



## Assessing and controlling the impact of hospital capacity planning on the waiting time

Nico Dellaert, Ezgi Cayiroglu & Jully Jeunet

**To cite this article:** Nico Dellaert, Ezgi Cayiroglu & Jully Jeunet (2016) Assessing and controlling the impact of hospital capacity planning on the waiting time, International Journal of Production Research, 54:8, 2203-2214, DOI: [10.1080/00207543.2015.1051668](https://doi.org/10.1080/00207543.2015.1051668)

**To link to this article:** <https://doi.org/10.1080/00207543.2015.1051668>



Published online: 08 Jun 2015.



Submit your article to this journal [↗](#)



Article views: 1051



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 3 View citing articles [↗](#)

## Assessing and controlling the impact of hospital capacity planning on the waiting time

Nico Dellaert<sup>a</sup>, Ezgi Cayiroglu<sup>a</sup> and Jully Jeunet<sup>b\*</sup>

<sup>a</sup>*School of Industrial Engineering, Technische Universiteit Eindhoven, Eindhoven, The Netherlands;* <sup>b</sup>*CNRS, Lamsade, université Paris Dauphine, place du Maréchal de Lattre de Tassigny, Paris Cedex 16, France*

(Received 29 July 2014; accepted 24 April 2015)

In the literature, tactical plans of elective patients aim at increasing hospital efficiency through a better resource utilisation, although hospitals claim that patient satisfaction, usually measured by the waiting time, is also important. In this regard, the purpose of this paper is to show how patient satisfaction can be associated with any tactical plan, by developing a method to compute exact waiting time distributions. We also present a procedure to calculate the exact levels of resource utilisation. Therefore, with our procedures, hospital managers can determine the operational performance of their tactical plan. We then explore two strategies to improve tactical plans in terms of waiting time: slack planning and smooth allocation. A case study based on data from a Dutch cardiothoracic surgery centre shows that slack planning leads to a trade-off between waiting time and hospital efficiency. When slack planning is combined to smooth allocation, additional improvements of the waiting time can be reached.

**Keywords:** service; hospital capacity planning; Markov processes; operational performance; waiting time; resource utilisation

### 1. Introduction

As an increasing number of surgical procedures are now safely performed on an outpatient basis, a lot of research focuses solely on the availability of operating theatres (OT) to schedule elective surgeries (recent surveys include Abdelrasol, Harraz, and Eltawil 2014; Cardoen, Demeulemeester, and Beliën 2010; Guerriero and Guido 2011). However, some types of surgery such as hip replacement for an elderly patient or cardiothoracic surgery procedures often require that patients spend several days in the Intensive Care Unit (ICU) post surgery. For these surgeries, integrating downstream resources such as beds or nursing care in the ICU leads to a better overall performance (Adan et al. 2009; Adan and Vissers 2002; Gupta 2007; Houdenhoven et al. 2007; Hulshof et al. 2013; McManus et al. 2003). In their recent survey, Cardoen, Demeulemeester, and Beliën (2010) recognise that although strides have been made, integration of hospital facilities such as ICU or Post-Anaesthesia Care Unit (PACU) still remain a major research issue.

Multiple resources can be incorporated at the three usual decision levels commonly adopted by hospitals. Strategic decisions involve long-term resource allocation to surgical specialities or to patient categories. For instance, in a recent paper Vanberkel et al. (2014) model the hospital as a queueing system to address the problem of choosing the number of patients in each diagnosis related group in order to maximise reimbursement of treatments, each of these treatments requiring several resources.

At the tactical level, the problem is to develop a master surgery schedule (MSS) that specifies the assignment of operating room time to different surgical specialities or individual surgeons on each day of a medium-term horizon. The impact of MSSs on the utilisation of downstream resources has been stressed by Beliën, Demeulemeester, and Cardoen (2006) through a simulation approach, and deterministic models have been proposed to build MSSs that account for additional resources such as nurses (Beliën and Demeulemeester 2008) or beds (Beliën, Demeulemeester, and Cardoen 2009). Adan et al. (2009) consider cardiothoracic surgeries divided into several categories with a number of expected patients in each category which is fixed, whereas these numbers are decision variables in Vanberkel et al. (2014). The problem – which is also under the scope of this paper – is to determine the number of patients in each category to be operated on for each day of the horizon. The authors use a mixed-integer programme formulation where the objective is to minimise deviations between expected resource consumption and target utilisation levels. Four resources are considered: OT, beds and nurses in the ICU and beds in the Medium Care Unit (MCU). The resulting plan is referred to as the tactical plan of surgeries in the following.

At the operational level, the problem is to schedule surgeries on a daily basis where patients are first assigned to days and surgery slots and then sequenced for a given day. Integration of a PACU in the sequencing problem has been made by Cardoen, Demeulemeester, and Beliën (2009a, 2009b).

---

\*Corresponding author. Email: [jully.jeunet@dauphine.fr](mailto:jully.jeunet@dauphine.fr)

Exact analytical approaches have been implemented to project the workload for downstream resources implied by patient admission or surgery plans at the tactical level, or to determine an admission policy based upon the resources utilisation. A number of references focus on bed requirements/occupancy only, like [Shaw and Marshall \(2007\)](#), [Garg, McClean, and Meenan \(2010\)](#), [Bekker and Koeleman \(2011\)](#) or [Vanberkel et al. \(2011b\)](#) who use a queueing model to calculate the exact distribution of recovering patients in a single ward resulting from a given MSS, where blocks of OT time in each period are assigned to specialities. The approach is then used to build MSSs achieving a balanced ward occupancy ([Vanberkel et al. 2011a](#)). [Fügener et al. \(2014\)](#) extend the works of [Vanberkel et al. \(2011b\)](#) and [Vanberkel et al. \(2011a\)](#) by developing an exact analytical approach to calculate the probability distribution for the number of patients in the ICU and multiple wards. Next, they propose an approach to develop a cyclic MSS that minimises downstream costs based on this distribution of patients.

Analytical approaches are also useful to obtain estimates of quantities such as the number of beds required to reach a given service level in the form of a probability of admission refusal for instance. They can therefore be used for dimensioning facilities at a strategic level. Related works include [Asaduzzaman, Chaussalet, and Robertson \(2010\)](#), [Asaduzzaman and Chaussalet \(2011, 2014\)](#), [de Bruin et al. \(2010, 2009\)](#) who integrate the interaction between the OT and the ICU and use the Erlang expression to estimate the ICU rejection probability.

Improvements of the service level can also be reached through strategies to reduce the waiting time, such as evenly balancing the load for bottlenecks ([Williams, Tai, and Lei 2010](#)) or levelling the bed occupancy ([Ma and Demeulemeester 2013](#)). Most strategies involve prioritisation of patients and consider either a single resource like OT ([Comas et al. 2008](#); [Creemers, Beliën, and Lambrecht 2012](#); [Patrick, Puterman, and Queyranne 2008](#)) or several resources ([Hulshof et al. 2013](#); [Persson and Persson 2009](#)).

The present paper deals with the management of the waiting list for cardiac surgery and builds on the functioning of the Thorax Centre described in [Adan et al. \(2009\)](#). A slightly modified version of these authors' formulation is used to generate a cyclic tactical plan of elective surgeries, where capacity shortages are allowed but penalised. The contribution of our paper is threefold.

