Journal of Medical and Applied Biosciences Volume 3, March 2011

www.cenresinpub.org
ISSN 2277-0054

OPTIMAL PATIENT SERVICE DESIGN IN A RADIODIAGNOSTIC MEDICAL FACILITY

C. O. Arimie*[1] and E. H. Etuk[2]
[1]Department of Radiology, University of Port Harcourt Teaching Hospital, Port Harcourt
[2]Department of Mathematics/Computer Science, Rivers State University of Science and Technology, Port Harcourt
E-mail: codarimie@yahoo.com

ABSTRACT

Queuing theory is a powerful quantitative tool that enables healthcare facilities to uncork chronic bottlenecks in the flow of patients but, healthcare systems have some peculiar features that impact on their modeling in the framework of the queuing theory. To redesign the radiodiagnostic facility of the University of Port Harcourt Teaching Hospital (U.P.T.H) using queuing theory, and to assess the impact of the time lag between patients arrival and the commencement of imaging procedures, occasioned by absences and lateness to work by healthcare personnel and other logistic problems, on the operating characteristics and the optimality criterion. The patients' service schedule was altered so that only outpatients were allowed service during the peak period of 8.00 a.m. to 1.00 p.m. daily. Arrival and service rates data were collected and used to fit the Poisson model. Prior to this time, data were not available on arrival rates and service times. The steady-state probabilities, measures of performance, the probability of an arrival joining the queue, and the joining rate were calculated. The rescheduling revealed that the system was M/M/c/GD/∞ /∞ as the patients' arrival and service rates data collected fitted the Poisson model. We discovered an improvement in the various measures of performance especially, the throughput of outpatients although, the time lag between arrival and commencement of service increased the queue length and waiting time. Patient service in the radiodiagnostic facility can be optimized using the queuing theory but, service must commence as the patients arrive in order to achieve optimal performance. Secondly, performance of a queuing system may just be specified in terms of probabilities.

Key words: Queuing, Optimization, Markov chain, Radiodiagnostic facility, Medical.

INTRODUCTION

The search for better ways to improve operational efficiency and production capacity for medical treatment has always bothered Healthcare managers. Queuing model offers an excellent tool to analyze and to improve the performance of healthcare systems. However, healthcare systems have some peculiar features that impact on their modeling in the framework of the queuing theory. Congestion occasioned by absences and lateness to work by healthcare personnel, and other logistic problems is a common feature in many public hospital based radiodiagnostic facilities in Nigeria. The diversity of the patient groups demanding service, and the diversity of the investigation mix poses a great challenge to effective and efficient design of the service delivery system. Whereas some investigations are simple and require short service times, others are complex and require longer service duration. Strategies must therefore, be developed to cope with the investigation mix, and ensure minimum delays and congestion at the facility.

In healthcare services, the demand for resources is to a large extent unscheduled. As a result, there is a permanent mismatch between demand for healthcare services and the available capacity (supply). Queuing models are useful for determining capacity levels (and the allocation of capacity) needed to respond to demands in a timely fashion[1]. The patient mix and the associated variability in the arrival stream and the hospital resources such as personnel, diagnostic rooms, waiting rooms, imaging machines, etc. in any radiodiagnostic process are inherently stochastic. Congestion at the facility could be attributed to a number of factors[2]. Delays and facility congestion have also been attributed to the quality and experience of the medical staff, and the allocation of medical capacity between distinct demand streams - outpatients, inpatients, and emergencies[3]. Queuing models usually assume time-independent (input) demand rates. In healthcare facilities, arrivals consist of acute (unscheduled) and elective (scheduled) patients. The long term steady-state probability distributions for queues are usually assumed to be independent of time. Green and Soares[4], and Ingolfsson et al.[5] have noted that in healthcare systems, we have time varying arrival rates and time varying server availability and time-dependent waiting times.

