Computers & Industrial Engineering 78 (2014) 299–312

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

# Patient flow modelling and performance analysis of healthcare delivery processes in hospitals: A review and reflections

Papiya Bhattacharjee *, Pradip Kumar Ray

*Department of Industrial and Systems Engineering, Indian Institute of Technology Kharagpur, Kharagpur 721302, India*

ARTICLE INFO

*Article history:*
Available online 9 May 2014

*Keywords:*
Patient flow
Performance
Queueing theory
Markov models
Simulation
Empirical modelling

ABSTRACT

Analysis of hospital processes is essential for development of improved methods, policies and decision tools for overall performance improvement of the hospital system. Amidst the current scenario of continuously increasing healthcare costs and scarcity of resources, optimal utilization of resources without hampering the quality of care has gained importance in any country. Modelling, analysis and management of patient flows, in this context, plays a key role in performance analysis and improvement of hospital processes as appropriate modelling of patient flows may help healthcare managers make decisions related to capacity planning, resource allocation and scheduling, appointment scheduling and for making necessary changes in the process of care. The concept of patient flow and its modelling has gained much attention in healthcare management literature over past few decades. In this paper, the existing approaches pertaining to modelling of patient flows in hospital systems have been classified and critically appraised focussing on the recent advancements in order to identify future research avenues. A generic framework for patient flow modelling and performance analysis of hospital systems that may serve as a guide for the practitioners dealing with similar kinds of problems to improve healthcare delivery has also been provided.



## 1. Introduction

Of late, the study and analysis of any healthcare system has become a necessity for its improved performance over time for which it has to meet a number of, often conflicting, objectives and goals, such as minimizing the cost of healthcare, maximizing the utilization of physical and human resources, improving the quality of care by providing efficient diagnostic systems, handling increasing number of patients effectively within a limited time span, arranging varieties of healthcare facilities in a single location, and improving overall healthcare system performance within limited and predetermined budget and time. Increasing cost of operations and maintenance, and healthcare cost to the patients due to the use of newer technologies, resources and methods have added newer dimensions as constraints to the problems of healthcare system (Brailsford & Vissers, 2011; Brandeau, Sainfort, & Pierskalla, 2004; Rais & Viana, 2010). The objectives of a healthcare system as mentioned above may be achieved effectively with establishment of appropriate healthcare planning and organization along with application of Operations Research – based approaches and other quantitative techniques (Brailsford & Vissers, 2011; Rais & Viana, 2010).

Out of the many subsystems of a healthcare system, hospitals are highly integrated service units attending to the needs of the patients under treatment. A typical hospital system consists of a number of interacting departments/sub-units (Outpatient Department (OPD), department of emergency, inpatient wards, Operating Theatre (OT), Intensive Care Unit (ICU)/Intensive Therapeutic Unit (ITU), diagnostic services, such as Pathology and Radiology, etc.) within a geographical area embedded in an organization. Each department is responsible for a single or at the most, a few related functions within the hospital. Such an integrated system dealing with different aspects of healthcare and its problems may have several possible combinations of service types (hence, varieties of care pathways), patient status, types of facilities as well as physician profiles/specializations available for treatment under varieties of constraints (primarily determined by availability of resources). It is imperative that hospital operational performance is required to be modelled for varieties of care pathways in respect of various system conditions and constraints. The overall performance of a hospital system is dependent on the performance of all the departments/sub-units which have different operational issues, and over

* Corresponding author. Tel.: +91 9831400985.
*E-mail addresses:* papiya@iem.iitkgp.ernet.in (P. Bhattacharjee), pkr@vgsom.iitkgp.ernet.in (P.K. Ray).

http://dx.doi.org/10.1016/j.cie.2014.04.016
0360-8352/

the years various performance measures, such as waiting time of patient at an emergency department, length of stay of an inpatient, utilization of operating rooms, bed utilization in ICU, etc. to mention a few, have been used to assess the performance of these operations (Cardoen, Demeulemeester, & Belien, 2010; Cayirli & Veral, 2003; Kim, Horwitz, Young, & Buckley, 1999; McClean & Millard, 1998).

An important factor that affects and is a determinant of the performance of healthcare delivery processes in a hospital system is the flow of patients through the system. 'Patient Flow' refers to the 'movement of patients through the whole process of care'. Hall, Belson, Murali, and Dessouky (2006) define patient flow analysis as 'the study of how patients move through the health-care system'. The rate of patient flow entering a system is affected by both seasonal and local factors as well as by the location of hospitals, and the kind of services offered (Alexopoulos, Goldsman, Fontanesi, Kopald, & Wilson, 2008). On the other hand, flow routings through the system are partly dependent on the process of care and decisions taken by physicians, and partly on the inherent uncertainties of healthcare processes (Cote, 2000; Harper, 2002). In this context, there may be two different perspectives for describing the movement of patients through the care process: 'clinical' and 'operational' (Cote, 2000). While patient flow refers to the progression of a patient's health status from a clinical perspective, it represents the movement of patients through various locations (or processes) of a healthcare system from an operational perspective. The flow of a patient through a hospital commences when the patient arrives at the hospital either through the outpatient department (OPD) or through the emergency department (ED) or sometimes admitted to inpatient wards when referred from some other healthcare system. Subsequently, the patient may travel through various routes depending upon the medical condition of the patient, decisions taken by the physicians, the effects of treatment and various constraints of the system. Patient flows are considered to be very complex because of the different pathways patients may take and the inherent uncertainty and variability of healthcare processes (Cote, 2000; Harper, 2002). Moreover, due to limited resource, queues may be formed for various services, and thus the patient flow may be represented by queueing networks (Creemers & Lambrecht, 2011; Hall et al., 2006). Performance of these queueing networks may be analyzed either through analytical queueing formulations, Markov chain analysis or by computer simulation of these networks. A careful and structured analysis of patient flows may reduce the congestions and significantly improve the performance of a hospital system (Hall, 2006; Marshall, Vasilakis, & El-Darzi, 2005; Zhao & Lie, 2008). It is a challenging task for a researcher to capture the inherent random nature and complexities of healthcare elements in patient flow. At various locations of the flow network, there is scope for operational improvements by reducing the delays experienced by patients at the junctions of the phases and also by reducing the idle times of the servers at each phase through effective capacity planning, resource allocation and scheduling and also through appointment scheduling. Another alternative for improving the performance of a healthcare delivery process is through making changes in the process of care. Such changes may include eliminating unnecessary or non-value-adding activities, eliminating duplicate activities, performing some of the activities in parallel, and identifying alternative process flows (McLaughlin & Hays, 2009). Modelling the patient flows may give important insights for such improvements.

Earlier review articles in recent times related to patient flow and performance modelling of healthcare systems in general and hospital systems in particular focus mainly on a single modelling technique, most of which concentrating on simulation technique (e.g., Brailsford, Harper, Patel, & Pitt, 2009; Eldabi, Paul, & Young, 2007; Fone et al., 2003; Green, 2006; Gunal & Pidd, 2010; Jun, Jacobson, & Swisher, 1999; Katsaliaki & Mustafee, 2011; Zonderland & Boucherie, 2012). In this paper, an attempt is made to review the frequently used modelling approaches specific to patient flow and performance analysis of healthcare processes in hospitals with a focus on the recent advancements and to bring out the reflections on the evolution of these techniques and to identify the appropriate modelling and analysis techniques according to the characteristics of patient flow and the type of problem to be solved. Critical appraisal of these techniques identifies key issues to be addressed while carrying out patient flow modelling and performance analysis. A generic framework for carrying out such an analysis is also presented. The remainder of the paper is organized as follows. Section 2 explains components of a patient flow network and its characteristics; Section 3 provides the aspects to be modelled and a review of the recent literature on patient flow modelling; Section 4 presents the performance measures used in the literature to assess the operational performance of various services of a hospital; Section 5 presents a generic framework for performance assessment through patient flow modelling. Concluding remarks are drawn in Section 6.

## 2. Components of a patient flow network and its characteristics

In the context of patient flow modelling, the aspects related to the building blocks of the patient flow network, the multiphase nature of the flows and the inherent uncertainties and complexities of patient flows need to be discussed before getting into the details of its modelling techniques.

### 2.1. Elements/building blocks of a patient flow network

As has been stated in the previous section, the flow of patients through a hospital system may be represented as queueing networks, and like any queueing network, elements or building blocks of patient flow networks include the network structure (which consists of an entrance point, an exit point, various single-server or multi-server nodes between these two terminal points, and the paths connecting the nodes), patients (the entities which are flowing through the network), and the resources (both physical and human) which act as servers at the nodes. Depending on whether the patient flow is being considered from a clinical or an operational perspective, the nodes represent either a health state (condition) of a patient (e.g., Xie, Chaussalet, & Millard, 2005) or a location in the hospital or a department (e.g., Cote & Stein, 2007).

### 2.2. Stages in patient flow

Any kind of patient flow, clinical or operational, through any type of process of care in a hospital, whether it is outpatient care, emergency care, surgical care or inpatient care, may, in general be considered as a multistage/multiphase process (Hulshof, Kortbeek, Boucherie, Hans, & Bakker, 2012). If the flow is being considered from an operational perspective, the stages represent various processes, such as registration, consultation and examinations (e.g., blood tests and X-ray). On the other hand, if the flow is being considered from a clinical perspective, the stages represent different stages of patient's health status. For instance, a patient admitted in a geriatric department may have health states like acute condition, condition requiring rehabilitative care and that requiring long stay (McClean & Millard, 1998). A particular patient may not need to undergo all the defined stages in a care process or a care pathway. Typical activities at each phase include arrival, wait, treatment/service, and leave/exit. The simplest of patient flows may have very few stages or even a single stage. For instance, a patient visiting a walk-in clinic for a follow-up may just consult the physician and leave the system. There may not exist any precedence relationship between some of the stages/phases, i.e., some of the phases (activities) may take place in any sequence or in parallel (Jiang & Giachetti, 2008). The problems in varieties of care pathways as prevalent in a hospital environment are required to be addressed separately in order to assess the performance and find improvement alternatives specific to each type of pathway.

### *2.3. Inherent uncertainties and complexities of patient flows*

Patient flows through any subsystem of a hospital have inherent uncertainties and complexities (Cote, 2000; Hall et al., 2006; Harper, 2002; Marshall et al., 2005). The uncertainties arise mainly due to the randomness in the inter-arrival times of unscheduled as well as scheduled patients, no-shows among the scheduled patients, randomness in service times at various stages of care, uncertainties in the transitions of a patient's health status and transitions among various care locations.

The complexities in patient flows are mainly due to the following characteristics:

(i) There are various possible care pathways for almost any subsystem of a hospital depending upon the inter-related factors, such as severity of illness, decisions taken by physicians, progression of patient's health status and various rules for routing and resource constraints.
(ii) There may be very large number of stages (some of them may be repeated) in some instances where the patient requires a combination of large number of services.
(iii) There may be various priority rules for patients to be seen by a physician or to be investigated at a radiology department. Also there are various rules governing admission and reservation policies while allocating beds to different types of patients.

