THE INTERNATIONAL JOURNAL OF
SCIENCE & TECHNOLEDGE

The Application of Queuing Theory on Patient Waiting Time in Ante-natal Care Clinic

Onoja Anthony
Master Degree Student, Department of Mathematical Statistics, Pan African University, Institute for Basic Sciences, Technology and Innovation, Nairobi, Kenya
Msugh Kembe
Professor, Department of Mathematics & Statistics, Statistics and Computer Sciences, Benue State University, Makurdi, Nigeria

Abstract:
Waiting lines have become an eyesore in our daily lives. Waiting in line popularly called, "queue" causes inconveniences to individuals (pregnant women) and significant costs to industries and public sectors. Expectant mothers queue for minutes, hours, days or months to receive Medical service – waiting before, during or after served. The effect of queuing about the time wasted by pregnant women to use clinical services is overwhelmingly becoming an important source of concern to most public Healthcare providers. Thus; keeping an Expectant mother waiting too long could result in danger to her health and the future baby as well as waiting cost. Providing too much service capacity to operate a system involves excessive cost. But not providing enough capacity service results in excessive waiting time and cost. In this study, the queuing characteristics at the ante-natal care clinic of the Jos University Teaching Hospital were analyzed using a Multi-Server Queuing Model and the waiting and service Costs were determined with the prime aim to evaluate the optimal service level. Data for this study was collected at the ante-natal care clinic for four weeks through initial observations, interviews and by administered questionnaires, the data was analyzed with the aid of TORA software and standard queuing formulas. The results of the analysis explicitly showed: average queue length, waiting time of pregnant women, as well as overworking of Doctors at the clinic could reduce at an optimal server level of six Doctors and a minimum total cost. As against the current server level of 4 Doctors at post with a high total cost which incur waiting and service costs. This model will help decision and other policies makers to solve the multi – server queuing problem where capacity and limited resource gives rises to longer waiting and services time.

Keywords: *Antenatal, queuing theory, waiting cost, service cost, servers*

1. Introduction

Queues are something that is inevitable in the day-to-day activities of life, it is a typical situation that happen in everyday life. They usually experienced at bus stops, hospitals, bank counters and so on.In a nutshell, queues form when the demand for service exceeds its supply. Wait time rely on the number of customers (a human being or objects) in a queuing system, the number of servers attending to the queue, and the amount of service time for each customer varies tremendously as the case may be. In healthcare institutions, the adverse effect of queuing in relation to time spent in a queue for patients to access treatments is increasingly becoming a huge source of concern for modern society that currently exposed to significant strides in technological advancement and the danger of keeping customers waiting could become a cost to them. The time wasted on the queue would have been effectively utilized elsewhere (opportunity cost of time spent in queuing). In a waiting line system, managers must know what level of service to offer. A small degree of service may be quite cheap, at least in the short while, but may result in high costs of customers' dissatisfaction, such as loss of future business. A high level of service will cost more to the provider and will incur a lower dissatisfaction costs. When considering improvements in services, the healthcare manager thinks about the cost of providing a given level of service against the prime costs from having patients (pregnant women) waiting. The ultimate goal of queuing is therefore to minimize the total cost of the system. The two core values worth note taking are costs associated with patients having to wait for service (waiting cost). Which include loss of business as some expectant mothers might not be ready to wait for service and may want to go to the nearest private Hospital organizations. Cost as the result of the delay in care or the value of the patient's time (opportunity cost of time spent in queuing) and decreased patient's satisfaction as well as the quality of care. While service cost is the cost of rendering service. These include salaries paid to employees, wages paid to employees or servers while they wait for service from other servers.

