ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

SCHOOL OF GRADUATE STUDIES

SCHOOL OF CIVIL AND ENVIRONMENTAL ENGINEERING

LOGISTICS PERFORMANCES OF HEALTHCARE SYSTEM USING QUEUE ANALYSIS, FOCUSING ON PATIENT FLOW: THE CASE OF ST. PAUL'S HOSPITAL, ETHIOPIA

By:

Abel Nigussie Tefera

A Thesis Submitted to the School of Graduate Studies in Partial fulfilment of the requirements for the degree of Master of Science in Road and Transport Engineering

Advisor:

Girma Gebresenbet (Professor)

May, 2020

Uppsala, Sweden

ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

SCHOOL OF GRADUATE STUDIES

SCHOOL OF CIVIL AND ENVIRONMENTAL ENGINEERING

LOGISTICS PERFORMANCES OF HEALTHCARE SYSTEM USING QUEUE ANALYSIS, FOCUSING ON PATIENT FLOW: THE CASE OF ST. PAUL'S HOSPITAL, ETHIOPIA

By:

Abel Nigussie Tefera

A Thesis Submitted to the School of Graduate Studies in Partial fulfilment of the requirements for the degree of Master of Science in Road and Transport Engineering

Approved by Board of Examiners:

Girma Gebresenbet (Professor)		25-07-2020
Advisor	Signature	Date
Bikila Teklu (Phd)		08/08/2020
Internal Examiner	Signature	Date
Techane Bosona (Phd)		03-08-2020
External Examiner	Signature	Date

UNDERTAKING

I, the undersigned, certify that this research work titled "Logistics performances of healthcare system using queue analysis, focusing on patient flow: the case of St. Paul's hospital, Ethiopia" is my original work performed under the supervision of my research advisor Professor Girma Gebresenbet and has not been presented elsewhere for assessment and for a degree in any other university. All sources of materials used for this thesis have also been duly acknowledged.

Signature:

Undertaker's Full Name: Abel Nigussie Tefera

Place of Undertaking: Uppsala, Sweden

Date: May, 2020

ACKNOWLEDGMENT

First and foremost, I would like to thank God Almighty for giving me the strength, knowledge, ability, determination and opportunity to undertake this research study and to persevere and complete it satisfactorily. Without his blessings, continuous grace and mercy with me throughout my life and ever more during the tenure of my research, this achievement would not have been possible.

Now, I would like to thank and express my deep and sincere gratitude to SLU Energy and Technology department and staffs, especially to my supervisor, Professor Girma Gebresenbet, for awarding me the scholarship to conduct this research work at Uppsala, for the patient guidance, encouragement, respected hospitality, kind and energetic approach and appreciated advice he has provided throughout my time as his student. I have been extremely lucky to have a supervisor, who cared so much about my work, whose insight and knowledge into the subject matter steered me through this research and who responded to my questions and queries so promptly. I thank you for being a father for me, while I was distant from family, which I will never have words to express!

My biggest thanks, to my family for all their support and continuous encouragement throughout my research work, without whom I would not have been able to complete this research, and would not have made it through my masters degree!

Finally, I would like to thank Mr. George William for proofreading this research paper and the St. Paul Hospital Management, physicians and staffs for their timely review of the research proposal during the issuance of ethical clearance as well for their kind cooperation and assistance during data collection.

Abstract

This study was conducted to assess the performance of patient flow at the healthcare system in Ethiopia by taking St. Paul's Hospital Millennium Medical College (SPHMMC), as a case study. As there is a very high demand for health service that exceeds the available capacity, the public healthcare centers are overwhelmed with the long queues or they are delivering the service with relatively very low consultation time. In the existing conditions, patients go as early as they can to the healthcare facilities, waiting in queue, even before the opening and had to wait long time for examination, consultation and diagnosis. However, due to high number of patients at the outpatient departments relative to the number of physicians, it results in an increased workload on the physicians and it shortens the patient consultation time, which has an impact on the patients' health. The main objective of this research was to study the logistic performances of the healthcare system using queuing analysis. This research used three key performance indicators namely, patient queue length, patient waiting time and consultation time length. The performance evaluation was conducted based on data from patients who visited 69 clinical, surgical and diagnosis departments at the outpatient clinics of the hospital. Queue analysis was performed to determine the operational characteristics using a queue scenario with Poisson arrival, exponential service, infinite population, First Comes First Served (FCFS) discipline and multiple server arrangement. The study showed that the patients' arrival rate highly exceeded the service rate, in each respective clinical department. The outpatient clinics at the SPHMMC achieved an average total waiting time of 92 minutes to get consultation and nearly 70% of the patients waited for more than 95 minutes. The consultation time was as low as 5.71 minute at the Medical clinic and 6.16 minute at the Ophthalmology clinic and around 60% of the patients saw the doctor for a time less than 10 minutes. Therefore, this research recommends addressing the gaps in human resources and logistical supplies, to implement and enforce a staggered patient scheduling and appointment system and to have serious intervention and control on the dual practice, to ensure a smooth clinic process and to reduce waiting times.

Keywords: Consultation Time, Healthcare Logistics, Outpatient, Queuing Theory, Waiting Time

Table of contents

List of tables

List of figures

Abbreviations

CT-Scan	Computerized Tomography Scan
ECG/EKG	Electrocardiogram
ECHO	Echocardiogram
ENT	Ear, Nose, Throat diseases and disorders
FDRE	Federal Democratic Republic of Ethiopia
MRI	Magnetic Resonance Imaging
OPD	Outpatient Department
OR	Operating Room
PAC	Pre-Anaesthesia Check up
PEP	Post-Exposure Prophylaxis
PITC	Provider Initiated Testing and Counselling
PMTCT	Prevention of Mother-To-Child Transmission
WHO	World Health Organization
SPHMMC	St. Paul's Hospital Millennium Medical College
SLU	Swedish University of Agricultural Sciences
TEE	Transoesophageal Echocardiography
VCT	Voluntary Counselling and Testing

1. Introduction

1.1. Background

According to the definition by WHO,

> ¨health refers to the overall state of a complete social, emotional, mental and physical well-being that considered as a resource for living a full life, and therefore, the access for the highest achievable level of health is the basic rights of every human being without the distinction of social or economic conditions¨ (WHO, 1948).

Health is not as simple as the absence of disease; however, it is the ability to get well from ailment and/or other problems.

Healthcare on the other hand, an act delivered by the health professionals, is the maintenance of health through prevention, diagnosis and treatment of mental and physical impairments. Healthcare systems exists to assure and help people, to maintain their optimal state. The improvement of health industry in a country is one of the indicators for economic development and prosperity, as it has been directly associated with the value of human resource.

Healthcare institution is a place, organization or agency, public or private that provides healthcare related facilities and services. These include medical, nursing, health screening or therapeutic services, medical drugs and medical laboratory services to the inpatients and outpatients.

Hospital logistics is one of the important infrastructures that facilitates healthcare services. Hospital logistics differentiate between three core flows: patient, material and information flows. The patient flow involves the movement of patients to get access for medical care and related resources, in the healthcare system. The material flow includes the supply of medical devices, pharmaceuticals and surgical consumables necessary to support the medical personnel and, unquestionably, patients at the hospitals. Information flows are vital to care delivery at all levels — the patient, the care team, the healthcare organization, and the surrounding political-economic environment, as the access for the patient's health record and other similar information is crucial to the physicians during diagnosis and consultation or treatments (Fanjiang, Grossman, Compton, & Reid, 2005).

A time spent at the hospitals are often the unpleasant experiences for patients and are the causes for frustration. This could arise, along with many other issues, from long waiting times for both scheduled as well as non-scheduled appointments; and delay in diagnosis and treatments. Waiting time, in this research, was the time spent by the patient in waiting before they get the desired consultation by the physician. The best service is the one that provides prompt deliveries without the need to wait for a long time, which, however, usually incur additional costs to the hospital as it requires addition of service facilities and technologies (Dachyar, Farizal, & Yafi, 2018).

These long waiting times for services and/or short consultation hours in healthcare systems can worsen the severity of disease, increase the socio-economic costs and potentially result in unnecessary suffering, straning the relationship between the physician and the patient, which leads to patient dissatisfaction (Aeenparast, Farzadi, & Maftoon, 2012; Bruni, Laupacis, Levinson, & Martin, 2010).

In developing countries, like Ethiopia, the number of trained physicians is much less than what the system demands. Thus, physicians are subjected to overloading, attending to outpatients in multiple centers and are required to show up at the inpatients in multiple hospitals. This result, the consultation hours being limited to only a short period of time, usually either mornings or afternoons. Moreover, even in top of this short period of service, regrettably, the service often becomes inactive due to late arrival of physicians and other interruptions (Babes & Sarma, 1991).

This research analysed the existing average consultation time at the outpatient department (OPD) and the extended waiting time that patients spent at the outpatient clinic, relative to the standard outpatient clinic industry average consultation time as a basis for performance measurement. An outpatient clinic is a department at the healthcare devoted to the diagnoses and consults of an outpatient (American Heritage Dictionary, 2007; Zhu, Heng, & Teow, 2009). Often however, shortening the patient processing time for various medical departments at the outpatient clinic of a hospital is needed to meet the requirements for the medical care system (Park, 2001), it is, then, crucial to measure the performance of the same.

According to FDRE Ministry of Health (2015), Ethiopia has 16,440 health posts, 3,547 health centers and 311 hospitals. Out of these figures, Addis Ababa shares about 0.59% of the health posts, 96 in total, around 25% of the health centers, 882, and 39 hospitals that count for 12.54% of the total. This study was conducted, to assess the performance of patient flow at the healthcare system in Ethiopia by taking one of the hospitals, namely, the St. Paul's Hospital Millennium Medical College (SPHMMC), as a case study.

The hospital was selected for case study due to the following reasons

- It is the second largest hospital of the country next to the Black Lion teaching and specialized hospital.
- The case study was planned to be contacted at Black Lion hospital primarily, as the hospital is under my parent university, the Addis Ababa University. Unfortunately, the hospital was not willing to cooperate in giving access to the required data.
- The case study hospital, SPHMMC, is a specialized, referral and teaching hospital with many clinical departments that incorporates different imaging and scanning laboratories too.
- The hospital had high patient flow with an observed situation of long waiting time at the OPD.
- The hospital was willing to cooperate on the study after thoroughly examining the research proposal and they had issued an ethical clearance document for the study.

SPHMMC was established by the late Emperor Haile Selassie, in 1968. The SPHMMC of Addis Ababa is the second largest hospital in Ethiopia, located in Swaziland St., at Gulele Sub city as seen on the Figure 1. According to SPHMMC

(2020), The hospital is a public, specialized, referral, and teaching hospital and has more than 2800 clinical, academic, administrative and supporting staffs to provide a healthcare service to the referral patients from every corner of the country. The hospital becomes a medical college in 2007 (Johansson, 2014) and have more than 700 beds for the inpatient clinics and served 1200 emergency and outpatient clinical patients per day, on average, during the study period.

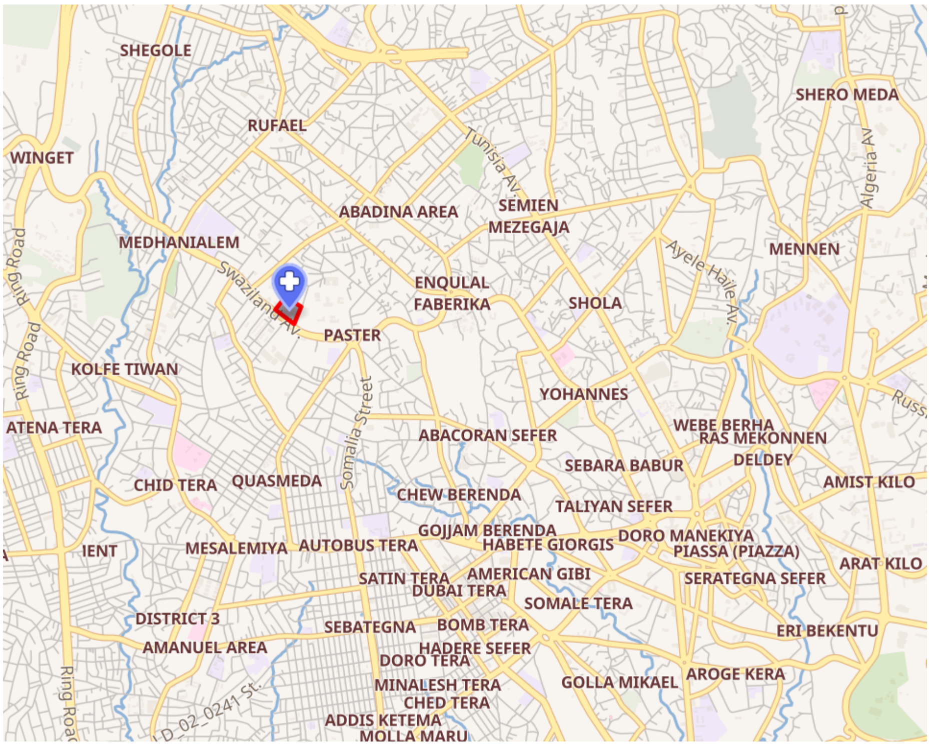

Figure 1. St. Paul hospital location map

1.2. Statement of the problem

The existed healthcare system, in Ethiopia, faces the need to re-design the service system for each clinical department by taking the needs and expectations of both the patients and professionals into consideration, whether in the private or public sector, for optimized service delivery. This optimization is necessary for fair, efficient and effective service delivery both in cost and quality aspects, and to progress further with regards to the main targets of the healthcare system (Kriegel, Jehle, Dieck, & Mallory, 2013).

