Queuing theory and the management of Waiting-time in Hospitals: The case of Anglo Gold Ashanti Hospital in Ghana

Sam Afrane
College of Art and Social Science, KNUST, Kumasi

Alex Appah
Anglo Gold Ashanti Hospital, Obuasi

Abstract
Queuing is a major challenge for healthcare services all over the world, but particularly so in developing countries. Application of queuing theory to enhance decision making to improve this problem is not commonly used by managers in developing countries in contrast to their counterparts in the developed world. This study investigates the application of queuing theory and modelling to the queuing problem at the out-patient department at AngloGold Ashanti hospital in Obuasi, Ghana. Using a descriptive, observational and ex-post facto case study approach, data was generated, analysed and used to model five capacity scenarios for the outpatient department. It is established that the optimum system performance can be achieved with eight doctors effectively at post from contrast to the prevailing five doctors that were effectively at post. In sum, the study establishes that applying queuing theory and modelling to queuing and capacity challenges can enhance decision making with regards to what will provide optimal performance.

Introduction
Queuing is a challenge for all healthcare systems. In the developed world, considerable research has been done on how to improve queuing systems in various hospital settings. This, unfortunately, has not been the case in developing countries like Ghana. This paper seeks to contribute to this subject by analysing the queuing situation in public hospitals in Ghana and also to bring its practical value to how decision making can be enhanced in hospitals.

Queuing theory is a potent mathematical approach to the analysis of waiting lines performance parameters in healthcare delivery systems (Ozcan, 2006). It has increasingly become a common management tool for decision making in the developed world. This vital tool is unfortunately minimally used in most healthcare systems in Ghana and other African countries. Review of extant literature establishes the use of queuing theory and modelling in

improving waiting time in various hospital settings (McQuarrie, 1983; Green 2006a; Siddahartan, Jones and Johnson, 1996). It has also been used in reducing costs relating to various aspects of healthcare (Keller and Laughhum, 1973) and generally improving system performance (Murray, 2000) in hospital systems. .

Long waiting queues are symptomatic of inefficiency in hospital services. Unfortunately this is the case in many public hospitals in Ghana and other developing countries. Capacity management decisions in Ghanaian hospitals are generally based on experience and rule of the thumb rather than with the help of strategic research model-based analysis. Like many big hospitals in Ghana, AngloGold Ashanti (AGA) Hospital in Obuasi, Ghana receives a large number of patients every day and this generally results in long patient waiting times. In response to this challenge, this paper analyses the queuing system of the outpatient department of AGA Hospital in order to develop a model that can help reduce the waiting time of patients. Specifically, the paper seeks to construct a structural model (or flowchart) of patient flow within the OPD and to model a queuing system using the queuing theory to minimize patient waiting times in the hospital.

Application of queuing theory to model hospital settings has been widely published (Ivalis and Millard, 2003; Adele and Barry, 2005; Vasanawala and Desser, 2005). Also, the use of queuing analysis and simulation to enhance performance at various hospital departments has been widely researched (Green, 2002), (Kim et al., 1999) and emergency departments (Green et al, 2006). In most healthcare settings, unless an appointment system is in place, the queue discipline is either first-in-first-out or a set of patient classes that have different priorities (as in an emergency department, which treats patients with life-threatening injuries before others). McQuarrie (1983) showed that it is possible to minimize waiting times by giving priority to clients who require shorter service times. Green (2006a), also provide models for queue disciplines while Siddhartan and Associates (1996) analyse the effect on patient waiting times when primary care patients use the Emergency Department. They propose a priority discipline for different categories of patients and then a first-in-first-out discipline for each category. Singh (2006) looked into minimizing total cost incurred and also minimizing the waiting costs by comparing the outputs for two nurses, three nurses and four nurses by evaluating the performance measures for each of the scenarios. In that study, it was found that scenario of 3 nurses was the optimal solution with optimum trade-off between the two types of cost involved in queuing models. In another study, Obamiro (2003) also applied the queuing theory in a study to determine the optimum number of nurses required in an antenatal clinic to reduce the time spent by pregnant women in the queue and the system. Queuing theory and modelling can thus be said to be useful modern tools for decision making on issues of capacity and resourcing.

