

## Application of Queuing Theory Models in Managing Patient Waiting Time: A Case Study of General Hospital Wudil, Kano State, Nigeria

Shamsuddeen Ahmad Sabo<sup>1\*</sup>, Umar Aliyu<sup>1</sup>, Umar Muhammad Sahalu<sup>2</sup>, Ahmed Shuaibu<sup>1</sup>

<sup>1</sup>Department of Statistics, Aliko Dangote University of Science and Technology, Wudil, Nigeria

<sup>2</sup>Department of Statistics, Shehu Idris College of Health Science and Technology Makarfi-Nigeria

\*Corresponding Author's Email & Phone No.: [shamsuddeenahmadsabo@gmail.com](mailto:shamsuddeenahmadsabo@gmail.com) ; +234-8067886262

### ABSTRACT

This study examined the application of queuing theory to manage patient waiting times at General Hospital Wudil, Kano State, Nigeria. The goal was to identify service inefficiencies and propose strategies for improving healthcare delivery. Primary data on patient arrivals, waiting times, and service durations were collected and analyzed using both  $M/M/1$  (single-server) and  $M/M/S$  (multi-server) queuing models. Preliminary model assumption checks, conducted through graphical plots, confirmed that both patient inter-arrival and service times follow exponential distributions. This supports the use of Poisson arrivals and exponential service times in the models. Furthermore, a Kolmogorov-Smirnov test indicates that both arrival and service times are indeed exponentially distributed, as the  $p$ -value is below the 5% significance level. The findings from the operating characteristics measures indicated that the single-server model resulted in high congestion, with a utilization rate of approximately 86%, an average queue length of about 5 patients, and a mean waiting time of nearly 14 minutes. This situation led to significant patient dissatisfaction and delays. In contrast, the multiple-server model exhibited improved performance, with utilization reduced to 77%, the average queue length decreasing to just 2 patients, and the average waiting time dropping to about 6 minutes. The study concludes that patient waiting time is strongly influenced by server capacity. Increasing the number of service points substantially reduces congestion and enhances efficiency. It recommends that hospital management adopt flexible staffing, expand service points, and implement queue management systems to optimize resource use and improve patient satisfaction.

**Keywords:** Queuing theory, Queuing management, Service efficiency, Patients satisfaction,  $M/M/1$  model, and  $M/M/S$  model.

### 1.0 INTRODUCTION

Queuing is a universal phenomenon that occurs in various aspects of everyday life, from banks and supermarkets to fuel stations and hospitals. It arises primarily due to the limited capacity of service facilities to meet fluctuating demand, often resulting in delays and congestion. These delays are not only common in daily life but also critical in more technical and industrial settings such as manufacturing systems, computer networks, telecommunications, fuel station and hospitals (Ugba et al., 2025). According to (Obamiro, 2013), queuing is often perceived as frustrating by customers, especially when delays are frequent or unpredictable. As a result, efficient queue management is essential to guaranteeing both operational effectiveness and a satisfying customer experience. Queuing models are crucial for developing and refining service systems to increase productivity and customer satisfaction, as (Kandemir & Cavas, 2007) pointed out. Recent studies emphasize the growing importance of queuing theory in managing congestion in service-oriented systems. For instance, (Li, et al., 2021) applied queuing models to optimize service flow in public transport systems, while (Adamu & Olalekan, 2020) demonstrated its relevance in improving customer experience in Nigerian banking halls. Similarly, (Isah & Yusuf, 2022) highlighted the

usefulness of queuing theory in minimizing waiting time and enhancing throughput in fuel dispensing stations during peak demand periods. (Abdulaziz & Muhammad, 2022) use the queuing concept to speed up patient care. According to (Shanmugasundaram & Umarani, 2015), waiting time is the amount of time a patient spends in line prior to the start of services. Sadly, this situation occurs at a lot of teaching hospitals. Every day, teaching hospitals receive a lot of patients, which typically leads to lengthy wait times for patients (Afrane & Appah, 2014). In every context where patients are involved, including hospitals, petrol stations, department stores, restaurants, and banks, waiting for service is a typical occurrence (Adewole, 2016). Whenever patients arrive at a service facility, some of them have to wait before they receive the desired service. It implies that the patient must wait for their turn. Patients may get dissatisfied because they are unable to receive the desired level of service when they must wait a long time in line to be served (Nkrumah et al., 2015). In a similar vein, some people pass away while awaiting their turn to receive care. When service demand exceeds available capacity, service is delayed. When capacity is not enough to meet demand, the system will generate a wait (Ahmed & Ali, 2021).

Segun (2020) used queuing theory to examine healthcare service performance at the Adekunle Ajasin University Health Center in Akungba-Akoko, Nigeria. In order to create a suitable queuing model, the study examined patient arrival, waiting, and service times. Direct observation was used to gather data over the course of three weeks during weekdays, and Microsoft Excel and Python-based simulation were used for analysis. The findings showed that there were times when the patient arrival rate was higher than the service rate, which resulted in long-term queue creation and more system congestion. The results indicate that service efficiency is greatly impacted by large patient input compared to service capacity. As a result, the study suggested ways to improve operational performance and shorten wait times for healthcare services.