The cost of waiting for space, facilities, equipment, and supplies depend on the patient waiting time. Using the estimation of the waiting cost helps decision makers to have the ability to evaluate the recommended number of servers by minimizing the total cost including the service cost and the waiting for every pregnant woman differs depending on what the expectant mother earns every hour. Some might have their cost of waiting in multiples of other people's value. Healthcare system at large and Hospitals, in precisely, made up a crucial part of the service sector. In recent times, Hospitals have become exceedingly successful in deploying medical and technical know-how to render more effective medical treatments. Though, they are clumsy (too) often rife with inefficiencies and delays, presenting a good ground for research in numerous scientific fields, especially queuing theory. The Ante-natal healthcare systems have faced congestion problems like many other organizations. However, due to the high rate of childbirth in many underdeveloped and developing countries, made the Ante-natal health care system ineffective, null and void, prudent sound system that can no longer serve public purposes. Also, time limitation and overcrowding in many public health care centers in the country, have resulted in congestion creating a frustrated bottleneck queues in many antenatal health care clinics, leading to dissatisfaction, balking, and reneging from queues by expectant mothers. The importance of this study cannot be over emphasis though this research work aimed at examining the possible delay at the antenatal healthcare service system of JUTH. The prime objective of analyzing queuing system is to understand the behavior of the underlying process extensively. The resultant performance variables and provide the necessary tools needed for the analysis so that the outcome can be used by the policy makers to increase efficiency, improve the quality of healthcare, as well as decrease the cost, not just in Jos University Teaching Hospital, but the nation at large.

2.Review Literature

Prior to this study, some relevant related works by past scholars on queuing systems were reviewed. Adeleye A. and Biradawa K. et.al, (2002), opined that the origin of the queuing theory dates back to 1909 when a Danish Mathematician, Statistician, and Engineer named Agner Krarup Erlang published his major paper relating to the study of congestion in telephone traffic between1878-1920. A.K. Erlang is appropriately called the father of queuing theory. Patients' view of health care has gained momentum over the past 20 years (Sitzia and Wood 1997). Patients' view of service quality is affected not only by the actual waiting time but also by the perceived waiting time. The act of waiting time has significantly influenced their view (Obamiro, 2010). More so, researchers, over time have found out that customer perception is affected not just by waiting time but also by customer expectations and view of the causes for the waiting (Taylor, 1994). Invariably, the critical issues in queue management are not only the actual amount of time the customer has to wait, but also the client's views of that wait (Davis and Heineke, 1994). There are two ways to increasing customer satisfaction concerning waiting time: through reducing entire waiting time, as well as through allowing customer's waiting experience (Katz, et al., and Larson, 1991; Obamiro, 2010). (Kembe, Onah, & Iorkegh, 2012) adopted the Multi-server queuing system to incorporate cost function into the system to determine the optimal total cost of the system. Different costs assigned to wait-time for prescriptions of different priorities, there are usually some selected number of nurses or doctors required to attain to the out-patients (pregnant women) in an ante-natal care clinic of a hospital by assigning costs to the healthcare experts, and the time during which an expectant mother is in the queue.

This study also calculated the number of servers (Doctors) required so that a given percentage of pregnant women do not exceed a given waiting time and the average number of expectant mothers in the queue do not surpass a given threshold. The model assumes that pregnant would leave without service if they wait above a particular period. Long waiting time for the expectant mothers and the overutilization of Medical Personnel have been the major challenges facing the Ante-natal care clinic of the Jos University Teaching Hospital (JUTH) located at Lamingo, Jos Plateau State.

3.Materials and Methods

This study adopts a direct observations and case study approach. In-depth review of JUTH Ante-natal care clinic, attendance records for four weeks, was made. Interviews with the management of JUTH, Doctors, and records Staff were conducted to validate the secondary data and to gather other relevant information. It was used to model patients (expectant mothers) arrival and length of stay. Questionnaires administered to collect information on daily arrival rates, patients' view on queuing at the hospital, waiting time to consult a Doctor, opportunity cost of waiting, etc. The survey population for this study was the entire out-patients of Ante-natal care clinic in JUTH, Jos Plateau State. In all, 200 Expectant mothers were interviewed for this study. The data were collected from the Antenatal care clinic of JUTH on designated weeks for the Antenatal out-patients visit to the Teaching Hospital by the expectant mothers. For unbiasedness, a systematic sampling method was employed where each week represents a stratum. The rate at which pregnant women come for an antenatal visit per hour and also the service rate of the same group of expectant women per hour also observed and recorded. The following assumptions were made for queuing system at the Ante-natal care clinic of JUTH, they are;

- The arrival of expectant mothers follows a Poisson Probability distribution at an average rate of λ pregnant women per hour.
- The queue discipline is First-Come, First-Served (FCFS) before any of the server there is no particular preference to any arrival.
- Service times are distributed exponentially, with an average of μ expectant mothers per hour.

