremely low, as opposed to almost all previous works. Hence, the current study proposed a monitoring tool for efficient queue management, addressing the severe queuing characteristics.

The second research objective is associated with the development of the proposed Queuing-DEA model along with the utility of the obtained analytical results, to assess the queue system for walk-in patients. Considering a few limitations of other OR techniques, DEA has been employed in order to provide a more detailed evaluation. The current study provides a novel application of DEA for queue management using patient level data. The number of personnel has been included as an output, as opposed to previous works which considered it as an input. The consultation time has been included as a 'non-discretionary' output. Very few studies have considered fixed variables and these have considered fixed inputs, with negligible applications in healthcare. The proposed Queuing-DEA model allows for determining the required number of personnel, using wait times and length of queue as inputs, keeping consultation time as a fixed output. The analytical results highlighted excessive wait times and extreme overload in the busiest departments of the designated hospital, resulting in a high requirement of personnel, and in some cases, to an unfeasible number. Hence the proposed DEA model is exceedingly useful as the queue can be controlled quickly by appropriate staff scheduling given target personnel, resulting in reduced wait times for patients.

The third objective is associated with constructing a dynamic framework for the sake of practical implementation of the proposed model. Considering previous works, very few studies have considered the implementation of results, and most of these have provided suggestions which might not be practical in case of developing countries. The current study has developed a dynamic framework using Excel, to encourage smooth and quick implementation. A few studies have used Excel but with respect to assessing an appointment system. The dynamic framework for the current study has been developed in Excel such that the complex technical functions and programming operates automatically in the background, requiring minimal effort from the users. Using real-time queue data, the dynamic framework will continuously monitor the queue system, updating automatically as new patients arrive, and display the target number of personnel. Hence, the management can respond to quickly to rapidly changing queue situation through dynamic and efficient staff scheduling, minimizing excessive queues.

The next section elaborates upon the contribution of the current study in terms of contribution in the research context, theoretical contribution and practical contribution. With regard to the research context, the absence of appointment systems is a significant issue faced by developing countries in addition to numerous challenges within healthcare delivery systems. Additionally, the arrival pattern of patients is variable with extreme overload, constraining the management to pre-plan effectively, leading to excessive wait times. The current study has specifically evaluated the outpatients' department within a large public hospital of a developing country, where all patients are walk-in, by identifying respective queuing characteristics.

Considering theoretical contribution, the current study extends the application of DEA modelling in the context of a queue problem, particularly by evaluating a queue system without appointment systems, from a developing country perspective; deviating from its more traditional applications in healthcare. Based on the crucial finding of inadequate scheduling, the proposed Queuing-DEA model included number of personnel as an output to determine the exact requirement of personnel at a specific time, using corresponding wait times and length of queue, with consultation time as a 'non-discretionary' output. Hence, the vital information regarding the exact number of personnel will allow the management to take prompt action by enhancing staff availability when actually required, to reduce excessive wait times for patients. Additionally, the proposed model is generic and can be applied to other similar public hospitals in other developing countries.

With regard to practical contribution, the proposed model has been applied to the three busiest departments of the designated hospital, pinpointing extreme wait times and overload, with an increased level of required personnel. Therefore, the proposed model indicates the severity of the queue situation at different intervals and provides guidance in terms of 'exact' target of personnel, to minimize the queue quickly. Furthermore, the current study has designed a dynamic framework specifically for the sake of implementation. The dynamic framework has been developed in Excel (using VBA coding), as it seemed more feasible with respect to implementation in resource-poor public hospitals of developing countries. The dynamic framework has been constructed such that minimal training and work is required for the users, and the whole system will re-run automatically in the background, displaying the required number of personnel. The dynamic framework will ensure constant supervision of the queue situation showing updated requirement of personnel, resulting in dynamic scheduling of personnel and reduced queues. Furthermore, the proposed model and the dynamic framework can be applied in similar public hospitals in other developing countries, where appointment systems do not exist for outpatients.

The following section mentions the limitations of the current study. Considering the modelling technique used, DEA can be sensitive to outlier data, where extreme values were disregarded in the current study when conducting further analysis. The efficiency scores are 'relative' and hence, need to be better elaborated especially to hospital administrators who are not aware of the technical aspects of DEA. Additionally, with regard to selection of inputs and outputs, the current study specifically evaluated the queuing problem, where an effort has been made to include the queuing variables which is suitable for the purpose of the study. Considering data availability, data was collected at a single public hospital due to lack of access to data.

Lastly, some future research avenues have been highlighted including an attempt to apply the proposed in other similar public hospitals of other developing countries where appointment systems do not exist; the actual implementation of the dynamic framework by holding detailed discussion with the administration team, as well as moving towards developing an integrated queue management for the entire hospital.

Concluding Remarks

The current study has been dedicated towards assessing the extremely crucial issue of long wait times of patients in busy public hospitals of developing countries, due to absence of appointment systems. The current study has utilized the efficiency assessment technique of DEA, extending the application of DEA modelling in the context of queue management. The proposed Queuing-DEA model provides vital information regarding the required number of personnel, to allow for optimal staff utilization and reduction in excessive queuing. Additionally, the current study proposed a dynamic framework, specifically for practical implementation of the proposed model, to continuously monitor the queue 'as it is' and indicating required number for personnel; to reduce excessive wait times and for dynamic staff scheduling. The current study has addressed a significant issue by highlighting the suffering of patients in terms of long wait times at a typical overcrowded public hospital of a developing country.

REFERENCES

Aboueljinane, L., Sahin, E. and Jemai, Z. (2013). A Review on Simulation Models Applied to Emergency Medical Service Operations, *Computers & Industrial Engineering*, 66(4), p.734-750.

Adeleke, R. A., Ogunwale, O.D. and Halid, O.Y. (2009). Application of Queuing Theory to Waiting Time of Out-Patients in Hospitals, *The Pacific Journal of Science and Technology*, 10(2), p.270-274.

Afonso, A. and St Aubyn, M. (2006). Relative Efficiency of Health Provision: A DEA Approach with Non-Discretionary Inputs. *ISEG-UTL Economics Working Paper*, 33.

Afrane, S. and Appah, A. (2014). Queuing theory and the management of Waiting-time in Hospitals: The Case of Anglo Gold Ashanti Hospital in Ghana, *International Journal of Academic Research in Business and Social Sciences*, 4(2), p.34-44.

Afsar, H.A. and Younus, M. (2004). Patient Referral at the Grass-roots Level in Pakistan, *Nature and Science*, 2(4), p.18-27.

Aharonson-Daniel, L., Paul, R.J. and Hedley, A.J. (1996). Management of Queues in Out-patient Departments: The Use of Computer Simulation, *Journal of Management in Medicine,* 10(6), p.50-58.

Ahmed, M.A. and Alkhamis, T.M. (2009). Simulation Optimization for an Emergency Department Healthcare Unit in Kuwait, *European Journal of Operational Research,* 198(3), p.936-942.

Akazili, J., Adjuik, M., Appiah, C. J. and Zere, E. (2008). Using Data Envelopment Analysis to Measure the Extent of Technical Efficiency of Public Health Centers in Ghana, *Bio Med Central Ltd.* 20(2), p.232-248.

Aksezer, Ç.S. (2016). A Nonparametric Approach for Optimal Reliability Allocation in Health Services. *International Journal of Quality & Reliability Management*, 33(2), p.284-294.

Allder, S., Silvester, K. and Walley, P. (2010). Managing Capacity and Demand across the Patient Journey. *Clinical medicine*, 10(1), p.13-15.

Al-Shammari, M. (1999).A Multi-criteria Data Envelopment Analysis Model for Measuring the Productive Efficiency of Hospitals, *International Journal of Operations & Production Management*, 19(9), p.879-891.

Amado, C.A.F. and Dyson, R.G. (2008). On Comparing the Performance of Primary Care Providers. *European Journal of Operational Research*, 185(3), p.915-932.

Amzat, J. and Razum, O. (2014). *Medical Sociology in Africa*, Switzerland: Springer.

Anderson, K., Zheng, B., Yoon, S.W. and Khasawneh, M.T. (2015). An Analysis of Overlapping Appointment Scheduling Model in an Outpatient Clinic, *Operations Research for Health Care*, 4, p.5-14.

Angelis, V.D., Felici, G. and Impelluso, P. (2003). Integrating Simulation and Optimization in Healthcare Centre Management, *European Journal of Operational Research*, 150(1), p.101-114.

Applanaidu, S.D., Samsudin, S., Ali, J., Dash, U. and Chik, A.R. (2014). Technical and Scale Efficiency of Public District Hospitals in Kedah, Malaysia: A Data Envelopment Analysis (DEA), *Journal of Health Management*, 16(3), p.327-335.

Ashton, R., Hague, L., Brandreth, M., Worthington, D. and Cropper, S. (2005). A Simulation-based Study of a NHS Walk-in Centre, *Journal of Operational Research Society*, 56(2), p.153-161.

Aston Business School. *Research Ethics: Issues to Consider and Address before Submitting an Application to the Research Ethics Committee*. Available at: <http://www1.aston.ac.uk/aston-business-school/research/ethics/issues-to-consider/#Points> [Accessed: 19 July 2013].

Au, L., Byrnes, G.B., Bain, C.A., Fackrell, M., Brand, C., Campbell, D.A. and Taylor, P.G. (2009). Predicting Overflow in an Emergency Department, *IMA Journal of Management Mathematics*, 20(1), p.39-49.

Azadeh, A., Keramati, A. and Songhori, M.J. (2009). An Integrated Delphi/VAHP/DEA Framework for Evaluation of Information Technology/Information System (IT/IS) Investments, *The International Journal of Advanced Manufacturing Technology*, 45(11-12), p.1233-1251.

Babes, M. and Sarma, G.V. (1991). Outpatient Queues at the Ibn-Rochd Health Centre, *Journal of the Operational Research Society*, 42(10), p.845-855.

Bailey, N. (1952). A Study of Queues and Appointment Systems in Hospital Outpatient Departments with Special Reference to Waiting Times, *Journal of the Royal Statistical Society,* 14, p.185–199.

Banker, R.D. and Morey, R.C. (1986). Efficiency Analysis for Exogenously Fixed Inputs and Outputs, *Operations Research*, 34(4), p.513-521.

Banker, R.D., Charnes, A. and Cooper, W.W. (1984). Some Models for Estimating Technical and Scale Efficiencies in Data Envelopment Analysis, *Management Science*, 30(9), p. 1078-1092.

Bard, J.F., Shu, Z., Morrice, D.J., Poursani, R. and Leykum, L. (2014). Improving Patient Flow at a Family Health Clinic, *Health care Management Science*, p.1-22.

Baril, C., Gascon, V., Miller, J. and Côté, N. (2016). Use of a Discrete-Event Simulation in a Kaizen Event: A Case Study in Healthcare, *European Journal of Operational Research*, 249(1), p.327-339.

Benneyan, J. C. (1997). An Introduction to Using Computer Simulation in Healthcare: Patient Wait Case Study, *Journal of the Society for Health System,* 5(3), p.1–15.

Benneyan, J.C., Horowitz, M.I. and Terceiro, M. (1994). Using Computer Simulation to Help Reduce Patient Waits, In *Proceedings of the 1994 Annual HIMSS Conference*,. 1, p.323-342.

Berg, B.P., Denton, B.T., Erdogan, S.A., Rohleder, T. and Huschka, T. (2014). Optimal Booking and Scheduling in Outpatient Procedure Centers. *Computers & Operations Research*, 50, p.24-37.

Bhattacharjee, P. and Ray, P.K. (2014). Patient Flow Modelling and Performance Analysis of Healthcare Delivery Processes in Hospitals: A Review and Reflections, *Computers and Industrial Engineering*, 78, p.299-312.

Bielen, F. and Demoulin, N. (2007). Waiting Time Influence on the Satisfaction-Loyalty Relationship in Services, *Managing Service Quality*, 17(2), p.174–193.

Biju, M.K. and Naeema, K. (2011). Application of Queuing Theory in Human Resource Management in Healthcare, In: *ICOQM-10*, p.1019-1027.

Bird, C. (2012). *Battling through the Malaria Season in the Democratic Republic of the Congo*, photograph, Available at: <http://www.theguardian.com/global-development/poverty-matters/2012/jan/18/battling-malaria-season-congo-shortages> [Accessed: 01 May 2015].

Blank, J.L.T. and Van Hulst, B.L. (2011). Governance and Performance: The Performance of Dutch Hospitals Explained by Governance Characteristics, *Journal of Medical Systems*, 35(5), p. 991-999.

Bogetoft, P. and Otto, L. (2011). *Benchmarking with DEA, SFA and R*. New York: Springer.

Brahimi, M. and Worthington, D.J. (1991). Queuing Models for Outpatient Appointment Systems: A Case Study, *Journal of Operational Research Society*, 42(9), p.733-746.

Brailsford, S.C., Lattimer, V.A., Tarnaras, P. and Turnbull J.C. (2004). Emergency and On-demand Healthcare: Modeling a Large Complex System, *Journal of Operational Research Society*, 55, p.34-42.

Brown, R., (2006). Mismanagement or Mismeasurement? Pitfalls and Protocols for DEA Studies in the Financial Services Sector, *European Journal of Operational Research*, 174(2), p.1100-1116.

