QUEUING THEORY AND PATIENT SATISFACTION: AN OVERVIEW OF TERMINOLOGY AND APPLICATION IN ANTE-NATAL CARE UNIT

BABAR VIJAY Y, RAUT YOGESH G, KHOT P.G.

Abstract : Waiting lines are experienced in our daily activities. Waiting in line or queue causes inconvenience to individuals (patients) and economic costs to firms and organizations. Patients (expectant mothers) wait for minutes, hours, days or months to receive medical service- waiting before, during or after being served. Queuing theory is a mathematical approach to the study of waiting in lines/queues. This paper presents the results of a study that evaluates the effectiveness of a queuing model in identifying the ante-natal queuing system efficiency parameters. It uses SPSS 17.0 software to analyze data collected from ante-natal care unit of Civil General Hospital, Wardha over a three-week period. The study showed that pregnant mothers spent less time in the queue and system in the first week than during the other succeeding two weeks. This implies that there are less average pregnant women in the queue and system in the first week than in the other weeks except on the third week when less expectant mother waited in the system.

Keywords : queuing theory, patient satisfaction, performance parameters, ante-natal clinic

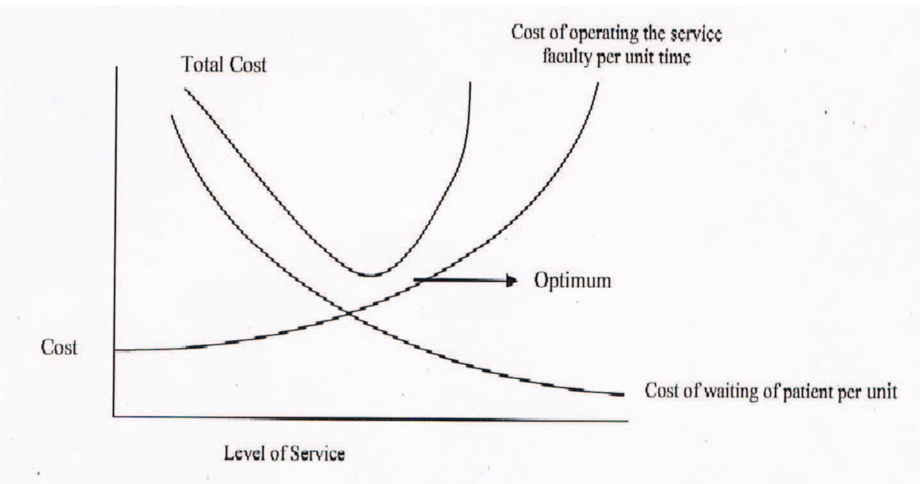
Introduction : Waiting in lines seems to be part of our everyday life. At the hospital, filling station, bus stop, or in the canteen, "waiting our turn". Queues form when the demand for a service exceeds its supply (Kandemir-Cavas and Cavas, 2007). In hospitals, patients can wait minutes, hours, days or months to receive medical service- waiting before, during or after being attended to. For many patients or customers, waiting in lines or queuing is annoying (Obamiro, 2003) or negative experience (Scotland, 1991). The unpleasant experience of waiting in line can often have a negative effect on the rest of a customer's experience with a particular firm. The way in which managers address the waiting line issue is critical to the long term success of their firms (Davis et al, 2003)

Managing the length of the line is one of the challenges facing most hospitals. A few of the factors that are responsible for long waiting lines or delays in providing service are: lack of passion and commitment to work on the part of the hospital staff (Belson, John Kolade Obamiro,1988) overloading of available doctors, doctors attending to patients in more than one clinic etc. These put doctors under stress and tension, hence tends to dispose off a patient without in-depth probing or treatment, which often leads to patient dissatisfaction (Babes and Sarma, 1991). This paper is based on the understanding that most of these difficulties can be managed by using queuing model to determine the waiting line performance such as: average arrival rate of expectant women, average service rate expectant women, system utilization factor, cost of service and the probability of a specific number of women in the system. The purpose of this study is to provide insight into the general background of queuing theory and its associated terminology, and how queuing theory can be used to model ante-natal of clinic of a Civil General Hospital, Wardha. The resultant performance variables can be used by the policy makers to increase efficiency, improve the quality of care, as well as decrease cost in hospital organizations and services.