- There is no limit to the number of queues (infinite calling source capacity)
- The service providers are working tirelessly to their full capacity
- The average arrival rate is greater than the mean service rate.
- Server, in this case, depicts only Doctors, not other Medical personals like Nurses.
- Service rate is independent of line length; service providers do not go fast due to the congested line.

2.1. Model Specification

This study adopted M/M/C: (FCFS/∞/∞) (birth – death process) multi-server queuing model. For this queuing system, it assumed that the arrivals follow a Poisson probability distribution at an average of λ customers (expectant mothers) per unit time. It also opined that they served on a First-Come, First-served protocol by any of the servers (in these case Doctors). The service times are exponentially distributed, with an average of μ customers (expectant mothers) per unit of time and number of servers C. with infinite system limit (∞) and total source limit (∞).

If there are n expectant mothers in the queuing system at any point in time, then the following two cases may arise:

- If $n < C$ (number of pregnant women in the system is less than the number of servers), then there will be no queue. However, $(C - n)$ number of servers will not be busy. The combined service rate will then be $\mu_n = n\mu$; $n < C$.
- If $n \geq C$,(the number of expectant mothers in the system exceeds or equivalent to the number of Doctors) then all servers will be busy, and the maximum number of Pregnant women in the queue will be $(n - C)$. The combined service rate will be $\mu_n = C\mu$; $n \geq C$

For the purpose of modeling, the arrivals (n) are the expectant mothers. As each arrive the clinic, she books for service. If service occurs instantly; she leaves the hospital or proceed to the queue. The doctors, in this case, are the servers(C). Medhi (2003) suggests that whenever there are $n(\geq C)$ in the system, then all the C channels are busy, and the interval ranging from two consecutive service completions is exponential with rate $C\mu$. Consider the figure 1. Below:

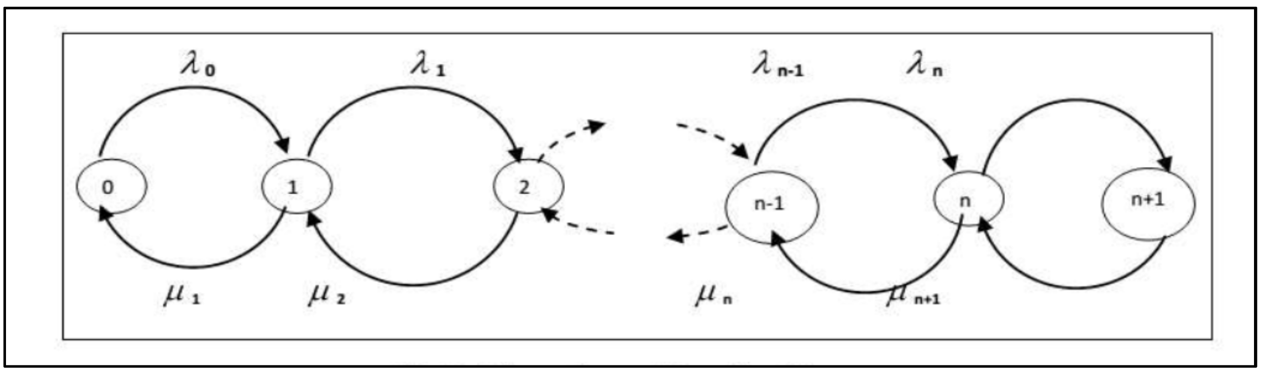

Figure 1
Source:(Obamiro, 2010)

Thus, we have a birth – death model having a constant arrival (birth) rate (λ) and state – ancillary service (death) rate as illustrated in figure 2. Service time (μ- duration of stay) determined as dismiss time, also, to admission time.