Bruin, A.M., Bekker, R. Van Zanten, L. and Koole, G.M. (2010). Dimensioning Hospital Wards Using Erlang Loss Model, *Annals of Operations Research*, 178(1), p.23-43.

Bruni, R. and Detti, P. (2014). A Flexible Discrete Optimization Approach to the Physician Scheduling Problem, *Operations Research for Health Care*, 3(4), p.191-199.

Bryman, A. and Bell, E., (2015*). Business Research Methods*, 4th ed, UK: Oxford University Press.

Bryman, A. (2006). Integrating Quantitative and Qualitative Research: How Is It Done? *Qualitative Research,* 6(1), p. 97-113.

Butler, T.W. and Li, L. (2005). The Utility of Returns to Scale in DEA Programming: An Analysis of Michigan Rural Hospitals, *European Journal of Operational Research*, 161(2), p.469-477.

Butt, S.E., Fredericks, T.K., Choi, S.D. and Kumar, A.R. (2002). Comparison of Commercial Stair-chairs Using Data Envelopment Analysis, In *Proceedings of the XVI Annual International Occupational Ergonomics and Safety Conference,* p.10-13.

Bwana, K.M. (2015). Measuring Technical Efficiency of Faith Based Hospitals in Tanzania: An Application of Data Envelopment Analysis (DEA), *Research in Applied Economics*, 7(1), p.1-12.

Cassum, L.A. and Shah, M. (2014). Comparative Assessment of Health Care Delivery Systems of Developing Countries: Pakistan versus Cuba, *International Journal of Science and Research,* 3(8), p.984-990.

Cayirli, T. and Gunes, E.D. (2014). Outpatient Appointment Scheduling in Presence of Seasonal Walk-ins, *Journal of the Operational Research Society*, 65, p. 512-531.

Cayirli, T. and Veral, E. (2003). Outpatient Scheduling in Healthcare: A Review of Literature, *Production and Operations Management*, 12(4), p.519-549.

Cayirli, T., Veral, E. and Rosen, H. (2008). Assessment of Patient Classification in Appointment System Design, *Production and Operations Management*, 17(3), p.338-353.

Chandra, D. (2015). Reducing Waiting Time of Outdoor Patients in Hospitals Using Different Types of Models: A Systematic Survey, *International Journal*, 3(1), p.81-87.

Chang, H., Cheng, M.A. and Das, S. (2004). Hospital Ownership and Operating Efficiency: Evidence from Taiwan, *European Journal of Operational Research*, 159(2), p.513-527.

Chao, X., Liu, L. and Zheng, S. (2003). Resource Allocation in Multi-site Service Systems with Inter-site Customer Flows, *Management Science*, 49(12), p.1739-1752.

Charnes, A. and Cooper, W.W. (1962). Programming with Linear Fractional Functionals, *Naval Research Logistics Quarterly*, 9 (3-4), p.181-185.

Charnes, A., Cooper, W.W. and Rhodes, E. (1978). Measuring the Efficiency of Decision Making Units, *European Journal of Operational Research*, 2(6), p.429-444.

Charnes, A., Cooper, W.W., Lewin, A. and Seiford, L. (eds.) (1994). *Data Envelopment Analysis: Theory, Methodology and Applications*, Boston, MA: Kluwer Academic Publishers.

Cheema, D.S. (2005). *Operations Research.* India: Laxmi Publications Ltd.

Chen, R.R. and Robinson, L.W. (2014). Sequencing and Scheduling Appointments with Potential Call-In Patients. *Production and Operations Management*, 23(9), p.1522-1538.

Chilingerian, J.A. (1995). Evaluating Physician Efficiency in Hospitals: A Multivariate Analysis of Best Practices, *European Journal of Operational Research*, 80(3), p.548-574.

Chilingerian, J.A. and Sherman, H.D. (1997). DEA and Primary care Physician Report Cards: Deriving Preferred Practice Cones from Managed care Service Concepts and Operating Strategies, *Annals of Operations Research*, 73, p.35-66.

Chuang, C.L., Chang, P.C. and Lin, R.H. (2011). An Efficiency Data Envelopment Analysis Model Reinforced by Classification and Regression Tree for Hospital Performance Evaluation, *Journal of Medical Systems*, 35(5), p.1075-1083.

Cleverly, W.O. and Cameron, A.E, (2007), Essentials of Health care Finance, Canada: Jones and Bartlett Publishers.

Cochran, J.K. and Roche, K. (2008). A Queuing-Based Decision Support Methodology to Estimate Hospital Inpatient Bed Demand, *Journal of the Operational Research Society*, 59(11), p.1471-1482.

Cook, W.D., Tone, K. and Zhu, J. (2014). Data Envelopment Analysis: Prior to Choosing a Model. *Omega*, 44, p.1-4.

Cooper, W.W., Seiford, L.M. and Tone, K. (2007). *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, 2nd ed, New York: Springer.

Cooper, W.W., Seiford, L.M., and Zhu, J. (2011). *Data Envelopment Analysis: History, Models and Interpretations,* In Handbook on Data Envelopment Analysis, 2nd ed, USA: Springer.

Cordero-Ferrera, J.M., Pedraja-Chaparro, F. and Santín-González, D. (2010). Enhancing the Inclusion of Non-Discretionary Inputs in DEA. *Journal of the Operational Research Society*, 61(4), p.574-584.

Cote, M.J. (1999). Patient Flow and Resource Utilization in an Outpatient Clinic, *Socio-Economic Planning Sciences*, 33(3), p.231-245.

Country Comparison to the World. Available at: <https://www.cia.gov/library/publications/the-world-factbook/fields/print_2103.html> [Accessed: 28 April 2015].

Creswell, J.W. and Plano-Clark, V.L. (2011). *Choosing a Mixed Methods Design*, In Designing and Conducting Mixed Methods Research. 2nd ed, USA: Sage.

Creswell, J.W. (2014). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. 4th ed, UK: Sage.

Dansky, K.H. and Miles, J. (1997). Patient Satisfaction with Ambulatory Healthcare Services: Waiting Time and Filling Time, *Hospital and Health Services Administration*, ProQuest Business Collection, 42(2), p.165-177.

Daultani, Y., Kumar, S. and Vaidya, O.S. (2016). Improving Out-Patient Flow at an Indian Ophthalmic Hospital. *Operations and Supply Chain Management*, 9(1), p.15-21.

De Vaus, D. (2001). *Research Design in Social Research*. UK: Sage.

Demir, E., Lebcir, R. and Adeyemi, S. (2013). Modelling Length of Stay and Patient Flows: Methodological Case Studies from the UK Neonatal Care Services. *Journal of the Operational Research Society*, 65(4), p.532-545.

Dhungel, B. (2014). *Challenges of Providing Primary Health care*. GRIN Verlag.

Dotoli, M., Epicoco, N., Falagario, M. and Sciancalepore, F. (2015). A Cross-Efficiency Fuzzy Data Envelopment Analysis Technique for Performance Evaluation of Decision Making Units Under Uncertainty, *Computers & Industrial Engineering*, 79, p.103-114.

Driouchi, A., (2014). *Labour and Health Economics in the Mediterranean Region: Migration and Mobility of Medical Doctors*. USA: IGI Global.

Dutta, A., Bandyopadhyay, S. and Ghose, A. (2014). Measurement and Determinants of Public Hospital Efficiency in West Bengal, India, *Journal of Asian Public Policy*, 7(3), p.231-244.

Easterby-Smith, M, Thorpe, R. and Jackson, P. (2012). *Management Research*. 4th ed, UK: Sage.

Emrouznejad, A. and De Witte, K. (2010). COOPER-framework: A Unified Process for Non-Parametric Projects, *European Journal of Operational Research*, 207(3), p.1573-1586.

Emrouznejad, A. and Cabanda, E. (eds.) (2014). *Managing Service Productivity: Uses of Frontier Efficiency Methodologies and MCDM for Improving Service Performance*, Berlin Heidelberg: Springer-Verlag.

Emrouznejad, A., Parker, B. and Tavares, G. (2008). Evaluation of Research in Efficiency and Productivity: A Survey and Analysis of the First 30 Years of Scholarly Literature in DEA, *Journal of Socio-Economics Planning Science*, 42(3), p.151-157.

Emrouznejad. A. and Thanassoulis. E (2014). *Performance Improvement Management (PIM-DEA) Software (PIMsoft).* Available at: <http://www.deasoftware.co.uk/> [Accessed: January 2013].

Ersoy, K., Kavuncubasi, S. Ozcan, Y.A., and Harris II, J.M. (1997). Technical Efficiencies of Turkish Hospitals: DEA Approach, *Journal of Medical Systems*, 21(2), p.67-74.

European Commission: Ethical Review in FP7. (2009). *Data Protection and Privacy Ethical Guidelines*. Available at: <ftp://ftp.cordis.europa.eu/pub/fp7/docs/privacy.doc> [Accessed 19 July 2013].

Farrell, M. (1957). The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society, Series A* (Part 3): p. 253-281.

Feldman, J., Liu, N., Topaloglu, H., and Ziya, S. (2014). Appointment Scheduling under Patient Preference and No-show Behavior, *Operations Research*, 62(4), p.794-811.

Fetter, R.B. and Thompson, J.D. (1966). Patients' Waiting Time and Doctors' Idle Time in the Outpatient Setting, *Health Services Research*, 1(1), p. 6-90.

Fleßa, S. (2009). *Costing of Health Care Services in Developing Countries: A Prerequisite for Affordability, Sustainability and Efficiency*. Frankfurt am Main: Lang.

Flokou, A., Kontodimopoulos, N. and Niakas, D. (2011). Employing Post-DEA Cross-Evaluation and Cluster Analysis in a Sample of Greek NHS Hospitals, *Journal of Medical Systems,* 35(5), p.1001-1014.

Fomundam, S. and Hermann, J.W. (2007), A Survey of Queuing Theory Applications in Healthcare, *Digital Repository at the University of Maryland*. Available at: <http://hdl.handle.net/1903/7222> [Accessed: 19 July 2013].

Gaimard, M., (2014). *Population and Health in Developing Countries (Demographic Transformation and Socio-Economic Development 2*, Dordrecht: Springer.

Gattoufi, S., Oral, M., Kumar, A. and Reisman, A. (2004). Content Analysis of Data Envelopment Analysis Literature and Its Comparison with that of Other OR/MS Fields, *Journal of the Operational Research Society*, 55(9), p.911-935.

GDP per capita (current US$), *Data: World bank.* Available at: <http://data.worldbank.org/indicator/NY.GDP.PCAP.CD> [Accessed: 28 April 2015].

George, J.A., Fox, D.R. and Canvin, R.W. (1983). A Hospital Throughput Model in the Context of Long Waiting Lists, *Journal of Operational Research Society*, 34, Case-Oriented Paper, p. 27-35.

Gerard, K. and Roderick, P. (2003). Comparison of Apparent Efficiency of Haemodialysis Satellite Units in England and Wales Using Data Envelopment Analysis, *International journal of technology assessment in health care*, 19(3), p.533-539.

Geyndt, W.D. (1995). Managing the Quality of Health care in Developing Countries, USA: World Bank.

Gholami, R., Higón, D.A. and Emrouznejad, A. (2015). Hospital Performance: Efficiency or Quality? Can We Have Both With It? *Dynamic frameworks with Applications*, 42(12), p.5390-5400.

Giachetti, R.E., Centeno, E.A., Centeno, M.A. and Sundaram, R. (2005). Assessing the Viability of an Open Access Policy in an Outpatient Clinic: A Discrete-event and Continuous Simulation Modelling Approach, In: *Proceedings of the 2005 Winter Simulation Conference.* M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines (eds).

Global Health Observatory (GHO) data, *World Health Organisation*, Available at: <http://www.who.int/gho/countries/en/> [Accessed: 15 January 2016].

Goddard, J. and Tavakoli, M. (2008). Efficiency and Welfare Implications of Managed Public Sector Hospital Waiting Lists, *European journal of operational research*, 184(2), p.778-792.

Golany, B. and Roll, Y. (1993). Some Extensions of Techniques to Handle Non-Discretionary Factors in Data Envelopment Analysis. *Journal of productivity Analysis*, 4(4), p.419-432.

Gomes, E.G. and Lins, M.P.E. (2008), Modelling Undesirable Outputs with Zero Sum Gains Data Envelopment Analysis Models, *Journal of Operational Research Society*, 59, p.616-623.

Gorunescu, F., McClean, S.I. and Millard, P.H. (2002) A Queuing Model for Bed-Occupancy Management and Planning of Hospitals. *Journal of the Operational Research Society*, 53(1), p.19-24.

Griffiths, J.D., Price-Lloyd, N., Smithies, M. and Williams, J.E. (2005). Modelling the Requirement for Supplementary Nurses in an Intensive Care Unit, *Journal of Operational Research Society*, 56(2), p.126-133.

Gul, M. and Guneri, A.F. (2012). A Computer Simulation Model to Reduce Patient Length Of Stay and to Improve Resource Utilization Rate in an Emergency Department Service System, *International Journal of Industrial Engineering*, 19(5), p.221-231.

Habib, N. (2012). *100 Dead from Suspected Tainted Heart Medicine in Pakistan*, photograph, Available at: <http://edition.cnn.com/2012/01/26/world/asia/pakistan-tainted-medicine/> [Accessed: 01 May 2015].