Research has shown that application of quantitative techniques to patient service management has produced formidable and desirable results. Green[6] propose a stationary independent period-by-period (SIPP) approach to determine how to vary staffing to meet changing demand. Green et al.[7] used a finite horizon dynamic programming to investigate the allocation of service capacity among several competing customer classes in a magnetic resonance imaging (MRI) medical facility. In another study, it was shown that priority queue discipline reduces the average waiting time for all patients[8]. The analysis of patient transfer from outpatient physicians to inpatient physicians yielded similar result[9]. Queuing theory was used to predict the optimal number of scheduled slots to be reserved for emergency computed tomography (CT) and ultrasonography (USS) in a radiology department[10]. Congestion at the facility increases the probability of reneging and discourages arrivals[11]. In systems where demand exceeds server capacity, reneging is the only way that a system attains a 'state of dysfunctional equilibrium'[12]. It is possible to redesign a queuing system to reduce reneging. A common approach is to separate patients by the type of service required[11]. In designing health care system using queuing theory, one problem is to determine the optimal design. Grassmann[13] distinguished between optimal design of queues and optimal control of queues. He showed that the optimal service rate is closely related to the variance, σ^2 of the queue length and the expected number of elements an arrival will encounter before joining the queue, L^*. He expressed the optimality condition,
$T' = rD' - C_s - C_wL' = 0$ in terms of σ^2 and L^* as
$r(L - L^*)(D/\mu) - C_s + C_w(\sigma^2/\mu) = 0$ where,
$D' = (L - L^*)(D/\mu)$, $L' = - \sigma^2/\mu$, and $\sigma^2 = \sum (i - L)^2 P_i$; $i = 0,1,2,\ldots\ldots$
T = net revenue per period, r = revenue for each customer served,
D = unconditional rate at which customers join the queue,
L = expected number of units in the system,
i = number of elements in the system, and
P_i = the probability of i elements in the system

Using this optimality criterion requires knowledge of the operating cost, waiting cost, and revenue of the system being modeled for the determination of the economic service rate (optimal service rate).

A design is said to be optimal if the mean service rate results in minimization of delays, facility congestion and cost, and maximization of patients' throughput and net revenue. Determining an optimal design in terms of cost and revenue is usually done on comparative basis with a view to selecting among alternative queuing models, one that yields the minimum cost service rate. Here, knowledge of the expected total systems cost per period, T_c is important. $T_c = W_c + F_c$ where W_c is the expected waiting cost per period, and F_c is the expected facility (service) cost per period. Now, $W_c = C_w.L_s = C_w(\lambda / \mu\text{-}\lambda)$ and $F_c = C_s\mu$ where, C_w is the cost of waiting per unit of time, C_s is the unit service cost, and L_s , λ and μ are as previously defined. The expected total system cost is therefore, expressed as $T_c = C_w(\lambda / \mu\text{-}\lambda) + C_s\mu$. Differentiating T_c with respect to μ and equating to zero allows one to solve for μ which is the minimum cost service rate[14]. That is, $\mu = \lambda + (\lambda C_w / C_s)^{1/2}$. Often times, it is difficult to have access to data on cost and revenue of the system being studied. Also, in healthcare services the objective of maximizing patients' throughput is considered more important than cost consideration, and it is difficult to estimate accurately the cost of a patient waiting for service. As a result, optimality criterion in healthcare service systems design may not be determined in terms of cost and revenue. It was the purpose of this study to redesign patient service in the radiodiagnostic facility of the University of Port Harcourt Teaching Hospital, using an $M/M/c/GD/\infty/\infty$ queuing model, to highlight the impact of the lag time between patients' arrival at the facility and the actual commencement of service, rescheduling different classes of patients, and segregating outpatients into priority groups according to the complexity and perceived duration of their cases, on the system's performance. In this notation, the symbols, M stand for 'Markov' indicating that the number of arrivals and the number of completed services in time **t** follow Poisson process which is a continuous time Markov chain[15, 16, 17], c stands for the number of service channels, GD stands for general queuing discipline, and ∞/∞ stands for unlimited expected number of patients in the queue and an infinite size of population from which the patients are drawn.