The arrival rate, service time and numbers of server's data used for the study that was obtained using the primary data source via; the observation method of data collections was adopted. The research instrument chosen was done mainly so that the queuing system can examine naturally. The study covers a period of four weeks in which the first four weeks taken into consideration.

2.2. TORA Optimization System Software

A Windows-based software that offers lots of automated or tutorial mode to operational research models such as linear programming, transportation models, queuing models, project planning, etc. the automatic method that was used in this study provides a final solution to real life large mathematical programming models.

2.2.1. Methods of System Parameters Evaluation

The following system performance parameters were employed in the study and defined as follows:

λ= arrival rate of pregnant woman per unit time.

λ_{eff}= **λ,** because there is no limit on the number of the system;

μ = service rate (length of stay) of an expectant woman per unit time;

C = Number of doctors (servers) in the model, there are C parallel servers.

ρ = Antenatal healthcare system utilization factor, i.e., the fraction of time servers (Doctors) are busy, this is equals $\frac{\lambda}{(C\mu)}$ since the recent study adopted a multiple – server model.

1) L_q : average number of pregnant women in the queue

$$= \left\{\frac{P^{c+1}\mu\lambda}{(C-1)!\,(C-P)^2}\right\} P_0 \tag{1}$$

2) L_s: average number of expectant mothers in the system

$$= L_q + \frac{\lambda}{(C\mu)} \tag{2}$$

3) W_q: Average waiting time of an expectant mother in the queue = $\frac{L_q}{\lambda}$
4) W_s: Expected waiting time of expectant mother in the system = $\frac{L_s}{\lambda}$
5) P_0 = possibility of 0 pregnant women existing in the system.
6) L_n = Expected some pregnant women waiting in line excluding those times when the line is empty.

$$P_n = \begin{cases} \frac{\lambda^n}{\mu(\mu)(\mu)\ldots(\mu)} P_0 = \frac{\lambda^n}{n!\mu^n} P_0 = \frac{P^m}{n! P_0}, & n < C \\ \frac{\lambda^n}{\mu(\mu)\ldots((C-1)\mu)(\mu^{n-C+1})} P_0 = \frac{\lambda^n}{C! C^{n-C}/n} P_0, & n \geq C \end{cases} \tag{3}$$

$$P_0 = \left\{\sum_{n=0}^{C-1} \frac{P^n}{n!} + \frac{P^C}{C!}\left(\frac{1}{1-\frac{P}{C}}\right)\right\}^{-1}, \quad \frac{P}{C} < 1 \tag{4}$$

2.3. Costs Determination in the Model

To evaluate and find the optimum number of servers in the system, two crucial costs must be considered in making these decisions:

- Service costs
- Waiting time costs of expectant mothers, economic analysis of these costs allows the management of JUTH make a trade-off between the incurring costs of providing better service and the decreased waiting time costs of Pregnant women derived from providing that service.

The objective is to minimize the total expected system cost: Expected service cost

$$E(CC) = C \tag{5}$$

Where, C = number of servers, C_s = service cost of each server
Expected waiting costs in the system

$$E(WC) = (\lambda W_s)C_w \tag{6}$$

Where, λ = some arrivals, W_s = average time a visitor spent on the system.

C_w = opportunity cost of waiting for the expectant mothers. Adding (5) and (6) we have: Expected Total costs = $min\{E(TC) = E(CC) + E(WC)$ (7)

Expected Total costs = $min\{E(TC) = SC_s + (\lambda W_s)C_w\}$ (8)

$$E(TC) = C_O + CC_S + (\lambda W_S)C_w \tag{9}$$

C_o:The fixed cost of operation system per hour
C_s :The marginal cost of a registration agent (doctor) per hour.
Note: for this scenario, an analytical solution doesnot exist, and it is necessary to solve the problem by grouping especially that to be a question of convex function. In other words, whether her curve is in form U**,** therefore, it just takes to estimate $E(TC)$ for the values of C grouping up until the cost cease to decrease.