- (1) We describe an exact analytical method developed in [Cayiroglu \(2009\)](#) to compute the probability distribution of patients in the ICU and in the MCU resulting from any tactical plan. Contrary to [Fügener et al. \(2014\)](#), who also develop this type of distribution of patients in the ICU and wards, we determine in addition the distribution of the utilisation of OT hours and nursing hours in the ICU. We thus quantify the impact of any tactical plan of elective surgeries on the consumption of four resources: OT, beds in the ICU and in the MCU, nursing care in the ICU.
- (2) Furthermore, the calculation of the queue length distribution is used to determine the average waiting time per category associated with any tactical plan. We thus link tactical plans to patient service.
- (3) We include two strategies – slack planning and smooth allocation – for modifying tactical plans in order to improve their performance in terms of waiting time and resource utilisation. Slack planning consists in increasing the number of expected patients so as to create additional surgery slots in the tactical plan to deal with fluctuations in patient arrival rates. This strategy was already used in [Adan et al. \(2011\)](#) together with a flexibility strategy to derive operational plans from a tactical plan where both emergency and elective patients were considered. Smooth allocation is a new strategy we propose here to evenly spread operations for each category in the tactical plan over the planning horizon. Contrary to most existing strategies to reduce wait times, our strategies avoid prioritisation as it is awkward to implement priorities in cardiac surgery for ethical considerations ([Walton et al. 2007](#)).

The paper is organised as follows. Section 2 describes the mathematical model to get the tactical plan of surgeries. In Section 3, we present the procedure to determine the queue length distribution. The steady state probabilities of the number of patients on each day are derived to compute the average waiting time and they are also used in Section 4 to calculate the exact levels of resource utilisation. Section 5 provides an extension to the mathematical model described in Section 2 to obtain a smoother allocation of patients in the tactical plan. In Section 6, data from the Thorax Centre are used to generate and to assess the tactical plans obtained with and without smooth allocation and slack planning, both in terms of waiting time and resource utilisation. Section 7 draws the main conclusions of this paper.

## 2. Mathematical model for the tactical plan

The model is based on the functioning of the Thorax Centre Rotterdam described in [Adan et al. \(2009\)](#). Patients are grouped in several categories that are homogeneous in terms of resource consumption. Some categories of patients are admitted to the MCU one day before surgery. Each patient category is associated with a deterministic operation duration, based on average observed values. After surgery, patients are transferred to the ICU where they benefit from specialised nursing care. The nursing hours required for each patient category on each post-operative day are deterministic. After recovery, patients may

Table 1. Notation.

| <i>Parameters</i> |   |
|-------------------|---|
| $N$               | Number of patient categories.   |
| $c$               | Category index, $c = 1, \dots, N$ .   |
| $T$               | Length of the cyclic planning horizon, in days.   |
| $t$               | Day index, $t = 1, \dots, T$ .  |
| $V_c$             | Target number of elective patients in category $c$ to be operated on during the horizon.  |
| $s_c$             | Surgery duration in hours for a patient in category $c$ .   |
| $l_c$             | Number of pre-operative days in the MCU for a patient in cat. $c$ .   |
| $r$               | Resource index, $r \in \{OT, IC, NH, MC\}$ .  |
| $L_{ICU}^{\max}$  | Maximum length of stay recorded in the ICU over all patient categories.   |
| $L_{MCU}^{\max}$  | Maximum length of stay recorded in the MCU over all patient categories.   |
| $P_{ICU,c,j}$     | Probability that a patient in category $c$ is in the ICU, $j$ days after surgery, $j = 0, 1, 2, \dots, L_{ICU}^{\max}$ .                            |
| $P_{MCU,c,j}$     | Probability that a patient in category $c$ is in the MCU, $j$ days after surgery, $j = 0, 1, 2, \dots, L_{MCU}^{\max}$ .                            |
| $w_{c,j}$         | Intensive care nursing workload (in hours) required for a patient in category $c$ , $j$ days after surgery.   |
| $K_{r,t}$         | Maximum capacity for resource $r$ on day $t$ (expressed in number of hours for OT and NH and in number of beds for IC and MC).                      |
| $A_{r,t}$         | Target utilisation of resource $r$ on day $t$ .   |
| $\alpha_r$        | Relative importance of resource $r$ as assessed by the stakeholders of the hospital.  |
| <i>Variables</i>  |   |
| $X_{c,t}$         | Number of patients in category $c$ planned for surgery on day $t$ , with $c = 1, \dots, N$ and $t = 1, \dots, T$ .                                  |
| $O_{r,t}$         | Over utilisation of resource $r$ on day $t$ , relative to its target level of utilisation, with $r \in \{OT, IC, NH, MC\}$ and $t = 1, \dots, T$ .  |
| $U_{r,t}$         | Under utilisation of resource $r$ on day $t$ , relative to its target level of utilisation, with $r \in \{OT, IC, NH, MC\}$ and $t = 1, \dots, T$ . |
| $E_{r,t}$         | Capacity shortage of resource $r$ on day $t$ , with $r \in \{OT, IC, NH, MC\}$ and $t = 1, \dots, T$ .  |

stay in the MCU for a few days. Lengths of stay in the ICU and in the MCU are probabilistic, with different distributions for each category. We thus consider four critical resources: operating theatre hours (OT), beds in the ICU (IC), nursing hours in the ICU (NH) and beds in the MCU (MC).

In the medium-run, the problem faced by hospital managers resides in the determination of a tactical plan for all elective patients, so as to optimally allocate the major resources while operating on all patients that are expected during a typical horizon. This problem is formulated as a mixed-integer programme where the objective is to minimise the weighted sum of deviations between the resource consumption and their target level of utilisation. The present formulation is a modified version of the model developed by [Adan et al. \(2009\)](#) where capacity shortages are allowed here but penalised. We thus add variables representing shortages for each resource in each day, both in the objective function and the capacity constraints of the model. We adopt the notation and definitions in Table 1.

Formally, the objective is to determine the values of variables  $\{X_{c,t}\}$  satisfying a number of constraints and for which the daily expected utilisation of each resource deviates as little as possible from the daily target. The target utilisation for each resource was determined by the hospital management, usually between 70 and 80 of the maximum capacity. The planning of patients  $\{X_{c,t}\}$  must also minimise capacity shortages. Thus, the objective function to be minimised can be written as

$$\sum_{r \in \{OT, IC, NH, MC\}} \alpha_r \sum_{t=1}^T (O_{r,t} + U_{r,t} + b \cdot E_{r,t}), \quad (1)$$

where  $b \geq 0$  is a constant penalising the capacity shortage. Weights  $\alpha_r$  in the objective function were decided by the stakeholders of the hospital and reflect the degree of flexibility of each resource. For instance,  $\alpha_{IC}$  was set to a greater value

than that of  $\alpha_{OT}$  because finding an extra bed in the ICU was considered as more difficult than calling for an additional surgeon (for a detailed discussion, see [Adan et al. 2011](#)).