In addition, there is variability in characteristics, such as arrival rates, inter-arrival times, lengths of stay and service times for various patient groups.

## 3. Modelling patient flows

Modelling of patient flows plays a key role in improvement of healthcare delivery by providing insights for process improvements, capacity planning, resource allocation and scheduling, and appointment scheduling. Patient flow modelling helps a hospital system carry out different kinds of analyses, such as (i) finding the most influential factors affecting the performance of the system in a given situation, (ii) evaluating the current appointment scheduling methods and various other alternatives to find out improved sequencing and scheduling rules, (iii) making resource allocation decisions for improvement in resource utilization through performance analysis of alternative resource scheduling and allocation policies, (iv) comparing alternative routings or pathways for a process so that better routing schemes (if there exist any) may be determined or from a broader perspective, evaluating the changes in the process of care so that better alternatives may be identified for improving the performance of the system, (v) estimating the cost of a treatment in terms of the lengths of stay of patients, (vi) exploring the interrelationships between parameters (performance) at various stages, and (vii) classification of patients into various groups so that different planning strategies may be applied to different groups (De Bruin, Van Rossum, Visser, & Koole, 2007; El-Darzi, Vasilakis, Chaussalet, & Millard, 1998; Garg, McClean, Meenan, El-Darzi, & Millard, 2009; Garg, McClean, Meenan, & Millard, 2010; Gillespie et al., 2011; Harrison & Escobar, 2010; Jiang & Giachetti, 2008; Matta & Patterson, 2007; McClean, Faddy, & Millard, 2005; McClean, McAlea, & Millard, 1998; McClean & Millard, 1998; Vasilakis & El-Darzi, 2001). It is imperative that hospital performance should be modelled and analyzed in relation to patient flow; a number of models representing the flows of patients in different hospital units under different conditions and constraints have been proposed by several researchers and practitioners over the years. Given the inherent nature of patient flow with characteristics such as stochastic arrivals, stochastic service times, and uncertainties in patient routings, the patient flow models have to be stochastic in nature (Harper, 2002).

### *3.1. Aspects to be modelled*

A number of aspects, such as the patient flow network structure and its configuration, patient arrival distributions, branching or transition probabilities, and service time distributions need to be considered in a systematic manner for developing a comprehensive approach for patient flow modelling of a given hospital system.

#### *3.1.1. Patient flow network structure and its configuration*

The network structure, the number of nodes and the number of servers at each of these nodes decide the basic modelling structure. Suppose a queueing network model is to be developed (applied), the network structure and configuration help in determining whether the model will be a single server model or a multi-server model and whether it will be an open network model or a closed network model (Gross & Harris, 2011). Similarly in a computer simulation model, the configuration is the primary feature which is incorporated in the model (Gunal, 2012; Lowery, 1996). While developing Markov process models (Ross, 2009), the structure helps in identifying the states and various possible transitions between the states.

#### *3.1.2. Arrival distribution*

The unscheduled or random arrivals of patients to an outpatient department and to other departments, such as emergency and inpatient department have mostly been modelled in the literature as homogenous and non-homogenous Poisson arrivals (e.g., Cote, 1999; Gorunescu, McClean, & Millard, 2002a; Gorunescu, McClean, & Millard, 2002b; Harrison, Shafer, & Mackay, 2005; Jiang & Giachetti, 2008; Yeon, Lee, & Jang, 2010). In case of scheduled arrivals, the probability distribution of the amount of time by which patients are early or late (i.e., patient tardiness) is required to be determined. Alexopoulos et al. (2008) mention that the scheduled arrivals do not satisfy the memoryless assumption, and point out the possibility of violation of the assumptions of Poisson process for unscheduled arrivals. They propose improved methods for modelling the arrivals of patients at a community clinic as non-homogenous Poisson process, and model the tardiness of patients as Johnson's distribution.

#### *3.1.3. Branching or transition probabilities*

During their movement through the system, the patients follow various routes which evolve during the process of care. Different classes of patients follow different routes where some of the paths may be common and the others different. At different nodes (or states) there may be various possible transitions to other states, and different classes of patients have different probabilities of transition between the states. These transition probabilities may be determined by analyzing previous data on patient routings (Cochran & Roche, 2009; Cote & Stein, 2007).

*3.1.4. Service time distribution*

The service time durations, such as the time taken for registration at the reception, duration of consultation with the physician, time taken for conducting lab tests, and duration of surgeries are random variables and their probability distributions need to be modelled. The assumption of exponential service times is common for both outpatient and inpatient services mainly because of the convenience for obtaining analytically tractable results using queueing theoretic models (De Bruin et al., 2007). Other distributions which are used to model service times of activities in a healthcare environment include triangular, gamma, lognormal, Weibull distributions, and a combination of two or three exponential terms (Ahmed & Alkhamis, 2009; Centeno, Dodd, Aranda, & Sanchez, 2010; Harrison, 2001; Rohleder, Lewkonia, Bischak, Duffy, & Hendijani, 2011; Spangler, Strum, Vargas, & May, 2004).

## 3.2. Specific modelling approaches

In this subsection, we present a review of the research papers appearing in the literature mainly over the past decade in respect of modelling and analysis of patient flows and performance of healthcare processes in hospitals. The modelling approaches may be broadly classified as analytical, simulation and statistical or empirical.

### 3.2.1. Analytical approaches

The analytical approaches used to capture patient flows are (i) Queueing theoretic models and (ii) Markov chains and compartmental models.

*3.2.1.1. Queueing theoretic models.* Queueing theory is being used since 1950's by many researchers and healthcare managers for representing and analyzing hospital systems (Baily, 1952; Brahimi & Worthington, 1991; Cochran & Roche, 2009; Green, 2006; Worthington, 2009). A number of analytical queueing theoretic models have been proposed over the years for analyzing and improving the performance of hospitals and their subsystems. The performance metrics, usually computed through analytical formulae developed in queueing theory include waiting time-related measures, congestion measures, measures of idle time, and utilization of the server(s). Closed form analytical results are usually restricted to steady-state Markovian queueing models (Gross & Harris, 2011). Once these assumptions are relaxed, the approximation approaches, numerical techniques or simulation has to be used to analyse such models. While a number of models with Markovian assumptions (Poisson arrival and exponential service time distributions), such as M/M/1, M/M/c and M/M/c/c (De Bruin et al., 2007; Green & Nguyen, 2001) are developed, other queueing models, such as M/PH/c/c, M/PH/c/N, G/G/1, and G/G/m are also proposed under non-Markovian situations (Creemers & Lambrecht, 2011; Gorunescu et al., 2002a; Gorunescu et al., 2002b; Jiang & Giachetti, 2008). The analytical formulae in queueing theory may be quite helpful in determining the relationships between the influential parameters and system outcomes for a typical patient flow. For instance, consider the widely used Erlang's loss formula (Gross & Harris, 2011) which holds for M/G/c/c model,

$$p_c = \frac{r^c/c!}{\sum_{i=0}^{c} r^i/i!} \quad (r = \lambda/\mu) \tag{1}$$

where, $\lambda$ and $\mu$ are the arrival and service rates respectively, c is the number of servers, $p_c$ is the probability that an arriving customer finds all the servers busy. Gorunescu et al. (2002a) represent the flow of patients through a hospital geriatric department using an M/PH/c/c queue, the service distribution is phase type (PH), and c denotes the total number of beds. An arriving patient, who finds all the beds occupied, is lost to the system which may not be the case in reality as these patients are usually admitted to other specialities or they come later when beds become available. Gorunescu et al. (2002a) use Eq. (1) to find the minimum number of beds required to keep the delay probability below a pre-specified value. As an extension of this work, Gorunescu et al. (2002b) represent a geriatric department using an M/PH/c/N queueing model, N being the maximum capacity of the system (which means that there is a waiting room for N-c patients), and analyse the effects of introducing extra, unstaffed back-up beds which can be used at the times of crises. Using the model, the effects of changing the arrival rate, the mean length of stay, and the bed allocation on the rejection probability, emptiness, and costs are illustrated. De Bruin et al. (2007) use Erlang Loss model (M/M/c/c) to represent emergency cardiac inpatient flows and investigate the bottlenecks in the emergency care process wherein the optimal allocation of beds, given a maximum number of refused admissions is determined, and the relationship between variation in arrivals and length of stay and occupancy rates is explored.

As has been found by several researchers, analytical queueing models are quite efficient in representing patient flows and analysing performance of the hospital systems in case such flows are of moderate complexity (Green, 2006). As the level of complexity increases, approach of simulation modelling is considered suitable mainly in cases where an integrated network of facilities need to be modelled. Moreover, in many healthcare environments (e.g., outpatient systems), the queueing system may not usually reach a steady state and hence steady state queueing results cannot be used in these situations (Brahimi & Worthington, 1991).

Apart from single-node queueing models, a number of applications of network queueing models may also be found in the literature (e.g., Cochran & Roche, 2009; Creemers & Lambrecht, 2011; Zonderland, Boer, Boucherie, de Roode, & van Kleef, 2009). As the healthcare processes can be represented by a number of interdependent work-stages, network queueing models may represent the stages in a network of departments or a network of facilities within any department of a hospital, and these models may help in analyzing the performance of any unit (facility) in relationship with the other units (facilities) of a hospital system.

Over the years, analytical queueing models have been used for a wide range of applications including appointment scheduling in an outpatient clinic or a radiology department, bed planning in inpatient wards or ICU, and resource allocation in an emergency department. The review of literature reveals that in most of the earlier articles, the analytical queueing models were used mainly to represent single-stage processes, such as an inpatient ward, where the ward is considered as a multi-server node or a physician's clinic where the physician(s) is considered as the server(s). Although there are fewer applications of analytical queueing network models in healthcare systems as compared to single node queueing models reported in the literature, the number of such applications has gradually increased in recent times (Zonderland & Boucherie, 2012). Also most of the queueing theoretic models have been applied for steady state analysis of a system as transient analysis using theoretical queueing models is quite cumbersome. Many hospital systems, such as outpatient appointment systems usually do not reach a steady state as they operate for a finite period of time in a day. A number of newer articles have considered the transient nature of appointment systems in analytical queueing models (e.g., Hassin & Mendel, 2008; Luo, Kulkarni, & Ziya, 2012).

Usually five dimensions, viz. (i) single server versus multi-server versus infinite server, (ii) exponential versus non-exponential (or Markovian versus Non-Markovian), (iii) steady-state versus time-dependent, (iv) single-node queues versus network of queues, and (v) extra features, such as priority systems, server vacations, batch arrivals, balking and reneging of customers are used to describe queues and queue modelling (Worthington, 2009). We use these dimensions in Table 1 (we do not mention the number of servers as it is not explicitly mentioned in many of the papers) to appraise a number of research papers pertaining to queueing models of patient flows which have appeared in the literature over the past decade.