Literature Review : Literature on queuing indicates that waiting in line or queue causes inconvenience to economic costs to individuals and organizations. Hospitals, airline companies, banks, manufacturing firms etc., try to minimize the total waiting cost, and the cost of providing service to their customers. Therefore, speed of service is increasingly becoming a very important competitive parameter (Katz, et al, 1991). Davis et al (2003) assert that providing ever-faster service, with the ultimate goal of having zero customer waiting time, has recently received managerial attention for several reasons. First, in the more highly developed countries, where standards of living are high, time becomes more valuable as a commodity and consequently, customers are less willing to wait for service. Second, this is a growing realization by organizations that the way they treat their customers today significantly impact on whether or not they will remain loyal customers tomorrow. Finally, advances in technology such as computers, internet etc., have provided firms with the ability to provide faster services. For these reasons hospital administrators, physicians and managers are continuously finding means to deliver faster services, believing that the waiting will affect after service evaluation negatively. Also, understanding the inefficiencies in the hospital and improving them is crucial for making health care policy and budgeting decisions (Wilson and Nguyen, 2004). Cochran and Bhati (2006) also argue that higher operational efficiency of the hospital is likely to help to control the cost of medical services and consequently to provide more affordable care and improve access to the public. Addressing the problems of queuing involve a trade-off between the costs of customers waiting time and the cost of providing faster service. This trade-off is illustrated in Figure 1. Researchers have argued that service waits can be controlled by two techniques: operations management or perceptions management (Katz et al, 1991). The operation management aspect deals with the management of how patients (customers), queues and servers can be coordinated towards the goal of rendering effective service at the least cost.

Patients' perception of health care has gained increasing attention over the past 20 years (Sitzia and Wood 1997). Patients' evaluation of service quality is affected not only by the actual waiting time but also by the perceived waiting time. The act of waiting has significant impact on patients' satisfaction. The amount of time customers must spend waiting can significantly influence their satisfaction (Davis and Vollman, 1990). Furthermore, research has demonstrated that customer satisfaction is affected not just by waiting time but also by customer expectations or attribution of the causes for the waiting (Taylor, 1994). Consequently, one of the issues in queue management is not only the actual amount of time the customer has to wait, but also the customer's perceptions of that wait (Davis and Heineke, 1994). Obviously, there are two approaches to increasing customer satisfaction with regard to waiting time: through decreasing actual waiting time, as well as through enhancing customer's waiting experience (Katz, Larson, and Larson, 1991; Davis and Heineke, 1994).

Fig 1. Costs Optimization of Queuing Systems



Queuing theory is basically a mathematical approach applied to the analysis of waiting lines within the field of operations management (Nosek and Wilson 2001). Any system in which arrivals place demand upon a finite capacity resource may be termed as a queuing system (Singh, 2006). In the case of ante-natal care unit, it can be found that pregnant women arrive or demand services randomly. The objective of queuing analysis and its application in health organizations is to "minimize costs" to the organization- both tangible and intangible. The rising cost of health care can be attributed not only to ageing population expensive and advanced treatment modalities but also to inefficiencies in health delivery. Queuing theory application is an attempt to minimize the cost of providing health care services through minimization of efficiencies and delays in the system (Singh, 2006).

Queuing theory uses queuing models or mathematical models and performance measures to assess and hopefully improve the flow of customers through a queuing system (Gorney, 1981; Bunday,1996;).A good patient flow means that the patient queuing is minimized while a poor patient flow means patients suffer considerable queuing delays (Hall, 1999). Queuing theory has many applications and has been used extensively by the service industries (Nosek and Wilson, 2001). A queuing system or waiting line phenomenon consists essentially of six major components: the population, the arrival, queues itself queue discipline, service mechanism, departure or exit.

The population source serves as where arrivals are generated. Arrivals of patients at the hospital may be drawn from either a finite or an infinite population. A finite population source refers to the limited size of the customer pool. Alternatively, an infinite source is forever.