The total number of patients in group  $c$  to be operated on over the  $T$ -day cycle must be equal to the target patient throughput  $V_c$ . Target values are obtained by rounding up the average number of patients based on historical data. Hence,

$$\sum_{t=1}^T X_{c,t} = V_c, \quad c = 1, \dots, N. \quad (2)$$

The expected utilisation of the OT by the patients must satisfy

$$\left. \begin{aligned} \sum_{c=1}^N s_c X_{c,t} &\leq K_{OT,t} + E_{OT,t} \\ &\leq A_{OT,t} + O_{OT,t} \\ &\geq A_{OT,t} - U_{OT,t} \end{aligned} \right\} t = 1, \dots, T. \quad (3)$$

The expected number of utilised beds in the ICU must satisfy the three following inequalities

$$\left. \begin{aligned} \sum_{c=1}^N \sum_{j=0}^{L_{ICU}^{\max}} p_{ICU,c,j} X_{c,t-j} &\leq K_{IC,t} + E_{IC,t} \\ &\leq A_{IC,t} + O_{IC,t} \\ &\geq A_{IC,t} - U_{IC,t} \end{aligned} \right\} t = 1, \dots, T. \quad (4)$$

In the above constraints, we used the convention that the subscript  $t - j$  in  $X_{c,t-j}$  should be treated modulo  $T$ : day 0 is the same as day  $T$ , day  $-1$  is the same as day  $T - 1$  and so on.

For the expected number of utilised nursing hours in the ICU, we must have

$$\left. \begin{aligned} \sum_{c=1}^N \sum_{j=0}^{L_{ICU}^{\max}} w_{c,j} p_{ICU,c,j} X_{c,t-j} &\leq K_{NH,t} + E_{NH,t} \\ &\leq A_{NH,t} + O_{NH,t} \\ &\geq A_{NH,t} - U_{NH,t} \end{aligned} \right\} t = 1, \dots, T. \quad (5)$$

Similarly, the expected number of utilised beds in the MCU must satisfy

$$\left. \begin{aligned} \sum_{c=1}^N \sum_{j=1}^{l_c} X_{c,t+j} + \sum_{c=1}^N \sum_{j=0}^{L_{MCU}^{\max}} p_{MCU,c,j} X_{c,t-j} &\leq K_{MC,t} + E_{MC,t} \\ &\leq A_{MC,t} + O_{MC,t} \\ &\geq A_{MC,t} - U_{MC,t} \end{aligned} \right\} t = 1, \dots, T. \quad (6)$$

It should be noted that according to the functioning of the Thorax Centre, patients are always hospitalised for a deterministic number of days in the MCU before surgery. As operating rooms on weekends are dedicated only to emergency patients, we have to require that

$$X_{c,t} = 0 \text{ and } X_{c,t+1} = 0, \quad t = 6 + 7(j - 1); \quad j = 1, \dots, (T/7); \quad c = 1, \dots, N. \quad (7)$$

Our tactical planning problem therefore consists in minimising the objective function in (1) subject to constraints (2) to (7) with the following integrality constraints

$$X_{c,t} \in \{0, 1, 2, \dots\}, \quad c = 1, \dots, N; \quad t = 1, \dots, T. \quad (8)$$

The decision variables consist of the number of patients in each category planned for surgery on each day, the over and under utilisation of each resource compared to its target level of utilisation on each day and the capacity shortage on each day for each resource. We refer to this model as the basic mathematical model. In the next section, we present a procedure to compute the waiting time associated with any tactical plan of surgery.

### 3. Queue length distribution and average waiting time

Tactical planning models in the literature like the one we use here consider an average number of patients to be operated on, based on historical data so wait time can hardly be optimised at this level. The waiting time is defined as the time interval between confirmation of the need for surgery and subsequent acceptance on the waiting list, and surgery.

Since the operations of each category of patient are assumed to be independent from each other, the waiting time calculation is done for each category separately. For this calculation, we need the steady-state probabilities of the patient queue for every patient category on every working day. This queue can be described as a Markov process, where the number increases by daily patient arrivals and decreases on operation days.

Patient arrivals in each category  $c$  on each day are assumed to follow a Poisson distribution with parameter  $\lambda_c$ . We let  $a_{c,t}$  be the number of patients in category  $c$  arriving on day  $t$ . We assume that the decisions about the operational plan are made  $L$  days in advance. For each category, we thus compare the number of patients in the queue on day  $t$  with the number of planned surgeries on day  $t + L$ ,  $X_{c,t+L}$ , in order to determine the number of patients actually undergoing surgery on day  $t + L$ . The probability that the number of patients of category  $c$  in the queue before beginning surgery on day  $t$  equals  $q_{c,t}^{\text{BS}}$  is denoted by  $Q_{c,t}^{\text{BS}}(q_{c,t}^{\text{BS}})$ , where superscript BS stands for Before Surgery. The number of patients in category  $c$  actually operated on day  $t + L$  equals  $\min(q_{c,t}^{\text{BS}}, X_{c,t+L})$ . If the number of patients of category  $c$  in the queue,  $q_{c,t}^{\text{BS}}$ , is less than the number of planned patients, we operate on all queued patients. Otherwise, the number of patients in category  $c$  that are not operated on is equal to  $q_{c,t}^{\text{BS}} - X_{c,t+L}$ . Thus, the number of patients of category  $c$  that are not operated on after the operation session on day  $t + L$  equals  $q_{c,t}^{\text{NS}} = q_{c,t}^{\text{BS}} - \min(q_{c,t}^{\text{BS}}, X_{c,t+L})$  with probability  $Q_{c,t}^{\text{NS}}(q_{c,t}^{\text{NS}})$ , where superscript NS stands for No Surgery. Finally, the number of patients of category  $c$  before the surgery session on day  $t + L$  is  $q_{c,t}^{\text{BS}} = a_{c,t-1} + q_{c,t-1}^{\text{NS}}$ .

We use an iterative method, the Power method (Bolch et al. 2006), until the values of  $Q_{c,t}^{\text{NS}}(q_{c,t}^{\text{NS}})$  and  $Q_{c,t}^{\text{BS}}(q_{c,t}^{\text{BS}})$  reach their steady state. Obtaining the steady state probabilities allows for the calculation of the average waiting time for each patient category.

We consider the waiting time of patients in each category  $c$ ,  $WT_c$ , to be composed of two elements,  $WT_c^{\text{BS}}$  and  $WT_c^{\text{NS}}$ , with  $WT_c = WT_c^{\text{BS}} + WT_c^{\text{NS}}$ , where  $WT_c^{\text{BS}}$  is the average waiting time before the first possible operation day and  $WT_c^{\text{NS}}$  is the additional expected waiting time caused by not having surgery on the first possible operation day. Thus,  $WT_c^{\text{NS}}$  considers the probabilities that there are more patients than the number of planned operations on the first possible operation day. Let  $d_{c,t}$  be the number of days between operation day  $t$  for category  $c$  and the next operation day  $t'$  for category  $c$  immediately following day  $t$ , with  $t' \geq t + 1$ . For all  $t = 1, \dots, T$ , the number of days  $d_{c,t}$  is thus defined as follows: if there is no operation of category  $c$  on day  $t$ , we set  $d_{c,t} = 0$ , otherwise we have  $d_{c,t} = t' - t$ . There are on average  $\lambda_c d_{c,t}$  arriving patients between two consecutive operation days  $t$  and  $t'$  and the average waiting time of a patient before the first possible operation day is equal to  $d_{c,t}/2$  days. When we consider the whole horizon  $T$ , we get for category  $c$

$$WT_c^{\text{BS}} = \left( 1/2\lambda_c \sum_{t=1}^T d_{c,t}^2 \right) / \left( \lambda_c \sum_{t=1}^T d_{c,t} \right). \quad (9)$$

Let us now determine  $WT_c^{\text{NS}}$ . If there are more queued patients than planned, the surplus  $q_{c,t}^{\text{BS}} - X_{c,t+L}$  will wait until the next operation day, thus for  $d_{c,t}$  days. For patient category  $c$ , we have

$$WT_c^{\text{NS}} = (1/(T\lambda_c)) \sum_{t=1}^T \sum_{q_{c,t}^{\text{BS}} > X_{c,t+L}} \left( q_{c,t}^{\text{BS}} - X_{c,t+L} \right) \cdot Q_{c,t}^{\text{BS}}(q_{c,t}^{\text{BS}}) \cdot d_{c,t}. \quad (10)$$

Now we have calculated the average total waiting time per category, we shall utilise the steady state probabilities to determine the resource utilisation levels.