*3.2.1.2. Markov chains and compartmental models.* Markov chains are stochastic processes with finite or countable state space in which the conditional distribution of any future state, given the past states and the present state, is independent of the past states and depends only on the present state (Markovian property, see Ross, 2009). In a semi-Markov process, the time between successive transitions follows a particular probability distribution. Patient flows (clinical and operational) in a hospital involve a number of states which evolve during the process of care, and may be represented by Markov chains. Such models may help in capacity planning and resource allocation decisions or estimating the cost for treatment by determining the number of patients in various phases at a given time in the future and performing the analyses of lengths of stay in inpatient environments or the time spent by a patient in various states (processes) during an outpatient visit.

Although Markov processes are associated with queueing theory as they are used as tools to analyse many queueing models, patient flows have not been represented as extensively using Markov models as by queueing theoretic models. Markov chain models have been widely used in the literature for modelling lengths of stay of patients for operations and planning of inpatient facilities (mostly geriatric departments) (e.g., Christodoulou & Taylor, 2001; Faddy & McClean, 2005; Garg et al., 2010; Gorunescu, Gorunescu, & Prodan, 2002c; Shaw & Marshall, 2007). In some situations, the flow of patients can be described as a Markov process having ordered states which starts from the first state and there is a sequential movement (with no backward flow allowed) through the states. There may be transition to absorption state from any state and each such transition has some probability associated with it. In such Markov processes, the time to absorption is described by a phase-type distribution of which Coxian and exponential distributions are special types. McClean and Millard (1998) use a Markov model for compartmental modelling of patient flows within a geriatric hospital where the three compartments denote the states of acute care, rehabilitative care, and long-stay care. By assigning costs to each of the three states (daily cost per patient in a state), the expected costs for treating the cohorts of patients are determined which may help healthcare planners by providing information for planning their budgets. By changing the values of the branching probabilities, the cost effectiveness of various strategies may be evaluated. Xie et al. (2005), Xie, Chaussalet, and Millard (2006) develop a continuous time Markov model for describing patient flow and for analyzing the lengths of stay of elderly people moving within and between residential home care (RC) and nursing home care (NC). The stay in both the care homes is modelled as a two-phase process (short stay and long stay). The actual state spaces of the Markov model are aggregated to form three compartments or classes (RC, NC, and 'discharge from these care homes'). A Markov model using Coxian phase-type distributions for modelling durations of stay of elderly patients is developed by Marshall and McClean (2004) for resource (bed) allocation decisions in the wards with the potential of the fitted distributions for grouping of patients according to duration of their stay assessed using relevant data like age, gender, admission reasons, destination on departure from hospital, and duration of stay in hospital from a department of geriatric medicine. By grouping the patients according to their lengths of stay, the details of patients with corresponding lengths of stay are examined to identify the common characteristics. This would help in management of care through better allocation of beds and other resources in wards for the care of elderly patients as it would be possible to estimate the duration of stay of a patient on admission by analysing their characteristics. Similar sort of studies are reported in the literature (e.g., McClean et al., 2005) wherein the factors influencing the lengths of stay are considered and the patients are clustered on the basis of lengths of stay or the number of phases involved (i.e., different patient pathways). Such clustering of patients may help in hospital management and improving efficiency by applying specific improvement strategies for a cluster. Marshall and Zenga (2009) discuss the problems and examine various approaches and recent developments associated with estimation of parameters while modelling healthcare systems by fitting Coxian phase-type distributions for modelling lengths of stay or patient waiting times. Garg et al. (2010) develop a non-homogeneous, discrete-time Markov model for admission scheduling and resource planning in an integrated care system (care is provided both in hospital and the community), and illustrate the model by an application to an elderly care system using data from historical records of the geriatric department.

Cote and Stein (2007) present a semi-Markov model to describe an individual patient's experience during a visit to 'doctor's office' (an outpatient clinic). The model is developed by using empirical data collected by tracking the visits of patients at a local family practice clinic. Here the performance measure of interest is sojourn time in each state i.e., the service time and waiting time combined, and the model provides first passage time (duration of a patient-care visit) predictions. This may help in operational decisions for the clinic, but the shortcoming of such model is that the performance measures, such as waiting time in the queue and queue length cannot be analyzed. Similarly, a Markov chain model has been presented by Wang, Li, and Howard (2011) to study the care delivery process within the patient rooms in an outpatient environment. A simple example depicting the application of a Markov chain for describing an outpatient system is given in Fig. 1. $p_{ij}$ ($i, j$ = 1,2,3,4,5) is the probability of transition from state i to state j, where state 5 is the absorbing state.

It is found that the earlier applications of Markov and semi-Markov chain models in hospital patient flows were mainly associated with modelling the inpatient clinical flows. Fackrell (2009), while presenting a brief review of the applications of phase-type distributions in healthcare, concludes that phase-type distributions have not been explored much in healthcare systems modelling, the applications being limited to only one class of patients (geriatric patients) and use of only special types of phase-type distributions (rather than general phase-type distributions). Fackrell (2009) also suggests the use of phase-type distributions for constructing integrated models for healthcare systems. However, many newer articles appear in the literature that illustrate the application of Markov and semi-Markov chain models in other environments and applications, such as capacity planning in an outpatient environment.

*3.2.2. Simulation modelling*

Different computer simulation modelling techniques, such as Discrete-Event Simulation, System Dynamics, and Agent-Based Simulation have been applied over the years to various healthcare problems including patient flow modelling and operational performance analysis (Eldabi et al., 2007; Fone et al., 2003; Gunal & Pidd, 2010; Jun et al., 1999; Katsaliaki & Mustafee, 2011). While system dynamics models essentially used for long-term forecasting or prediction, are more suitable when the system is to be analyzed at the macro level, discrete-event simulation technique is more suitable when individual patient flows are to be analyzed and decisions are to be taken at the operational level (Katsaliaki & Mustafee,

**Table 1**
Queueing models of hospital systems.

| | | Authors | Type of analysis/application and special features of the model if any | Analysis tools and approaches |
|---|---|---|---|---|
| *Single node queueing models* | | | | |
| Markovian models | Steady state | 1. Wang (2004) | Theoretical paper considering a patient queue in general (deterioration of patient condition over time and patient priority considered) | Analytical formulation |
| | | 2. Green & Nguyen, 2001; Cochran and Roche (2008) | Hospital bed planning | AQF |
| | | 3. De Bruin et al. (2007) | Bed allocation (bed blocking taken into consideration) | AQF, Markov process analysis |
| | Time dependent | De Bruin et al. (2007) | Bed allocation (bed blocking taken into consideration, Time varying arrival rates represented by dividing the day into a number of time intervals) | AQF, Markov process analysis |
| Non-Markovian models | Steady state | 1. Lehtonen, Kujala, Kouri, and Hippelainen (2007) | Analysis of impact of different process interventions on productivity of operating theatre | DES |
| | | 2. Zonderland, Boucherie, Litvak, and Vleggeert-Lankamp (2010) | Surgery planning | AF and N |
| | | 3. Gorunescu et al. (2002a) | Bed allocation (Loss model) | AQF |
| | | 4. Gorunescu et al. (2002b) | Bed allocation (waiting space considered) | AQF |
| | | 5. Li, Beullens, Jones, and Tamiz (2009) | Bed planning | AQF |
| | | 6. Zhang et al. (2009) | Operating room performance analysis (priority classes) | DES |
| | | 7. Creemers and Lambrecht (2009) | Theoretical paper, to assess waiting list performance measures in any appointment driven service system including healthcare (batch service, vacation model) | Matrix analytic methods |
| | | 8. De Bruin, Bekker, van Zanten, and Koole (2010) | Bed allocation (bed pooling, blocking) | AQF |
| | | 9. Persson and Persson (2010) | Operating room planning | DES |
| | Time Dependent | 1. Cayirli, Veral, and Rosen (2008) | Outpatient appointment scheduling | DES |
| | | 2. Hassin and Mendel (2008) | Outpatient appointment scheduling | AQF |
| | | 3. Harper and Shahani (2002) | Bed planning (patient categories) | DES |
| | | 4. Masterson et al. (2004) | Bed planning and operating policies (Patient categories) | DES |
| *Network queueing models* | | | | |
| Markovian models | Steady state | 1. Vasilakis and El-Darzi (2001) | Bed planning, understanding the factors affecting bed crisis | DES |
| | | 2. Koizumi, Kuno, and Smith (2005) | Allocation of mental health resources (open network, blocking) | A F and N |
| | | 3. Cochran and Bharti (2006) | Bed capacity planning and analysis (blocking) | AQF |
| | | 4. Chaussalet, Xie, and Millard (2006) | Bed planning (closed network) | AQF |
| | | 5. Xie, Chaussalet, and Rees (2007) | Bed planning (semi-open network) | AQF |
| | | 6. Asaduzzaman, Chaussalet, and Robertson (2010) | Bed planning (overflow and rejection probabilities) | AQF |
| | | 7. Bretthauer, Heese, Pun, and Coe (2011) | Capacity allocation in inpatient care units (blocking) | A F and N |
| | Time dependent | 1. Koizumi et al. (2005) | For allocation of mental health resources | DES |
| | | 2. Kopach et al. (2007) | Analysing the factors affecting open access scheduling of an outpatient clinic | DES |
| Non-Markovian models | Steady state | 1. Swisher et al. (2001) | Staffing, facility design and scheduling of a physician clinic (Patient categories considered) | DES |
| | | 2. Testi, Tanfani, and Torre (2007) | Analysing different sequencing rules in surgical activities | DES |
| | | 3. VanBerkel and Blake (2007) | Surgical capacity planning | DES |
| | | 4. Jiang and Giachetti (2008) | Reducing patient cycle time in an outpatient facility by parallelization of activities (multiple patient classes) | AF and N |
| | | 5. Oddoye et al. (2009) | Resource management in hospitals | DES |
| | | 6. Osorio and Bierlaire (2009) | Analysing bed blocking and their impact on various units of a hospital (Blocking) | A F and N |
| | | 7. Cochran and Roche (2009) | Capacity planning of an emergency department (multiple classes of patients) | AQF |
| | | 8. Zonderland et al. (2009) | Redesigning of processes (rescheduling of appointments and reallocation of tasks) in a hospital preanesthesia evaluation clinic | AQF |
| | | 9. Creemers and Lambrecht (2011) | Illustrated using an example hospital department involving consultation, surgery and recovery | AQF, New expressions to model service outages |