Also the number of messengers required to transport patients or specimens in a hospital by assigning costs to the messenger and to the time during which a request is in queue was determined by Gupta and Associates (1971). In a queuing network, a patient may have to go through several nodes and consequently several queues in order to obtain the desired service. Nodes where the ratio of demand to available service capacity is relatively high become bottlenecks. Such bottlenecks increase overall patient waiting times even though other nodes may have low utilization. One of the major elements in improving efficiency in the delivery of health care services is patient flow. Good patient flow means that patient queuing is

minimized and poor patient flow means that patients suffer considerable queuing delays (Hall et. al, 2006). Effective resource allocation and capacity planning are determined by patient flow because it informs the demand for health care services (Murray, 2000). Queuing theory provides exact or approximate estimation of performance measures for such systems based upon specific probability assumptions. In a hospital, these assumptions rarely hold, and so results are approximated (Cochran & Bharti, 2006).

RESEARCH METHODOLOGY
The study adopts a descriptive, observational and ex post facto case study approach. In depth review of hospital OPD attendance records from January to June, 2010 was made. Interviews with management, doctors, and records staff were conducted to validate the secondary data and to gather information required to construct the structural model of the routings in and out of OPD. Direct observations were used to model patient average arrival and length of stay. Questionnaires were also used to gather information on daily arrival rates, patients' view on queuing at the hospital, waiting time to consult a doctor, etc.

The survey population of the study was the entire out-patients of AGA Hospital in Obuasi during the period of the study. In all, 475 patients were surveyed for this study. Purposive sampling, a non-probability sampling technique in which the researcher selects a group of people because they have particular traits that the researcher wants to study was used. The out-patient department was selected because it had the greatest queuing challenge compared to the other units in the hospital.

RESULTS
The survey revealed that patients arrive at the outpatient department and drop their hospital card in a box and wait their turn. A nurse then attends to them and allocates a consulting room. From the consulting room, patients are sent to the laboratory and the radiology department for laboratory tests or medical imaging. Those who do not require these for diagnosis are given treatment and proceed to the data entry clerk with their prescriptions, entries are done and they collect their drugs from the pharmacy and go back home. Those who return from the laboratory and radiology department also go through this same process. Patients who require admission are directly admitted to the ward from the consulting room and care continues from there. This process is sketched in Fig 1.

Flow Chart for AGA OPD
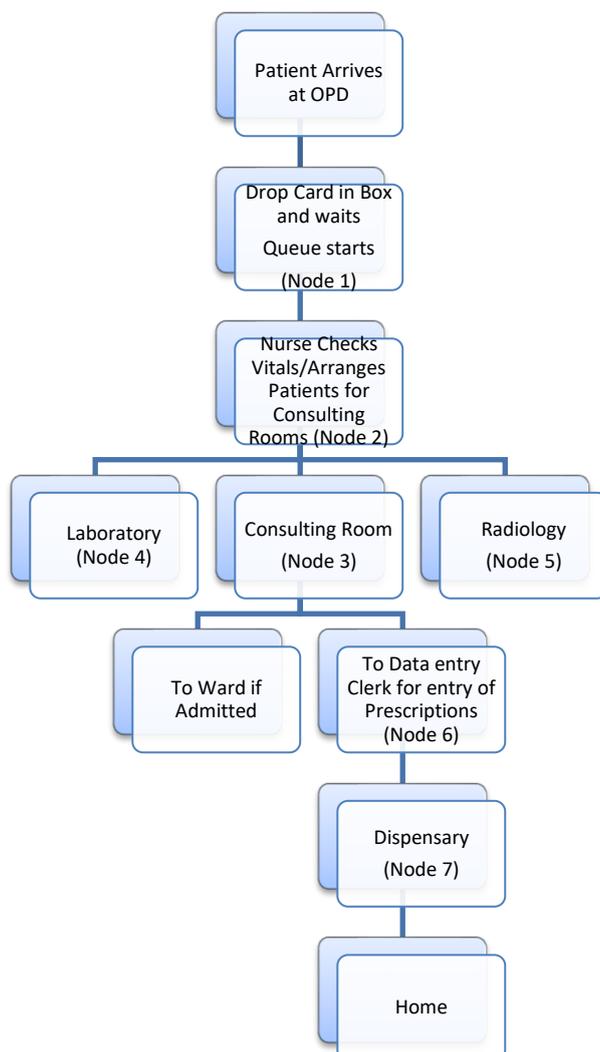
The Average Daily Attendance (ADA) was estimated using both secondary and primary data from the OPD. Secondary data from January 2012 to June 2012 gave a daily average of 455 patients [(421.45 + 488.69 + 456.45 + 421.70 + 427.81 + 500.77) ÷ 6] (See Fig 2). Looking at the secondary data, the distribution of the attendance was busiest on Mondays and Fridays. Accordingly, using a worst case scenario and within the limit of resource capacity, an observational study was done on the other four days to establish the OPD attendance (i.e., 6th and 9th June, 2012; and 13th and 16th of July 2012). The ADA derived from the data was 475 patients [(464 + 498 + 531 + 449) ÷ 4]. Noting that the difference between the secondary and primary values was not that significant, (barely 4 per cent), the ADA from the observational study (i.e., worst case scenario) was used.

