International Journal of Mathematical Sciences and Optimization: Theory and Applications
Vol. 9, No. 2, pp. 90 - 101
https://doi.org/10.5281/zenodo.10202665

Queuing Model for Hospital Congestion with Application

F. O. Okoth [1*], A. W. Okoth [2], K. L. Nyongesa [3], J. L. Sirengo [4]

1-4. Department of Mathematics, Masinde Muliro University of Science and Technology, P.O Box 190 - 50100, Kakamega, Kenya.
* Corresponding author: fokoth22@gmail.com

Article Info
Received: 03 October 2023 Revised: 15 November 2023
Accepted: 17 November 2023 Available online: 25 November 2023

Abstract

The emergence of Covid-19 posed a great health challenge worldwide. Health care facilities were stretched beyond limit, with no space to admit new critical patients. This motivated this study, which sought to understand the dynamics around queues, particularly in hospitals. The main objective was to analyze the queuing process between the Emergency Department (ED) and Internal Wards (IW) at Moi Teaching and Referral Hospital, in order to understand patient flow with the view to minimizing congestion. The study modelled the process as a queuing system with heterogeneous server pools, where the pools represent the wards and servers are beds. The system was analysed under various queue-architectures and routing policies, in search for fairness and optimum operational performance so as to enhance the level of access to health care in the facility. The existing models such as Kendall's, Erlang's, Little's Law and De Bruin's models were used to analyze various ward/unit operating characteristics and sufficient bed count was determined to guarantee certain access standards to care. The results indicated that there were long queues arising from blockage which led to higher chances of patients waiting for long hours for services in the facility. To solve the problem the study proposes reassignment of beds within the Hospital wards so that any arriving patient and those waiting for services are examined and allocated beds immediately. This will reduce the long waiting time for services and give the hospital sufficient time in daily operations.

Keywords: Hospital, Queuing Model, Congestion, Queue Discipline, Blocking.
MSC2010: 00A71.

1 Introduction

A hospital is an institution for health care. It should be able to provide patient stay with efficient and optimal service. Hospitals have various medical units specializing in different areas of medicine, e.g., internal, surgery, intensive care and obstetrics among others. In large hospitals, there are several similar medical units operating in parallel. The research focused on the Emergency Department (ED) and its interface with twenty eight Internal Wards (IWs) of Moi Teaching and Referral Hospital (MTRH). The ED caters for immediate threats to health and provides emergency medical services. Thus proper functioning of the ED is of utmost importance, as overcrowding causes delays in attending to new patients, resulting in inefficiency, [1].

The Government of Kenya introduced the 'Big Four Development Agenda' in 2017 namely food security, affordable housing, manufacturing and most importantly provision of universal health care to its ever surging population which then stood at 52.57 million [2]. Hospitals are therefore, increasingly becoming aware of the need to use their resources as efficiently as possible in order to continue to assure the survival and prosperity of their institutions [3]. Moi Teaching and Referral Hospital serves residents of Western Kenya region, with a population of approximately 24 Million, therefore long queues and congestion at the ED is inevitable due to the strained resources like beds and medical staff. Referral hospitals, typically rife with inefficiencies and delays, present a propitious ground for numerous research projects in Science and Operations Research fields on how such operational bottlenecks can be addressed. Previous studies in queue analysis in a hospital set up did not adequately address the aspect of blocking associated with insufficiency of facilities such as beds in key departments. It is against this background that the study set out to analyze the situation using existing queuing models and propose a better model that would help ease congestion at the facility. This would be achieved by having an adequate model of patient flows to and between the different departments of the Hospital.

Further, accurate estimation and forecasting of parameters were prerequisites for consistent service levels and efficient operation. Though a lot had been done in statistical inference and forecasting, comparatively little had been devoted to queuing processes, particularly queuing for services in health care settings [4]. By considering this, the study would be significant in the following ways: Firstly to guide hospital management in formulating policies that would result in enhanced patient service in the ED; secondly, to serve as a basis for further research on use of simulation modeling in other sectors of the economy, and thirdly, to add to the already existing knowledge on the use of simulation modeling in health care settings.