**Table 1** (*continued*)

| | | Authors | Type of analysis/application and special features of the model if any | Analysis tools and approaches |
|---|---|---|---|---|
| Non-Markovian models | Time dependent | 1. Harper and Gamlin (2003) | Evaluation of various appointment schedules for an outpatient clinic (patient categories) | DES |
| | | 2. Komashie and Mousavi (2005) | Improving the performance of an emergency department | DES |
| | | 3. Sinreich and Marmor (2005) | Analysing emergency department performance | DES |
| | | 4. Cochran and Bharti (2006) | Bed capacity planning and analysis (blocking) | DES |
| | | 5. Wijewickrama and Takakuwa (2006) | Evaluating doctor schedules and appointment schedules in an outpatient (Patient classes) | DES |
| | | 6. Gunal and Pidd (2007) | A whole hospital model represented as a combination of smaller models presented only conceptually | DES |
| | | 7. Matta and Patterson (2007) | Evaluating scheduling practices, process flow and resource levels of an outpatient department | DES |
| | | 8. Kolker (2008) and Kolker (2009) | Performance analysis of emergency department and ICU respectively | DES |
| | | 9. Ahmed and Alkhamis (2009) | Optimal staffing allocation for an emergency department | DES |
| | | 10. Chand et al. (2009) | Performance analysis and improvement of an outpatient clinic | DES |
| | | 11. Glowacka, Henry, and May (2009) | Evaluating various sequencing rules for outpatient scheduling | DES |
| | | 12. Santibáñez, Chow, French, Puterman, and Tyldesley (2009) | Scheduling, resource allocation and process improvement in an outpatient environment | DES |
| | | 13. Harper, Powell, and Williams (2010) | Nurse resource planning and allocation/nurse staffing (patient categories) | DES |
| | | 14. Koo et al. (2010) | Evaluating patient scheduling and staff allocation alternatives in an endoscopy unit | DES |
| | | 15. Rohleder et al. (2011) | Improving the performance of an orthopaedic clinic (patient classes) | DES |
| | | 16. Griffin, Xia, Peng, and Keskinocak (2012) | Bed allocation (Patient classes, blocking) | DES |
| | | 17. Izady and Worthington (2012) | Staffing A&E department | DES |

(AQF = Analytical Queueing Formulae, AF and N = Analytical Formulations and Numerical solutions or approximations, DES = Discrete Event Simulation).

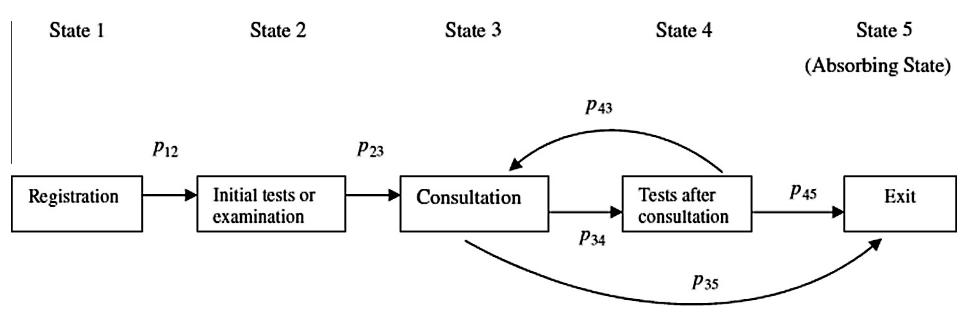


**Fig. 1.** An outpatient system as a Markov process.

2011). Here we consider only Discrete Event Simulation techniques. When the patient flows are highly complex, the solutions to the patient flow-related problems may not be achieved through analytical models as they may become invalid for a specific situation under consideration. An important characteristic of simulation modelling is that it allows to evaluate various scenarios so that what-if analyses can be performed and improvement initiatives be taken. However, unlike analytical models, simulation modelling is time consuming and requires a lot of data and expertise.

A number of researchers have used the technique of simulation for studying and analyzing outpatient environments (Chand, Moskowitz, Norris, Shade, & Willis, 2009; Koo, Jang, Nielsen, & Kolker, 2010; Swisher, Sheldon, Jacobson, Jun, & Balci, 2001). Chand et al. (2009) analyse the patient flow at an outpatient primary care clinic by identifying the sources of variability and improvement factors using simulation. Gunal and Pidd (2006) present a discrete event simulation model for the patient flows of an accident and emergency (A&E) department for analyzing the effect of multitasking and other factors affecting the performance measures of interest, such as the percentage of patients staying in A&E for more than 4 h. A simulation study to increase the throughput of an endoscopy centre is presented by Centeno et al. (2010) wherein the causes of low throughput are identified and possible changes in the operational policies are examined with the help of a fractional factorial design in order to improve performance in terms of flow time, waiting time, throughput, and utilization of the resources. Rohleder et al. (2011) use discrete event simulation to diagnose the causes of poor patient flow (high waiting times and patient congestion) and to identify improvement measures in an outpatient orthopedic clinic. Findlay and Grant (2011) analyse the operational policies of an outpatient primary clinic serving a military population with the consideration of several alternatives to identify the procedural changes that may improve the performance of the system. The arrivals to the clinic usually occur in batches with stochastic arrival times and batch sizes. Among a number of alternatives, the authors propose a model allowing both appointments and walk-in systems resulting in maximum performance improvement of the system. Studies using simulation for performance improvement of inpatient facilities aiding in decisions such as bed planning and allocation and operating room planning are also reported in the literature (e.g., Harper & Shahani, 2002; Ridge, Jones, Nielsen, & Shahani, 1998 and Persson & Persson, 2010). Holm and Dahl (2010) use discrete-event simulation for patient flow modelling of an emergency department to find the effects of increase in the number of patients on the performance and to determine the requirements of additional resources with respect to increased flow of patients in the department.

The review of literature reveals that simulation modelling has been applied extensively to model the patient flows in outpatient environments and emergency departments mainly because of the complexities of patient flows and the time-dependent characteristics of such systems. In recent years, there is a trend towards modelling larger multi-facility integrated systems and models representing multiple performance measures across several dimensions (e.g., Gunal & Pidd, 2007; Matta & Patterson, 2007). Matta and Patterson develop a stratification framework and an evaluation construct to address the problem of multiple responses in simulation experiments of outpatient clinics and illustrate the methodology by applying it to a discrete-event simulation model of a real-life, large-scale oncology centre. Gunal and Pidd (2007) present a conceptual framework to develop a whole hospital discrete event simulation model to analyse its performance by developing separate discrete event simulation models for each of the three main units, viz. the department of Accident and Emergency (A&E), outpatient clinics, and inpatient units. The A&E sub-model and the outpatient sub-model, which are the generators of emergency and elective patients respectively, are connected to the inpatient sub-model. Some recent articles (e.g., Lee, Min, Ryu, & Yih, 2013) have developed simulation models that consider two different levels of appointment scheduling problem, i.e., long-term appointment schedules together with daily appointment schedules.

*3.2.3. Statistical/empirical modelling*

Empirical models are entirely based on observations of the system characteristics and experimentations on the system for analyzing the relationship between the performance-related factors and the influencing variables and parameters related to patient flows. This type of modelling of patient flows is in its nascent stage. There are very few studies which apply entirely empirical modelling techniques for capturing patient flows, and in recent times, such statistical or empirical models are reported in the literature (Adeyemi & Chaussalet, 2008; Adeyemi & Chaussalet, 2009; Adeyemi, Chaussalet, & Demir, 2011; Adeyemi, Chaussalet, Xie, & Asaduzaman, 2010). These models essentially consider between and within patients' heterogeneity and may capture both operational and clinical patient flows. Adeyemi and Chaussalet (2008), Adeyemi and Chaussalet (2009) represent individual patient experience through the process of care by modelling their pathways using a multinomial logit random effects model (Hartzel, Agresti, & Caffo, 2001) and extract useful information on patient pathways such as identifying the pathways resulting in high probabilities of death or survival, and those resulting in short or long lengths of stay for a neonatal unit of a hospital. The sensitivity of the models as developed, to different random effects distribution assumptions are analyzed by Adeyemi and Chaussalet (2008). Patient-specific frailties are incorporated and modelled as random effects in Adeyemi and Chaussalet (2009). Both operational flows (that are physical in nature and are observable) and clinical flows (which are latent and unobservable) have been modelled. The information extracted from such models may help in allocation of scarce resources directly affecting the performance of the hospital system.

The current trend shows that these types of models are mainly used to extract information regarding pathways, such as the pathways having high probability of leading to death or about resource usage for different classes of patients and about utilization of a resource so that appropriate capacity planning and resource allocation strategies may be applied. These models are not applicable for determining waiting time measures.

## 4. Performance criteria and indicators

Performance measurement and evaluation of any system is important for continual improvements in the system, and in order to manage the performance of a system, it must be quantified and measured (Sink, 1985). Performance of a hospital system mainly assessed from the perspectives of both the service providers and the customers may have important criteria, such as efficiency, productivity, and quality of care to be considered. Presence of stakeholders with different objectives and goals and hence, having different perceptions of performance, various dimensions of performance, and dynamics of performance over time (short term as well as long term) make the performance assessment of healthcare systems a challenging task (Li & Benton, 1996; Matta & Patterson, 2007). The challenge is to find a balance or tradeoff between service for patients and efficiency for providers (Brailsford & Vissers, 2011). This paper is mainly concerned with the operational performance, specifically the ones related to time and volume (Hopp & Lovejoy, 2012). Various modelling techniques as discussed in the previous section as well as different optimization techniques are used by the researchers and practitioners for operational performance modelling of hospital systems; each department of a hospital (OPD, Emergency, OT, Inpatient Wards, ICU, Pathology, Radiology etc.) has different operational issues, and over the years various performance measures have been proposed to assess the performance of these operations (e.g., Cochran & Roche, 2009; Fei, Meskens, & Chu, 2010; Klassen & Yoogalingam, 2009). These measures include waiting times of customers, congestion measures, utilization of physicians and other resources, cost-based measures in terms of lengths of stay (Cardoen et al., 2010; Cayirli & Veral, 2003), etc. The operational performance measures which are widely used to assess the care delivery processes in the subunits viz., (i) outpatient department, (ii) emergency department, (iii) operating theatre, (iv) inpatient unit (which includes wards, ICU and other units where patients stay for treatment) and (v) diagnostic facilities (e.g., Pathology and Radiology) are summarized in Table 2.

