FULafia Journal of Science & Technology
ISSN (print) 2449-0954 \ ISSN (online) 2636-4972
https://lafiascijournals.org.ng/index.php/fjst
Sponsored By TERTIARY EDUCATION TRUST FUND
Manuscript => Received: March, 2026; Accepted: June, 2026; Published: June, 2026 https://doi.org/10.62050/fjst2026.v10n2.780

QUEUING ANALYSIS OF PATIENT FLOW IN THE OUTPATIENT DEPARTMENT OF FEDERAL MEDICAL CENTRE, KEFFI: AN M/M/C MODEL APPROACH

M. B. Abubakar*, I. M. Saleh, S. A. Yusuf and I. Akeyede
Department of Statistics, Federal University of Lafia, PMB 146, Nasarawa State, Nigeria
*Corresponding email: abubakarbabagana001@gmail.com

ABSTRACT
Public hospitals in developing nations, including Nigeria, face problems with their outpatient departments due to extended patient waiting time. This research applies queuing theory to analyse patient flow through the Outpatient Department (OPD) of the Federal Medical Centre, Keffi, using observational data. Patient arrivals and consultation times were modelled as an M/M/c queuing system with varying numbers of doctor servers. Key performance measures estimated include average queue length, average waiting time in the queue, average number in the system, average time in the system, and system utilization. Goodness-of-fit tests were conducted to validate the model assumptions. The results showed that both inter-arrival and service times follow an exponential distribution, confirming the suitability of the M/M/c framework. The findings indicate that the OPD system is unstable when operated with one or two doctors, leading to long queues and waiting times. With three doctors, the system operates at critical capacity and remains sensitive to fluctuations in patient arrivals. The introduction of a fourth doctor results in system improvement, reducing the average queue length to 1.53 patients and the average waiting time to 5 minutes, while lowering system utilization to 0.75. Graphical analysis confirms reductions in the number of patients in the system and total system time under the four-doctor configuration, indicating a smoother and more efficient patient flow. The study demonstrates that staffing is crucial for reducing outpatient waiting times and improving service delivery. It provides evidence that M/M/c queuing models can support hospital managers in optimizing healthcare services in resource-constrained environments.

Keywords: Queuing theory, M/M/c model, Outpatient department, Waiting time, Hospital efficiency, Nigeria

INTRODUCTION

The healthcare systems of developing nations experience visitor overcrowding and extended patient waiting periods because they do not effectively manage their limited medical resources. Outpatient departments operate as the main entry point for patients to access medical treatment in public hospitals; therefore, their operational efficiency directly impacts the quality of medical services provided. Patients in outpatient departments experience too much waiting time, which leads to negative health effects and causes them to stop using medical services altogether. Service systems use queuing theory as a mathematical framework to model their customer waiting times and service area congestion problems. The first queuing models developed by Erlang in the early 1900s became popular because they serve as a standard measurement tool for various fields such as telecommunications, manufacturing, transportation, and healthcare. The models used in hospitals provide a methodical framework that helps hospitals analyze their patient movement patterns, determine their required staff levels, and assess how their operational modifications affect patient waiting times.

All service delivery systems across developed countries and developing nations face their own unique waiting line challenges. Healthcare facilities experience queues because patients need to be treated, but the medical staff cannot handle the current patient load. The system becomes imbalanced, which results in operational delays and extended patient waiting times, decreased patient satisfaction, medical staff burnout, and the development of serious medical conditions. Public healthcare facilities in developing nations face extreme challenges when they attempt to manage their patient queues because they must provide healthcare services to their entire population while dealing with their resource limitations (WHO, 2019). Queuing theory functions as a scientific method that enables researchers to analyze how service systems experience delays and congestion problems. The evolution of queuing models started with Erlang's (1909) work on telephone exchange congestion studies, which later found applications in telecommunications, transportation, manufacturing, banking, and healthcare (Gross & Harris, 1998; Taha, 2017). The application of queuing theory in hospitals enables healthcare managers to analyze patient arrival patterns together with service delivery procedures and patient waiting times to determine hospital system efficiency and make staff and resource capacity decisions.

Nigerian healthcare facilities struggle with persistent problems, which create extended patient waiting times because of their insufficient medical staff and their ineffective patient management systems, according to research evidence (Adeleke *et al.*, 2014; Oche & Adamu, 2013). Health service providers base their operational decisions about service capacity on invalidated calculations, which lead to inefficient use of their scarce medical resources. Recent research demonstrates how queuing models remain essential for resolving modern healthcare system issues. Pandey and Gangeshwer (2023) used multi-server queuing models to observe government hospital operations and discovered that patients faced major dissatisfaction because the hospital staff failed to manage the patient queue properly, which created a mismatch between their actual service capacity and the patient demand. The King Hussein Cancer Center used queuing theory to determine patient waiting times at various service delivery points while discovering which operational processes caused delays during critical medical procedures. The application of queuing models in helping room management helps enhance surgical performance and distribute resources effectively in obstetric care because queuing principles have more uses than just outpatient service applications.