2 Literature Review

Studies have been done on waiting for services in many sectors like business, hospital, banking among others. For instance Yaduvanshi *et al* [5] did research on Application of Queuing Theory to Optimize Waiting-Time in Hospital Operations. Since waiting time is inherent to the healthcare service sector in India and a major challenge faced by almost every big hospital is queuing, long waiting time was said to be a reflection of inefficiency in hospital operations. The outpatient departments had long queues compared to other departments in hospital operations. The study comprised of in-depth analysis of outpatient departments from several dimensions. In many hospitals that serves large populations across India, the out-patient departments of Fortis Escorts Hospital in Jaipur, India is managed using experience and rule of thumb rather than strategic research-based techniques such as queuing theory. The Fortis Escorts Hospital in Jaipur receives a large number of patients each day which results in longer waiting time for patients due to long queues. To address this challenge, a strengths, weaknesses, opportunities, and threats analysis was conducted for the outpatient department of Fortis Escorts Hospital Jaipur (FEHJ) which resulted into dissecting the queuing problem and coming out with solutions knowing where the hospital operations could excel and where there was a scope of improvement to make the working and processes better. Further, after examining the problem analytically and applying queuing theory, measures were suggested to improve the delay points and make the Outpatient department more efficient so as to gain higher patient satisfaction rating. This study analysed the waiting probability and bed requirements so as to solve the problem especially on inpatient services.

Further, Ponkshe *et al* [6] in the essay queuing theory in the healthcare sector, provided an outline of how queueing theory is used and applied in the healthcare industry. The study on mathematical waiting lines or queues is known as queuing theory. The subsequent essay also made an effort to clarify the Strengths Weakness, Opportunities and Threats analysis of the queueing theory's readiness by relating to the COVID-19 pandemic in India and elsewhere. Also the different ways that the theory allows for quicker allocation of resource and very low consumption, proving to be an outstanding gem was discussed. It was also discovered that various proposed queuing models have not yet been adopted by hospital administrators due to a lack of real-world validation. As a result, it is necessary to investigate the relevance and implications of queuing theory by putting a straightforward queuing model to the test at a busy hospital in India. The essay aimed to assess this theory, examples of its application in healthcare institutions around the globe, and its benefits. This study was motivated by COVID-29 pandemic to asses the congestion of Moi Teaching and Referral Hospital by analysing waiting probabilities and number of allocated bed in each ward.

Kalwar *et al* [7], did research on Applications of Queuing Theory and Discrete Event Simulation in Health Care Units of Pakistan. The study indicated that patients are delayed at the public healthcare units long before they are served by medical staff. The delay is the major problem which can be obviously due to poor design or mismanagement of queuing system. For all these problems, queuing theory is the best tool but nowadays, simulation methodology has taken over the side of queuing theory very precisely. This review therefore contributes in highlighting the problems of healthcare in Pakistan and so it focused on the solution provided by the previous researchers as well. Their paper presented the overall picture of healthcare delivery system of Pakistan. Since, the issues of intensive care units (ICUs) and emergency departments (EDs) at the hospitals have specific issues of their own. The major disadvantage of this research paper is that it did not present the in-depth understanding of problems of each facility. When the system (healthcare facility) is congested, patients wait more in the queues and system in order to get served. AT ICUs and EDs patients are in critical conditions and if the patients are made to wait in that condition, anything can happen. It was suggested that review of problems of ED and ICUs should also be reviewed specifically so that the problems can highlighted for the greater good of nation. After the deep review of literature, it was clear that none of the literature reviews discussed the healthcare problems in the respective country and the review of contribution of related methodology at the same time. The contribution of present research in highlighting the problems of healthcare delivery system in Pakistan and review of queuing theory and queuing simulation cannot be ignored.

Dong & Perry [8], carried out a research on Queueing Models for Patient-Flow Dynamics in Inpatient Wards. In the research paper, they proposed a queueing model that takes into account the most salient features of queues associated with patient-flow dynamics in inpatient wards, including the need for a physician's approval to discharge patients and subsequent discharge delays. Fundamental quantities, such as the (effective) mean hospitalization time and the traffic intensity, become functions of the queueing model's primitives. Therefore, begin by characterizing these quantities and quantifying the impacts that the discharge policy has on the average bed utilization and maximal throughput. They introduced a deterministic fluid model to approximate the non stationary patient-flow dynamics. The fluid model is shown to possess a unique periodic equilibrium, which is guaranteed to be approached as time increases so that long-run performance analysis can be carried out by simply considering that equilibrium cycle. Consequently, evaluating the effects of policy changes on the system's performance and optimizing long-run operating costs are facilitated considerably. The effectiveness of the fluid model is demonstrated via comparisons to data from a large hospital and simulation experiments. This study determines the inpatient flow through analysis of bed requirements in each ward so as to minimize congestion in the facility.