## 5. A Generic framework for patient flow modelling and performance analysis of healthcare delivery processes in hospitals

In this section, we present a detailed generic framework for patient flow modelling and performance analysis of the processes involved during the delivery of care within a hospital system depicting all the logical steps involved in such modelling so as to help the researchers and practitioners in this field to proceed systematically towards identification of problems and improvement of a hospital system under consideration. On selection of the unit of analysis or a group of units (i.e., a subsystem of the hospital), a set of sequential steps need to be followed for modelling and analysis of the system. The proposed framework has four specific phases, viz. (i) Preliminary understanding and data collection, (ii) Data analysis, (iii) Patient Flow modelling and (iv) Performance Analysis. The details of these phases and the steps involved are described in this section. The details of the framework in terms

**Table 2**
Key performance indicators to assess the operational performance of various hospital sub-units.

| Subsystem | Analysed performance measures | References |
|---|---|---|
| OPD | Patients' average/expected waiting time in the system | Harper and Gamlin (2003), Cayirli, Veral, and Rosen (2006), Kaandorp and Koole (2007), Jerbi and Kamoun (2011) and Rohleder et al. (2011) |
| | Physician's average idle time | Cayirli et al. (2006) |
| | Physician's average overtime | Cayirli et al. (2006) |
| | Mean and standard deviation of patient waiting time | Chand et al. (2009) |
| | Clinic throughput | Kopach et al. (2007) |
| | Utilization of physicians' time | Chand et al. (2009) and Jerbi and Kamoun (2011) |
| | Time spent by patients in the system | Rohleder et al. (2011) |
| | Cost of patient waiting time, physician idle time and physician overtime | Klassen and Yoogalingam (2009) and Glowacka et al. (2009) |
| Emergency department | Waiting time from arrival till receiving care | Gunal and Pidd (2006), Ahmed and Alkhamis (2009) and Cochran and Roche (2009) |
| | Total time spent in the system from arrival to discharge or admission (Patient cycle time/completion time) | Gunal and Pidd (2006), Mayhew and Smith (2008) and Jiang and Giachetti (2008) |
| | Patient throughput | Ahmed and Alkhamis (2009) |
| | Resource utilization | Cochran and Roche (2009) |
| | Overflow probability, Percentage of patients who leave without treatment | Cochran and Roche (2009) |
| Operating theatre | Number of operations performed over a given time period or Throughput of operating theatre | Testi et al. (2007), Lehtonen et al. (2007), Zhang et al. (2009) and Santibáñez et al. (2009) |
| | Overrun hours | Testi et al. (2007) |
| | Expected number of postponements and cancellations | Testi et al. (2007) and Zonderland et al. (2010) |
| | Utilization of operating rooms, beds and other resources | Testi et al. (2007), Zhang et al. (2009) and Santibáñez et al. (2009) |
| | Patient waiting time | Zhang et al. (2009) and Santibáñez et al. (2009) |
| | Productivity of operating theatre | Lehtonen et al. (2007) |
| Inpatient care | Bed occupancy or utilization | Harper and Shahani (2002), De Bruin et al. (2007), Li et al. (2009) |
| | Average number of refused admissions | Harper and Shahani (2002), De Bruin et al. (2007) |
| | Length of stay | Christodoulou and Taylor (2001), Harper and Shahani (2002) and De Bruin et al. (2007) |
| | Blocking probability, Average number of jobs at a queue, Average number of blocked jobs at a node | Osorio and Bierlaire (2009) |
| | ICU diversions | Kolker (2009) |
| Diagnostic facilities | Indirect waiting time of patients (Access time) | Joustra et al. (2010), Patrick and Puterman (2007) |
| | Direct waiting time of patients or cost associated with waiting | Gocgun, Bresnahan, Ghate, and Gunn (2011) and Huang and Marcak (2013) |
| | Overtime of session | Huang and Marcak (2013) |
| | Utilization of scanning machine and other resources | Kolisch and Sickinger (2008) and Huang and Marcak (2013) |
| | Turnaround time (TAT), TAT outlier percentage | Holland, Smith, and Blick (2005) |
| | Number of tests performed per day | Patrick and Puterman (2007) |

of phases and steps involved, methodology, and sources of data, required tools and techniques is described with the help of a flow chart as shown in Fig. 2.

The details of all the phases are given below.

### *5.1. Preliminary Understanding and Data Collection Phase*

This phase consists of the two interrelated steps viz. preliminary understanding and data collection.

(i) **Preliminary understanding:** This step involves observation of various processes and operations of the system and discussion/interviews with management and clinical staff. Detailed flow/process charts are to be prepared and validated by the management and clinical staff. The process chart of a system qualitatively describes different patient routings, depicts the possible relationships (interactions) between various units of a hospital system, and identifies the stages in the whole process of care. Information regarding the configuration of the system, resources available and other conditions and constraints, performance metrics of interest, broad classification of patients and priorities of these classes are to be obtained through discussion with the concerned personnel and/or reference to hospital information systems and published documents/records. Discussion is also required to get an initial idea about the problems, their causes, and options for improvement.

(ii) **Data collection:** This step involves collection of relevant data to determine characteristics of the system, such as arrival pattern of patients, service rates of various services involved in the process of care, and the transition probabilities between various nodes. If the arrivals are unscheduled, data on arrival patterns are required for determining the arrival distributions for patients (combined and for all categories separately). For scheduled arrivals, the data are required to determine the probability distribution of the length of time the patients are early or late to their scheduled appointment times. The service times at all the stages are also required to be collected. The arrival and service time data (i.e., lengths of stay) for inpatients may usually be obtained from hospital records. In case of outpatient arrivals, the arrival times and service times are usually not recorded, and may not be obtained from hospital records. These are required to be collected through direct observation of events, such as times of arrivals, starting of a service, and end of a service. A track record of the total number of patients entering into the system and subsequently their pathways through the system components/subsystems and discharge from the system may be required to determine the transition probabilities from one state to another at different stages. The frequency of each possible pathway may help in identifying the busiest resources (Adeyemi & Chaussalet, 2009). The data related to performance measures of the system, such as waiting times, cycle time, idle

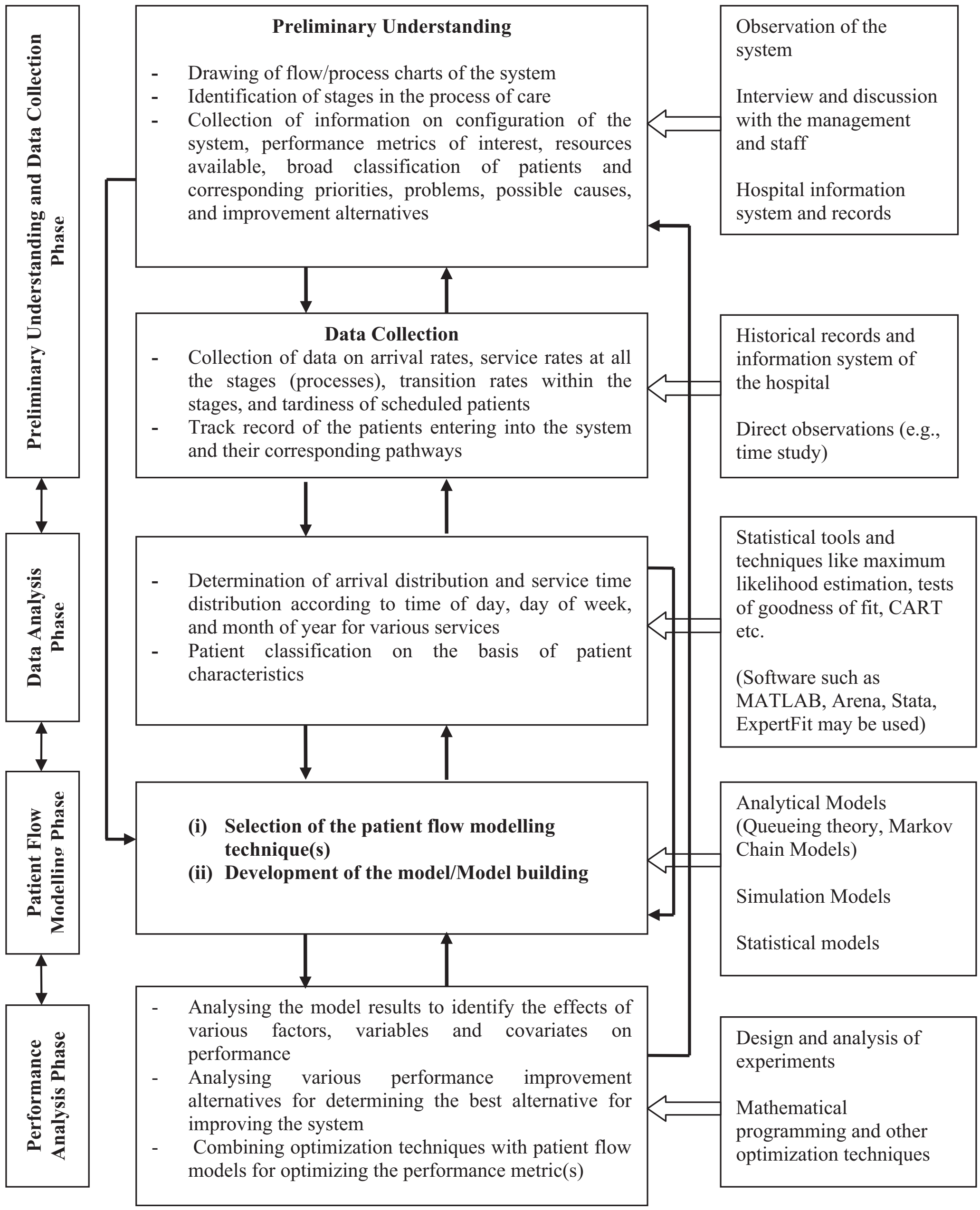


**Fig. 2.** A generic framework for patient flow modelling and performance analysis of healthcare delivery processes in hospitals.

time, overtime, and utilization of resources are also required to be recorded for identifying the problems and for validating the patient flow models to be developed.

*5.2. Data analysis phase*

The objectives of data analysis are to determine the nature of distributions followed by patient arrivals and service times and the related parameters, to determine the transition probabilities and in some cases, to classify the patients into various categories. Preliminary statistical analyses (using tools, such as histograms and box plots) pertaining to input variables and performance metrics may be required before their actual characteristics are known. Patients with similar characteristics may be grouped using techniques in survival analysis like Classification and Regression Trees (CART) (Breiman, Friedman, Olshen, & Stone, 1984). The variability in pattern of arrivals by time of day, day of week, and month of year for various services and different patient classes need to be considered. Either theoretical or empirical distributions are used for modelling. The parameters of theoretical distributions are estimated using methods, such as the method of matching moments or quantiles, the method of maximum likelihood, and the method of least squares, and the goodness-of-fit is tested using statistical tests, such as $\chi^2$ test, Kolmogorov–Smirnov (K–S) test and Anderson–Darling test.

*5.3. Patient flow modelling phase*

This phase consists of the following two interrelated steps, viz. selection of modelling technique and model building.

(i) **Selection of modelling technique**: At this step, the appropriate technique(s) for modelling the patient flow is to be selected depending upon the problems faced by a system

**Table 3**
Selection of modelling techniques based on patient flow characteristics and type of problem to be solved or analyses to be performed.

| Patient flow modelling technique | Characteristics of the patient flows and the type of problem to be solved |
|---|---|
| Analytical queueing theoretic models | (i) More suitable for modelling simple patient flows with less number of stages, homogeneous patients and Markovian inter-arrival and service time distributions<br>(ii) More suitable for performance analysis of steady state systems<br>(iii) Applicable for determining performance measures related to waiting time, congestion and utilization |
| Markov chains and compartmental models | (i) Suitable for modelling both clinical and operational patient flows<br>(ii) More appropriate for capacity planning and resource allocation decisions<br>(iii) Analysing performance measures related to waiting time is not possible |
| Discrete event simulation | (i) Is an appropriate technique when patient flows are quite complex in terms of number of stages, number of classes of patients, patient priorities and routing probabilities<br>(ii) May be easily used for transient analysis of performance of the system under consideration<br>(iii) Suitable for almost any type of application related to capacity planning, resource allocation and scheduling |
| Statistical or empirical models | (i) Suitable for modelling both clinical and operational patient flows<br>(ii) More appropriate for extracting information about patient pathways for different patient classes<br>(iii) Analysing performance measures related to waiting time is not possible |

under consideration, types of analyses required to be performed, the complexity of patient flows, and the level of detail and generality required. The criteria for selection of modelling techniques based on the characteristics of patient flows and the types of problem to be solved or analyses to be performed is provided in Table 3.

