Dear Author,

Here are the proofs of your article.

- You can submit your corrections **online**, via **e-mail** or by **fax**.
- For **online** submission please insert your corrections in the online correction form. Always indicate the line number to which the correction refers.
- You can also insert your corrections in the proof PDF and **email** the annotated PDF.
- For fax submission, please ensure that your corrections are clearly legible. Use a fine black pen and write the correction in the margin, not too close to the edge of the page.
- Remember to note the **journal title**, **article number**, and **your name** when sending your response via e-mail or fax.
- **Check** the metadata sheet to make sure that the header information, especially author names and the corresponding affiliations are correctly shown.
- **Check** the questions that may have arisen during copy editing and insert your answers/ corrections.
- **Check** that the text is complete and that all figures, tables and their legends are included. Also check the accuracy of special characters, equations, and electronic supplementary material if applicable. If necessary refer to the *Edited manuscript*.
- The publication of inaccurate data such as dosages and units can have serious consequences. Please take particular care that all such details are correct.
- Please **do not** make changes that involve only matters of style. We have generally introduced forms that follow the journal's style.
  Substantial changes in content, e.g., new results, corrected values, title and authorship are not allowed without the approval of the responsible editor. In such a case, please contact the Editorial Office and return his/her consent together with the proof.
- If we do not receive your corrections **within 48 hours**, we will send you a reminder.
- Your article will be published **Online First** approximately one week after receipt of your corrected proofs. This is the **official first publication** citable with the DOI. **Further changes are, therefore, not possible.**
- The **printed version** will follow in a forthcoming issue.

Please note

After online publication, subscribers (personal/institutional) to this journal will have access to the complete article via the DOI using the URL: http://dx.doi.org/[DOI].
If you would like to know when your article has been published online, take advantage of our free alert service. For registration and further information go to: http://www.link.springer.com.

Due to the electronic nature of the procedure, the manuscript and the original figures will only be returned to you on special request. When you return your corrections, please inform us if you would like to have these documents returned.

Metadata of the article that will be visualized in OnlineFirst

Please note: Images will appear in color online but will be printed in black and white.

ArticleTitle	Solving the operating room scheduling problem with prioritized lists of patients	
Article Sub-Title		
Article CopyRight	Springer Science+Business Media New York (This will be the copyright line in the final PDF)	
Journal Name	Annals of Operations Research	
Corresponding Author	Family Name	**Rey**
	Particle	
	Given Name	**Pablo A.**
	Suffix	
	Division	Departamento de Ingeniería Industrial
	Organization	Universidad de Chile
	Address	Av. República 701, Santiago, Chile
	Email	pablo.rey.cl@gmail.com
Author	Family Name	**Durán**
	Particle	
	Given Name	**Guillermo**
	Suffix	
	Division	Departamento de Ingeniería Industrial
	Organization	Universidad de Chile
	Address	Av. República 701, Santiago, Chile
	Division	Departamento de Matemática, Instituto de Cálculo
	Organization	Universidad de Buenos Aires
	Address	Buenos Aires, Argentina
	Division	
	Organization	CONICET
	Address	Buenos Aires, Argentina
	Email	gduran@dm.uba.ar
Author	Family Name	**Wolff**
	Particle	
	Given Name	**Patricio**
	Suffix	
	Division	Departamento de Ingeniería Industrial
	Organization	Universidad de Chile
	Address	Av. República 701, Santiago, Chile
	Email	wolffpatricio@gmail.com

Schedule	Received
	Revised
	Accepted

Abstract	The scheduling of surgical interventions directly impacts the number of patients that can be treated with given operating room resources. Medical centres often do not respond satisfactorily to the demand for interventions, and the shortcomings of traditional manual scheduling approaches contribute to the growth of waiting lists. In addition to the timetabling aspect, operating room scheduling methods must determine the order in which patients should be treated as a function of their relative priorities. This paper develops and compares two optimization models and two algorithms for scheduling interventions over a defined period that satisfy patient priority criteria. The four mathematical methods were studied under a range of different scenarios using real data from a public hospital in Chile. Improvements in operating room utilization rates using the proposed formulations ranged from 10 to 15 % over the current manual techniques, but the choice of method in any given real application will depend on the scenarios likely to be encountered.
Keywords (separated by '-')	Operating room - Integer programming - Exhaustive enumeration - Patient priority
Footnote Information	