Hadad, S., Hadad, Y. and Simon-Tuval, T. (2013). Determinants of Healthcare System's Efficiency in OECD Countries, *The European Journal of Health Economics*, 14(2), p.253-265.

Harper, P.R. and Gamlin, H.M. (2003). Reduced Outpatient Waiting Times with Improved Appointment Scheduling: A Simulation Modelling Approach, *OR Spectrum,* 2(2), p.207-222.

Harper, P.R. and Shahani, A.K. (2002). Modelling for the Planning and Management of Bed Capacities in Hospitals, *Journal of Operational Research Society,* 53(1), p.11-18.

Harrison, J., Spaulding, A. and Mouhalis, P. (2015). The Efficiency Frontier of For-Profit Hospitals. *Journal of Health Care Finance*, 41(4), p. 1-15.

Harrison, J.P., Coppola, M.N. and Wakefield, M. (2004). Efficiency of Federal Hospitals in the United States, *Journal of Medical Systems*, 28(5), p.411-422.

Hashmi, F., Chaudhry, T., Ahmad, K. (2010). An Evaluation of Referral System for Retinopathy of Prematurity in Leading Health Centers across Karachi, Pakistan, *Journal of the Pakistan Medical Association*, 60(10), p.840-844.

Hassin, R. and Mendel, S. (2008). Scheduling Arrivals to Queues: A Single-Server Model with No-shows, *Management Science,* 54(3), p.565-572.

Hassin, R. and Mendel, S. (2008). Scheduling Arrivals to Queues: A Single-Server Model with No-Shows, *Management Science*, 54(3), p.565-572.

Herwartz, H. and Strumann, C. (2014). Hospital Efficiency under Prospective Reimbursement Schemes: An Empirical Assessment for the Case of Germany. *The European Journal of Health Economics*, 15(2), p.175-186.

Hill-Smith, I. (1989). Mathematical Relationship between Waiting Times and Appointment Interval for Doctors and Patients, *Journal of the Royal College of General Practitioners*, 39(329), p. 492-494.

Hollingsworth, B. and Peacock, S.J. (2008). *Efficiency Measurement in Health and Healthcare*, Oxon: Routledge.

Hollingsworth, B. (2003). Non-Parametric and Parametric Applications Measuring Efficiency in Healthcare, *Healthcare Management Science*, 6(4), p.203-218.

Hollingsworth, B. (2008). The Measurement of Efficiency and Productivity of Healthcare Delivery, *Health Economics*, 17(10), p.1107-1128.

Hollingsworth, B. (2012). Revolution, Evolution or Status-quo? Guidelines for Efficiency Measurement in Healthcare, *Journal of Productivity Analysis*, 37(1), p.1-5.

Hollingsworth, B. and Parkin, D. (1995). The Efficiency of Scottish Acute Hospitals: An Application of Data Envelopment Analysis, *IMA Journal of Mathematics Applied in Medicine and Biology*, 12(3-4), p.161-173.

Hollingsworth, B., Dawson, P.J. and Maniadakis, N. (1999). Efficiency Measurement of Healthcare: A Review of Non-Parametric Methods and Applications, *Healthcare Management Science*, 2(3), p.161-172.

Huang, Y.G. and McLaughlin, C.P. (1989). Relative Efficiency in Rural Primary Health care: An Application of Data Envelopment Analysis, *Health Services Research*, 24(2), p.143-158.

Huarng, F. (2003). The Impact of Walk-in Ratio on Outpatient Department, *2003 International Conference on Technology & Management with Special Symposium on Advanced Interactive Systems.*

Huarng, F. and Lee, M.H. (1996). Using Simulation in Outpatient Queues: A Case Study, *International Journal of Healthcare Quality Assurance*, 9(6), p.21-25.

Ishizaka, A., and Nemery, P. (2013). *Multi-criteria Decision Analysis: Methods and Software*. UK: John Wiley and Sons.

Isken, M.W., Ward, T.J. and McKee, T.C. (1999). December. Simulating Outpatient Obstetrical Clinics. In *Proceedings of the 31st Conference on Winter Simulation: Simulation---a Bridge to the Future,* 2, p. 1557-1563.

Jacobs, R., Smith, P.C. and Street, A. (2006). *Measuring Efficiency in Healthcare: Analytic Techniques and Health Policy*, UK: Cambridge University Press.

Jafari, Y., Madahi, R. and Khodayifar, S. (2010). Evaluating the Performance of Multiple Comparable Queuing by Using DEA, *2nd National Conference on Data Envelopment Analysis*.

Jalal, S. and Inam-ul-Haq, (2014). Revisiting the Three Different Tiers of the Health System of Pakistan and Their Implications for the Achievement of MDGs by Pakistan, *Journal of Pakistan Medical Association*, 64(2), p.195-200.

Jamison, D.T., Breman, J.G., Measham, A.R., Alleyne, G., Claeson, M., Evans, D.B., Jha, P., Mills, A. and Musgrove, P. (2006). *Disease Control Priorities in Developing Countries*, 2nd ed, New York: Oxford University Press.

Jehu-Appiah, C., Sekidde, S., Adjuik, M., Akazili, J., Almeida, S.D., Nyonator, F., Baltussen, R., Asbu, E.Z. and Kirigia, J.M. (2015). Ownership and Technical Efficiency of Hospitals: Evidence from Ghana Using Data Envelopment Analysis, *Cost Effectiveness and Resource Allocation*, 12(9), p.1-13.

Johnson, R.B. and Onwuegbuzie, A.J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come, *Educational Researcher: American Educational Research Association*, 33(7), p.14-26.

Jun, J.B., Jacobson, S.H. and Swisher, J.R. (1999). Application of Discrete-event Simulation in Healthcare Clinics: A Survey, *Journal of the Operational Research Society*, 50(2), p.109-123.

Kang, H., Nembhard, H.B. and DeFlitch, C. (2014). Identifying Emergency Department Efficiency Frontiers and the Factors Associated with Their Efficiency Performance, *Proceedings of the 2014 Industrial and Systems Engineering Research Conference*, Guan, Y and Liao, H. (eds).

Kao, E.P. and Tung, G.G. (1981). Bed Allocation in a Public Health Care Delivery System. *Management Science*, 27(5), p.507-520.

Kawaguchi, H, Tone, K. and Tsutsui, M. (2014). Estimation of the Efficiency of Japanese Hospitals Using a Dynamic and Network Data Envelopment Analysis Model, *Healthcare Management Science*, 17, p.101-112.

Keshtkaran, A., Barouni, M., Ravangard, R. and Yandrani, M. (2014). Economic Efficiency of Radiology Wards Using Data Envelopment Analysis: Case Study of Iran, *Health*, 6(5), p.311.

Kilic, B., Kalaca, S., Unal, B., Phillimore, P. and Zaman, S. (2015). Health Policy Analysis for Prevention and Control of Cardiovascular Diseases and Diabetes Mellitus in Turkey, *International Journal of Public Health*, 60(1), p.47-53.

Kirigia, J.M., Emrouznejad, A., Cassoma, B., Asbu, E.Z. and Barry, S. (2008). A Performance Assessment Method for Hospitals: The Case of Municipal Hospitals in Angola, *Journal of Medical Systems,* 32 (6), p.509-519.

Kirigia, J.M., Emrouznejad, A., Sambo, L.G., Munguti, N. and Liambila, W. (2004). Using Data Envelopment Analysis to Measure the Technical Efficiency of Public Health Centers in Kenya, *Journal of Medical Systems*, 28(2), p.155-166.

Kirigia, J.M., Emrouznejad, A., Vaz, R.G., Bastiene, H. and Padayachy, J. (2007). A Comparative Assessment of Performance and Productivity of Health Centres in Seychelles, *International Journal of Productivity and Performance Management*, 57(1), p.72–92.

Kittelsen, S.A., Winsnes, B.A., Anthun, K.S., Goude, F., Hope, Ø., Häkkinen, U., Kalseth, B., Kilsmark, J., Medin, E., Rehnberg, C. and Rättö, H. (2015). Decomposing the Productivity Differences between Hospitals in the Nordic Countries, *Journal of Productivity Analysis*, 43(3), p.281-293.

Klassen, K.J. and Rohleder, T.R. (1996). Scheduling Outpatient Appointments in a Dynamic Environment, *Journal of Operations Management*, 14(2), p.83-101.

Klassen, K.J. and Yoogalingham, R. (2009). Improving Performance in Outpatient Appointment Services with a Simulation Optimization Approach, *Production and Operations Management,* 18(4), p.447-458.

Kortbeek, N., Braaksma, A., Burger, C.A.J., Bakker, P.J.M. and Boucherie, R.J. (2015). Flexible Nurse Staffing Based on Hourly Bed Census Predictions, *International Journal of Production Economics*, 161, p.167-180.

Kose, T., Uckun, N. and Girginer, N. (2014). An Efficiency Analysis of the Clinical Departments of a Public Hospital in Eskisehir by Using DEA, *Global Journal on Advances in Pure and Applied Sciences*, 4, p.252-258.

Kozlowski, D. and Worthington, D. (2015). Use of Queue Modelling in the Analysis of Elective Patient Treatment Governed by a Maximum Waiting Time Policy, *European Journal of Operational Research*, 244(1), p.331-338.

Kyriacou, A., Muinelo-Gallo, L. and Roca-Sagalés, O. (2015). *On The Redistributive Efficiency of Fiscal Policy,* (63276), University Library of Munich, Germany.

Lakshmi, C and Sivakumar, A.I. (2013). Application of Queuing Theory in Healthcare: A Literature Review, *Operations Research for Healthcare*, 2(1-2), p.25-39.

Lane, D.C., Monefeldt, C. and Rosenhead, J.V. (2000). Looking in the Wrong Place for Healthcare Improvements: A System Dynamics Study of an Accident and Emergency Department, *Journal of Operational Research Society*, 51(5), p.518-531.

Laskowski, M., McLeod, R.D., Friesen, M.R., Podaima, B.W. and Alfa, A.S. (2009). Models Emergency Departments for Reducing Patient Waiting Times, *PLOS One*, 4(7), p.1-12.

Lee, H. and Kim, C. (2014). Benchmarking of Service Quality with Data Envelopment Analysis, *Dynamic Frameworks with Applications*, 41(8), p.3761-3768.

Lee, N. and Lings, I. (2008). *Doing Business Research: A Guide to Theory and Practice*. UK: Sage.

Lehaney, B., Clarke, S.A. and Paul, R.J. (1999). A Case of an Intervention in an Outpatients Department, *Journal of Operational Research Society*, 50(9), p. 877-891.

Lewis, H.F., Sexton, T.R. and Dolan, M.A. (2011). An Efficiency-Based Multicriteria Strategic Planning Model for Ambulatory Surgery Centers, *Journal of Medical Systems*, 35(5), p. 1029-1037.

Li, X., Beullens, P, Jones, D. and Tamiz, M. (2009). An Integrated Queuing and Multi-objective Bed Allocation Model with Application to a Hospital in China, *Journal of Operational Research Society*, 60, p.330-338.

Lindlbauer, I., Schreyögg, J. and Winter, V. (2016). Changes in Technical Efficiency after Quality Management Certification: A DEA Approach Using Difference-In-Difference Estimation with Genetic Matching in the Hospital Industry. *European Journal of Operational Research*, 250(3), p.1026-1036.

Linna, M. and Häkkinen, U. (1999). *Determinants of Cost efficiency of Finnish Hospitals: A Comparison of DEA and SFA*. Helsinki University of Technology, Department of Engineering Physics and Mathematics, Systems Analysis Laboratory.

Liu, J.S., Lu, L.Y., Lu, W.M. and Lin, B.J. (2013b). A survey of DEA Applications. *Omega*, 41(5), p.893-902.

Liu, J.S., Lu, L.Y.Y., Lu,W. and Lin, B.J.Y. (2013a). Data Envelopment Analysis 1978-2010: A Citation-based Literature Survey, *Omega*, 41(1), p.3-15.

Liu, L., and Liu, X., (1998). Block Appointment Systems for Outpatient Clinics with Multiple Doctors, *Journal of Operational Research Society*, 49(12), p.1254-1259.

Mandelbaum, A., Momcilovic, P. and Tseytlin, Y. (2012). On Fair Routing from Emergency Departments to Hospital Wards: QED Queues with Heterogeneous Servers, *Management Science*, 58(7), p.1273-1291.

Mankowska, D.S., Meisel, F. and Bierwirth, C. (2014). The Home Healthcare Routing and Scheduling Problem with Interdependent Services, *Healthcare Management Science*, 17, p.15-30.

Manzi, A, Magge H, Hedt-Gauthier, B.L., Michaelis, A.P., Cyamatare, F.R., Nyirazinyoye, Hirschhorn, L.R. and Ntaganira, J. (2014). Clinical Mentorship to Improve Pediatric Quality of Care at the Health Centers of Rural Rwanda: A

Qualitative Study of Perceptions and Acceptability of Health Workers, *BMC Health Services Research*, 14(275), p.1-9.

Map of Pakistan, (2010). *Pakistan in Pictures*, photograph, Available at: <https://thepeopleofpakistan.wordpress.com/2010/02/17/pakistan-in-pictures/> [Accessed 24.04.2015].

Márquez, F.P.G. and Lev, B. eds., (2015). *Advanced Business Analytics*. Switzerland: Springer.