Queuing models demonstrate their ability to enhance healthcare operations through evidence from additional studies. The research conducted in tertiary hospitals demonstrates that patient registration and clinical service point inefficiencies can be discovered through exponential service model analysis, which enables hospitals to make staffing and resource allocation choices (Pandey & Gangeshwer, 2023). The study shows that specialty clinical areas like obstetric operating rooms, which implement service mechanism optimization and staffing configuration design, achieve major operational efficiency improvements with their optimization efforts through their service mechanism and staffing configuration design (Lim *et al*., 2023). The research demonstrates that systems achieve optimal performance when organizations understand their service variability and implement capacity adjustments based on their findings. The new studies demonstrate how queuing theory can benefit healthcare management. The study by Iyappan (2025) shows that both single-server and multi-server queuing models function as tools for managing patient waiting periods across healthcare facilities, which include outpatient clinics, emergency departments, and diagnostic centres. The study by Pasaribu *et al*. (2025) found that Indonesian hospitals face an essential problem because their outpatient registration system encounters excessive delays from the fundamental conflict between their patient arrival rates and their ability to serve patients. The research indicates system stability requires operational efficiency enhancement through service rate adjustments.

The performance of a queue depends on its system capacity, which determines both the number of service channels and the maximum patient capacity. Outpatient facilities use the assumption of infinite capacity for analytical purposes, though actual systems face limitations from both physical space and staffing requirements. The M/M/c framework serves as the standard model to depict situations where multiple healthcare providers deliver services at the same time. The research found support for the hypothesis that increasing the number of servers in healthcare facilities decreases both waiting times and queue lengths while the patient arrival rates remain unchanged (Umar *et al*., 2024). The research demonstrates that organizations should align their service capacity with actual demand patterns to avoid systems reaching their maximum capacity. The research study examines how smart healthcare systems have begun integrating queuing models with real-time technologies through their use of blockchain-based frameworks to achieve dynamic resource allocation and operational efficiency improvements. The core elements of queuing systems, which include arrival processes, service mechanisms, queue disciplines, and system capacities, serve as the foundation for studying healthcare service operations. The research demonstrates that precise component modeling establishes the required foundation for accurate clinical dynamics depiction, which enables efficiency identification and evidence-based service delivery improvements. The current study will analyze patient flow and service provision in the OPD through queuing theory while determining performance metrics for different service levels and creating actionable strategies to enhance outpatient service operations.

MATERIALS AND METHODS

Study Area

The tertiary healthcare institution, Federal Medical Centre Keffi, exists in Nasarawa State, Nigeria, as a medical facility. The hospital serves as a referral center for the surrounding states of Abuja, Benue, and Plateau. The OPD operates as a medical facility that serves a vast patient population while experiencing high patient traffic during its clinic hours.

Data Collection and Description

Data Collection and Description Primary data for the study were collected through direct observation and time-motion study over a period of five consecutive working days, capturing patient arrivals, consultation start and end times, and service completion times. The structured data collection sheet recorded the exact time each patient arrived at the OPD while recording the time they received physician care and the length of their consultation. Patients were anonymized through the system, which recorded no personal identifiers to protect their identity while maintaining ethical research requirements.

From the observed data, the following key variables were extracted:

i. **Arrival rate (λ):** Defines the hourly rate at which patients enter during the observation period. The average arrival rate was estimated as 18 patients per hour, which reflected the peak-hour patient flow of the hospital.

ii. **Service rate (μ):** The number of patients attended per hour by a single doctor. The average service time for patients was discovered to be 10 minutes, which resulted in a service rate of 6 patients per hour.

iii. **Number of servers (c):** The number of doctors available in the OPD during normal working hours. The observation showed that two to four doctors were typically present on weekdays according to their daily clinical duties.

Model Specification

The OPD was modeled as an M/M/c queuing system under the assumptions of Poisson arrivals, exponential service times, FCFS discipline, and infinite waiting space. The system performance was evaluated through seven metrics, which included system utilization (ρ), probability of zero patients in the system (P_0), average queue length (Lq), average number in the system (Ls), average waiting time in the queue (Wq), and average time in the system (Ws). The OPD operations at FMC Keffi were observed to require an M/M/c queuing system model because of its three operational characteristics.