Ann Oper Res
DOI 10.1007/s10479-016-2172-x

S.I.: CLAIO 2014

Solving the operating room scheduling problem with prioritized lists of patients

Guillermo Durán[1,2,3] · **Pablo A. Rey**[1] · **Patricio Wolff**[1]

Abstract The scheduling of surgical interventions directly impacts the number of patients that can be treated with given operating room resources. Medical centres often do not respond satisfactorily to the demand for interventions, and the shortcomings of traditional manual scheduling approaches contribute to the growth of waiting lists. In addition to the timetabling aspect, operating room scheduling methods must determine the order in which patients should be treated as a function of their relative priorities. This paper develops and compares two optimization models and two algorithms for scheduling interventions over a defined period that satisfy patient priority criteria. The four mathematical methods were studied under a range of different scenarios using real data from a public hospital in Chile. Improvements in operating room utilization rates using the proposed formulations ranged from 10 to 15 % over the current manual techniques, but the choice of method in any given real application will depend on the scenarios likely to be encountered.

1

2

3

4

Keywords Operating room · Integer programming · Exhaustive enumeration · Patient priority

✉ Pablo A. Rey
pablo.rey.cl@gmail.com

Guillermo Durán
gduran@dm.uba.ar

Patricio Wolff
wolffpatricio@gmail.com

[1] Departamento de Ingeniería Industrial, Universidad de Chile, Av. República 701, Santiago, Chile

[2] Departamento de Matemática, Instituto de Cálculo, Universidad de Buenos Aires, Buenos Aires, Argentina

[3] CONICET, Buenos Aires, Argentina

1 Introduction

A surgery operating room (OR) is a space designed and equipped specially for carrying out anaesthetic procedures and surgical interventions. Scheduling operations at a medical centre is a highly complex process (Santibáñez et al. 2007). The availability of operating room resources and how they are scheduled has a direct impact on the number of patients that can be treated as well as patient waiting times (Reveco and Weber 2011). Public health systems are frequently unable to immediately satisfy total demand for elective (non-urgent) surgery, with the result that waiting lists are often lengthy. Decisions as to which set of patients should be operated on and when depend on available resources, waiting times and various biomedical criteria [disease progression, pain or dysfunction and disability (Testi et al. 2006)].

The main objective of the present article is to develop and propose methods for determining operating room schedules at a public children's hospital in Chile.

The scheduling process must take into account a series of considerations relating to the characteristics of both the patients waiting for surgery and the hospital. The solutions generated provide the basis of an operating room resource use plan for a given period that includes the specification of the order of patient interventions to be performed.

The purpose of the proposed methods is to optimize the use of the operating room resource while complying with the relative priority ordering of the patients to be operated on. The systematization of intervention scheduling and patient prioritization by these models and algorithms also affords an opportunity to improve transparency and achieve greater equity in the assignment of surgery resources (Santibáñez et al. 2007).

2 Literature review

The operating room scheduling process involves a number of complications arising from the large number of factors that must be taken into account (Cardoen et al. 2009; Jebali et al. 2006). Among these factors are the many types of surgical interventions performed, operation duration times, relative patient priority, hospital capacity, length of hospital stay, operating room hours and patient ages (Dexter and Macario 2002).

The operating room planning problem, and more generally, operating theatre planning (which includes recovery rooms as well as operating rooms), has been the subject of many studies. Two thorough surveys of this literature have been recently published (Cardoen et al. 2010; Guerriero and Guido 2011).

These planning processes are often divided into three stages. Gupta et al. (2007), for example, in a general context, separates the planning decisions into capacity allocation, booking control and surgery sequencing levels, and proposes dynamic programming models for each one.