Rotich (2016) evaluated the relationship between waiting costs and service costs by employing an M/M/s queuing model to estimate service efficiency in the emergency unit, based on patient arrival and service data. The findings indicated that the six ICU beds currently in use resulted in severe congestion, with an anticipated daily queue of 34 patients and significant waiting costs. Queuing theory has been extensively applied to study waiting lines and service performance in various systems. Previous research has demonstrated its effectiveness in assessing key performance metrics, such as waiting time, server utilization, and delay probability (Prieger, 2005; Okeke, 2003). Empirical applications, such as fuel station operations and traffic flow management, illustrate how queuing models have been used to enhance system efficiency and reduce congestion (Abebi et al., 2019). Increasing the number of active service pumps decreased waiting and service times, according to Odior's (2013) analysis of fuel station queuing patterns in Benin City. In Nigeria, (Kembe et al., 2017) created a Vehicle Queue Management System that employed real-time data to enhance client allocation and forecast service times, leading to increased flow efficiency and satisfaction. After observing petrol stations in Indonesia, (Lestari et al., 2016) and (Arighi, 2017) came to the conclusion that optimal attendant utilization (70–80%) balanced tiredness and efficiency. Sarla et al. (2020) examined lines outside urban gas stations using sensor-based data, concluding that inadequate wait management exacerbated traffic congestion and suggesting integrated digital monitoring and scheduling. Queue management, according to Lee (2019), is a collection of guidelines designed to improve customer flow and expedite the queuing process. Unmanaged queues are detrimental to the gainful operation of service systems and results in a lot of other managerial hitches (Chase, et al 2001). Customer satisfaction has been observed as a key strategy of every business and a benchmark against which many

organizations have set their standards (Burodo et al 2021).

Khaskheli et al. (2020) investigated the optimization of service capacity in the outpatient departments (OPDs) of two public hospitals in Sindh, Pakistan. The study, based on data collected over a two-week period, analyzed patient arrival and service rates, the number of servers, staff salaries, and patient waiting costs using a multi-server M/M/c queuing model. Recent research has also applied queuing models to various service systems. For example, (Simamora et al., 2025) utilized an M/G/4 queuing model, calibrated with field data, to study a congested fuel station in Mumbai. Their findings highlighted the effectiveness of queuing models in improving operational performance in real-world service settings, as empirical calibration of service distributions reduced average waiting time by nearly 22%. Dhibar (2025) examined the performance of a Markovian M/M/1 double-orbit queueing system with balking and imperfect service under both crisp and fuzzy environments. Using the probability generating function approach, the study derived explicit expressions for the queue size distribution and several important performance measures. Key metrics such as queue length, waiting time, and throughput were obtained analytically and their behavior was further illustrated through numerical experiments.

## 2.0 METHODOLOGY

This study adopted a quantitative research design using queuing theory models to analyze patient waiting time at the General Hospital Wudil, Kano State, Nigeria. Primary data were collected from the hospital's outpatient department through direct observation and structured recording of patient arrival, waiting, and service times over a four-week period. A total of 91 patients were included using purposive sampling to capture typical service flow conditions during peak hours. Data analysis was conducted using R statistical software version 4.5.0, applying both M/M/1 (single server) and M/M/S (multi-server) models to estimate performance metrics such as utilization rate, average queue length, waiting time, and system capacity. Model assumptions were validated using graphical exponential fits and the Kolmogorov Smirnov test to confirm exponential distributions for inter-arrival and service times. The analysis focused on steady-state conditions to ensure comparability of performance measures across different service configuration (Tyagi, et al., 2023)

### 2.1 Queuing Theory Models

The models adopted in this work are  $(M/M/1):(\infty/FCFS)$  and  $(M/M/S):(\infty/FCFS)$ , single and multi-server



queuing model respectively. These queuing systems assume that the arrival of patients follow a Poisson distribution at an average of  $\lambda$  patient per unit of time and are served on a first-come first-served basis. The service times, follows exponential distribution, with an average of  $\mu$  patient per unit of time and number of servers  $S$ . If there are  $n$  patients in the queuing system at any point in time, then the following two cases may arise:

1. If  $n < S$ , (number of patients in the system is less than the number of servers), then there will be no queue however,  $(S - n)$  a number of servers will not be busy. The combine service rate will be.  $\mu_n = n\mu$
2. If  $n \geq S$ , (number of patients in the system is more than or equal to the number of servers), then all servers will be busy and the maximum number of patients in the queue will be  $(S - n)$ . The combine service rate will be.

$$\mu_n = s\mu.$$

Kendall's notation is a standardized system in queuing theory used to describe and classify different types of queuing models. Introduced by British statistician David G. Kendall in 1953, this notation provides a concise way to represent the characteristics of queuing systems, facilitating analysis and comparison across various applications such as telecommunications, Banking, Hospitals, and customer service. The model has a general structure of  $\{A/B/C\};\{d/e/f\}$ . Where;  $A$  = Arrival (or inter-arrival) distribution,  $B$  = Departure (or service time) distribution,  $C$  = Number of service channels in the system (server),  $D$  = Service discipline,  $E$  = Maximum number of customers allowed in the system and  $F$  = Calling source of population (Gunes et al., 2021)

### 2.2 Operating Characteristics

The operating characteristics of a queuing system are key performance measures that describe how the system behaves under steady-state conditions. These metrics help assess the efficiency, service quality, and customer experience at the service facility. The average arrival rate, denoted by  $\lambda$ ,