To achieve the expected excellency of the operational aspect and to ensure the clinical based quality at the hospitals, it is necessary to manage the patient flow in the outpatient department (Mardiah & Basri, 2019). Bahadori, Mohammadnejhad, Ravangard, and Teymourzadeh (2014) argues that it is necessary to accurately plan and manage the outpatient departments to handle the changes and challenges in the healthcare system.

In the current system the, patients arrive as early as they can to the healthcare facility, wait in the queue, even before the facility is open, and have to wait longer for examination, consultation and diagnosis. However, due to high number of patients at the OPD relative to the number of physicians, there is an increased workload on the physicians and the patient consultation time is shortened, which has an impact on the patients' health as well as their satisfaction. Longer consultation time is an indication of better care, as it helps the physician identify the specific problem with the patients' health and detect patients with chronic cases, to prescribe fewer drugs, to promote healthier lifestyle and to deal the psychosocial problems (Petek Ster, Svab, & Zivcec Kalan, 2008; Pollock & Grime, 2003; Wilson, McDonald, Hayes, & Cooney, 1992).

Unfortunately, this is not the case for many developing countries like Ethiopia. In addition, the researcher found a scientific gap on addressing the issues of longer waiting time and, hence, shorter consultation time in the OPDs' at the healthcare system in Ethiopia. This research, therefore, seeks to bring new practical value that could help the Ethiopian healthcare system on decision making, by analysing the queuing performance of the patient flow at the case hospital in question.

1.3. Research Question

To address the issues raised in the statement of the problems, the following dominant research questions were established: -

1. How are the outpatient arrival and service pattern at SPHMMC OPD?
2. What is the level of the existing service utilization with respect to the minimum expected standard at SPHMMC OPD?
3. What advantages can we obtain with a queuing approach in managing the patient waiting times to improve the healthcare performance?
4. What measures needed to be implemented to reduce the length of queues in hospitals and increase patients' satisfaction?

1.4. Research Hypothesis

Based on the objectives of the study, the following research hypotheses were tested.

- Due to the longer waiting time at SPHMMC OPD, the consultation time is very low relative to the standard.

1.5. Research aims and objectives

The main objective of this research work was to study the logistical performances of the healthcare system using queuing analysis, Whilst, the specific objectives were to:

1. Determine the mean number of patient arrivals per hour (λ) in SPHMMC OPD.
2. Determine the mean number of patients served per hour (μ) in SPHMMC OPD.
3. Determine the average time a patient spent waiting in the queue before seen by a physician in SPHMMC OPD.
4. Analyse the waiting line of patients in SPHMMC OPD.
5. Analyse and recommend service optimization options for SPHMMC OPD.

1.6. Scope of the study

Although, this research was intended to evaluate the logistical performance of healthcare system, the research was limited to a case study analysis and focuses only on the outpatient clinics, out of many, as OPD have the greatest queueing encounter relative to the other. The scope was also limited to focus only on the patient flow logistics and did not consider both the material and information logistics flow of the healthcare system. The patient is required to pass through other parts of the hospital before joining the waiting queue at the respective physician, like triage, registration and card departments, this research didn’t consider the time the patient spent there, since it was considered to be negligible as compared to the time patients wait for consultation.

1.7. Ethical considerations

Ethical approval was requested and obtained from the hospital ethical clearance committee. The permission of the various heads of the outpatient departments was also sought as the data collection was done and all the research data endured confidential throughout and will remain after the completion of the study, too.

2. Literature review

One of the broadly used descriptions of logistics is that of The Council of Supply Chain Management Professionals (CSCMP) who describes logistics as

> “that part of the supply chain process that plans, implements, and controls the efficient flow and storage of goods, services, and related information from the point of origin to the point of consumption in order to meet customers’ requirements” (Council of Supply Chain Management Professionals, 2015; Feibert, Andersen, & Jacobsen, 2019).

Hospital logistics, as discussed by Frichi et al. (2018), addresses different healthcare quality aspects, particularly those related with patient and physicians satisfaction and standards. They also argued that a well-managed hospital logistics could improve the healthcare efficiency by reducing the patient waiting time. To support this premises they mentioned the inadequacy in resource synchronization between resources and needs as the main cause for waiting time , as it attributes to the lack of planning, coordination and communication in healthcare logistics as discussed by Melo (2012).

Service quality performance in a healthcare is the patient’s perspective based on their own experience and determined by waiting times, staff interactions, perceived medical service quality, and the overall communication (Eitel, Rudkin, Malvehy, Killeen, & Pines, 2010). According to previous studies, waiting time (Watson, Marshall, & Fosbinder, 1999) and time spent during evaluation, and consultation time, (Hobbs, Kunzman, Tandberg, & Sklar, 2000) directly affect the patient satisfaction.

In recent years, outpatient medical services are increasingly becoming the main components in the healthcare system due to the greater attention given to the preventive medical practices and shorter period of stay at the hospitals (Cayirli & Veral, 2003). The services at the outpatient clinic includes the registration, medical examinations and the received prescriptions (Chand, Moskowitz, Norris, Shade, & Willis, 2009; Dachyar et al., 2018). Time management at the OPD is vital to the patient, especially to those with bad injuries and to those who seeks diagnosis for the critical illnesses (Aziati & Hamdan, 2018; Mittal, Chatterjee, Hasnain, & Varshney, 2016).

Though it is not realized in the existing public healthcare systems, due to high patient flow towards a limited resources, treating patients in a timely manner is an entirely accepted objectives of the healthcare system (Au-Yeung, Harrison, & Knottenbelt, 2006). Bahadori et al. (2014) showed in their research that in OPD patient waiting time (Aeenparast et al., 2012; Huarng & Lee, 1996; Scanzano et al., 2005) is the main quality assurance indicator, among many others, and the waiting and service times are considered for healthcare system improvement (Scanzano et al., 2005). From the studies conducted by many researchers, Alabduljabbar, Madhi, Medan, Alaqeel, and Alsubaie (2018), identified four causes as main factors for long waiting time; High number of patients, Understaffing at the healthcare facility, long and complicated registration process and system and equipment aging.

Yaduvanshi, Sharma, and More (2019) show in their research, there have been waiting time studies on those seeks consultation at the hospitals (Park, 2001) and found that the characteristics of the healthcare providers, the characteristics of the consultation and the characteristics of the patients (Hwang, 2006), as the influential factors.

The key task in performance analysis and measurement is to quantify the gap between reality and anticipations in reference to certain standards and guidelines (Akachi & Kruk, 2017), to reveal the weak links where the efforts for improvement should concentrate (Frichi et al., 2018).

Patient Flow

Patient flow is amongst the major essentials that needs to be considered during healthcare system improvement actions. The management of patient flow in the outpatient department (OPD) requires to address the three basic aspects: the patient arrival, the queue and service processes (Mardiah & Basri, 2019).

Patient flow management is about forecasting the demand and then organizing the required resources to meet the demand. Therefore, the hospital capacity planning needs to follow the procedure of forecasting the demand and the matching the delivery of resources to the predicted demand. To arrive at reasonably accurate demand prediction there needs to be an understanding of the arrival variations throughout the courses of the day, that begins with getting the daily historical rate of incoming patients over the hours for the week (Eitel et al., 2010).

Even though, the mortality rate due to medical errors is very high, these deaths are unmeasured and the discussion regarding their prevention is limited to a number of confidential forums, like a hospital based committee, and the lesson learnt from these incidents are not distributed beyond the institution (Makary & Daniel, 2016). These researches and reports show that the healthcare system is not as safe as we all expect it to be. Yet it is found to be difficult to have a conclusive idea about the epidemiology of these errors, we can see from the existing conditions that most errors are the results of actions taken by physicians that are forced to give the healthcare assistance through excessive clinical workloads.

The work load is found to be very high for the physicians resulting in increased fatigue and more medical errors, these scenarios have likely increased with the current high economic pressure on hospitals and restrictions of physicians work hours (Michtalik, Yeh, Pronovost, & Brotman, 2013). After their research, they concluded that, there is a minimum of around 40% of reported unsafe workloads.

In addition, nearly a quarter of the hospitalists report that excessive workloads adversely affected the patient outcomes as it hinders full discussion of treatment options, while 22% of the physicians reported they ordered potentially unnecessary medical tests, consultations or procedures due to a lack of adequate time to evaluate their patient in person.

Queue theory

Queueing theory is a mathematical way of studying the waiting lines, the queues, and was developed to predict the waiting time and the queue lengths. A queueing system in the healthcare environment consists of arriving patients and one or more physicians, providing service (Eitel et al., 2010). This theory was originated in the field of research by the Danish engineer and mathematician Agner Krarup Erlang, to develop models that could describe the telephone exchange.

Queue theory is generally considered as a branch of operation research as the outputs are mostly used to make business decisions regarding to the resource allocation. Queueing models are popular among researchers since they provide reasonably accurate system performance evaluations due to their analytical nature and provision of quick answers to the “what-if” analyses (Prabakaran & Kumar, 2019).

In a healthcare system, the patient queuing effect in relation to the time spent for treatment is gradually becoming a concern to modern society (Agarwal & Singh, 2018). Queues are formed when the patient's arrivals rate exceeds the rate of service delivery (Bahadori et al., 2014). Quantitative tools, like queuing models, can help to make decisions regarding resource planning, resource utilization and scheduling, as these are all affected by the flow of patients, which the queue performance measures, such as the time spent in the system and the traffic intensity, which have direct correlations with the patient flow characteristics (Mardiah & Basri, 2019).

Bailey (1952), documents that queueing theory is valuable to make adjustments between patient waiting time and the healthcare system idle time and service utilization rates (Zonderland & Boucherie, 2012). Over the years, the variety of applications examined can be summarised into waiting time, utilization analysis, system design and problem solving (Cayirli & Veral, 2003; Creemers & Lambrecht, 2007; Fomundam & Herrmann, 2007; Lakshmi & Iyer, 2013; Luo, Wu, Gopukumar, & Zhao, 2016; Moore, 1977; Palvannan & Teow, 2012; Preater, 2002;

Wang, Guan, Koong, & Koong, 2016; D. Worthington, 1991; D. J. Worthington, 1987).

Queuing theory has increasingly become a common decision making management tool in the developed world, though it is seldom used in many African countries (Afrane & Appah, 2014), including Ethiopia. Green (2006) showed the application of queuing theory in the healthcare system, by discussing the relationship amongst delays, number of servers and utilisations using the basic M/M/s model (Lakshmi & Iyer, 2013) .

Fomundam and Herrmann (2007), in their research, showed waiting time and service utilization can be used as analytical tool in predicting the healthcare facility configurations effect on the delay of delivering services and resource utilization (Aziati & Hamdan, 2018). Their research argued that, in queueing systems, minimizing the time that patients have to wait and maximizing the utilization of physicians are contradicting issues.