Fig 1- Graphical Illustration of Daily Attendance (Jan – Jun 2012)
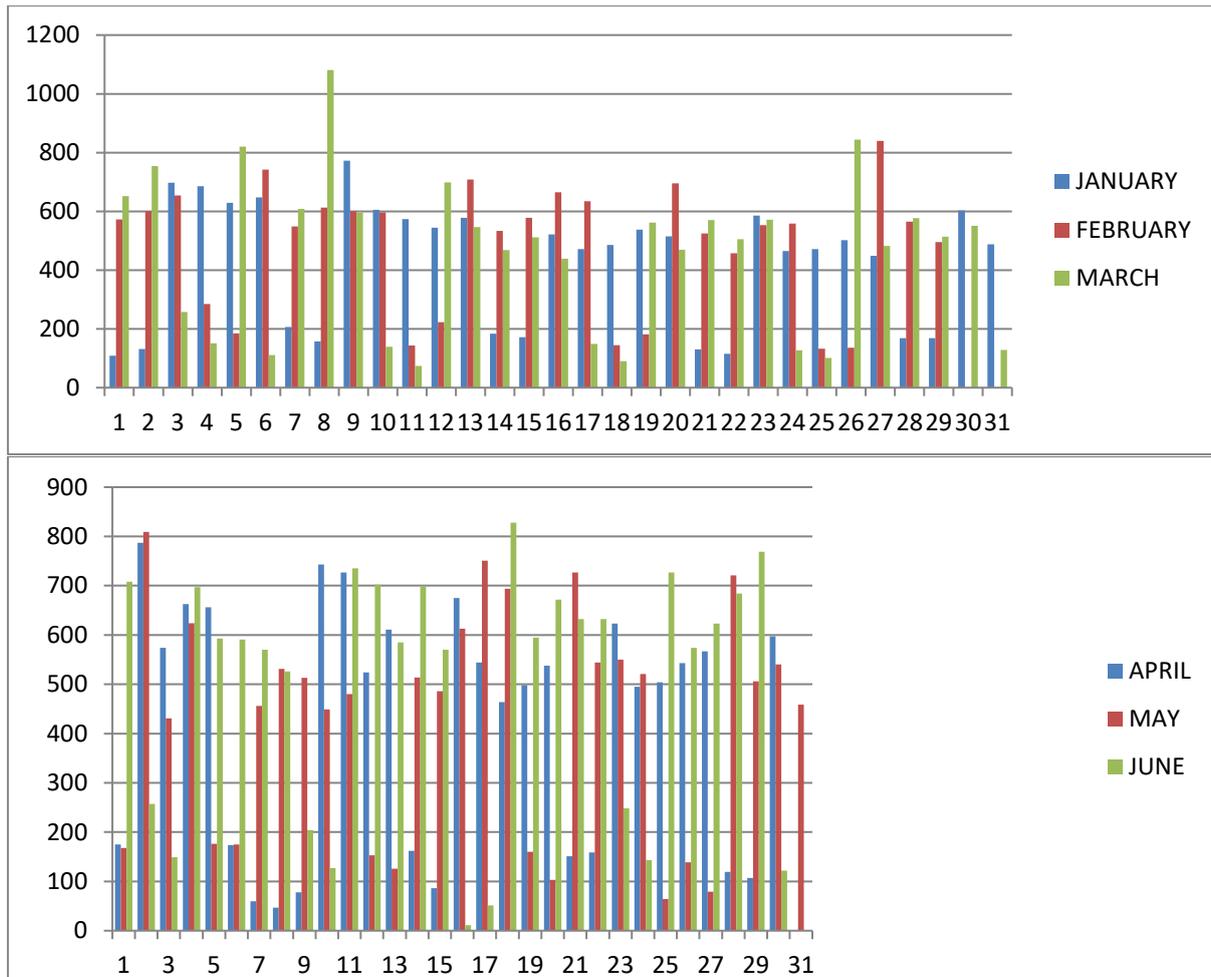
3.1 SERVICE-RELATED OPD INDICATORS