2.3.1. Cost of Waiting Time (WC)

- One unit of waiting time by a pregnant woman estimated by their monthly earned allowance.
- The mean waiting time cost per hour of the expectant mothers:

$$C_W = \sum W_i C_i \tag{10}$$

Where: C_i : Hourly earned income of an expectant mother W_i : Weight of the category i extracted from the total

4.Results

The TORA optimization window base ® software as stated earlier was used to evaluate the performance measures of the Multi-server queuing system at Lamingo, Jos University Teaching Hospital (JUTH). Using the following input data below:

Parameter M/M/C:FCFS/∞/∞	Value
Arrival rate(λ)	11 pregnant women per hour
Service rate(μ)	3 pregnant women per hour
Number of servers	4,5,6,7 and 8 depending on the scenario

Table 1: input parameters for TORA optimization window based ® software

Scenario	C	Lambda	Mu	L'da eff	P_0	Ls	Lq	Ws	Wq
1	4	11.00000	3.00000	11.00000	0.00909	12.70584	9.03918	1.15508	0.82174
2	5	11.00000	3.00000	11.00000	0.02090	4.85703	1.19036	0.44155	0.10821
3	6	11.00000	3.00000	11.00000	0.02419	3.99659	0.32992	0.36333	0.02999
4	7	11.00000	3.00000	11.00000	0.02516	3.76942	0.10275	0.34267	0.00934
5	8	11.00000	3.00000	11.00000	0.02545	3.69888	0.03221	0.33626	0.00293

Table 2: performance measures of M/M/C/FCFS/∞/∞

Scenario	(C)	λ	Ws	λWs	Co	Cs	Cw	CCs	(λWs)Cw	E(TC)
1	4	11.00	1.155	12.705	#1388.89	#3645.83	#7066.67	#14583.32	#89782.00	#105,754.22
2	5	11.00	0.442	4.862	#1388.89	#3645.83	#7066.67	#18229.17	#34358.13	#53,976.19
3	6	11.00	0.363	3.993	#1388.89	#3645.83	#7066.67	#21874.99	#28217.20	#51,481.09
4	7	11.00	0.343	3.773	#1388.89	#3645.83	#7066.67	#25520.83	#26662.53	#53,572.26
5	8	11.00	0.336	3.696	#1388.89	#3645.83	#7066.67	#29166.67	#26118.40	#56,673.96

Table 3: Summary of the Marginal Costs, Service Costs, Opportunity Cost and the Number of Servers for the ante-natal Care unit of JUTH

Performance Measures	4 Doctors	5 Doctors	6 Doctors	7 Doctors	8 Doctors
Arrival rate (λ)	11	11	11	11	11
Service rate (μ)	3	3	3	3	3
System Utilisation	91.7%	73.3%	61.1%	52.4%	45.8%
L_S	12.706	4.857	3.997	3.769	3.699
L_Q	9.039	1.190	0.330	0.103	0.032
$W_s - in$ hours	1.155	0.442	0.363	0.343	0.336
$W_q - in$ hours	0.822	0.108	0.030	0.009	0.003
P_0	0.009%	0.021%	0.024%	0.025%	0.025%
Total System Cost/hr	#105,754.22	#53,976.19	#51,481.09	#53,572.26	#56,673.96

Table 4: Summary analysis of the M/M/C: FCFS/∞/∞ queuing model of antenatal care clinic of JUTH.

Figure. 2, 3, 4, 5 and 6 depicts the line plots of the Utilization factor (ρ) against Average pregnant women waiting time in the system (W_s). Average Number of pregnant women in the system (L_s) against the probability of the system being idle (P_o), Expected service Cost against Level of service, Expected waiting Cost against Level of service and the Expected Total Cost (TC) against the Level of service.