The queue discipline is the sequence in which customers or patients are processed or served. The most common discipline is first come, first served (FCFS). Other disciplines include last come, first served (LCFS) and service in random order (SIRO). Customers may also be selected from the queue based on some order of priority (Taha, 2005).

The service mechanism describes how the customer is served. It includes the number of servers and the duration of the service time-both of which may vary greatly and in a random fashion (Nosek and Wilson, 2001). The number of lines and servers determines the choice of service facility structures. The common service facility structures are: single-channel, single – phase; single-channel, multiphase; multi-channel, single phase and multi-channel, multiphase.

The departure or exit occurs when a customer is served. The two possible exit scenarios as mentioned by Davis (2003) are: (a) the customer may return to the source population and immediately become a competing candidate for service again; (b) there may be a low probability of re-service.

Application Of Queuing Theory In Ante-Natal Unit :

The health systems ability to deliver safe, efficient and smooth services to the patients did not receive much attention until mid 1990's (Singh, 2006). Several key reimbursement changes, increasing critiques and cost pressure on the system and increasing demand of quality and efficacy from highly aware and educated patients due to advances in technology and telecommunications, have started putting more pressure on the healthcare managers to respond to these concerns (Singh, 2006).

Queuing theory manages patient flow through the system. If patient flow is good, patients flow like a river, meaning that each stage is completed with minimal delay. When the system is broken, patients accumulate like a reservoir (Hall, 1991). Healthcare systems resemble any complex queuing network in that delay can be reduced through: (i) Synchronization of work among service stages, (ii). Scheduling of resources (e.g. doctors and nurses) to match patterns of arrival and, (iii) constant system monitoring (e.g. treating number of patients waiting by location, diagnostic grouping) linked to immediate actions (Hall, 1991). Recently, application of stochastic methods has increased in analyzing clinical problems (Kandemir-Cavas and Cavas, 2007). Queuing theory, as the most common application of the stochastic process, examines queues or waiting lines dealing with random input and servicing processes (Wu, 1998).

Materials And Methods :

Model Specification

The study adopted M/M/C: FCFS/∞/∞, where;

M=Markovian (or poisson) arrivals and exponential service time.

C= Multi-server;

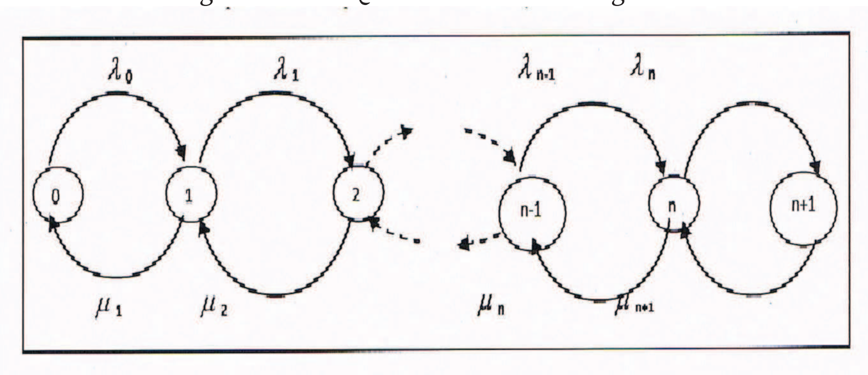
FCFS = First come, first served;

∞ = Infinite system limit;

∞ = Infinite source limit.

For the purpose of modeling, the arrivals (n) are the pregnant women. As each reaches the clinic, she books for service. If service is rendered immediately she leaves the clinic or otherwise joins the queue. The doctors are, of course, the servers (C). Medhi (2003) asserts that if there are n (≥C) in the system, then all the C channels are busy and the interval between two consecutive service completions is exponential with rate cμ.

Fig. 2. Poisson Queues Transition Diagram



Thus, we have a birth-death model having constant arrival (birth) rate (λ) and state- dependent service (death) rate as illustrated in figure 2. Service time (μ- duration of stay) was calculated as discharge time plus admission time.

The arrival rate, service time and numbers of servers were data used for the study which was collected using observation method. This research instrument was adopted so that the queuing system can be examined naturally. The study covers a period of three weeks in which the first three days of each week were taken into consideration. Monday to Wednesday of each week were considered because they are the busiest days of the week.