3. Methodology

Queuing is an occurrence formed when peoples or things undergoes through a process of arriving at the queue, entering to the queue and waiting for a turn to get a service as they are met with delays within the service system. Depending on the number of service channels, the queue length can be limited or unlimited. To analyse the waiting time the queue line is divided into the following parts (Heizer & Render, 2008; Mardiah & Basri, 2019).

i. Arrival: This is the inputs to the system and have distinctive size and statistical distribution behaviour

ii. Discipline: This is the characteristics of the queue

iii. Service facility: There are four types of queuing model, which are single channel single phase system, multiple channel single phase system; single channel multiphase system and multiple channel multiple phase systems. This research is a multiple channel single phase system type, as the patients go to more than one server and the considers only one phase of consultation or diagnosis.

The basic notations to describe the queue analysis were created by the statistician David Kendall (Kendall, 1953). To make readers familiar with the notations, known as the Kendall Notation (A/B/C/D/E/F), they are described in brief as under.

A: Arrival Process: In this research, taken as Poisson arrival

B: Service Process: In this research, taken as exponentially distributed

C: Service Mechanism: In this research, taken as the number of physicians or beds or machines, depending on the clinic type.

D: System Capacity: The maximum number of patients in the system at any specific time. In this research, taken as to be infinite.

E: Population: In this research, taken as the total outpatients went to SPHMMC**.**

F: Queue Discipline: In this research, taken as First comes First Served

Kendall's Notation use some standard code letters for the arrival process A and the service process B. In this research, we use M to represent the Markovian arrival and service rate distributions.

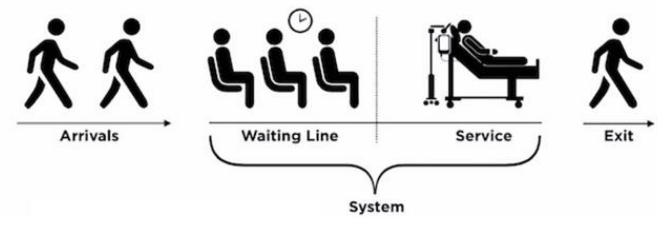

Figure 2. M/M/1 Queuing Model

The multi-server queue M/M/s is the model used most in analysing service stations with more than one server such as banks, checkout counters in stores, check-in counters in airports, and the like. The customers' arrival was assumed to follow a Poisson process, and service times were assumed to have exponential distribution. The number of servers S, were also considered to provide service independently to one another. The research also assumed that the arriving customers form a single queue and the one at the head of the waiting line entered into service as soon as a server was free. No server stayed idle as long as there were customers to be served. Note that the service rate μ was the same for all servers. Clearly, the arrival rate did not change with the number of customers in the system (i.e., λ was the constant arrival rate).

Whether there was a single doctor or there were multiple doctors on the OPD, each doctor gave service and had control for decisions regarding the process flow to his/her patients only, not to/on other doctors' patients (Yeon, Lee, & Jang, 2010). Therefore, we took the number of physicians, in each clinic at the OPD, as the number of the servers in that respective clinic. Therefore, we considered M/M/s queueing model for the analysis.

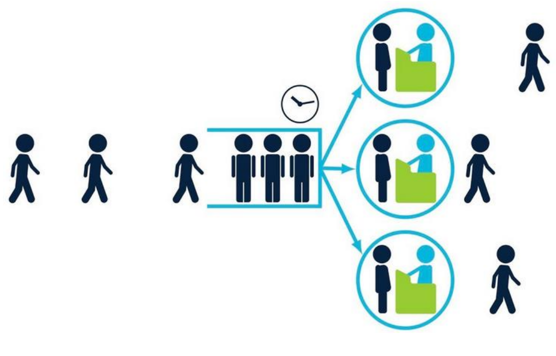

Figure 3. M/M/s Queue

The measurement scales for the queueing system included the average number of customers in the queue (Lq), the average number of customers in the entire system including those being served (Ls), the average waiting time in the queue (Wq), and the average waiting time in the entire system (Ws) (Cho, Kim, Chae, & Song, 2017) and the consultation time length. For this research, the three key performance indicators for the measurement of the hospital outpatient clinic were patient **queue length**, **patient-waiting time** and **Consultation time** length.

To conduct the performance analysis, the research focused on the two fundamental parameters, from the collected data for queueing analysis, the **arrival rate (λ)** and the **service rate (μ)**. Using these arrival rates and the service rates data, the average waiting time in a queue (Wq) were calculated. Almost all queueing models assumed an exponential distribution for the service time and Poisson distribution for the patient arrival per unit of time (Cho et al., 2017). The Poisson process distribution considers an independent and exponentially distributed probabilistic inter-arrival times for the patients that originate from the large population. An exponential distribution service delivery system uses a probability distribution that is convenient for getting systematically traceable results.

The trend in the OPD queueing at the hospital showed that patients arrived randomly and the queue capacity was assumed to be unlimited (Yin, 2010). The number of physicians at the respective outpatient clinic, the number of machines at the diagnostic imaging tests and scans centers represented the number of servers in the healthcare system. Therefore, the research assumed the same probability distributions for this paper and the patients got the healthcare service in the manner they arrived, simply in First Come First Served (FCFS) queue discipline.

Since the arrivals of the patients followed scattered nature and did not occur at a regular interval (Agarwal & Singh, 2018), this research assumed the Poisson distribution arrival. Patient arrivals were considered to be random, due to their nature that each patient's arrival was independent of one another and their occurrence couldn't be predicted exactly. In the Poisson distribution assumption, the constant λ, which is independent of any random property of the queue, was used.

Therefore, P (the arrival probability between time "t" and "t + δt") = $\lambda\ \delta t$

For a small δt interval, the probability of "n" arrivals in time "t" is defined as:

$$P_n(t)=(\lambda t)^n.e^{-\lambda t}/n! \qquad \text{Eq. 1}$$

Where: n= a random discrete variable, that obeys a Poisson-distribution, representing the number of arrivals in the given time interval

n= time at specified unit

This research paper defines terms in the calculation to be similar with the way they defined in most literatures and books. Accordingly, λ was defined to be the patient arrival rate expressed in numbers per hour, μ to be the service rate and both with the same unit. This research assumed steady state system consideration, in which the patient arrival rate, the service rate and the number of servers (Physicians/ number of beds/ number of machines, in our case) were assumed to stay constant over time. Wq was the expected time that a patient could spent in line. The expected time that the patient spent in the system, including the time spends in line and in service, was represented by Ws.

The research considered the M/M/s queue in which patients arrived to the respective clinical department at a rate λ, and each patient got service from one server among the many and waited in queue when all the servers were occupied. The maximum service rate is the product of the number of servers S and the individual server service rate μ (Zonderland & Boucherie, 2012).

The next critical challenge was to select the clinical departments that needed special focus or those that were very fundamental for the performance analysis. The hospital on the case study, SPHMMC, had 35 Medical Clinics, 20 Diagnostic Imaging Tests and Scans Centers, 4 Minor Operation Centers and 10 Major Operation Centers. Each are listed under according to their classification.

Table 1. Medical Clinics at SPHMMC

Item No.	Clinic*	Item No.	Clinic
1	Adult Emergency	19	PAC
2	Anti-natal	20	Palliative
3	Anti-rabies	21	Paediatric Emergency
4	Cervical Ca Screening	22	Paediatric OPD
5	Chest Clinic	23	PEP
6	Dental Clinic	24	Physiotherapy
7	Dermatology	25	PITC
8	Endoscopy	26	PMTCT
9	ENT	27	Post-natal
10	Family Planning	28	Psychiatry
11	Gynaecology Emergency	29	Referral Surgical
12	Haematology	30	Regular Gynaecology OPD
13	Kidney Transplant	31	Surgical

14	Medical	32	Sexual Assault
15	Nephrology	33	Staff Clinic
16	Neurology	34	Student Clinic
17	Oncology	35	VCT
18	Ophthalmology		

Table 2. Diagnostic Imaging Tests and Scans centers at SPHMMC

Item No.	Test Center *	Item No.	Test Center
1	Angiography	11	Gastroscopy
2	Angiograph screening	12	Ligation
3	Audiometry	13	MRI
4	Bronchoscopy	14	Optometry
5	Colonoscopy	15	Pathology
6	CT scan	16	Polypectomy
7	Dialysis	17	TEE
8	ECG	18	Ultrasound
9	ECHO	19	Valvotomy screening
10	Endourology	20	X-ray

Table 3. Minor Operation Centers at SPHMMC

Item No.	Minor Operation Center *
1	E & C
2	ENT
3	OR
4	Ophthalmology

Table 4. Major Operation Centers at SPHMMC

Item No.	Major Operation Center *	Item No.	Major Operation Center *
1	Constructive	6	Neuro-Surgery
2	ENT	7	Ophthalmology
3	General Surgery	8	Orthopaedic
4	Gynaecology/ Obstetrics	9	Paediatric
5	Maxillo facial	10	Urology

* Listed in Alphabetical order

Even though, the data were collected for all clinics and departmental centers listed above, it was necessary to focus on certain key departments and clinics to get a representative sample for measuring the characteristics performances. Therefore, the Pareto principle was applied, to screen and select clinics and departments from each of the categories tabled above.

The Pareto´s principle is a concept that was first applied in economics and then became one of the governing rules in different areas. Named after Vilfredo Pareto, the 19th century Italian economist, the principle is as: 80 percent of effects always come from 20 percent of the causes. Since he published these findings, the magical ratio of 80:20 (or the "80-20 rule") has been applied in many of our day-to-day encounters. Accordingly, the principle could be interpreted as, 20 percent of the products covers 80 percent the company's profits; 20 percent of the roads accommodate 80 percent of the traffic; and 80 percent of food production is generated from 20 percent of the crops. This is why the 80-20 phenomenon is the often cited as a universal baseline for most of the distributions.

With regard to this research, the consideration was that 80 percent of the patients flows into 20 percent of the clinics and we can apply this into each of the categories. From this consideration and calculation, the busiest clinical departments, the vital few, which aim to give consultation for the trivial many were sorted out. The Pareto analysis used to categorize data and to identify which clinical departments´ processes have the most effect on a quality of care outcome of the hospital. Since it is critical to devote efforts in the vital few clinical departments, it is valuable to use

Pareto analysis when there are many clinical departments that contributes to the hospital performance.

Therefore, the following 13 clinical departments, 4 Diagnostic Imaging Tests and Scans centers, 4 Minor Operation centers and 6 Major Operation centers were selected for performance analysis based on their respective patient flow numbers recorded for eighteen months.

Table 5. Medical Clinics selected for analysis at SPHMMC

Item No.	Clinic	Percent share of the patient flow
1	Medical	17.59
2	Surgical	12.50
3	Ophthalmology	8.91
4	Paediatric OPD	5.76
5	Regular Gynaecology OPD	5.74
6	Anti-natal	5.17
7	Psychiatry	4.37
8	ENT	3.60
9	Adult Emergency	3.58
10	Neurology	3.11
11	Dental Clinic	3.08
12	Gynaecology Emergency	2.98
13	Dermatology	2.75

Table 6. Diagnostic Test centers selected for analysis at SPHMMC

Item No.	Test Center	Percent share of the patient flow
1	X-ray	33.05
2	Ultrasound	31.74
3	MRI	9.62
4	CT scan	7.86

Table 7. Minor Operation centers selected for analysis at SPHMMC

Item No.	Minor Operation Center	Percent share of the patient flow
1	OR	42.87

2	Ophthalmology	31.23
3	MRI	23.18
4	E & C	7.86

Table 8. Major Operation centers selected for analysis at SPHMMC

Item No.	Major Operation Center	Percent share of the patient flow
1	General Surgery	23.98
2	Ophthalmology	21.12
3	Orthopaedic	16.54
4	Urology	7.65
5	Neuro-Surgery	7.48
6	Paediatric	5.59

Population

The study population, for this research, were all outpatients treated and consulted at the SPHMMC.

Samples

Sampling is the process of selecting a predetermined number of observations for further statistical analysis to draw conclusions applicable to the larger population. To determine the sample of our study, we had used the following formula:

$$n = \frac{z^2 p(1-p)}{E^2} = (\frac{z_c}{E}\sigma)^2 \quad \text{Eq 1}$$

Where:

Z = Z-score (1.96 for 95% confidence level)

p = Healthcare service utilization rate (as proportion)

E = Margin of Error (Confidence Interval); assumed to be 5%

Zc = cumulative z value

σ = Sample standard deviation

NB: The above sample size calculation provided us, the recommended number samples required to estimate the true proportion mean with the required margin of error and confidence level. We took a 50% sample proportion out of the total population and 100,000 outpatients as a total population, as recommended by statistical books when there is uncertainty on the exact population and likely sample proportion.