#### 4. Distributions of the resource utilisation levels

In the previous section, we obtained the steady state probabilities for the number of queueing patients in each category on each day. In addition to the average waiting time calculations, these steady state probabilities can also be used to determine the bed utilisation levels in the ICU and in the MCU, the nursing hours in the ICU and the OT hours. We develop here in details only the reasoning to get the distribution of occupied beds in the ICU for a similar method is employed to obtain the distribution of consumption of the three other resources. An example of these computations is provided in the Appendix 1. The algorithm to determine the distribution for the number of occupied beds in the ICU consists of the four following steps.

*Step 1* Let  $S_c(n, t)$  be the probability that  $n$  operations of category  $c$  take place on day  $t$ . For each day  $t = 1, \dots, T$  and for each category  $c = 1, \dots, N$ , we have

$$S_c(n, t) = Q_{c,t-7}^{\text{BS}}(n), \quad \forall n = 0, \dots, X_{c,t} - 1,$$

$$S_c(X_{c,t}, t) = 1 - \sum_{k=0}^{X_{c,t}-1} Q_{c,t-7}^{\text{BS}}(k). \quad (11)$$

If the number  $n$  of patients in the queue is less than the number of planned patients, then we operate on all  $n$  patients with the probability there are  $n$  patients in the queue. If the number  $n$  of patients in the queue exceeds or equals the number of planned patients  $X_{c,t}$ , we operate on  $X_{c,t}$  patients (with the probability that there are at least  $X_{c,t}$  patients in the queue).

*Step 2* Let  $P_c(n, t, j)$  be the probability that there are  $n$  patients of category  $c$  still in the ICU  $j$  days after their surgery, that took place on day  $t$ . This probability can be easily derived by considering a binomial distribution based on the probability that an individual patient is in the ICU on this specific day. For each category  $c = 1, \dots, N$ , for all  $t = 1, \dots, T$ , and for all  $j = 0, \dots, L_{\text{ICU}}^{\text{max}}$ , we have

$$P_c(n, t, j) = \begin{cases} \sum_{k=n}^{X_{c,t}} \binom{k}{n} S_c(k, t) p_{\text{ICU},c,j}^n (1 - p_{\text{ICU},c,j})^{k-n} & \text{if } n \leq X_{c,t}, \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

If  $n \leq X_{c,t}$ , having  $n$  patients of category  $c$  staying in the ICU  $j$  days after their surgery that took place on day  $t$  results from the following combinations:

- $n$  patients receive surgery on day  $t$  and all of them are still in the ICU  $j$  days after surgery, with probability  $p_{\text{ICU},c,j}^n$ ;
- $n + 1$  patients receive surgery on day  $t$ , with  $n$  patients still in the ICU  $j$  days after surgery (with probability  $p_{\text{ICU},c,j}^n$ ) and one patient has left before, with probability  $1 - p_{\text{ICU},c,j}$ ;
- $n + 2$  patients receive surgery on day  $t$ , with  $n$  patients still in the ICU  $j$  days after surgery and two patients have left before, with probability  $(1 - p_{\text{ICU},c,j})^2$ , etc.

The same reasoning applies until we reach the maximum number of patients  $X_{c,t}$  that can undergo surgery on day  $t$ .

*Step 3* Let  $TP_c(n, t, j)$  be the probability that on day  $t$  there are exactly  $n$  patients of category  $c$  in the ICU who will stay at least  $j$  more days. Let  $R_c(n, t)$  be the probability there are exactly  $n$  patients of category  $c$  in the ICU on day  $t$ . Obviously, for all  $n$  and for all  $t = 1, \dots, T$ , we have for each category  $c = 1, \dots, N$

$$R_c(n, t) = TP_c(n, t, 0). \quad (13)$$

The set of  $n$  patients of category  $c$  that are in the ICU on day  $t$  consists of patients who underwent surgery in the last  $L_{\text{ICU}}^{\text{max}}$  periods. In order to calculate  $R_c(n, t)$ , we would need a multiplication of probabilities of all combinations of patient numbers and lengths of stay in the ICU that would lead to get the exact number of patients. Instead of describing the combinations explicitly, we adopt a recursive reasoning: to have exactly  $n$  patients in the ICU on day  $t$  staying at least for  $j$  days, we must have  $n$  patients in the ICU on day  $t - 1$  staying at least for  $j + 1$  days and zero patients receiving surgery on day  $t$ ; or  $n - 1$  patients in the ICU on day  $t - 1$  (staying at least for  $j + 1$  days) and one patient receiving surgery on day  $t$  staying for  $j$  days; or  $n - 2$  patients in the ICU on day  $t - 1$  and two patients receiving surgery on day  $t$  and staying for  $j$  days, etc. We thus have for all  $n$  and for all  $t = 1, \dots, T$

$$TP_c(n, t, j) = \sum_{k=0}^n TP_c(n - k, t - 1, j + 1) \cdot P_c(k, t, j). \quad (14)$$

As a starting point, we let for all possible  $n$  values

$$TP_c(n, t - L_{\text{ICU}}^{\text{max}}, L_{\text{ICU}}^{\text{max}}) = P_c(n, t - L_{\text{ICU}}^{\text{max}}, L_{\text{ICU}}^{\text{max}}). \quad (15)$$

*Step 4* Once the probabilities of bed utilisation level in the ICU for each category have been determined in this way, they are combined to obtain the overall utilisation of beds in the ICU. Formally, we let  $R(n, t)$  be the probability there are exactly  $n$  patients in the ICU on day  $t$ . Again, we use a temporary variable, namely  $TR(n, t, c')$ , for the recursive calculations, where  $c'$  means that categories  $c = 1, \dots, c'$  are taken into account in the computation. For all number  $n$  of patients and for all

$t = 1, \dots, T$ , we have

$$\begin{aligned} R(n, t) &= TR(n, t, N), \\ TR(n, t, c') &= \sum_{k=0}^n TR(n-k, t, c'-1)R_{c'}(k, t). \end{aligned} \quad (16)$$

As a starting point, we let for all possible  $n$  values

$$TR(n, t, 1) = R_1(n, t). \quad (17)$$

Finally,  $TR(n, t, N)$  describes the probability that we have exactly  $n$  patients in the ICU on day  $t$ . A similar procedure is employed to get the bed utilisation level in the MCU, the nursing hours in the ICU as well as the OT hours.