1. 
$$\lambda = \frac{\text{Total arrivals}}{\text{Total time}}$$

2. 
$$\mu = \frac{1}{\text{Average service time}}$$
3. 
$$\text{Average service time} = \frac{\text{Total service time}}{\text{Total arrival}}$$
4. 
$$\rho = \frac{\lambda}{\mu} \text{ or } \rho = \frac{\lambda}{s\mu}$$
5. 
$$P_0 = 1 - \rho \text{ or } P_0 = \left[ \sum_{n=0}^{s-1} \frac{(\rho s)^n}{n!} + \frac{(\rho s)^s}{s!(1-\rho)} \right]^{-1}$$
6. 
$$L_s = \frac{\lambda}{\mu - \lambda} \text{ or } L_s = \frac{\rho(\rho s)^s P_0}{s!(1-\rho)^2} + \rho s$$
7. 
$$L_q = L_s - \frac{\lambda}{\mu} \text{ or } L_q = \frac{\rho(\rho s)^s P_0}{s!(1-\rho)^2}$$
8. 
$$W_s = \frac{1}{\mu - \lambda} \text{ or } W_s = \frac{(\rho s)^s P_0}{s!(1-\rho)^2 s\mu} + \frac{1}{\mu}$$
9. 
$$W_q = \frac{\lambda}{\lambda(\lambda - \mu)} \text{ or } W_q = \frac{(\rho s)^s P_0}{s!(1-\rho)^2 s\mu}$$
10. 
$$P = \frac{\rho P_0}{s!(1-\rho)} \text{ or } P = \frac{(\rho s)^s P_0}{s!(1-\rho)}$$
11. 
$$P' = 1 - P$$
12. 
$$P_n = \rho^n P_0 \text{ or } P_n = \begin{cases} \frac{1}{n!} \left( \frac{\lambda}{\mu} \right)^n P_0, & \text{if } n < s \\ \frac{1}{s! s^{n-s} \left( \frac{\lambda}{\mu} \right)^n P_0}, & \text{if } n \geq s \end{cases}$$
13. 
$$P(n \geq s) = \frac{1}{s!} \left( \frac{\lambda}{\mu} \right)^s \frac{s\mu P_0}{s\mu - \lambda}$$

### 2.3 Notations Used

The following notations are used in this paper which includes:

1.  $\lambda$  (Average Arrival Rate): The average number of patients arriving at the system per unit time.
2.  $\mu$  (Average Service Rate): The average number of patients that can be served per unit time.
3.  $\rho$  (Traffic Intensity or Utilization Factor): The ratio of arrival rate to service rate, i.e.,  $\rho = \lambda / \mu$ . It

- indicates the proportion of time the service facility is busy.
4. System Capacity ( $S$ ): The maximum number of servers that can be accommodated in the system
  5.  $P_0$  (Probability of Zero patients in the System): The probability that there are no patients the system (idle server).
  6.  $L_s$  (Average Number of patients in the System): The expected number of patients in both the queue and service facility.
  7.  $L_q$  (Average Number of patients in Queue): The expected number of patients waiting in the queue for service.
  8.  $W_s$  (Average Time a patients Spends in the System): The expected total time a patients spends in the system, including waiting and service time.
  9.  $W_q$  (Average Waiting Time in Queue): The expected time a patients spends waiting before service begins.
  10.  $P$  The probability that an arriving patient has to wait is the probability that all servers (service

- channels) are busy at the time of arrival, forcing the patient to queue. A high value of  $P$  indicates that most arriving patient experience delays, which may be a sign of insufficient service capacity.
11.  $P'$  The probability that an arriving patient does not have to wait (i.e., it finds at least one server free and enters service immediately) is the complement of the probability of queuing on arrival. It reflects the efficiency and responsiveness of the service system.
  12.  $P_n$  (Probability of n patients in the System): The probability that there are exactly  $n$  patients in the system at a given time.
  13. The probability that all servers are simultaneously busy indicates the likelihood that an arriving patient will find no available service channel and must therefore join the queue. This situation occurs when the number of patients in the system is greater than or equal to the number of servers.

### 3.0 RESULT AND DISCUSSION

This section presents the analysis and discussion of results including the sociodemographic features of patients, queuing metrics and model assumption checking.

### 3.1 Socio Demographic Features of Patients

Table 1: Patients' Socio-Demographic Information

GENDER	FREQUENCY	PERCENTAGE
MALE	38	41.8
FEMALE	53	58.2
TOTAL	91	
AGE	FREQUENCY	PERCENTAGE
≤ 30	28	30.8
31-50	31	34.1
51-60	12	13.2
ABOVE 60	20	21.9
TOTAL	91	
HIGHEST QUALIFICATION	FREQUENCY	PERCENTAGE
ISLAMIYYAH SCHOOL	20	22.0
PRIMARY SCHOOL	20	22.0
SECONDARY/TERTIARY SCHOOL	51	56.0
TOTAL	91	
MARITAL STATUS	FREQUENCY	PERCENTAGE
SINGLE	23	25.3
MARRIED	44	48.4
DIVORCED	24	26.4
TOTAL	91	
EMPLOYMENT STATUS	FREQUENCY	PERCENTAGE
CIVIL SERVANT	26	28.6