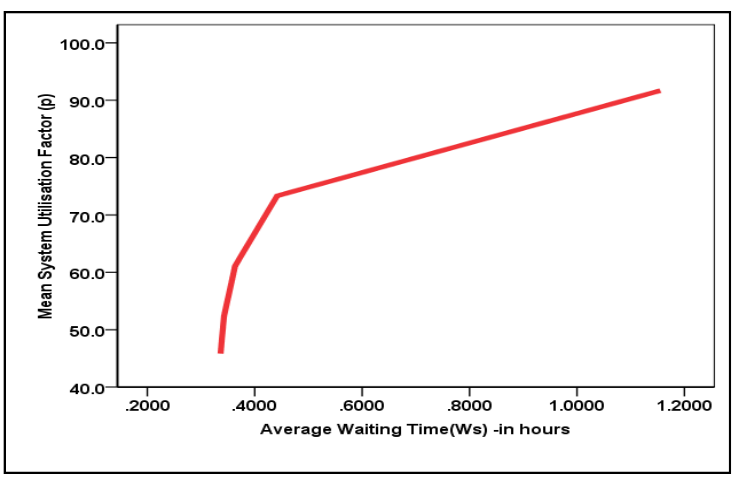

Figure 2: Utilisation Factor (p) against Average Pregnant Women waiting time in the System (Ws)

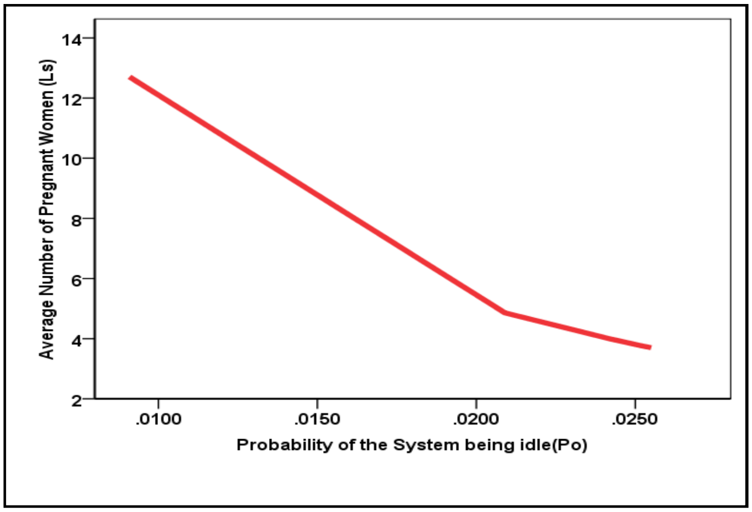

Figure 3: Average Number of Pregnant Women in the System (Ls) against Probability of the System being idle (Po)

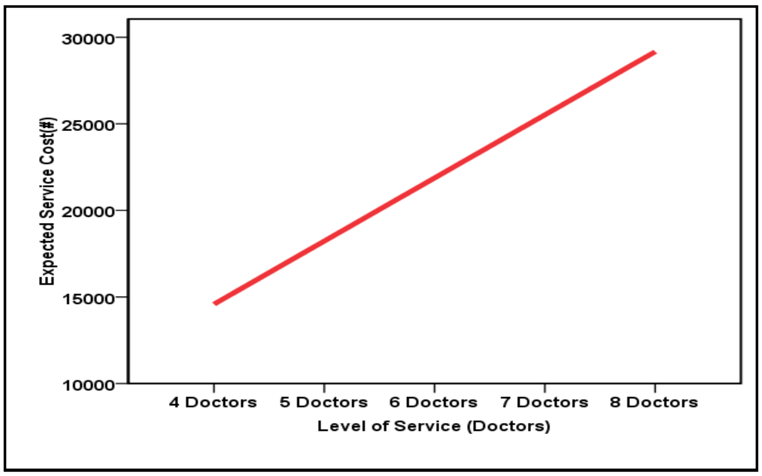

Figure 4: Expected Service Cost (#) against Level of Service (Doctors)

