Omega 96 (2020) 102120

Contents lists available at ScienceDirect

Omega

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

The problem of the last bed: Contextualization and a new simulation framework for analyzing physician decisions☆

Cristina Azcarate[a], Laida Esparza[b], Fermin Mallor[a,*]

[a] Department of Statistics and Operational Research, Public University of Navarre, Campus Arrosadía, 31006 Pamplona, Spain
[b] Complejo Hospitalario de Navarra, C/ Irunlarrea 3, 31008 Pamplona, Spain

ARTICLE INFO

Article history:
Received 19 March 2018
Accepted 16 September 2019
Available online 18 September 2019

Keywords:
ICU
LoS
Medical discharge decisions
Simulation in healthcare

ABSTRACT

Faced with a full Intensive Care Unit (ICU), physicians need to decide between turning away a new patient in need of critical care and creating a vacancy by prematurely discharging a current occupant. This dilemma is widely discussed in the medical literature, where the influencing factors are identified, the patient discharge process described and the patient health consequences analyzed. Nevertheless, the existing mathematical models of ICU management practices overlook many of the factors considered by physicians in real-world triage decisions.

This paper offers a review of the medical and mathematical literature on patient discharge decisions, and a proposal for a new simulation framework to enable more realistic mathematical modeling of the real-world patient discharge process. Our model includes a) the times at which discharge decisions are made and setup times for patient transfer from the ICU to a general ward and preparation of an ICU bed for an incoming patient, in order to capture the impossibility of an immediate switch of patients; b) advance notice of the number of patients due to arrive from elective surgery requiring intensive postoperative care and potentially triggering the need for early discharges to avoid surgery cancelations; and c) patient health status (to reflect the dependency of physicians' discharge decisions on health indicators) by modeling length of stay with a phase-type distribution in which a medical meaning is assigned to each state.

A simulation-based optimization method is also proposed as a means to obtain optimal discharge decisions as a function of the health status of current patients, the bed occupancy level and the number of planned arrivals from elective surgery over the following days. Optimal decisions should strike a balance between patient rejection and LoS reduction.

This new simulation framework generates an optimal discharge policy, which closely resembles real decision-making under a cautious discharge policy, where the frequency of early discharge increases with the ICU occupancy level. This is a contrast with previous simulation models, which consider only the triage of the last bed, disregarding the pressures on physicians faced with high bed occupancy levels.

1. Introduction. The ritual of the last bed

A hospital intensive care unit (ICU) provides continuous surveillance and highly specialized care to acute patients, whose conditions are life-threatening and require comprehensive care. The resources in ICUs are limited and constitute an important part of hospital budgets. Higher patient expectations and an aging population are further increasing pressure on limited ICU resources. According to [54], ICU costs amounted to $4300 per day in the US in 2010 and the total annual cost of critical care medicine was $108 billion. In a previous study, the same authors estimated critical care as 13.4% of total hospital costs, 4.1% of national health expenditures, and 0.66% of GDP [52]. In highly developed European health-care systems, the average cost per ICU patient is around €1200 per day and €17,000 per admission [41,112]. Efficient use of these resources in ICU management and, by extension, general hospital management, is therefore essential. According to the latest guidelines of the Society of Critical Care Medicine (SCCM) for ICU admission, discharge, and triage [87], further research is needed on all aspects of critical care rationing in order to address current shortcomings. Scarcity of resources can threaten or impede critical

☆ This paper was processed by Associate Editor Bish.

* Corresponding author.
E-mail addresses: cazcarate@unavarra.es (C. Azcarate), laida.esparza.artanga@navarra.es (L. Esparza), mallor@unavarra.es (F. Mallor).

https://doi.org/10.1016/j.omega.2019.102120

care provision; and the problem can only worsen with misuse of resources. There is an urgent need, therefore, to address this problem which has both current and future implications.

The hospitalization bed is used to represent a structure or quantify an activity in the healthcare context. Its use as a unit of measurement has recently come under criticism, however, because it is based on capacity rather than activity. Nonetheless, it is still widely used as a hospital management parameter both at strategic and operational level.

Different ICU bed management strategies will obviously have different impacts on ICU service quality. Given that most ICU costs are fixed—that is, independent of the level of occupancy—low occupancy should be avoided, since it implies the underuse of an extremely costly service [52]. A management policy aimed at high occupancy, however, could result in delayed or refused admission to the ICU, both of which are associated with a poorer prognosis [15,17] and a higher risk of mortality [11,58,123]. In 1993, Teres [113] stated that one of the great ethical dilemmas affecting Intensive Care Services was the admission of patients under almost full occupancy, which he described as the **ritual of the last bed**, noting the need for policies to maximize the use of resources and minimize costly triage processes. Increasing ICU occupancy levels and access block rates are leading to full or even overwhelmed ICUs [28,51]. Thus, it is no longer a question of how to allocate the last bed but how to proceed when there is none. The average ICU occupancy rate in the US is 90% [94], where it is reported that 90% of ICUs are unable to provide beds when needed [46].

The undesirable consequences of bed shortages include the following:

- *The triage of patients admitted to the ICU.* New patients requiring ICU admission have usually suffered trauma or surgery. In the event of a bed shortage, space must be re-allocated by transferring current patients to units with lower staffing and care levels. Triage decisions are important not only in terms of resource management but also in terms of intensive care outcomes. Nonetheless, despite the recommendations and implications for triage, there are few ethical decision-making scales for physicians to rely upon [9,108]. The authors of [66] report greater severity in patients discharged from the ICU during high occupancy. Unscheduled or early releases from the ICU have been associated with a greater probability of readmission [7], longer LoS in hospital, and higher mortality [40,106].
- *The triage of patients for possible ICU admission.* According to [107], the status of patients admitted to an ICU during a bed shortage tends to be more critical. Mortality rates in patients denied ICU admission are also extensively reported in medical literature [84,99,105]. The authors of [65] observe a lack of systematic ICU admission criteria and propose an econometric model to quantify the effect of ICU admission on patient outcomes.
- *Referral to other centers.* Inter-hospital transfer is associated with two potential problems: the risk inherent to the transfer [38,73,120] and delay in the proper treatment of time-dependent diseases [17,22,97].
- *Cancelation of scheduled surgeries requiring postoperative stays in the ICU.* Hospitals have large daily flows of patients from the operating theater to the ICU. The combined volume of scheduled and unscheduled arrivals leads to bed shortages and surgery cancelations [57,67,83], with negative, and even fatal consequences, for patients awaiting surgery, and an increase in the administrative pressure involved in the modification of waiting lists [21].
- *Stress in medical staff due to work overload.* The negative effects of workload saturation in the health services have been extensively discussed in the medical literature. In [60], the relationship between workload and length of stay (LoS) is assessed in an analysis involving 203 hospitals. A lack of ICU beds calls for rapid triage and referral of patients, which have been shown to increase stress levels [5,31,53,119], work overload and medical errors [1,111,116] in health staff, and thereby poorer outcomes for ICU patients.

Optimal ICU bed management is therefore crucial to providing high-quality healthcare to sick patients. Such a task is not easy, however, given the complexity of a system which involves highly specialized medical staff and equipment, different types of patients, knock-on effects from other hospital departments—such as operating theatres, emergency departments, wards, etc.—and which evolves stochastically over time. Operations research can provide valuable insights into this problem by developing simulation models that accurately reproduce ICU performance under different management policies and assess the influence of various parameters and external factors, such as elective surgeries. In addition, optimal ICU bed management policies can be obtained by combining simulation models with optimization techniques.

The main contributions of this paper can be summarized as follows:

- Review of the medical and mathematical literature on discharge decisions, underlining the main factors influencing real-practice medical decision-making and found lacking in previous ICU mathematical models.
- New simulation framework including all these factors in order to address the shortcomings of previous models and thus provide a useful tool for the analysis of medical discharge decisions.
- Mathematical model of all the factors and features considered by physicians faced with discharge decisions, such as the patientś evolving health status and the information available at the time of decision making, etc., which are found lacking in previous ICU models. Our proposal is for a new approach whereby discharge decisions are modeled as a function of the patients' current health status, the bed occupancy level and the number of planned arrivals from elective surgery over the following days.
- Proposal for a simulation-based optimization technique for obtaining optimal ICU discharge policies, considering a bi-criteria optimization model aimed at minimizing both the percentage of patient rejections and the LoS reduction.
- The optimal discharge policies thus obtained are structurally different from those generated by previous ICU models, and closely reflect real-world medical decision-making, which follows a cautious discharge policy.

The remainder of the paper is organized as follows. Section 2 provides an overview of medical and mathematical literature, with particular attention to discharge decisions, the use of simulation models in the ICU context and the probabilistic representation of ICU LoS. Section 3 discusses classic ICU simulation models, highlighting neglected aspects that would contribute to an accurate representation of the real patient discharge process and the factors considered by physicians faced with discharge decisions. This section presents a proposal for the mathematical modeling of all these features and their incorporation into a new simulation framework, which is then used to obtain optimal ICU management policies using simulation-based optimization methodology. Section 4 presents the experimental design and results. The paper closes with the conclusions and some final remarks.

2. Medical and mathematical literature review

The math-based approach to solving healthcare problems has typically focused on developing complex analytical models, some of which require assumptions that do not hold in practice. This departure from the representation of real practice in healthcare systems may explain the low rate of success obtained when implementing the findings [14]. A higher success rate can be achieved by reviewing the mathematical and medical literatures, and identifying the key factors, processes, personal attitudes, and behaviors for inclusion in mathematical models. The non-incorporation of any of these factors can compromise model validity and result in the rejection of the findings by healthcare policy-makers. The following sections therefore review both the medical and mathematical literatures to find research on the decision-making processes of physicians challenged by ICU bed pressure.

2.1. Discharge decisions

As mentioned in the latest SCCM guidelines [87], transfer from the ICU occurs, ideally, when a patient no longer meeting ICU care criteria fulfills the clinical criteria for a lower level of care. The decision is hampered by the absence of clear and objective metrics to determine which patients will continue to benefit from critical care. The guidelines also state that if a patient is clinically stable, and thus no longer requires ICU monitoring and treatment, he/she can safely be transferred to a lower-acuity area. However, the patient transfer process is conditioned not only by patient health factors but also by certain teamwork and organizational issues [74]. In a bed shortage scenario, one of the proposed solutions is to triage current ICU patients. In 2013, the Ethics Section of the European Society of Intensive Care Medicine presented a set of general triage principles under which a patient could justifiably be released from the ICU in order to admit another patient [109]. Nonetheless, despite these recommendations and the implications of triage, there are scarcely any decision-making scales for use in these circumstances [108]. Whatever the intervening factors, non-scheduled releases could be considered premature or inadvisable because the persistence of patients' severity and organic dysfunction at the time of release can compromise their final outcome [24,86]. Furthermore, and as various studies testify, ICU bed shortages are a growing problem [42,43,114].

In theory, a patient should be sufficiently stable in order to be considered for transfer to a less intensive care environment (such as an intermediate care unit or medical/surgical ward); the necessary stability assessment should be based ideally on plentiful clinical data. In practice, in the absence of predictive models of patient dynamics, clinicians must make these transfer decisions based entirely on clinical judgment [18]. Physicians are aware of the health risks involved in shortening one patient's LoS in order to admit another when the ICU is full. As stated in the introduction, these risks include a more prolonged total length of stay in hospital, and a higher risk of mortality [10,20,42,68,90,93,101,106], or readmission to the ICU [39,40,93], which is also linked to higher mortality risk and prolonged LoS [19,29,30,62,102]. In fact, the transfer of patients from ICU to a general ward is among the riskiest care transitions [69].

Recognizing the clinical risks associated with early ICU discharge, many hospitals are accounting for readmission probabilities in their discharge strategies [37,122]. Action to reduce demand-driven ICU discharge may feature in healthcare performance improvement projects, but the published research on this topic is scant [87]. The potential impact of these demand-driven discharge decisions on patient welfare presents ethical issues for the hospital and is undesirable. A deeper understanding of discharge practices could therefore ultimately improve ICU resource availability [74].

Thus, patient discharge decisions and their consequences are already implicit in much of the medical literature but, although the importance of ICU management decisions is mentioned in the mathematical literature, they are not formally modeled [23,104]. As noted in [3], researchers attempting to model and understand patient flow through a hospital typically fail to consider physicians' decision-making. Patient discharge is commonly assumed to take place regardless of the state of the system. Very few mathematical models include this decision-making process, where it is variously referred to as "bumping", "demand-driven discharge", "premature discharge" or "early discharge".

The authors of [26] develop a stochastic ICU model including patient bumping as a response to overcrowding. They consider both scheduled and unscheduled arrivals. The arrival of a new patient at a full ICU triggers the discharge of the patient with the least expected remaining LoS. The authors propose a Markov chain model for evaluating this discharge policy for different patient arrival patterns and capacity/load scenarios. Unfortunately, a patient's expected remaining LoS is rather difficult to estimate. The authors of [18] describe a lowest cost criterion for selecting a patient for early discharge in order to admit a new arrival when the ICU is full. They also discuss various cost function inputs, including mortality risk and readmission risk at different occupancy levels. Finally, they propose a patient criticality measure based on an increase of the readmission load score. They assume a memoryless geometric LoS distribution. A similar study, using non-memoryless LoS distributions, is carried out in [56], where the readmission probability and expected LoS following readmission are considered when selecting a patient for premature discharge. A dynamic programming model is used in [72] to study admissions and premature discharge decisions in ICUs with bed management policies based on bed reservation for more critically-ill patients. Arriving patients are divided into two classes: those with a higher probability of survival after ICU admission, who cannot be kept waiting or referred to another unit, and those with a lower probability of survival who will be held for transfer to another unit if no ICU beds are available.