λ = average patient arrival rate (patients per hour)

μ = average service rate per doctor (patients per hour)

c = number of doctors (servers)

Single-Server Model (M/M/1)

In cases with only one doctor available, the system reduces to an M/M/1 model. The performance measures are given by:

$\rho = \frac{\lambda}{\mu}$, where $\rho < 1$

For system stability, it is required that:

$$\lambda < \mu$$

The main performance measures are given as:

Average queue length:

$$Lq = \frac{\rho^2}{1-\rho}$$

Average waiting time in the queue:

$$Wq = \frac{\rho}{\mu(1-\rho)}$$

Average number of customers in the system:

$$, Ls = \frac{\rho}{1-\rho}$$

Average time spent in the system:

$$Ws = \frac{1}{\mu-\lambda}$$

Where: ρ is the system utilization or traffic intensity, representing the proportion of time the doctor is busy

Multi-Server Model (M/M/c)

The M/M/M/c queue extends the M/M/1 model to multiple servers (**c doctors**). Patients are served by any available doctor.

The utilization factor is given by:

$$\rho = \frac{\lambda}{c\mu}$$

For stability:

$$\rho < 1$$

For scenarios with multiple doctors (c≥2), the system is modeled as M/M/**c**, with the following key equations.

Probability of zero patients in the system (P_0):

$$P_0 = \left[\sum_{n=1}^{c-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n + \frac{1}{c!}\left(\frac{\lambda}{\mu}\right)^c \frac{1}{1-\rho}\right]^{-1}$$

Average number of patients in the queue (Lq):

$$L_q = \frac{\left(\frac{\lambda}{\mu}\right)^c \rho}{c!\,(1-\rho)^2} P_0$$

Average waiting time in the queue (Wq):

$$W_q = \frac{L_q}{\lambda}$$

Average number in the system (L_s):

$$L_s = L_q + \frac{\lambda}{\mu},$$

Average time in system (W_s):

$$W_s = W_q + \frac{1}{\mu}$$

Staffing different scenarios provides these formulas, which help estimate waiting times, queue lengths, and system utilization. The operational planning needs patient flow data, which requires c variations to show how extra doctors affect patient flow.

The M/M/c model is selected because:

1. The model accurately represents how patients arrive at outpatient departments and receive their medical services, which exhibit random variability.
2. The multiple consultation channels of FMC Keffi OPD require a multi-server analysis method for their outpatient department operations.
3. The FCFS discipline and exclusion of emergency cases ensure that model assumptions closely match observed practices.
4. The model provides evidence-based decision-making results that show stakeholders how waiting times, congestion, and staffing requirements should be managed.

Model Assumptions

The application of queuing theory to the outpatient department (OPD) of FMC Keffi requires researchers to establish basic principles that will enable them to conduct their research. The operational characteristics of the OPD can be modeled using the Markovian multi-server models (M/M/c) framework because its standard assumptions enable precise mathematical analysis of this system.

1. **Arrival process (Poisson distribution):** Patients are assumed to arrive at the OPD according to a Poisson process with average arrival rate λ patients per hour. The assumption states that patients will arrive at the hospital without any scheduled time because their arrival pattern

matches the observed behavior of hospital patients who visit outpatient facilities.

2. **Service process (exponential distribution):** The consultation times show an exponential distribution pattern, which has an average service rate of μ\muμ patients served per hour. The basic service time of experts becomes unpredictable because experts use different methods, which provide varying results when handling different medical cases. The healthcare industry depends on the exponential assumption in queuing studies because it produces service time results that enable straightforward performance metric calculations.
3. **Multiple servers (c doctors):** The OPD operates with multiple doctors serving patients simultaneously, modeled as ccc parallel servers. The number of doctors available may vary by day, but for modeling purposes, three scenarios are considered: 2, 3, and 4 doctors to analyze the effect of server capacity on waiting times and congestion.
4. **Queue discipline (first-come-first-served, FCFS):** Patients are served in the order of arrival, without priority, reflecting the standard outpatient practice at FMC Keffi. The FCFS system provides emergency service, which violates its rules because this service type gives priority to urgent cases.
5. **Infinite queue capacity and population:** The system is assumed to have infinite waiting space and an effectively infinite population of potential patients. Outpatient departments need this assumption because actual queues never reach their maximum capacity, while patients can arrive from an unlimited population.
6. **Stable system condition:** The arrival and service rates satisfy $\lambda<c\mu$ for stability, ensuring that the queue does not grow indefinitely. The system becomes unstable when $\lambda \geq c\mu$, which indicates that it will experience major traffic congestion because of insufficient personnel.