Average Daily Arrival Rate

The outpatient clinic is open from 7.30 am till 4.30pm giving a total of 9 hours per day. It is important to note that after 4.30 pm there is still an emergency out-patient clinic but for the purposes of this study that aspect was not included. Using the information above the average patient arrival rate is given by: 475 patients per day divided by 9 hours per day resulting 53 patients per hour.

Service Rate

From interviews with the doctors, one doctor averagely uses 5 minutes to complete the care of one patient. This implies that a doctor can see 12 patients per hour.

Service Capacity

A total of five doctors and three medical assistants attend to patients at the OPD. Hence, for the purpose of this study the number of doctors (servers) is considered to be eight.

Average Time It Takes Before Seeing a Doctor

Out of 145 questionnaires administered, 141 responded to this question and the data generated is shown in Table 2.

Table 1- Average Time Taken before to Seeing a Doctor

	Number of Respondents	Percentage (%)

36

Average Time Taken Before Seeing a Doctor (Hours)		
0 - 0.5	4	2.84
0.5 - 1.0	24	17.02
1.0 – 2.0	47	33.33
2.0 – 3.0	58	41.13
>3.0	8	5.67
Total Number of Respondents	141	100

Source: Field Survey June/July 2012

From the above data, the modal average time it takes to see a doctor on arriving for care at the hospital is between 2 to 3 hours. The median time also falls within 2-3 hours. Therefore most people will spend an average of 2.5 hours before seeing a doctor at the OPD.

Patients Views on Queuing at the Hospital
Questionnaires were administered to determine the views of patients in relation to the effects of queuing on them at the hospital. Out of 143 respondents, 98% experienced negative effects with queuing at the hospital. Out of this number, 43% felt frustration with the service, 38% tiredness from waiting, 15% anxious about their health and 4% a desire not to use the facility again. With regards to satisfaction with the services provided, out of 140 respondents, 67.6 % were satisfied, 26.06% undecided, and 6.34% unsatisfied. An overwhelming majority of 93% of respondents to the questionnaire blamed insufficient staff as the main cause of the delays.

Queuing is therefore a problem of concern at the hospital, as is the case in several other hospitals in the country. There is therefore a need to take steps to improve the queuing situation at the hospital. One such way is by application of queuing theory and modelling to rectify the queuing situation, as has been done in other places discussed in the literature review. The study accordingly goes on to model five scenarios of server capacity at the hospital applying the queuing theory to establish whether this can inform an optimal scenario setting that will reduce the queuing burden.

3.2 USING QUEUING MODEL (QM) FOR WINDOWS SOFTWARE TO MODEL SCENARIOS
Five scenarios of server capacity was considered, comparing capacity utilization, average number in the queue, average number in the system, average waiting time in the queue and average time in the system by patients using Software: QM for Windows. Based on the input data below, sensitivity analysis was carried out and the results are presented in Table 3 and Fig. 2.