(ii) **Model Building**: After the selection of the appropriate modelling technique(s), the patient flows in the system under consideration are modelled incorporating the aspects as observed during data collection phase and using the parameters computed after analysis of data. For queueing models, the structure and configuration may be determined using the conceptual model and the data collected. While fitting Markov models to lengths of stay of patients, the number of states may be determined using Akaike's or Bayesian information criteria (Xie et al., 2005). Similarly for simulation modelling, the characteristics observed and the data collected and analysed are used to build the model.

### 5.4. Performance analysis phase

After formulating the models, the next step is to analyse various performance metrics to find the relationships between the performance metrics and the factors affecting the performance either using the analytical formulae or numerical solutions or approximations of analytical formulations or through simulation of various possible scenarios. Once the patient flows of the system are modelled and the current performance of a hospital system is assessed, alternatives may be required to be identified to improve or optimize the performance of the system. Varieties of factorial experiments may be designed to explore various decision alternatives. Each scenario usually consists of a number of levels for the factors to be tested. The factors may be related to operational changes or changes in the process of care, appointment scheduling rules, or allocation and scheduling of resources. The results as obtained from the performance analysis give insights for making changes in the existing system. A number of optimization tools and techniques, such as linear programming, integer programming and goal programming are recommended for performance optimization of hospital systems (Kuo, Schroeder, Mahaffey, & Bollinger, 2003; Oddoye, Jones, Tamiz, & Schmidt, 2009; Zhang, Murali, Dessouky, & Belson, 2009). Integration of the optimization techniques with the patient flow modelling techniques may help in improving the overall performance of a hospital system. The most widespread trend is to combine a simulation model with an optimization algorithm to find the optimal value of design parameters in order to optimize the performance of the simulated system (also known as simulation optimization techniques) (Ahmed & Alkhamis, 2009; De Angelis, Felici, & Impelluso, 2003; Klassen & Yoogalingam, 2013; Ozcan, Tànfani, & Testi, 2011). The combination of these modelling and optimization methods may lead to design of decision support systems which may help the managers in improving the performance of the system. The basic idea in this context is to analyze the patient flows as modelled and to optimize one or more performance measures of interest subject to resource or budget constraints by combining the optimization techniques with the modelling techniques for patient flows. Depending upon the type of the problem and the objectives of the study, optimization methods may be combined with the patient flow modelling techniques either iteratively or before modelling or after modelling.

## 6. Concluding remarks

A systematic approach towards modelling and analysis of patient flows helps in identifying various problems and alternatives for addressing these problems and improving the performance of a hospital system. In this paper, the recent literature pertaining to patient flow modelling and performance analysis of hospital systems is reviewed, and a generic framework for such performance modelling and analysis is provided. This framework attempts to provide a systematic approach which includes the overall process of performance improvement of a hospital system starting from understanding the system, identifying the problem areas, data collection, selection of modelling technique, model building, assessing the performance and finally making improvements in the system, and can be referred to for a systematic step-by-step procedure towards improving the operational performance of any hospital system.

Review of various patient flow modelling techniques and their applications as reported in the literature reveal that these techniques have long been used to address the issues and gain insights related to performance analysis of subsystems of a hospital. Simulation is the most widely used technique in general owing to the flexibility in modelling patient flow complexities and time-dependent behavior of a system. Moreover, simulation is suitable for almost any type of analyses to be performed and any kind of operational performance measures to be computed. It is found that the analytical queueing models have mostly been used to represent a single stage of a care process (i.e., single node models), and steady state systems. Although closed form solutions may not be obtained for real-life queueing situations which generally have characteristics, such as non-Poisson and time-varying arrivals and non-exponential service time distributions, such characteristics need to be incorporated in the patient flow model. As an approach towards theoretical contribution, finding exact or approximate numerical solutions for queueing systems incorporating specific characteristics of hospital patient flows is an important current research area (in this context, the models developed by Creemers & Lambrecht, 2011 providing expressions for modelling service outages is worth mentioning) so that the real system may be modelled more closely. Moreover, there is enough scope for applying available analytical queueing network results which have not yet been applied in healthcare context. Markov chains and compartment models are mainly suitable for modelling inpatient length of stay and clinical patient flows. Statistical modelling of patient flows seems to be in nascent stage, and more interesting applications are expected to come up in future. Whatever may be the technique for patient flow modelling, the current requirement and trend is to model integrated multi-facility systems and multi-stage patient flows (e.g., Matta & Patterson, 2007 and Lin, Patrick, & Labeau, 2014) so that the patient flow for the whole process of care may be represented and interaction between successive stages may be explored. Further, another important area for future research and methodological developments is the integration of optimization techniques with the patient flow modelling techniques as well as the combined application of different patient flow modelling techniques for different purposes while addressing a problem related to the operations of a hospital system.

## References


Adeyemi, S., & Chaussalet, T. J. (2008). A random effects sensitivity analysis for patient pathways model. In: S. Puuronen, M. Pechenizkiy, A. Tsymbal, & D. J. Lee (Eds.), *Proceedings of 21st IEEE International Symposium on Computer-Based Medical Systems* (pp. 536–538). Jyväskylä.

Adeyemi, S., Chaussalet, T., & Demir, E. (2011). Nonproportional random effects modelling of a neonatal unit operational patient pathways. *Statistical Methods and Applications, 20*(4), 507–518.

Adeyemi, S., & Chaussalet, T. J. (2009). Models for extracting information on patient pathways. In S. McClean et al. (Eds.), *Intelligent patient management, SCI 189* (pp. 171–182). USA: Springer-Verlag.

Adeyemi, S., Chaussalet, T. J., Xie, H., & Asaduzaman, M. (2010). Random effects models for operational patient pathways. *Journal of Applied Statistics, 37*(4), 691–701.

Ahmed, M. A., & Alkhamis, T. M. (2009). Simulation optimization for an emergency department healthcare unit in Kuwait. *European Journal of Operational Research, 198*(3), 936–942.

Alexopoulos, C., Goldsman, D., Fontanesi, J., Kopald, D., & Wilson, J. R. (2008). Modeling patient arrivals in community clinics. *Omega, 36*(1), 33–43.

Asaduzzaman, M., Chaussalet, T. J., & Robertson, N. J. (2010). A loss network model with overflow for capacity planning of a neonatal unit. *Annals of Operations Research, 178*(1), 67–76.

Baily, N. T. J. (1952). A study of queues and appointment systems in hospital out-patient departments, with special reference to waiting-times. *Journal of the Royal Statistical Society. Series B (Methodological), 14*(2), 185–199.

Brahimi, M., & Worthington, D. J. (1991). Queueing models for out-patient appointment systems – a case study. *Journal of the Operational Research Society, 42*(9), 733–746.

Brailsford, S. C., Harper, P. R., Patel, B., & Pitt, M. (2009). An analysis of the academic literature on simulation and modelling in health care. *Journal of Simulation, 3*(3), 130–140.

Brailsford, S., & Vissers, J. (2011). OR in healthcare: A European perspective. *European Journal of Operational Research, 212*(2), 223–234.

Brandeau, M. L., Sainfort, F., & Pierskalla, W. P. (Eds.). (2004). *Operations research and health care: A handbook of methods and applications*. USA: Springer-Verlag.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and Regression Trees. *Wadworth and Brooks/Cole*. Monterey.

Bretthauer, K. M., Heese, H. S., Pun, H., & Coe, E. (2011). Blocking in healthcare operations: A new heuristic and an application. *Production and Operations Management, 20*(3), 375–391.

Cardoen, B., Demeulemeester, E., & Belien, J. (2010). Operating room planning and scheduling: A literature review. *European Journal of Operational Research, 201*(3), 921–932.

Cayirli, T., & Veral, E. (2003). Outpatient scheduling in health care: A review of literature. *Production and Operations Management, 12*(4), 519–549.

Cayirli, T., Veral, E., & Rosen, H. (2006). Designing appointment scheduling systems for ambulatory care services. *Health Care Management Science, 9*(1), 47–58.

Cayirli, T., Veral, E., & Rosen, H. (2008). Assessment of patient classification in appointment system design. *Production and Operations Management, 17*(3), 338–353.

Centeno, M. A., Dodd, H. R., Aranda, M., & Sanchez, Y. (2010). A simulation study to increase throughput in an endoscopy center. In: B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, & E. Yiicesan (Eds.), *Proceedings of the 2010 Winter Simulation Conference* pp. (2462–2473).

Chand, S., Moskowitz, H., Norris, J. B., Shade, S., & Willis, D. R. (2009). Improving patient flow at an outpatient clinic: Study of sources of variability and improvement factors. *Health Care Management Science, 12*(3), 325–340.

Chaussalet, T. J., Xie, H., & Millard, P. H. (2006). A closed queueing network approach to the analysis of patient flow in health care systems. *Methods of Information in Medicine, 45*(5), 492–497.

Christodoulou, G., & Taylor, G. J. (2001). Using a continuous time hidden Markov process, with covariates, to model bed occupancy of people aged over 65 years. *Health Care Management Science, 4*(1), 21–24.

Cochran, J. K., & Bharti, A. (2006). Stochastic bed balancing of an obstetrics hospital. *Health Care Management Science, 9*(1), 31–45.

Cochran, J. K., & Roche, K. (2008). A queuing-based decision support methodology to estimate hospital inpatient bed demand. *Journal of the Operational Research Society, 59*(11), 1471–1482.

Cochran, J. K., & Roche, K. T. (2009). A multi-class queuing network analysis methodology for improving hospital emergency department performance. *Computers and Operations Research, 36*(5), 1497–1512.

Cote, M. J. (1999). Patient flow and resource utilization in an outpatient clinic. *Socio-Economic Planning Sciences, 33*(3), 231–245.

Cote, M. J. (2000). Understanding patient flow. *Decision Line, 33*, 8–10.

Cote, M. J., & Stein, W. E. (2007). A stochastic model for a visit to the doctor's office. *Mathematical and Computer Modelling, 45*(3–4), 309–323.

Creemers, S., & Lambrecht, M. (2009). An advanced queueing model to analyze appointment-driven service systems. *Computers and Operations Research, 36*(10), 2773–2785.

Creemers, S., & Lambrecht, M. (2011). Modeling a hospital queueing network. In R. J. Boucherie & N. M. Van Dijk (Eds.), *Queueing networks: A fundamental approach, international series in operations research and management science 154* (pp. 767–798). USA: Springer-Verlag.