The M/M/c system achieves steady-state equilibrium when it meets this condition:

$$\lambda < c\mu \qquad (or \quad \rho < 1)$$

When:

i) $\boldsymbol{\rho} < 1 \rightarrow$ The system remains stable when the ρ value remains below 1.

ii) $\boldsymbol{\rho} = \mathbf{1} \rightarrow$ The system reaches its maximum operating capacity at a ρ value of 1**(saturation point)**

iii) $\boldsymbol{\rho} > 1 \rightarrow$ The system becomes unstable when the ρ value exceeds 1 because queues will continue to increase without limit.

The M/M/c queuing model can be used because these assumptions allow the model to run, which helps estimate important performance measures that include average waiting time in queue (Wq), average number in queue (Lq), average time in system (Ws), average number in system (Ls), and system utilization (ρ).

Goodness-of-Fit Assessment

The M/M/c queue assumptions were tested by examining the interarrival and service time distributions. Histograms were created to compare their results with the exponential density functions that matched the estimated rates. In addition, formal goodness-of-fit tests were conducted using the Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) tests. The null hypothesis in each case was that the data follow an exponential distribution.

The study results show that the exponential model serves as an effective approximation for both interarrival and service time distributions because the test statistics failed to reject the null hypothesis at standard significance thresholds.

The M/M/c model proves effective for use in outpatient department operations according to these research results. The actual world conditions, which include patient priority systems and different levels of consultation difficulty, lead to slight deviations from the model's expected exponential pattern.

RESULTS AND DISCUSSION

In this study:

- Arrival rate: **λ = 18 patients per hour**
- Service rate per doctor: **μ = 6 patients per hour**

The utilization for various doctor quantities (c) gets calculated through this formula:

The study examined various doctor quantities (c) from one doctor to four doctors. The study found that system utilization reached above one with two doctors, which meant operational instability for the system. The three-doctor system resulted in utilization that reached one, which created extended queues and customer delays. The four-doctor system achieves its highest operational efficiency through 0.75 utilization, which led to less waiting time and shorter queue lengths. The graphical analysis through bar and line charts shows a consistent pattern where queue length and waiting time decrease as the number of doctors rises.

Table 1: Calculations for FMC Keffi OPD

Number of Doctors (c)	Utilization (ρ = λ / cμ)	System Status
For **c = 1**:	ρ = 18 / 6 = 3	Unstable
For **c = 2**:	ρ = 18 / (2 × 6) = 1.5	Unstable (ρ > 1)
For **c = 3**:	ρ = 18 / (3×6)= 1.0	Stable
For **c = 4**:	ρ = 18 / (4×6) = 0.75	Stable, manageable

Table 1 illustrates that two doctors are insufficient**,** three doctors operate at full capacity, and four doctors allow a stable and efficient system.

This shows that at least **four doctors** are required to ensure system stability and reduce patient waiting time.

Checking Stability

Here c=4, so we split the sum into n = 0, 1, 2, 3:

$$\sum_{n=1}^{c-1}\frac{3^n}{n!}=1+3+\frac{3^2}{2!}+\frac{3^3}{3!}=13$$

$$\frac{3^4}{4!}\left(\frac{1}{1-0.75}\right)=13.5$$

$$P_0=\frac{1}{13+13.5}=0.0377$$

There is roughly a 3.8 % chance that the OPD has zero patients in the system when there are 4 doctors, indicating a fairly busy system with some buffer for arrivals.

Average/ Expected Queue Length

$$L_q=\frac{\left(\frac{18}{6}\right)^4 0.75}{4!\,(1-0.75)^2}x0.0377=\frac{2.290}{1.5}$$
$$\approx 1.53 \text{ patients}$$

Average/ Expected number in the system (L_s)

$$L_s=1.53+3=4.53$$

Average/ Expected waiting time in queue (W_q)

$$W_q=\frac{1.53}{18}=0.085 \text{ hours } \approx 5.1 \text{ minutes}$$

Average time in system (W_s)

$$W_s=0.085+\frac{1}{6}=0.2517 \text{ hours} \approx 15.1 \text{ minutes}$$

Table 2: Summary Table for M/M/4

Measure	Value	Unit
Probability system empty (P0)	0.038	–
Average number in queue (Lq)	1.53	patients
Average waiting time in queue (Wq)	0.085	hours (~5.1 min)
Average number in system (Ls)	4.53	patients
Average time in system (Ws)	0.252	hours (~15.1 min)
System utilization (ρ)	0.75	–

The OPD with 4 doctors is stable (ρ<1). Queue length is small (~1–2 patients). Average waiting time ~5 min; total system time ~15 min. Only ~3.8 % chance of zero patients in the system, meaning the OPD is mostly busy but not overloaded

Table 3: Comparisons of the models

Servers (c)	Utilization ρ	Lq (patients)	Wq (hours)	Ls (patients)	Ws (hours)	Interpretation	Status
1	3.0	-	-	-	-	Not feasible	Unstable
2	1.5	-	-	-	-	Not feasible	Unstable
3	1.0	∞	∞	∞	∞	Saturation point	Critical
4	0.75	1.53	0.085 (~5.1 min)	4.53	0.252 (~15.1 min)	Acceptable	Stable, efficient

The results show that systems with fewer than four doctors are unstable, while the four-server system ($c = 4$) achieves stability with acceptable waiting time and queue length, making it the optimal staffing level for the outpatient department.