## 5. Slack planning and smooth allocation

The levels of resource utilisation as optimised in the tactical plan are considered as an indicator of hospital efficiency but hospitals claim that waiting time is also important. However, the allocation of patients resulting from this optimisation may lead to long waiting times if planned operations are much less than the number of queued patients or if they are too clustered. To allow for a faster admission of queued patients, we apply the slack planning strategy that consists in overestimating the number of patients on the basis of which the tactical plan is computed. This leads to additional slots at the tactical level that can be utilised at the operational level. We already used this strategy in [Adan et al. \(2011\)](#). Three slack planning options are considered. The large slack planning option consists in planning a number of patients per category such that less than 5% of patients have to wait more than one cycle. In the medium slack planning, this percentage is set to 10%. We also consider the no slack planning option where the target throughput of patients for each category is left equal to the average number of arrivals based on historical data. The slack planning strategy thus amounts to modifying the values of target throughputs  $V_c$  in constraint (2) of the MIP to obtain the tactical plan (see Section 2).

Besides, to avoid having too many patients clustered over the horizon, we propose in this paper, a smooth allocation strategy which is expected to reduce the variance in the time between patient arrivals and the first possible operation slot. To get a smooth allocation of patients, the horizon is divided for each category into smoothing cycles, each containing same number of patients. We impose that smoothing cycle lengths are multiples of seven days and are no less than seven days. By doing so, we limit the number of additional constraints in the basic mathematical model (Section 2), hence, the degradation of the objective function. For instance, with a target throughput of patients  $V_c = 3$  and  $T = 28$ , we would have ideally one patient to be operated on every  $T/V_c = 9.33 \simeq 9$  days. This would be translated in the basic mathematical model by moving cycle constraints requiring that at least one patient undergoes surgery in each time interval  $\{1, \dots, 9\}, \{2, \dots, 10\}, \dots, \{19, \dots, 28\}$ . If the cycle is set to 14 days instead of 9 days, then at least one surgery takes place in each interval  $\{1, \dots, 14\}, \{2, \dots, 15\}, \dots, \{15, \dots, 28\}$ , with a number of constraints that drops from 20 to 15.

The smoothing cycle length,  $\theta_c$ , expressed in number of days for each category  $c$  is therefore defined as

$$\theta_c = 7 \lceil (T/V_c) / 7 \rceil. \quad (18)$$

To formulate the moving cycle constraint, we let  $v_c$  be the cycle index with  $v_c = 1, \dots, T/\theta_c$  and  $m_c$  the moving cycle index, with  $m_c = 1, \dots, T - \theta_c + 1$ . For each category  $c$ , the moving cycle constraint is written as

$$\sum_{k=m_c}^{m_c-1+\theta_c} X_{c,k} \geq \lfloor V_c \theta_c / T \rfloor, \quad \forall m_c = 1, \dots, T - \theta_c + 1. \quad (19)$$

It should be noted that, when  $T = 28$ , the right-hand side of constraint (19) always equals 1 for  $V_c \leq 7$  and is strictly greater than 1 for values of  $V_c$  from 8 patients.

As the lowest integer value of the number of operations per smoothing cycle is considered in (19), another constraint is added for each category to ensure that cumulated operations until the next cycle meet the requirements:

$$\sum_{k=1}^{v_c \theta_c} X_{c,k} \geq \lfloor v_c V_c \theta_c / T \rfloor, \quad \forall v_c = 1, \dots, T/\theta_c. \quad (20)$$

In our example with  $V_c = 3$  and  $T = 28$ , constraint (20) guarantees that the number of planned patients is at least equal to 1 in time interval  $\{1, \dots, 14\}$  and at least equal to  $V_c = 3$  over the whole horizon  $\{1, \dots, 28\}$ . Obviously, the number of planned patients in each time interval will take a minimum value due to the minimisation process and the total number of planned patients will be equal to  $V_c$  as required by Equation (2).

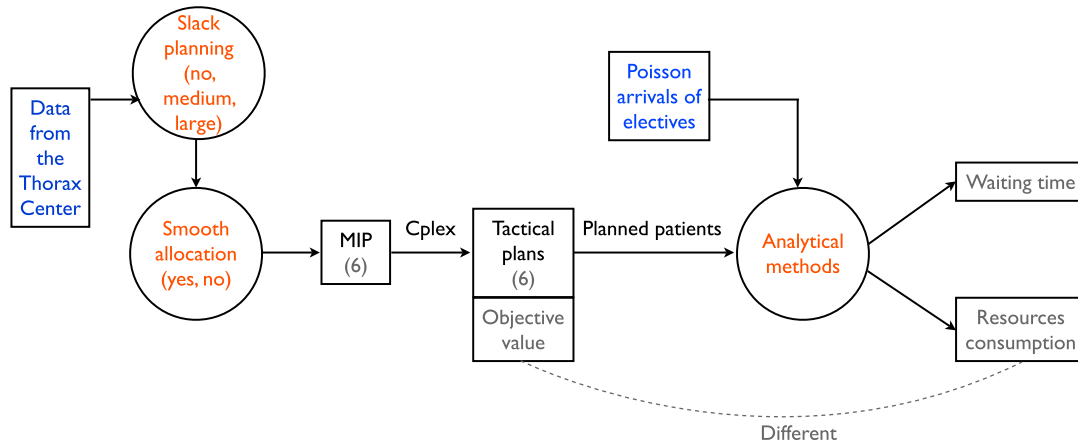


Figure 1. Summary of the approach.

Table 2. Average waiting time in days.

| Slack planning | Basic model | Smooth allocation |
|----------------|-------------|-------------------|
| No             | 29.91       | 29.28             |
| Medium         | 9.33        | 8.50              |
| Large          | 5.98        | 5.37              |

The smooth allocation model results from the addition of constraints (19) and (20) to the basic mathematical model presented in Section 2, with smoothing cycle lengths computed according to (18).

Combining the three slack planning options to both the basic mathematical model and the smooth allocation model results in six tactical plans that are compared in terms of waiting time and resource utilisation in the next section.

## 6. Case study

Our incentive is to assess the impact of slack planning and smooth allocation on both the waiting time and the resource consumption that we are now able to compute exactly, using the methods we developed in Sections 3 and 4. Figure 1 summarises the approach.

We used the data of the Thorax Centre available in the paper of Adan et al. (2009) with same values for all parameters defined in Table 1 ( $T = 28$ ,  $N = 8$ , etc.). Three basic mathematical models were obtained by applying the three slack planning options. We added to these basic models constraints (19) and (20) for smooth allocation described in Section 5, with smoothing cycles determined according to Equation (18). We thus obtained six mathematical programs that we solved using Cplex12.1, with a fixed maximum computation time of 10 min. The solving process led to six tactical plans with optimality gaps ranging from 8.31 to 28.82%. Because running Cplex for an hour did not bring any significant improvement of the objective values, we considered the tactical plans obtained after 10 min as good enough to be exploited.

Arrivals of patients were simulated using a Poisson process with different parameters according to the categories of patients (values of  $\lambda_c$  are available in Adan et al. (2009)) over 100 cycles of 28 days. The expected waiting time was computed for each category of patients, using Equations (9) and (10). The average waiting time over all categories was obtained by adding up the expected waiting time of each category and multiplying by its relative frequency. Table 2 displays the average waiting time for the basic mathematical model and the smooth allocation model, each of them being combined with the three slack planning options. For instance, with the basic model and no slack planning, the average waiting time over all categories equals 29.91 days, which means that the average number of patients in queue is larger than the number that can be operated on over 28 days.

