AN INVESTIGATION OF MULTI-SERVER QUEUING ANALYSIS TO ASSESS HOSPITAL HEALTHCARE SYSTEMS' OPERATIONAL EFFECTIVENESS

Abstract- Prolonged wait (queue) times in medical outpatient departments are a growing concern in Nigerian hospitals/clinics, due to a variety of consequences such as overcrowding, patients leaving in anger without being attended to, and being stressed for not staying too long in the system. The primary goal of this paper is to research various techniques or methods for reducing long queues. Patients who wait for minutes, hours, days, or months to receive medical services may incur waiting costs. The time spent in the queue could have been better used elsewhere. This paper aims to determine an optimal server level while keeping total system costs to a minimum, including expected service costs and waiting costs in a multi-server system, to reduce patient congestion in the hospital. Data for the study was collected in two ways. The secondary method was first used to identify the most congested OPD among the numerous OPDs considered in the study. The performance measures costs were then calculated using primary data. The performance measures of the queuing system were calculated using TORA optimization software. MS Excel was used to calculate the costs and plot the charts. Based on the results of the analysis, it was suggested that one physician be added to the hospital's medical OPD to reduce patient overcrowding and wait times. As a result, this call for refocusing is issued to improve overall patient care in our cultural context while also meeting the needs of patients in our society.

Keywords: Multi-server, utilization factor, reneging, waiting costs, service costs, and servers.

1.0 Introduction

According to the WHO definition, health is the overall state of complete social, emotional, mental, and physical well-being that is regarded as a resource for living a full life; thus, access to the highest achievable level of health is a basic right of every human being regardless of social or economic circumstances (WHO, 1948). Health is not simply the absence of disease; rather, it is the ability to recover from illness and/or other issues.

Healthcare, on the other hand, is an act performed by medical professionals. Is the preservation of health through prevention, diagnosis, and treatment of mental and physical impairments? Healthcare systems exist to ensure and assist people in maintaining their optimal state.

An outpatient department is a healthcare department dedicated to diagnosing and consulting with outpatients (American Heritage Dictionary, 2007; Zhu, Heng, and Teow, 2009). Outpatients are patients who come to see a doctor for treatment and then return home without being admitted to the hospital.

The number of qualified doctors in a nation like Nigeria is far less than what the system needs. As a result, physicians are overburdened by having to attend to outpatients in multiple centers as well as inpatients in multiple hospitals. As a result, consultation hours are limited to only a short period, usually in the mornings or afternoons.

Moreover, despite this short duration of operation, physician tardiness and other disruptions often result in the service being rendered inactive (Babes & Sarma, 1991).

A crucial area of research in operations research and healthcare management is multi-server queuing analysis for evaluating hospital healthcare systems. To enhance patient flow and shorten wait times, queuing theory—a mathematical technique for simulating and assessing waiting lines—has been applied extensively in healthcare settings. As the health sector's growth is closely tied to the value of human resources, it serves as one measure of a nation's economic development and success. As a result of his research, a Danish engineer by the name of Erlang created queuing theory in 1909 (Winston, 1991). After experimenting, he found that the demand for phone traffic varies. Subsequently, He issued a report on

equipment delays linked to automatic dialling. Shortly after his published work and at the end of World War II, Erlang expanded the scope of his early work to include more general issues and business applications of waiting lines, as well as numerous industries. One of the many benefits of modelling the service sector as a queuing system is that, as opposed to making precise predictions about real-world systems, this approach allows for the diagnosis of issues and the identification of restrictions. The process of joining waiting lines, or queues, and their mathematical analysis are the focus of queuing theory. Models are created within this theoretical framework to determine wait times and line lengths (Sundarapandian, 2009). In scientific circles, there has been a great deal of discussion about queuing because lineups are a problem in almost every society. Queuing situations arise in a variety of professional and personal contexts. They usually involve people or things that are waiting in a queue for services while following certain behavioural guidelines (Burodo, Suleiman, & Shaba, 2019). Queues are likely to occur if there is rivalry for scarce resources (Koko, Burodo, & Suleiman, 2018). When people who are looking for assistance—also known as customers—arrive at a location where assistance is provided but are unable to receive prompt attention, lines form (Suleiman, Burodo, & Ahmed, 2022).