On the other hand, it is common practice in health services to estimate the required number of beds as given by Equation (2.1),

$$Bed\ requirement = (X_i \times ALOS_i)/OR_i \tag{2.1}$$

Where;
X_i = Average no. of daily admissions
OR_i = Average bed occupancy rate
$ALOS_i$ = Average length of stay
However, as De Bruin *et al* [9] mention in the model, only based on average numbers, is not capable of describing the complexity and dynamics of the in-patient flow. Moreover, reported occupancy levels are generally based on the average midnight census (for billing purposes), which results in underestimation of the bed requirements. More recently, queuing models have provided better means of estimating the necessary number of beds based on sound performance measures. The

$M/G/\infty$ queue was used as a model for the casualty ward of a hospital [10]. They showed that in steady state, the bed occupancy rate follows a Poisson distribution with mean $\lambda\mu$, where λ denotes the daily admission rate and µ denotes the average duration of stay. Using this model, the authors determine the required number of beds in order to guarantee that a given target percentage of arrivals receive a bed immediately.

Winston [11], also used the $M/G/\infty$ system to model the queue of patients needing alternative levels of care in acute care facilities whose treatment is completed and who are waiting to be transferred to an Extended Care Facility (ECF). These patients are kept in the hospital due to unavailability of beds in the ECF and reduce the hospital utilization. The authors' model allows managers to predict the effect of certain policy changes on appropriate access measures. For instance, the cost-benefit trade-off of opening an additional extended care facility within a region is compared to that of assigning a higher priority to patients going to ECF from Acute Care Facility than to those coming from other sources.

Instead of using an infinite capacity queue, Weiss and McClain [12], used an M/G/c queue with a state-dependent arrival rate to address the long hospital-wait list problem. He experiments with various management actions such as increasing the number of beds or decreasing mean service times through appropriate means. Gans *et al* [13] developed a queuing model for bed occupancy management and planning of hospitals. Performance measures, such a Mean bed occupancy and the probability of rejecting an arriving patient due to hospital overcrowding, are computed. These quantities enable hospital managers to determine the number of beds needed in order to keep the fraction of delays under a threshold, and also to optimize the average cost per day by balancing the costs of empty beds against those of delayed patients. Although service times, unlike inter-arrival times, do not usually have an exponential distribution, such an assumption is often made in order to simplify the analysis greatly. For instance, De Bruin *et al* [9] used the M/M/c/c queue, referred to as the Erlang Loss model, to investigate the emergency in-patient of cardiac patients in a university medical Centre in order to determine the optimal bed allocation so as to keep the fraction of refused admissions under a target limit. The authors find the relation between the size of a hospital unit, occupancy rate, and TAR, cancellation rate of 5. Another queuing network model applied to a hospital setting is that of Laguna and Marklund [14], who studied a specific obstetrics hospital consisting of 8 sub units with 4 different patient arrival streams. The transfer of patients between the different compartments creates delay in some of the units. As can be seen, the application of queuing models to health care is growing more popular as hospital management teams are gaining awareness of the advantages of these operational research techniques in addressing such issues as determining optimal bed counts and making policy decisions with regards to resource allocation. Research in applying queuing networks with blocking is rarer in the literature due to the mathematical complexities involved in computing performance measures associated with such systems. As a result, hospitals with interacting subunits are often studied through simulations, because they are able to incorporate much more detail than is affordable by analytical methods.

The study also borrows bed optimization theory from Danjuma *et al* [15] who did research on Asset Optimization Problem In A Financial Institution. Their paper looked at how a financial institution could optimally allocate its total wealth among three assets namely; treasury, security and loan in a stochastic interest rate setting. The optimal investment strategy was derived through the application of a stochastic optimization theory for the case of constant relative risk aversion (CRRA) utility function. From the results it was seen that the optimal investment strategy was to shift the financial institution investment away from the risky assets (security and loan) toward the riskless asset (treasury). Also the investment in security and loan was observed to be more risky as the volatility increased. The results further showed that there is increased investment in the risky assets as the investor became less risk averse. In this study, the optimal utilization of beds was key to observations

International Journal of Mathematical Sciences and Optimization: Theory and Applications
Vol. 9, No. 2, pp. 90 - 101
https://doi.org/10.5281/zenodo.10202665

3 Materials and Methods

For the investigations of elementary queuing system, let us denote a system by;

$$A/B/m/K/n/D \quad [16] \tag{3.1}$$

where;
A: Distribution function of the inter-arrival times,
B: Distribution function of service times,
m: Number of servers,
K: Capacity of the system, the maximum number of customers in the system including the one being serviced,
n: Population size, number of sources of customers,
D: Service discipline.