Results are striking: switching from the basic situation (no slack planning and no smooth allocation) to the situation where both strategies are used with the large slack planning option, the waiting time decreases by a factor of six (from about 30 days to 5 days). As expected, smooth allocation of patients reduces the waiting time since operations are planned more frequently and more regularly. But slack planning leads to a much stronger decrease in the waiting time as it allows for more

Table 3. Average under utilisation, over utilisation and capacity shortage per day.

| Resource                  | Slack | Under utilisation |        | Over utilisation |        | Capa. shortage |        |
|---------------------------|-------|-------------------|--------|------------------|--------|----------------|--------|
|                           |       | No smooth         | Smooth | No smooth        | Smooth | No Smooth      | Smooth |
| OT<br>(hours)             | No    | 1.217             | 1.219  | 0.353            | 0.354  | 0.000          | 0.000  |
|                           | Med.  | 1.741             | 1.677  | 0.878            | 0.813  | 0.000          | 0.000  |
|                           | Large | 2.065             | 1.927  | 1.202            | 1.064  | 0.000          | 0.000  |
| IC<br>(number<br>of beds) | No    | 0.677             | 0.677  | 0.319            | 0.319  | 0.001          | 0.001  |
|                           | Med.  | 0.713             | 0.704  | 0.355            | 0.347  | 0.001          | 0.001  |
|                           | Large | 0.764             | 0.744  | 0.406            | 0.386  | 0.002          | 0.002  |
| NH<br>(hours)             | No    | 8.928             | 8.934  | 3.912            | 3.918  | 0.000          | 0.000  |
|                           | Med.  | 9.378             | 9.267  | 4.364            | 4.253  | 0.000          | 0.000  |
|                           | Large | 9.953             | 9.794  | 4.938            | 4.780  | 0.000          | 0.000  |
| MC<br>(number<br>of beds) | No    | 1.753             | 1.767  | 2.061            | 2.075  | 0.005          | 0.005  |
|                           | Med   | 1.922             | 1.914  | 2.232            | 2.224  | 0.008          | 0.007  |
|                           | Large | 1.998             | 1.939  | 2.308            | 2.249  | 0.011          | 0.008  |

flexibility in the admission possibilities whereas smooth allocation constrains the surgeries to take place evenly along the horizon. Additional improvements of the waiting time can be reached by combining the two strategies.

It should be noted that exact resource consumptions at the operational level differ from their expected levels as optimised in the tactical plan. Exact probability distributions for each resource utilisation are computed using the method described in Section 4. Over and under utilisations compared to target levels (resource deviations, in the following) are then derived as well as the capacity shortage for each resource. Table 3 displays for each resource the average under utilisation, over utilisation and capacity shortage per day over 100 cycles of 28 days. For instance, when none of the strategies are used, there is on average an under utilisation per day of 0.677 beds in the ICU compared to the target level, an over utilisation of 0.319 beds per day and a shortage of 0.001 beds per day.

Except for scenarios with no slack planning, smooth allocation always leads to plans with less resource deviations than the basic model. In combination with medium and large slack planning, smooth allocation decreases the capacity shortage of beds in the MCU. Slack planning creates a trade-off between waiting time and hospital efficiency, as evidenced by Tables 2 and 3: increasing the number of planned patients substantially reduces the waiting time whereas shortages and deviations between resource utilisation and their target level are augmented. Slack planning creates operation slots that contribute to decrease in the size of the queue at the operational level. With more flexibility in the use of planned slots, there is also a greater chance to get higher resource deviations in some periods. To improve the waiting time at the expense of a decrease in the hospital efficiency, the best option is definitely to combine the two strategies.

## 7. Conclusion

In this paper, we have shown how the probability distribution of waiting times can be computed from any tactical plan of surgeries. For the waiting time computations, we used the steady state probabilities of the number of patients in each category on each day. These probabilities have also been exploited to compute the exact levels of resource utilisation. To improve the waiting time, we considered two strategies: slack planning and smooth allocation.

Results showed that slack planning strongly reduces the waiting time but is detrimental to the hospital efficiency. A better patient satisfaction and a lower degradation of the hospital efficiency can be reached by combining the two strategies. Thus, the best option will depend on the relative preferences of the stakeholders for these two criteria. Hospital executives have to prioritise their performance goals, and the current tendency is to adopt a greater competitor focus than customer focus. Hospitals continue to undergo the market pressure that has gradually transformed health care from a social good to an economic good. Nevertheless, the hospitals' attention to patient satisfaction is critical to maintaining the support of stakeholders like medical staff and patients.

A logical extension of our model would be to include emergency patients. Under some strict conditions, we could use similar formulae for this group and still have analytical results. This would require that we have reserved some operating time per day for this group that we use all remaining time from the elective patients for the emergency operations (disregarding

their categories). However using the empty operating slots for other types of patients, described as the flexibility options in Adan et al. (2011) leads to a better performance. The downside of these valuable rules is that our model would no longer be analytically tractable and that we would need a simulation model for that. Therefore, we leave these options to a future paper.

### Disclosure statement

No potential conflict of interest was reported by the authors.