2.0 Literature Review

A study on the application of the multi-server queuing model to analyse the OPD queuing system during the COVID-19 pandemic was piloted by Ali et al. in 2021. Their objective was to suggest the ideal degree of service for both the outpatient department and the reception queue system. The ABC Public Hospital in Hyderabad's reception and outpatient department served as the sites for data collection. The dataset contained information on arrival times, patient service times, the number of doctors and receptionists, their salaries, and patient waiting costs. Patient arrivals and services were analyzed by the method of input analyzer of the Rockwell Arena software. To compute performance metrics, TORA optimization software was used. Various costs associated with the queuing system were computed using MS Excel, and corresponding graphs were generated. The study recommended an increase of one receptionist and one doctor to optimize the queuing system and patient flow. Additionally, it suggested reducing patients' waiting costs to a greater extent.

The goal of Kazemi et al. (2017)'s study was to use simulation models to assess the effectiveness of emergency departments in two different hospitals. They utilized queuing

performance measures to compare single-server and multi-server systems and concluded that multi-server systems can reduce waiting times and enhance patient satisfaction.

Segun (2020) investigated ways to simulate the performance of healthcare service delivery using queuing theory at Adekunle Ajasin University in Akungba-Akoko, Nigeria. The study's objectives were to determine patient waiting, arrival, and service times in the AAUA Health context and to design a suitable queuing system utilizing a simulation approach. Three weeks of weekday data collection were done using analytical and modelling methodologies in this study project. The current patient scenario wait system was modelled and replicated using PYTHON software. The AAUA Health Centre was advised by the findings to enhance the calibre of its offerings by drawing attention to certain problems, such as extended wait times during peak hours. Nor and Binti (2018) computed patient waiting, arrival, and service times in the outpatient department using the Queuing Theory Model and Simulation. A public health clinic in southern Malaysia served as the study's location, and both descriptive and simulation methods were used.

The ARENA software was used for modelling and simulation, yielding insights into patient waiting times and service efficiency.

Kumar et al. (2020) and Brailsford et al. (2008) utilized queuing models to analyse patient flow in hospital emergency departments and outpatient departments, respectively. Queuing theory offers a mathematical framework for analysing waiting lines, which is particularly beneficial in healthcare settings to improve patient flow and resource utilization.

Khan et al. (2021) studied how to use a multi-server queuing model to improve OPD and reception performance. The study suggested hiring two additional doctors based on performance measures calculated using Rockwell Arena and TORA optimization software the research aimed to enhance healthcare delivery and reduce patient waiting times.

Kembe *et al*. (2012) employed a multi-server queuing model to examine queuing characteristics at the Federal Medical Centre, Makurdi's Riverside Specialist Clinic. Their results showed that to shorten patient wait times and enhance resource efficiency, physicians' service capacity levels should be raised.

In 2020, Mittal and Sharma looked into the probabilistic model used to estimate how long COVID-19 patients would have to wait in queue. We can predict hospital patient wait times during disease outbreaks by using the Queuing Model.

3.0 Methods

Data for this study were collected from the Federal Medical Centre, Keffi, Nasarawa State. Secondary data were used for selecting the study area, spanning one month (September 2023), while primary data were collected through direct observations, personal interviews, and questionnaire methods over four weeks (Monday through Friday). Data on arrival time, waiting time, number of servers, service time, utilization rate, etc. were collected to gauge performance.

Assumptions made for the queuing system adhered to queuing theory principles.

i. The length of the line is not restricted
ii. The size of the population is uncountable. This assumption implies that the input source is unlimited; the assumption is also warranted in cases where the number of customers is finite but after being served customer rejoins the input source.
iii. The arrival rate distribution is approximately by a Poisson distribution**.**
iv. There is no balking
v. There is no reneging. This assumption implies that customers stay in line until served**.**
vi. The queue discipline is first-come first-served (FIFO)**.**
vii. The service time distribution is approximated by an exponential distribution.