METHODOLOGY

The patients' service schedule was altered so that only out patients were allowed service during the peak periods of 8.00 a.m. to 1.00 p.m. Of every working day (Monday through Friday) for 6 months (July to December, 2007). Outpatients who presented after 12 noon were rescheduled for the next day. Data of the time of arrival, time of entering into service, and the time a service was completed were collected prospectively from all outpatients who presented for medical x-ray imaging in each month of the period of study. These times were recorded for each of the patients sampled. The facility had two functional diagnostic rooms hence, the data were organized into hourly time periods and presented in tabular form as the structure of hourly arrival, and hourly departure from each of the two diagnostic rooms designated rooms 1 and 2 respectively. These were recorded for 6 consecutive months. The cases were divided into two groups according to their complexity and perceived service duration. All chests and the extremities were done in room 1 while other cases were done in room 2. We considered x-ray imaging of outpatients only as work on all the demand streams and imaging modalities was considered to be too cumbersome for the purpose of this study.

Frequency distribution tables were generated for the arrival rate, and service rates (departure rates) and the expected frequencies computed for each of the 6 months. The arrival and service rates data were then fitted to the Poisson model, and the goodness of fit test conducted using the chi-square statistic to ascertain the adequacy of the model. The mean arrival rate, λ and the mean service rate, μ for diagnostic rooms 1 and 2 were calculated and the values were used to determine the utilization factor, ρ (i.e. the fraction of time the service facility is busy), the steady state probabilities, and the operating characteristics (measures of performance) of the system for each month. In this facility, registration and documentation of patients' data commences at 8.00a.m. but imaging procedures actually commences at 9.00a.m. Hence, a sample size, N = (5hours x No. of days worked in the month) hours was used to fit the arrival rate

Data to the Poisson model, $f(x) = \lambda^x e^{-\lambda} / x!$ and a sample size, N = (4hours x No. of days worked in the month) hours was used to fit the service rate data of diagnostic rooms 1 and 2 to the Poisson model, $f(x) = \mu_i^x e^{-\mu_i} / x!$ respectively, i = 1, 2.

THE MODEL

The model is essentially **M/M/2/GD/∞/∞** with modification to the formulae for the measures of performance occasioned by the increase in the queue length due to the time lag between patient arrival at the facility and actual commencement of imaging service. The modified queue length,

$$L_{q_m} = L_{q_{t=0}} + L_{q_{t=1,2,......,n}}$$ Where,

$L_{q_{t=0}}$ is the expected (mean) number of patients waiting in queue before commencement of service.

$L_{q_{t=1,2,....,n}}$ is the computed steady-state expected (mean) number of patients waiting in queue (where n = hourly time periods of service). This and the steady-state probabilities were computed using the TORA primer optimization software.

The other steady-state measures of performance and steady-state probabilities were computed using the following formulae:

1. $W_{q_m} = L_{q_m} / \lambda$ 2. $L_s = L_{q_m} + \lambda / \mu$

3. $$W_s = W_{q_m} + 1/\mu \quad \text{Where} \quad \mu = \frac{\mu_1 + \mu_2}{c}$$

L_{q_m} = mean number of patients waiting in the queue

W_{q_m} = mean waiting time

L_s = mean number of patients in the system (i.e. patients in queue plus those being served)

W_s = mean transit time from entering to leaving the system (including the waiting time).

4. $$P(n \geq c) = \frac{(c\rho)^c \pi_0}{c!(1-\rho)}$$ This is the steady-state probability that all servers are busy (i.e. the probability that an arrival has to wait), where π_0 is the probability of zero patient in the queue and zero server is busy, $\rho = \lambda / c\mu$ and c = number of service channels.