All these models depict *"aggressive"* discharge policies [77]: whereby no action is taken until there are no remaining beds for an incoming patient, at which point an instantaneous exchange of patients takes place. However, physicians consider a *"cautious"* policy to be more representative of their decision-making in practice. A cautious strategy dictates that the frequency of early discharge must increase with ICU occupancy. Thus, patient triage is not delayed until the last bed, but begun in high occupancy situations in anticipation of scheduled and urgent admissions. Such advance discharge planning enables ICU physicians to avoid extreme occupancy situations and discharge to take place at conventional hours, avoiding night shifts, to ensure sufficient staffing at the patient's destination and avoid emergency bed management issues.

Adopting this anticipation and management approach, the authors of [77] propose a queuing control problem to obtain efficient bed management policies, with service rates dependent on occupancy levels. They propose bi-objective optimization problems to minimize both patient rejection and LoS reduction. However, an instantaneous switch of patients for an ICU bed is also assumed.

2.2. Simulation in health and in ICU

Hospitals are highly complex stochastic systems involving a large number of interacting agents. At the same time, hospital managers face growing pressures to increase the quality and quantity of hospital services using limited resources [59]. Optimal system logistics management requires tools for interpreting the behavior of the system and predicting different scenario outcomes [117]. In this context, simulation emerges as the most suitable analytical tool, since it is a powerful quantitative instrument for the analysis of complex systems, and commonly used in combination with other statistical and optimization techniques.

The specialist literature contains numerous bibliographical references relating to the use of simulation models for decision making in the healthcare context. Since the first work was published in 1965 [35], these models have been used to analyze various problems, such as patient flow, bed-planning, waiting list management, health service design, medical staff scheduling, operating theater management, etc. The reader can refer to [13,36,50,61,85,95,124] for reviews of the use of simulation models in healthcare.

ICU sizing and management optimization are other classic problems often addressed with simulation modeling. A review of the use of operations research methods, including simulation, in ICU management appears in [6]. Some medical journals include simulation studies aimed at providing mathematical solutions to ICU capacity problems [23,83,89,91,103,110,125] and the need to optimize the distribution of beds and elective admissions [67,115,121]. The mathematical literature also includes simulation models for analyzing ICU capacity problems [75,81,98] and ICU admission and discharge processes [63]; for comparing bed allocation rules using bi-objective optimization [64]; for bed management optimization making a distinction between emergency and elective surgery patients [47]; for analyzing changes in the patient-flow circuit with the use of intermediate care wards [79,100]; for adjusting staffing to current bed occupancy [49]; and for assessing bed occupancy and patient transfers to other ICU facilities due to resource shortage [110]. The ultimate aim of all these models is to reconcile bed availability with bed occupancy in order to minimize the number of rejections from ICU admission while keeping bed occupancy at a manageable level. Although some studies suggest early discharge as a bed management tool [98], they do not include it in their models.

The need to include the discharge decision-making process in order to construct a valid simulation model was noted in [8] and [78], where the simulation model is embedded in an optimization framework to calibrate a parametric set of patient discharge rules, which attempt to mimic physician's decisions. The authors of [4] perform a sensitivity analysis of the effects of such discharge decisions on ICU performance indicators: ICU rejection rates and LoS. Their discharge decision models are implemented in a simulation framework with no time-consuming discharge processes. Operational rules for the practical implementation of the optimal discharge strategies obtained in [77] are assessed by simulation modeling in [76].

2.3. Modeling ICU LoS using phase-type distributions

A critical issue when attempting to construct a valid simulation model is the probabilistic representation of ICU LoS. An overview of LoS and patient flow modeling techniques is provided in [80]. Although the simplest models assume an exponential distribution [63,83,104], several studies have shown that LoS distributions are usually heavily skewed to the right (see, for example, [96,118]) and, accordingly, non-exponential distributions have been used: Weibull distributions in [98]; lognormal distributions in [23] and [81]; and Pearson VI in [49]. The authors of [78] propose regression models based mainly on lognormal and Weibull distributions, including variables with the capacity to explain some of the LoS variability, such as the Apache index or number of infections.

While the above-mentioned probability distributions can be used successfully to model LoS, they are not suitable for describing ICU patient health status dynamics, and thus provide inadequate support for real-world patient-discharge decision-making; thus, phase-type distributions are a better alternative. A phase-type distribution is the distribution of the time to absorption in a finite Markov chain where one state is absorbing and the remaining states are transient. Phase-type distributions can approximate any positive-valued distribution, as they are dense in the field of all positive-valued distributions. Since their introduction by Neuts [88], phase-type distributions have been used in a wide range of stochastic modeling applications, including telecommunications, finance, tele-traffic, biostatistics, queuing theory, drug kinetics, reliability theory, and survival analysis. They have also been used to model LoS in health services such as hospital wards [45,118], geriatric units [34], maternity units [55], and in capacity planning for stroke patients [82]. The author of [32] discusses the modeling of healthcare systems with phase-type distributions. Specific models with phase-type distributions, such as the Coxian phase-type distribution and the hyper-exponential distribution, are used to model LoS. The Coxian is used in [25] to model LoS in neonatal care, which includes three care levels (special care, high dependency and intensive care); in [44] it is used to predict LoS for elderly patients in hospital and community care services; and in [18] to model ICU LoS. The hyper-exponential distribution is employed in [12,48] to model ICU LoS. The authors of [33] note that the dynamic nature of hospital stays cannot be captured except by phase-type modeling. Nevertheless, these studies do not interpret phase states in terms of the patient's physical recovery.

A summary of the literature review presented in this section is included in Appendix A. Table A1 focuses on the medical literature concerning ICU bed management issues. Table A2 focuses on papers featuring ICU mathematical models, listing, in each case, the study objective, quantitative tools employed and factors considered in the discharge process modeling.

3. A simulation framework for modeling ICU processes and physicians' decision-making

3.1. A critique of classical ICU simulation models

Previous discrete-event simulation models [47,49,63,64,75,79,81,98,100,110] use simple queuing theory models (M/G/c/c or G/G/c/c) as a mathematical representation of ICU dynamics (see Fig. 1), the main characteristics of which are the client (patient) arrival process, service time (LoS) and the queue discipline (except where there is no waiting room, as in most ICUs). The research on the construction of ICU simulation models has therefore focused on analyzing and achieving a better representation of LoS and the patient arrival process. Within this queuing theory framework, an early discharge decision only makes sense in the event of a new patient arriving when all the beds are occupied. According to physicians, however, this is not the case in real practice, where it is more common to adopt the so-called cautious policy mentioned in [77], which is aimed at avoiding the rejection of new patient arrivals by allowing the rate of early discharges to increase progressively prior to full ICU occupancy. Thus, the analysis of discharge decision-making requires simulation models that will improve upon those based on G/G/c/c queuing theory, by having the capacity to reproduce the key characteristics of the real-world patient discharge process and the information environment in which decisions are made.

The following are the key aspects to be considered when modeling an ICU with the purpose of analyzing physicians' decision-making with respect to patient discharge.

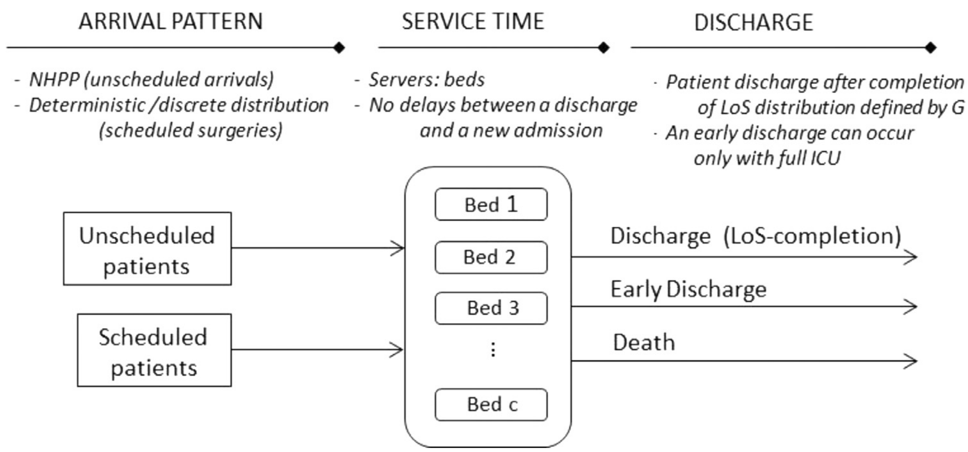

Fig. 1. G/G/c/c queuing model representation of an ICU.

- *Discharge and admission process times*

The replacement of a current ICU patient with a new arrival is a complex and non-automatic process requiring a number of conditions: a free bed in the ward to which the patient is being transferred; consent of the ward staff; and the availability of a family member to assist in the care of the patient. All these requirements are essential and cannot be improvised. At the same time, discharge can be delayed by bed shortages in other hospital departments [74]. Admission to the ICU is often delayed because it is full, possibly due to ICU beds being occupied by patients awaiting ward beds; a situation referred to as discharge delay, “bed-block” or outflow limitation [16,71]. Furthermore, the admission of a patient to a recently freed up bed also requires setup time (cleaning, disinfection, etc.).

- *Patient health status*

The ICU patient stability level (henceforth referred to as *health status*) is a holistic concept summarizing the severity of illness, the dependence of organic functions on machines (mechanical ventilation, continuous renal replacement therapy), complications such as infections, etc. Patient health status does not evolve linearly, or even monotonically, during a stay in the ICU. As noted in [26], in practice, a doctor’s choice of ICU patient for discharge depends on a stability assessment. The physician must first physically examine the patient and consult the clinical progress report to check for stability for transfer to a lower level of care. Mathematical modeling of individual patient health status would enable the simulation of physician discharge decisions based on health indicators rather than expected remaining LoS, LoS spent in the ICU, the probability of readmission, expected LoS following readmission or a combination of the above, as found in previous models [18,56,104]. A reduction of LoS will occur in the final stage of the recovery process.

- *The integration of scheduled arrivals from surgery in physicians’ discharge decisions*

ICU physicians usually have reliable advance information regarding the number of patients due to arrive from elective surgery requiring intensive postoperative care. This enables them to estimate short-term future bed requirements and plan discharges in accordance. The influence of scheduled surgeries on ICU bed occupancy levels, early patient discharge and LoS has been reported in several studies [3,62,67,92]. Using techniques ranging from simulation modeling to econometrics and survival analysis, all the cited works mention the effect of occupancy levels and elective surgeries on discharge decisions, but none attempts to model the discharge decision-making process.

- *Discharge decision times*

Patient discharge decisions typically take place at only a few scheduled times of day; as in the twice-a-day routine featured in our case study (8 a.m. and 4 p.m.). Therefore, the patient discharge decision process is essentially a periodic rather than a continuous review process, as allowed by classic queuing theory models.

The consideration of discharge decision times and discharge and admission process times prevents the instantaneous replacement of a current patient with a new incoming one. Therefore, the decision-making involved in the possible discharge of patients prior to the arrival of new ones must also be considered. This often results in a disconnection between the simulation model and the real-world environment. The analysis of discharge decisions also requires consideration of ICU bed occupancy, patient health status and scheduled arrivals from surgery, as described earlier.

The following subsection describes an ICU simulation framework to enable the analysis of real-world discharge decision-making. Table 1 summarizes the notation used in the rest of Section 3.

3.2. The new simulation framework

The proposed simulation model incorporates the common elements of all ICU simulations, such as differentiation between unscheduled and elective surgery patients, classification of patients by type of illness and other personal characteristics, and the modeling of their arrival patterns by non-homogeneous Poisson Processes for unscheduled arrivals and deterministic or discrete probabilistic patterns for arrivals from elective surgery [78]. The resource unit is the bed, including equipment (monitors, ventilators, etc.) and sufficient medical and nursing staff to care for the occupant.

In addition, however, the simulation model must be enhanced by incorporating the discharge process and the discharge decisions, together with the information these require, as discussed in the previous section.

Modeling patient health status. We propose to model the LoS of each type of patient with a phase-type distribution in order to represent the underlying dynamics of the recovery process. Different states within the distribution are associated with different health

Table 1
List of notations.

Notation	Description
A	Absorbing state of the underlying Markov chain of the Phase-type distribution
$B_k(y,\boldsymbol{X})$	Binary function to model discharge decisions for a patient in state k, when y beds are occupied and the number of planned arrivals from elective surgery are described by $\boldsymbol{X}$
$\boldsymbol{\beta}$	Row-vector of the coefficients on the logistic function and vector of decision variables in the optimization problem defined in expression (4)
$\boldsymbol{\beta}^{(i)}$	Value of decision variables in iteration i of the simulation-based optimization methodology to solve the optimization problem defined in expression (5)
c	Number of beds in the ICU
DS	Set of states of the underlying Markov chain of the Phase-type distribution in which a patient could be early discharged from the ICU
ε	Upper bound in the ε-constraint method
E[LoS]	Expected LoS
$E[LoS]^{(i)}$	Simulation estimation of E[LoS] for solution $\boldsymbol{\beta}^{(i)}$
λ	Poisson Process rate for non-scheduled ICU-arrivals
N	Number of patients not checked for early release, with health status in state $k \in$ DS
$p_k(y, \boldsymbol{X})$	Probability of discharge for a patient in state $k \in$ DS when y beds are occupied and future planned arrivals from surgery are described by $\boldsymbol{X}$
P_R	Percentage of patients rejected due to full ICU-occupancy
$P_R^{(i)}$	Simulation estimation of percentage of patients rejected for solution $\boldsymbol{\beta}^{(i)}$
r	High ICU-occupancy threshold
$S = \{1, 2, \ldots, A\}$	Set of states of the underlying Markov Chain of the Phase-type distribution
S_j	State of patient j
u	Uniform (0,1) number
$\mathbf{W} = (w_1, \ldots, w_T)$	Planned surgeries (from day 1 today T in the planning horizon) that will require patientś ICU admission
$\mathbf{X} = (x_1, \ldots, x_h)$	Information about planned surgeries for current day and several days ahead h), which physicians know and use when making discharge decisions
y	Number of occupied beds in ICU
$\mathbf{Z}$	Vector of logistic function predictors

acuity levels, to prevent the early discharge of a patient who is insufficiently recovered. Survival and exitus patients should also have different phase-type distributions. The absorbing state in survival patients represents physiological stability (henceforth referred to as full recovery) and the ideal moment for discharge from the ICU, but transfer to another hospital department is also possible prior to the absorbing state. The absorbing state for exitus patients is death, or, in some cases of irreversible terminal illness, transfer to another ward. In real practice, exitus patients are not considered for early discharge; thus, only survival patients can be discharged in states other than the absorbing state.