Table 2- Input Parameters for QM for Windows models

Parameter M/M/1 (exponential service times)	Value
Arrival rate (Λ)	53 patients per hour
Service rate (mu)	12 patients per hour

Number of servers (doctors)	5, 7, 8, 10 and 12 depending on scenario

Table 3- Sensitivity Analysis of 5 Model Scenarios

| Scenarios	Avg. Service Utilization	Avg. no in the queue (Lq)	Average no. in system (Ls)	Avg. time in queue (Wq)	Avg. time in system (Ws)
Scenario 1 (5 doctors)	0.7833	1.8888	5.8054	2.4112	7.4112
Scenario 2 (7 doctors)	0.5595	0.1572	4.0739	0.2007	5.2007
Scenario 3 (8 doctors)	0.4896	0.0510	3.9677	0.00651	5.00651
Scenario 4 (10 doctors)	0.3917	0.0049	3.9216	0.0063	5.0063
Scenario 5 (12 doctors)	0.3264	0.0004	3.9171	0.0005	5.0005
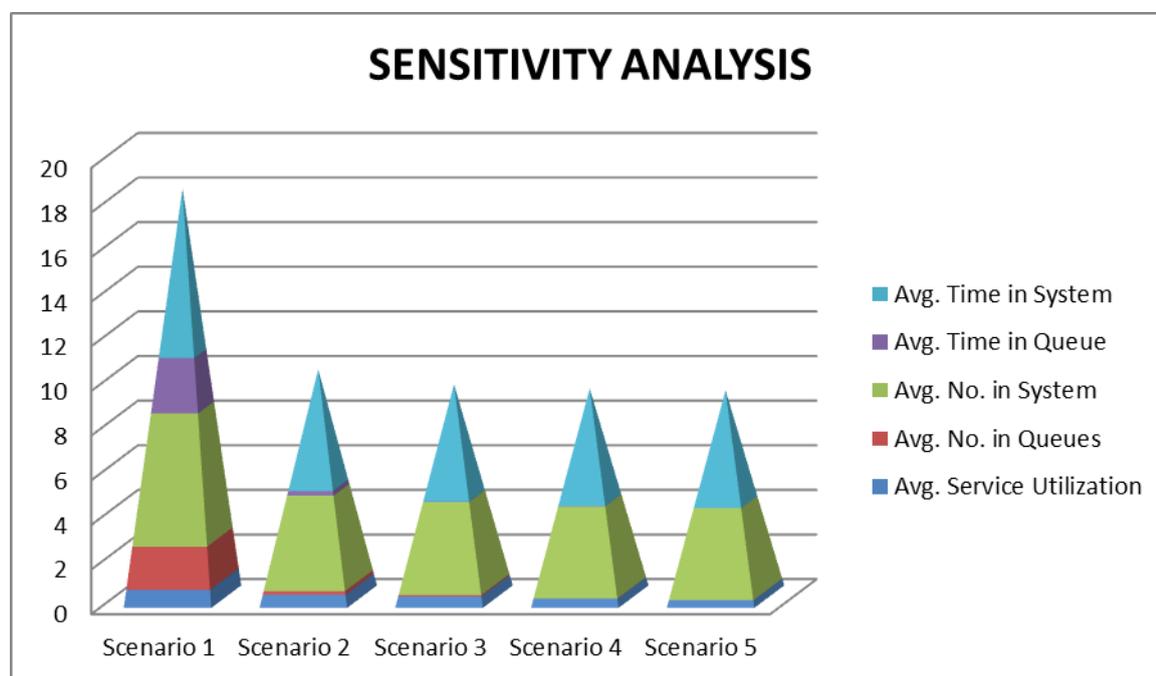
Fig 2 - Graphical Plot of Results of Sensitivity Analysis

At AGA hospital, there are a total of eight doctors (servers) that provide service to patients at the OPD department. However, five of these doctors, start the day by spending an average of two hours seeing cases on the wards before starting their work at the outpatient clinic. They also take one hour lunch break. Accordingly, they spend effectively five hours providing

service at the outpatient department. The other three doctors also take one hour lunch break and so effectively consult for eight hours. From the above, in practice the total number of effective hours of work at the OPD department by the eight doctors is 49 hours [given by (5x5) + (8x3)]. This will effectively represent 5.44 doctors at post (given by 49/9) or by approximation 5 doctors at post.