3.1 The M/M/S Model

In this study, we employ the (M/M/s: FCFS/∞) - multi-server queuing model. This model assumes that customer (patient) arrivals follow a Poisson probability distribution, with an average arrival rate of ג customers per unit of time. Patients are served on a first-come, first-served basis by any available server (in this case, physicians). Service times follow an exponential distribution. S servers provide a mean service rate of μ customers (patients) per

unit of time. When evaluating number of clients, n, in the waiting system at any particular time, two scenarios can occur:

1. If the number of customers (n) is less than the number of servers (S), denoted as n<S, there will be no queue. However, (S-n) servers will remain idle, resulting in a combined service rate of μn.
2. If the number of customers (n) is greater than or equal to the number of servers (S), denoted as n≥S, all servers will be occupied, and the maximum number of customers in the queue will be (n-S). The combined service rate will then be Sμ.

Therefore, the M/M/S model provides a framework for analysing queuing systems, considering factors such as arrival rates, and service times, to better understand system performance and efficiency, and consider the number of servers.

$$\boldsymbol{p_0} = \left[\left[\sum_{\mathbf{n=0}}^{\mathbf{s-1}} \frac{(\boldsymbol{\lambda/\mu})^{\mathbf{n}}}{\mathbf{n!}}\right] + \frac{(\boldsymbol{\lambda/\mu})^{\mathbf{s}}}{\mathbf{s!}}\left(\frac{\mathbf{s\mu}}{\mathbf{s\mu-\lambda}}\right)\right]^{\mathbf{-1}} \tag{3.0}$$

Now the other features of the multiple-channel system can be found out.

The mean number of customers in the system represented by $L_s$will be,

$$\boldsymbol{L_s} = \frac{\boldsymbol{\lambda.\mu\left(\frac{\lambda}{\mu}\right)^{s}}}{\boldsymbol{(s-1)!(s-2)^2}}\boldsymbol{p_0} + \frac{\boldsymbol{\lambda}}{\boldsymbol{\mu}} \tag{3.1}$$

While the expected (average) number of customers waiting in the queue L_q is,

$$\mathbf{L_q} = \frac{\boldsymbol{\lambda.\mu\left(\frac{\lambda}{\mu}\right)^{s}}}{\boldsymbol{(s-1)!(s\mu-\lambda)^2}} * \mathbf{p_0} \tag{3.2}$$

To check the arrival of patients, the necessary parameter, is the average time a customer spends in the system defined as,

$$\boldsymbol{W_S} = \frac{\boldsymbol{L_S}}{\boldsymbol{\lambda}} = \frac{\boldsymbol{\mu\left(\frac{\lambda}{\mu}\right)^{s}}}{\boldsymbol{(s-1)!(s\mu-\lambda)^2}} * \boldsymbol{p_0} + \frac{\mathbf{1}}{\boldsymbol{\mu}} \tag{3.3}$$

Before a patient is served, the patient is expected to wait in the queue defined as,

$$W_q = \frac{L_q}{\lambda} = \frac{\mu\left(\frac{\lambda}{\mu}\right)^s}{(s-1)!(s\mu-\lambda)^2} \cdot p_0$$

(3.4)

With the chances of having to wait given by the proportion defined in form of

$$s$$

Probability as; $p(n \geq s) = \frac{\mu.\left(\frac{\lambda}{\mu}\right)}{(s-1)!(s\mu-\lambda)} \cdot p_0$ (3.19)

The utilization factor (ρ).

$$\rho = \frac{\lambda}{\mu s} \quad (3.20)$$

The likelihood of a patient arriving and encountering no queue is exceedingly low. This occurrence occurs when the service rate surpasses the arrival rate. In practical terms, this signifies idle servers, which can incur costs for the facility. The probability of a customer or patient accessing service without any wait is determined by the analysis of parameters aimed at determining the minimum number of servers required to accommodate patients' needs without any servers being idle. This analysis relies on the average number of idle servers. The utilization rate of the servers is defined as, thereby reflecting the efficiency of the M/M/s model through the ratio,

$$= \frac{Average\ number\ of\ customers\ served}{total\ number\ of\ customers}$$

3.2 Integration of Costs into the Model

The study incorporates three types of costs: Expected service cost, expected waiting cost, and expected total system cost.