Let $S = \{1, 2, \ldots, A\}$ be the set of states of the underlying Markov chain, where A is the absorbing state. Let DS be the set of states in which a patient could be selected for early discharge from the ICU ($DS \subseteq S$). In the simulation model, a change of state in the Markov chain for any patient is an event which changes the health status of the respective patient and requires simulation of time spent in the new state. When faced with a patient discharge decision, the physician will know the state S_i of each patient i. If $S_i = A$, the patient will be discharged; if $S_i \in DS$, then patient i will become a candidate for early discharge; and if $S_i \in (DS \cup A)^c$, patient i is not eligible for early discharge from the ICU.

Fig. 5 represents a phase-type distribution with 5 states and the absorbing state for modeling the LoS of certain types of patients. For illustrative purposes, we allow the possibility of discharge for patients in state 5, that is, not having achieved full recovery, but able to be moved to an intermediate-care room with little risk of health consequences (that is, $DS = \{5\}$). However, the discharge of patients in states 1, 2, 3 or 4 ($(DS \cup A)^c = \{1,2,3,4\}$) is not allowed, because they may require respiratory assistance or have an infection or other clinical condition that could be exacerbated by discharge, thereby increasing the mortality risk.

Maximum likelihood estimation is the most common approach for fitting phase-type distributions [27]. Phase-type distributions can be fitted using Function phtMCMC2 of the "Phase Type" package implemented in [125], which performs Bayesian inference of the rate parameters of the latent continuous-time Markov chain, where the generator has some fixed structure.

Simulating advance notice of scheduled surgeries. Periodically, the ICU is notified of planned surgeries that will require the patient's admission. We denote this information by $\boldsymbol{W} = (w_1, \ldots, w_T)$, where w_i represents the number of planned arrivals from surgery for day i within the surgery-planning time horizon T. The value of T can differ between hospitals. As an example, in the hospital which employs one of the co-authors [8], this happens at the end of Friday mornings, when notification of the surgery schedule for the whole of the following week is given. Thus, $\boldsymbol{W} = (w_1, \ldots, w_7)$, where w_1 denotes the number of planned arrivals from surgery for Monday and w_7 those for the following Sunday. On Monday morning, physicians know the number of patients due to arrive from elective surgery for Monday through to Sunday, while, on Thursday morning, they know only those for Thursday through to Sunday. Thus, on day j, ICU physicians have notice of planned arrivals $(w_j, \ldots, w_T)$ and discharge decisions will be made based on planned arrivals from surgery on current day j and several days ahead. This information is denoted by $\boldsymbol{X} = (x_1, \ldots, x_h)$.

The arrival of this information counts as an event in the simulation model and requires simulation of arrivals from elective surgery during the next period. This information will be used to make patient discharge decisions.

Representing the decision-making process. The decision-making takes into account the number of occupied beds (denoted by y), patients' health status (denoted by states S_i for each patient $i = 1, \ldots, y$), and the number of planned admissions from elective surgery, denoted by vector $\boldsymbol{X}$. At the time of a discharge decision, all patients with health status described as "full recovery" (absorbing state A) begin their discharge process. The bed occupancy level y is updated, and all patients i with health status $S_i = k, \ \forall k \in DS$, are sequentially considered for early discharge, starting with those with the *best* health status. The discharge decision can be modeled as a binary function of y and vector $\boldsymbol{X}$, that is,

$$B_k(y, \boldsymbol{X}) = \begin{cases} 1 & if\ patient\ in\ state\ k\ is\ early\ discharged \\ 0 & otherwise \end{cases} \quad (1)$$

When the result is an early discharge, the bed occupancy level y is updated and the next patient in state $k \in DS$ is considered for early discharge. Observe that a coherent discharge policy must verify the following monotonicity relations:

$B_k(y, \boldsymbol{X}) \geq B_{k'}(y, \boldsymbol{X}), \ \ \forall$ health status $k' \in DS$ considered worse than k

$$B_k(y, \boldsymbol{X}) \geq B_k\left(y', \boldsymbol{X}\right), \ \ \forall y' < y$$

$$B_k(y, \boldsymbol{X}) \geq B_k(y, \boldsymbol{X'}), \quad \forall X \geq X' \ \text{(componentwise)}$$

The binary representation of the patient discharge decision can be shifted to a probabilistic framework. We define $p_k(y,\ \boldsymbol{X})$ as the probability of discharge for a patient in state $k \in DS$ when y beds are occupied and future admissions from elective surgery are described by $\boldsymbol{X}$. That is, $p_k(y, \boldsymbol{X}) = P\{B_k(y, \boldsymbol{X}) = 1\}$. In this case the discharge decision is implemented in the simulation model as follows:

Sequentially and in decreasing order of health status, each patient in state $k \in DS$ is subjected to a discharge/no discharge test which is performed by drawing a random number u from a uniform (0,1) distribution and comparing it with the patient's probability of discharge, $p_k(y,\ \boldsymbol{X})$:

If $u \leq p_k(y, \boldsymbol{X})$ The patient is early discharged

If $u > p_k(y, \boldsymbol{X})$ The patient remains in the ICU

As before, when the result is an early discharge, the bed occupancy level y is updated, discharge probabilities for subsequent patients are recalculated, and the test procedure is iterated with the remaining patients in state $k \in DS$. The monotonicity conditions for discharge are also applied here. Fig. 2 outlines the simulation of the discharge decisions.

Other details and the practical implementation of the simulation model. Patients arrive from elective surgery during certain time windows; one in the early afternoon for interventions performed before noon and another for those performed in the afternoon/evening. The times at which discharge decisions are made are considered as new events in the discrete event simulation model. These events trigger a decision as to how many and which patients are to be discharged.

Once discharge has been decided, the setup time for transfer to a ward begins, and the future event (effective patient discharge and the beginning of the preparation of the bed for the new arrival) is generated. Patients who die in the ICU are discharged immediately or following organ extraction in the case of donations (the event indicating the end of this process is also generated). Once the setup time (or organ extraction) is complete, the patient leaves the ICU, and the event indicating the end of the setup time to prepare the bed for a new arrival is generated. Once this occurs, the bed is ready for a new patient.

Summing up, the discrete-event simulation model has now been enhanced by including the following events: patient arrivals; discharge decision times; effective patient discharge from ICU; end of preparation time for a newly available bed; reception of the surgery schedule; and patient health status transitions. Fig. 3 outlines the simulation model.

Construction of the simulation model requires the collection of data for estimating model inputs, which are summarized in Table 2. Databases with electronic records of ICU patients include arrival and personal and medical details for each patient. A statistical analysis of these data enables determination of the best-fit probability distribution for the arrival patterns and the LoS (phase-type distributions) for each type of patient. The number of beds is obtained from ICU facility data. ICU physicians provide management information including details of the procedure, teamwork, communication with other hospital departments, and organizational factors relating to actual discharge practices, for use in modeling periodic advance notice of arrivals from elective surgery, discharge setup-times, decision times and probabilities of early discharge.

3.3. Patient discharge policies

A discharge policy is a set of rules to guide patient discharge decisions in any ICU situation (bed occupancy level, patient health status and surgery schedule). In the previous section, discharge policies were denoted by the set of binary variables $\{B_k(y,\ \boldsymbol{X})\}$ or the set $\{p_k(y,\ \boldsymbol{X})\}$. The new simulation model is valid for testing different discharge policies by measuring ICU performance by the key performance indicators (KPI), based on, for example, the patient rejection and early discharge rates [77].

By comparing the KPIs of several simulated discharge policies, ICU physicians should, in theory, be able to identify the best ICU management strategy. In reality, however, the continuity of the variables $\{p_k(y,\ \boldsymbol{X})\}$ makes this impossible, because the number of possible discharge policies is huge, or even infinite. We propose an optimization model for generating optimal discharge policies, which is solved by combining the simulation model with an optimization procedure.

The discharge probabilities should depend, as explained in Section 3.2, on patient health status $k \in DS$, the number y of occupied beds, and the number of planned arrivals from elective surgery denoted by $\boldsymbol{X}$. That is, the probability is a function of all these values: $p_k(y, \boldsymbol{X}) = f_k(y, \boldsymbol{X})$.

We propose a logistic function to link the discharge probability to its influencing factors:

$$p_k(y, \boldsymbol{X}) = \frac{1}{1 + \exp\left(-\boldsymbol{\beta}(1, \boldsymbol{Z})'\right)} \quad (2)$$

where $\boldsymbol{\beta}$ is the row-vector of the coefficients and $\boldsymbol{Z}$ is the row-vector of the predictors, built upon the bed occupancy level y and on planned arrivals from surgery for h days ahead $\boldsymbol{X} = (x_1,\ \ldots, x_h)$.

Several possible formulations fit this general logistic framework. For example, the following formulation with $Z_{in} = x_i \ \ 1_{\{y=n\}}$ includes the variable bed occupancy level, which interacts with the surgical patient arrival schedule.

$$\boldsymbol{\beta}(1, \boldsymbol{Z})' = \beta_0 + \sum_{i=1}^{h} \sum_{n=r}^{c} \beta_{in} \ \ Z_{in} = \beta_0 + \sum_{i=1}^{h} \sum_{n=r}^{c} \beta_{in} \ \ x_i \ \ 1_{\{y=n\}} \quad (3)$$

Table 2
Input data required to run the simulation model.

Simulation model inputs
Patient data
Patient type by nature of illness and other personal characteristics
Emergency patient arrival pattern
Arrival pattern for elective surgery patients
LoS for each type of patient: phase-type distributions
ICU facility data
Number of beds
ICU management data
Renewal period for advance notice of elective surgery schedule
Discharge decision times and probabilities of early discharge decisions
Discharge setup times

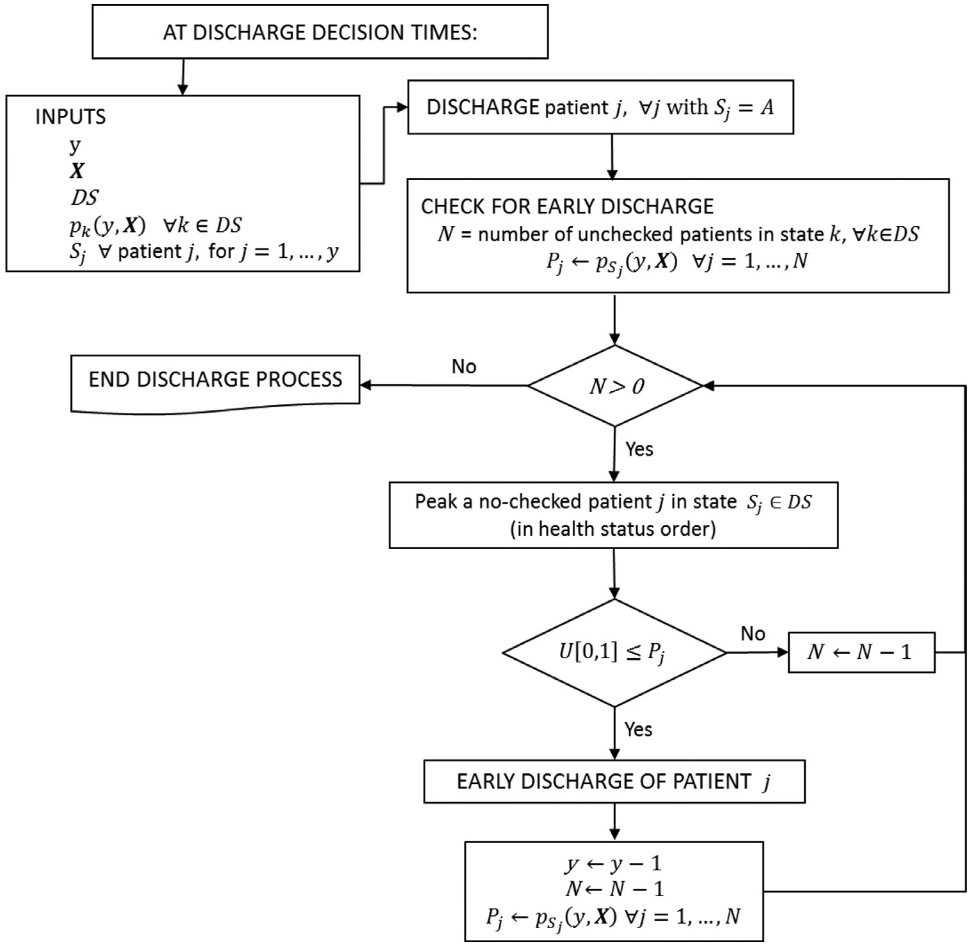

Fig. 2. Simulation of discharge decisions.

where c is the number of beds in the ICU and r is the high-occupancy threshold for possible early discharges.

The logistic link function enables assessment of the odds of shortening the length of stay of a patient according both to the ICU occupancy level and the number of planned arrivals from elective surgery.