De Angelis, V., Felici, G., & Impelluso, P. (2003). Integrating simulation and optimization in health care centre management. *European Journal of Operational Research, 150*(1), 101–114.

De Bruin, A. M., Bekker, R., van Zanten, L., & Koole, G. M. (2010). Dimensioning hospital wards using the Erlang Loss model. *Annals of Operations Research, 178*(1), 23–43.

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.

Eldabi, T., Paul, R. J., & Young, T. (2007). Simulation modelling in healthcare: Reviewing legacies and investigating futures. *Journal of the Operational Research Society, 58*(2), 262–270.

El-Darzi, E., Vasilakis, C., Chaussalet, T., & Millard, P. H. (1998). A simulation modelling approach to evaluating length of stay, occupancy, emptiness and bed blocking in a hospital geriatric department. *Health Care Management Science, 1*(2), 143–149.

Fackrell, M. (2009). Modelling healthcare systems with phase-type distributions. *Health Care Management Science, 12*(1), 11–26.

Faddy, M. J., & McClean, S. I. (2005). Markov chain modelling for geriatric patient care. *Methods of Information in Medicine, 44*(3), 369–373.

Fei, H. N., Meskens, N., & Chu, C. (2010). A planning and scheduling problem for an operating theatre using an open scheduling strategy. *Computers and Industrial Engineering, 58*(2), 221–230.

Findlay, M. & Grant, H. (2011). An application of discrete-event simulation to an outpatient healthcare clinic with batch arrivals. In: S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, & M. Fu (Eds.), *Proceedings of the 2011 Winter Simulation Conference* pp. (1166–1177).

Fone, D., Hollinghurst, S., Temple, M., Round, A., Lester, N., Weightman, A., et al. (2003). Systematic review of the use and value of computer simulation modelling in population health and health care delivery. *Journal of Public Health Medicine, 25*(4), 325–335.

Garg, L., McClean, S., Meenan, B., El-Darzi, E., & Millard, P. (2009). Clustering patient length of stay using mixtures of Gaussian models and phase type distributions. In *Proceedings of 22nd IEEE Symposium on Computer-Based Medical Systems (CBMS 2009)* (pp. 1–7). New Mexico, USA: Albuquerque.

Garg, L., McClean, S., Meenan, B., & Millard, P. (2010). A non-homogeneous discrete time Markov model for admission scheduling and resource planning in a cost or capacity constrained healthcare system. *Health Care Management Science, 13*(2), 155–169.

Gillespie, J., McClean, S., Scotney, B., Garg, L., Barton, M., & Fullerton, K. (2011). Costing hospital resources for stroke patients using phase-type models. *Health Care Management Science, 14*(3), 279–291.

Glowacka, K. J., Henry, R. M., & May, J. H. (2009). A hybrid data mining/simulation approach for modelling outpatient no-shows in clinic scheduling. *Journal of the Operational Research Society, 60*(8), 1056–1068.

Gocgun, Y., Bresnahan, B. W., Ghate, A., & Gunn, M. L. (2011). A Markov decision process approach to multi-category patient scheduling in a diagnostic facility. *Artificial Intelligence in Medicine, 53*(2), 73–81.

Gorunescu, M., Gorunescu, F., & Prodan, A. (2002c). Continuous-time Markov model for geriatric patients behavior. Optimization of the bed occupancy and computer simulation. *Korean Journal of Computational and Applied Mathematics, 9*(1), 185–195.

Gorunescu, F., McClean, S. I., & Millard, P. H. (2002a). A queueing model for bed occupancy management and planning of hospitals. *Journal of Operational Research Society, 53*(1), 19–24.

Gorunescu, F., McClean, S. I., & Millard, P. H. (2002b). Using a queueing model to help plan bed allocation in a department of geriatric medicine. *Health Care Management Science, 5*(4), 307–312.

Green, L. (2006). Queueing analysis in healthcare. In R. W. Hall (Ed.), *Patient flow: Reducing delay in healthcare delivery* (pp. 281–307). USA: Springer-Verlag.

Green, L. V., & Nguyen, V. (2001). Strategies for cutting hospital beds: The impact on patient service. *Health Services Research, 36*(2), 421–442.

Griffin, J., Xia, S., Peng, S., & Keskinocak, P. (2012). Improving patient flow in an obstetric unit. *Health Care Management Science, 15*(1), 1–14.

Gross, D., & Harris, C. M. (2011). *Fundamentals of queueing theory* (3rd ed.). John Wiley and Sons. Authorized Reprint by Wiley India (P.) Ltd., New Delhi.

Gunal, M. M. (2012). A guide for building hospital simulation models. *Health Systems, 1*(1), 17–25.

Gunal, M. M. & Pidd, M. (2006). Understanding accident and emergency department performance using simulation. In: L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, & R. M. Fujimoto (Eds.), *Proceedings of the 2006 Winter Simulation Conference* (pp. 446–452). Monterey, CA.

Gunal, M. M. & Pidd, M. (2007). Interconnected DES models of emergency, outpatient, and inpatient departments of a hospital. In: S. G. Henderson, B. Biller, M. H. Hsieh, J. Shortle, J. D. Tew, & R. R. Barton (Eds.), *Proceedings of the 2007 Winter Simulation Conference* (pp. 1461–1466). Washington, DC.

Gunal, M. M., & Pidd, M. (2010). Discrete event simulation for performance modelling in healthcare: a review of the literature. *Journal of Simulation, 4*(1), 42–51.

Hall, R. W. (Ed.). (2006). *Patient flow: Reducing delay in healthcare delivery*. USA: Springer-Verlag.

Hall, R., Belson, D., Murali, P., & Dessouky, M. (2006). Modeling patient flows through the healthcare system. In R. W. Hall (Ed.), *Patient flow: Reducing delay in healthcare delivery* (pp. 1–44). USA: Springer-Verlag.

Harper, P. R. (2002). A framework for operational modelling of hospital resources. *Health Care Management Science, 5*(3), 165–173.

Harper, P. R., & Gamlin, H. M. (2003). Reduced outpatient waiting times with improved appointment scheduling: A simulation modelling approach. *OR Spectrum, 25*(2), 207–222.

Harper, P. R., Powell, N. H., & Williams, J. E. (2010). Modelling the size and skill-mix of hospital nursing teams. *Journal of the Operational Research Society, 61*(5), 768–779.

Harper, P. R., & Shahani, A. K. (2002). Modelling for planning and management of bed capacities in hospitals. *The Journal of the Operational Research Society, 53*(1), 11–18.

Harrison, G. W. (2001). Implications of mixed exponential occupancy distributions and patient flow models for health care planning. *Health Care Management Science, 4*(1), 37–45.

Harrison, G. W., & Escobar, G. J. (2010). Length of stay and imminent discharge probability distributions from multistage models: Variation by diagnosis, severity of illness, and hospital. *Health Care Management Science, 13*(3), 268–279.

Harrison, G. W., Shafer, A., & Mackay, M. (2005). Modelling variability in hospital bed occupancy. *Health Care Management Science, 8*(4), 325–334.

Hartzel, J., Agresti, A., & Caffo, B. (2001). Multinomial logit random effects models. *Statistical Modelling, 1*(2), 81–102.

Hassin, R., & Mendel, S. (2008). Scheduling arrivals to queues: A single-server model with no-shows. *Management Science, 54*(3), 565–572.

Holland, L. L., Smith, L. L., & Blick, K. E. (2005). Reducing laboratory turnaround time outliers can reduce emergency department patient length of stay. *American Journal of Clinical Pathology, 124*, 672–674.

Holm, L. B. & Dahl, F. A. (2010). Simulating the influence of a 45% increase in patient volume on the emergency department of Akershus University hospital. In: B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, & E. Yiicesan (Eds.), *Proceedings of the 2010 Winter Simulation Conference* (pp. 2455–2461).

Hopp, W. J., & Lovejoy, W. S. (2012). *Hospital operations: Principles of high efficiency health care* (1st Edition). Upper Saddle River, New Jersey, USA: FT Press.

Huang, Y. L., & Marcak, J. (2013). Radiology scheduling with consideration of patient characteristics to improve patient access to care and medical resource utilization. *Health Systems, 2*(2), 93–102.

Hulshof, P. J. H., Kortbeek, N., Boucherie, R. J., Hans, E., & Bakker, P. J. M. (2012). Taxonomic classification of planning decisions in health care: a review of the state of the art in OR/MS. *Health Systems, 1*(2), 129–175.

Izady, N., & Worthington, D. (2012). Setting staffing requirements for time dependent queueing networks: The case of accident and emergency departments. *European Journal of Operational Research, 219*(3), 531–540.

Jerbi, B., & Kamoun, H. (2011). Multiobjective study to implement outpatient appointment system at Hedi Chaker Hospital. *Simulation Modelling Practice and Theory, 19*(5), 1363–1370.

Jiang, L., & Giachetti, R. E. (2008). A queueing network model to analyze the impact of parallelization of care on patient cycle time. *Health Care Management Science, 11*(3), 248–261.

Joustra, P. E., Wit, J. D., Struben, V. M. D., Overbeek, B. J. H., Fockens, P., & Elkhuizen, S. G. (2010). Reducing access times for an endoscopy department by an iterative combination of computer simulation and Linear Programming. *Health Care Management Science, 13*(1), 17–26.

Jun, J. B., Jacobson, S. H., & Swisher, J. R. (1999). Application of discrete-event simulation in health care clinics: A survey. *Journal of the Operational Research Society, 50*(2), 109–123.

Kaandorp, G. C., & Koole, G. (2007). Optimal outpatient appointment scheduling. *Health Care Management Science, 10*(2), 217–229.

Katsaliaki, K., & Mustafee, N. (2011). Applications of simulation within the healthcare context. *Journal of the Operational Research Society, 62*(8), 1431–1451.

Kim, S. C., Horwitz, I., Young, K. K., & Buckley, T. A. (1999). Analysis of capacity management of the intensive care unit in a hospital. *European Journal of Operational Research, 115*(1), 36–46.

Klassen, K. J., & Yoogalingam, R. (2009). Improving performance in outpatient services with a simulation optimization approach. *Production and Operations Management, 18*(4), 447–458.

Klassen, K. J., & Yoogalingam, R. (2013). Appointment system design with interruptions and physician lateness. *International Journal of Operations and Production Management, 33*(4), 394–414.

Koizumi, M., Kuno, E., & Smith, T. E. (2005). Modeling patient flows using a queuing network with blocking. *Health Care Management Science, 8*(1), 49–60.

Kolisch, R., & Sickinger, S. (2008). Providing radiology health care services to stochastic demand of different customer classes. *OR Spectrum, 30*, 375–395.

Kolker, A. (2008). Process modeling of emergency department patient flow: Effect of patient length of stay on ED diversion. *Journal of Medical Systems, 32*(5), 389–401.