Martin, E., and Haugene, K. (2003). Proposals to Reduce Over-crowding, Lengthy Stays and Improve Patient Care: Study of the Geriatric Department in Norway's Largest Hospital, In: *Proceedings of the 2003 Winter Simulation Conference*, Chick, S. Sánchez, P.J., Ferrin, D and Morrice, D.J. (eds).

Matta, M.E. and Patterson, S.S. (2007). Evaluating Multiple Performance Measures across Several Dimensions at a Multi-facility Outpatient Center, *Healthcare Management Science,* 10(2), p.173-194.

May, J.H., Spangler, W.E. and Strum, D.P. (2011). The Surgical Scheduling Problem: Current Research and Future Opportunities, *Production and Operations Management*, 20(3), p.392-405.

Mayhew, L. and Smith, D. (2008). Using Queuing Theory to Analyse the Government's 4-h Completion Time Target in Accident and Emergency Departments, *Healthcare Management Science*, 11(1), p.11-21.

Maynard, J (2014). *Poor Sanitation and Hygiene Killing Pregnant Women and New-borns in Developing Countries: WHO*, photograph, Available at: <http://www.techtimes.com/articles/22149/20141214/poor-sanitation-hygiene-killing-pregnant-women-newborns-nations-who.htm> [Accessed on: 24.04.2015].

McCarthy, R.L., Schafermeyer, K.W. and Plake, K.S., (2012). *Introduction to Healthcare Delivery: A Primer for Pharmacists*. 5th ed, USA: Jones and Bartlett Learning.

McDermott, C. and Stock, G.N (2007). Hospital Operations and Length of Stay Performance, *International Journal of Operations & Production Management*, 27(9), p.1020-1042.

McGuire, F. (1994). Using Simulation to Reduce Length of Stay in Emergency Departments, In: *Proceedings of the 1994 Winter Simulation* Conference, Tew, J.D., Manivannan, S., Sadowski, D.A. and Seija, A.F. (eds).

Meghani, S.T., Sehar,S. and Punjani, N.S. (2014). Comparison and Analysis of Healthcare Delivery System: Pakistan versus China, *International Journal of Endorsing Health Science Research,* 2(1), p.46-50.

Mehandiritta, R. (2011). Applications of Queuing Theory in Healthcare, *International Journal of Computing and Business Research*, 2(2).

Mensah, J., Asamoah, D. and Tawiah, A.A. (2014). Optimizing Patient Flow and Resource Utilization in Out Patient Clinic: A Comparative Study of Nkawie Government Hospital and Aniwaa Health Center. *The Journal of Applied Business and Economics*, 16(3), p.181-188.

Microsoft Excel (2013), Microsoft Corporation.

Mital, K.M. (2010). Queuing Analysis for Outpatient and Inpatient services: A Case Study, *Management Decision,* 48(3), p.419-439.

Mitropoulos, P., Mitropoulos, I. and Giannikos, I. (2013). Combining DEA with Location Analysis for the Effective Consolidation of Services in the Health Sector. *Computers & Operations Research*, 40(9), p.2241-2250.

Mitrović, Z., Vujošević, M. and Savić, G. (2016). Data Envelopment Analysis for Evaluating Serbia's Health Care System. *Management*, 1820-0222(75), p.39-46.

Muñiz, M., Paradi, J., Ruggiero, J. and Yang, Z. (2006). Evaluating Alternative DEA Models Used to Control for Non-Discretionary Inputs, *Computers & Operations Research*, 33(5), p.1173-1183.

Muñiz, M.A. (2002). Separating Managerial Inefficiency and External Conditions in Data Envelopment Analysis, *European Journal of Operational Research*, 143, p.625–643.

Naqvi, H.A., Hussain, S., Zaman, M. and Islam, M. (2009). Pathways to Care: Duration of Untreated Psychosis from Karachi, Pakistan. *PLOS ONE,* 4(10), p.1-5. Available at:

Nishtar, S., Boerma, T., Amjad, S., Alam, A.Y., Khalid, F., Ul-Haq, I. and Mirza, Y.A. (2013). Pakistan's Health System: Performance and Prospects After the 18th Constitutional Amendment. Series: Health Transitions in Pakistan 1, *The Lancet*, 381(9884), p.2193-2206.

Nunamaker, T.R. (1983). Measuring Routine Nursing Service Efficiency: A Comparison of Cost per Day and Data Envelopment Analysis Models, *Health Services Research*, 18(2), p.83-208.

O'Keefe, R.M. (1998). Investigating Outpatient Departments: Implementable Policies and Qualitative Approaches, *Journal of Operational Research Society*, 36(8), p.705-712.

O'Neill, L., Rauner, M., Heidenberger, K. and Kraus, M. (2008). A Cross-National Comparison and Taxonomy of DEA-based Hospital Efficiency Studies, *Socio-Economic Planning Sciences*, 42(3), p.158-189.

Oche, M.O. and Adamu, H. (2014). Determinants of Patient Waiting Time in the General Outpatient Department of a Tertiary Health Institution in North Western Nigeria, *Annals of Medical and Health Sciences Research*, 3(4), p.588-592.

OECD (2010), *HealthCare Systems: Efficiency and Policy Settings*, OECD Publishing, Available at: <http://dx.doi.org/10.1787/9789264094901-en> [Accessed: 28 April 2015].

Osman, I.H., Anouze, A.L. and Emrouznejad, A. (2014), *Handbook of Research on Strategic Performance Management and Measurement Using Data Envelopment Analysis*. USA: IGI Global.

Osman, I.H., Berbary, L.N., Sidani, Y., Al-Ayoubi, B. and Emrouznejad, A. (2011). Data Envelopment Analysis Model for the Appraisal and Relative Performance Evaluation of Nurses at an Intensive Care Unit, *Journal of Medical System,* 35(5), p.1039-1062.

Ouellette, P. and Vierstraete, V. (2004). Technological Change and Efficiency in the Presence of Quasi-Fixed Inputs: A DEA Application to the Hospital Sector, *European Journal of Operational Research*, 154(3), p. 755-763.

Ozcan, Y.A. (2008). *Performance Measurement Using Data Envelopment Analysis (DEA)*, Healthcare Benchmarking and Performance Evaluation: An Assessment Using Data Envelopment Analysis (DEA), USA: Springer, p. 15-41.

Pakistan: Country Profile/Statistics. *World Health Organization*. Available at: <http://www.who.int/countries/pak/en/> [Accessed: 23 July 2013].

Parkin, D. and Hollingsworth, B. (1997). Measuring Production Efficiency of Acute Hospitals in Scotland, 1991-94: Validity Issues in Data Envelopment Analysis, *Applied Economics*, 29(11), p.1425-1433.

Patwardhan, A., Davis, J., Murphy, P. and Ryan, S.F. (2013). Comparison of Waiting and Consultation Times in Convenient Care Clinics and Physician Offices A Cross-Sectional Study, *Journal of Primary care & Community Health*, 4(2), p.124-128.

Paul, J.A. and Lin, L. (2012). Models for Improving Patient Throughput and Waiting at Hospital Emergency Departments. *The Journal of Emergency Medicine*, 43(6), p.1119-1126.

Peters, D.H., El-Saharty, S., Siadat, B., Janovsky, K. and Vujicic, M. (2009), *Improving Health Service Delivery in Developing Countries: From Evidence to Action*, Washington: The World Bank

Pol, L.G. and Thomas, R.K. (2002). The Demography of Health and Health care, 2nd ed, USA: Kluwer.

Post, T. (2001). Performance Evaluation in Stochastic Environments Using Mean-variance Data Envelopment Analysis, *Operations Research*, 49(2), p.281–292.

Prior, D. (2006). Efficiency and Total Quality Management in Healthcare Organizations: A Dynamic Frontier Approach, *Annals of Operations Research*, 145(1), p.281-299.

Proudlove, N.C., Black, S., and Fletcher, A. (2007). OR and the Challenge to Improve the NHS: Modelling for Insight and Improvement in In-patient Flows, *Journal of the Operational Research Society*, 58(2), p.145-158.

Puig-Junoy, J. (2000). Partitioning Input Cost Efficiency into its Allocative and Technical Components: An Empirical DEA Application to Hospitals, *Socio Economic Planning Sciences*, 34(3), p.199-218.

Ragan, C.T.S and Lipsey R.J., (2005). *Economics,* 11th edi, Canada: Pearson.

Ramanathan. R. (2003). An Introduction to Data Envelopment Analysis: A Tool for Performance Measurement, New Delhi: Sage.

Ramanathan, R. (2005). Operations Assessment of Hospitals in the Sultanate of Oman, *International Journal of Productions and Operations Management*, 25(1), p.39-54.

Ramirez-Valdivia, M.T., Maturana, S. and Salvo-Garrido, S. (2011). A Multiple-stage Approach for Performance Improvement of Primary Healthcare Practice, *Journal of Medical Systems*, 35(5), p.1015-1028.

Rao, R.V. (2013). *Decision-Making in the Manufacturing Environment Using Graph Theory and Fuzzy Multiple Attribute Decision Making Methods*. London: Springer-Verlag.

Ridge, J.C., Jones, S.K., Nielsen, M.S. and Shahani, A.K. (1998). Capacity Planning for Intensive Care Units, *European Journal of Operational Research*, 105(2), p.346-355.

Rising, E., Baron, R. and Averill, B. (1973). A System Analysis of a University Health Service Outpatient Clinic, *Operations Research*, 21(5), p.1030–1047.

Rouse, P., Harrison, J. and Turner, N. (2011). Cost and Performance: Complements for Improvement, *Journal of Medical Systems*, 35(5), p.1063-1074.

Rozner, S. (2013). *Developing and Using Key Performance Indicators: A Toolkit for Health Sector Managers*. Bethesda, MD: Health Finance & Governance Project, Abt Associates Inc.

Ruenson, P., Host, M., Rainer, A. and Regnell, B. (2012). *Case Study Research in Software Engineering: Guidelines and Examples*. New Jersey: John Wiley and Sons.

Ruggeiro, J. (2004). Data Envelopment Analysis with Stochastic Data, *Journal of the Operational Research Society*, 55(9), p.1008-1012.

Ruggiero J. (1996). On the Measurement of Technical Efficiency in the Public Sector, *European Journal of Operational Research*, 90(3), p.553–565.

Ruggiero, J. (1998). Non-Discretionary Inputs in Data Envelopment Analysis, *European Journal Operational Research*, 111(3), p.461-469.

S.C. Ray (1991). Resource Use Efficiency in Public Schools: A Study of Connecticut Data, *Management Science,* 37(12), p.1620-1628.

Saen, R.F. (2009). Using Data Envelopment Analysis for Ranking Suppliers in the Presence of Non-discretionary Factors, *International Journal of Procurement Management*, 2(3), p.229-243.

Safdar, K.A., Emrouznejad, A. and Dey, P.K. (2016). Assessing the Queuing Process Using Data Envelopment Analysis: An Application in Health Centres. *Journal of Medical Systems*, 40(1), p.1-13.

Salinas-Jimenez, J. and Smith, P. (1996). Data Envelopment Analysis Applied to Quality in Primary Healthcare, *Annals of Operations Research*, 67(1), p.141-161.

Samorani, M. and LaGanga, L.R. (2015). Outpatient Appointment Scheduling Given Individual Day-Dependent No-Show Predictions, *European Journal of Operational Research*, 240(1), p.245-257.

Santibanez, P., Chow, V.S., French, J., Puterman, M.L. and Tyldesley, S. (2009). Reducing Patient Wait Times and Improving Resource Utilization at British Columbia Agency's Ambulatory Care Unit through Simulation, *Health Care Management Science,* 12(4), p. 392-407.

Saremi, A., Jula, P., ElMekkawy, T. and Wang, G.G. (2015). Bi-Criteria Appointment Scheduling of Patients with Heterogeneous Service Sequences, *Dynamic Frameworks with Applications*, 42(8), p. 4029-4041.

Saunders, M., Lewis, P. and Thornhill, A. (2009). *Research Methods for Business Students*. 5th ed, UK: Pearson Education.

Saunders, M., Lewis, P. and Thornhill, A. (2012). *Research Methods for Business Students*. 6th ed, UK: Pearson Education.

Seiford, L.M. (1996). Data Envelopment Analysis: The Evolution of the State of Art (1978-1995), *Journal of Productivity Analysis*, 7(2-3), p.99-137.

Seung-Chul, K. and Ira, H. (2000). Flexible Bed Allocation and Performance in the Intensive Care Unit, *Journal of Operations Management*, 18(4), p.427-443.

Shahhoseini, R., Tofighi, S., Jaafaripooyan, E. and Safiaryan, R. (2011). Efficiency Measurement in Developing Countries: Application of Data Envelopment Analysis for Iranian Hospitals, *Health Services Management Research*, 24(2), p.75-80.

Shaikh, B., Rabbani, F., Safi, N., Dawar, Z. (2010). Contracting of Primary Health care Services in Pakistan: Is Up-scaling a Pragmatic Thinking, *Journal of the Pakistan Medical Association*, 60(5), p. 387-389.

Sheikh, M., Mahamoud, A. and Househ, M. (2015). *Transforming Public Health in Developing Nations*, USA: IGI Global.