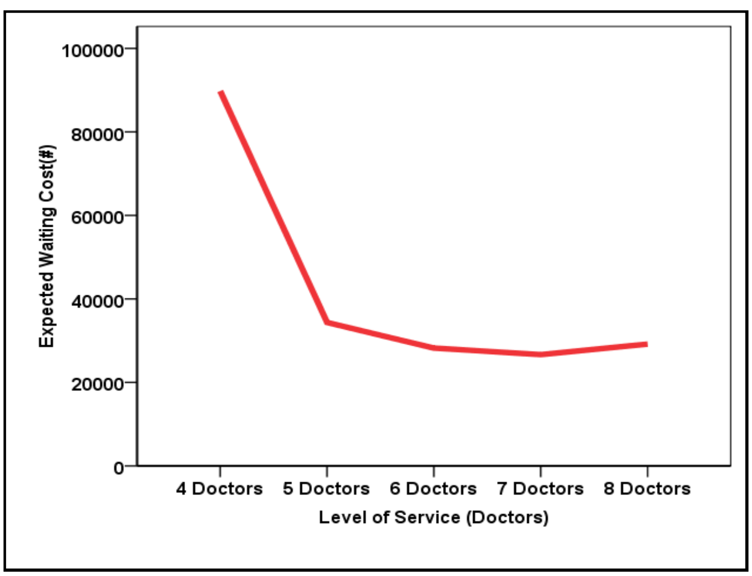

Figure 5: Expected Waiting Cost (#) against Level of Service (Doctors)

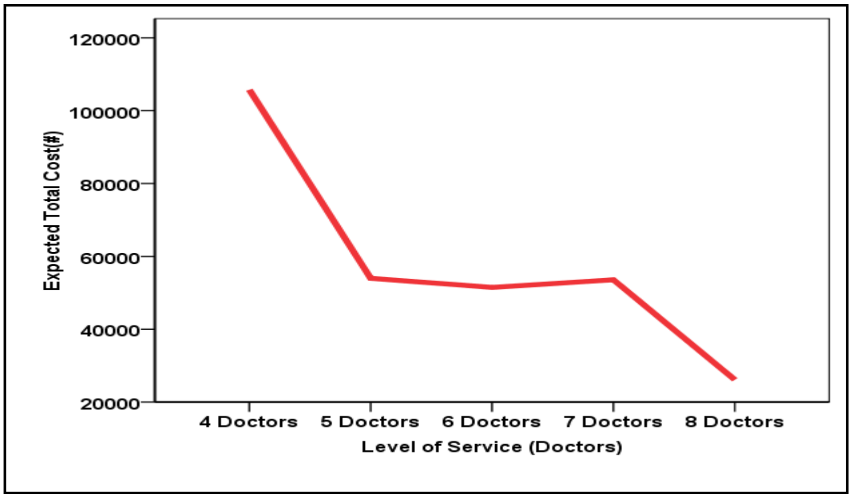

Figure 6: Expected Total Cost (#) against Level of Service (Doctors)

5.Discussion of Results

The Figures 2,3,4,5 and 6 above; depicts the extremely marginal changes for scenario 2,3,4,5 and 6 on the average waiting time. Average time spends in the system, the length of a queue, the number of expectant mothers in the system, system utilization factor (%), the probability of no expectant mother in the queue (P_o) and the total system cost. All these collaborates the above analytical graphs. One needs to be cautious of the costs involved in achieving these minimal changes. Employing more Doctors will mean taking on more costs. An excellent, outstanding balance between the number of Doctors, costs, and optimal system performance is essential for sustainability. Hence, the conclusion that scenario 3 with six Doctors at the post is best for optimal performance with a total system cost of #51,481.09. This is in-line with the findings of (Kembe et al., 2012). It is also worth noting that the expectant mothers' waiting time, cost and congestion in the system is less at this optimal server level.