Kolker, A. (2009). Process modeling of ICU patient flow: Effect of daily load leveling of elective surgeries on ICU diversion. *Journal of Medical Systems, 33*(1), 27–40.

Komashie, A. & Mousavi, A. (2005). Modeling emergency departments using discrete event simulation techniques. In: M. E. Kuhl, N. M. Steiger, F. B. Armstrong, & J. A. Joines (Eds.), *Proceedings of the 2005 Winter Simulation Conference* (pp. 2681–2685).

Koo, P. H., Jang, J., Nielsen, K. B. & Kolker, A. (2010). Simulation-based patient flow analysis in an endoscopy unit. In: *Proceedings of 2010 IEEE workshop on Healthcare, Management* (pp. 1–6).

Kopach, R., DeLaurentis, P. C., Lawley, M., Muthuraman, K., Ozsen, L., Rardin, R., et al. (2007). Effects of clinical characteristics on successful open access scheduling. *Health Care Management Science, 10*(2), 111–124.

Kuo, P. C., Schroeder, R. A., Mahaffey, S., & Bollinger, R. R. (2003). Optimization of operating room allocation using linear programming techniques. *Journal of American College of Surgeons, 197*(6), 889–895.

Lee, S., Min, D., Ryu, J. H., & Yih, Y. (2013). A simulation study of appointment scheduling in outpatient clinics: Open access and overbooking. *Simulation, 89*(12), 1459–1473.

Lehtonen, J. M., Kujala, J., Kouri, J., & Hippelainen, M. (2007). Cardiac surgery productivity and throughput improvements. *International Journal of Health Care Quality Assurance, 20*(1), 40–52.

Li, L. X., & Benton, W. C. (1996). Performance measurement criteria in health care organizations: Review and future research directions. *European Journal of Operational Research, 93*(3), 449–468.

Li, X., Beullens, P., Jones, D., & Tamiz, M. (2009). An integrated queuing and multi-objective bed allocation model with application to a hospital in China. *Journal of the Operational Research Society, 60*(3), 330–338.

Lin, D., Patrick, J., & Labeau, F. (2014). Estimating the waiting time of multi-priority emergency patients with downstream blocking. *Health Care Management Science, 17*(1), 88–99.

Lowery, J. C. (1996). Introduction to simulation in health care. In: J. M. Charnes, D. J. Morrice, D. T. Brunner, & J. J. Swain (Eds.), *Proceedings of the 1996 Winter Simulation Conference* (pp. 78–84).

Luo, J., Kulkarni, V. G., & Ziya, S. (2012). Appointment scheduling under patient no-shows and service interruptions. *Manufacturing and Service Operations Management, 14*(4), 670–684.

Marshall, A. H. & Zenga, M. (2009). Recent developments in fitting Coxian Phasetype distributions in healthcare. In: L. Sakalauskas, C. Skiadas, & E. K. Zavadskas (Eds.), *Proceedings of The XIII International Conference "Applied Stochastic Models and Data Analysis"* (pp. 482–485). Vilnius.

Marshall, A. H., & McClean, S. I. (2004). Using coxian phase-type distributions to identify patient characteristics for duration of stay in hospital. *Health Care Management Science, 7*(4), 285–289.

Marshall, A., Vasilakis, C., & El-Darzi, E. (2005). Length of stay-based patient flow models: Recent developments and future directions. *Health Care Management Science, 8*(3), 213–220.

Masterson, B. J., Mihara, T. G., Miller, G., Randolph, S. C., Forkner, M. E., & Crouter, A. L. (2004). Using models and data to support optimization of the military health system: A case study in an intensive care unit. *Health Care Management Science, 7*(3), 217–224.

Matta, M. E., & Patterson, S. S. (2007). Evaluating multiple performance measures across several dimensions at a multi-facility outpatient center. *Health Care Manage Science, 10*(2), 173–194.

Mayhew, L., & Smith, D. (2008). Using queuing theory to analyse the Government's 4-h completion time target in Accident and Emergency departments. *Health Care Management Science, 11*(1), 11–21.

McClean, S., Faddy, M. & Millard, P. (2005). Markov Model-Based Clustering for Efficient Patient Care. In: A. Tsymbal, & P. Cunningham (Eds.), *Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems* (pp. 467–472). Dublin.

McClean, S. I., McAlea, B., & Millard, P. H. (1998). Using a Markov reward model to estimate spend-down costs for a geriatric department. *Journal of the Operational Research Society, 49*(10), 1021–1025.

McClean, S. I., & Millard, P. H. (1998). A three compartment model of the patient flows in a geriatric department: a decision support approach. *Health Care Management Science, 1*(2), 159–163.

McLaughlin, D. B., & Hays, J. M. (2009). *Healthcare Operations Management*. New Delhi, India: PHI Learning Private Limited.

Oddoye, J. P., Jones, D. F., Tamiz, M., & Schmidt, P. (2009). Combining simulation and goal programming for healthcare planning in a medical assessment unit. *European Journal of Operational Research, 193*(1), 250–261.

Osorio, C., & Bierlaire, M. (2009). An analytic finite capacity queueing network model capturing the propagation of congestion and blocking. *European Journal of Operational Research, 196*(3), 996–1007.

Ozcan, Y. A., Tànfani, E. & Testi, A. (2011). A simulation-based modeling framework to deal with clinical pathways. In: S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, & M. Fu (Eds.), *Proceedings of the 2011 Winter Simulation Conference* (pp. 1190–1201).

Patrick, J., & Puterman, M. L. (2007). Improving resource utilization for diagnostic services through flexible inpatient scheduling: A method for improving resource utilization. *Journal of the Operational Research Society, 58*, 235–245.

Persson, M. J., & Persson, J. A. (2010). Analysing management policies for operating room planning using simulation. *Health Care Management Science, 13*(2), 182–191.

Rais, A., & Viana, A. (2010). Operations research in healthcare: A survey. *International Transactions in Operational Research, 18*(1), 1–31.

Ridge, J. C., Jones, S. K., Nielsen, M. S., & Shahani, A. K. (1998). Capacity planning for intensive care units. *European Journal of Operational Research, 105*(2), 346–355.

Rohleder, T. R., Lewkonia, P., Bischak, D. P., Duffy, P., & Hendijani, R. (2011). Using simulation modeling to improve patient flow at an outpatient orthopedic clinic. *Health Care Management Science, 14*(2), 135–145.

Ross, S. M. (2009). *Stochastic processes* (2nd ed.). John Wiley and Sons. Authorized Reprint by Wiley India (P.) Ltd., New Delhi.

Santibáñez, P., Chow, V. S., French, J., Puterman, M. L., & Tyldesley, S. (2009). Reducing patient wait times and improving resource utilization at British Columbia Cancer Agency's ambulatory care unit through simulation. *Health Care Management Science, 12*(4), 392–407.

Shaw, B., & Marshall, A. H. (2007). Modelling the flow of congestive heart failure patients through a hospital system. *Journal of the Operational Research Society, 58*(2), 212–218.

Sink, D. S. (1985). *Productivity management: Planning, measurement and evaluation, control and improvement*. John Wiley and Sons Inc.

Sinreich, D., & Marmor, Y. (2005). Emergency department operations: The basis for developing a simulation tool. *IIE Transactions, 37*(3), 233–245.

Spangler, W. E., Strum, D. P., Vargas, L. G., & May, J. H. (2004). Estimating procedure times for surgeries by determining location parameters for the Lognormal model. *Health Care Management Science, 7*(2), 97–104.

Swisher, J. R., Sheldon, H., Jacobson, S. H., Jun, J. B., & Balci, O. (2001). Modeling and analyzing a physician clinic environment using discrete-event (visual) simulation. *Computers and Operations Research, 28*(2), 105–125.

Testi, A., Tanfani, E., & Torre, G. (2007). A three-phase approach for operating theatre schedules. *Health Care Management Science, 10*(2), 163–172.

VanBerkel, P. T., & Blake, J. T. (2007). A comprehensive simulation for wait time reduction and capacity planning applied in general surgery. *Health Care Management Science, 10*(4), 373–385.

Vasilakis, C., & El-Darzi, E. (2001). A simulation study of the winter bed crisis. *Health Care Management Science, 4*(1), 31–36.

Wang, Q. (2004). Modeling and analysis of high risk patient queues. *European Journal of Operational Research, 155*(2), 502–515.

Wang, J., Li, J. & Howard, P. K. (2011). Modelling and analysis of care delivery services within patient rooms. In: *2011 IEEE International Conference on Automation Science and Engineering* (pp. 537–542).

Wijewickrama, A. K. A. & Takakuwa, S. (2006). Simulation analysis of an outpatient department of internal medicine in a unive4rsity hospital. In: L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, & R. M. Fujimoto (Eds.), *Proceedings of the 2006 Winter Simulation Conference* (pp. 425–432).

Worthington, D. (2009). Reflections on queue modelling from the last 50 years. *Journal of the Operational Research Society, 60*(1), S83–S92.

Xie, H., Chaussalet, T. J., & Millard, P. H. (2005). A continuous time Markov model for the length of stay of elderly people in institutional long-term care. *Journal of the Royal Statistical Society A, 168*(1), 51–61.

Xie, H., Chaussalet, T. J., & Millard, P. H. (2006). A model-based approach to the analysis of patterns of length of stay in institutional long-term care. *IEEE Transactions on Information Technology in Biomedicine, 10*(3), 512–518.

Xie, H., Chaussalet, T. & Rees, M. (2007). A Semi-open queueing network approach to the analysis of patient flow in healthcare systems. In: *Proceedings of Twentieth IEEE International Symposium on Computer-Based Medical Systems (CBMS'07)* (pp. 719–724).

Yeon, N., Lee, T. & Jang, H. (2010). Outpatient appointment scheduling with milti-doctor sharing resources. In: B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, & E. Yiicesan (Eds.), *Proceedings of the 2010 Winter Simulation Conference* (pp. 3318–3329).

Zhang, B., Murali, P., Dessouky, M. M., & Belson, D. (2009). A mixed integer programming approach for allocating operating room capacity. *Journal of the Operational Research Society, 60*(5), 663–673.

Zhao, L., & Lie, B. (2008). Modeling and simulation of patient flow in hospitals for resource utilization. In *Proceedings of 49th Scandinavian conference on simulation and modeling*. Oslo University College.

Zonderland, M. E., Boer, F., Boucherie, R. J., de Roode, A., & van Kleef, J. W. (2009). Redesign of a University Hospital Preanesthesia evaluation clinic using a queuing theory approach. *Anesthesia and Analgesia, 109*(5), 1612–1621.

Zonderland, M. E., & Boucherie, R. J. (2012). Queueing networks in healthcare systems. In R. Hall (Ed.), *Handbook of healthcare system scheduling, international series in operations research and management science 168* (pp. 201–243). USA: Springer-Verlag.

Zonderland, M. E., Boucherie, R. J., Litvak, N., & Vleggeert-Lankamp, C. L. A. M. (2010). Planning and scheduling of semi-urgent surgeries. *Health Care Management Science, 13*(3), 256–267.