A similar threefold division more closely approximating the process in the present case study is the one discussed in Guerriero and Guido (2011), Marques et al. (2011) and Santibáñez et al. (2007). At the first or strategic level, known as case mix planning, the available time is distributed in aggregate terms among the various surgical specialties. At the second or tactical level, known as the block scheduling problem, a master surgery schedule (MSS) is constructed to assign specific time slots and operating rooms for a given period to each specialty based on the resources defined at the first level.

Finally, at the third or operating level (the level at which the problem in this paper arises), the sessions and times for the interventions, surgical specialists involved and other operating decisions relating to a short time horizon are defined in accordance with the MSS just defined. Some authors further divide this level into two problems. The first one (advanced scheduling) assigns patients to sessions and operating rooms, followed by the second one (allocation scheduling) in which patients are sequenced within the sessions. These problems have been studied in the literature both separately and integratedly (Cardoen et al. 2010).

In the present article the problem addressed occurs at the advanced scheduling stage. The allocation scheduling problem is solved directly by the application of rules for sequencing the patients assigned to a given session. These rules are set out in Sect. 4.

A number of studies propose integer programming approaches to tackle variations on the advanced scheduling problem. The first model developed here below (Sect. 4.1) is similar in terms of the type of variables and constraints to those presented in Guinet and Chaabane (2003) and Marques et al. (2011), although the specific conditions in the latter two cases differ slightly. One such difference is in their objective functions, for whereas Guinet and Chaabane (2003) minimizes the assignment costs, which include waiting costs (estimated as the hospitalization cost for the waiting time) plus possible overtime costs, Marques et al. (2011) maximizes the operating room time use. Unlike either of these works, however, in the first model and the algorithms of the present study (Sects. 4.3, 4.4) the objective is to satisfy the patient priorities "strictly", as is explained in what follows.

2.1 Prioritization of patients

Prioritizing patients on a waiting list for elective surgery has been extensively studied (Hilkhuysen et al. 2005; MacCormick et al. 2003; Min and Yih 2010; Mullen et al. 2003; Oudhoff et al. 2007; Testi et al. 2006; Valente et al. 2009). The importance of this task is stressed in Oudhoff et al. (2007) due to its effectiveness in reducing the negative consequences of long waits for certain operations. Although there is little evidence on what is the most appropriate ethical basis for patient prioritization, there is a general consensus on the central importance of including clinical criteria (Siciliani and Hurst 2005). In recent decades, countries such as Australia, Canada, Wales, Italy and New Zealand have implemented different prioritization systems for surgery patients (Hadorn and Holmes 1997; Testi et al. 2006; Noseworthy et al. 2003; Russell et al. 2003). A critical analysis of various systems currently used in practice is found in Mullen et al. (2003).

In Min and Yih (2010), the authors consider patient priority in an assignment based on the trade-off between the costs of performing an operation and the costs of postponing it. A prioritization method based on the available patient information is essential to the operating room scheduling process for determining the relative positions of patients on the waiting list. In the present study we use the approach presented in Testi et al. (2006), where the authors demonstrate the advantages of using a measure known as need-adjusted-waiting-day (*NAWD*). To implement this method, two factors must first be established: the patients' biomedical category, which is based on maximum waiting time for the required intervention, and the number of waiting days between the day the patient was diagnosed and the scheduled day of the intervention. The *NAWD* for each patient is then calculated according to the following formula:

$$NAWD_i = Pond_i \cdot te_i, \quad Pond_i \in \{1, 2, 4, 12, 48\} \tag{1}$$

where te_i is the number of waiting days of patient i and $Pond_i$ is a factor related to the patient's biomedical category. The more urgent is the diagnosis, the greater should be the value of $Pond_i$. Once the *NAWD* values have been calculated for each patient they are sorted in decreasing order, the resulting patient ordering then constituting the prioritized waiting list.