Exponentially distributed random variables are notated by M, meaning Markovian or memoryless. Hence M/M/1 denotes a system with Poisson arrivals, exponentially distributed service times and a single server. M/G/m denotes an m-server system with Poisson arrivals and generally distributed service times. M/M/r/K/n stands for a system where the customers arrive from a finite source with n-elements where they stay for an exponentially distributed time, the service times are exponentially distributed, the service is carried out according to the requests arrival by r servers and the system capacity K.

Further, the queue system is usually described in shorten form by using some characteristics. These characteristics can be represented by Kendall's notation which was initially a three factor notation A/B/C. Later D, E and F were also included in the model to make it A/B/C/D/E/F. The notation of the queuing model is presented in Table 1;

Table 1: Kendall's queuing model Notation

Symbol	Explanation
A	Distribution of arrival time
B	Distribution of service time
C	Number of servers (agents available)
D	Capacity of the system
E	Calling population
F	Queue discipline

Examples of some special notations for various probability distributions describing arrivals and departures include: M - Arrival or departure distribution that is a Poisson process, E - Erlang distribution, G - General distribution, GI - General independent distribution.

For the application of the queuing models to any situation we should describe first the input and the output process. In our ED the input process is the patient's arrival and the output process is considered the patient's discharge in the hospital unit.

Using this M/M/k model, it is known that the system is in a steady state if the following relation is fulfilled:

$$\frac{\lambda}{k\mu} < 1 \tag{3.2}$$

Where; k is the number of servers/channels/wards, λ is the mean arrival rate for the system and μ is the mean service rate for each channel/ward which is also equal to ALOS.

To optimize the process, the probability (P_k) that an entering patient must queue for treatment which means that all beds are busy is obtained. In order to calculate these probabilities, The relations shown in Equation (3.2) was used. In the case of M/M/k model to calculate the probability that no patients is in the ED, the condition that overall sum of probabilities must be 1 is used. Therefore;

International Journal of Mathematical Sciences and Optimization: Theory and Applications
Vol. 9, No. 2, pp. 90 - 101
https://doi.org/10.5281/zenodo.10202665

$$P_0 + P_1 + P_2 + P_3 + \cdots + P_k = 1 \tag{3.3}$$

When the probabilities for k servers is substituted Equation 3.3 is given by:

$$P_0 + P_0\Big(\frac{\lambda}{\mu}\Big) + P_0\Big(\frac{(\frac{\lambda}{\mu})^2}{2!}\Big) + \cdots + P_0\Big(\frac{(\frac{\lambda}{\mu})^{k-1}}{(k-1)!}\Big) + P_0\Big(\frac{(\frac{\lambda}{\mu})^k}{k!}\frac{k\mu}{(k\mu-\lambda)}\Big) = 1 \tag{3.4}$$

$$P_0\Big[1 + \frac{\lambda}{\mu} + \Big(\frac{(\frac{\lambda}{\mu})^2}{2!}\Big) + \cdots + \Big(\frac{(\frac{\lambda}{\mu})^{k-1}}{(k-1)!}\Big) + \Big(\frac{(\frac{\lambda}{\mu})^k}{k!}\frac{k\mu}{(k\mu-\lambda)}\Big)\Big] = 1 \tag{3.5}$$

$$P_0 = \frac{1}{\Big[1 + \frac{\lambda}{\mu} + \Big(\frac{(\frac{\lambda}{\mu})^2}{2!}\Big) + \cdots + \Big(\frac{(\frac{\lambda}{\mu})^{k-1}}{(k-1)!}\Big) + \Big(\frac{(\frac{\lambda}{\mu})^k}{k!}\frac{k\mu}{(k\mu-\lambda)}\Big)\Big]} \tag{3.6}$$

The geometric series is convergent and introducing the sum of the series, we have:

$$P_0 = \Big[\sum_{i=0}^{k-1}\frac{(\frac{\lambda}{\mu})^i}{i!} + \frac{(\frac{\lambda}{\mu})^k}{k!}\frac{k\mu}{(k\mu-\lambda)}\Big]^{-1} \tag{3.7}$$

Therefore, the probability that no patient is in the ED is as stated in Equation (3.7) above.