Sheikhzada, Y., Roudsari, A.V., Vahidi, R.G., Emrouznejad, A. and Dastgiri, S. (2012). Public and Private Hospital Services Reform Using Data Envelopment Analysis to Measure Technical, Scale, Allocative and Cost Efficiencies, *Health Promotion Perspectives*, 2(1), p.28-41.

Sherman, H.D. and Zhu, J. (2006). *Data Envelopment Analysis Explained*, Service Productivity Management: Improving Service Performance Using Data Envelopment Analysis (DEA), USA: Springer, p.49-88.

Silva, F. and Serra, D. (2008). Locating Emergency Services with Different Priorities: The Priority Queuing Covering Location Problem, *Journal of the Operational Research Society*, 59(9), p. 1229-1238.

Silvester, K., Lendon, R., Bevan, H., Steyn, R. and Walley, P. (2004). Reducing Waiting Times in the NHS: Is Lack of Capacity the Problem? *Clinician in Management*, 12, p. 1-7.

Sinclair, J.A.C. (2009). *China's Healthcare Reform,* photograph, Available at: <http://www.chinabusinessreview.com/chinas-healthcare-reform/> [Accessed: 01 May 2015].

Song, H., Tucker, A.L. and Murrell, K.L. (2015). The Diseconomies of Queue Pooling: An Empirical Investigation of Emergency Department Length of Stay, *Management Science*, 61(12), p.3032-3053.

Spinks, J. and Hollingsworth, B. (2009). Cross-Country Comparisons of Technical Efficiency of Health Production: A Demonstration of Pitfalls, *Applied Economics*, 41(4), p.417-427.

Tan, J., (2013). *Healthcare Information Technology Innovation and Sustainability: Frontiers and Adoption*. USA: IGI Global.

Tang, J., Yan, C. and Cao, P. (2014). Appointment Scheduling Algorithm Considering Routine and Urgent Patients, *Dynamic Frameworks with Applications*, 41(10), p.4529-4541.

Thanassoulis, E. (2001). *Introduction to the Theory and Application of Data Envelopment Analysis: A Foundation Text with Integrated Software*. USA: Kluwer Academic Publishers.

Thanassoulis, E., Boussofiane, A. and Dyson, R.G. (1996). A Comparison of DEA and Ratio Analysis as Tools for Performance Measurement, *Omega*, 24(3), p.229-244.

Thanassoulis, E., Portela, M.C.S. and Graveney, M. (2014). Using DEA to Estimate Potential Savings at GP Units at Medical Specialty Level. *Socio-Economic Planning Sciences*, 48(1), p.38-48.

The Research Ethics Guidebook: A Resource for Social Scientists. *Conducting Your Research: Data Storage and Data Security*. Available at: <http://www.ethicsguidebook.ac.uk/Data-storage-and-data-security-308> [Accessed 19 July 2013].

Tsai, P.F. and Molinero, C.M. (2002). A Variable Returns to Scale Data Envelopment Analysis Model for the Joint Determination of Efficiencies with an Example of the UK Health Service, *European Journal of Operational Research*, 141(1), p.21-38.

Van Ineveld, M., van Oostrum, J., Vermeulen, R., Steenhoek, A. and van de Klundert, J. (2015). Productivity and Quality of Dutch Hospitals During System Reform. *Health Care Management Science*, p.1-12.

Velasquez, K. (2014). *Global Environmental Public Health*, photograph, Available at: <http://phlessons.com/345/2014/03/26/mobile-medicine-clinics-in-the-developing-world/> [Accessed: 29 April 2015].

Venkatasubramanian, K.V. (2015). *India's Healthcare System Lies in Tatters*, photograph, Available at: <http://www.millenniumpost.in/NewsContent.aspx?NID=97975> [Accessed: 29 April 2015].

Wagner, J.M., Shimshak, D.G. and Novak, M.A. (2003). Advances in Physician Profiling: The Use of DEA, *Socio-Economic Planning Sciences*, 37(2), p.141-163.

Wang, Q. (2004). Modelling and Analysis of High-risk Patient Queues. *European Journal of Operational Research*, 155(2), p.502-515.

Wei, J., Shen, L., Yang, H.B., Qin, J.B., Huang, W., Zhang, J.J., Gong, Q.Y., Li, X.X. and Yang, T.B. (2015). Development and Validation of a Chinese Outpatient Satisfaction Questionnaire: Evidence from 46 Public General Hospitals and 5151 Outpatients. *Public Health*, 129(11), p.1523-1529.

Weng,S.J., Wu, T., Blackhurst J. and Mackulak, G. (2009). An Extended DEA Model for Hospital Performance Evaluation and Improvement, *Health Services and Outcomes Research Methodology*, 9(1), p.39-53.

Wijewickrama, A. and Takakuwa, S. (2005). Outpatient Appointment Scheduling in a Multi-Facility System, In: *Proceedings of the 2005 Winter Simulation Conference,* ed. M.E. Kurl, N.M. Steiger, F.B. Armstrong, J.A. Joines, 2264-2273. Piscataway, New Jersey: IEEE.

Wijewickrama, A. and Takakuwa, S. (2008), Outpatient Appointment Scheduling in a Multi Facility System, In*: Proceedings of the 2008 Winter Simulation Conference,* ed. S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler, 1563-1571. Piscataway, New Jersey: IEEE.

World Health Statistics (2010). *Monitoring the Building Blocks of Health Systems: A Handbook of Indicators and Their Measurement Strategies*, Available at: <http://www.who.int/healthinfo/systems/WHO_MBHSS_2010_section1_web.pdf > [Accessed: 5 May 2016].

World Health Statistics (2014). *World Health Organisation*. Available at: <http://www.who.int/gho/publications/world_health_statistics/2014/en/> [Accessed: 28 April 2015].

World Health Statistics (2015). *World Health Organisation.* Available at: <http://www.who.int/gho/publications/world_health_statistics/2015/en/> [Accessed: 15 January 2016].

Worthington, A.C. (2004). Frontier Efficiency Measurement in Healthcare: A Review of Empirical Techniques and Selected Applications, *Medical Care Research and Review*, 61(2), p.135-170.

Worthington, D. (2009). Reflections on Queuing Modeling from the Last 50 Years, *Journal of the Operational Research Society*, 60, p. 83-92.

Wright, P.D., Bretthauer, K.M. and Cote, M.J. (2006). Re-examining the Nursing Scheduling Problem: Staffing Ratios and Nursing Shortages, *Decision Sciences*, 37(1), p.39-70.

Yang, Z. and Paradi, J.C. (2003). Benchmarking Competitive Banking Units Using 'Handicapped' DEA. *Omega*, (submitted for publication).

Yankovic, N. and Green, L.V. (2011), Identifying Good Nursing Levels: A Queuing Approach, *Operations Research*, 59(4), p.942-955.

Yearly Statistics Report (2013). Statistics Department, Designated Hospital, Pakistan.

Yeboah, E.K. and Thomas, M.E. (2010). A Cost-effective Way of Reducing Outpatient Clinic Waiting Times: How We Did It, *Internal Journal of Healthcare Administration,* 7(1).

Yin, R.K. (2014). *Case Study Research: Design and Methods*. 5th ed, USA: Sage.

Zere, E., McIntyre, D., and Addison, T. (2001). Technical Efficiency and Productivity of Public Sector Hospitals in Three South African Provinces, *South African Journal of Economics*, 69(2), p.336-358.

Zhang, Y., Puterman, M.L., Nelson, M., and Atkins, D. (2012). A Simulation Optimization Approach to Long-Term Care Capacity Planning, *Operations Research*, 60(2), p. 249-261.

Zhu, J. (2009). *Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets*. 2nd ed, New York: Springer.

Zhu, Z., Heng, B.H. and Teow, K.L. (2012). Analysis of Factors Causing Long Patient Waiting Time and Clinic Overtime in Outpatient Clinics, *Journal of Medical Systems*, 36(2), p.707-713.

APPENDICES

Appendix 1: Ethical Approval for Data Collection from Aston University

Aston Business School

Komal Safdar

ABS Research Student

Date: 19th November 2013

Dear Komal,

I am pleased to be able to inform you that the ABS Research Ethics Committee has approved your ethics application. For future reference please quote 33:10/13.

Aston University

Content has been removed for copyright reasons

Appendix 2a: Information Sheet for Interview Respondents

Name of the research project: **Better Queue Management Using DEA: An Application in a Large Public Hospital of a Developing Country**

My name is **Komal Aqeel Safdar**. I am a PhD student at Aston University, Birmingham, UK. I would like to request you to take part in this short interview which is related to my thesis.

The main aim of this research is to study that how long the patients have to wait to receive a service and how it can be reduced.

I will be very grateful if you could spare some time to provide me with information regarding the procedures of patients visiting and what is your opinion of the reasons patients have to wait and how it can be improved. Your response will prove to be extremely valuable for my research as it will provide information about the functioning of the system and how it can be improved, to reduce the long waiting times of patients. This interview will not take more than 10 minutes.

I can assure you that the information gathered through this interview will be used for research purposes ONLY and will be kept confidential at all times. Your participation is entirely voluntary and even if you choose to take part in this study, you can withdraw at any time you want. If you choose to participate, you will be requested to sign a consent form before proceeding with interview. During the interview, you can withdraw at any time. The interview will not take more than 10 minutes.

This research study has been approved by the Ethical Review Committee of Military Hospital and the Ethics Committee of Aston University.

If you require any further information about the research study, please do not hesitate to contact me. My email address is safdarka@aston.ac.uk.

Thank you for reading the information sheet.

If you wish to participate, please turn to the next page to sign the consent form.

Appendix 2b: Consent Form for Interview Respondents

Full title of Project: **Better Queue Management Using DEA: An Application in a Large Public Hospital of a Developing Country**

Name: Komal Aqeel Safdar

Position: PhD Student at Aston University, Birmingham, UK

Contact Address of Researcher: safdarka@aston.ac.uk

	Please tick box
I confirm that I have read and understand the information sheet for the above study and have had the opportunity to ask questions.	☐
I understand that my participation is voluntary and that I am free to withdraw at any time, without giving reason.	☐
I agree to take part in this study.	☐

	Please tick box	
	Yes	No
I agree to the interview being recorded through note-taking	☐	☐
I agree to the information being shared with the supervisors	☐	☐
I agree to the use of anonymised quotes in publications	☐	☐

Date: ---

Signature: ---------------------------------------

Appendix 3a: Information Sheet for Volunteers

Name of the research project: **Better Queue Management Using DEA: An Application in a Large Public Hospital of a Developing Country**

My name is **Komal Aqeel Safdar**. I am a PhD student at Aston University, Birmingham, UK. I would like to request you to take part in the data collection for my thesis, which includes the waiting times of patients.

The main aim of this research is to study that how long the patients have to wait to receive a service and how it can be reduced.

I will be extremely grateful if you could spare some time to help me note down the time when patients move from one stage to another. This information will prove to be extremely valuable for my research as it will help in understanding the current system and how it can be improved, in order to reduce the long waiting times of patients.

I can assure you that the information gathered through this interview will be used for research purposes ONLY and will be kept confidential at all times. Your anonymity will be maintained. Your participation is entirely voluntary and you can withdraw any time during this activity. If you choose to participate, you will be requested to sign a consent form before proceeding with data collection. For your convenience, a time sheet will be provided to you to note down the times.

This research study has been approved by the Ethical Review Committee of Military Hospital and the Ethics Committee of Aston University.

If you require any further information about the research study, please do not hesitate to contact me. My email address is safdarka@aston.ac.uk.

Thank you for reading the information sheet.

If you wish to participate, please turn to the next page to sign the consent form.