5. The probability that room 1 is idle $= \pi_0 + [\ \mu_1 / (\mu_1 + \mu_2)\ ]\pi_1$

6. The probability that room 2 is idle $= \pi_0 + [\ \mu_2 / (\mu_1 + \mu_2)\ ]\pi_1$

The rate of joining the queue given that already there are **i** patients in the system was calculated using the relation $\lambda_i = \alpha_i \lambda$, (Grassmann,1979) where λ_i = the rate at which new patients join the queue, λ = the mean arrival rate, and α_i = the probability of a patient joining the queue given that there are already i patients in the system. α_i was calculated using the formula

$$\alpha_i = 1 - \left(\pi_0 + \sum_{n=1}^{15} \pi_n \right).$$

RESULTS

The goodness of fit test showed that the model was adequate as the arrival and service rates data fitted the Poisson model. For the arrival rate data, $\chi^2 = 3.6866 < \chi^2_{10,\ 0.05} = 18.307$ thus implying a good fit of the data to the Poisson model at 0.05 level of significance. For the departure rate (from room1) data, $\chi^2 = 10.2364 < \chi^2_{10,\ 0.05} = 18.307$ while for the departure rate (from room2) data, $\chi^2 = 4.7357 < \chi^2_{6,\ 0.05} = 12.592$ thus implying a good fit of the data to the Poisson model at 0.05 level of significance. The arrival rate, λ for the period of study = 11.87 patient/hour. The service rate for room 1, μ_1 =11.07 patient/hour and the service rate for room 2, μ_2 = 3.77 patients/hour. The mean service rate, μ = 7.42 patients/hour. The utilization factor, $\rho = \lambda / (\mu_1 + \mu_2) = 0.79987 \cong 0.8$, implying that the facility was busy 80% of the time. The probability that room 1 is idle = 0.24388 and the probability that room 2 is idle = 0.15638. The probability that an arrival has to wait, $P(n \geq c) = 0.71162$. The mean number of patients waiting before commencement of service at 9.00 a.m., $L_{qt=0} = 9.75591$ and the steady-state mean number waiting in queue for service, $L_{qt=1,2,\ldots,n} = 2.84131$. Therefore, the effective queue length, $L_{qm} = 9.75591 + 2.84131 = 12.59722$ patients. The other performance measures, W_{qm} = 1.06127 hours, L_s = 14.19695 patients, and W_s =

1.19604 hours. The probability of a patient joining the queue given that there are already i patients in the system, α_i = 0.0312 and the rate of joining the queue given that there are already i patients in the system, λ_i = 0.37034patients per hour ($\cong$ 1patient in 3hours). There was improvement in the outpatient throughput consequent on the redesigning of the facility as shown in table 5.

Table 1: Structure of hourly arrival of patients to the facility

Month of study	No. of days	8 - 9 a.m.	9 - 10 a.m.	10 - 11 a.m.	11 - 12 p.m.	12 -1 p.m.	TOTAL	MEAN ARRIVAL RATE, λ
July	20	215	298	327	273	186	1299	12.99
August	23	247	273	290	298	232	1340	11.65
Sept.	20	198	315	319	267	193	1292	12.92
Oct.	23	234	264	282	235	189	1204	10.47
Nov.	22	206	301	350	274	186	1317	11.97
Dec.	19	139	246	297	278	125	1085	11.42
TOTAL	127	1239	1697	1865	1625	1111	7537	11.87

Table 2: Structure of hourly departure of patients from diagnostic room 1

Month of study	No. of days	9 - 10 a.m.	10 - 11 a.m.	11 - 12 noon	12 - 1 p.m.	Total	Mean Service Rate, μ_1
July	20	286	272	248	184	990	12.38
Aug.	23	290	301	236	181	1008	10.96
Sept.	20	265	293	250	177	985	12.31
Oct.	23	248	266	272	107	893	9.71
Nov.	22	198	290	279	215	982	11.16
Dec.	19	223	237	214	91	765	10.07
Total	127	1510	1659	1499	955	5623	11.07