RETIRED	17	18.7
SELF EMPLOYMENT	17	18.7
OTHER BUSINESS	31	34.1
TOTAL	91	

Table 1 presents the socio-demographic characteristics of the 91 patients included in the study. In terms of gender distribution, 38 patients (41.8%) were male, while 53 patients (58.2%) were female. This indicates a higher representation of female patients in the study population. From a queuing perspective, the predominance of females suggests that healthcare demand and waiting lines in the facility may be more influenced by female patients, which could affect service planning, especially in gender-sensitive healthcare services. Regarding age distribution, 28 patients (31.3%) were aged 30 years and below, 31 patients (34.4%) were between 31–50 years, 11 patients (12.5%) were between 51–60 years, while 20 patients (21.9%) were above 60 years. The age structure shows that the majority of patients fall within the economically active age group (31–50 years), followed by younger adults ( $\leq 30$  years). This has implications for queuing patterns, as younger and middle-aged patients may have different arrival times and service expectations compared to elderly patients, who often require longer consultation times. With respect to educational qualification, 20 patients (21.9%) had attended Islamiyyah School, 20 patients (21.9%) had attained primary education, while the majority, 51 patients (56.3%), had secondary or tertiary education. The relatively high educational attainment among patients

may influence their awareness of health services and their tolerance for waiting times, thereby shaping queuing behavior in the facility. For marital status, 23 patients (25.5%) were single, 44 patients (48.4%) were married, and 24 patients (26.0%) were divorced. The dominance of married patients may reflect family-related healthcare decisions, which could impact arrival rates, particularly during peak clinic hours when family members attend together. In terms of employment status, 26 patients (28.1%) were civil servants, 17 patients (18.8%) were retired, and 17 patients (18.8%) were self-employed, while 31 patients (34.4%) engaged in other forms of business. This distribution indicates that a substantial proportion of patients are engaged in either formal or informal employment. From a queuing perspective, employed individuals may prefer specific consultation times (e.g., mornings or after work hours), thereby contributing to fluctuations in patient arrival rates and potential peak queues. The demographic profile reveals that the study population is predominantly female, largely within the age range of 31–50 years, and mostly with at least secondary-level education. Furthermore, the majority are married and economically active. These socio-demographic factors are likely to influence patient arrival patterns, service demand, and waiting times, which are critical elements in queuing analysis.

### 3.2 Queue Management

Table 2: Patients' Queue Management Satisfaction

STATEMENT	STRONGLY AGREE	AGREE	UNDECIDED	DISAGREE	STRONGLY DISAGREE
Queues are managed properly at the hospital	7(7.4)	33(35.9)	4(4.8)	40(44)	7(8)
Hospital staff are of great help in helping patients' queues	0(0)	25(28.1)	23(25.0)	34(37.5)	9(9.4)
I am satisfied with how patients are handled at the hospital while waiting in queue	0(0)	17(18.8)	11(12.5)	34(37.5)	29(31.3)

Table 2 presents patients' perceptions of queue management within the hospital. Regarding the statement "Queues are managed properly at the hospital," only 7 patients (7.4%) strongly agreed, and 33 patients (35.9%) agreed. In contrast, the majority, 40 patients (44.0%), disagreed, and 7 patients (8.0%) strongly disagreed. This suggests that a substantial proportion of patients feel that queues are not effectively managed in the hospital.

For the statement "Hospital staff are of great help in assisting patients in queues," 25 patients (28.1%) agreed, while 23 patients (25.0%) were undecided. However, a larger proportion of 34 patients (37.5%) disagreed, and 9 patients (9.4%) strongly disagreed. This indicates that many patients perceive inadequate staff support in managing queues, although a notable percentage remained neutral.

Concerning the statement "I am satisfied with how patients are handled at the hospital while waiting in queue," only 17 patients (18.8%) agreed, while 34 patients (37.5%) disagreed, and 29 patients (31.3%) strongly disagreed. This reflects a general level of dissatisfaction with the way patients are managed during waiting times.

Overall, the results reveal that patients perceive queue management at the hospital as ineffective, with limited staff involvement and a high level of dissatisfaction

concerning how queues are handled. From a queuing theory perspective, this indicates potential inefficiencies in the system, such as long waiting times, poor service discipline, or inadequate allocation of service capacity. Addressing these concerns would require both structural improvements in queue design and better staff engagement in managing patient flow.

### 3.3 Model Fit Assumptions

Before applying the queuing models, it is necessary to verify whether the data meet the fundamental assumptions of queuing theory, particularly the distributions of inter-arrival times and service times. The exponential distribution is a key assumption for the  $M/M/1$  and  $M/M/S$  models, where arrivals follow a Poisson process and service times are exponentially distributed.

Figure 1 presents the exponential fit for inter-arrival times, while the Kolmogorov-Smirnov (K-S) test results are shown in Table 3. The test statistic was  $D = 0.2352$  with a corresponding  $p$ -value = 0.1126. Since the  $p$ -value is greater than the 5% level of significance (0.05), we fail to reject the null hypothesis that the data follow an exponential distribution. This indicates that the inter-arrival times are consistent with an exponential distribution, thereby supporting the assumption of a Poisson arrival process.