The sample size for each clinical department was calculated using the above formula, and a minimum of 384 hours of outpatients' secondary data were selected to determine the arrival rate, service rate and other related data to perform the patients flow and queuing network performance variables.

Table 9. Sample size calculations

Z_c	p	E	$\boldsymbol{\sigma}$	$\boldsymbol{n}$
1.96	0.5	5%	100,000	384

Data collection

Initially the secondary data that were required to analyse the patients' arrival patterns and their average waiting times for treatment, were collected from November 2019 to January 2020, from all the medical clinics in the hospital at the case study. Patients queue data, both new and repeated, who visited all the 69 clinical, surgical and diagnosis departments at the outpatient clinics of the hospital, the last 18 months from October 2019, were analysed to evaluate the hospital performance. After the initial diagnosis and treatment in the first visit, if the patient returns back to the hospital for check-up, having symptoms of sickness, having complication from previous treatments, or any other related issues, then from the next arrival onwards that patient is categorized as repeated or follow-up patient.

In addition, the number of physicians in each clinical department of the hospital were collected. As it is in most healthcare settings, with no appointment system in place, the existing system in the hospital was in such a way that the physicians met, for the medical procedure, with both the new and repeated patients in First Comes First Served (FCFS) basis. On the existed conditions, follow up patients are appointed to come for check-up or repeated consultation on scheduled date rather than specific time-based appointment on that day. The data was analysed using STATA statistical software and Microsoft Excel Spreadsheet.

Calculation of arrival rate

The arrival rate was calculated as the average number of arrivals per unit time. The calculation was performed by dividing the total number of customers, in this research case patients, per day to the total working hours per day, usually eight hours, unless stated to be different based on the special characteristics of the clinic in analysis. For the early arrivals, the earliest time that the facility opened for the waiting and the average number of patients during service opening hour was considered. This research was limited to use this approach due to the unavailability of patient arrival data collected or registered on the system, to record the arrival time of the patient.

If the arrival time registry had done, we could simply take the ratio of the difference between the first patient's arrival time and the last patient's arrival time to the inter arrival time to be the number of patients arrived during that time interval. In this paper, for the sake of simplicity, an assumption of no baulking (no patients refused to join the waiting line because it was too long), no reneging (no patients left the waiting line due to long waiting for service), or no jockeying (no patients switch between waiting lines by perceiving they will get served faster by so doing) taken into consideration.

Though the arrival pattern of patients who are looking for healthcare services, at the hospital, is always random in nature, they demand to get immediate services by that time. If the hospital service facility is working at its peak capacity when the patients arrive, they are required to wait with patience until their turn. This is where a queue, the number of patients waiting to be served, is formed due to the difference between patient arrival and the time taken for the service delivery. Due to the medical services special characteristics, the exact prediction of patients arrival and the time it will take for consultation is very difficult (Cho et al., 2017). Thus, the absolute objective of queueing theory is to attain an economic equilibrium between the service cost and the patients' waiting time (ibid.).

Calculation of service rate

In most developing countries hospitals, including Ethiopia, digital data recording system that includes Patients consultation durations was not found to be applicable. Therefore, it was not possible to get the consultation hour and the service rate, for each patient or for whole patients arrived and consulted per day. Therefore, the calculation of average service rate was done by dividing the total serviced patients per day to the total working hour per day. The effective consultation hour was then calculated by considering the average time lost during transition, taken by interviewing each clinic physicians, from this patient to the next patient.

Consulting time varies from country to country and even from hospital to hospital, as it depends on the patient and physician characteristics (Ahmad, Khairatul, & Farnaza, 2017). The American Academy of Family Physicians, after the survey conducted in 2013, a physician spends nearly 93.2 hours weekly, in healthcare activities, which is the equivalent of consulting 19 patients per day and a family physician spends around 22 minutes per encounter with a patient (Bernstein, 2014). Therefore, an assumption of 19 patients consulted per day was taken as a reference for service rate calculation. For emergency service clinics 15 minutes of average consultation time as recommended by Samuel, Aldeen, Gravenor, and Malik (2015) was taken. For gynaecology and obstetrics emergency OPD an assumption of 390 minutes was taken as the average treatment and diagnosis time as discussed by Ocak, Bekdas, Duran, Göksügür, and Küçükbayrak (2013). All physicians were assumed as punctual at their duty station. In addition, the consultation time difference among physicians was considered as insignificant, hence similar consultation time between patients.

The number of servers

This research paper assumed the number of physicians in the specified clinic, the number of beds in surgical and delivery wards, and the number of machines in the testing and diagnosis labs to be the number of servers as these are the main constraints throughput, respectively.

The queue system

Patients arriving at the hospital for outpatient treatment are expected to be assessed for the kind of diagnosis or consultation they need before going to get their treatment card. This assessment is carried out at the triage of the respective clinic. Triage is where the process of assessing and categorizing of patients according to their medical care demand, irrespective of their sex, age, socioeconomic condition and other similar factors (Qureshi, 2010). The basis behind this triage intervention is analogous to the ‘front-loading’ effect of the product development process. Front-loading is the art of bringing the problem-identification and problem-solving in the product developments to the earlier stages (Thomke & Fujimoto, 2000).

Once the patient got assessed at the triage s/he is then expected to get the treatment card from the central system. Newly arriving patients apply for a new card and get the same along with a new medical record filing document while repeated patients get their medical record file by showing their existing card number. Once patients are done with the process of getting the medical record file, they will wait for their turn for consultation or examination, at the respective clinic. This process is described diagrammatically as follows.

Figure 4. Patient Flow Process at the OPD

Data Processing and Analysis

For data analysis and interpretation of the results, computer software, namely Microsoft Excel and STATA Statistical software were used. The collected secondary data was entered in Excel Spreadsheets to organize, sort and compile and then exported to STATA for descriptive statistical analysis to be carried out. The figures and tables, from both applications, were interpreted accordingly to arrive at meaningful conclusions and recommendations.

The queue system performance parameters used in this research were defined as follows:

λ: Arrival rate of patients at the subjected OPD per hour;

μ: Service rate of the subjected OPD per hour;

s: Number of servers in the respected OPD for consultation.

S Russell, W Taylor, Castillo, and Vidyarthi (2011) discusses the formulas for determining the operational characteristics of a queue scenario with Poisson arrival, exponential service, infinite population, First Comes First Served (FCFS) discipline and multiple server as follows

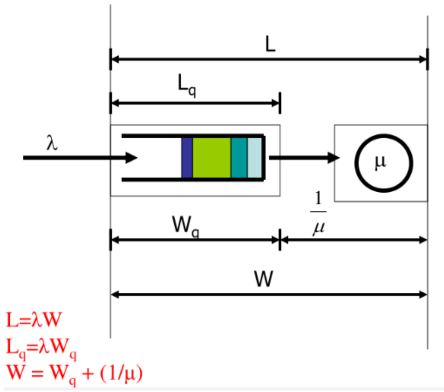

Figure 5. Little's law

The probability of no patients in the system (all physicians are idle) is

$$P_0 = \frac{1}{\left[\sum_{n=0}^{n=s-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n\right] + \frac{1}{s!}\left(\frac{\lambda}{\mu}\right)^s \left(\frac{s\mu}{s\mu-\lambda}\right)}$$ ……………………. Eq 2

The probability of n customers in the queuing system is

$$P_n = \begin{cases} \frac{1}{s!s^{n-s}}\left(\frac{\lambda}{\mu}\right)^n P_0, \text{ for } n > s \\ \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n P_0, \text{ for } n \le s \end{cases}$$ …………...………… Eq 3

The probability that a customer arriving in the system must wait for service (i.e., the probability that all the servers are busy) is

$$P_W = \frac{1}{s!}\left(\frac{\lambda}{\mu}\right)^s \frac{s\mu}{s\mu-\lambda} P_0$$ …………...………… Eq 4

Avg. no. of patients in system

$$L = \frac{\lambda\mu(\lambda/\mu)^s}{(s-1)!(s\mu-\lambda)^2} P_0 + \frac{\lambda}{\mu}$$ …………...………… Eq 5

Avg. time patients spend in system

$$W = \frac{L}{\lambda} \qquad \text{……………...…………. Eq 6}$$

Avg. no. of patients in queue

$$L_q = L - \frac{\lambda}{\mu} \qquad \text{……………...…………. Eq 7}$$

Avg. time patients spend in queue

$$W_q = W - \frac{1}{\mu} = \frac{L_q}{\lambda} \qquad \text{……………...…………. Eq 8}$$

Service Utilization

$$\rho = \frac{\lambda}{s\mu} \qquad \text{…………….....…………. Eq 9}$$

Queue Intensity (measure of how busy a system is)

$$\rho = \lambda/\mu \qquad \text{…………....……..……. Eq 10}$$

In the existing system clinics technically limits the number of patients to be served by considering the remaining possible consultation time and this resembled the finite population queue system. However, this analysis considered the patients population to be infinite by considering the regulation that forces to give medical care to all patients that arrived at the clinic before the official business closure time. In the mean time it was difficult to get data that showed the exact time the clinic prohibits patients from joining the system and the personals at the clinic denied these incidents as it could lead to administrational penalties.

In the finite population model the patient arrival rate depends on the number of patients being served and waiting. While in the infinite population model the patient arrival rate is not affected by the number of patients being served and waiting. Taking this assumption also helped to give optimization options or alternatives at the worst-case scenario.

4. Results and Discussions

4.1. Results

For the selected medical clinics at SPHMMC the analysis result is tabulated as follows from table 10 up to table 14.

In Table 10, the queue performances are calculated at each selected thirteen clinical departments. The total working hour in the week is calculated by multiplying the number of working days in a week by the total working hour in a day for each specific clinic. The average number of patients per day was calculated by taking the total number of patients during the sampling hours and dividing it with the total number of days that maintains this sampling hour. The number of servers was directly brought from the data and represented the available number of physicians in each respective clinic except for surgical and Gynaecology Emergency departments, for both the number of available patient beds considered.

Therefore, the average arrival rate for patients, λ was calculated by dividing the total number of patients in the day to the total hour of patient flow. The total hour of patient flow includes the earliest time the patient goes to the clinic and started to wait until the very last minute of the working time or the time at which the clinic closed its gate to the new comers of the clinic, whichever happened, including the lunchtime at which the service ceased to operate. Even though the ministry of health's administration procedure was to give service for anyone who showed at the clinic at the working hours, some clinics technically prohibit patients from getting card for the consultation of examination before one to two hours.

The average service rate was calculated by dividing the average number of patients per day to the product of the total working hour of the clinic to the number of servers in the clinic. The remaining queue performance indicators were calculated based on the equations listed in the methodology, data processing and analysis.

Table 10. Queue performance at clinical departments

Clinical Department	Total working hour in a week	Average number of patients per day	Number of Servers in the clinic, s, (physicians or as stated)	Average patient's arrival rate, λ (hr^{-1})	Average service rate hr^{-1}, per server (μ)	Average number of patients in the system (L)	Average number of patients in Queue (Lq)	Average Time Spent, in min, in System (W)	Average time waiting, in min, in line (Wq)
Medical	48	351	6	35. 10	7.31	68.13	63.33	116.46	108.25
Ophthalmology	40	189	4	18.90	5.91	41.13	37.57	117.50	107.35
Surgical	40	78	9 beds	11.14	1.08	23.40	13.11	126.00	70.61
Paediatric OPD	40	98	4	9.80	3.06	17.21	14.01	105.39	85.80
Regular Gynaecology OPD	40	138	4	11.50	4.31	26.84	24.18	140.05	126.14
Anti-natal	40	132	8	14.67	2.06	21.03	13.92	86.02	56.93
Psychiatry	40	72	2	10.29	4.50	7.20	4.91	42.00	28.66
ENT	40	76	2	7.60	4.75	16.16	14.56	127.54	114.91
Adult Emergency	168	40	10	2.50	1.2	2.00	-	48.01	-
Neurology	24	134	3	13.40	5.58	24.21	21.81	108.41	97.66
Dental Clinic	40	65	7	6.5	1.16	14.31	8.71	132.06	80.37
Gynaecology Emergency	168	13	7 beds	0.81	0.67	0.65	-	48.00	-
Dermatology	40	62	2	6.89	3.88	5.70	3.93	49.67	34.19

For table 11, the existing average service time was calculated by dividing the total daily working time by the number of patients and considering the time lost between patient transitions. The time lost between patients’ transition was considered based on the data collected by interviewing the physicians in each clinical department. The average standard service time was taken as recommended by Bernstein (2014). The system stability and utilization are parameters to evaluate the overall system capacity and efficiency. Stable systems were those with an existing service time of equal to or greater than the average standard service time and unstable otherwise. Utilization describes the extent to which the existing service time varies with the average standard service time. If they are somewhat reasonably close to each other, in a stable system, then the system was well utilized. If the system were stable and the existing service time was much greater than the average standard service time, the system was considered as underutilized.