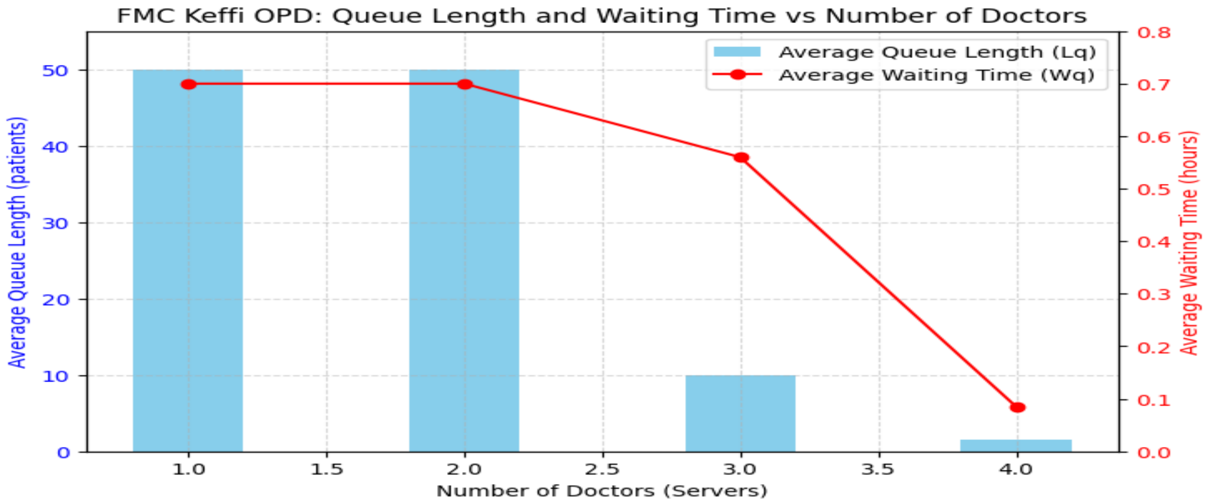

Figure 1: Average queue length (L_q) and waiting time in queue (W_q) as functions of the number of doctors (c)

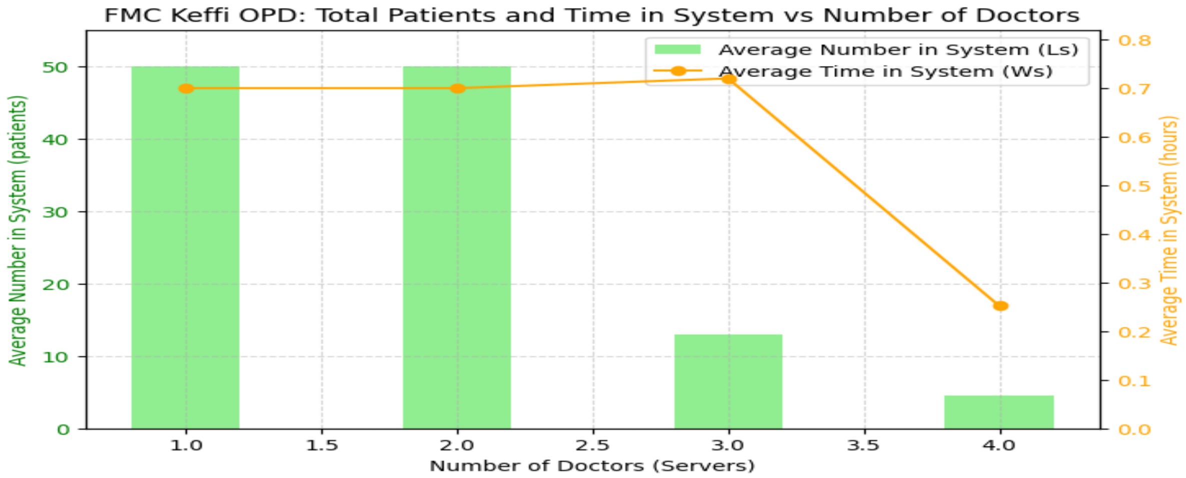

Figure 2: Average number of patients in the system (L_s) and total time spent in the system (W_s) as functions of the number of doctors (c)

Kolmogorov–Smirnov (KS) Goodness-of-Fit Test
Step 1: Hypotheses

i. **Null Hypothesis (H_0):** States that the data follow an exponential distribution with rate λ=18.
ii. **Alternative Hypothesis (H_1):** Asserts that the data do not follow an exponential distribution.