### References

- Abdelrasol, Z., N. Harraz, and A. Eltawil. 2014. "Operating Room Scheduling Problems: A Survey and a Proposed Solution Framework." *Transactions on Engineering Technologies* 717–731.
- Adan, I., J. Bekkers, N. P. Dellaert, J. Jeunet, and J. Vissers. 2011. "Improving Operational Effectiveness of Tactical Master Plans for Emergency and Elective Patients Under Stochastic Demand and Capacitated Resources." *European Journal of Operational Research* 213: 290–308.
- Adan, I., J. Bekkers, N. Dellaert, J. Vissers, and X. Yu. 2009. "Patient Mix Optimization and Stochastic Resource Requirements: A Case Study in Cardiothoracic Surgery Planning." *Health Care Management Science* 12: 129–141.
- Adan, I., and J. Vissers. 2002. "Patient Mix Optimisation in Hospital Admission Planning: A Case Study." *International Journal of Operations and Production Management* 22: 445–461.
- Asaduzzaman, Md, and T. J. Chaussalet. 2011. "An Overflow Loss Network Model for Capacity Planning of a Perinatal Network." *Journal of the Royal Statistical Society* 174: 403–417.
- Asaduzzaman, Md, and T. J. Chaussalet. 2014. "Capacity Planning of a Perinatal Network with Generalised Loss Network Model with Overflow." *European Journal of Operational Research* 232: 178–185.
- Asaduzzaman, Md, T. J. Chaussalet, and N. J. Robertson. 2010. "A Loss Network Model with Overflow for Capacity Planning of a Neonatal Unit." *Annals of Operations Research* 178: 67–76.
- Bekker, R., and P. M. Koeleman. 2011. "Scheduling Admissions and Reducing Variability in Bed Demand." *Health Care Management Science* 14: 237–249.
- Beliën, J., and E. Demeulemeester. 2008. "A Branch-and-price Approach for Integrating Nurse and Surgery Scheduling." *European Journal of Operational Research* 189: 652–668.
- Beliën, J., E. Demeulemeester, and B. Cardoen. 2006. "Visualizing the Demand for Various Resources as a Function of the Master Surgery Schedule: A Case Study." *Journal of Medical Systems* 30: 343–350.
- Beliën, J., E. Demeulemeester, and B. Cardoen. 2009. "A Decision Support System for Cyclic Master Surgery Scheduling with Multiple Objectives." *Journal of Scheduling* 12: 147–161.
- Bolch, G. S. Greiner, H. de Meer, and K. S. Trivedi. 2006. *Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications*. 2nd ed. Wiley.
- Cardoen, B., E. Demeulemeester, and J. Beliën. 2009a. "Optimizing a Multiple Objective Surgical Case Sequencing Problem." *International Journal of Production Economics* 119: 354–366.
- Cardoen, B., E. Demeulemeester, and J. Beliën. 2009b. "Sequencing Surgical Cases in a Day-care Environment: An Exact Branch-and-price Approach." *Computers and Operations Research* 36: 2660–2669.
- Cardoen, B., E. Demeulemeester, and J. Beliën. 2010. "Operating Room Planning and Scheduling: A Literature Review." *European Journal of Operational Research* 201: 921–932.
- Cayiroglu, E. 2009. "Tactical Strategies for Scheduling Elective and Emergency Patients Under Uncertainty. Operations Management and Logistics." Master's Thesis, Department of Technology Management, Technische Universiteit Eindhoven.
- Comas, M., X. Castells, L. Hoffmeister, R. Roman, F. Cots, J. Mar, S. Gutiérrez-Moreno, and M. Espallargues. 2008. "Discrete-event Simulation Applied to Analysis of Waiting Lists. Evaluation of a Prioritization System for Cataract Surgery." *Value in Health* 7: 1203–1213.
- Creemers, S., J. Beliën, and M. Lambrecht. 2012. "The Optimal Allocation of Server Time Slots Over Different Classes of Patients." *European Journal of Operational Research* 219: 508–521.
- de Bruin, A. M., R. Bekker, L. van Zanten, and G. M. Koole. 2010. "Dimensioning Hospital Wards Using the Erlang Loss Model." *Annals of Operations Research* 178: 23–43.
- van Dijk, N. M., and N. Kortbeek. 2009. "Erlang Loss Bounds for OT-ICU Systems." *Queueing Systems* 63: 253–280.
- Fügner, A., E. W. Hans, R. Kolisch, N. Kortbeek, and P. T. Vanberkel. 2014. "Master Surgery Scheduling with Consideration of Multiple Downstream Units." *European Journal of Operational Research* 239: 227–236.
- Garg, L., S. McClean, and B. Meenan. 2010. "A Non-homogeneous Discrete Time Markov Model for Admission Scheduling and Resource Planning in a Cost or Capacity Constrained Healthcare System." *Health Care Management Science* 13: 155–169.
- Guerriero, F., and R. Guido. 2011. "Operational Research in the Management of the Operating Theatre: A Survey." *Health Care Management Science* 14: 89–114.
- Gupta, D. 2007. "Surgical Suites' Operations Management." *Production and Operations Management* 16: 689–700.

- Houdenhoven, M. Van, J. M. Van Oostrum, E. W. Hans, G. Wullink, and G. Kazemier. 2007. "Improving Operating Room Efficiency by Applying Bin-packing and Portfolio Techniques to Surgical Case Scheduling." *Anesthesia and Analgesia* 105: 707–714.
- Hulshof, J. H., R. J. Boucherie, E. W. Hans, and J. L. Hurink. 2013. "Tactical Resource Allocation and Elective Patient Admission Planning in Care Processes." *Health Care Management Science*. 152–166.
- Ma, G., and E. Demeulemeester. 2013. "A Multilevel Integrative Approach to Hospital Case Mix and Capacity Planning." *Computers and Operations Research* 40: 2198–2207.
- McManus, M. L., M. C. Long, A. Cooper, J. Mandell, D. M. Berwick, and M. Pagano. 2003. "Variability in Surgical Caseload and Access to Intensive Care Services." *Anesthesiology* 98: 1491–1496.
- Patrick, J., M. L. Puterman, and M. Queyranne. 2008. "Dynamic Multipriority Patient Scheduling for a Diagnostic Resource." *Operations Research* 6: 1507–1525.
- Persson, M., and J. A. Persson. 2009. "Health Economic Modelling to Support Surgery Management at a Swedish Hospital." *Omega* 853–863.
- Shaw, B., and A. H. Marshall. 2007. "Modelling the Flow of Congestive Heart Failure Patients through a Hospital System." *Journal of the Operational Research Society* 58: 212–218.
- Vanberkel, P. T., R. J. Boucherie, E. W. Hans, and J. L. Hurink. 2014. "Optimizing the Strategic Patient Mix Combining Queueing Theory and Dynamic Programming." *Computers and Operations Research* 43: 271–279.
- Vanberkel, P. T., R. J. Boucherie, E. W. Hans, J. L. Hurink, W. A. M. van Lent, and W. H. van Harten. 2011a. "Accounting for Inpatient Wards when Developing Master Surgical Schedules." *Anesthesia and Analgesia* 112: 1472–1479.
- Vanberkel, P. T., R. J. Boucherie, E. W. Hans, J. L. Hurink, W. A. M. van Lent, and W. H. van Harten. 2011b. "An Exact Approach for Relating Recovering Surgical Patient Workload to the Master Surgical Schedule." *Journal of the Operational Research Society* 62: 1851–1860.
- Walton, N. A., D. K. Martin, E. H. Peter, D. M. Pringle, and P. A. Singer. 2007. "Priority Setting and Cardiac Surgery: A Qualitative Case Study." *Health Policy* 3: 444–458.
- Williams, P., G. Tai, and Y. Lei. 2010. "Simulation Based Analysis of Patient Arrival to Health Care Systems and Evaluation of an Operations Improvement Scheme." *Annals of Operations Research* 178: 263–279.

## Appendix 1. An example of computations to get the distribution of occupied beds in the ICU

Let us consider a small example for a given category  $c$  for which we have  $V_c = 2$  and  $L_{ICU}^{\max} = 2$ . We want to compute the probabilities that the number of occupied beds by patients in category  $c$  on day  $t$  equals 0, 1 and 2 ( $n = 0, 1, 2$ ). Preliminary computations involve the determination of probabilities  $P_c(n, t-k, k)$ , for all  $n = 0, \dots, V_c$  and  $k = 0, \dots, L_{ICU}^{\max}$ , using Equations (11) and (12). The computations of  $R_c(n, t)$  are presented and developed in Table A1.