1. Average service cost: This encompasses the wages paid to physicians by the hospital for their services.
2. Average opportunity cost: This represents the cost incurred by customers while waiting at the hospital, as they are unable to engage in activities that generate income.
3. Average total system cost: This combines the mean service cost and mean waiting cost.

Conducting an economic analysis of these costs will assist hospital management in striking a balance between waiting costs and service costs. This involves considering potential increases in service costs resulting from improved service provision and decreases in waiting costs incurred by patients waiting at the hospital.

$$E(Sc) = SCs \quad (3.5)$$

Where S= number of servers, Cs= service cost of each server.

Cost paid by the customers/patients due to waiting in the system

$$E(Wc) = (\lambda\ Ws)^* Cw \quad (3.6)$$

Where, λ = number of arrivals, Ws= Average time an arrival spends in the system.

4.0 Result

This paper analyzes the selection of the study area, the professions involved, associated waiting costs for patients, service costs, system costs, etc., aiming to determine the optimal cost structure. The results will inform management about the minimum number of servers

required to reduce or minimize waiting costs for patients while simultaneously ensuring services are provided at minimum service costs in the OPDs of the hospitals under study.

4.1 Selection of the Study Area

The researcher considered eleven OPDs in the case study hospital for selection of the study area. Data on the number of patient arrivals at various outpatient departments were collected from the research and statistics department for one month (September 2023). The collected data showed that the medical outpatient department experienced the highest congestion among the eleven OPDs (See fig. 1), prompting its selection as the study area due to its maximum arrival rate of patients.

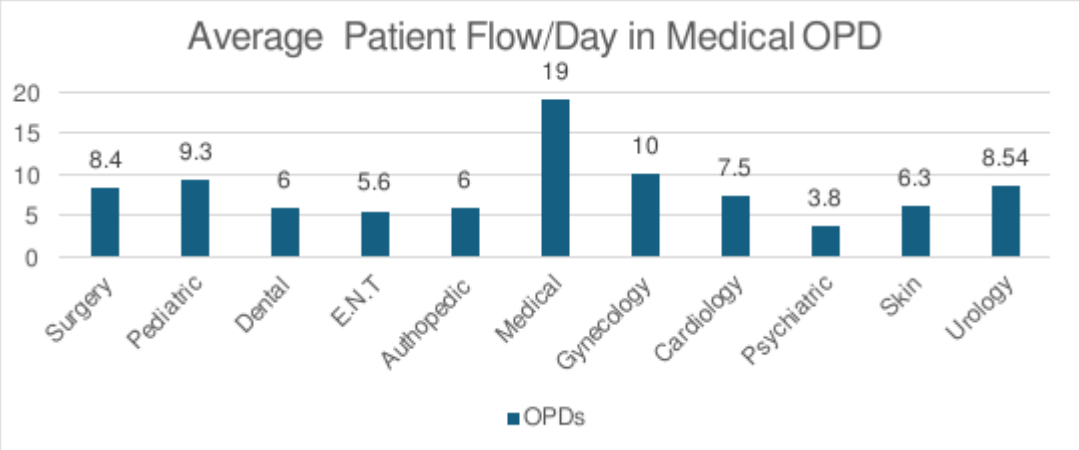


Fig. 1 Average per day of patients' arrival at Medical OPD

The selection of the Medical Out-Patients Department in the hospital for research and optimization was driven by the high volume of arrivals. This department was chosen due to observed issues of lengthy queues and extended waiting times experienced by patients. Recognizing that patient waiting time correlates with opportunity costs, it became imperative to analyze and enhance the queuing system within the Medical Out-Patient Department.