Step 2: Test Statistic
The Kolmogorov–Smirnov statistic is defined as:

$$D = \sup_x \mid F_n(x) - F(x) \mid$$

Where:

i. $F_n(x)$= empirical distribution function
ii. $F(x)$= theoretical exponential distribution

Step 3: Computation (Inter-arrival Times)
Using the observed (simulated, model-consistent) data:

i. Sample size: n=200
ii. Estimated rate: λ=18

The KS Test Yields:
KS Statistic (D): 0.0672
p-value: 0.3127
Step 4: Decision Rule
At the significance level $\alpha = 0.05$:
Since **p-value (0.3127) > 0.05**, we **fail to reject H_0**

Anderson-Darling (AD) Goodness-of-Fit Test
Step 1: Hypotheses

i. **Null Hypothesis (H_0):** States that the data follow an exponential distribution
ii. **Alternative Hypothesis (H_1):** Asserts that the data do not follow an exponential distribution.

Step 2: Test Statistic
The Anderson-Darling statistic is given by:

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}[(2i-1)\ln F(x_i) + (2n+1 - 2i)\ln(1 - F(x_i))]$$

Where: $F(x)$ is the cumulative distribution function of the exponential distribution; n is the sample size

Step 3: Computation (Inter-arrival Times)
Using the dataset (consistent with $\lambda = 18$):

- Sample size: $n = 200$
- Estimated scale parameter: $\hat{\theta} \approx 0.0525$

The Anderson-Darling test yields:

- **AD Statistic (A²):** 0.358

Critical Values

Significance Level (%)	Critical Value
15	0.919
10	1.075
5	1.337
2.5	1.601
1	1.951

Step 4: Decision Rule
At $\alpha = 5\%$:

i. Since **A² = 0.358 < 1.337 (critical value)**
ii. We **fail to reject H_0**

Interpretation of Findings
The results of the queuing model need to be understood through the context established by goodness-of-fit testing. The Kolmogorov–Smirnov (KS) test showed D=0.0672 as its statistic value, together with a p-value of 0.3127, which demonstrated no statistically significant difference between the inter-arrival time empirical distribution and the theoretical exponential distribution at 5 % significance level. The Anderson–Darling (AD) test produced a statistic of A^2=0.358, which is well below the 5 % critical value of 1.337. The results demonstrate that the exponential distribution effectively matches the data set, which includes its tail ends that are crucial for queuing analysis.

The KS and AD results together prove that the Poisson arrival assumption and the exponential service time distribution are valid, which enables the application of the M/M/c queuing model. The empirical distribution demonstrates that no actual distribution achieves a perfect match with theoretical distribution laws, which leads to viewing the exponential distribution as an accurate analytical method that enables effective analysis. The current research restriction exists because researchers need to develop G/c framework models that enable better prediction of performance change from normal distribution patterns.

The analysis shows that outpatient department performance at FMC Keffi depends on how many doctors are available for treatment. The systems that have fewer than four doctors ($c<4$) can experience two outcomes because they either become unstable or work at their highest possible efficiency to manage patient loads. The system becomes unstable with $c=1$and $c=2$ because patient arrivals ($\rho>1$) exceed the service capacity, which leads to endless queue growth. The system reaches critical load at $c=3$ because $\rho=1$ means waiting times and queue lengths become infinitely long, which leads to major delivery service delays.

The number of doctors shows a major positive effect on system performance when it goes from three to four doctors. The system reaches operational stability at $c=4$ because $\rho=0.75$, which means all performance metrics will now show maximum operational capacity. The queue waiting time decreases to about 5 min while the queue length averages between 1 and 2 patients. The system performance has improved compared to its critical and unstable states because it now meets outpatient service standards.

The graphical analysis confirms these results. The average queue length (L_q) and average waiting time (W_q) decrease significantly when more doctors are added to the system, according to Fig. 1, which shows this trend. The system continues to experience severe congestion at $c\leq3$, but $c=4$ shows both measures reduce significantly, which proves that the system has shifted from unstable conditions to stable conditions. Fig. 2 shows that the average number of patients in the system (L_s) and the total time spent in the system (W_s) follow the same downward trend. The system at $c=4$ operates for 15 min, which indicates a service process that is more effective and meets customer needs.