Using this representation of the discharge probabilities, the optimization problem involves finding the best values for the parameters $\boldsymbol{\beta}$ in the proposed logistic formulation in order to optimize two conflicting objective functions: the minimization of patient rejection due to full occupancy and the minimization of LoS reduction. The first is denoted by *Min* P_R and the second is formulated as the maximization of expected LoS, *Max E[LoS]*. The decision variables are the vector $\boldsymbol{\beta} \in \mathbb{R}^n$. The optimization problem can be formulated as follows:

$$\begin{aligned} &\min_{\beta} P_R \\ &\max_{\beta} \quad E[LoS] \\ &\text{Subject to} \begin{cases} p_k(y, \boldsymbol{X}) = \frac{1}{1+\exp\left(-\boldsymbol{\beta}(1,\boldsymbol{Z})'\right)} \\ \beta \in \mathbb{R}^n \end{cases} \end{aligned} \tag{4}$$

where $\boldsymbol{\beta}(1, \boldsymbol{Z})'$ is defined in (3).

We transform this bi-objective optimization problem into a single-objective problem by using the ε-constraint method:

$$\begin{aligned} &\mathrm{P}[\varepsilon] \ \max_{\beta} \ E[LoS] \\ &\text{Subject to} \begin{cases} p_k(y, \boldsymbol{X}) = \frac{1}{1+exp\left(-\boldsymbol{\beta}(1,\boldsymbol{Z})'\right)} \\ P_R \le \varepsilon \\ \beta \in \mathbb{R}^n \end{cases} \end{aligned} \tag{5}$$

This optimization problem cannot be solved by means of non-linear optimization methods because neither the objective function *E[LoS]* nor P_R in ε-constraint can be expressed in terms of the decision variables $\boldsymbol{\beta}$. Thus, we solve this problem with simulation-based optimization methodology, in which the simulation model proposed in the above section is used as an evaluator of possible solutions to the optimization problem, as follows. The optimization procedure starts with an initial set of values for the vector of decision variables $\boldsymbol{\beta}$ (denoted by $\boldsymbol{\beta}^{(0)}$), and an initial value for the patient discharge probabilities. Every iteration i involves the following procedure. The ICU is simulated with patient discharge probabilities calculated using $\boldsymbol{\beta}^{(i)}$. The output of this simulation enables assessment of the objective function $\mathrm{E[LoS]}^{(i)}$ and verification of the upper bound constraint of $\mathrm{P_R}^{(i)} \le \varepsilon$ on the probability of patient-rejection (that is, whether $\boldsymbol{\beta}^{(i)}$ is a feasible or infeasible solution). Using this information and its own search method, the optimization procedure decides the next solution $\boldsymbol{\beta}^{(i+1)}$ to be evaluated by running the ICU-simulation model with the patient

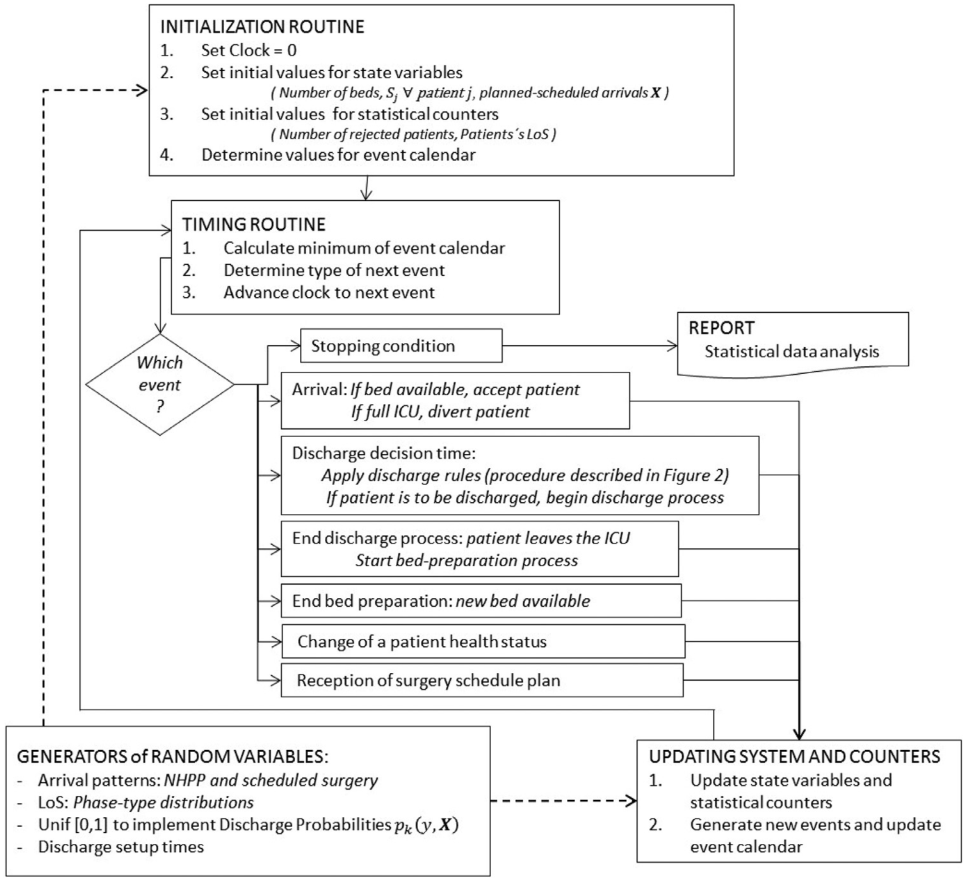

Fig. 3. ICU simulation model.

discharge probabilities calculated with $\boldsymbol{\beta}^{(i+1)}$. This process continues until the stopping conditions of the optimization method are met, thus providing the best solution, $\boldsymbol{\beta}^*$. Fig. 4 depicts the simulation-optimization procedure. Observe that the above description and Fig. 4 provide an overall view of the optimization procedure, which could be implemented with various adequately-parameterized heuristic algorithms. In the illustrative example presented in Section 4, this optimization problem is solved using a search-engine based on the scatter-search heuristic procedure [70].

4. Experimental results

This simulation-optimization framework for generating operative discharge policies is illustrated for a 20-bed ICU based on a real case analyzed by the authors in previous studies [8,78], with arrival patterns, LoS, discharge process and discharge decisions as described below.

Arrival process. Two types of patients are considered, each with a different arrival pattern: Markovian for unscheduled patients and non-Markovian for elective surgery patients. Unscheduled patient arrivals occur daily, round the clock, according to a Poisson Process with $\lambda = 1.11$ patients/day. Elective surgeries are usually performed Monday to Friday, and elective patients arrive in the late morning or early afternoon. There are no elective surgeries at weekends. The daily number of arrivals from elective surgery has the following random distribution: no arrivals with a probability of 0.2, one patient with a probability of 0.6, and two arrivals with a probability of 0.2. It is assumed that ICU physicians have reliable, 3-day advance notice of the number of patients due to arrive from elective surgery.

Discharge process. The simulation model assumes that discharge decisions are made twice a day (8 am. and 4 pm.), as in the real-world environment. Discharge setup times are also considered: Unif [2,4] distribution, in hours, representing the time required for the discharged patient's transfer process (which includes administrative paperwork, the resolution of bed-block issues, and contacting the family) and Unif [1,2] hours to prepare the bed for a new arrival.

Patient LoS. Phase-type distributions with 5 states and the absorbing state are used to model patients' LoS. The states of the phase-type distributions represent different patient health statuses, and only patients in state five are assumed sufficiently recovered to be considered for premature discharge (DS = {5}). Different transition probability matrices are used to model the LoS of each type of patient and exitus/non-exitus cases. Fig. 5 depicts the states of the phase-type distribution and transition probabilities for non-exitus unscheduled patients. The expected times in phases 1–5 are 1, 6, 2, 2 and 2 days, respectively (exponential distributions). Fig. 6 shows the transition matrices for the two types of patients and the exitus/non-exitus cases.

Discharge decisions. Discharge decisions are modeled in terms of early discharge probabilities, represented by the following linear part of the logistic function where y represents the bed occupancy level ($y = 0,1,\ldots,20$), 18 beds is the high occupancy threshold, and the number of planned elective surgery arrivals is known 3-days

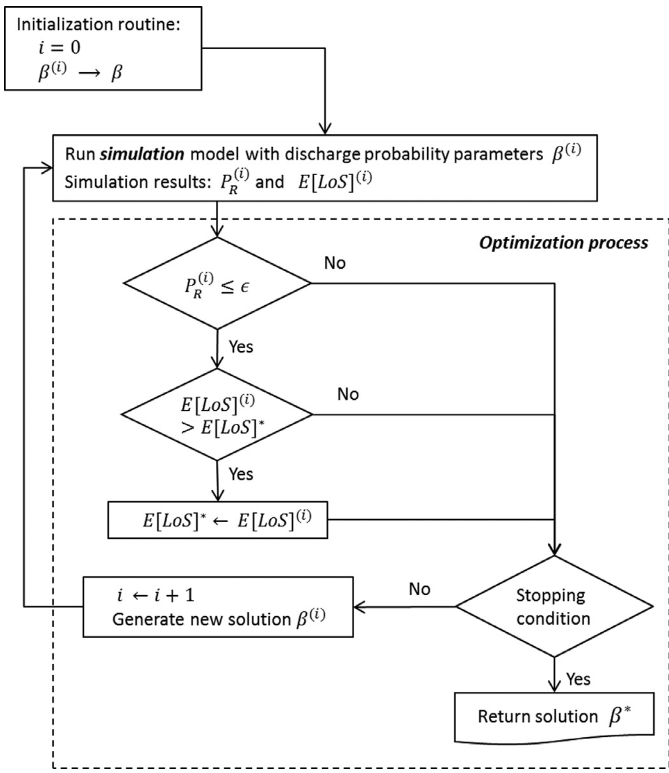

Fig. 4. Simulation-optimization procedure to generate patient discharge probabilities.

in advance.

$$\begin{aligned}\boldsymbol{\beta}(1,Z)' &= \beta_0 + \sum_{i=1}^{3}\sum_{n=18}^{20} \beta_{in} \ \ Z_{in} \\ &= \beta_0 + \beta_1 \ x_1 \ 1_{\{y=20\}} + \beta_2 \ x_1 \ 1_{\{y=19\}} + \beta_3 \ x_1 \ 1_{\{y=18\}} \\ &\quad + \beta_4 \ x_2 \ 1_{\{y=20\}} + \beta_5 \ x_2 \ 1_{\{y=19\}} + \beta_6 \ x_2 \ 1_{\{y=18\}} \\ &\quad + \beta_7 \ x_3 \ 1_{\{y=20\}} + \beta_8 \ x_3 \ 1_{\{y=19\}} + \beta_9 \ x_3 \ 1_{\{y=18\}} \end{aligned} \tag{6}$$

Optimization problem. Parameters $\boldsymbol{\beta} = (\beta_0, \ \ldots, \beta_9)$ are calculated by solving the optimization problem defined in (5), where $\boldsymbol{\beta}$ constitutes the vector of decision variables. The simulation model is implemented in ARENA software, and the optimization problem is solved with OptQuest software [70]. At every iteration, OptQuest determines one value for the parameter-vector of the discharge probabilities $\boldsymbol{\beta}$, which is sent to ARENA for assessment. ARENA returns the estimated values for the patient rejection percentage and expected LoS. With this information and the list of solutions already explored, OptQuest determines the next parameter values to be assessed by ARENA. The process continues until the stopping criteria are satisfied. The length of each simulation run is 200 years with one year as a warm-up period. Each solution is simulated independently at least three times. This simulation run length provides accurate KPI estimations.

Table 3
Optimal solutions for the $\boldsymbol{\beta}$ parameters, obtained by solving the optimization problem defined in (5), with the linear part of the logistic function defined in (6), and considering three levels of patient rejection due to a full ICU (ε = 6%, 4.5%, 3%).

Decision variables	Percentage of patient rejections		
	ε = 6%	ε = 4.5%	ε = 3%
$\boldsymbol{\beta}_0$	-17.2	-10	-10
$\boldsymbol{\beta}_1$	7.45	8.65	18.4
$\boldsymbol{\beta}_2$	4.65	6.45	18.4
$\boldsymbol{\beta}_3$	0	3.8	18.4
$\boldsymbol{\beta}_4$	7.45	6.45	18.4
$\boldsymbol{\beta}_5$	2.3	6.45	18.4
$\boldsymbol{\beta}_6$	0.25	2.65	15.55
$\boldsymbol{\beta}_7$	0.25	6.45	15.55
$\boldsymbol{\beta}_8$	0.2	0.5	12.75
$\boldsymbol{\beta}_9$	0	0.5	4.1
E[LoS]	***9.21***	***9.08***	***8.86***

Table 3 depicts the optimal solutions of the P[ε] optimization problem, considering three patient rejection levels due to a full ICU: ε = 6%, 4.5%, 3%.

Without early discharges, expected LoS and the percentage of patient rejection have the following values: $E[LoS] = 9.22$ days and P_R = 6.29%. These would be the KPIs for a discharge policy $B_A(y, \boldsymbol{X}) = p_A(y, \boldsymbol{X}) = 1$, where A is the absorbing state and $B_k(y, \boldsymbol{X}) = p_k(y, \boldsymbol{X}) = 0 \ \ \forall k \neq A$. Note that reduction of the percentage of patient rejection to 4.5% requires an average LoS reduction of 1.5% (from 9.22 to 9.08). A greater reduction in the percentage of patient rejection (to 3%) requires a greater LoS reduction (4%).