4.1.2 Profession distribution at case hospital

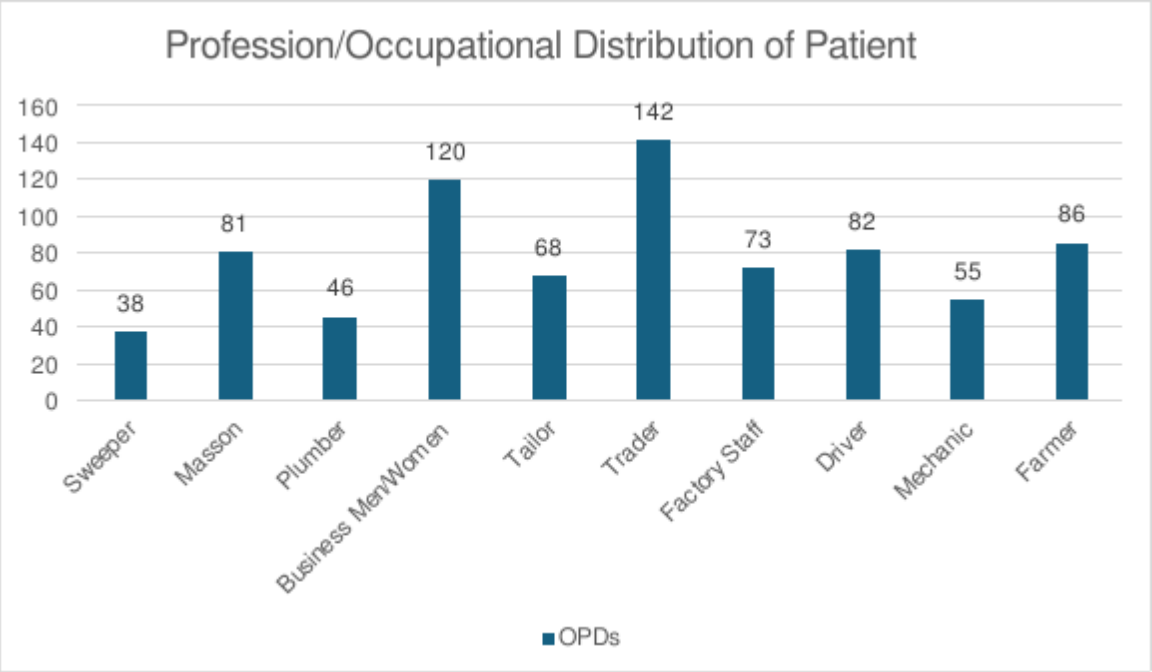


Fig. 2 Occupation distribution of the customer (patients) arriving at the OPD

Information regarding the professions of the customers was also gathered. From the collected data, it was observed that patients visiting the OPD represented a diversity of ten (10) professions (refer to Fig. 2). Fig. 2 illustrates that the largest group of patients belonged to the trading profession, totaling 142 individuals. Additionally, the average waiting cost for patients was calculated to be #119.34 (refer to Fig. 3). Notably, the highest waiting cost recorded was #214.3 for traders, attributed to their extended idle time throughout the day.