Table 12 was calculated in a similar approach to Table 10 and the only difference to the approach was that the test centers was considered in place of the clinical departments.

In Table 13, the minor operation center was considered as a single department even though there were many operational departments underneath due to the fact that all those operational departments use the same and single operational room. This scenario applied to the major operation centers too, in Table 14.

Table 11. Service time performance at selected clinical departments

Clinical Department	Existing Average service time in minute	Average standard service time in minute	System stability and utilization
Medical	5.71	22.00	Unstable
Ophthalmology	6.16		Unstable
Surgical	35.38		Stable, under utilized
Paediatric OPD	16.59		Unstable
Regular Gynaecology OPD	10.91		Unstable
Anti-natal	25.59		Stable, well utilized
Psychiatry	10.83		Unstable
ENT	9.63		Unstable
Adult Emergency	44.26	15.00	Stable, under utilized
Neurology	8.25	22.00	Unstable
Dental Clinic	47.69	22.00	Stable, under utilized
Gynaecology Emergency	46.50	15.00	Stable, under utilized
Dermatology	11.48	22.00	Unstable

Table 12. Queue Performance at selected test centers

Test Center	Average number of patients per day	Number of machines in the clinic (s)	Average patient's arrival time, λ (hr^{-1})	Average service rate, hr^{-1}, per server (μ)	Average number of patients in the system (L)	Average patients in Queue (Lq)	Average Time Spent, in min, in System (W)	Average time, in min, waiting in line (Wq)
X-ray	98	2	4.45	2.08	4.90	2.72	65.97	36.58
Ultrasound	76	6	8.44	1.58	4.17	-	29.64	-
MRI	137	1	13.70	17.13	10.50	9.70	45.99	42.48
CT scan	104	1	10.40	13.00	17.60	16.80	101.54	96.92

Table 13. Queue Performance at Minor Operation Center

<table>
<tr><th>Minor Operation Center</th><th>Average number of patients per day</th><th>Number of beds in the clinic (s)</th><th>Average patient's arrival time, λ (hr^{-1})</th><th>Average service rate, hr^{-1}, per server (μ)</th><th>Average number of patients in the system (L)</th><th>Average patients in Queue (Lq)</th><th>Average Time Spent, in hr, in System (W)</th><th>Average time, in hr, waiting in line (Wq)</th></tr>
<tr><td>OR</td><td rowspan="4">19</td><td rowspan="4">4</td><td rowspan="4">2.71</td><td rowspan="4">0.59</td><td rowspan="4">8.46</td><td rowspan="4">3.89</td><td rowspan="4">187.09</td><td rowspan="4">86.04</td></tr>
<tr><td>Ophthalmology</td></tr>
<tr><td>ENT</td></tr>
<tr><td>E & C</td></tr>
</table>

Table 14. Queue Performance at Major operation centers

Major Operation Center	Average number of patients per day	Number of beds in the clinic (s)	Average patient's arrival time, λ (hr^{-1})	Average service rate per server (μ)	Average number of patients in the system (L)	Average patients in Queue (Lq)	Average Time Spent, in min, in System (W)	Average time, in min, waiting in line (Wq)
General Surgery								
Urology	19	4	2.71	0.59	8.46	3.89	187.18	86.04
Paediatric								
Ophthalmology								
Orthopaedic	63	8	9.00	0.98	31.50	22.35	209.97	149.02
Neuro-Surgery								

4.2. Discussions

Although waiting time is a worldwide phenomenon that affects healthcare organisations throughout the globe (Ghazali et al., 2011), in Ethiopia there is still much to be done in order to reduce patient waiting time in public hospitals. On this study, it is clearly shown that the patients' arrival rate highly exceeds the service rate, in each respective clinical department. On average, there was about 351 patients in the medical clinic with an average arrival rate of 35.10 patients per hour, which is the highest among the clinics. The lowest rate of arrivals was in the Gynaecology and Adult emergency clinics and this could be why there was no overcrowding and serious performance issues within these clinics. If these were not the case, a crowded or poorly structures emergency clinic could be dangerous.

Arrival time pattern of the patients showed that around 15 to 25% patients in most clinical departments and around 50% of the patients at the surgical clinics comes as early as 6 am in the morning and waited there until the clinic opened for service. However, the hospital administration argued as they gave service to anyone arrived during the working time; practically the clinics stopped to accept patients as early as one hour before the closure of the clinic and consulted only those who were already on the queue.

This could be because the public healthcare in Ethiopia is in a state of excess demand, where the demand for subsidised healthcare far outstrips the supply; given the fact that public healthcare in the country is almost free at the point of delivery.

For instance, SPHMMC hospital provided services free of any sums of charges to about 75% of its patients to help the economically under privileged populations (Gizaw & Gebremedhin, 2011; Johansson, 2014). Furthermore, the huge fee differential between public and private healthcare also contributes to this kind of unbalanced demand. Seeking the healthcare at the private sectors, commonly not available for complicated cases, entails much higher payments, which is not, at almost all times, within the economic means of average citizens.

Regarding the service capacity per hour, the research showed there was as high as 7.31 patients per hour load at the medical clinic and 5.91 patients per hour at the ophthalmology clinic. While as expected, it is very low, 0.67 patients per hour in the gynaecology emergency and this would give enough time for the consultation. This very low service capacity per hour was expected on the emergency departments because they were projected to operate on relatively enough number of servers for the worst-case scenarios as most of the emergency cases could be cases that leaded to fatalities. The service capacity was higher at the medical and ophthalmology clinics due to a relatively higher number of patients with respect to the number of physicians. As each physician was expected to cover many patients per unit time to cover the high number of patients with relatively lower stuff numbers, they had higher service rate.

A study conducted in 2011 at the outpatient clinics in Malaysia showed that the average waiting time to get consultation was 60 minutes (Ahmad et al., 2017; Ghazali et al., 2011) and according to Ghazali et al. (2011) it could go as high as 64 minutes. A study conducted in India showed that the average waiting time for consultation ranges from the minimum 7.5 minutes at the paediatric OPD to a maximum 61 minutes at the cardiac OPD (Yaduvanshi et al., 2019). In Nigeria, the average waiting time was found to be 73 minutes as per the study conducted in University College Hospital Ibadan (Bamgboye, Erinoso, & Ogunlesi, 1992). In comparison, the outpatient clinics at the SPHMMC achieved an average total waiting time of 92 minutes to be consulted and nearly around 70% of the patients waited more than 95 minutes. This indicates that, as we hypothesised, a high percentage of patients were in an extended long waiting time to see the physician.

Meanwhile, the available consultation time per patient in the emergency OPD departments was more than the placed average standard consultation duration, 15 minutes, as it was found to be 44.26 minutes in the Adult Emergency and 46.50 minutes in the Gynaecology Emergency. However, the consultation time, in most of the clinics, as hypothesised in this paper, was found to be lower than the 22-minute consultation time specified as standard. The average consultation time was as low as 5.71 minute at the Medical clinic and 6.16 minute at the Ophthalmology clinic, while the average in the OPD clinics was 15.98 minutes and around 60% of the patients saw the doctor for a time less than 10 minutes. Nevertheless, the Anti-natal care clinic was found to be the only clinic that meets the standard, which even had a higher consultation time.

In the current healthcare environment, patient waiting time is the main aspect in performance measurement. Patient satisfaction is becoming an increasingly important variable in healthcare quality assessment and it’s difficult to satisfy patients with longer waiting times (M Oche & Umar, 2011; Mackey & Cole, 1997).

A satisfied patient believes that the healthcare facility do understand his/her needs and demands (Ny Net & Chompikul, 2007).

This research identified that this long waiting time and hence shorter consultation time was due to long queues caused by an early arrival of patients', even way before opening of the clinic, for the service and extended arrival of patients during the lunch break time, hence arriving at the same time in large numbers. The imbalanced number of patients in relation to the available number of physicians or diagnostic equipment, according to their respective clinic, was the other reason for longer waiting time and shorter consultation duration. Addissie, Jira, and Bekele (1998) recommends to avoid traffic jams and early arrival of patients, as it could lead to creation of large batches of patients at the opening of the system.

Lack of proper staffing, low deployment, in quantity, of physicians in the clinics and beds and equipment in the diagnosis and surgical clinics, is a known contributor to lengthy waiting time (Ahmad et al., 2017; MO Oche & H Adamu, 2013) and shorter consultation duration. Ethiopia is a country with the lowest physician density that arises from the serious shortages of healthcare professionals (Johansson, 2014). This claim is supported with the fact that, though, WHO recommends doctor to population ratio to be one per 1000, and according to the reports Ethiopia have less than one per 36,000 (Abera, Alemayehu, & Herrin, 2017). This is due to high physician attrition rates, rapid population growth, low physicians production rate, increased post-graduate enrolment (Yifru Berhan, 2008; Health, 2010) and migration of professionals (Johansson, 2014; Melka, 2016; Yigzaw, 2009). Y. Berhan (2008) discussed the migration of physicians to oversees or to joint NGOs/ Private Sectors was as high as 73 %, as from the year 1987 to 2006 (Johansson, 2014).

Many patients early arrive early before service opening which was another factor. According to the research data, from 15% up to 50% of the patients in most clinics arrived early before the opening of the respective clinic. The dynamic nature of service operations at the hospitals makes it difficult to make predictions on the number of patients arrivals at any moment and to assure the immediate availability of the service (Yaduvanshi et al., 2019). Ahmad et al. (2017) argues that even despite the warnings and the reminders given to the patients to arrive at the scheduled time, early arrivals were common and have significant impact on waiting time length (Atnafu, Haile Mariam, Wong, Awoke, & Wondimeneh, 2015; Su & Shih, 2003).

Physician dual practice was the other factor for the shorter consultation and longer waiting times. Dual practice refers to full-time salaried physicians at the public sector practicing simultaneously in the private firms and clinics (Ferrinho & Van Lerberghe, 2004; Socha & Bech, 2011). Public-on-private dual practice is like physician retention mechanism at the public sector, particularly in low- and middle-income countries and is often upheld as a means to complement a lower government salary grade (Abera et al., 2017; Bir & Eggleston, 2003; Ferrinho et al., 1998; Macq, Ferrinho, De Brouwere, & Van Lerberghe, 2001; Russo, McPake, Fronteira, & Ferrinho, 2014).

Unfortunately, this practice reduced the current quality of healthcare at the public sector as the physicians were diverting their attention and resources to the private and preferred to spend less time around the public hospitals (Abera et al., 2017; Bir & Eggleston, 2003; Ferrinho & Van Lerberghe, 2004; Kuhn & Nuscheler, 2013; Socha & Bech, 2011). Many patients also accused their Physicians of diverting them to public clinics where they worked or to the private medical scheam that was conducted on the hospital after 4:00 pm, either by direct referral or by more subtle acts, like manipulating increased public sector waiting times in order to stimulate demand for their those services.

Agarwal and Singh (2018), argued as unnecessary waste of time in the healthcare centres may lead to patients' health complications that could worsen the cases to eventual death, which could be avoided. Among the several alternative ways, towards improving the situations in patient flow, and thus reducing waiting time for the patients, they recommend adjusting the system by increasing the number of servers and managing the arrival rate.

The number of servers can be increased by deploying more physicians, beds and equipment, in each respective clinic, accordingly, which shall present an immediate improvement in healthcare services, and shall eliminate the crowd in the waiting-room, and thus increase the service efficiency. The arrival rate can be controlled by arranging a schedule time for the follow-up patients to come at the non-peak hours preferably at the afternoon. Implementing a thicket numbering will help the patients to be aware of the number of patients ahead of them and the time they will wait before the service.

The findings of this research study indicated that for a patient to see a physician for a consultation time a lower than 10 minutes, he or she had to wait for about two hour. Atnafu et al. (2015) discussed that, at the many healthcare facilities in Ethiopia as well as in other countries, waiting time could exceed two hours (Addissie et al., 1998; Bachmann & Barron, 1997; Mahomed & Bachmann, 1998), and this result was coherent with the findings of this research.