The patient order on a prioritized waiting list can be used to define the weights or coefficients of a weighted sum representing the "cost" or "benefit" of a given priority plan. This is done in Min and Yih (2010) for defining penalties per time unit of postponement of a patient's operation. The penalties decline as patients move up on the prioritized waiting list. In the case that inspired the present study, the prioritized waiting list is satisfied in "strict" order, meaning that preference is given to operating on the first patient on the list before any of the others. In this case, the weighted sum used as the objective is not a "cost" but rather serves as a mechanism for obtaining the appropriate selection by strictly satisfying the waiting list priorities.

3 The surgery scheduling problem in public hospitals in Chile

Chile's hospitals constitute part of a number of different entities making up the country's network of public health facilities. They are classified by the complexity of the services they offer. High and medium complexity hospitals have operating rooms where both urgent and elective surgeries are performed. Each facility is generally devoted to certain medical specialties and carries out elective surgery interventions in operating rooms at scheduled hours. The assignment of operations is based on historical factors and the availability of medical personnel. At the hospitals investigated for this study, an available operating room for a surgical specialty is considered to include the physical space itself, medical supplies, anaesthetists, medical equipment and a medical team. Scheduling of the operating rooms must specify the patients to be operated on during the assigned time blocks, the order in which interventions are to be performed, the doctors performing the operations and, where there are multiple operating rooms, the one in which each operation is to be performed.

The planning process generally used in Chile is similar to the three-stage "Surgical Planning Process" described in Santibáñez et al. (2007). The first stage, denoted *surgical mix*, defines the share of OR time assigned to each specialty. The second stage, called *block scheduling*, assigns time blocks or morning and afternoon sessions for each specialized OR. Finally, the third stage—and the one we will focus on here—consists in assigning patients and scheduling the corresponding interventions and necessary resources.

The design considerations for the proposed scheduling methods are presented below.

1. The scheduling period is a work week (five working days).
2. Operations are performed Monday to Friday, with each day divided into a morning session of 8 am to 1 pm and an afternoon session of 2 pm to 5 pm.
3. The hospital contains a specific set of operating rooms. Each operating room is unique and specially adapted for certain types of interventions.
4. Each specialty is assigned one, more than one or no operating room per session for the scheduling of its interventions. In the hospitals studied, no specialty was assigned more than two operating rooms simultaneously. These assignments are decided in the stage previous to the block scheduling.
5. There is a prioritized patient list for operations by specialty. Patient priority is defined as described in Sect. 2.1.
   The relative priorities of the patients are determined on the basis of medical and waiting-time factors. Patients whose characteristics are such that they cannot be assigned within the scheduling period must be excluded from the set of scheduled patients.

6. Each surgical specialty manages its patients independently, implying the number of waiting lists is equal to the number of specialties. The hospital has a given group of doctors to perform the scheduled operations. Each doctor is a specialist in at least one of the specialties performed at the hospital. The hospital has detailed information on the hours of availability of its personnel.
7. A patient cannot be operated on more than once in the same scheduling period. This is a limitation of the model but is consistent with the practices followed at the hospitals studied. For each patient the days and sessions he or she is available for his or her operation are known.
8. Each procedure requires two doctors. If a patient specifies a particular doctor, shifts must be sought in which that doctor is accompanied by a secondary doctor.
9. High priority patients are assigned preferably to the early part of the scheduling period (i.e., early in the week).
10. Provision is made for “special” patients, who may be given such status if a complicating factor (e.g., latex allergy, under 1 year of age) requires their operation to be scheduled in the morning session.
11. Time extensions to the regular session hours are allowed for scheduling operating rooms. This improves efficiency but implies a commitment by the hospital to cover the extra costs involved. These scheduled extensions have a maximum duration of 10 min.

4 Proposed scheduling methods

In this section we develop four methods for solving the problem of determining which patients will be operated on and when their operations will be performed. The indicator to be optimized is the capacity utilization of the operating room, defined as the percentage of available operating room time that is effectively scheduled (the exact formula is given in Sect. 5 below).