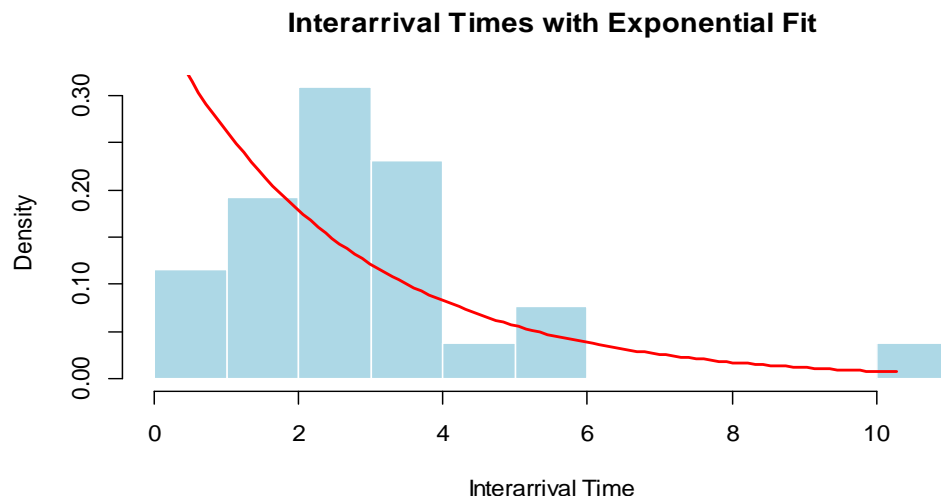


Fig 1: Inter arrival Time Exponential Fit

Table 3: One-sample Kolmogorov-Smirnov test of inter arrival time

Test Statistic	P-value
$D = 0.2352$	0.1126

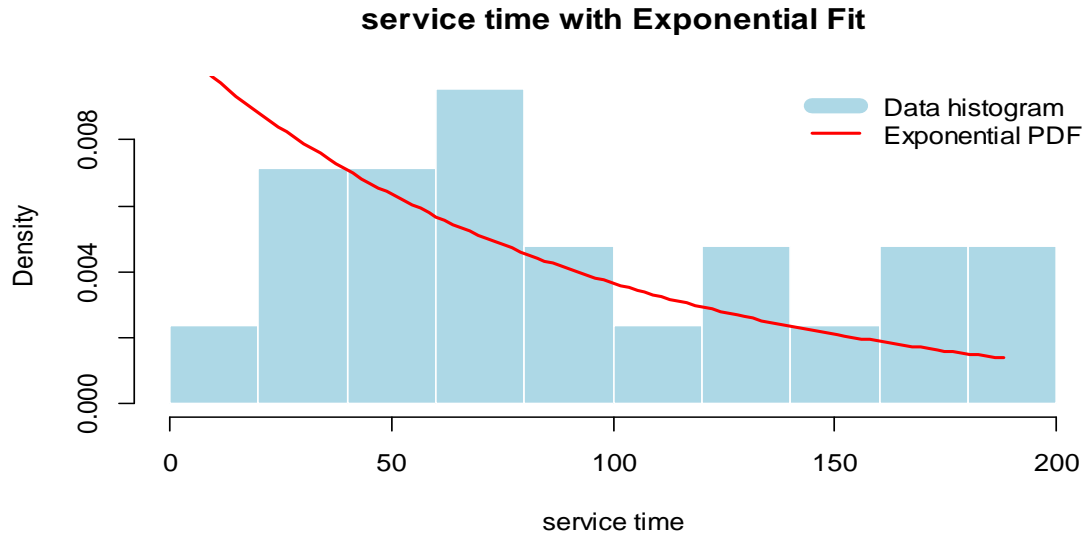


Fig 2 : Service Time Exponential Fit

Table 4: One-sample Kolmogorov-Smirnov test of service time

Test Statistic	P-value
D = 0.2132317	0.2564936

Similarly, Figure 2 illustrates the exponential fit for service times, with the corresponding K–S test results presented in Table 4. The test statistic was  $D = 0.2132$  with a  $p$ -value = 0.2565. In addition, the  $p$ -value exceeds the 5% significance level, providing no evidence against the null hypothesis. This implies that service times also follow an exponential distribution.

### 3.4 Queuing Metrics

Table 5: Summary Results of Performance Measures Single Server.

S/N	PARAMETER	VALUE	S/N	PARAMETER	VALUE
1	$(\lambda)$	0.398524	6	$(L_s)$	6.277121
2	$\frac{1}{\mu}$	2.164445	7	$(L_q)$	5.414538
3	$(\mu)$	0.4620123	8	$(W_s)$	15.75092
4	$(\rho)$	0.862583	9	$(W_q)$	13.58648
5	$(P_0)$	0.137417	10	$(P_n)$	0.7145926

Table 6: Summary results of performance measures of multiple server.

S/N	PARAMETER	VALUE	S/N	PARAMETER	VALUE
1	$(\lambda)$	0.3538928	8	$(W_s)$	10.46729
2	$\frac{1}{\mu}$	4.328096	9	$(W_q)$	6.139194
3	$(\mu)$	0.2310485	10	$(P)$	0.9635768
4	$(\rho)$	0.7658409	11	$(P')$	0.03642316
5	$(P_0)$	0.1326049	12	$(P_n)$	0.497877
6	$(L_s)$	3.704299	13	$(P(n \geq s))$	0.69523
7	$(L_q)$	2.172617			

Tables 5 and 6 presents the summary results of the queueing characteristics using the single and multi-servers (more than one doctor). The results obtained in table 5 and 6 are summarized in table 7.

### 3.5 Comparison of Single and Multi-Server Models

Table 7: Comparison of performance measures between single server and multiple server.

S/N	PERFORMANCE MEASURES	PARAMETERS	M/M/1	M/M/2
1	Arrival rate	$(\lambda)$	0.3985	0.3539
2	Service rate	$(\mu)$	0.4620	0.2310
3	Utilization factor	$(\rho)$	0.8626	0.7658
4	Average number of patients in the system	$(L_s)$	6.2771	3.7043
5	Average number of patients in the queue	$(L_q)$	5.4145	2.1726
6	Average waiting time of a patient in the system	$(W_s)$	15.7509	10.4673
7	Average waiting time of a patient in the queue	$(W_q)$	13.5865	6.1392

Table 7 compares the performance of a single-server ( $M/M/1$ ) system and a two-server ( $M/M/2$ ) system based on patient arrival and service data. For the  $M/M/1$  model, the utilization factor ( $\rho = 0.8626$ ) indicates that the single server is busy approximately 86% of the time, leaving very little idle capacity. Consequently, the system experiences a relatively high average queue length ( $L_q = 5.4145$  patients) and long waiting times ( $W_q = 13.5865$  minutes). The average time a patient spends in the system ( $W_s = 15.7509$

minutes) is also notably high, reflecting considerable delays in service delivery.