If we took into account the traveling time to and from the hospital, getting access to healthcare within the system would actually be a daunting task. Though, patients prefer to have longer time with physicians, many research works like Deveugele, Derese, van den Brink-Muinen, Bensing, and De Maeseneer (2002), and Ogden et al. (2004) argued there are no universal guideline in placing the best length of time for consultation. Ghazali et al. (2011) showed in their works that, the average consultation time in Malaysia was 15 minutes (Ahmad et al., 2017).

Consultation time is dependent on individual patients' characteristics and their illnesses level, type of visit, physician's characteristics and workload. However, Slower and longer consultation with physicians is expected to identify psychosocial problems, more accurate exploration and presenting of complaints, prescribe less and offer more advices towards preventative cares and physicians that consult for a time less than 7 minutes, likely miss most of these points (Wilson & Childs, 2002).

Observational evidences suggest that the consultation time is directly interrelated with the service quality (Chen, Farwell, & Jha, 2009; Mechanic, McAlpine, & Rosenthal, 2001). Though, patients prefer longer consultation as discussed above, the same literatures debated the conflicting objective of reduction on the patients waiting times and trade-off between speed and quality becomes a main concern (Srivatsa Srinivas & Marathe, 2020). According to Anand, Paç, and Veeraraghavan

(2011) queuing theory have been implemented in the past to design such systems and to get the optimal speed-quality trade-off.

Accordingly, there needs to have a proper assignment in the number of physicians, beds and equipment to cope with the tasks by considering, speed-quality trade-off. These recommendations considered an acceptable emergency room wait time of 15 min (Ocak et al., 2013), and 30 minutes of waiting time as recognized by the institute of medicine (O'Malley, Fletcher, Fletcher, & Earp, 1983; M. Oche & H. Adamu, 2013). For the test and operation centers, the numbers of servers were calculated by assuming a 30 minutes average waiting time only, since the service time was dependent on the test and surgical procedure type.

For clinical departments where it found to be impractical to assign the required number of servers to satisfy the minimum requirement for the OPD standard consultation time, the recommendation was done in a stage-by-stage improvement on the sector. Thus, the approach to the solution was using two alternatives, by taking the maximum and minimum time from the average consultation time found from other studies which is 10-15 minutes (Ahmad et al., 2017; Britt, Valenti, & Miller, 2002; Cape, 2002).

5. Conclusions and Recommendations

5.1. Conclusions

Findings from this study have showed that nearly three-quarter of the patients waited for more than 90 minutes and there were shorter consultation times, relative to the OPD standard, in most of the clinical departments. Therefore, this research concludes the following key points as main contributing factors:

- The imbalanced number of patients with respect to the available number of physicians or diagnostic equipment,
- Patient flow jams and early arrival of patients even way before the service opening,
- Physicians dual practice,
- Physicians absenteeism, late coming and/or early leaving
- Lack of proper scheduling and appointment systems

5.2. Recommendations

To make the healthcare system more accessible and effective, there is an urgent need towards addressing the existing significant gaps in the human resources, equipment and logistics so as to efficiently minimize the existing long waiting and short consultation times, to ensure effective, efficient and operational healthcare delivery system.

As the current operating philosophy in the healthcare system is “First Comes First Served” and unless the healthcare places an appointment system to implement a patient scheduling scheme, patients believe that a physician can see them earlier if they come earlier. Implementing and enforcing a staggered appointment system for patients and improving the clinics patient scheduling system, will ensure a smooth clinic process and will reduce the waiting times. This scheduling shall also consider implementing specific time-based appointments for the follow-up patients, in place of the existing day-based appointment. This intervention is expected to be cost effective and for its long-term efficiency there needs to be a close follow-up in its implementations.

Regarding to the number of servers, the following recommendations were made to reform the number of physicians, beds and equipment to cope with the tasks. These recommendations considered an acceptable waiting and consultation times as discussed earlier.

Table 15. Recommended number of servers at clinical departments by assuming 22 min consultation time

Clinical Department	**Average number of patients per day**	**Existing number of servers in the clinic, s, (physicians or as stated)**	**Existing average service rate, hr^{-1}, per server (μ)**	**Recommended number of servers in the clinic, s, (physicians or as stated)**	**Change in number of servers in the clinic**
Medical	351	6	7.31	18	+12
Ophthalmology	189	4	5.91	11	+7
Surgical	78	9 beds	1.08	6	-3
Paediatric OPD	98	4	3.06	5	+1
Regular Gynaecology OPD	138	4	4.31	7	+3
Psychiatry	72	2	4.50	4	+2
ENT	76	2	4.75	4	+2
Adult Emergency	40	10	1.2	5	-5
Neurology	134	3	5.58	7	+4
Dental Clinic	65	7	1.16	4	-3
Gynaecology Emergency	13	7 beds	0.67	5 beds	-2
Dermatology	62	2	3.88	4	+2

Table 16. Expected service improvement at clinical departments with targeted 22 minutes consultation time

Clinical Department	Improved average service rate, hr^{-1}, per server (μ)	Average number of patients in the system (L)	Average patients in Queue (Lq)	Average Time Spent, in min, in System (W)	Average time waiting, in min, in line (Wq)	Average consultation time in minute
Medical	2.44	32.23	17.83	55.09	30.48	22.12
Ophthalmology	2.15	18.29	9.49	58.07	30.13	23.94
Surgical	1.63	8.73	3.93	67.17	30.25	21.92
Paediatric OPD	2.45	8.96	4.96	54.83	30.34	20.99
Regular Gynaecology OPD	2.46	12.68	7.08	55.15	30.88	21.35
Psychiatry	2.25	6.69	3.49	55.72	29.05	24.17
ENT	2.38	7.19	3.99	56.77	31.50	22.26
Adult Emergency	0.33	3.62	-	150.21	-	171.09
Neurology	2.39	12.53	6.93	56.10	31.02	22.07
Dental Clinic	2.03	6.52	3.32	60.18	30.64	25.54
Gynaecology Emergency	0.34	2.32	-	185.07	-	204.30
Dermatology	1.94	6.89	3.34	60.03	29.06	25.97

Evaluating from the standard consultation time perspective, the above tables shows that there were over assignments to the surgical (3 excess beds), Adult Emergency (5 excess physicians), Dental (3 excess physicians) and Gynaecology Emergency clinics (2 excess beds). Therefore, the hospital shall adjust accordingly by conducting other related assessments to efficiently utilize the very scarce resource of the country. However, major improvement was needed on the Medical clinic (12 additional physicians required), Ophthalmology (7 additional physicians required), Neurology (4 additional physicians required) and Regular Gynaecology (3 additional physicians required). Nevertheless, deployment of this much physicians per hospital is almost impractical, considering the scarcity of the physicians in the country.

As Ethiopia is ranked as 180th in the WHO ranking, out of 191 countries, of health systems, which was a lower rank than its neighbours like from Somalia (179th), Eretria (158th), Djibouti (157th), Kenya (140th), Sudan (134th) it is difficult to expect a consultation time that is closer to the expected standard consultation time. Therefore, we need to have another reference for a stage-by-stage improvement on the sector.

First, let us consider 15 minutes of consultation time as a reference and by excluding clinics those already meet the standard consultation time.

Table 17. Recommended number of servers at clinical departments by assuming 15 min consultation time

Clinical Department	**Average number of patients per day**	**Existing number of servers in the clinic, s, (physicians or as stated)**	**Existing average service rate, hr^{-1}, per server (μ)**	**Recommended number of servers in the clinic, s, (physicians or as stated)**	**Change in number of servers in the clinic**
Medical	351	6	7.31	13	+7
Ophthalmology	189	4	5.91	8	+4
Paediatric OPD	98	4	3.06	4	-
Regular Gynaecology OPD	138	4	4.31	5	+1
Psychiatry	72	2	4.50	3	+1
ENT	76	2	4.75	3	+1
Neurology	134	3	5.58	5	+2
Dermatology	62	2	3.88	3	+1

Table 18. Expected service improvement at clinical departments with targeted 15 minutes consultation time

Clinical Department	**Improved average service rate, hr^{-1}, per server (μ)**	**Average number of patients in the system (L)**	**Average patients in Queue (Lq)**	**Average Time Spent, in min, in System (W)**	**Average time waiting, in min, in line (Wq)**	**Average consultation time in minute**
Medical	3.38	28.43	18.03	48.60	30.82	15.28
Ophthalmology	2.95	15.94	9.54	50.59	30.27	16.32
Paediatric OPD	3.06	8.20	5.00	50.19	30.60	16.59
Regular Gynaecology OPD	3.45	11.03	7.03	47.96	30.57	14.39
Psychiatry	3.00	5.98	3.58	49.83	29.83	17.50
ENT	3.17	6.23	3.83	49.16	30.22	15.95
Neurology	3.35	10.78	6.78	48.28	30.37	14.91
Dermatology	2.58	6.08	3.41	52.93	29.71	18.23

Again, by evaluating from the 15 minute consultation time target, the above tables show that there is still a major improvement needed on the Medical clinic (7 additional physicians required) and Ophthalmology (4 additional physicians required) clinics while Neurology requires an additional of 2 physicians and the remaining Paediatric OPD which requires an additional single server each. The Paediatric OPD only requires patient arrival management without any additional server and can meet the required consultation time.

For the clinics that requires an additional server, we can consider an optimization with the minimum recommended consultation time, which is 10 minutes, as shown in the following tables.

Table 19. Recommended number of servers at clinical departments by assuming 10 min consultation time

Clinical Department	**Average number of patients per day**	**Existing number of servers in the clinic, s, (physicians or as stated)**	**Existing average service rate, hr^{-1}, per server (μ)**	**Recommended number of servers in the clinic, s, (physicians or as stated)**	**Change in number of servers in the clinic**
Medical	351	6	7.31	9	+3
Ophthalmology	189	4	5.91	6	+2
Regular Gynaecology OPD	138	4	4.31	4	-
Psychiatry	72	2	4.50	2	-
ENT	76	2	4.75	2	-
Neurology	134	3	5.58	4	+1
Dermatology	62	2	3.88	2	-

Table 20. Expected service improvement at clinical departments with targeted 10 minutes consultation time

Clinical Department	**Improved average service rate, hr^{-1}, per server (μ)**	**Average number of patients in the system (L)**	**Average patients in Queue (Lq)**	**Average Time Spent, in min, in System (W)**	**Average time waiting, in min, in line (Wq)**	**Average consultation time in minute**
Medical	4.88	25.30	18.10	43.24	30.94	9.81
Ophthalmology	3.94	14.37	9.57	45.63	30.39	11.24
Regular Gynaecology OPD	4.31	10.17	6.97	44.23	30.31	10.91
Psychiatry	4.50	7.20	4.91	42.00	28.66	10.83
ENT	4.75	5.50	3.90	43.42	30.79	9.63
Neurology	4.19	9.94	6.74	44.52	30.19	11.33
Dermatology	3.88	5.08	3.31	44.27	28.79	11.48

From the 10-minute consultation time target optimization, we can see that the Medical clinic requires 3, the Ophthalmology clinic 4 and the Neurology clinic requires 1 additional physician to fulfil the minimum requirement. While, the remaining clinics can be optimized by only managing the patient arrival. As we expected, this is relatively the minimum first step towards improved service delivery and can be taken as a first step towards phase-by-phase optimization.

The next step is optimization of the test and surgical centers. It is very difficult to put an average consultation time for both centers as the time it took for the test is dependent on the machine type, model, test complication and other factors. Similarly, it is also difficult to standardize the average time it took for a surgical procedure. In most cases, surgical procedures usually take hours and then after the patient is expected to go into recovery for two to three hours before admitting back to the hospital floor. Therefore, for both the test and the surgical centers, due to the difficulty of placing an average diagnosis and operation period, this research only considered an average waiting time of 30 minutes for the optimization.