Procedures Of System Parameters Estimation :
The system performance parameters used in the study were defined as follows:
Λ: Arrival rate of expectant mother per hour;
$\lambda_{eff} = \lambda$, because there no limit on the number in the system;
μ: Service rate (Length of stay) of expectant mother per hour;
C: Number of doctors (servers).
In this model, there are c parallel servers
P : Ante-natal care system utilization factor = λ/C μ, since the recent study adopted - a multiple- Server model:
Lq: Average number of pregnant mother in the queue =

$$\frac{P^{C+1}}{(C-1)!\,(C-P)^{2}} Po \qquad (1)$$

Ls: Average number of pregnant mother in the system = Lq +p (2)
Wq: Waiting time of pregnant mother in the queue = Lq/ λ (3)
Ws: Waiting time of expectant mother in the system = LS/ λ (4)
Pn = probability of n expectant mothers existing in the system.

$$\begin{cases} \dfrac{\lambda^{n}}{\mu(2\mu)(3\mu)\ldots(n\mu)} Po = \dfrac{\lambda^{n}}{n!\mu^{n}} Po = \dfrac{P^{m}}{n!Po}, & n < C \\ \dfrac{\lambda^{n}}{\mu(2\mu)\ldots(C-1)\mu\,(C\mu)^{n-C+1}} Po = \dfrac{\lambda^{n}}{C!\,C^{n-C}\mu^{n}} Po & n \geq C \end{cases} \qquad \ldots\ldots\ldots\ldots\ldots\ (5)$$

Po = Possibility of 0 expectant mothers existing in the system. (6)

$$P_o \left\{ \sum_{n=0}^{c-1} \frac{P^{n}}{n!} + \frac{P^{c}}{c!}\left(\frac{1}{1-\frac{P}{c}} \right) \right\}^{-1}, \frac{P}{c} < 1$$

Results And Discussion : Data Presentation
A model of ante-natal unit flow was then analyzed using TORA optimization system(window based software) and standard queuing formulae. Tora is windows-based software that offers automated or tutorial mode to some operations models such as linear programming, transportation models, queuing models project planning etc. The automated mode which was used in this study provides final solution to real life large mathematical programming models.

As mentioned earlier, the ante-natal unit was modeled as a multi channel, single phase system of identified parallel servers that attend to pregnant women randomly according to exponentially distributed service times using first come, first served queuing discipline. The summary of average arrival rate and average service rate per hour plus number of servers are shown below (Table 1.) in the TORA input for the model:

Table 1. Summary of average arrival rate and average service rate per hour Plus number of servers

Scenario	Lambda	Mu (μ)	Number of Servers	System Limit	Source Limit
Scenario 1	28.00	14.00	5	infinity	infinity
Scenario 2	21.00	12.00	3	infinity	infinity
Scenario 3	14.00	9.00	3	infinity	infinity

Results
Queuing Output Analysis
Scenario1-(M/M/5): (FCFS/infinity/infinity)
Lambda = 28.00000;
Mu = 14.00000;
Lambda eff = 28.00000;
Rho/c = 0.40000;
Ls = 2.03980;
Lq = 0.03980;
Ws = 0.07285;
Wq = 0.00142.

Table 2. Results of M/M/5: FCFS/infinity/infinity

n	Probability, Pn	Cumulative, Pn
1	0.13433	0.13433
2	0.26866	0.40299
3	0.17910	0.85075
4	0.08955	0.94030
5	0.03582	0.97612
6	0.01433	0.99045
7	0.00573	0.99618
8	0.00229	0.99847
9	0.00092	0.99939
10	0.00037	0.99976

Output Analysis
Scenario2-(M/M/3): (FCFS/infinity/infinity)
Lambda = 21.00000;
Mu = 12.00000;
Lambda eff = 21.00000;
Rho/c = 0.58333;
Ls = 2.21712;
Lq = 0.46712;
Ws = 0.10558;
Wq = 0.02224.