We also observe an increase in the beta parameters for the number of occupied beds. Monotonicity of the beta parameters also occurs when the patient rejection level decreases. The impact of the number of patient arrivals grows as the hori-

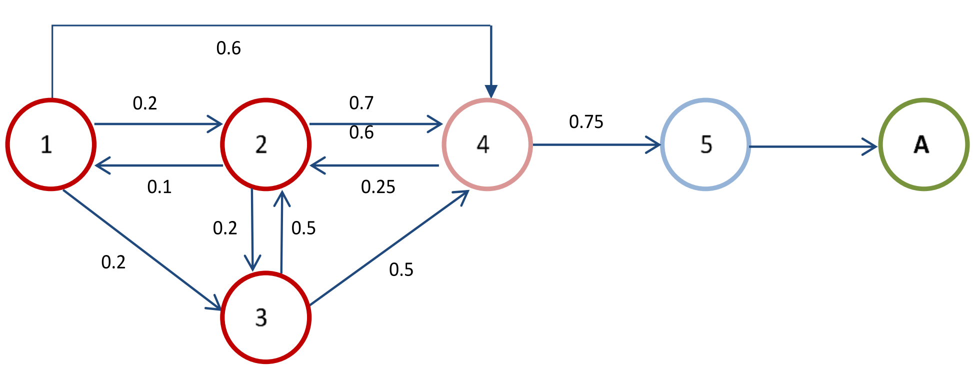

Fig. 5. Non-exitus outpatient LoS phase-type distribution: States of the phase-type distribution and transition probabilities.

$$\begin{pmatrix} 0 & 0.15 & 0 & 0.85 & 0 \\ 0 & 0 & 0.15 & 0.85 & 0 \\ 0 & 0.15 & 0 & 0.85 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Non-exitus elective patients

$m_1=1, m_2=6, m_3=2, m_4=1, m_5=1$

$$\begin{pmatrix} 0 & 0.4 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Exitus elective patients

$m_1=7, m_2=6, m_3=6$

$$\begin{pmatrix} 0 & 0.2 & 0.2 & 0.6 & 0 \\ 0.1 & 0 & 0.2 & 0.7 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0.25 & 0 & 0 & 0.75 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Non-exitus outpatients

$m_1=1, m_2=6, m_3=2, m_4=2, m_5=2$

$$\begin{pmatrix} 0 & 0.45 & 0 & 0 & 0 \\ 0.65 & 0 & 0.35 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Exitus outpatients

$m_1=1, m_2=2, m_3=1$

Fig. 6. Phase-type distributions: Transition matrices for the two patient types, exitus/non-exitus cases and expected time spent in each phase (m_i).

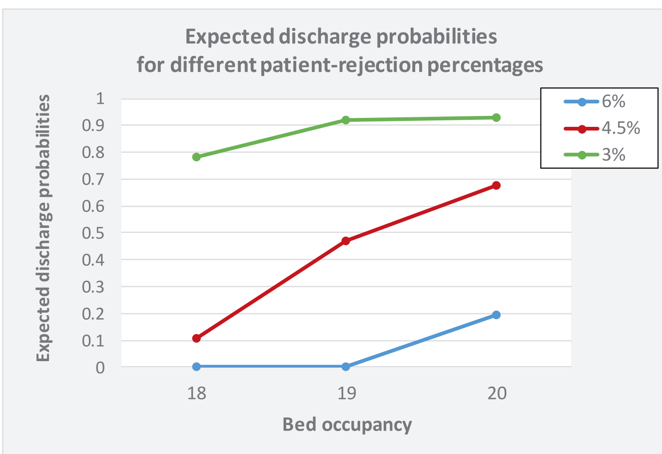

Fig. 7. Expected discharge probabilities for different bed occupancy levels (18,19,20) and percentages of patient rejection (6%, 4.5% and 3%).

zon draws nearer, as reflected in the monotonicity of the beta parameters.

The probability of early discharge is given by the following expression:

$$\frac{1}{1+exp\left(-\left(\beta_0+\beta_1 x_1\ 1_{\{y=20\}}+\beta_2\ x_1\ 1_{\{y=19\}}+\beta_3\ x_1\ 1_{\{y=18\}}+\ldots+\beta_9\ x_3\ 1_{\{y=18\}}\right)\right)} \quad (7)$$

For example, for $P_R = 4.5\%$, the probability of a patient's early discharge with 19 occupied beds and one, zero and two arrivals from elective surgery expected within the next three days is

$$p_5(19,(1,0,2)) = \frac{1}{1+exp(10-6.45\times 1-0.5\times 2)} = \frac{1}{1+\exp(2.55)} = 0.072 \quad (8)$$

Similarly, with 20 and 18 occupied beds, the early discharge probabilities are 1 and 0.005, respectively. Observe that these discharge probabilities only apply to patients who are sufficiently recovered, that is, in state 5 of the phase-type distribution shown in the illustrative example.

Table 4 shows the probabilities of discharge for a patient in state 5 of the phase-type distribution for the three levels of patient rejection, $\varepsilon = 6\%, 4.5\%, 3\%$, with 18, 19 and 20 occupied beds, and the number of arrivals from elective surgery expected within the next 3 days is x_i, $i = 1,2,3$, with $x_i \in \{0,1,2\}$, following the stochastic pattern mentioned in the description of the arrival process. Given the monotonicity of the discharge probabilities in the beta parameters, the previous assertions regarding the monotonicity of the betas are again applicable: discharge probabilities increase with the number of occupied beds; as the patient rejection rate decreases, the discharge probabilities increase; and the impact of the number of patient arrivals in the nearest time horizons is also reflected in the monotonicity of the discharge probabilities.

The cautious discharge policy prevails. The *Expected discharge probabilities* for the different occupancy levels (see Table 5 and Fig. 7) are derived from the results in Table 4 and the probability calculus of each elective planned-arrival scenario. The probability of an early patient discharge increases as more ICU beds are occupied and is increasingly ordered as referral probabilities decrease. Thus, the obtained probabilistic discharge policy matches the *cautious* policy, which is accepted by physicians as being the closest to their usual practice.

Discharge probability must not be seen by physicians as a roulette wheel, generating a random number indicating whether or not to discharge a patient (in other words, it is not a coin-tossing policy) but rather a full-occupancy risk score, which would indicate the need to reject a new patient arrival. A frequentist interpretation of probability would be more appropriate here. This point is further discussed in the Discussion and Conclusions section.

As is usual in medical studies, we use the odds ratio (OR) as a measure of the association between an influencing factor and the occurrence of an outcome of interest—the discharge of a patient in our case. The OR represents the odds that a patient's discharge will occur given exposure to a particular value of an influencing-factor, compared to the odds of it occurring given exposure to a reference value of the same influencing-factor. The odds ratio can also be used to determine whether a particular exposure is a risk factor for the discharge of a patient and to compare the magnitude of various risk factors for that patient's discharge. OR = 1 means that the exposure does not affect the odds of a patient discharge; OR > 1 means that the exposure is associated with higher odds of

Table 4
Probability of discharging a patient in state 5 of the phase-type distribution, when 18, 19 and 20 beds are occupied and the number of planned arrivals from elective surgery within the next 3 days is X_i, $i = 1,2,3$, with $X_i \in \{0, 1, 2\}$, for three levels of patient rejection.

Planned arrivals from elective surgery within the next 3 days			Percentage of patient rejection								
			6%			4.5%			3%		
			Bed-occupancy level								
X1	*X2*	*X3*	*20*	*19*	*18*	*20*	*19*	*18*	*20*	*19*	*18*
0	*0*	*0*	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0	*0*	*1*	0.000	0.000	0.000	0.028	0.000	0.000	0.996	0.940	0.003
0	*0*	*2*	0.000	0.000	0.000	0.948	0.000	0.000	1.000	1.000	0.142
0	*1*	*0*	0.000	0.000	0.000	0.028	0.028	0.001	1.000	1.000	0.996
0	*1*	*1*	0.000	0.000	0.000	0.948	0.045	0.001	1.000	1.000	1.000
0	*1*	*2*	0.000	0.000	0.000	1.000	0.072	0.002	1.000	1.000	1.000
0	*2*	*0*	0.091	0.000	0.000	0.948	0.948	0.009	1.000	1.000	1.000
0	*2*	*1*	0.114	0.000	0.000	1.000	0.968	0.015	1.000	1.000	1.000
0	*2*	*2*	0.142	0.000	0.000	1.000	0.980	0.024	1.000	1.000	1.000
1	*0*	*0*	0.000	0.000	0.000	0.206	0.028	0.002	1.000	1.000	1.000
1	*0*	*1*	0.000	0.000	0.000	0.994	0.045	0.003	1.000	1.000	1.000
1	*0*	*2*	0.000	0.000	0.000	1.000	0.072	0.005	1.000	1.000	1.000
1	*1*	*0*	0.091	0.000	0.000	0.994	0.948	0.028	1.000	1.000	1.000
1	*1*	*1*	0.114	0.000	0.000	1.000	0.968	0.045	1.000	1.000	1.000
1	*1*	*2*	0.142	0.000	0.000	1.000	0.980	0.072	1.000	1.000	1.000
1	*2*	*0*	0.994	0.000	0.000	1.000	1.000	0.289	1.000	1.000	1.000
1	*2*	*1*	0.996	0.000	0.000	1.000	1.000	0.401	1.000	1.000	1.000
1	*2*	*2*	0.996	0.001	0.000	1.000	1.000	0.525	1.000	1.000	1.000
2	*0*	*0*	0.091	0.000	0.000	0.999	0.948	0.083	1.000	1.000	1.000
2	*0*	*1*	0.114	0.000	0.000	1.000	0.968	0.130	1.000	1.000	1.000
2	*0*	*2*	0.142	0.001	0.000	1.000	0.980	0.198	1.000	1.000	1.000
2	*1*	*0*	0.994	0.004	0.000	1.000	1.000	0.562	1.000	1.000	1.000
2	*1*	*1*	0.996	0.004	0.000	1.000	1.000	0.679	1.000	1.000	1.000
2	*1*	*2*	0.996	0.005	0.000	1.000	1.000	0.777	1.000	1.000	1.000
2	*2*	*0*	1.000	0.036	0.000	1.000	1.000	0.948	1.000	1.000	1.000
2	*2*	*1*	1.000	0.043	0.000	1.000	1.000	0.968	1.000	1.000	1.000
2	*2*	*2*	1.000	0.052	0.000	1.000	1.000	0.980	1.000	1.000	1.000

Table 5
The cautious discharge policy prevails: Calculus of the expected discharge probabilities with 18, 19 and 20 occupied beds, for three levels of patient rejection (6%, 4.5% and 3%).

Probabilities of each planned- arrival scenario	Planned arrivals from elective surgery within the next 3 days			Percentage of patient rejection								
				6%			4.5%			3%		
				Bed-occupancy level								
	X1	X2	X3	20	19	18	20	19	18	20	19	18
0.0720	*0*	*0*	*0*	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.1131	*0*	*0*	*1*	0.000	0.000	0.000	0.028	0.000	0.000	0.996	0.940	0.003
…	…	…	…	…	…	…	…	…	…	…	…	…
0.0274	*0*	*2*	*1*	0.114	0.000	0.000	1.000	0.968	0.015	1.000	1.000	1.000
0.0091	*0*	*2*	*2*	0.142	0.000	0.000	1.000	0.980	0.024	1.000	1.000	1.000
0.1131	*1*	*0*	*0*	0.000	0.000	0.000	0.206	0.028	0.002	1.000	1.000	1.000
0.0309	*1*	*0*	*1*	0.000	0.000	0.000	0.994	0.045	0.003	1.000	1.000	1.000
…	…	…	…	…	…	…	…	…	…	…	…	…
0.0309	*1*	*2*	*1*	0.996	0.000	0.000	1.000	1.000	0.401	1.000	1.000	1.000
0.0103	*1*	*2*	*2*	0.996	0.001	0.000	1.000	1.000	0.525	1.000	1.000	1.000
0.0091	*2*	*0*	*0*	0.091	0.000	0.000	0.999	0.948	0.083	1.000	1.000	1.000
0.0389	*2*	*0*	*1*	0.114	0.000	0.000	1.000	0.968	0.130	1.000	1.000	1.000
…	…	…	…	…	…	…	…	…	…	…	…	…
0.0103	*2*	*2*	*1*	1.000	0.043	0.000	1.000	1.000	0.968	1.000	1.000	1.000
0.0034	*2*	*2*	*2*	1.000	0.052	0.000	1.000	1.000	0.980	1.000	1.000	1.000
Expected discharge probabilities				***0.1933***	***0.0013***	***0.0000***	***0.6771***	***0.4693***	***0.1086***	***0.9275***	***0.9212***	***0.7826***

a patient discharge; and OR $<$ 1 means that exposure is associated with lower odds of a patient discharge.

For example, in the case of bed occupancy level $y = 18$,

$$\frac{P(\text{Discharge}|Y = 18, (n, x_2, x_3))/P(\text{No Discharge}|Y = 18, (n, x_2, x_3))}{P(\text{Discharge}|Y = 18, (n-1, x_2, x_3))/P(\text{No Discharge}|Y = 18, (n-1, x_2, x_3))} = \exp(\beta_3)$$

Thus, with patient rejection at 4.5% and bed occupancy at 18, one new arrival within the next 24 h multiplies the odds of a patient discharge by 44.7. These odds drop to 1 in the 6% patient-rejection scenario.

When bed occupancy increases from 18 to 19 and (x_1,0,0) scheduled arrivals are planned within the following few days, the estimated OR is

$$\frac{P(\text{Discharge}|Y=19, (x_1, 0, 0))/P(\text{No Discharge}|Y=19, (x_1, 0, 0))}{P(\text{Discharge}|Y=18, (x_1, 0, 0))/P(\text{No Discharge}|Y=18, (x_1, 0, 0))} = \exp(x_1(\beta_2 - \beta_3))$$

Then, with patient rejection at 4.5%, a change in ICU bed occupancy from 18 to 19 multiplies the odds of a patient discharge by 14.15, assuming (1,0,0) scheduled arrivals for the next three days. The estimated OR values increase to 104.6 when (2,0,0) scheduled arrivals are expected. In the 6% patient-rejection scenario, the exposure does not affect the odds.