Appendix 3b: Consent Form for Volunteers

Full title of Project: **Better Queue Management Using DEA: An Application in a Large Public Hospital of a Developing Country**

Name: Komal Aqeel Safdar

Position: PhD Student at Aston University, Birmingham, UK

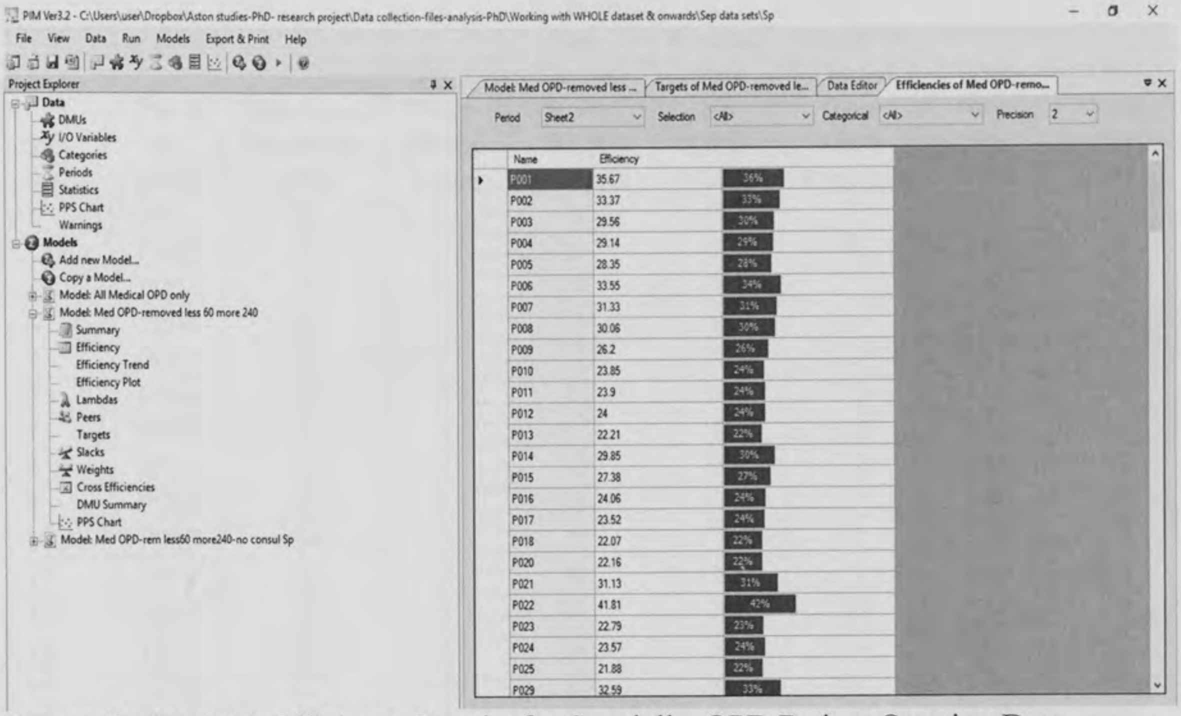
Contact Address of Researcher: safdarka@aston.ac.uk

	Please tick box
I confirm that I have read and understand the information sheet for the above study and have had the opportunity to ask questions.	☐
I understand that my participation is voluntary and that I am free to withdraw at any time, without giving reason.	☐
I agree to take part in this study.	☐

Date: ---

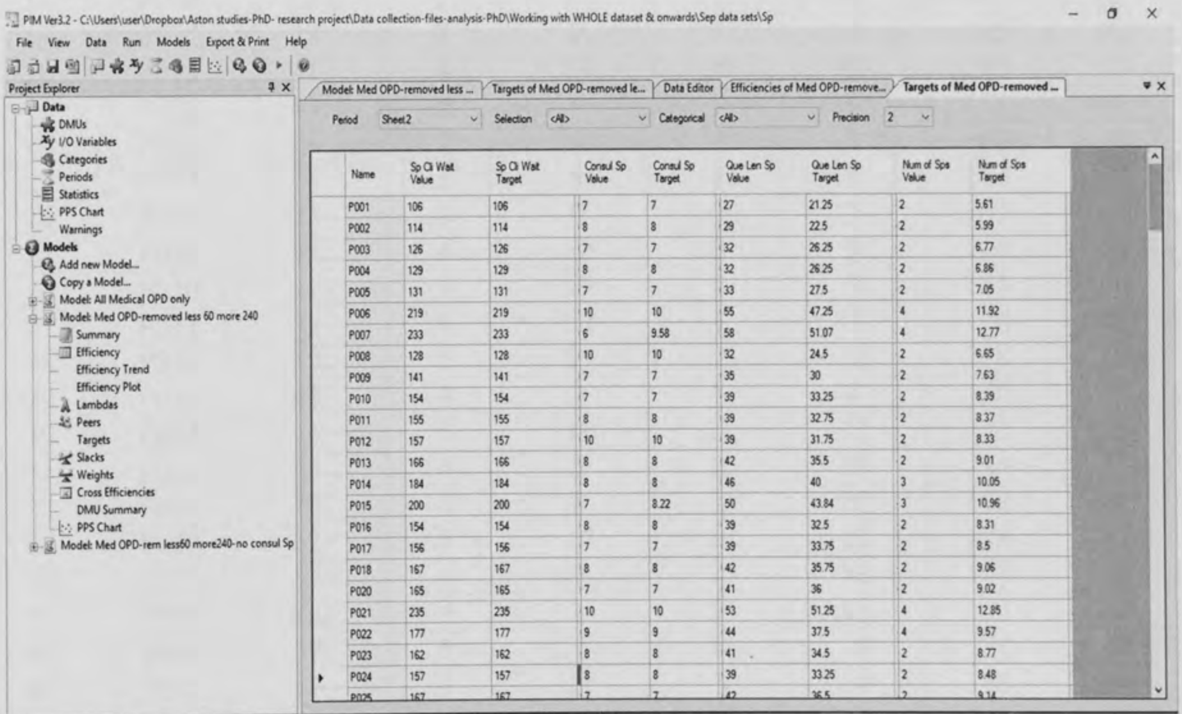
Signature: ---------------------------------------

Appendix 4: Screenshots from PIM-DEA Software of DEA Analysis for the Specialist OPD



Name	Efficiency	
P001	35.67	36%
P002	33.37	33%
P003	29.56	30%
P004	29.14	29%
P005	28.35	28%
P006	33.55	34%
P007	31.33	31%
P008	30.06	30%
P009	26.2	26%
P010	23.85	24%
P011	23.9	24%
P012	24	24%
P013	22.21	22%
P014	29.85	30%
P015	27.38	27%
P016	24.06	24%
P017	23.52	24%
P018	22.07	22%
P020	22.16	22%
P021	31.13	31%
P022	41.81	42%
P023	22.79	23%
P024	23.57	24%
P025	21.88	22%
P029	32.59	33%

Appendix 4a: DEA Efficiency Results for Specialist OPD Patient Queuing Data



Name	Sp Cli Wait Value	Sp Cli Wait Target	Consul Sp Value	Consul Sp Target	Que Len Sp Value	Que Len Sp Target	Num of Sps Value	Num of Sps Target
P001	106	106	7	7	27	21.25	2	5.61
P002	114	114	8	8	29	22.5	2	5.99
P003	126	126	7	7	32	26.25	2	6.77
P004	129	129	8	8	32	26.25	2	6.86
P005	131	131	7	7	33	27.5	2	7.05
P006	219	219	10	10	55	47.25	4	11.92
P007	233	233	6	9.58	58	51.07	4	12.77
P008	128	128	10	10	32	24.5	2	6.65
P009	141	141	7	7	35	30	2	7.63
P010	154	154	7	7	39	33.25	2	8.39
P011	155	155	8	8	39	32.75	2	8.37
P012	157	157	10	10	39	31.75	2	8.33
P013	166	166	8	8	42	35.5	2	9.01
P014	184	184	8	8	46	40	3	10.05
P015	200	200	7	8.22	50	43.84	3	10.96
P016	154	154	8	8	39	32.5	2	8.31
P017	156	156	7	7	39	33.75	2	8.5
P018	167	167	8	8	42	35.75	2	9.06
P020	165	165	7	7	41	36	2	9.02
P021	235	235	10	10	53	51.25	4	12.85
P022	177	177	9	9	44	37.5	4	9.57
P023	162	162	8	8	41	34.5	2	8.77
P024	157	157	8	8	39	33.25	2	8.48
P025	167	167	7	7	42	36.5	2	9.14

Appendix 4b: DEA Target Values for Specialist OPD Patient Queuing Data

Appendix 5: Complete DEA Results for the Specialist OPD

Ser No	Pat # in data	Efficiency Specialist OPD	Num of Sps Value	Num of Sps Target from DEA analysis	Consul time Sp OPD Value	Consul time Sp OPD Target from DEA analysis	Wait time Sp OPD	Que Len Sp OPD
1	P001	38	2	5	7	7	106	27
2	P002	36	2	5	8	8	114	29
3	P003	30	2	7	7	7	126	32
4	P004	30	2	7	8	8	129	32
5	P005	29	2	7	7	7	131	33
6	P006	34	4	12	10	10	219	55
7	P007	31	4	13	6	10	233	58
8	P008	35	2	6	10	10	128	32
9	P009	27	2	8	7	7	141	35
10	P010	24	2	8	7	7	154	39
11	P011	24	2	8	8	8	155	39
12	P012	25	2	8	10	10	157	39
13	P013	23	2	9	8	8	166	42
14	P014	30	3	10	8	8	184	46
15	P015	27	3	11	7	8	200	50
16	P016	25	2	8	8	8	154	39
17	P017	24	2	8	7	7	156	39
18	P018	22	2	9	8	8	167	42
19	P020	22	2	9	7	7	165	41
20	P021	31	4	13	10	10	235	53
21	P022	43	4	9	9	9	177	44
22	P023	23	2	9	8	8	162	41
23	P024	24	2	8	8	8	157	39
24	P025	22	2	9	7	7	167	42
25	P029	33	4	12	8	9	224	56
26	P030	41	4	10	9	9	183	46
27	P031	42	4	10	7	7	174	44
28	P032	40	5	13	12	12	238	48
29	P034	43	4	9	8	8	173	43
30	P035	46	5	11	12	12	211	53
31	P036	83	4	5	10	10	124	31
32	P037	48	4	8	10	10	162	41
33	P038	52	5	10	12	12	189	47
34	P039	57	4	7	10	10	144	36
35	P040	100	4	4	13	13	153	38
36	P041	57	5	9	10	10	171	43
37	P042	35	4	11	5	14	239	34
38	P043	39	4	10	5	13	211	31
39	P044	36	4	11	6	14	223	33
40	P045	41	4	10	3	12	193	29
41	P046	36	4	11	3	14	226	33
42	P047	43	4	9	6	11	182	28
43	P048	43	5	12	5	13	223	38

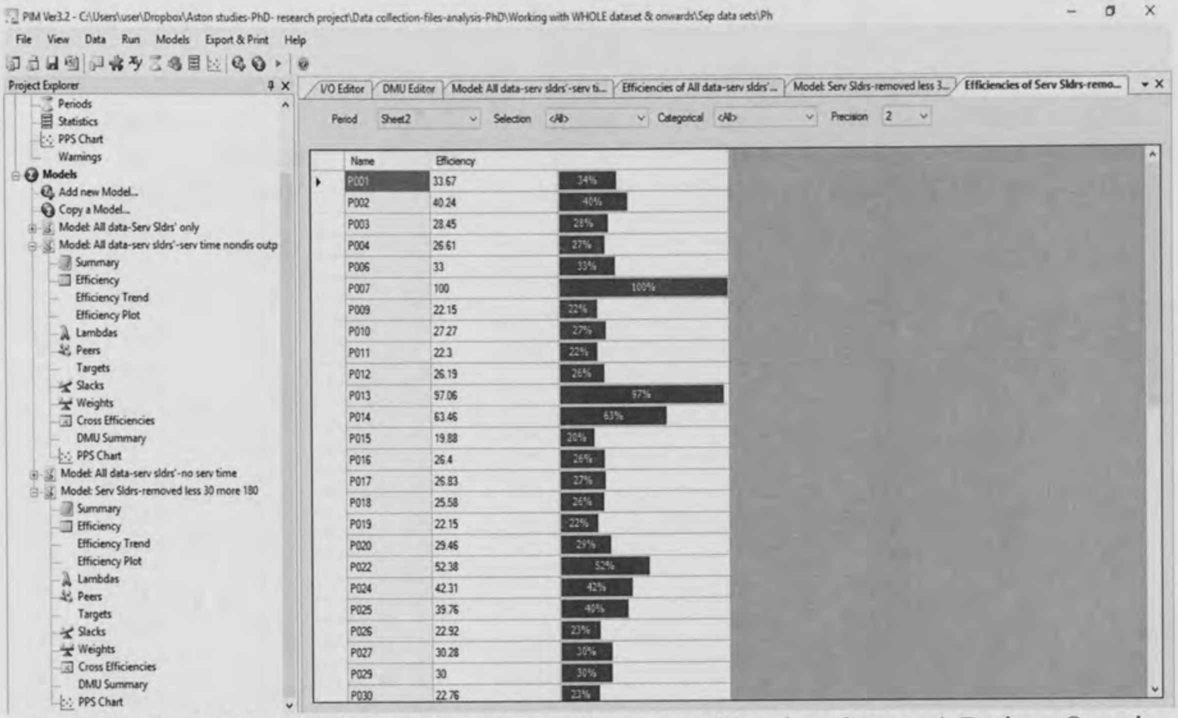
44	P049	45	5	11	1	14	227	33
45	P050	39	4	10	2	12	200	31
46	P051	43	4	9	5	12	191	28
47	P052	45	5	11	2	14	228	33
48	P053	56	5	9	2	9	170	31
49	P054	44	4	9	5	9	173	32
50	P055	43	4	9	6	11	180	29
51	P056	48	5	10	1	13	219	31
52	P057	44	4	9	4	11	192	27
53	P058	50	5	10	4	12	198	31
54	P059	54	5	9	1	12	191	28
55	P060	55	5	9	4	11	177	29
56	P061	59	5	9	3	10	166	27
57	P062	55	5	9	3	10	176	29
58	P063	41	4	10	2	11	190	31
59	P064	73	4	6	2	5	102	21
60	P065	100	4	4	3	3	73	16
61	P066	100	4	4	2	3	73	18
62	P067	35	2	6	5	7	112	18
63	P068	47	2	4	6	6	87	14
64	P069	30	2	7	7	8	133	20
65	P070	22	2	9	6	11	181	29
66	P071	24	2	8	5	10	160	26
67	P075	66	4	6	6	7	119	19
68	P076	43	4	9	6	11	181	29
69	P078	53	4	8	6	9	147	24
70	P079	38	2	5	6	7	104	16
71	P081	66	4	6	6	7	118	19
72	P083	84	4	5	6	6	94	15
73	P084	76	4	5	7	7	109	16
74	P085	52	5	10	6	12	211	29
75	P086	44	4	9	6	11	178	29
76	P088	42	4	10	6	9	180	34
77	P089	25	2	8	5	8	149	28
78	P090	49	4	8	6	9	158	28
79	P091	51	4	8	7	8	150	27
80	P092	35	4	11	5	11	213	41
81	P093	34	4	12	4	12	225	41
82	P094	32	4	12	7	13	238	43
83	P095	40	4	10	5	10	189	36
84	P096	41	5	12	6	11	229	45
85	P097	60	5	8	5	8	157	29
86	P098	50	4	8	8	9	155	26
87	P099	45	2	4	5	5	86	16
88	P100	53	4	8	6	8	145	26
89	P101	74	4	5	5	5	102	19
90	P102	49	4	8	6	8	154	28
91	P103	54	5	9	7	10	178	31
92	P104	53	5	9	7	10	181	32
93	P105	92	4	4	6	6	89	16
94	P106	54	4	7	8	8	142	24
95	P107	49	4	8	6	8	154	28
96	P108	42	4	10	5	9	181	35