In M/M/n queue an entering patient must queue for service exactly when n or more patients are already in the system. Using Erlang's C formula, this study obtains;

$$\begin{aligned} P_w &= \frac{\frac{A^N}{N!}\frac{N}{N-A}}{\left(\sum_{i=0}^{N-1}\frac{A^i}{i!}\right) + \frac{A^N}{N!}\frac{N}{N-A}} \\ &= \sum_{k=29}^{\infty} p_k = 1 - \sum_{k=0}^{28} p_k \\ P_w &= p_0\frac{(\frac{\lambda}{\mu})^k}{k!}\frac{k\mu}{(k\mu-\lambda)} \end{aligned} \tag{3.8}$$

Equation (3.8) is the probability that an arriving patient must wait for service, denoted by p_w. The length of the queue in the case of M/M/n is given by:

$$L_q = p_0\frac{(\frac{\lambda}{\mu})^k}{k!}\sum k\Big(\frac{\lambda}{k\mu}\Big)^k \tag{3.9}$$

Calculating the sum of the series the following is obtained

$$L_q = p_0\frac{(\frac{\lambda}{\mu})^k}{k!}\frac{k\mu\lambda}{(k\mu-\lambda)^2} \tag{3.10}$$

Simplifying the equation the solution is given by,

$$L_q = p_0\frac{(\frac{\lambda}{\mu})^k}{(k-1)!}\frac{\mu\lambda}{(k\mu-\lambda)^2} \tag{3.11}$$

Equation (3.11) is the number of patients in the queue or simply the queue length, denoted by L_q. Now the waiting time for service on a queue by use of the Little's Law, the mean waiting time in the queue can be obtained from:

$$W_q = \frac{L_q}{\lambda} \tag{3.12}$$

International Journal of Mathematical Sciences and Optimization: Theory and Applications
Vol. 9, No. 2, pp. 90 - 101
https://doi.org/10.5281/zenodo.10202665

The average treatment/service time per patient, denoted by W_t is

$$W_t = \frac{1}{\mu} \tag{3.13}$$

Thus the total waiting time in system can be obtained from:

$$W = W_q + W_t \tag{3.14}$$

The overall number of patients in the ED is an average

$$L = W\lambda \tag{3.15}$$

As pointed out by De Bruin *et al.* [9], Equation (2.1) is not adequate in describing the dynamics of the in-patient flow.Moreover, reported occupancy levels are generally based on the average midnight census (for billing purposes), which results in underestimation of the bed requirements. More recently, queuing models have provided better means of estimating the necessary number of beds based on sound performance measures. Pike *et al* [17], used the M/G/∞ queue as a model for the casualty ward of a hospital. They showed that in steady state, the bed occupancy rate follows a Poisson distribution with mean $\lambda\mu$, where λ denotes the admission rate (arrival rate) and μ denotes the average duration of stay per ward. Using this model, the authors determined the required number of beds in order to guarantee that a given target percentage of arrivals receive a bed immediately.

Notice that the denominator in Equation (2.1) is a factor of the ALOS. In this case the number of beds required depends on the average daily admissions and ALOS which gives a multiplicative model as presented in Equation (3.16).

$$Y_i = ALOS_i * X_i; \quad i = 1, 2, \cdots k \tag{3.16}$$

Where; Y_i is estimated bed requirements, ALOS is average length of stay, X_i is the average daily admissions.

4 Data analysis

Secondary data obtained from Moi Teaching and Referral hospital was analyzed through model fitting. The current operating characteristics of the system were first studied using Kendall's notation represented by A/B/C/D/E/F. Equations (2.1) and (3.16) were then used to do a comparative analysis of bed requirements for the facility. The results are presented in the following section.

5 Results

Results on operating characteristics are presented in Table 2. The table gives a summary of the operating characteristics for the multi-server ED- IW system

International Journal of Mathematical Sciences and Optimization: Theory and Applications
Vol. 9, No. 2, pp. 90 - 101
https://doi.org/10.5281/zenodo.10202665