6.Conclusion and Recommendation

Approaching a time of significant challenges in the healthcare sector, enhancing satisfaction is becoming increasingly important because satisfaction is a measure of quality. Existing Knowledge of the use of queuing model to determine system parameters is of outermost priority to healthcare providers who seek to attract, keep and provide quality healthcare to a patient in the ever-competitive "marketplace." Queuing theory is a renowned and tested mathematical approach to the analysis of waiting lines. The queuing characteristics at the Ante-natal care clinic of JUTH were analyzed using a Multi-server Queuing Model and the Waiting and service Costs obtained to (or "intending to") determining the optimal service level. The results of the analysis demonstrated that average queue length, waiting time of pregnant women as well as over usage of Doctors can reduce when the capacity service level of Doctors at the clinic is increased from four to six at a lesser total cost which incur the waiting cost and service costs. The operation Directors can recognize the trade-off that must take place between the cost of providing excellent service and the value of customers waiting time. Service improves the cost of time spent waiting in the line decreases. And be done by expanding the service facilities or using models that consider cost optimization. The queuing modeling approach and adopted software (TORA) in this study are quite understandable and flexible. They can be used to model different healthcare units such as the outpatient session, intensive care unit, blood banks, etc. more so, the model can also be adopted by the Independent National Electoral Commission (INEC) and other service organizations such as banks, restaurants, and telecommunication.

7.References

i. Adeleye A., Madawaki M.D., Biradawa K. (2002), A Quantitative Business Technique
ii. Adedayo, O. A., O., Obamiro, J. K.,(2006)Operations Research in Decision Analysis and Production Management. Lagos: Pugmark Nig.Ltd.,
iii. Adele, M., Barry, S.,(2005).Modeling patient flow in hospital using queuing theory, Unpublished Manuscript,
iv. Babes, M., Serma, G. V.,(1991); outpatient Queues at the Ibn-Rochd Health Centre, Journal of Operational Research, 42(10): 1086 – 1087.
v. Belson, G.V., (1988), Waiting Times Scheduled Patient in the Presence of Emergency Request, Journal of Operational Management 2(3),
vi. Cochran, K.J., Bharti, A.,(2006), A Multi-stage Stochastic Methodology for Whole Hospital Bed Planning Under Peak Loading, International Journal of Industrial and Systems Engineering, 1 (1/2): 8-35,
vii. Davis, M.M., Heineke, J.,(1998),How Disconfirmation, Perception and Actual Waiting Times Impact, Unpublished Manuscript,

viii. Davis, M.M., Vollmann, T.E.,(1990), A Framework for Relating Waiting Time and Customer Satisfaction in a Service Operation, Journal of Services Marketing, 4(1) 61-69,
ix. Hall, W.R.,(2006), Patient Flow: The New Queuing Theory for Healthcare, OR/Ms, Today,
x. Ivalis, S., Millard, P.,(2003), Healthcare modeling operating the black box, British Journal of Healthcare Management, 8(7): 251-255,
xi. Katz, K., Larson, K., Larson, R.,(1991), Prescription for the waiting in line Blues: Entertain, Enlighten, and Engage", Sloan Management Review (Winter): 44-53,
xii. Kembe, M.M., Onah, E.S., Lorkegh, S. A.,(2012), A Study of Waiting and Service Costs of a Multi-server Queuing Model in a Specialist Hospital, International Journal of Scientific and Technology Research, 5(2): 2277-8616,
xiii. Medhi, J.,(2003), Stochastic Models in Queuing Theory, Amsterdam: Academic Press, Second Edition.
xiv. Nosek, A.R., Wislon, P. J.,(2001), Queuing Theory and Customer Satisfaction: A Review of Terminology, Trends, and Applications to Pharmacy Practice, Hospital Pharmacy, 36(3): 275-279,
xv. Nosek, R.A., and Wilson, J.P.,(2001)," Queuing Theory and Customer Satisfaction: A Review of Terminology, Trends, and Applications to Pharmacy practice. Hospital Pharmacy", 36, 275-276, http://www.drum.lib.edu.
xvi. Obamiro, J.K.(2003), Queuing theory and Patient Satisfaction, An overview of Terminology & Application in Ante – natal Care Unit.http://www.upg-bulletin-se.ro
xvii. Singh, V.,(2007), Use of Queuing Models in Health Care, Department of Health Policy and Management, the University of Arkansas for medical science.
xviii. Sitzia J., Wood N.,(1997), Patient Satisfaction: A Review of Issues and Concepts, Soc Sci Med, 45:1829-1843.
xix. Taylor, S.,(1994), Waiting for Service: The Relationship between Delays and Evaluation of Service, Journal of Marketing, 58(2): 56-69,