Table 3. Results of M/M/3: FCFS/infinity/infinity

n	Probability, Pn	Cumulative, Pn
0	0.15564	0.15564
1	0.27237	0.42802
2	0.23833	0.66634
3	0.13902	0.80537
4	0.08110	0.88646
5	0.04731	0.93377
6	0.02760	0.96137
7	0.00939	0.98685
8	0.00939	0.98685
9	0.00548	0.99233
10	0.00320	0.99553

Queuing Output
Scenario3-(M/M/3): (FCFS/infinity/infinity)
Lambda = 14.00000;
Mu = 9.00000;
Lambda eff = 14.00000;
Rho/c = 0.51852;
Ls = 1.83240;
Lq = 0.27685;
Ws = 0.13089;
Wq = 0.01977.

Table 4. Results of M/M/3: FCFS/infinity/infinity

Probability, Pn		Cumulative, Pn
0	0.19730	0.19730
1	0.30691	0.50422
2	0.23871	0.74293
3	0.12378	0.86670
4	0.06418	0.93088
5	0.03328	0.96416
6	0.01726	0.98142
7	0.00895	0.99036
8	0.00464	0.99500
9	0.00241	0.99741
10	0.00125	0.99866

Queuing Output Analysis

Table 5. Comparative Analysis of Application of Queueing Theory in Ante-natal Care Unit

S*	C	Lam-bda	Mu	L'daeff	Po	Ls	Lq	Ws	Wq
1	5	28.000	14.00	28.00	0.13	2.03	0.03	0.07	0.001
2	3	21.000	12.00	21.00	0.15	2.21	0.4	0.10	0.022
3	3	14.000	9.000	14.00	0.19	1.83	0.27	0.13	0.019

Discussion of Results

Queuing theory was first developed for studying queuing phenomena in commerce, telephone traffic, transportation and business-industrial servicing etc (Gross and Haris, 1985). Application of queuing theory to model hospital settings has been published (Ivahs and Millard, 2003; Adele and Berry' 2005; Vasanawala and Desser 2005). Although the application to ante-natal clinic was not found in literature but these three (3) weeks experience illustrates that queuing theory may be used to accurately model ante-natal unit system in a large hospital operating at or near capacity.

In scenario 1, where the arrival rate (Lambda effective) = 28 patients; number of servers is 5 and service rate (Mu) is 14 patients. The system performance parameters are as follows; Lq = 0.03980. This implies there are 0.03980 pregnant women in the queue waiting to be served by doctors.

Ls = 2.03980. This measures the average number of pregnant women in the system. That is, there are 2.03980 pregnant women in the system.

Wq = 0.00142, meaning that pregnant women spent 0.00142hour(0.85 minutes) on the queue waiting to be attended to by the doctor.

Ws = 0.07285 this means that pregnant women spent 0.07285hour (4.37 minutes) in the system. The time spent before joining the queue, waiting in the queue to be served and time spent after being served before departure.

The probability of 0-13 expectant women in the system was also determined.

Scenario 2: λ, which is the average arrival rate are 21 i.e. expectant mothers per hour with 3 doctors who attended to 12 pregnant women in an hour. The average number of expectant mothers in the queue and system is 0.46712 and 2.21712 respectively. The expectant mothers spent 0.0224hour (1.34 minutes) and 0.10558(6.33 minutes) in the queue and system respectively. In this case, server utilization factor is 58.33%, which is better than the case in the first day.

In scenario 3, the arrival rate per hour is 14 expectant mother, 9 of them were served within an hour by 3 doctors. The number of expectant mothers in the queue and system is 0.27685 and 1.8340 respectively. The expectant mother spent 0.01977(1.19 minutes) and 0.13089(7.85 minutes) in the queue and system. The system server busy factor is 51.85%.

Considering the comparative analysis of the three scenarios, pregnant women spent less time in the queue and system in the first week than in the other two weeks, likewise there are less average of them in the queue and system on the same first week than in the other weeks except on the third week when less expectant mother waited in the system.

Limitations of the Study

Despite its contributions, this study also has its own limitations. Most particularly, we recognize that the data was gathered during three weeks in one hospital, this may limit the generalization of our findings. As a result, a further study is recommended to cover some selected hospitals for a longer period of time.