Table 2: Summary operating characteristics of the multi-server ED- IW

Wards	$\rho = \frac{\lambda}{k\mu}$	P_0	L_q	L	W_q	W	P_w
Amani	0.0503	0.2445	$1.2535e^{+01}$	$1.4084e^{+01}$	$4.8331e^{-03}$	0.1043	$9.3076e^{-01}$
Umoja	0.0547	0.2164	$1.4961e^{+01}$	$1.5305e^{+01}$	$4.8076e^{-03}$	0.1134	$9.7771e^{-01}$
CCU	0.3919	$1.7170e^{-05}$	$8.0197e^{-01}$	10.9724	$5.9352e^{-03}$	0.0812	$1.0000e^{+00}$
Isolation Ward	3.2093	-					
Faraja	0.0895	0.0816	$4.2743e^{-01}$	$2.5057e^{+01}$	$3.1633e^{-02}$	0.1856	$9.5649e^{-01}$
Riley Mother & Baby	0.1329	0.0242	$1.3274e^{+01}$	$1.7201e^{+01}$	$9.8234e^{-03}$	0.1275	$9.7992e^{-01}$
Neema	0.0489	0.2544	$1.0539e^{+01}$	$1.3687e^{+01}$	$2.1861e^{-03}$	0.1014	$9.5886e^{-01}$
Upendo	0.1067	0.0504	$2.5374e^{+01}$	$2.9885e^{+01}$	$3.3580e^{-03}$	0.2214	$9.8108e^{-01}$
Tumaini	0.0949	0.0701	$1.2473e^{+01}$	$1.6576e^{+01}$	$1.5151e^{-03}$	0.1228	$8.7786e^{-01}$
Subira	0.0902	0.0801	$2.2329e^{+01}$	$2.5248e^{+01}$	$3.8727e^{-03}$	0.1870	$7.9149e^{-01}$
ICU	0.2415	0.0012	$2.7731e^{+00}$	$2.7611e^{+01}$	$2.0523e^{-02}$	0.0500	$9.0520e^{-01}$
HDU	32.5847						
Ada	0.2555	0.0008	$9.9838e^{-10}$	7.1535	$7.3887e^{-12}$	0.0529	$1.5740e^{-06}$
Pw II Peads	0.4735	$1.7482e^{-06}$	0.0003	13.2572	$1.9432e^{-06}$	0.0981	$9.0833e^{-01}$
Pw II Nbu	9.7966						

In Table 2 the probability that there is no patient in the ED or no queues is represented by P_0. Clearly, the probabilities are too low showing presence of patients waiting for service at the department. The waiting probabilities at the wards are represented by P_w, which are high, ranging between 79% to 100% except ADA ward which had less than 1%. This is again an indicator that there is waiting at the wards as well. Queue lengths were represented by L_q for instance; Amani = 12.535, Umoja = 14.961, CCU = 0.8020. The total waiting time is represented by (W) in days. Clearly, this time affects the length of the queue. When patients take long to wait for services, there are high chances of a corresponding long queue in that ward. The overall number of patients in the system from the analysis is represented by L where the wards with higher values of L could result in long waiting time for services.

On the other hand, the analysis of bed requirements in the facility was done using Equation (2.1) and (3.16). The analysis was done because the study found out that the probability of waiting for services in the wards except ADA was ranging between 79% to 100%, which is quite high (See Table 2). Apart from bed requirements, the study also established that for the problem to be resolved in the facility, each nurse/doctor should have the same workload by taking care equal number of patients at any given time. As the number of nurses and doctors is usually proportional to standard capacity, this criterion is equivalent to keeping bed occupancy rates equal among the wards. However, if one maintains occupancy levels equal then, by Little's law, wards with shorter ALOS will have a higher turnover rate (discharges) and thus admit more patients per bed. But the load on the ward staff is not spread uniformly over a patient's stay, as treatment during the first days of hospitalization requires much more time and effort from the staff than in following days [18]. In addition, patients admissions and discharges consume more doctors' and nurses' time and effort. Therefore, even if the occupancy among wards is kept equal, the ward admitting more patients per bed ends up having a larger load on its staff. This means that natural alternative fairness criterion is balancing incoming load, or flux - namely, the number of admitted patients per bed per certain time unit among the wards. The results of bed requirements at Moi Teaching and Referral Hospital are presented in Table 3.

It is clear from Table 3 that Equation (3.16) gives a better estimate of bed requirements than Equation (2.1) which grossly overestimates the same. Therefore, the model for bed requirement as presented by Equation (3.16) is fair, efficient and cost effective as it results into equitable reassignment and/or optimal allocation of beds among the wards and units as opposed to acquisition of more beds, which in this case would not be necessary. These results are also graphically depicted in Figures 1 and 2.