Table 21. Recommended number of servers at test centers

Test Center	**Average number of patients per day**	**Existing number of machines in the clinic (s)**	**Existing average service rate, hr^{-1}, per server (μ)**	**Recommended number of machines in the clinic (s)**	**Change in number of servers in the clinic**
X-ray	98	2	2.08	2	-
Ultrasound	76	6	1.58	6	-
MRI	137	1	17.13	2	+1
CT scan	104	1	13.00	2	+1

Table 22. Expected improved service performance at test centers

Test Center	**Improved average service rate, hr^{-1}, per server (μ)**	**Average number of patients in the system (L)**	**Average patients in Queue (Lq)**	**Average Time Spent, in min, in System (W)**	**Average time waiting, in min, in line (Wq)**
X-ray	2.08	4.90	2.72	65.97	36.58
Ultrasound	1.58	4.17	-	29.64	-
MRI	8.56	8.12	6.52	35.54	28.54
CT scan	6.50	7.56	5.96	43.59	34.36

From this system optimization, there will be an improvement in the average waiting time at MRI center from the existing 42.48 minutes into 28.54 minutes, which is around 33% reduction. Moreover, deploying one additional test machine at the CT scan center will improve the existing 96.92 minutes of average waiting time into 34.36 minutes, which in turn is a progress on the waiting time by around 65%.

Table 23. Recommended number of servers at Minor Operation Center

Minor Operation Center	**Average number of patients per day**	**Existing number of beds in the clinic (s)**	**Existing average service rate, hr^{-1}, per server (μ)**	**Recommended number of beds in the clinic (s)**	**Change in number of servers in the clinic**
OR	19	4	0.59	5	+1
Ophthalmology					
ENT					
E & C					

Table 24. Expected service improvement at Minor Operation Center

Minor Operation Center	Improved average service rate, hr^{-1}, per server (μ)	Average number of patients in the system (L)	Average patients in Queue (Lq)	Average Time Spent, in min, in System (W)	Average time waiting, in min, in line (Wq)
OR Ophthalmology ENT E & C	0.48	7.16	1.45	158.36	32.04

An addition of one surgical server on the minor operation center will improve the existing service by reducing the average waiting time from 86.04 minutes into 32.04 minutes, which is about 63% reduction and accordingly the average operation time will improve from 101.05 minutes into 126.32 minutes, which is a 25% improvement.

Table 25. Recommended number of servers at Major Operation Center

Major Operation Center	Average number of patients per day	Existing number of beds in the clinic (s)	Existing average service rate, hr^{-1}, per server (μ)	Recommended number of beds in the clinic (s)	Change in number of servers in the clinic
General Surgery Urology Paediatric	19	4	0.59	5	+1
Ophthalmology Orthopaedic Neuro-Surgery	63	8	0.98	11	+3

Table 26. Expected service improvement at Major Operation Center

Major Operation Center	Improved average service rate, hr^{-1}, per server (μ)	Average number of patients in the system (L)	Average patients in Queue (Lq)	Average Time Spent, in min, in System (W)	Average time waiting, in min, in line (Wq)
General Surgery Urology Paediatric	0.48	7.16	1.45	158.36	32.04
Ophthalmology Orthopaedic Neuro-Surgery	0.72	17.32	4.75	115.45	31.65

An addition of one surgical server on the General surgery operation center will improve the existing service by reducing the average waiting time from 86.04 minutes into 32.04 minutes, which is about a 63% improvement and accordingly a 25% average time improvement, which is from 101.14 minutes into 126.32 minutes. Meanwhile, the ophthalmology surgical operation center requires an additional three servers to improve the waiting time from 149.02 minute into 31.65 minute, which is around a 79% improvement. In the meantime, the surgical operation time would improve from the existing 60.95 minute into 83.80 minute.

The other consideration, towards creating a better healthcare system, needs to be implementing and having a serious intervention on the dual practice, and regulate the existing poor implementations and lack of any significant control to ensure the positive outcomes and minimize its drawbacks.

Applying the queue theory from time to time will assure and support the healthcare managers and decision makers to apply a more effective and efficient scheduling system, in some probabilistic assumptions, and to standardize the consultation time duration and to minimize the time wasted in waiting.

In the meantime, the hospital should take measures to reduce the patient boredom during this lengthy waiting time, within the constraints of the public healthcare delivery system. These measures can be using digital signage at the waiting area, as it is a proven method in effectively minimizing the perceived patient wait time. The patient will feel like they were there for a shorter waiting time because they were being entertained.

References

Abera, G. G., Alemayehu, Y. K., & Herrin, J. (2017). Public-on-private dual practice among physicians in public hospitals of Tigray National Regional State, North Ethiopia: perspectives of physicians, patients and managers. *BMC Health Services Research, 17*(1), 713. doi:10.1186/s12913-017-2701-6

Addissie, A., Jira, C., & Bekele, F. (1998). Patient waiting and service time in the outpatient department of Jimma hospital. *Ethiopian Journal of health Sciences, 8*(1), 15-21.

Aeenparast, A., Farzadi, F., & Maftoon, F. (2012). Waiting time for specialist consultation in Tehran. *Arch Iran Med, 15*(12), 756-758. doi:0121512/AIM.007

Afrane, S., & Appah, A. (2014). Queuing theory and the management of Waiting-time in Hospitals: The case of Anglo Gold Ashanti Hospital in Ghana.

Agarwal, R., & Singh, B. K. (2018). An analytical study of queues in medical sector. *Opsearch, 55*(2), 268-287. doi:10.1007/s12597-017-0324-7

Ahmad, B. A., Khairatul, K., & Farnaza, A. (2017). An assessment of patient waiting and consultation time in a primary healthcare clinic. *Malays Fam Physician, 12*(1), 14-21. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/28503269

Akachi, Y., & Kruk, M. E. (2017). Quality of care: measuring a neglected driver of improved health. *Bull World Health Organ, 95*(6), 465-472. doi:10.2471/BLT.16.180190

Alabduljabbar, R., Madhi, R. B., Medan, H., Alaqeel, N., & Alsubaie, M. (2018). *UrNext : A Proposed Location-aware Patient Queuing Solution using BLE Technology.* Paper presented at the 21st Saudi Computer Society National Computer Conference, NCC 2018.

American Heritage Dictionary. (2007). *The American heritage medical dictionary*: Houghton Mifflin Harcourt.

Anand, K. S., Paç, M. F., & Veeraraghavan, S. (2011). Quality–speed conundrum: Trade-offs in customer-intensive services. *Management Science, 57*(1), 40-56.

Atnafu, A., Haile Mariam, D., Wong, R., Awoke, T., & Wondimeneh, Y. (2015). Improving Adult ART Clinic Patient Waiting Time by Implementing an Appointment System at Gondar University Teaching Hospital, Northwest Ethiopia. *Advances in Public Health, 2015*, 892464. doi:10.1155/2015/892464

Au-Yeung, S. W. M., Harrison, P. G., & Knottenbelt, W. J. (2006). *A queueing network model of patient flow in an accident and emergency department.*

Paper presented at the ESM 2006 - 2006 European Simulation and Modelling Conference: Modelling and Simulation 2006.

Aziati, A. H. N., & Hamdan, N. S. B. (2018). *Application of queuing theory model and simulation to patient flow at the outpatient department.* Paper presented at the Proceedings of the International Conference on Industrial Engineering and Operations Management.

Babes, M., & Sarma, G. V. (1991). OUTPATIENT QUEUES AT THE IBN-ROCHD-HEALTH-CENTER. *Journal of the Operational Research Society, 42*(10), 845-855. doi:10.1038/sj/jors/0421003

Bachmann, M. O., & Barron, P. (1997). Why wait so long for child care? An analysis of waits, queues and work in a South African urban health centre. *Tropical Doctor, 27*(1), 34-38. doi:10.1177/004947559702700113

Bahadori, M., Mohammadnejhad, S. M., Ravangard, R., & Teymourzadeh, E. (2014). Using queuing theory and simulation model to optimize hospital pharmacy performance. *Iran Red Crescent Med J, 16*(3), e16807. doi:10.5812/ircmj.16807

Bailey, N. T. J. (1952). A STUDY OF QUEUES AND APPOINTMENT SYSTEMS IN HOSPITAL OUT-PATIENT DEPARTMENTS, WITH SPECIAL REFERENCE TO WAITING-TIMES. *Journal of the Royal Statistical Society Series B-Statistical Methodology, 14*(2), 185-199. Retrieved from <Go to ISI>://WOS:A1952YH06700003

Bamgboye, E., Erinoso, H., & Ogunlesi, A. (1992). The waiting time at the children's emergency room, University College Hospital, Ibadan. *Nigerian Journal of Paediatrics, 19*(1), 0-14.

Berhan, Y. (2008). Medical doctors profile in Ethiopia: production, attrition and retention. *Ethiop Med J, 46*(suppl 1), 1-77.

Berhan, Y. (2008). Medical doctors profile in Ethiopia: production, attrition and retention. In memory of 100-years Ethiopian modern medicine & the new Ethiopian millennium. *Ethiop Med J, 46 Suppl 1*, 1-77.

Bernstein, L. (2014, May 22, 2014 at 9:09 p.m. GMT+2). How many patients should your doctor see each day? *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/to-your-health/wp/2014/05/22/how-many-patients-should-your-doctor-see-each-day/

Bir, A., & Eggleston, K. (2003). Physician dual practice: access enhancement or demand inducement. *Tufts University Department of Economics Working Paper, 11*.

Britt, H., Valenti, L., & Miller, G. (2002). Time for care. Length of general practice consultations in Australia. *Australian Family Physician, 31*(9), 876.

Bruni, R. A., Laupacis, A., Levinson, W., & Martin, D. K. (2010). Public views on a wait time management initiative: a matter of communication. *BMC Health Serv Res, 10*, 228. doi:10.1186/1472-6963-10-228

Cape, J. (2002). Consultation length, patient-estimated consultation length, and satisfaction with the consultation. *Br J Gen Pract, 52*(485), 1004-1006.

Cayirli, T., & Veral, E. (2003). Outpatient scheduling in health care: a review of literature. *Production and operations management, 12*(4), 519-549.

Chand, S., Moskowitz, H., Norris, J. B., Shade, S., & Willis, D. R. (2009). Improving patient flow at an outpatient clinic: study of sources of variability and improvement factors. *Health care management science, 12*(3), 325-340.

Chen, L. M., Farwell, W. R., & Jha, A. K. (2009). Primary care visit duration and quality: does good care take longer? *Archives of internal medicine, 169*(20), 1866-1872.

Cho, K. W., Kim, S. M., Chae, Y. M., & Song, Y. U. (2017). Application of queueing theory to the analysis of changes in outpatients' waiting times in hospitals introducing EMR. *Healthcare Informatics Research, 23*(1), 35-42. doi:10.4258/hir.2017.23.1.35

Council of Supply Chain Management Professionals. (2015). Definition of logistics management. Retrieved from https://cscmp.org/about-us/supply-chain-management-definitions

Creemers, S., & Lambrecht, M. (2007). Modeling a healthcare system as a queueing network: The case of a Belgian hospital. *Available at SSRN 1093618*.

Dachyar, M., Farizal, & Yafi, M. M. (2018). *Analysis of Outpatient Service Queue of Public Hospital in Jakarta.* Paper presented at the MATEC Web of Conferences.

Deveugele, M., Derese, A., van den Brink-Muinen, A., Bensing, J., & De Maeseneer, J. (2002). Consultation length in general practice: cross sectional study in six European countries. *BMJ, 325*(7362), 472. doi:10.1136/bmj.325.7362.472

Eitel, D. R., Rudkin, S. E., Malvehy, M. A., Killeen, J. P., & Pines, J. M. (2010). Improving service quality by understanding emergency department flow: a White Paper and position statement prepared for the American Academy of Emergency Medicine. *J Emerg Med, 38*(1), 70-79. doi:10.1016/j.jemermed.2008.03.038

Fanjiang, G., Grossman, J. H., Compton, W. D., & Reid, P. P. (2005). *Building a better delivery system: a new engineering/health care partnership*: National Academies Press.

FDRE Ministry of Health. (2015). *Health Sector Transformation Plan*. Retrieved from https://www.globalfinancingfacility.org/ethiopia-health-sector-transformation-plan-201516-201920

Feibert, D. C., Andersen, B., & Jacobsen, P. (2019). Benchmarking healthcare logistics processes–a comparative case study of Danish and US hospitals. *Total Quality Management and Business Excellence, 30*(1-2), 108-134. doi:10.1080/14783363.2017.1299570

Ferrinho, P., Lerberghe, W. V., Julien, M. R., Fresta, E., Gomes, A., Dias, F., . . . Bäckström, B. (1998). How and why public sector doctors engage in private practice in Portuguese-speaking African countries. *Health Policy and Planning, 13*(3), 332-338.

Ferrinho, P., & Van Lerberghe, W. (2004). Inês Fronteira, Fátima Hipólito, and André Biscaia. 2004.". *Dual Practice in the Health Sector: Review of the Evidence." Human Resources for Health, 2*(1), 14.