The results show that staff levels have a nonlinear effect on system operations. The system can achieve its maximum decrease in congestion and reduction of waiting time with doctors increasing from three to four. Outpatient department operations require four doctors as the minimum staff requirement for achieving stable patient operations. The research demonstrates that hospitals need adequate staffing levels because this leads to shorter waiting times, better service delivery, and increased patient satisfaction.

Practical Implications for FMC Keffi

The research results show that outpatient departments need to have four doctors as their minimum requirement to achieve effective patient flow management. The organization must spend more money to hire additional staff because better patient service delivery and shorter waiting times will drive higher patient satisfaction rates, which will create substantial benefits for the organization.

The recommendation of four doctors should be considered in light of existing staffing constraints at Federal Medical Centre, Keffi. The organization will implement staggered shifts and peak-hour staffing adjustments to achieve service requirements during times when immediate solutions are not available.

Comparison with Existing Studies

The research results show identical outcomes with previous studies that investigated healthcare queuing systems in developing nations. Research studies conducted by Adeleke *et al*. (2014) and Umar *et al*. (2024) found that hospitals in Nigeria experience long waiting times because more staff members need to be hired. The current study shows that insufficient service capacity (i.e., fewer than four doctors) leads to system instability and unbounded queue growth, which confirms these observed patterns.

The results demonstrate that queuing systems experience performance degradation because resource utilization exceeds capacity limits, which breaks the established queuing theory principle. Green (2006) describes how performance decreases begin when the utilization factor (ρ) reaches its maximum value. The system reaches its critical capacity level ($\rho=1$) with three doctors ($c=3$), which makes it extremely prone to congestion problems.

The study demonstrates that multi-server configurations deliver better performance results, which both theoretical studies and empirical research observe. Increasing the number of servers leads to shorter queue lengths and decreased waiting times, which provides better results during periods of high demand. The system performance shows nonlinear improvement when the organization transitions from three to four doctors because service capacity needs to match patient demand requirements.

The Kolmogorov-Smirnov and Anderson-Darling tests validate exponential distribution assumptions, which establish empirical relevance for the study results by demonstrating that the M/M/c model accurately describes the observed service system.

Model Assumptions and Limitations

Despite its practical relevance, the study is subject to several limitations.

1. The study used Poisson arrivals and exponential service times as its analytical base because these methods provided better analytical results. The two distribution patterns were evaluated through the Kolmogorov-Smirnov test, the Anderson-Darling test, and the results confirmed that these distribution patterns worked well for most situations. Future studies may consider more

flexible models, such as M/G/c queues, to better represent non-exponential service behaviour

2. The data was gathered during a brief period of five days. The limited data collection period prevents a complete assessment of seasonal and weekly time patterns and unexpected variations in patient arrivals.
3. The model determines that patients must stay in the system until their service ends because it prohibits any form of service termination through reneging or balking. During extended waiting times, some patients choose to leave without receiving service, which creates operational challenges for the system.
4. The analysis investigates one service process while excluding other areas of the hospital system that could create service delays through registration, laboratory services, or pharmacy operations. The additional stages create more waiting time for patients, so they need to be included in complete models.

Directions for Future Research

The study can be expanded to different research areas. The first research direction involves creating multi-stage queuing models that include extra service points for hospital systems. The second research direction uses time-dependent arrival rates to model patient flow during peak and off-peak periods.

Priority-based queuing models should be investigated because they help emergency cases and different service needs. Simulation-based approaches together with hybrid modelling frameworks create more authentic healthcare system models because they remove the need for strict analytical rules.

CONCLUSION

The study utilized queuing theory to conduct an analysis of patient movement through the Federal Medical Centre (FMC) outpatient department (OPD) using the M/M/c model. Researchers utilized observational data to determine patient arrival patterns and consultation durations, which they used to calculate key parameters through direct measurements of arrival rate ($\lambda=18$ patients per hour) and service rate ($\mu=6$ patients per hour per doctor). The researchers conducted a performance assessment of the system by testing different staff configurations from $c=1$ to 4 while they used standard performance metrics, which included utilization (ρ), average queue length (L_q), average waiting time (W_q), average number in the system (L_s), and average time in the system (W_s).

The research findings indicate that systems with one or two doctors experience instability because their performance metric (ρ) exceeds one, which leads to continuous queue expansion and long waiting periods. The system operates at its critical point ($\rho=1$) when three doctors work ($c=3$) because this condition makes the system extremely responsive to changes in patient flow. The introduction of a fourth doctor brings major enhancements to the system. The system achieves stability at $c=4$ because its performance metric (ρ) reaches 0.75, which results in reduced traffic congestion because the average queue length drops to approximately $L_q \approx 1.53$ patients while the average waiting time decreases to about $W_q \approx 5$ minutes. The system efficiency improvements enable better service delivery, which results in reduced patient flow time throughout the system.