International Journal of Mathematical Sciences and Optimization: Theory and Applications
Vol. 9, No. 2, pp. 90 - 101
https://doi.org/10.5281/zenodo.10202665

Table 3: Bed Requirement Estimates

Wards	Current Bed Allocation	Average Daily Admissions	Average Daily Occupancy	ALOS	Bed Requirement Estimates (De Bruins Equation)	Bed Requirement Estimates (the study Equation)
Amani	92	8.4444	2.8724	11.000	32.3383	92.8884
Umoja	96	8.9667	2.7947	9.6667	31.0154	86.6784
CCU	14	1.4667	0.0576	8.3333	212.1953	12.2225
Isolation Ward	5	0.0222	0.0061	67.6667	246.2624	1.5022
Faraja	40	7.8778	0.7022	6.6667	74.7920	52.5189
Riley Mother & Baby	125	38.7111	1.1944	1.0000	32.4105	38.7111
Neema	60	8.5222	1.9333	11.3333	49.9584	96.5846
Upendo	32	4.7778	0.4836	9.3333	92.1336	44.5926
Tumaini	35	4.5889	0.5756	10.6667	85.0390	48.9484
Subira	35	4.6111	0.5989	11.3333	87.2583	52.2590
ICU	21	3.1889	0.1330	6.3333	151.8516	20.1963
HDU	3	0.1111	0.0003	1.3333	493.7654	0.1481
Ada	16	0.1333	0.0818	100.000	162.9584	13.3300
Pw II Peads	6	1.9000	0.0200	5.3333	506.6635	10.1333
Pw II Nbu	3	0.0667	0.0003	6.3333	1408.1037	0.4224

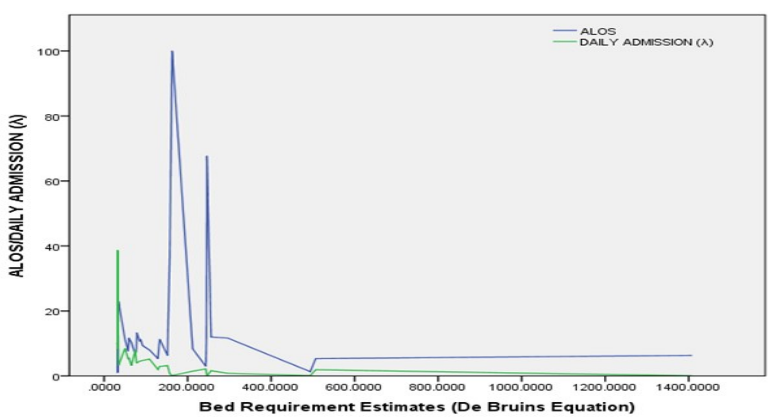

Figure 1: Impact of ALOS and Daily Admissions on De Bruins Bed Requirement Equation

Results in Figure 1 show that bed requirements increase with decrease in admissions and average length of stay (ALOS). Notice that these results were based on Equation (2.1). The impression given here is that of negative correlation between admissions and bed requirement. This does not happen in real life situations as the bed requirement is always directly proportional to the number of admissions. From the figure, when daily admissions and ALOS increase, the number of

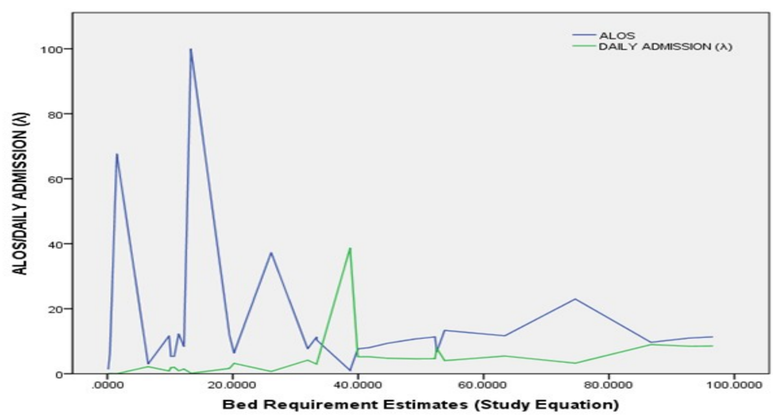

Figure 2: Impact of average length of stay (ALOS) and daily admissions on the Study Bed Requirement Equation

beds required also increase. This scenario is based on the Study's Improved Model as presented by Equation (3.16). This scenario presents a positive correlation between admissions and bed requirement, which is the expected case in reality.