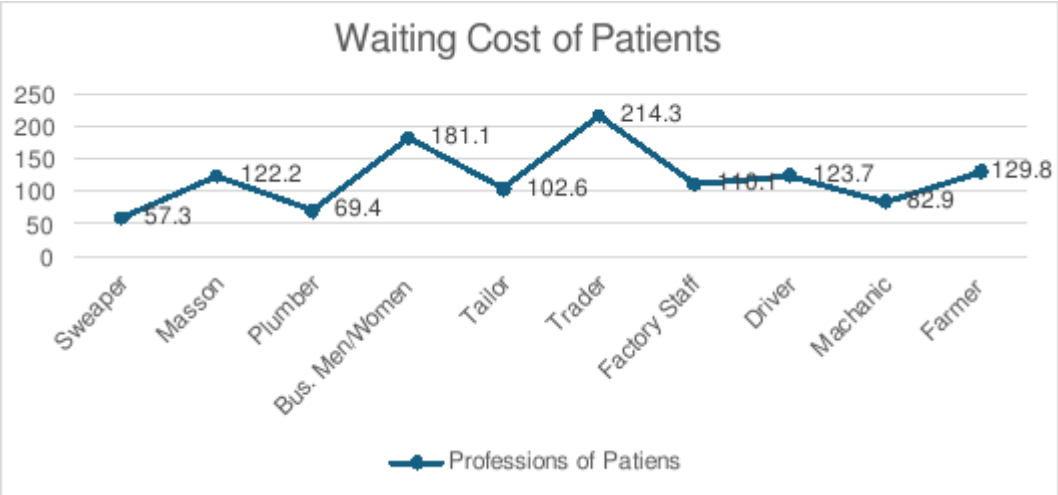


Fig. 3. Waiting costs of customers arriving at the outpatient Department

1.5 Service cost/salary of physicians at medical OPD

In this study, the service cost refers to the expenses that may change following the implementation of the optimal scenario, comprising solely the physicians' salaries. Therefore, the salaries of the physicians currently serving at the medical Out-Patient Department (OPD) of the hospital were regarded as the anticipated cost. The medical OPD currently employs five (5) physicians, and their respective salaries are detailed in Table 1.

Table: 1 physician's salaries at the case hospital

No of Physicians	Salaries of Physician Per month (₦)	Salaries of Physician Per Day (₦)	Salaries of Physician Per Hour (₦)
1	700,000.00	23,233.33	3,888.88
2	584,000.00	19,466.66	3,244.44
3	356,000.00	11,866.66	1,977.77

4	473,000.00	15,766.66	2,627.77
5	441,000.00	14,700	2,450.00
Average	510,800	17,006.66	2,837.77

Mean of the monthly wages of physician's turns out to be #510,800 at medical OPD.

2.1 Overview of the Medical Out-Patients Department (OPD) at the Case Hospital

The Medical Out-Patients Department (OPD) at the case hospital frequently experienced queues extending outside the physician hall, as patients awaited their turn for consultation A The queuing mechanism was optimised using a multi-server queuing paradigm.

Evaluation of Performance Metrics at the Medical OPD of the Hospital.

The Medical OPD emerged as the busiest and most congested area within the hospital, with a high influx of patients seeking consultation with physicians compared to other OPDs. Table 2 provides an overview of the performance metrics of the queuing analysis at the Medical outpatients department of the case hospital.

Performance metrics were computed across five different scenarios, aiming to determine the optimal number of physicians necessary to minimize patient waiting times. In each scenario, the addition of one physician was considered, maintaining the same arrival and service rates for patients.

The mean arrival rate of customers per hour was calculated as 28.8, 30.4, 24.3, 27.9, and 26.1, respectively, with five physicians. Correspondingly, the average service rate of patients was computed as 57.9, 67.3, 52.5, 57.6, and 53.7 patients per hour. Upon assessing the

utilization factor, values of 95%, 50%, 34%, 25%, and 20% were obtained, indicating that the addition of one physician could reduce congestion by 50% from the initial 95%, and so forth.

Table 2 presents a detailed breakdown of the performance measures observed within the queuing analysis of the Medical outpatient department at the case hospital.