5. Discussion and conclusions

This paper offers a thorough review of the medical and the mathematical literature with particular attention to physicians' ICU bed management practices. The medical literature highlights the existing ICU saturation problem and its patient health implications, and describes the triage of current ICU patients. Early patient discharge has become an ICU management tool with physicians trying to strike a balance between the rate of patient rejection due to a full ICU and the degree of LoS reduction for patients already admitted. The medical literature also describes the patient discharge process, which involves physicians from other units and the patient's family and can take several hours. This description is important because it prevents consideration of the substitution of one patient with another in an ICU bed from becoming an instantaneous process. However, most of the mathematical models found in the literature treat such an exchange of patients as an instantaneous event. While focusing on obtaining an accurate representation of the stochasticity of the patient arrival process and patient LoS, these models neglect the entire discharge process. This is no minor issue because it implies a poor representation of the discharge decisions made prior to the arrival of new patients in need of the beds made available by discharging others. None of the queuing theory models or other stochastic models which appear in the mathematical literature includes this anticipation. Thus, the management policies examined in mathematical models are aggressive; since they consider only the triage of the last bed and disregard the pressure on physicians working close to full capacity. Any model of ICU physician discharge decisions must include the bed requirements of planned arrivals from elective surgery, which make it necessary for early discharge decisions to be made in advance of full occupancy.

The simulation model presented in this paper bridges the gap between the real-world patient discharge process and how it is modeled in the mathematical literature. The new framework presented here not only reflects the discharge process as it actually occurs but also takes into account that not every patient in the ICU is eligible for early discharge. In the real world, patient discharge is based exclusively on clinical criteria and the possibility of discharge is considered only when the patient is considered stable and under no health risk if transferred to a lower level of care. When mathematically modeling the underlying dynamics of the recovery process, LoS is represented by a phase-type distribution in which patient health status is indicated by the different states, some of which contraindicate patient discharge. This type of model is closer to clinical reality since individual discharge decisions are based on health indicators rather than other criteria, such as the expected remaining LoS or the probability of readmission, which do not determine whether the patient is stable enough for transfer to a unit with lower staffing.

This new simulation framework enables a more accurate analysis of discharge policies by testing them in a more realistic environment. It also serves to obtain optimal discharge policies (or rather, efficient ones, given that there are two conflicting objectives), by parameterizing them as a function of the number of occupied beds and the number of planned arrivals from surgery.

In this study, physicians are assumed to have reliable advance notice of the number of patients due to arrive from elective surgery. Occasionally, however, some such patients may not ultimately require an ICU-bed for one of several reasons (for example, a change in the patient's health status —nosocomial infection, peri–operative myocardial infarction or even death—, technical or operational problems in the operating theater, a shortage of theater time, a change in the surgery schedule or the cancelation of elective surgeries to accommodate emergency surgery) [2]. By way of example, in the hospital which employs one of the co-authors, cancelations are usually filled by drawing patients with similar needs from the waiting list and the percentage of cancelations of planned surgeries within a 24-h period is less than 1%. The probability of cancelation could be easily incorporated into the simulation model. Sensitivity analysis could be used to assess how the optimal discharge policy is influenced by the fact that cancelations create uncertainty about the number of surgical arrivals.

The discharge policies indicate the patient discharge probabilities in any given ICU situation and for any ICU patient health status required to meet a certain target rejection rate and minimize LoS reduction. These probabilities are interpreted by physicians as pressure to discharge patients early in order to avoid saturation. The greater the probability, the greater the need for an early discharge in order to comply with the target rejection rate.

A frequentist interpretation of probability would say that, in current conditions, physicians should resort to early discharge at a rate equal to that probability. Thus, a discharge probability score of 0.4 would mean that early discharge should take place on 40% of the occasions that current conditions prevail.

The interpretation of the discharge probability score as "pressure-to-discharge" brings the optimal discharge policy into line with the cautious policy used by physicians in practice. Therefore, this result validates our proposed simulation framework, which reproduces real-world ICU patient flow, tested on real input data. This simulation framework generates optimal strategies matching those considered best and commonly used by physicians, which outperform the optimal strategies generated by queuing-based simulation models.

As mentioned in the introduction, the dilemma of the last bed extends to all beds when doctors are under pressure due to ICU saturation.

Acknowledgments

This paper was supported by grant MTM2016-77015-R (AEI, FEDER). The authors wish to acknowledge the useful comments received from the editor and two anonymous referees.

Appendix A.

Tables A1 and A2.

Table A1
Medical literature concerning issues relating to ICU physicians' bed management decisions.

MEDICAL LITERATURE REVIEW							
Publication	Period	Country	Type of ICUs	Research objective	Research design	Conclusion	Category
Baker et al. [7]	2006–2007	USA	22-bed adult neurosciences ICU	To determine whether high ICU inflow volumes are associated with unplanned readmissions	Retrospective comparative analysis. Statistical analysis	Days with high ICU inflow volumes were significantly associated with subsequent unplanned readmissions	Health risk-Readmission
Beck et al. [10]	1996–2000	Germany	9-bed general ICU	To assess the effects of discharge TISS scores, discharge time and type of discharge facility on ultimate hospital mortality after intensive care	Retrospective cohort study. Statistical analysis	Premature ICU discharge was associated with increased mortality. Intermediate care reduced mortality in patients prematurely discharged	Health risk- Hospital mortality
Chen et al. [18]	1995–1996	Canada	3 ICUs of two teaching hospitals and 4 ICUs of 4 community hospitals	To determine the clinical features and outcomes of patients readmitted to the ICU during the same hospital stay and the causes for these readmissions	Retrospective, multicenter, cohort study. Statistical analysis	Readmitted patients have a high risk of hospital death	Health risk- Hospital mortality
Chrusch et al. [20]	1989–1996	Canada	Tertiary care teaching hospital	To determine whether a bed shortage was leading to premature patient discharge and subsequent early readmission or death	Prospective cohort study. Statistical analysis	Increased patient occupancy is associated with an increased risk of early death or readmission post ICU-discharge. Overloading ICU capacity may affect physician decision-making, resulting in premature discharge	Health risk- Readmission and hospital mortality
Daly et al. [24]	1989–1998	United Kingdom	13-bed ICU at Guy's hospital and 19 UK ICUs (Riyadh ICU program users group, RIPUG)	To reduce mortality after discharge	Multiple-center, prospective cohort study. Statistical analysis	The discharge mortality of at risk patients can be reduced by 39% if they remain in ICU for a further 48 h. The discharge triage model for identifying patients at risk from too early discharge may help doctors to make the difficult clinical decision of whom to discharge to make room for a patient requiring urgent admission	Health risk
Durbin et al. [29]	18-month period (1993)	USA	8-bed medical and 16-bed surgical ICU in a 650-bed university hospital	To determine the characteristics of patients requiring readmission to an ICU	Retrospective, case-control chart review	Readmission to an ICU carries a high risk of mortality and increased length of stay and may represent premature discharge in at least 30% of patients	Health risk- Hospital Mortality and length of stay
Elliot et al. [30]	2007	Australia	12-bed general ICU in a 500-bed tertiary referral hospital	To identify and describe the experiences and perceptions of nurses regarding the contributing factors of ICU readmissions		Early discharge of clinically unstable patients creates issues around workload and challenges ward staff. It also increases the likelihood of patients being readmitted	Health risk-Readmission

(*continued on next page*)

Table A1 (*continued*)

MEDICAL LITERATURE REVIEW							
Publication	Period	Country	Type of ICUs	Research objective	Research design	Conclusion	Category
Franklin et al. [37]	1979–1980, (1year)	USA	Medical ICU	To identify those patients most likely to be readmitted to a Medical ICU	Retrospective cohort study. Statistical analysis	To delineate diseases, medications, and complications which may predict the high-risk discharge from MICU	Define better discharge strategies
Frost et al. [39]	1997–2007	Australia	24-bed ICU	To develop a prediction model using an inception cohort of patients surviving an initial ICU stay to determine the risk of readmission to the ICU during the same hospital stay	Statistical analysis	Discharge after-hours was associated with a higher risk of readmission to the ICU	Health risk-Readmission
Gantner et al. [40].	2005–2012	Australia	Data from the Australian and New Zealand Intensive Care Society Adult Patient Database	To examine trends over time in discharge timing and contemporary associations with mortality and readmission	Multiple-center, retrospective cohort study Statistical analysis	After-hours discharge remains an important independent predictor of hospital mortality and readmission	Health risk- Readmission and hospital mortality
Goldfrad et al. [42]	1988–90 and 1995–98	United Kingdom	26 ICUs in the first period and 62 ICUS in the second period	To determine discharge at night as a proxy measure to investigate pressure	Multiple-center, prospective cohort study. Statistical analysis	Night discharges are increasing in the UK	Premature discharge in lack of bed
Goldhill et al. [43]	1992–1996	United Kingdom	24 ICUs	To identify priorities for ICU intervention and research	Multiple-center, retrospective cohort study. Statistical analysis	Many patients die after discharge from ICU and this mortality may be decreased by minimizing inappropriate early discharge to the ward	Premature discharge in lack of bed
Kim et al. [66]	2013–14 (15 months)	USA	36-bed medical and 21-bed surgical ICU at a teaching hospital	To study whether workload has an impact on a direct measure of the health status of discharged patients	Multiple-center, retrospective cohort study. Econometric model	More acutely ill patients are discharged when ICU occupancy levels are high	Health risk
Kramer et al. [68]	2002–2010	USA	One hundred five ICUs at 46 hospitals	To examine the association between ICU readmission rates and case-mix-adjusted outcomes	Multiple-center, retrospective cohort study. Statistical analysis	Patients readmitted to ICUs have increased hospital mortality and lengths of stay	Health risk- Readmission and hospital mortality
Moreno et al. [86]	1997–1998	Spain	Database of the EURICUS-II study, 44 ICUs, 10 European countries and 4621 patients	To examine whether post-ICU discharge mortality is associated with the presence and severity of organ dysfunction/failure just before ICU discharge	Multiple-center, prospective cohort study. Statistical analysis	It is better to delay the discharge of a patient with organ dysfunction/failure from the ICU, unless adequate monitoring and therapeutic resources are available in the ward	Health risk
Nates et al. [87]	2016			To update the Society of Critical Care Medicine's guidelines for ICU admission, discharge, and triage, providing a framework for clinical practice, the development of institutional policies, and further research	Revision of the literature to develop these guidelines	These recommendations provide a comprehensive framework to guide practitioners in making informed decisions during the admission, discharge, and triage process as well as in resolving issues of no beneficial treatment and rationing	Define better discharge strategies

(*continued on next page*)

Table A1 (*continued*)

MEDICAL LITERATURE REVIEW

Publication	Period	Country	Type of ICUs	Research objective	Research design	Conclusion	Category
Ouanes et al. [90]	1998–2008	French	4 ICUs	To identify independent risk factors for early post-ICU readmission or death and to construct a prediction model	Multiple-center, retrospective cohort study. Statistical analysis	Independent risk factors were indicators of patients' severity and discharge at night (a marker of bed shortage)	Health risk- Readmission and hospital mortality
Priestap et al. [93]	2001–2004	Canada	31 ICUs of community and teaching hospitals	To determine the impact of night-time discharge on patient outcome	Multiple-center, retrospective cohort study. Statistical analysis	Patients discharged at night have a higher risk of mortality than those discharged during the day	Health risk- Hospital mortality
Rodríguez-Carvajal et al. [101]	2000–2005	Spain	A 10-bed general ICU in a community hospital	To determine the frequency and to evaluate the relationship between premature discharge and post-ICU hospital mortality	Retrospective cohort study. Statistical analysis	Premature discharges appear to be common in our setting and have a significant impact on mortality	Health risk- Hospital mortality
Rosenberg et al. [102]	2000	USA		To evaluate the causes, risk factors, and mortality rates associated with unexpected ICU-readmissions	Review article	ICU readmission is associated with dramatically higher hospital mortality	Health risk- Hospital mortality
Singh et al. [106]	2004–2006	Australia	ICU in a tertiary care teaching hospital	To assess the frequency of after-hours discharge of patients and its effect on in-hospital mortality	Retrospective cohort study. Statistical analysis	Discharge after-hours was associated with a higher risk of in-hospital mortality than discharge during work hours	Health risk- Hospital mortality
Sprung et al. [109]	2013	Europe		To provide an updated consensus statement on the principles and recommendations for patient triage	Review article	Consensus was reached for most general and specific ICU triage principles and recommendations	Discharge strategies
Tobin et al. [114]	1992–2002	Australia	A 16-bed ICU in a 400-bed tertiary referral hospital	To determine the impact of time of discharge on subsequent hospital mortality	Retrospective cohort study. Statistical analysis	More patients are being discharged in the afternoon and night suggesting increasing pressure on ICU beds. Patients discharged on these shifts have a higher mortality risk	Premature discharge in lack of bed
Yoon et al. [122]	2000–2002	Korea	34 beds of two medical-surgical units	To evaluate the effect of intensivists' discharge decision-making on readmission to ICU	Prospective and retrospective cohort study. Statistical analysis	The readmission rate was lower when intensivists participated in the discharge decision-making	Discharge strategies

Table A2
ICU-modeling literature: Study objectives, quantitative tools used and elements included in the modeling of the discharge process.