97	P109	50	4	8	6	9	155	27
98	P110	24	2	8	8	10	171	25
99	P111	26	2	8	8	10	160	23
100	P112	23	2	9	6	10	169	27
101	P113	26	2	8	6	9	148	24
102	P114	31	2	6	7	8	127	20
103	P115	39	4	10	5	11	197	34
104	P116	39	4	10	6	12	198	32
105	P117	40	4	10	6	13	212	30
106	P118	39	4	10	7	13	216	31
107	P120	33	2	6	7	8	121	18
108	P121	48	5	10	6	13	232	31
109	P122	42	4	10	6	11	186	30
110	P123	41	4	10	8	12	200	29
111	P124	40	4	10	6	13	235	30
112	P125	46	4	9	5	9	168	29
113	P126	82	5	6	5	6	116	21
114	P127	43	4	9	9	12	187	28
115	P128	100	4	4	7	7	89	14
116	P129	60	4	7	8	8	146	20
117	P130	63	5	8	7	10	167	24
118	P131	41	4	10	6	12	192	29
119	P132	30	3	10	7	13	198	30
120	P133	19	2	10	6	13	223	31
121	P134	33	3	9	6	11	180	29
122	P135	50	4	8	5	10	160	24
123	P136	57	4	7	8	9	145	21
124	P137	81	5	6	6	7	121	19
125	P138	100	4	4	8	8	103	14
126	P139	100	4	4	7	7	92	13
127	P140	100	4	4	5	5	79	12
128	P143	21	2	10	7	12	222	29
129	P145	21	2	10	6	12	192	29
130	P149	26	2	8	4	9	153	24
131	P150	40	4	10	8	13	225	30
132	P151	39	4	10	8	13	232	31
133	P153	27	2	7	5	9	144	22
134	P154	45	4	9	6	10	173	28
135	P155	39	4	10	6	12	201	33
136	P156	75	4	5	5	7	107	16
137	P157	76	4	5	7	7	107	16
138	P158	40	4	10	5	12	198	31
139	P159	21	2	10	5	12	230	29
140	P160	60	4	7	6	8	129	21
141	P161	22	2	9	6	11	176	28
142	P162	22	2	9	5	11	176	28
143	P163	28	2	7	6	9	139	22
144	P164	60	4	7	8	8	134	20
145	P165	63	4	6	8	8	132	19
146	P167	71	4	6	7	7	114	17
147	P168	44	4	9	6	11	178	29
148	P169	21	2	10	6	12	230	29
149	P170	43	4	9	6	11	184	29

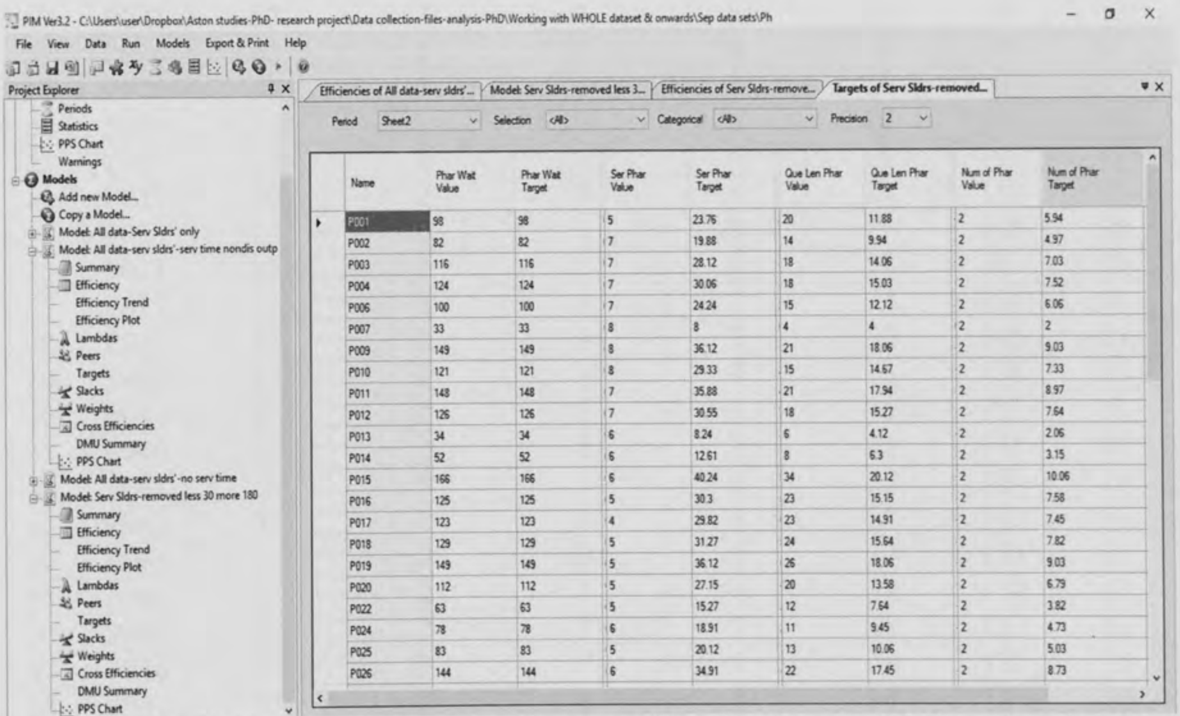
150	P171	63	5	8	7	10	159	24
151	P172	41	4	10	6	12	205	29
152	P173	62	4	6	6	8	126	20
153	P174	43	4	9	7	11	184	29
154	P175	54	4	7	7	9	144	23
155	P176	38	4	10	6	12	203	33
156	P177	41	4	10	7	11	191	31
157	P178	58	4	7	5	8	134	22
158	P179	43	4	9	6	12	196	28
159	P180	39	4	10	7	12	202	32
160	P181	33	4	12	7	15	239	38

(Key: Se No: Serial Number; Pat #: Patient Number; Num of Sps: Number of specialists; Consul time: Consultation time; Sp OPD: Specialist OPD; Que Len: Queue Length)

Appendix 6: Screenshots from PIM-DEA Software of DEA Analysis for the Pharmacy (Serving Counter)



Appendix 6a: DEA Efficiency Results for Pharmacy (Serving Counter) Patient Queuing Data



Appendix 6b: DEA Target Values for Pharmacy (Serving Counter) Patient Queuing Data

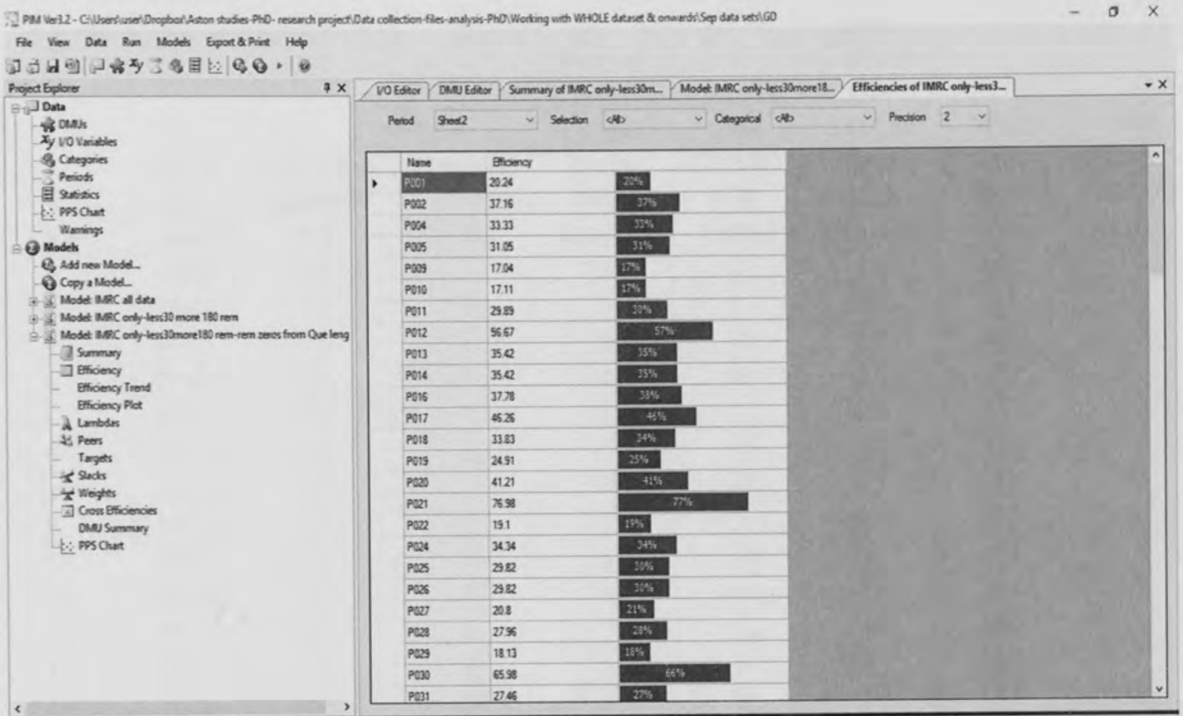
Appendix 7: Complete DEA Results for Pharmacy (Serving Counter)

Ser No	Pat # in data	Efficiency Phar	Num of Pharm Value	Num of Pharm Target from DEA Analysis	Service time Phar Value	Service time Phar Target from DEA Analysis	Wait time Phar	Que Len Phar
1	P001	34	2	6	5	24	98	20
2	P002	40	2	5	7	20	82	14
3	P003	28	2	7	7	28	116	18
4	P004	27	2	8	7	30	124	18
5	P006	33	2	6	7	24	100	15
6	P007	100	2	2	8	8	33	4
7	P009	22	2	9	8	36	149	21
8	P010	27	2	7	8	29	121	15
9	P011	22	2	9	7	36	148	21
10	P012	26	2	8	7	31	126	18
11	P013	97	2	2	6	8	34	6
12	P014	63	2	3	6	13	52	8
13	P015	20	2	10	6	40	166	34
14	P016	26	2	8	5	30	125	23
15	P017	27	2	7	4	30	123	23
16	P018	26	2	8	5	31	129	24
17	P019	22	2	9	5	36	149	26
18	P020	29	2	7	5	27	112	20
19	P022	52	2	4	5	15	63	12
20	P024	42	2	5	6	19	78	11
21	P025	40	2	5	5	20	83	13
22	P026	23	2	9	6	35	144	22
23	P027	30	2	7	6	26	109	17
24	P029	30	2	7	6	27	110	15
25	P030	23	2	9	9	35	145	20
26	P031	35	2	6	6	23	95	15
27	P032	25	2	8	7	32	134	21
28	P034	23	2	9	7	35	145	21
29	P035	21	2	10	7	38	158	23
30	P037	80	2	3	7	10	42	5
31	P040	25	2	8	7	32	133	21
32	P041	30	2	7	6	26	109	19
33	P042	33	2	6	6	24	101	16
34	P043	22	2	9	6	37	153	23
35	P044	25	2	8	6	32	130	17
36	P045	28	2	7	7	28	117	16
37	P046	27	2	7	9	30	122	15
38	P047	27	2	8	6	30	124	16
39	P049	32	2	6	6	25	103	16
40	P050	30	2	7	6	26	109	17
41	P051	23	2	9	7	35	145	23
42	P052	28	2	7	6	29	118	19
43	P053	20	2	10	5	41	168	26

44	P054	57	2	4	5	14	58	8
45	P055	19	2	10	6	41	171	24
46	P058	30	2	7	5	27	111	14
47	P059	19	2	10	5	41	171	23
48	P063	28	2	7	6	28	116	17
49	P064	25	2	8	6	32	130	21
50	P065	61	2	3	6	13	54	8
51	P066	35	2	6	6	23	93	15
52	P067	39	2	5	5	21	85	12