Table 3: Structure of hourly departure of patients from diagnostic room 2

Month of study	No. of days	9 - 10 a.m.	10 - 11 a.m.	11 - 12 noon	12 - 1 p.m.	Total	Mean Service Rate, μ_2
July	20	67	70	82	90	309	3.86
Aug.	23	75	97	86	74	332	3.61
Sept.	20	70	81	73	83	307	3.84
Oct.	23	69	79	91	72	311	3.38
Nov.	22	84	77	79	95	335	3.81
Dec.	19	73	80	78	89	320	4.21
Total	127	438	484	489	503	1914	3.77

Table 4: The steady-state probabilities of the system

n	π_n	n	π_n	n	π_n	n	π_n	n	π_n
0	0.11119	9	0.02980	18	0.00399	27	0.00054	36	0.00007
1	0.17788	10	0.02384	19	0.00319	28	0.00043	37	0.00006
2	0.14228	11	0.01907	20	0.00256	29	0.00034	38	0.00005
3	0.11381	12	0.01525	21	0.00204	30	0.00027	39	0.00004
4	0.09103	13	0.01220	22	0.00163	31	0.00022	40	0.00003
5	0.07281	14	0.00976	23	0.00131	32	0.00018	41	0.00002
6	0.05824	15	0.00780	24	0.00105	33	0.00014	42	0.00002
7	0.04658	16	0.00624	25	0.00084	34	0.00011	43	0.00002
8	0.03726	17	0.00499	26	0.00067	35	0.00009	44	0.00001

Table 5: Record of outpatients x-rayed between January and December, 2007

JANUARY	825	JULY	1,299
FEBRUARY	826	AUGUST	1340
MARCH	960	SEPTEMBER	1292
APRIL	850	OCTOBER	1204
MAY	961	NOVEMBER	1317
JUNE	986	DECEMBER	1085
TOTAL	5408		7537

DISCUSSION

Queuing theory is a powerful quantitative tool that enables healthcare facilities to uncork chronic bottlenecks in the flow of patients and to determine capacity levels (and the allocation of capacity) needed to respond to demands in a timely fashion (minimizing the delay). However, we discovered in the course of this study that queuing theory is not monolithic. Different models are needed to solve different types of problems. Efficiency of a queuing model may be evaluated by the measures of performance especially, when the cost and revenue structure of the system is inaccessible. In this wise, optimal design of queues would, at best, be described as optimal control of queues. Our result showed that rescheduling, and segregating the patients into different service channels impacted positively on the systems performance (service in room1 is 3 times faster than in room2). This is in agreement with previous studies[8, 9, 18]. The time lag between patient arrival to the facility and the actual commencement of imaging procedures impacted negatively on the performance of the facility as it increased the expected queue length and waiting time. The calculated rate (0.37034patients/hour) at which new patients will join the queue suggests an inefficient system despite the overall improvement in the patients' throughput. However, we believe that if the facility can commence service at 8.00 a.m. as the patients arrive, the queue length and waiting time will reduce considerably. Assuming that the effective service rate remains the same as computed (μ = 7.42 patients/hr.), the queue length and the expected number of patients in the system will reduce to $L_{qm} \cong 3$ patients and $L_s \cong 5$ patients respectively. And, the probability of a new patient joining the queue given the number of patients, i already in the system, and the rate of joining the queue will increase as shown in the computation below:

Given $i \geq L_s = 5$ patients, α_i = the steady state probabilities of 6, 7, 8, ….., patients in the system. That is,

$\alpha_i = \pi_6 + \pi_7 + \pi_8 + \ldots\ldots.. = 1 - (\pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5)$

= 1 - 0.709 = 0.291. The rate of joining the queue, λ_i become

$\lambda_i = 0.291\lambda = 0.291 \times 11.87 = 3.45417$patients / hr. ($\cong$ 4patients / hr.).