We consider first the probability of having zero patients in the ICU on day  $t$ ,  $R_c(0, t)$ , which is given by  $TP_c(0, t, 0)$  (see Equation (13)). The corresponding event (a) to this situation is detailed in Table A1 and consists in having zero patients in the ICU on day  $t-1$  staying at least one more day (which is written as  $TP_c(0, t-1, 1)$ ) and having zero patients undergoing surgery on day  $t$  with a zero-day stay in the ICU. Thus, we have  $TP_c(0, t, 0) = TP_c(0, t-1, 1) \cdot P_c(0, t, 0)$  as described by Equation (14). The probability  $TP_c(0, t-1, 1)$  itself corresponds to the probability of having zero patients in the ICU on day  $t-2$  staying at least two more days multiplied by the probability of having zero patients receiving surgery on day  $t-1$  with a one-day stay in the ICU. We thus have  $TP_c(0, t-1, 1) = TP_c(0, t-2, 2) \cdot P_c(0, t-1, 1)$ . Since patients do not stay in the ICU more than two days ( $L_{ICU}^{\max} = 2$ ), the number of patients in the ICU on day  $t-2$  only comes from surgeries that took place on day  $t-2$ . As described by Equation (15), we thus have  $TP_c(0, t-2, 2) = P_c(0, t-2, 2)$ . We indicated this equality in Table A1, by putting the first letter of the notation  $TP_c()$  between brackets. Thus,  $(TP_c(0, t-2, 2))$  means that  $(TP_c(0, t-2, 2)) = P_c(0, t-2, 2)$ . We finally have  $TP_c(0, t, 0) = P_c(0, t, 0) \cdot P_c(0, t-1, 1) \cdot P_c(0, t-2, 2)$  which means that in order to have zero patients in the ICU on day  $t$ , we must have zero patients left from the surgeries that took place on days  $t-1$  and  $t-2$ .

Turning to  $n = 1$ , one occupied bed in the ICU on day  $t$  results from the combination of events (c) and (d), where (c) corresponds to zero patients in the ICU on day  $t-1$  staying at least one more day (with probability  $TP_c(0, t-1, 1)$ ) and one patient receiving surgery on day  $t$  with a zero-day stay in the ICU (with probability  $P_c(1, t, 0)$ ). The advantage of the recursive equation (14) is the spare of some computations as  $TP_c(0, t-1, 1)$  was computed when we considered  $n = 0$ . Event (d) corresponds to the situation where there is one patient in the ICU on day  $t-1$  staying at least one more day (with probability  $TP_c(1, t-1, 1)$ ), and zero patients receiving surgery on day  $t$ . Using Equation (14), the probability  $TP_c(1, t-1, 1)$  is also decomposed as shown in Table A1 in the rows corresponding to events (e) and (f). Finally, to get one occupied bed in the ICU on day  $t$ , we must have: zero patients in the ICU on day  $t-1$  and one patient receiving surgery on day  $t$ ; or one patient in the ICU on day  $t-1$  staying at least one more day and zero patients receiving surgery on day  $t$ .

Having  $n = 2$  patients in the ICU on day  $t$  results from the combination of events (g), (h) and (i), with (g): zero patients in the ICU on day  $t-1$  with probability  $TP_c(0, t-1, 1)$  (previously computed) and two patients receiving surgery on day  $t$  or (h): one patient in the ICU on day  $t-1$  staying at least one more day with probability  $TP_c(1, t-1, 1)$  (previously computed) and one patient receiving surgery on day  $t$  or (i): two patients in the ICU on day  $t-1$  staying at least one more day with probability  $TP_c(2, t-1, 1)$  and zero patients receiving surgery on day  $t$ .

To illustrate Step 4, let us consider three categories of patients ( $N = 3$ ) and  $n = 0, 1$ . Starting with  $n = 0$ , we want to compute  $R(0, t)$  which, from Equation (16), is equal to  $TR(0, t, 3)$ . For  $c' = 1$  we have  $TR(0, t, 1) = R_1(0, t)$  as stated by Equation (16). Probability  $R_1(0, t)$  was computed in Step 3. For  $c' = 2$ , we have  $TR(0, t, 2) = TR(0, t, 1) \cdot R_2(0, t)$ , meaning that we get zero patients in the first two categories only when there are zero patients in category 1 and zero patients in category 2. Note that  $TR(0, t, 1)$  was determined for  $c' = 1$ . For  $c' = 3$ , we have  $TR(0, t, 3) = R(0, t) = TR(0, t, 2) \cdot R_3(0, t)$  meaning that we get zero patients in the ICU when there are zero patients in both categories 1 and 2, and zero patients in category 3. Note that  $TR(0, t, 2)$  was computed for  $c' = 2$ .

Turning to  $n = 1$ , we want to determine  $R(1, t) = TR(1, t, 3)$ . Following Equation (16), for  $c' = 1$ , we have  $TR(1, t, 1) = R_1(1, t)$ . For  $c' = 2$ , we have  $TR(1, t, 2) = TR(1, t, 1) \cdot R_2(0, t) + TR(0, t, 1) \cdot R_2(1, t)$ , which means that having in the ICU one patient in one

Table A1. Example of computations to obtain the distribution of occupied beds in the ICU.

| $n$   | $t-2$                                   | $t-1$ | $t$ | Event |
|---|---|-------|-----|-------|
| 0   | $TP_c(0,t,0)$                           |       |     |       |
|   | $TP_c(0,t-1,1) \times P_c(0,t,0)$       |       |     | (a)   |
|   | $(T)P_c(0,t-2,2) \times P_c(0,t-1,1)$   |       |     | (b)   |
| 1   | $TP_c(1,t,0)$                           |       |     |       |
|   | $TP_c(0,t-1,1) \times P_c(1,t,0)$       |       |     | (c)   |
|   | $+ TP_c(1,t-1,1) \times P_c(0,t,0)$     |       |     | (d)   |
|   | $(T)P_c(0,t-2,2) \times P_c(1,t-1,1)$   |       |     | (e)   |
|   | $+ (T)P_c(1,t-2,2) \times P_c(0,t-1,1)$ |       |     | (f)   |
| 2   | $TP_c(2,t,0)$                           |       |     |       |
|   | $TP_c(0,t-1,1) \times P_c(2,t,0)$       |       |     | (g)   |
|   | $TP_c(1,t-1,1) \times P_c(1,t,0)$       |       |     | (h)   |
|   | $TP_c(2,t-1,1) \times P_c(0,t,0)$       |       |     | (i)   |
|   | $(T)P_c(0,t-2,2) \times P_c(2,t-1,1)$   |       |     | (j)   |
|   | $(T)P_c(1,t-2,2) \times P_c(2,t-1,1)$   |       |     | (k)   |
|   | $(T)P_c(2,t-2,2) \times P_c(2,t-1,1)$   |       |     | (l)   |
| $P_c(n,t,j)$ : proba. of $n$ patients operated on $t$ in the ICU $j$ days after surgery       |   |       |     |       |
| $TP_c(n,t,j)$ : proba. of $n$ patients in the ICU on $t$ who will stay at least $j$ more days |   |       |     |       |
| $(T)P_c()$ means that $(T)P_c()=P_c()$ .  |   |       |     |       |

of the first two categories results from having one patient in the previous category and zero patients in category 2 or having zero patients in the previous category and one patient in category 2. Note that  $TR(1, t, 1)$  was determined for  $n = 1$  and  $c' = 1$  and  $TR(0, t, 1)$  was also determined when we considered  $n = 0$  and  $c' = 1$ . For  $c' = 3$ , we get  $R(1, t) = TR(1, t, 3) = TR(1, t, 2) \cdot R_3(0, t) + TR(0, t, 2) \cdot R_3(1, t)$ , meaning that we have one patient in the ICU when there is one patient in the ICU in either category 1 or 2 and zero patients in category 3, or when there are zero patients in the first two categories and one patient in category 3. Again,  $TR(1, t, 2)$  and  $TR(0, 2)$  were determined previously. This example illustrates the spare of computations that the recursive method allows for, as we do not need to enumerate all combinations.