Moreover, only a unit (ante-natal clinic) of the hospital was considered in this study. That is, a queuing network model should be considered to analyse the patient flow of the entire hospital. Additionally, this study takes into consideration the actual waiting time but ignores the effect of perception of waiting time on patient satisfaction.

References

1. Adedayo, O. A., Ojo, O., Obamiro, J. K., Operations Research in Decision Analysis and Production Management. Lagos: Pumark Nig. Ltd., 2006.
2. Adele, M., Barry, S., Modeling patient flow in hospital using queuing theory, Unpublished Manuscript, 2005.
3. Babes, M., Serma, G. V., Out-patient Queues at the Ibn-Rochd Health Centre, Journal of Operational Research, 42 (10): 1086-1087, 1991.
4. Belson, G. V., Waiting Times Scheduled Patient in the Presence of Emergency Request, Journal of Operational Management 2(3), 1988.
5. Bunday, B.D., An Introduction to Queuing Theory, New York: Halsted Press, 1996
6. Chaska, L. N. et all., Use of Medical Services and Satisfaction with Ambulatory Care among a Rural Minnesota Population, Public Health Reports, 95 (1): 44-52, 1980.
7. Cochran, K.J., Bharti, A., 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, 2006.
8. Davis, M. M., Aquilano, J. N., Chase, B. R., Fundamentals of Operations Management. Boston: McGraw- Hill Irwin, Fourth Edition, 2003.
9. Davis, M.M., Heineke, J., How Disconfirmation, Perception and Actual Waiting Times Impact, Unpublished Manuscript, 1998.
10. Davis, M.M., Vollmann, T.E., A framework for Relating Waiting Time and Customer Satisfaction in a Service Operation, Journal of Services Marketing, 4 (1) 61-69, 1990.
11. Gorney, L, Queuing Theory: A Problem Solving Approach, New York: Petrocelli Booko, Inc., 1981.
12. Gross, D., Harris, C.M., Fundamentals of Queuing Theory. New York: John Wiley, Second Edition, 1985.
13. Hall, W. R., Patient Flow: The New Queuing Theory for Healthcare, OR/Ms, Today, 2006.
14. Ivalis, S., Millard, P., Health care modeling operating the black box, British Journal of Health Care Management, 8(7):251-255, 2003.
15. Kandemir-Caues, C., Cauas, L., An Application of Queuing Theory to the Relationship between Insulin Level and Number of Insulin Receptors, Turkish Journal of Biochemistry, 32 (1): 32-38, 2007.
16. Katz, K., Larson, K., Larson, R., Prescription for the waiting in line Blues: Entertain, Enlighten and Engage", Sloan Management Review (winter): 44-53, 1991.
17. Nosek, A.R., Wislon, P. J., Queuing Theory and Customer Satisfaction: A Review of Terminology, Trends and Applications to Pharmacy Practice, Hospital Pharmacy, 36(3):275-279, 2001.
18. Scotland, R., Customer Service: A Waiting Game, Marketing, pp 1-3, 1991.
19. Sitzia J., Wood N., Patient Satisfaction: A Review of Issues and Concepts, Soc Sci Med, 45: 1829–1843, 1997.
20. Taylor, S., Waiting for Service: The Relationship between Delays and Evaluation of Service, Journal of Marketing, 58(2): 56-69, 1994.
21. Young, G.J., Meterko, M., Desal, R. K., Patient Satisfaction with Hospital Care: Effects of Demographic and Institutional Characteristics, MedicereCcare, 38(3): 325-334, 2000.

***

15, Swami Samarth Nagari,Near New Arts & Commerce College
Nalwadi, Wardha -442001 ,(MS), India
Asst. Professor in Biostatistics, Dept. Community Medicine
J.N.M.C. Sawangi (M), Wardha, kushbabar@yahoo.com

Wanjari Fail, Near Jijau B.Ed.,College Yavatmal – 445001 (MS), India,
Asst. Professor in Biostatistics, Dept. Community Medicine
J.N.M.C. Sawangi (M), Wardha, yogesh.raut2@rediffmail.com

Professor, Dept. of Statistics,
PG Teaching Department of Statistics,R.S.T.M. Nagpur University
Amravati Road, Nagpur