International Journal of Mathematical Sciences and Optimization: Theory and Applications
Vol. 9, No. 2, pp. 90 - 101
https://doi.org/10.5281/zenodo.10202665

6 Conclusions

The probabilities of having no patient at the ED was too low. Additionally the probabilities of waiting in each ward were high indicating that patients had higher chances of waiting for services. Further more, queue lengths and waiting times were also long indicating that there was congestion in the system, hence reducing its efficiency. Interestingly, the model estimating bed requirements gave a number approximately equal to the existing number of beds. This implies that the available beds are sufficient, except that they needed more appropriate assignment to meet the demands of the facility. This is in agreement with Green [3] who opined that most organizations do not require extra resources but proper allocation of existing resources would lead to optimal resource utilization and delivery of services. Additionally, Bhavani and Jayalalitha [19] in their study advised that seasons needed to be put in to consideration when assigning resources for optimal service because some seasons were expected naturally to have a higher number of customers than others. This would apply to the current study, particularly during disease outbreaks.

7 Recommendations

The study focused on inpatient analysis. It is recommended that similar studies could be extended to other departments such as the outpatient department to shade light on the dynamics around such departments.

References

[1] Litvak, E., Long, M. C., Cooper, A. B., & McManus, M. L. (2001). Emergency department diversion: causes and solutions. *Academic emergency medicine: official journal of the Society for Academic Emergency Medicine*, **8**(11), 1108-1110.

[2] Ng'ang'a, J. (2019). Kenya's Big Four Agenda Gets Global Recognition. *Kenya News Agency*.

[3] Green, L. (2004). Capacity Planning and Management in Hospitals: Operations Research and Health Care. *International Series in Operations Research & Management Science,* 70, 15-41.

[4] Armony, M., Israelit, S., Mandelbaum, A., Marmor, Y. N., Tseytlin, Y., & Yom-Tov, G. B. (2015). On patient flow in hospitals: A data-based queueing science perspective. *Stochastic systems*, **5**(1), 146-194.

[5] Yaduvanshi, D., Sharma, A., & More, P. (2019). Application of queuing theory to optimize waiting-time in hospital operations. Operations and Supply Chain Management: *An International Journal*, **12**(3), 165-174.

[6] Ponkshe, S., Bhasin, S., Hooghan, S., Vazifdar, S., & Khetan, S. QUEUING THEORY IN THE HEALTHCARE SECTOR.

[7] Kalwar, M. A., Marri, H. B., Khan, M. A., & Khaskheli, S. A. (2021). Applications of queuing theory and discrete event simulation in health care units of Pakistan. *International Journal of Science and Engineering Investigations*, **10**(109), 6-18.

[8] Dong, J., & Perry, O. (2020). Queueing models for patient-flow dynamics in inpatient wards. *Operations Research*, **68**(1), 250-275.

[9] De Bruin, A. M., Van Rossum, A. C., Visser, M. C., & Koole, G. M. (2007). Modeling the emergency cardiac in-patient flow: an application of queuing theory. *Health Care Management Science*, **10**(2), 125-137.

[10] Koizumi, N., Kuno, E., & Smith, T. E. (2005). Modeling patient flows using a queuing network with blocking. *Health care management science,* **8**(1), 49-60.

[11] Winston, W. L. (1991). *Operations research: applications and algorithms.* Brooks/Cole.

[12] Weiss, E. N., & McClain, J. O. (1987). Administrative days in acute care facilities: a queueing-analytic approach. *Operations Research*, **35**(1), 35-44.

[13] Gans, N., Koole, G., & Mandelbaum, A. (2003). Telephone call centers: Tutorial, review, and research prospects. *Manufacturing & Service Operations Management*, **5**(2), 79-141.

[14] Laguna, M., & Marklund, J. (2018). *Business process modeling, simulation and design.* Chapman and Hall/CRC.

[15] Danjuma, T., Onah, E. S., & Aboiyar, T. (2019). Asset Optimization problem in a financial Institution. *International Journal of Mathematical Sciences and Optimization: Theory and Applications*, **2019**(2), 581-591.

[16] Kendall, D. G. (1951). Some problems in the theory of queues. *Journal of the Royal Statistical Society: Series B (Methodological)*, **13**(2), 151-173.

[17] Pike, M. C., Proctor, D. M., & Wyllie, J. M. (1963). Analysis of admissions to a casualty ward. *British journal of preventive & social medicine*, **17**(4), 172.

[18] Schlechter, K. (2015). Hershey Medical Center to open redesigned emergency room. *The Patriot-News.*

[19] Bhavani, L. K and Jayalalitha, G. (2021). Applying Queuing Theory to Enhance the Service Provided by a Restaurant. *Annals of R.S.C.B., ISNN: 1583-6258, Vol. 25, Issue 6, 2021, Pages 4479-4484*