From the results of the sensitivity analysis for the modelling, scenario 2 with five doctors at post produces inferior performance in comparison to the others with average number in the system being 5.8054 (Ls), average time in queue (Wq) 2.4112, average number in the queue (Lq) 1.8888, and average time in the system (Ws) being 7.4112. Scenario 3 (8 doctors at post) provides optimal performance of the system with average number in the system (Ls) being 3.9677, average time in queue (Wq) 0.00651, average number in the queue (Lq) 0.0510, and average time in the system (Ws) being 5.00651. The change in time spent in the queue from seven doctors (Scenario 2) to eight doctors at post (Scenario 3) is very appreciable. The average number in the queue (Lq) drops from 0.1572 with 7 doctors to 0.0510 with 8 doctors. Similarly, average time in the queue (Wq) drops from 0.2007 to 0.00651 respectively.

DISCUSSION

The average server utilization is 0.5595 in scenario 2 with 7 doctors and 0.4896 for scenario 3 with 8 doctors. This means that patients spend less time in the queue and the system utilization is quite good. This is in line with the findings by Bailey (1954), which established that in outpatient and inpatient clinics, when the number of servers is below a certain threshold, a clinic develops an infinite queue whereas when it is slightly above this threshold, waiting time and queues are lower. Figure 2 depicts the extremely marginal changes for scenarios 3, 4, and 5 with respect to the average number in the queue and the average time in the queue and this also collaborates the above. One needs to be mindful of the costs involved in achieving these marginal changes. Hiring more doctors will mean taking on more costs. A good balance between the number of doctors, costs, and optimal system performance is important for sustainability, hence the conclusion that scenario 3 with 8 doctors at post is best for optimal performance. This study, however, did not look into costs and it would be interesting to include the cost dimension in another study in future.

In similar studies, using similar software, Singh (2006) looked into minimizing total cost incurred and also minimizing the waiting costs by comparing the outputs for two nurses, three nurses and four nurses by evaluating the performance measures for each of the scenarios. In that study, it was found that scenario of three nurses was the optimal solution with optimum trade-off between the two types of cost involved in queuing models. Queuing theory and modelling can thus be said to be useful modern tools for decision making on issues of capacity and resourcing.

FINDINGS AND RECOMMENDATIONS

The study has established that at the OPD department at AGA hospital, the current situation effectively is one of 5 doctors at post considering their effective working time. There are several nodes (6) in the patient flow at the OPD. The average daily arrival rate is approximately 53 patients per hour and the service rate averages 12 patients per hour. Primary data analysis determined average waiting time before seeing a doctor to be 2.5 hours. Using the queuing

model for windows software, five scenarios modelled (for 5, 7, 8, 10 and 12doctors), established that optimum system performance will be achieved with eight doctors effectively at post from the start to the close of work in stark contrast to the prevailing situation that effectively has five doctors at post.

Based on the findings of this study, and with objective of addressing the queuing challenge at the OPD at AGA hospital, it is recommended that first, the hospital should have a system such that eight doctors effectively consult in the OPD from the start to the close of work for optimal performance instead five currently at post. Secondly, the hospital should introduce a user friendly computerized database system that will reduce the number of nodes in the system and consequently the queuing problem at the OPD. Thirdly, what is needed to be considered ultimately is not only doctors but also the cost implications and the needed supporting staff. A similar study should therefore be conducted that will consider the other dimensions such as costs, support staff, and not just waiting time as was the case in this study.

CONCLUSION

Providing patients with timely access to appropriate medical care is an important element of healthcare delivery and increases patient satisfaction. "When" care is received, is often as important as "what" care is received. This study establishes that patients are generally dissatisfied with long waiting times and experience negative effects as a result. It is further established that queuing theory and modelling is an effective tool that can be used to make decisions on staffing needs for optimal performance with regards to queuing challenges in hospitals. This study should therefore be replicated in other hospitals in Ghana and other countries in order to inform hospital administrators more on the usefulness of the application of queuing theory and modelling as a tool for improved decision making with regards to the queuing challenges that are faced by hospitals.