Publication	Type of literature	Objective	Quantitative tools	Elements considered in modeling the discharge process					
				Early discharge	Occupancy level	Health status	Discharge decision times	Discharge process times	Planned scheduled arrivals
Anderson et al. [3]	Math	How occupation affects discharge rate and LoS	Survival analysis	Y	Y	N	N	N	Y
Azcárate et al. [4]	Math	Calibrate the simulation model	Optimization Simulation	Y	Y	N	Y	N	N
Barado et al. [8]	Medical	Capacity planning	Simulation	Y	N	N	Y	N	N
Bowers [12]	Math	Capacity planning	Simulation	N	N	N	N	N	N
Chan et al. [18]	Math	Discharge policy that minimizes cost associated with early discharge	Dynamic optimization	Y*	Y	Y	N	N	N
Costa et al. [23]	Medical	Capacity planning	Simulation	N	N	N	N	N	N
Dobson et al. [26]	Math	To predict performance (in terms of bumping) under different arrival pattern and capacity	Markov chain	Y*	Y	Y*	Y	N	N
Griffiths et al. [47]	Math	Bed-management	Simulation	N	Y	N	N	N	N
Griffiths et al. [48]	Math	Bed-management	Queuing theory	N	N	N	N	N	N
Griffiths et al. [49]	Math	Minimize nursing staff cost	Simulation	N	N	N	N	N	N
Hosseinifardet al. [56]	Math	Minimize ICU-load resulting from readmission of early-discharge patients	Stochastic dynamic problem solved by optimization–simulation	Y*	N	Y*	N	N	N
Kc et al. [62].	Math	Estimate the impact of occupancy on LoS and readmission	Statistical analysis	Y*	N	N	N	N	N
Kim et al. [63]	Math	Capacity planning	Queuing theory Simulation	N	N	N	N	N	N
Kim et al. [64]	Math	Minimize the number of canceled surgeries by means of bed-reservation schemes	Simulation	N	N	N	N	N	N
Kolker [67]	Medical	Maximize the number of elective surgeries scheduled per day in order to reduce ICU-diversion	Simulation	N	N	N	N	N	Y
Li et al. [72]	Math	To determine the best policy to allocate beds to different classes of patients by reducing premature discharge costs	Dynamic Programming	Y*	Y	Y	N	Y	N

(*continued on next page*)

Table A2 (*continued*)

Publication	Type of literature	Objective	Quantitative tools	Elements considered in modeling the discharge process					
				Early discharge	Occupancy level	Health status	Discharge decision times	Discharge process times	Planned scheduled arrivals
Litvack et al. [74]	Math	Capacity planning	Queuing theory	N	N	N	N	N	N
Mallor et al. [76]	Math	Minimize patient rejection and LoS reduction	Queuing theory Optimization	Y	Y	Y	N	N	N
Mallor et al. [77]	Math	Minimize patient rejection and LoS reduction	Queuing theory Optimization-based simulation	Y	Y	Y	N	N	N
Mallor et al. [78]	Math	Capacity planning	Simulation	Y	Y	N	N	Y	N
Marmor et al. [79]	Math	Capacity planning	Simulation	N	N	N	N	N	N
Masterson et al. [81]	Math	Capacity planning	Simulation	N	N	N	N	N	N
McManus et al. [83]	Medical	Capacity planning	Queuing theory	N	N	N	N	N	N
Nguyem et al. [89]	Medical	Capacity planning	Simulation	N	N	N	N	N	N
Ridge et al. [98]	Math	Capacity planning	Simulation	N	N	N	N	N	N
Pearson et al. [91]	Medical	Capacity planning	Simulation	N	N	N	N	N	N
Rodrigues et al. [100]	Math	Capacity planning	Simulation	N	N	N	N	N	N
Shahani et al. [103]	Medical	Capacity planning	Simulation	N	N	N	N	N	N
Shmueli et al. [104]	Math	Determine the best admission policy to maximize the expected survival benefit	Queuing theory	N	N	N	N	N	N
Steins et al. [110]	Medical	Capacity planning	Simulation	N	Y	N	N	N	Y
Troy et al. [115]	Medical	Capacity planning	Simulation	N	N	N	N	N	Y
Yang et al. [121]	Medical	Determine the best elective-admission policy to minimize surgery cancelations	Simulation	N	Y	N	N	N	N
Zhuet al. [124]	Medical	Capacity planning	Simulation	N	Y	N	N	N	N

Early discharge: *Y*:* only if full ICU.
Health status: *Y*:* through probability properties of LoS and/or readmission probability.

References

[1] Ahmed AH, Giri J, Kashyap R, Singh B, Dong Y, Kilickaya O, et al. Outcome of adverse events and medical errors in the intensive care unit: a systematic review and meta-analysis. Am J Med Qual 2013;20:1–8.

[2] Al Talalwah N, Mclltrot KH. Cancellation of surgeries: integrative review. J Perianesth Nurs 2019;34(1):86–96.

[3] Anderson D, Price C, Golden B, Jank G, Wasil E. Examining the discharge practices of surgeons at a large medical center. Health Care Manag Sci 2011;14:338–47.

[4] Azcárate C, Mallor F, Barado J. Calibration of a decision-making process in a simulation model by a bicriteria optimization problem. In: Proceedings of the 2012 winter simulation conference; 2012. p. 782–91.

[5] Azoulay EL, Timsit JF, Sprung CL, Soares M, Rusinova K, Lafabrie A, et al. Conflicus study investigators and for the ethics section of the European society of intensive care medicine. Prevalence and factors of intensive care unit conflicts: the conflicus study. Am J Respir Crit Care Med 2009;180:853–60.

[6] Bai J, Fügener A, Schoenfelder J, Brunner JO. Operations research in intensive care unit management: a literature review. Health Care Manag Sci 2018;21:1–24.

[7] Baker DR, Pronovost PJ, Morlock LL, Geocadin RG, Holzmueller CG. Patient flow variability and unplanned readmissions to an intensive care unit. Crit Care Med 2009;37:2882–7.

[8] Barado J, Guergue JM, Esparza L, Azcarate C, Mallor F, Ochoa S. A mathematical model for simulating daily bed occupancy in an intensive care unit. Crit Care Med 2012;40:1098–104.

[9] Barnato AE, Kahn JM, Rubenfeld GD, McCauley K, Fontaine D, Frassica JJ, et al. Prioritizing the organization and management of intensive care services in the United States: the promis conference. Crit Care Med 2007;35: 1003–1011.

[10] Beck DH, McQuillan P, Smith GB. Waiting for the break of dawn? The effects of discharge time, discharge TISS scores and discharge facility on hospital mortality after intensive care. Intensive Care Med 2002;28:1287–93.

[11] Bing-Hua YU. Delayed admission to intensive care unit for critically surgical patients is associated with increased mortality. Am J Surg 2014;208:268–74.

[12] Bowers J. Balancing operating theatre and bed capacity in a cardiothoracic centre. Health Care Manag Sci 2013;16:236–44.

[13] Brailsford S, Harper PR, Patel B, Pidd M. An analysis of the academic literature on simulation and modelling in health care. J Simul 2009;3:130–40.

[14] Brailsford S, Vissers J. OR in healthcare: a European perspective. Eur J Oper Res 2011;212:223–34.

[15] Cardoso LT, Grion CM, Matsuo T, Anami EH, Kauss IA, Seko L, Bonametti AM. Impact of delayed admission to intensive care units on mortality of critically ill patients: a cohort study. Crit Care 2011;15:R28.

[16] Chaboyer W, Thalib L, Foster M, Elliott D, Endacott R, Richards B. The impact of an ICU liaison nurse on discharge delay in patients after prolonged ICU stay. Anaesth Intensive Care 2006;34:55–60.

[17] Chalfin DB, Trzeciak S, Likourezos A, Baumann BM, Dellinger RPThe DELAY-ED study Group. Impact of delayed transfer of critically ill patients from the emergency department to the intensive care unit. Crit Care Med 2007;35:1477–83.

[18] Chan CW, Farias VF, Bambos N, Escobar G. Optimizing intensive care unit discharge decisions with patient readmissions. Oper Res 2012;60:1323–41.

[19] Chen LM, Martin CM, Keenan SP, Sibbald WJ. Patients readmitted to the intensive care unit during the same hospitalization: clinical features and outcomes. Crit Care Med 1998;26:1834–41.

[20] Chrusch CA, Olafson KP, McMillan PM, Roberts DE, Gray PR. High occupancy increases the risk of early death or readmission after transfer from intensive care. Crit Care Med 2009;37:2753–8.

[21] Colmenero M. El ritual de la falta de camas. Med Intensiva 2011;35:139–42.

[22] Combes A, Luyt CE, Trouillet JL, Chastre J, Gibert C. Adverse effect on a referral intensive care unit's performance of accepting patients transferred from another intensive care unit. Crit Care Med 2005;33:705–10.

[23] Costa AX, Ridley SA, Shahani AK, Harper PR, De Senna V, Nielsen MS. Mathematical modelling and simulation for planning critical care capacity. Anaesthesia 2003;58:320–7.

[24] Daly K, Beale R, Chang RW. Reduction in mortality after inappropriate early discharge from intensive care unit: logistic regression triage model. BMJ 2001;322:1274–6.

[25] Demir E, Lebcir R, Adeyemi S. Modelling length of stay and patient flows: methodological case studies from the UK neonatal care services. J Oper Res Soc 2014;65:532–45.

[26] Dobson G, Lee HH, Pinker E. A model of ICU bumping. Oper Res 2010;58:1564–76.

[27] Donnelly C, McFetridge LM, Marshall AH, Mitchell HJ. A two-stage approach to the joint analysis of longitudinal and survival data utilising the Coxian phase-type distribution. Stat Methods Med Res 2018;27(12):3577–94.

[28] Duke GJ, Buist MD, Pilcher D, Scheinkestel CD, Santamaria JD, Gutteridge GA, Cranswick PJ, Ernest D, French C, Botha JA. Interventions to circumvent intensive care access block: a retrospective 2-year study across metropolitan Melbourne. Med J Aust 2009;190:375–8.

[29] Durbin CG, Kopel RF. A case-control study of patients readmitted to the intensive care unit. Crit Care Med 1993;21:1547–53.

[30] Elliott M, Crookes P, Worrall-Carter L, Page K. Readmission to intensive care: a qualitative analysis of nurses' perceptions and experiences Heart and Lung. J Acute Crit Care 2011;40:99–309.

[31] Embriaco N, Papazian L, Kentish-Barnes N, Pochard F, Azoulay E. Burnout syndrome among critical care healthcare workers. Curr Opin Crit Care 2007;13:482–8.

[32] Fackrell M. Modelling healthcare systems with phase-type distributions. Health Care Manag Sci 2009;12:11–26.

[33] Faddy M, Graves N, Pettitt A. Modeling length of stay in hospital and other right skewed data: comparison of phase-type, gamma and log-normal distributions. Value Health 2009;12:309–14.

[34] Faddy MJ, McClean SI. Using a multi-state model to enhance understanding of geriatric patient care. Aust Health Rev 2007;31:91–7.

[35] Fetter RB, Thompson JD. The simulation of hospital systems. Oper Res 1965;13:689–711.

[36] Fone D, Hollinghurst S, Temple M, Round A, Weightman, Roberts K, et al. Systematic review of the use and value of computer simulation modelling in population health and healthcare delivery. J Public Health Med 2003;25:325–35.

[37] Franklin C, Jackson D. Discharge decision-making in a medical ICU: characteristics of unexpected readmissions. Crit Care Med 1983;11:61–6.

[38] Fried MJ, Bruce J, Colquhoun R, Smith G. Inter-hospital transfers of acutely ill adults in Scotland. Anaesthesia 2010;65:136–44.

[39] Frost S, Tam V, Alexandrou E, Hunt L, Salamonson Y, Davidson P, Parr MJ, Hillman KM. Readmission to intensive care: development of a nomogram for individualising risk. Crit Care Resuscitation 2010;12:83–9.

[40] Gantner D, Farley K, Bailey M, Huckson S, Hicks P, Pilcher D. Mortality related to after-hours discharge from intensive care in Australia and New Zealand, 2005–2012. Intensive Care Med 2014;40:1528–35.

[41] Geitona M, Androutsou L, Theodoratou D. Cost estimation of patients admitted to the intensive care unit: a case study of the Teaching University Hospital of Thessaly. J Med Econ 2010;13:179–84.

[42] Goldfrad C, Rowan K. Consequences of discharges from intensive care at night. Lancet 2000;355:1138–42.

[43] Goldhill DR, Summer A. Outcome of intensive care patients in a group of British intensive care units. Crit Care Med 1998;26:1337–45.

[44] Gordon AS, Marshal A, Zenga M. To predict elderly patient LoS in hospital and community care services. Health Care Manag Sci 2018;21:269–80.

[45] Gorunescu F, MacClean SI, Millard PH. A quequeing model for bed-occupancy management and planning of hospitals. J Oper Res Soc 1990;53:19–24.

[46] Green LV. How many hospital beds? Inquiry 2003;39:400–12.

[47] Griffiths JD, Knight V, Komenda I. Bed management in a Critical Care Unit. IMA J Manag Math 2013;24:137–53.

[48] Griffiths JD, Price-Lloyd N, Smithies M, Williams J. A queueing model of activities in an intensive care unit. IMA J Manag Math 2006;17:277–88.

[49] Griffiths JD, Price-Lloyd N, Smithies M, Williams JE. Modelling the requirement for supplementary nurses in an intensive care unit. J Oper Res Soc 2005;56:126–33.

[50] Günal MM, Pidd M. Discrete event simulation for the performance modelling in health care: a review of the literature. J Simul 2010;4:42–51.

[51] Hall RW. Patient flow: reducing delays in healthcare delivery. International series in operations research & management science. Springer; 2006.

[52] Halpern NA, Pastores SM. Critical care medicine in the United States 2000–2005: an analysis of bed numbers, occupancy rates, payer mix, and costs. Crit Care Med 2010;38:65–71.