In contrast, the  $M/M/2$  model with two servers shows improved performance. The utilization factor per server reduces to  $\rho = 0.7658$ , and both the average number of patients in the queue ( $L_q = 2.1726$ ) and the waiting time in the queue ( $W_q = 6.1392$  minutes) are significantly lower than in the single-server case. Similarly, the average time a patient spends in the system decreases to  $W_s = 10.4673$  minutes.

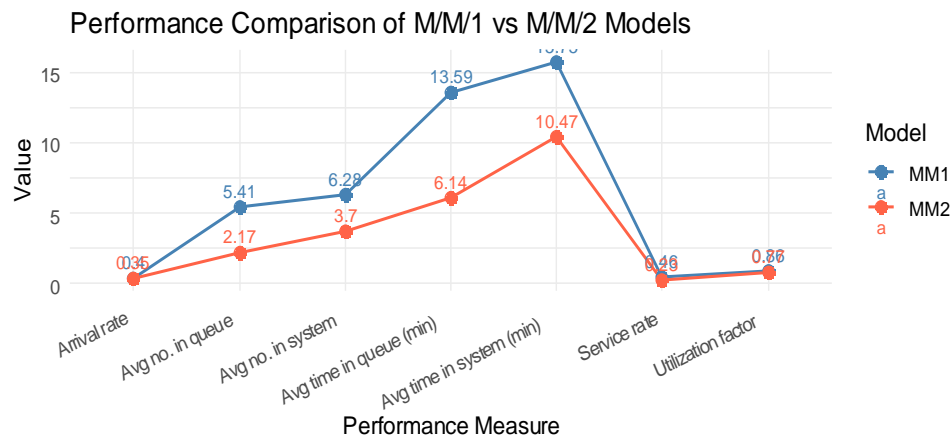


Fig 3: Performance Comparison of M/M/1 Vs M/M/2 Models

The comparison was further shown graphically in Fig 3. Overall, the comparison demonstrates that the two-server system (M/M/2) is more efficient in handling patient arrivals than the single-server system (M/M/1). While both models experience relatively high utilization, the multiple-server configuration reduces congestion, lowers waiting times, and enhances patient flow. This suggests that adopting a multi-server setup could substantially improve queue management and patient satisfaction in the hospital.

#### 4.0 CONCLUSION

This study concludes that patient waiting times at General Hospital Wudil are significantly influenced by the existing queuing structure. The single-server (M/M/1) system resulted in severe congestion and longed waiting times, while the introduction of a multi-server (M/M/2) configuration substantially improved service efficiency, reduced waiting times, and enhanced patient satisfaction, a similar result by (Daramola et al., 2024). However, the high probability of queuing upon arrival indicates that service demand still exceeds available capacity, especially during peak hours. Hospital administrators should adopt flexible staffing strategies and introduce digital queue management systems to monitor patient flow in real time. Expanding service points and increasing the number of healthcare providers during busy periods will further enhance operational performance. The study was limited to one hospital with a relatively small sample size; hence, findings may not be generalizable across all healthcare settings. Data collection was also limited to a specific observation period, which may not fully capture seasonal or random fluctuations in patients' arrivals. Further studies should incorporate simulation-based approaches and cost-

benefit analysis to evaluate the financial implications of additional servers. Moreover, hybrid models such as M/G/1 or M/D/1 could be applied to account for variability in service times and patient behavior, offering a broader insight into hospital queuing dynamics.

#### ACKNOWLEDGEMENT

The authors sincerely acknowledge the valuable contributions of the reviewers and the editorial team for their insightful comments and constructive suggestions, which have greatly enhanced the quality of this paper.

#### CONFLICT OF INETEREST

The authors declare no conflict of interest.

#### REFERENCES

- Afrane, S., & Appah, A. (2014). Queuing theory and the management of waiting time in hospitals: The case of Anglo Gold Ashanti Hospital in Ghana. *International Journal of Academic Research in Business and Social Sciences*, 4(2), 34–44. <https://doi.org/10.6007/IJARBS/v4-i2/590>
- Ahmed, O. I. A., & Abdelgadir, M. E. A. (2021). Applying waiting queue models to solve queue problems during the Corona pandemic in Khartoum State hospitals (Al-Hikma Hospital as a model). *International Journal of Innovation, Creativity and Change*, 15(3), 42–57.
- Abebi, S., Radfar, R., Rafati, M., & Hadadi, M. (2009). Using hybrid model with probability parameters to analysis queuing systems with layout constraints. In 2009 International Conference on Computers & Industrial Engineering (pp. 1243-1248). IEEE.
- Anokye, M., Abdul-Aziz, A. R., Annin, K., & Oduro, F. T. (2013). Application of queuing theory to vehicular traffic at signalized intersection in Kumasi-Ashanti Region, Ghana. *American International Journal of Contemporary Research*, 3(7), 23–29.