REFERENCES

Adele, M., & Barry, S. (2005). *Modeling patient flow in hospital using queuing theory.* Unpublished Manuscript.

Albin, S. L., Barrett, J., Ito, D., & Mueller, J. E. (1990). A queuing network analysis of a health center. *Queuing Systems, 7*, 51-61.

Bailey, N. T. J., (1954).Queuing for medical care, Applied Statistics, Vol.3, 137-145.

Brahimi, M., & Worthington, D. J. (1991). Queuing models for out-patient appointment systems - a case study. *The Journal of Operational Research Society, 42*, 733-746.

Cahill, W., & Render, M. (1999). Dynamic Simulation Modeling of ICU Bed Availability. Proceedings of the 1999 Winter Simulation Conference.

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.

Davies, R., & Davies, H. T. (2005). Modelling Patient Flows and Resource Provisions in Health Systems. *Omega - Oxford Pergamon Press, 22*, 123.

Green, L. V. (2002). How many Hospital Beds? *Inquiry, 37(4)*, 400-412.

Green, L. V. (2006a). *Queuing analysis in healthcare, in Patient Flow: Reducing Delay in Healthcare Deliver* (Hall, R.W.ed ed.). New York: Springer.

Green, L. V. (2006b). Using queuing theory to to increase the effectiveness of emergency department provider staffing. *Academic Emergency Medicine, 13*, 61-68.

Green, L. V., Soares, J., Giulio, J., & Green, R. (2006). Using Queuing Theory to Increase the Effectiveness of Emergency Department Provides Staffing. *Academic Emergency Medicine, 13(1)*, 61-68.

Gupta, I., Zoreda, J., & Kramer, N. (1971). Hospital manpower planning by use of queuing theory. *Health Services Research, 6*, 76-82.

Hall, R., Belson, D., Murali, P., & Dessouky, M. (2006). Modeling patient flows through the health care system, in Patient Flow: Reducing Delay in Healthcare Delivery. In *Hall, R.W ed* (Hall, R.W. ed. ed., pp. 1-44). New York: Springer.

Ivalis, S., & Millard, P. (2003). Health care modeling operating the black box. *British Journal of Health Care Management*(8(7)), 251-255.

Keller, T. F., & Laughhunn, D. J. (1973). An application of queuing theory to a congestion problem in an outpatient clinic. *Decision Sciences, 4*, 379-394.

Kim, S., Horowitz, I., Young, K., & Buckly, T. (1999). Analysis of Capacity Management of the Intensive Care Unit in a Hospital. *European Journal of Operational Research, 105*, 346-355.

McQuarrie, D. G. (1983). Hospital utilization levels. The application of queuing theory to a controversial medical economic problem. *Minnesota Medicine, 66*, 679-86.

Murray, S. C. (2000). Understanding the Patient Flow. *Decision Line, March*, 8-10.

Obamiro, J. (2003). *Application of Queing Model in Determining the Optimum number of Service Facility in Nigerian Hospitals, M Sc. Project submitted to Department of Business Administration, University of Ilorin.*

Ozcan, Y. A. (2006). *Quantitative methods in health care management; Techniques and applications* (First edition ed.). Jossey-Bass Publications.

Roche, K. T., Cochran, J. K., & Fulton, I. A. (2007). Improving patieny safety by maximizing fast-track benefits in the emergency department - a queuing network approach, in Proceedings of the 2007 Industrial Engineering Research Conference, pp 619-624.

Rosenquist, C. J. (1987). Queing analysis: a useful planning and management technique for radiology. *Journal of Medical Systems, 11*, 413-419.

Siddhartan, K., Jones, W. J., & Johnson, J. A. (1996). A priority queuing model to reduce waiting times in emergency care. *International Journal of Health Care Quality Assurance, 9*, 10-16.

Singh, V. (2006). *Use of Queuing Models in Health Care Decision Analysis.* Department of Health Policy and Management, University of Arkansas for Medical Science.

Vasanawala, S. S., & Desser, T. S. (2005). Accomodation of requests for emergency US and CT: Applications of queuing theory to scheduling of urgent studies. *Radilogy*(235), 244-249.