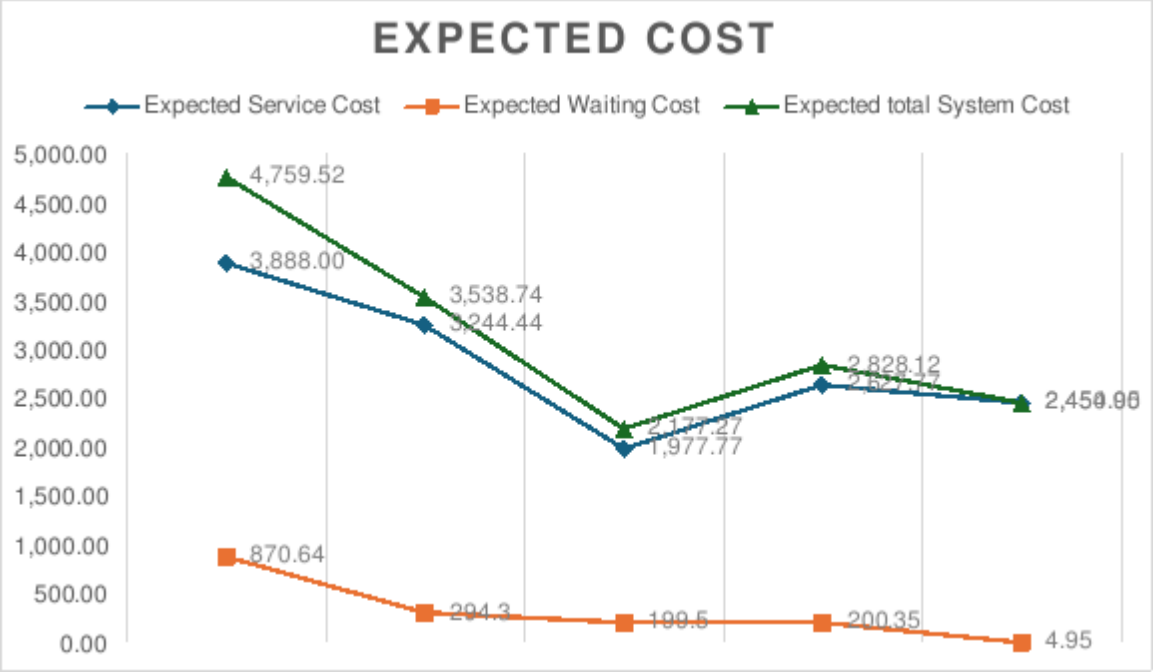
Table 2: Performance measures of the queuing system of medical OPD of case hospital

Scenarios	**1**	**2**	**3**	**4**	**5**
Physicians	1	**2**	3	4	5
Arrival rate (arrived)	577	**608**	486	558	521
Service rate (served)	563	**597**	477	552	512
Total arrival time (hr)	32	**32**	32	32	32
Total service time (aims)	2363	**2149**	2226	2310	2289
Expected service cost/hr	3,888.88	**3,244.44**	1,977.77	2,627.77	2,450.00
Expected waiting cost/hr	870.64	**294.30**	199.50	200.35	4.95
Expected total system cost/hr	4759.52	**3,538.74**	2,177.27	2,828.12	2,454.95
L'daeff	577	**608**	486	558	521
Utilization factor (Rho)	95%	**50.9%**	34%	25%	20%
Probability (P_0)	0.35884	**0.36116**	0.36100	0.3639	0.3647

Ls (Number of customers in the system/hr)	24.02487	**9.01843**	7.01887	6.01087	5.01758
Lq (Number of patients in the queue/hr)	20.4211	**4.23016**	1.02106	0.31013	0.06001
Ws (Time spent by patients in the system/hr)	.301781	**0.09682**	0.082101.	0.071810	0.00190
Wq (Time spent by patients in the queue/hr	.28200	**0.03000**	0.01000	0.00000	0.00000

Assuming two physicians were on duty, the utilization factor (Rho) would decrease at initial percent of 95% to 50.9% in scenario 2 compared to scenario 1, as illustrated in table 2. This signifies that the servers (physicians) would experience a 50.9% reduction in workload. Similarly, in scenario 3, the servers (physicians) would be 34% less occupied, and so forth. Additionally, it is projected that 20.42 patients would remain in the waiting line, with 24.02 patients in the system (refer to table 2). This suggests that with an increase of two servers (physicians), the number of patients waiting in line would decrease by 9.018.

The anticipated mean service cost per hour, average waiting cost per hour, and average total system cost per hour were computed as #3,244.44, #294.30, and #3,538.74, respectively (refer to figure 4). The performance metrics derived for the second scenario were deemed optimal since the average total system cost was observed to be at its minimum (refer to figure 4).