- Burodo, M. S., Suleiman, S., & Yusuf, G. (2021). An assessment of queue management and patient and satisfaction of some selected hospitals in North-Western Nigeria. *International Journal of Mathematics Statistics Invention (IJMSI)*, 9(8), 14-24.
- Chakraborty, P., & Madane, A. R. (2023). Vehicle queue management system at fuel stations. In K. Balasubramanian, V. Gunasekaran, & A. K. Singh (Eds.), *Proceedings of the International Conference on Communication, Electronics and Digital Technology* (pp. 617–628). Springer Nature Singapore. [https://doi.org/10.1007/978-981-99-1553-6\\_51](https://doi.org/10.1007/978-981-99-1553-6_51)
- Daramola, K., Yahaya, A., & Abdullahi, U. K. (2024). Evaluating single and multi-server exponential queuing models: A case study of Access Bank PLC in Anyigba, Kogi State, Nigeria. *Malaysian Journal of Science and Advanced Technology*, 4(4), 482-488. <https://doi.org/10.56532/mjsat.v4i4.360>
- Dhibar, S. (2025). MOHFDQ: A metaheuristic approach to optimizing hospital patient registration with a fuzzy double-orbit queueing model. *Swarm and Evolutionary Computation*, 98, 102090. <https://doi.org/10.1016/j.swevo.2025.102090>
- Gunes, F., Bayrakli, S., & Zaim, A. H. (2021). Smart cities and data analytics for intelligent transportation systems: An analytical model for scheduling phases and traffic lights at signalized intersections. *Applied Sciences*, 11(15), 6816. <https://doi.org/10.3390/app11156816>
- Jalalvand, F., Teimoury, E., Makui, A., Aryanezhad, M. B., & Jolai, F. (2011). A method to compare supply chains of an industry. *Supply Chain Management: An International Journal*, 16(2), 82-97.
- Janos, S. (2010). *Queuing Theory and its Application: A Personal view*. 8th International Conference on Applied Informatics, Eger: Hungary, 1, 9-30.
- Khaskheli, S. A., Marri, H. B., Nebhwani, M., Khan, M. A., & Ahmed, M. (2020). Comparative study of queuing systems of medical outpatient departments of two public hospitals. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, 1913, 2702–2720.
- Kalwar, M. A., Marri, H. B., Khan, M. A., & Khaskheli, S. A. (2021). Applications of queuing theory and discrete event simulation in health care units of Pakistan. *International Journal of Science and Engineering Investigations*, 10(109), 6-18
- Mala & Varma S.P. (2016). Minimization of Traffic Congestion by Using Queuing Theory. *IOSR Journal of Mathematics*, 12(1), 116-122.
- Michael, O. S. (2020). Performance Modelling of Healthcare Service Delivery in Adekunle Ajasin University, Akungba-Akoko, Nigeria Using Queuing Theory. *Journal of Advances in Mathematics and Computer Science*, 35(3), 119-127.
- Nkrumah, S., Yeboah, F. B., & Adiwokor, E. (2015). Client satisfaction with service delivery in the health sector: The case of Agogo Presbyterian Hospital. *International Journal of Business*, 3(2), 45–55.
- Okeke, A. O. (2003). *Quantitative methods for business decisions*. Enugu: Macro Academic Publishers.
- Obamiro, J. K. (2013). Effects of waiting time on patient satisfaction: Nigerian hospitals experience. *International Journal of Economic Behavior*, 3(1), 117–126.
- Onoja, A. A., Oguche, A. A., & Salihu, I. (2017). Application of queueing theory to customers purchasing premium motor spirit (PMS) at a filling station. *Statistics and Mathematical Sciences*, 3(10), 45–52.
- Odiar, A. O. (2013). Application of queuing theory to petrol stations in Benin-City area of Edo state, Nigeria. *Nigerian Journal of Technology*, 32(2), 325-332.
- Rotich, T. K. (2016). Utility analysis of an emergency medical service model using queuing theory. *British Journal of Mathematics & Computer Science*, 19(1), 1-18.
- Tyagi, M., Tyagi, P. K., Singh, S., Sathpathy, S., Kant, S., Gupta, S. K., & Singh, R. (2023). Impact of application of queuing theory on operational efficiency of patient registration. *medical journal armed forces india*, 79(3), 300-308. <https://doi.org/10.1016/j.mjafi.2021.06.028>
- Ugba, A., Wombo, A. B., & Agbulu, O. N. (2025). Assessment Of Soybean Production Skills Needed By Agricultural Science Teachers For Teaching Secondary School Students In Benue State, Nigeria.
- Simamora, R., Sianturi, R., Junarti, J., & Simarmata, G. (2025). Optimization of Service at Public Fuel Filling Stations (SPBU) Using Multiple Channel Single Phase Queue Model. *CAUCHY: Jurnal Matematika Murni dan Aplikasi*, 10(1), 415-427.

## APPENDICES

Table A1: The data for queuing process at 8:00am to 9:00am, on Monday using single server/consultant

CUSTOMERS	A.T (min:sec)	I.T (min:sec)	W.T (min:sec)	S.T (min:sec)	D.T (min:sec)
1	02:22	00:00	16:45	2:14	20:21
2	05:32	3:10	12:24	3:07	21:03
3	07:44	2:12	10:27	3:01	21:22
4	09:33	1:90	19:29	2:30	31:29
5	11:01	1:70	18:58	2:13	32:12
6	13:05	2:04	23:32	2:07	38:44
7	16:34	3:29	20:57	2:21	39:12
8	18:05	5:11	21:20	1:07	40:32
9	25:33	10:28	19:44	2:02	47:19
10	28:47	3:14	19:20	1:34	49:01
11	33:20	5:13	16:14	1:16	50:5
12	36:11	3:31	14:14	2:01	52:26
13	37:04	0:93	19:30	2:14	58:48
14	38:12	1:08	17:55	1:23	56:9
15	39:53	1:41	18:22	1:16	59:31
16	41:23	1:70	19:22	1:24	62:09
17	43:17	3:70	18:47	1:54	63:18
18	45:34	2:17	15:26	3:01	64:01
19	46:45	1:12	16:17	2:19	65:21
20	49:16	3:11	12:11	3:12	66:39
21	52:19	3:03	12:57	2:57	67:33
22	54:33	2:14	19:24	3:34	76:91
23	56:56	2:23	16:55	2:58	78:69
24	57:25	0:69	18:44	3:24	78:93