CONCLUSION

This study evaluated the application of queuing theory to the problem of patient service design in a healthcare facility, and we came to the conclusion that patient service can be optimized using the queuing theory but, service must commence as the patients arrive in order to achieve optimal performance. Secondly, performance of a queuing system may just be specified in terms of probabilities. As it is not certain that these results can be generalized, a further research in this area is recommended.

REFERENCES

1. Creemers, S., Lambrecht M., and Vandaele N., (2007). Queuing Models in Healthcare. *Research Center for Operations Management,* **...** *https://lirias.kuleuven.be/bitstream/...*
2. Stafford, E. F. and Aggarwal, S. C. , (1979). Managerial Analysis and Decision-making in Outpatient Health Clinics. *The Journal of the Operational Research Society,* vol. 30, 905 - 915.

3. Arimie, C. O. D., (2011). Optimizing Patient Service in a Fixed Capacity Medical Imaging Facility. *Nigerian Journal of Medical Imaging and Radiation Therapy,* Vol. 1 No. 2, 27-35.
4. Green, L. and Soares, J., (2007). Computing Time-Dependent Waiting Time Probabilities in M(t)/M/s(t) Queuing Systems, *M&SOM Manufacturing & Service Operations Management,* 9, 1, 54-61.
5. Ingolfsson, A., Haque, A. and Umnikov, A., (2002). Accounting for Time-Varying Queuing Effects in Workforce Scheduling, *European Journal of Operational Research,* 139, 585-597.
6. Green, L., (2006). Queuing Analysis in Healthcare, in Hall, R. W. ed., *Patient Flow: Reducing Delay in Healthcare Delivery.* Springer, New York. 281-307.
7. Green, L. V., Savin, S., and Wang, B., (2006). Managing Patient Service in a Diagnostic Medical Facility. (http://www.accessmylibrary.com). 4/13/2007.
8. Siddhartan, K., Jones, W. J., and Johnson, J. A., (1996). A priority Queuing Model to Reduce Waiting Times in Emergency Cases. *International Journal of Health Care Quality Assurance, vol. 9, 10 - 16.*
9. Worthington, D., (1991). Hospital Waiting List Management Models. *The Journal of the Operational Research Society, vol.42, 833 - 843.*
10. Vasanawala, S. S. and Desser, T. S., (2005). Accommodation of requests for emergency US and CT: Application of Queuing Theory to Scheduling of Urgent Studies. (http://www.radiology.rsnajnls.org/cgi). 11/21/2007.
11. Fomundam, S. and Herrmann, J., (2007). A Survey of Queuing Theory Applications in Health Care. (http://www.isr.umd.edu). 12/5/2007.
12. Hall, R., Belson, D., Murali, P., and Dessouky, M., (2006). Modeling Patient Flows through the Health Care System, in: *Patient Flow: Reducing Delay in Health Care Delivery,* Hall, R. W. ed., Springer, New York.1-44.
13. Grassmann, W. K., (1979). The Economic Service Rate. *The Journal of the Operational Research Society,* Vol. 30, 149-155.
14. Fabrycky, W. J., Ghare, P. M., and Torgersen, P. E., (1987). *Applied Operations Research and Management Science*. Prentice-Hall of India Private Ltd., New Delhi. 414-429.
15. Veerarajan, T., (2007). *Probability, Statistics and Random Processes,* 2nd edition. Tata McGraw-Hill Publishing Company Ltd., New Delhi. 8.3.
16. Taha, H. A., (2004). *Operations Research: An Introduction,* 7th edition. Prentice-Hall of India Private Ltd., New Delhi. 598.
17. Winston, W. L., (1994). *Operations Research: Applications and Algorithms,* 3rd edition. Duxbury Press, Wadsworth, Inc. Belmont, California. 1112-1113.
18. Brahimi, M. and Worthington, D. J., (1991). Queuing Models for Outpatient Appointment Systems - A Case Study. *The Journal of the Operational Research Society,* Vol. 42, 733-746.