Fomundam, S., & Herrmann, J. W. (2007). *A survey of queuing theory applications in healthcare*. Retrieved from

Frichi, Y., Jawab, F., Boutahari, S., Zehmed, K., Moufad, I., Akoudad, K., & Laaraj, N. (2018). *Hospital logistics: An effective tool in improving the* *quality of care*. Paper presented at the Proceedings of the International Conference on Industrial Engineering and Operations Management.

Ghazali, R. J. D. M., Manaf, N. H. A., Abdullah, A. H. A., Bakar, A. A., Salikin, F., Umapathy, M., . . . Ismail, W. I. W. (2011). Hospital waiting time: the forgotten premise of healthcare service delivery? *International journal of health care quality assurance*.

Gizaw, R., & Gebremedhin, D. L. T. (2011). Implementing a Balanced Scorecard at St. Paul's Hospital, Addis Ababa Ethiopia. *Ministerial Leadership Initiative for Global Health*.

Green, L. (2006). Queueing analysis in healthcare. In *Patient flow: reducing delay in healthcare delivery* (pp. 281-307): Springer.

Health, F. D. R. o. E. M. o. (2010). Health Sector Development Program IV: 2010/11-2014/15. In: Ministry of Health Addis Ababa.

Heizer, J. H., & Render, B. (2008). *Operations management* (Vol. 1): Pearson Education India.

Hobbs, D., Kunzman, S. C., Tandberg, D., & Sklar, D. (2000). Hospital factors associated with emergency center patients leaving without being seen. *The American journal of emergency medicine, 18*(7), 767-772.

Huarng, F., & Lee, M. H. (1996). Using simulation in out-patient queues: a case study. *Int J Health Care Qual Assur, 9*(6), 21-25. doi:10.1108/09526869610128232

Hwang, J.-I. (2006). Factors influencing consultation time and waiting time of ambulatory patients in a tertiary teaching hospital. *Journal of Korean Society Quality Assurance Health Care, 12*(1), 6.

Johansson, S. C. (2014). *Migration of Ethiopian Doctors.*

Kendall, D. G. (1953). Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain. *The Annals of Mathematical Statistics*, 338-354.

Kriegel, J., Jehle, F., Dieck, M., & Mallory, P. (2013). Advanced services in hospital logistics in the German health service sector. *Logistics Research, 6*(2), 47-56. doi:10.1007/s12159-013-0100-x

Kuhn, M., & Nuscheler, R. (2013). *Saving the public from the private? Incentives and outcomes in dual practice*. Retrieved from

Lakshmi, C., & Iyer, S. A. (2013). Application of queueing theory in health care: A literature review. *Operations research for health care, 2*(1-2), 25-39.

Luo, J., Wu, M., Gopukumar, D., & Zhao, Y. (2016). Big Data Application in Biomedical Research and Health Care: A Literature Review. *Biomed Inform Insights, 8*, 1-10. doi:10.4137/BII.S31559

M Oche, O., & Umar, A. (2011). Patient waiting time in a tertiary health institution in Northern Nigeria. *3*, 78-82.

Mackey, T., & Cole, F. (1997). Patient waiting time in a nurse managed clinic. *The Internet Journal of Advanced Nursing Practice, 1*(1).

Macq, J., Ferrinho, P., De Brouwere, V., & Van Lerberghe, W. (2001). Managing health services in developing countries: between ethics of the civil servant and the need for moonlighting: managing and moonlighting. *Human Resources for Health Development Journal, 5*(1-3), 17-24.

Mahomed, H., & Bachmann, M. O. (1998). Block appointments in an overloaded South African health centre: quantitative and qualitative evaluation. *International Journal of Health Care Quality Assurance*.

Makary, M. A., & Daniel, M. (2016). Medical error-the third leading cause of death in the US. *BMJ, 353*, i2139. doi:10.1136/bmj.i2139

Mardiah, F. P., & Basri, M. H. (2019). The Analysis of Appointment System to Reduce Outpatient Waiting Time at Indonesia's Public Hospital. *Human Resource Management Research, 3*(1), 27-33. doi:10.5923/j.hrmr.20130301.06

Mechanic, D., McAlpine, D. D., & Rosenthal, M. (2001). Are patients' office visits with physicians getting shorter? *New England Journal of Medicine, 344*(3), 198-204.

Melka, A. E. (2016). *Experiences with and attitude towards medicine use among Ethiopian immigrants in Norway.*

Melo, T. (2012). *Hospital logistics: Lessons learned from using Operations Research to improve resource management.* Paper presented at the Workshop on The paths of Combinatorial Optimization Applications in Portugal st.

Michtalik, H. J., Yeh, H. C., Pronovost, P. J., & Brotman, D. J. (2013). Impact of attending physician workload on patient care: a survey of hospitalists. *JAMA Intern Med, 173*(5), 375-377. doi:10.1001/jamainternmed.2013.1864

Mittal, P., Chatterjee, A., Hasnain, S. S., & Varshney, S. (2016, 2016). *Application of Queuing Model to Determine Waiting Time of Patients at the Blood Collection Centre at the Opd of a Tertiary Care Teaching Hospital*.

Moore, B. J. (1977). Use of queueing theory for problem solution in Dallas, Tex., Bureau of Vital Statistics. *Public health reports, 92*(2), 171.

Ny Net, S., & Chompikul, J. (2007). Patient Satisfaction with Health Services at the Out-Patient Department Clinic of Wangmarnyen Community Hospital, Sakeao Province, Thailand. *Journal of Public Health and Development, 5*(2), 33-42.

O'Malley, M. S., Fletcher, S. W., Fletcher, R. H., & Earp, J. A. (1983). Measuring patient waiting time in a practice setting: A comparison of methods. *The Journal of ambulatory care management, 6*(3), 20-27.

Ocak, T., Bekdas, M., Duran, A., Göksügür, S. B., & Küçükbayrak, B. (2013). The Obstetrics Gynecology and Children's Hospital Emergency Room waiting time before hospitalization. *African health sciences, 13*(4), 1162-1169. doi:10.4314/ahs.v13i4.43

Oche, M., & Adamu, H. (2013). 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), 588-592.

Oche, M., & Adamu, H. (2013). Determinants of patient waiting time in the general outpatient department of a tertiary health institution in north Western Nigeria. *Ann Med Health Sci Res, 3*(4), 588-592. doi:10.4103/2141-9248.122123

Ogden, J., Bavalia, K., Bull, M., Frankum, S., Goldie, C., Gosslau, M., . . . Vasant, K. (2004). "I want more time with my doctor": a quantitative study of time and the consultation. *Family Practice, 21*(5), 479-483.

Palvannan, R. K., & Teow, K. L. (2012). Queueing for healthcare. *J Med Syst, 36*(2), 541-547. doi:10.1007/s10916-010-9499-7

Park, S.-H. (2001). Analysis of factors delaying on waiting time for medical examination of outpatient on a hospital. *Journal of Korean Society Quality Assurance Health Care*, *8*(1), 56.

Petek Ster, M., Svab, I., & Zivcec Kalan, G. (2008). Factors related to consultation time: experience in Slovenia. *Scand J Prim Health Care, 26*(1), 29-34. doi:10.1080/02813430701760789

Pollock, K., & Grime, J. (2003). GPs' perspectives on managing time in consultations with patients suffering from depression: a qualitative study. *Fam Pract, 20*(3), 262-269. doi:10.1093/fampra/cmg306

Prabakaran, R., & Kumar, K. (2019). Application of queuing theory in hospital management. *International Journal of Innovative Technology and Exploring Engineering*, *8*(8), 1788-1790. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-85067880030&partnerID=40&md5=fce2fe7064210f02dda5da477b10b402

Preater, J. (2002). Queues in Health. *Health Care Management Science, 5*(4), 283-283. doi:10.1023/A:1020334207282

Qureshi, N. (2010). Triage systems: a review of the literature with reference to Saudi Arabia. *EMHJ-Eastern Mediterranean Health Journal, 16 (6), 690-698, 2010*.

Russo, G., McPake, B., Fronteira, I., & Ferrinho, P. (2014). Negotiating markets for health: an exploration of physicians’ engagement in dual practice in three African capital cities. *Health policy and planning, 29*(6), 774-783.

S Russell, R., W Taylor, B., Castillo, I., & Vidyarthi, N. (2011). *Operations Management Creating Value Along the Supply Chain*.

Samuel, P., Aldeen, A., Gravenor, S., & Malik, S. (2015). 28 Analysis of Emergency Department Consultation Times. *Annals of Emergency Medicine*, *4*(66), S10.

Scanzano, P., Caracci, G., De Michelis, G., Casertano, L., Duranti, G., Mantuano, C., . . . Splendori, F. (2005). [Review and evaluation of waiting time indicators used in Italy]. [Valutazione degli indicatori utilizzati in Italia per il monitoraggio delle liste d'attesa.]. *Ig Sanita Pubbl, 61*(2), 163-212. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/17211958

Socha, K. Z., & Bech, M. (2011). Physician dual practice: a review of literature. *Health policy, 102*(1), 1-7.

SPHMMC. (2020). Saint paul's Hospital millennium medical college. Retrieved from https://sphmmc.edu.et/

Srivatsa Srinivas, S., & Marathe, R. R. (2020). Using queue to signal quality in primary care: The need for further empirical investigation. *International Journal of Health Planning and Management*, *35*(1), 394-396. doi:10.1002/hpm.2907

Su, S., & Shih, C.-L. (2003). Managing a mixed-registration-type appointment system in outpatient clinics. *International journal of medical informatics, 70*(1), 31-40.

Thomke, S., & Fujimoto, T. (2000). The effect of “front-loading” problem-solving on product development performance. *Journal of Product Innovation*

Management: AN INTERNATIONAL PUBLICATION OF THE PRODUCT DEVELOPMENT & MANAGEMENT ASSOCIATION, 17(2), 128-142.

Wang, A., Guan, S., Koong, K. S., & Koong, L. Y. (2016). An M/M/1 and M/M/2 queuing solution to improve virtual scheduling in healthcare. *International Journal of Electronic Healthcare, 9*(1), 42-59. doi:10.1504/ijeh.2016.078744

Watson, W. T., Marshall, E. S., & Fosbinder, D. (1999). Elderly patients' perceptions of care in the emergency department. *J Emerg Nurs, 25*(2), 88-92. doi:10.1016/s0099-1767(99)70152-2

WHO, W. H. O. (1948). World health organization constitution. *Basic documents, 1*, 22.

Wilson, A., & Childs, S. (2002). The relationship between consultation length, process and outcomes in general practice: a systematic review. *Br J Gen Pract, 52*(485), 1012-1020.

Wilson, A., McDonald, P., Hayes, L., & Cooney, J. (1992). Health promotion in the general practice consultation: a minute makes a difference. *BMJ (Clinical research ed.), 304*(6821), 227-230. doi:10.1136/bmj.304.6821.227

Worthington, D. (1991). Hospital Waiting List Management Models. *Journal of the Operational Research Society, 42*(10), 833-843. doi:DOI 10.1038/sj/jors/0421002

Worthington, D. J. (1987). Queueing models for hospital waiting lists. *J Oper Res Soc, 38*(5), 413-422. doi:10.1057/jors.1987.69

Yaduvanshi, D., Sharma, A., & More, P. V. (2019). Application of queuing theory to optimize waiting-time in hospital operations. *Operations and Supply Chain Management, 12*(3), 165-174. doi:10.31387/oscm0380240

Yeon, N., Lee, T., & Jang, H. (2010). *Outpatients appointment scheduling with multi-doctor sharing resources.* Paper presented at the Proceedings of the 2010 Winter Simulation Conference.

Yigzaw, D. Y. (2009). *Brain drain in Ethiopia: perception of Ethiopians living in Norway*.

Yin, C. (2010). *Outpatient queue business simulation based on acceptable waiting time.* Paper presented at the 2010 International Conference on Computer Design and Applications, ICCDA 2010.

Zhu, Z. C., Heng, B. H., & Teow, K. L. (2009). Simulation study of the optimal appointment number for outpatient clinics. *International Journal of Simulation Modelling, 8*(3), 156-165. doi:10.2507/ijsimm08(3)3.132

Zonderland, M. E., & Boucherie, R. J. (2012) Queuing networks in healthcare systems. In*: Vol. 168. International Series in Operations Research and Management Science* (pp. 201-243).