25	57:54	0:29	22:49	1:01	81:04
26	58:11	0:57	28:45	1:54	88:1

Table A2: The data for queuing process at 8:00am to 9:00am, on Wednesday using single server/consultant.

CUSTOMERS	A.T (min:sec)	I.D (min:sec)	W.T (min:sec)	S.T (min:sec)	D.T (min:sec)
1	00:58	0:00	13:14	1:46	15:18
2	03:20	3:02	12:32	2:17	18:09
3	05:22	2:02	13:16	1:58	20:16
4	08:30	3:08	11:57	2:30	22:17
5	09:20	1:30	15:37	1:57	26:14
6	11:01	2:21	14:29	3:14	29:04
7	14:20	3:23	18:05	1:32	30:59
8	15:59	1:35	16:23	1:34	33:21
9	18:14	2:55	14:36	1:58	35:09
10	20:55	2:41	16:25	2:03	39:23
11	23:48	2:33	17:48	1:40	42:36
12	25:28	2:20	18:10	1:53	45:31
13	28:22	3:38	17:00	1:52	47:18
14	29:35	1:09	18:30	2:13	50:18
15	31:56	2:21	17:47	1:56	50:59
16	34:05	3:34	17:30	2:51	57:26
17	37:48	3:23	18:54	2:04	58:06
18	38:47	1:39	19:53	2:12	60:17
19	40:35	2:28	20:39	2:09	63:24
20	43:44	3:09	18:34	3:02	64:58
21	45:51	2:07	20:19	2:02	67:32
22	48:12	3:01	22:22	1:54	71:48
23	53:42	5:30	19:25	2:31	74:58
24	54:58	1:16	22:07	3:16	79:41
25	56:20	2:08	24:08	1:28	81:22
26	58:47	2:21	22:01	3:11	83:19

Table A3: The data for queuing process at 8:00am to 9:00am, on Thursday using multiple server/consultant.



CUSTOMERS	A.T (min:sec)	I.T (min:sec)	W.T (min:sec)	S.T (min:sec)	D.T (min:sec)	I.D
1	4:22	0:00	23:05	4:02	31:29	1
2	8:08	4:26	21:00	4:07	33:15	2
3	11:34	3:26	19:54	4:58	35:46	1
4	15:32	4:28	19:16	6:01	40:39	2
5	18:54	3:32	18:10	5:06	42:10	1
6	20:64	2:00	18:50	6:08	45:12	2
7	21:33	1:19	20:41	3:57	45:31	1
8	22:01	1:17	20:54	3:47	46:02	2
9	26:54	4:53	19:07	3:49	48:10	1
10	31:57	5:03	17:23	5:57	54:37	2
11	32:16	1:00	20:05	3:28	54:49	1
12	43:26	2:10	18:22	5:32	58:00	2
13	36:41	2:15	17:58	4:08	58:07	1
14	41:27	2:26	16:29	3:22	61:18	2
15	44:01	3:14	19:06	3:21	65:28	1
16	45:29	1:28	18:08	2:60	66:17	1
17	49:49	4:20	19:11	4:29	73:29	2
18	52:45	3:36	15:16	4:11	74:12	1
19	55:12	3:07	19:07	5:17	79:36	2
20	56:11	1:39	22:27	5:47	84:25	2
21	59:46	3:35	21:29	4:22	85:37	1

Table A4: The data for queuing process at 8:00am to 9:00am, on Monday using multiple server server/consultant.

CUSTOMERS	A.T (min:sec)	I.T (min:sec)	W.T (min:sec)	S.T (min:sec)	D.T (min:sec)	I.T (min:sec)
1	00:00	0:00	16:45	6:04	22:49	2
2	08:21	8:21	15:36	4:56	28:53	2
3	12:32	4:11	18:23	3:44	34:39	1
4	13:22	0:50	13:32	5:54	32:48	1
5	21:24	8:02	14:15	7:20	42:59	1
6	24:43	3:02	13:22	8:41	46:46	1
7	34:59	9:16	15:32	8:23	58:54	2
8	36:51	1:51	19:17	5:02	61:10	1
9	39:42	2:31	20:56	6:51	67:29	1
10	42:54	3:12	22:54	2:11	67:59	2
11	43:55	0:21	14:26	10:32	68:53	2
12	45:14	3:21	22:41	4:14	72:09	1
13	47:00	2:26	20:42	6:21	74:03	1
14	52:03	5:03	18:32	6:11	76:46	2
15	55:18	3:15	11:31	11:26	78:15	1
16	56:12	1:34	20:42	2:11	79:05	1
17	58:32	2:20	16:32	7:12	82:16	2



18

59:44

1:12

21:14

6:22

87:20

1

Key: A.T: Arrival Time, I.T: Inter-arrival Time, W.T: Waiting Time, S.T: Service Time, D.T: Departure Time, I.D: Identification (Server)