[53] Halpern NA, Pastores SM, Oropello JM, Kvetan V. Critical care medicine in the United States: addressing the intensivist shortage and image of the specialty. Crit Care Med 2013;41:2754–61.

[54] Halpern NA, Pastores SM. Critical care medicine beds, use, occupancy, and cost in the United States: a methodological review. Crit Care Med 2015;43:2452–9.

[55] Harper PR, Knight VA, Marshall AH. Discrete conditional phase-type models utilizing classification trees: application to modelling health service capacities. Eur J Oper Res 2012;219:522–30.

[56] Hosseinifard SZ, Abbasi B, Minas JP. Intensive care unit discharge policies prior to treatment completion. Oper Res Health Care 2014;3:168–75.

[57] Houdenhoven M, Oostrum JM, Wulink G, Hans E, HurinK JL, Bkker J, Kazemier G. Fewer intensive care unit refusals and a higher capacity utilization by using a cyclic surgical case schedule. J Crit Care 2008;23:222–6.

[58] Iapichino G, Gattinoni L, Radrizzani D, Simini B, Bertolini G, Ferla L, Mistraletti G, Porta F, Miranda DR. Volume of activity and occupancy rate in intensive care units. Association with mortality. Intensive Care Med 2004;30:290–7.

[59] Isken MW, Rajagopalan B. Data mining to support simulation modeling of patient flow in hospitals. J Med Syst 2002;26:179–97.

[60] Jaeker JB, Tucker A. Past the point of speeding up: the negative effects of workload saturation on efficiency and patient severity. Manage Sci 2017;63:1042–62.

[61] Katsaliaki K, Mustafee N. Applications of simulation within the healthcare context. J Oper Res So 2011;62:1431–51.

[62] Kc D, Terwiesch C. An econometric analysis of patient flows in the cardiac intensive care unit. Manuf Serv Oper Manag 2012;14:50–65.

[63] Kim SC, Horowitz I, Young K, Buckley TA. Analysis of capacity management of the intensive care unit in a hospital. Eur J Oper Res 1999;115:36–46.

[64] Kim SC, Horowitz I, Young K, Buckley TA. Flexible bed allocation and performance in the intensive care unit. J Oper Manag 2000;18:427–43.

[65] Kim S-H, Chan CW, Olivares M, Escobar G. ICU admission control: an empirical study of capacity allocation and its implication for patient outcomes. Manag Sci 2015;61:19–38.

[66] Kim S-H, Pinker EJ, Rimar J, Bradley E. Impact of severity-adjusted workload on health status of patients discharged from an ICU, Los Angeles: University of Southern California; 2015. Working paper.
[67] Kolker A. Procces modeling of ICU patient flow: effect of daily load leveling of elective surgery on ICU diversion. J Med Syst 2009;33:27–40.
[68] Kramer AA, Higgins TL, Zimmerman JE. The association between ICU readmission rate and patient outcomes. Crit Care Med 2013;41:24–33.
[69] Kripalani S, Jackson AT, Schnipper JL, Coleman EA. Promoting effective transitions of care at hospital discharge: a review of key issues for hospitalists. J Hosp Med 2007;2:314–23.
[70] Laguna M., Marti R. OptQuest callable library. 2006, 10.1007/0-306-48126-X_7.
[71] Levin PD, Sprung CL. The process of intensive care triage. Intensive Care Med 2001;27:1441–5.
[72] Li J, Dong M, Zhao W. Admissions optimisation and premature discharge decisions in intensive care units. Int J Prod Researc 2015;53:7329–42.
[73] Ligtenberg JJ, Arnold LG, Stienstra Y, van der Werf TS, Meertens JH, Tulleken JE, et al. Quality of interhospital transport of critically ill patients: a prospective audit. Crit Care 2005;9:R446–51.
[74] Lin F, Chaboyer W, Wallis M. A literature review of organisational, individual and teamwork factors contributing to the ICU discharge process. Aust Crit Care 2009;22:29–43.
[75] Litvack N, van Rijsbergen M, Boucherie RJ, van Mhoven M. Managing the overflow of intensive care patiens. Eur J Oper Res 2008;185:998–1010.
[76] Mallor F, Azcárate C, Barado J. Optimal control of ICU patient discharge: from theory to implementation. Health Care Manag Sci 2015;18:234–50.
[77] Mallor F, Azcárate C, Barado J. Control problems and management policies in health systems. Application to intensive care units. Flex Serv Manuf J 2016;28:62–89.
[78] Mallor F, Azcárate C. Combining optimization with simulation to obtain credible models for intensive care units. Ann Oper Res 2014;221:255–71.
[79] Marmor YN, Rohleder TR, Cook DJ, Huschka TR, Thompson JE. Recovery bed planning in cardiovascular surgery: a simulation case study. Health Care Manag Sci 2013;16(4):314–27.
[80] Marshall A, Vasilakis C, El-Zardi E. Length of stay-based patient flow models: recent developments and future directions. Health Care Manag Sci 2005;8:213–20.
[81] Masterson BJ, Mihara TG, Miller G, Randolph SC, Forkner E, Crouter AL. Using models and data to support optimization of the military health system: a case study in an intensive care unit. Health Care Manag Sci 2004;7:217–24.
[82] McClean S, Qiao Y, Fullerton K. STOPGAP: stroke patient management and capacity planning. Oper Res Health Care 2015;6:78–86.
[83] McManus ML, Long MC, Cooper A, Litvak E. Queuing theory accurately models the need for critical care resources. Anesthesiology 2004;100:1271–6.
[84] Metcalfe MA, Sloggett A, McPherson K. Mortality among appropriately referred patients refused admission to Intensive-Care Units. Lancet 1997;350:7–11.
[85] Mielczarek B, Uzialko-Mydlikowska J. Application of computer simulation modeling in the health care sector: a survey. Simulation 2012;88:197–216.
[86] Moreno R, Reis Miranda D, Matos R, Fevereiro T. Mortality after discharge from intensive care: the impact of organ system failure and nursing workload use at discharge. Intensive Care Med 2001;27:999–1004.
[87] Nates JL, Nunnally M, Kleinpell R, Blosser S, Goldner J, Birriel B, Fowler CS, Byrum D, Miles WS, Bailey H, Sprung CL. ICU admission, discharge, and triage guidelines: a framework to enhance clinical operations, development of institutional policies, and further research. Crit Care Med 2016;44:1553–602.
[88] Neuts MF. Probability distributions of phase-type. In: Florin Emeritus H, editor. Liber amicorum. Louvain: Department of Mathematics, University of Louvain; 1975. p. 173–206.
[89] Nguyen JM, Six P, Parisot R, Antonioli D, Nicolas F, Lombrail P. A universal method for determining intensive care unit bed requirements. Intensive Care Med 2003;29:849–52.
[90] Ouanes I, Schwebel C, Francais A, Bruel C, Philippart F, Vesin A, et al. A model to predict short-term death or readmission after intensive care unit discharge. J Crit Care 2012;27:422–9.
[91] Pearson GA, Reynolds F, Stickley J. Calculating the need for intensive care beds. Arch Dis Child 2012;97:943–6.
[92] Price C. Applications of operations research models to problems in health care. University of Maryland. College Park.; 2009. Dissertation.
[93] Priestap FA, Martin CM. Impact of intensive care unit discharge time on patient outcome. Crit Care Med 2006;34:2946–51.
[94] Pronovost PJ, Needham DM, Waters H, Birkmeyer CM, Calinawan JR, Birkmeyer JD, Dorman T. Intensive care unit physician staffing: financial modelling of the Leapfrog standard. Crit. Care Med. 2004;32:1247–53.
[95] Rais A, Viana A. Operations research in healthcare: a survey. Intern Transactions in Operational Research 2011;18:1–31.
[96] Rauner MS, Zeiles A, Schaffhauser-Linzattti MM, Hornik K. Modelling the effects of the Austrian inpatient reimbursement system on length-of-stay distributions. OR Spectrum 2003;25:183–206.
[97] Renaud B, Brun-Buisson C, Santin A, Coma E, Noyez C, Fine MJ, et al. Outcomes of early, late, and no admission to the intensive care unit for patients hospitalized with community-acquired pneumonia. Acad Emerg Med 2012;19:294–303.
[98] Ridge JC, Jones SK, Nielsen MS, Shahani AK. Capacity planning for intensive care units. Eur J Oper Res 1998;105:346–55.
[99] Robert R, Reignier J, Tournoux-Facon C, Boulain T, Lesieur O, Gissot V, et al. Refusal of intensive care unit admission due to a full unit: impact on mortality. Am J Respir Crit Care Med 2012;185:1081–7.
[100] Rodrigues F, Zaric G, Stanford D. Discrete event simulation model for planning level 2 “step-down” bed needs using NEMS. Operations Research for Health Care; 2018 https://doi.org/10.1016/j.orhc.2017.10.001.
[101] Rodriguez-Carvajal M, Mora D, Doblas A, Garcia M, Dominguez P, Tristancho A, et al. Impacto de las altas no programadas en la mortalidad hospitalaria tras la estancia en una unidad de cuidados intensivos. Med Intensiva 2011;35:143–9.
[102] Rosenberg AL, Watts C. Patients readmitted to ICUs: a systematic review of risk factors and outcomes. Chest 2000;118:492–502.
[103] Shahani AK, Ridley SA, Nielsen MS. Modelling patient flows as an aid to decision making for critical care capacities and organisation. Anaesthesia 2008;63:1074–80.
[104] Shmueli A, Sprung CL, Kaplan EH. Optimizing admissions to an intensive care unit. Health Care Manag Sci 2003;6:131–6.
[105] Simchen E, Sprung CL, Galai N, Zitser-Gurevich Y, Bar-Lavi Y, Levi L, et al. Survival of critically ill patients hospitalized in and out of intensive care. Crit Care Med 2007;35(2):449–57.
[106] Singh MY, Nayyar V, Clark PT, Kim C. Does after-hours discharge of ICU patients influence outcome? Crit Care Resus 2010;12(3):156–61.
[107] Sinuff T, Kahnamoui K, Cook DJ, Luce JM, Levy MM. Rationing critical care beds: a systematic review. Crit Care Med 2004;32(7):1588–97.
[108] Sprung CL, Baras M, Iapichino G, Kesecioglu J, Lippert A, Hargreaves C, et al. The Eldicus prospective, observational study of triage decision making in European intensive care units: part I–European intensive care admission triage scores. Crit Care Med 2012;40(1):125–31.
[109] Sprung CL, Danis M, Iapichino G, Artigas A, Kesecioglu J, Moreno R, et al. Triage of intensive care patients: identifying agreement and controversy. Intensive Care Med 2013;39(11):1916–24.
[110] Steins K, Walther SM. A generic simulation model for planning critical care resource requirements. Anaesthesia 2013;68(11):1148–55.
[111] Steyrer J, Schiffinger M, Huber C, Valentin A, Strunk G. Attitude is everything?: The impact of workload, safety climate, and safety tools on medical errors: a study of intensive care units. Health Care Manage Rev 2013;38:306–16.
[112] Tan SS, Bakker J, Hoogendoorn ME, Kapila A, Martin J, Pezzi A, Pittoni G, Spronk PE, Welte R, Hakkaart-van Roijen L. Direct cost analysis of intensive care unit stay in four European countries: applying a standardized costing methodology. Value Health 2012;15:81–6.
[113] Teres D. Civilian triage in the intensive care unit: the ritual of the last bed. Crit Care Med 1993;21:598–606.
[114] Tobin A, Santamaría J. After-hours discharges from intensive care are associated with increased mortality. MJA 2006;184:334–7.
[115] Troy PM, Rosenberg L. Using simulation to determine the need for ICU beds for surgery patients. Surgery 2009;146:608–20.
[116] Valentin A, Capuzzo M, Guidet B, Moreno R, Metnitz B, Bauer P, et al. Errors in administration of parenteral drugs in intensive care units: multinational prospective study. BMJ 2009;338:b814.
[117] vanSambeek JR, Cornelissen FA, Bakker PJ, Krabbendam JJ. Models as instruments for optimizing hospital processes: a systematic review. Int J Qual Health Care Assur 2010;23(4):356–77.
[118] Vasilakis C, Marshall AH. Modelling nationwide hospital length of stay: opening the black box. J Oper Res Soc 2005;56:862–9.
[119] Verdon M, Merlani P, Perneger T, Ricou B. Burnout in a surgical ICU team. Intensive Care Med 2008;34(1):152–6.
[120] Wiegersma JS, Droogh JM, Zijlstra JG, Fokkema J, Ligtenberg JJ. Quality of interhospital transport of the critically ill: impact of a mobile intensive care unit with a specialized retrieval team. Crit Care 2011;15(1):R75.
[121] Yang M, Fry MJ, Raikhelkar J, Chin C, Anyanwu A, Brand J, et al. A model to create an efficient and equitable admission policy for patients arriving to the cardiothoracic ICU. Crit Care Med 2013;41(2):414–22.
[122] Yoon KB, Koh SO, Han DW, Kang OC. Discharge decision-making by intensivists on readmission to the intensive care unit. Yonsei Med J 2004;45:193–8.
[123] Young MP, Gooder VJ, McBride K, James B, Fisher ES. Inpatient transfers to the intensive care unit: delays are associated with increased mortality and morbidity. J Gen Intern Med 2003;18:77–83.
[124] Zhang C, Grandits T, Härenstam KP, Hauge JB, Meijer S. A systematic literature review of simulation models for non-technical skill training in healthcare logistics. Adv Simul 2018;3:15.
[125] Zhu Z, Hen BH, Teow KL. Estimating ICU bed capacity using discrete event simulation. Int J Health Care Qual Assur 2012;25(2):134–44. [Web1] Aslett L https://cran.r-project.org/web/packages/PhaseType/PhaseType.pdf accessed in March 2019.