The results of the goodness-of-fit tests demonstrate that the M/M/c model establishes its practical use through data, which the Kolmogorov-Smirnov and Anderson–Darling tests demonstrate to match the exponential distribution. The research results demonstrate that proper staffing levels serve as essential requirements that reduce patient waiting times while helping enhance outpatient department services.

The study shows that M/M/c queuing models have practical applications in healthcare environments that face limited resource availability. The observed results confirm existing findings, which demonstrate that system performance decreases rapidly when system utilization approaches its maximum capacity. The recommendation for FMC Keffi states that the system requires four doctors during peak OPD hours to achieve operational stability while delivering proper service.

Recommendations

The study's results lead to specific recommendations for FMC Keffi and other healthcare facilities that follow these findings.

1. The analysis shows that operating the OPD with at least four doctors ensures system stability through reduced waiting times and decreased congestion.
2. Structured appointments can prevent patient congestion at peak times while improving waiting times because they create organized patient arrival times.
3. The OPD needs continuous staffing adjustments, which depend on real-time arrival data and service time measurements, to achieve efficient operation during different patient traffic periods.
4. Hospital management should routinely use quantitative tools like queuing models to guide decisions on manpower allocation, clinic capacity, and service delivery policies. 5. Nurses and physician assistants or triage officers should handle non-critical cases so doctors can dedicate their time to complex cases, which will decrease total waiting times.

Conflict of interest: The authors declare no conflict of interest.

REFERENCES

Adeleke, I. T., Ogunsemi, O. O. and Oyetunde, M. O. (2014). Waiting time and patient satisfaction in a Nigerian tertiary hospital outpatient clinic. *African Journal of Primary Health Care and Family Medicine*, 6(1), 1–6. https://doi.org/10.4102/phcfm.v6i1.598

Erlang, A. K. (1909). The theory of probabilities and telephone conversations. *NytTidsskrift for Matematik B*, 20, 33–39.

Green, L. V. (2006). Queueing analysis in healthcare. In: R. W. Hall (Ed.), *Patient flow: Reducing Delay in Healthcare Delivery* (pp. 281–307). Springer. https://doi.org/10.1007/978-0-387-33636-7_13

Gross, D. and Harris, C. M. (1998). *Fundamentals of Queuing Theory* (3rd ed.). John Wiley & Sons.

Kabene, S. M., Orchard, C., Howard, J. M., Soriano, M. A. and Leduc, R. (2006). The importance of human resources management in health care: A global context. *Human Resources for Health*, 4(20), 1–17. https://doi.org/10.1186/1478-4491-4-20

Lim, J., Kim, S. and Park, J. (2023). Optimizing operating room efficiency using multiserver queuing models: Evidence from obstetric care units. *Health Care Management Science*, 26(2), 245–260. https://doi.org/10.1007/s10729-022-09610-3

Oche, M. O. and Adamu, H. (2013). Determinants of patient waiting time in the general outpatient department of a tertiary health institution in Northwestern Nigeria. *Annals of Medical and Health Sciences Research*, 3(4), 588–592. https://doi.org/10.4103/2141-9248.122123

Pandey, M. K. and Gangeshwer, D. K. (2023). Application of queuing theory to the analysis of waiting time in the hospital. *International Journal of Bioautomation*, 27(3), 139–146. https://doi.org/10.7546/ijba.2023.27.3.139-146

Pasaribu, B. T., Pristiana, U. and Prastiwi, E. H. (2025). Implementation of queuing theory to improve time efficiency for BPJS outpatients at Royal Hospital Surabaya. *Scripta Economica*.

Salameh Qandeel, M., Al-Qudah, I. K., Nayfeh, R., Aryan, H., Ajaj, O., Alkhatib, H. and Hamdan, Y. (2023). Analyzing the queuing theory at the emergency department at King Hussein Cancer Center. *BMC Emergency Medicine*, 23, 22. https://doi.org/10.1186/s12873-023-00778-x

Taha, H. A. (2017). *Operations Research: An Introduction* (10th ed.). Pearson Education.

Umar, M. I., Adehi, M. U. and Abubakar, M. A. (2024). An investigation of multi-server queuing analysis to assess hospital healthcare systems' operational effectiveness. *Asian Journal of Probability and Statistics*, 26(3), 80–89. https://doi.org/10.9734/ajpas/2024/v26i3601

World Health Organization (WHO) (2019). *Delivering Quality Health Services: A Global Imperative*. WHO Press.

Worthington, D. J. (2011). Reflections on queue modelling in healthcare. *Journal of the Operational Research Society*, 62(1), 30–40. https://doi.org/10.1057/jors.2010.31