Scenarios

Fig. 4. Cost calculations of the medical outpatients department of case hospital

4.3 Discussion

The investigation used a multi-server queuing model to evaluate the queuing dynamics at the Federal Medical Centre in Keffi, Nasarawa State, to establish the ideal service level by assessing waiting times and service costs. The outcomes of the investigation reveal that by increasing the service capacity level of physicians at the hospital, both the expected queue length and waiting time for patients can be minimized, along with the overutilization of physicians. This optimization can be achieved at minimal total costs, which encompass service costs and waiting costs. Consequently, patients would experience reduced wait times

compared to current scenarios, leading to decreased waiting costs. This would be advantageous for healthcare providers and the healthcare system as a whole.

4.4Conclusion

The medical outpatient department (OPD) at the medical center exhibited higher congestion levels in patient flow compared to other OPDs at the Federal Medical Centre in Keffi. Based on the findings, it is recommended that the medical OPD increase its physician capacity to alleviate congestion among patients. Consequently, patient queues would be reduced or minimized, and the overall cost associated with patient waiting times would be decreased due to reduced wait times compared to the initial scenario.

4.5 Future Research

The present study did not incorporate the time spent by patients at the reception and in the laboratory due to time constraints. Subsequent research endeavours should consider these parameters to provide a more comprehensive understanding of the queuing dynamics within the healthcare facility.

Ethical Approval:

As per international standards or university standards written ethical approval has been collected and preserved by the author(s).

Consent

As per international standards or university standards, respondents' written consent has been collected and preserved by the author(s).

References

Akande, T. M., & Ayodele, O. O. (2019). Evaluation of patient queuing system at a tertiary hospital in Osogbo, Nigeria. Nigerian Journal of Clinical Practice, 22(2), 268-275. https://doi.org/10.4103/njcp.njcp_372_18

Alabi, O. M., Ogunseye, O. A., & Oyediran, O. S. (2020). Queuing analysis of healthcare system in Nigeria: A case study of University College Hospital, Ibadan. Journal of Applied Science and Environmental Management, 24(10), 1649-1655. https://doi.org/10.4314/jasem.v24i10.21

Burodo M.S., Suleiman S. and Yusuf G. (2021). An assessment of Queue management and Patient Satisfaction of Some Selected Hospitals in North-Western Nigeria, *International Journal of Mathematics and Statistics Invention (IJMSI), 9(8), 14-24.*

Burodo M.S., Suleiman S. and Shaba Y. (2019), Queuing Theory and ATM Service Optimization: Empirical Evidence from First Bank Plc, Kaura Namoda Branch, Zamfara State *American Journal of Operations Management and Information Systems; 4(3): 80-86. doi: 10.11648/j.ajomis.20190403.12*

Kembe, M. M, Onah, E. S & Iorkegh, S. (2012). A Study of Waiting and Service Costs of A Multi-Server Queuing Model In A Specialist Hospital. *International Journal of Scientific & Technology Research.*

Koko M.A., Burodo M.S. and Suleiman S. (2019). Queuing Theory and Its Application Analysis on Bus Services Using Single Server and Multiple Servers Model. American *Journal of Operations Management and Information Systems. 3(4), 81-85.doi: 10.11648/j.ajomis.20180304.12*

Kumar, S., & Bano, S.,(2020) Comparison and Analysis of Health Care Delivery Systems: Pakistan versus Bangladesh,
Journal of Hospital & Medical Management, vol. 3, no. 1, pp. 1–7, 2017.

Kumar, A., Goyal, S. K., & Singh, S. P. (2017). Multi-server queuing model for analysis of outpatient department in a hospital. Journal of Industrial Engineering and Management, 10(2), 366-382.

Nor A.H.A & Binti N.S. H (2018) Application of Queuing Theory Model and Simulation to Patient Flow at the Outpatient Department. Proceedings of the International Conference on Industrial Engineering and Operations Management Bandung, Indonesia, March 6-8.

Wang, X., Hao, L., Zhang, Y. & Guo, X. (2018). Queuing theory in healthcare: A systematic review. *Journal of healthcare engineering, 2018.*

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