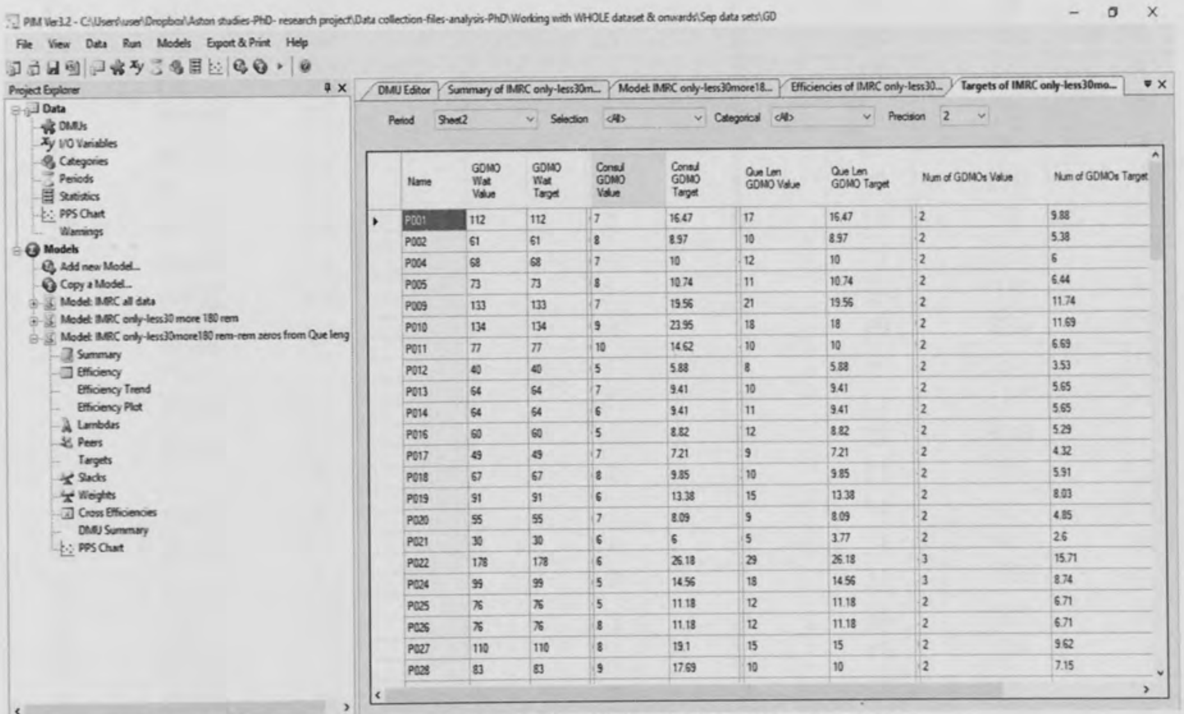
(Key: Se No: Serial Number; Pat #: Patient Number; Phar: pharmacy; Num of Pharm: Number of pharmacists; Que Len: Queue Length)

Appendix 8: Screenshots from PIM-DEA Software of DEA Analysis for the GDMO Clinic



Name	Efficiency	
P001	20.24	20%
P002	37.16	37%
P004	33.33	33%
P005	31.05	31%
P009	17.04	17%
P010	17.11	17%
P011	29.89	30%
P012	56.67	57%
P013	35.42	35%
P014	35.42	35%
P016	37.78	38%
P017	46.26	46%
P018	33.83	34%
P019	24.91	25%
P020	41.21	41%
P021	76.98	77%
P022	19.1	19%
P024	34.34	34%
P025	29.82	30%
P026	29.82	30%
P027	20.8	21%
P028	27.96	28%
P029	18.13	18%
P030	65.98	66%
P031	27.46	27%

Appendix 8a: DEA Efficiency Results for GDMO Clinic Patient Queuing Data



Name	GDMO Wait Value	GDMO Wait Target	Consul GDMO Value	Consul GDMO Target	Que Len GDMO Value	Que Len GDMO Target	Num of GDMOs Value	Num of GDMOs Target
P001	112	112	7	16.47	17	16.47	2	9.88
P002	61	61	8	8.97	10	8.97	2	5.38
P004	68	68	7	10	12	10	2	6
P005	73	73	8	10.74	11	10.74	2	6.44
P009	133	133	7	19.56	21	19.56	2	11.74
P010	134	134	9	23.95	18	18	2	11.69
P011	77	77	10	14.62	10	10	2	6.69
P012	40	40	5	5.88	8	5.88	2	3.53
P013	64	64	7	9.41	10	9.41	2	5.65
P014	64	64	6	9.41	11	9.41	2	5.65
P016	60	60	5	8.82	12	8.82	2	5.29
P017	49	49	7	7.21	9	7.21	2	4.32
P018	67	67	8	9.85	10	9.85	2	5.91
P019	91	91	6	13.38	15	13.38	2	8.03
P020	55	55	7	8.09	9	8.09	2	4.85
P021	30	30	6	6	5	3.77	2	2.6
P022	178	178	6	26.18	29	26.18	3	15.71
P024	99	99	5	14.56	18	14.56	3	8.74
P025	76	76	5	11.18	12	11.18	2	6.71
P026	76	76	8	11.18	12	11.18	2	6.71
P027	110	110	8	19.1	15	15	2	9.62
P028	83	83	9	17.69	10	10	2	7.15

Appendix 8b: DEA Target Values for GDMO Clinic Patient Queuing Data

Appendix 9: Complete DEA Results for GDMO Clinic

Ser No	Pat # in data	Efficiency GDMO Cli	Num of GDMOs Value	Num of GDMOs Target from DEA Analysis	Consul time GDMO Cli Value	Consul time GDMO Cli Target from DEA Analysis	Wait time GDMO Cli	Que Len GDMO Cli
1	P001	20	2	10	7	16	112	17
2	P002	37	2	5	8	9	61	10
3	P004	33	2	6	7	10	68	12
4	P005	31	2	6	8	11	73	11
5	P009	17	2	12	7	20	133	21
6	P010	17	2	12	9	24	134	18
7	P011	30	2	7	10	15	77	10
8	P012	57	2	4	5	6	40	8
9	P013	35	2	6	7	9	64	10
10	P014	35	2	6	6	9	64	11
11	P016	38	2	5	5	9	60	12
12	P017	46	2	4	7	7	49	9
13	P018	34	2	6	8	10	67	10
14	P019	25	2	8	6	13	91	15
15	P020	41	2	5	7	8	55	9
16	P021	77	2	3	6	6	30	5
17	P022	19	3	16	6	26	178	29
18	P024	34	3	9	5	15	99	18
19	P025	30	2	7	5	11	76	12
20	P026	30	2	7	8	11	76	12
21	P027	21	2	10	8	19	110	15
22	P028	28	2	7	9	18	83	10
23	P029	18	2	11	5	18	125	23
24	P030	66	2	3	9	9	38	5
25	P031	27	3	11	8	19	124	18
26	P032	37	2	5	5	9	61	10
27	P033	34	3	9	4	15	101	20
28	P034	27	3	11	9	24	130	17
29	P035	30	2	7	8	11	76	12
30	P036	40	3	7	7	14	85	12
31	P037	25	2	8	8	14	91	13
32	P038	31	2	7	7	11	74	12
33	P039	36	3	8	7	14	95	15
34	P040	24	2	8	12	17	96	13
35	P041	36	3	8	6	14	95	16
36	P042	88	2	2	9	9	35	5
37	P043	41	3	7	8	13	83	12
38	P045	47	2	4	5	7	48	7
39	P046	45	2	4	7	7	50	8
40	P047	38	2	5	7	11	61	8
41	P048	100	2	2	10	10	38	5
42	P049	30	2	7	6	11	76	12

43	P051	38	3	8	6	14	90	13
44	P052	27	2	8	8	14	86	12
45	P054	25	2	8	6	14	90	13
46	P055	22	2	9	9	16	104	15
47	P056	27	2	7	8	16	84	11
48	P057	38	2	5	9	14	64	7
49	P058	33	2	6	5	10	68	11
50	P059	31	2	6	6	13	73	10
51	P061	21	2	9	11	18	107	15
52	P062	40	3	8	8	14	86	12
53	P064	100	3	3	8	8	35	4
54	P065	31	3	10	7	19	109	15
55	P066	53	3	6	7	14	66	8
56	P067	34	3	9	6	15	99	15
57	P068	48	3	6	9	12	71	10
58	P069	39	3	8	6	13	88	13
59	P070	47	3	6	5	11	72	12
60	P072	71	3	4	7	7	48	7
61	P073	36	3	8	10	16	95	13
62	P074	37	3	8	6	14	92	19
63	P075	15	2	13	4	22	151	29
64	P076	24	2	8	7	14	93	16
65	P078	100	2	2	8	8	31	4
66	P079	20	3	15	6	25	171	26
67	P080	33	3	9	8	15	102	15
68	P081	60	3	5	9	12	58	7
69	P082	38	2	5	8	11	61	8
70	P083	18	2	11	7	18	123	19
71	P084	24	2	8	5	14	96	14
72	P085	44	3	7	7	11	77	12
73	P086	54	3	6	8	10	63	9
74	P088	52	3	6	8	10	65	10
75	P089	30	3	10	11	20	116	16
76	P090	93	3	3	6	7	37	5
77	P091	32	2	6	3	10	70	13
78	P093	22	3	14	7	23	153	26
79	P094	29	3	10	8	17	116	20
80	P095	27	2	7	9	13	83	12
81	P097	33	2	6	5	10	69	13
82	P098	23	2	9	7	15	99	16
83	P099	80	3	4	7	7	43	6
84	P100	37	3	8	7	14	93	14
85	P102	59	2	3	7	8	39	5
86	P103	33	2	6	7	13	69	9
87	P104	33	2	6	5	10	69	13
88	P105	57	2	3	7	7	40	6
89	P106	29	2	7	8	14	80	11
90	P107	33	2	6	7	11	69	10
91	P108	93	3	3	7	7	37	5
92	P109	31	2	6	6	11	73	11
93	P110	26	2	8	6	13	88	14
94	P111	80	3	4	5	10	48	5
95	P112	28	3	11	8	19	120	17

96	P113	49	2	4	6	7	46	7
97	P114	52	2	4	7	8	44	6
98	P115	26	2	8	7	13	88	14
99	P116	44	2	5	5	8	51	8
100	P117	74	3	4	7	7	46	7
101	P118	75	3	4	6	9	46	6
102	P119	39	3	8	5	13	88	13
103	P121	54	3	6	7	10	63	9
104	P123	54	3	6	5	10	63	9
105	P124	30	2	7	6	11	76	12
106	P125	100	3	3	5	5	34	5
107	P126	83	3	4	6	6	41	6
108	P127	31	2	6	6	11	72	11
109	P129	26	3	12	6	20	133	21
110	P130	19	3	16	6	26	178	28
111	P131	31	3	10	5	16	110	17
112	P132	27	3	11	6	19	127	20
113	P133	29	3	11	5	18	119	18
114	P134	26	3	12	6	20	133	20

(Key: Se No: Serial Number; Pat #: Patient Number; Num of GDMOs: Number of General Duty Medical Officers; Consul time: Consultation time; GDMO Cli: GDMO Clinic; Que Len: Queue Length)

Appendix 10: Visual Basic for Applications (VBA) Codes and Spreadsheet Formulae in Excel for Dynamic Framework

Appendix 10a: VBA Code for Running the Queuing-DEA Model

Appendix 10b: VBA Macro (Macro 2) Mentioned in Appendix 10a

Appendix 10c: VBA Macro (Macro 3) Mentioned in Appendix 10a

Column/ Cells	Category	Formulae in Worksheet 1 'DEAefficiency few values'
Column F	Wait time	(C2-B2)*24*60
Column G	Length of Queue	COUNTIF(C2:C2,">"&B3
Cells K4:M7	Constraints for DEA model	-
K4,L4,M4	Constraint for input 'Wait time'	SUMPRODUCT(INDEX(F2:G51,0,1),H2:H51) ≤ INDEX(F2:G51,N2,1)
K5,L5,M5	Constraint for input 'Length of Queue'	SUMPRODUCT(INDEX(F2:G51,0,2),H2:H51)≤ INDEX(F2:G51,N2,2)
K6,L6,M6	Constraint for output 'Number of specialists'	SUMPRODUCT(INDEX(D2:E51,0,1),H2:H51)≥ O3*INDEX(D2:E51,N2,1) *(O3: efficiency of 'each' unit when VBA macro runs)*
K7,L7,M7	Constraint for output 'Consultation time'	SUMPRODUCT(INDEX(D2:E51,0,2),H2:H51)≥ INDEX(D2:E51,N2,2)
		Formulae in Worksheet 2 'DEAslack'
P8	Objective function for slack model	SUM(P4:P7)
P4,P5,P6,P7	Slack values for two inputs and two outputs as calculated by VBA	See VBA codes in Appendix 10a, 10b and 10c
Cells K4:M7	Constraints for DEA Slack model	
K4,L4,M4	Constraint for input 'Wait time'	SUMPRODUCT(INDEX(F2:G51,0,1),H2:H51)+P4 ≤ INDEX(F2:G51,N2,1)
K5,L5,M5	Constraint for input 'Length of Queue'	SUMPRODUCT(INDEX(F2:G51,0,2),H2:H51)+P5 ≤ INDEX(F2:G51,N2,2)
K6,L6,M6	Constraint for output 'Number of specialists'	SUMPRODUCT(INDEX(D2:E51,0,1),H2:H51)-P6 ≥ INDEX(I2:I51,N2,1)*INDEX(D2:E51,N2,1)
K7,L7,M7	Constraint for output 'Consultation time'	SUMPRODUCT(INDEX(D2:E51,0,2),H2:H51)-P7 ≥ INDEX(D2:E51,N2,2)

Appendix 10d: Formulae set up within the Excel Spreadsheet for Worksheet 1 'DEAefficiency few values' and Worksheet 2 'DEAslack'