In more precise terms, the problem is to determine which patients will be operated on and in which session. Once this is decided by either of the methods, the order of these operations within each session is specified by the following two conditions:

- If in a given morning session a special patient is assigned, that patient is scheduled to be operated on first. As noted above in the design considerations, special patients can only be assigned to a morning session.
- All of the other patients are then ordered by age (youngest first).

The four proposed methods consist of an integer linear programming model, a variant on that model, an algorithm based on a feasibility model and a constructive algorithm. They are described individually in the following subsections.

4.1 Integer linear programming model (ILP1)

The first of the two integer linear programming models, designated ILP1, assigns operations to patients. The variables expressing this principal assignment are binary, and for each patient indicate a specific OR in an appropriate time block and the doctors who will perform the operation. Another set of variables penalizes operation duration time extensions, thus modelling the sessions as soft restrictions. The objective is to obtain an assignment that maximizes compliance with patient priority in the strict sense.

The various indexes, parameters, variables and constraints of the integer linear programming model and its objective function are presented below.

4.1.1 Indexes and parameters

The indexes of the ILP model are:

- $doc1$: principal doctors
- $doc2$: secondary doctors
- i: session
- p: patient
- pab: operating rooms

The parameters of the ILP model are:

- Dur_i: length of session in minutes
- ST_i: maximum operation duration time extension in minutes
- $Dura_p$: duration of operation to be performed on patient p in minutes. Includes preparation and cleanup as well as actual surgery time.
- Pri_p: patient priority
  The values of Pri_p are determined by the lexicographic rule that "operating on a given patient is preferred to operating on all other lower-priority patients combined". If integers are used for the Pri_p term, it will grow exponentially as priority increases. Thus, if N is the number of patients, the value of the term for patient p on the priority list will be

$$Pri_p = 2^{N-p} \tag{2}$$

  Recall that the list is sorted in decreasing order of the patients' individual *NAWD* values as defined in Sect. 2.1).
- $MN_i = \begin{cases} 1, & \text{if session } i \text{ is a morning session} \\ 0, & \text{otherwise} \end{cases}$
- $ESP_p = \begin{cases} 1, & \text{if patient } p \text{ is a special patient} \\ 0, & \text{otherwise} \end{cases}$
- $Coin_{doc1,doc2} = \begin{cases} 1, \text{ if doctor } doc1 = doc2 \\ 0, \text{ otherwise} \end{cases}$
- $f1_p^{doc1} = \begin{cases} 1, & \text{if doctor } doc1 \text{ can perform the intervention} \\ & \text{on patient } p \\ 0, & \text{otherwise} \end{cases}$
- In the hospitals studied there were differences between the various operating rooms, some of which had special characteristics for particular types of interventions.
  $f_p^{pab} = \begin{cases} 1, & \text{if operation on patient } p \text{ can be carried out in operating room} pab \\ 0, & \text{otherwise} \end{cases}$
- $Asig_p^{doc1} = \begin{cases} 1, & \text{if patient } p \text{ is assigned to } doc1 \\ 0, & \text{otherwise} \end{cases}$
- $d1_i^{doc1} = \begin{cases} 1, & \text{if doctor } doc1 \text{ works on session} i \\ 0, & \text{otherwise} \end{cases}$
- $d2_i^{doc2} = \begin{cases} 1, & \text{if doctor } doc2 \text{ works on session} i \\ 0, & \text{otherwise} \end{cases}$
- $cor_i^{pab} = \begin{cases} 1, & \text{if operating room } pab \text{ is available on session} i \\ 0, & \text{otherwise} \end{cases}$

- $disp_i^p = \begin{cases} 1, & \text{if patient } p \text{ is available for an operation on session} i \\ 0, & \text{otherwise} \end{cases}$

4.1.2 Variables

The decision variable of the ILP model is

$$t_i^{p,doc1,doc2,pab} = \begin{cases} 1, & \text{if patient } p \text{ is operated on by doctors } doc1 \text{ and } doc2 \\ & \text{in operating room } pab \text{ on session} i \\ 0, & \text{otherwise} \end{cases}$$

The variable indicating whether a time extension has been used is

$$x_i = \begin{cases} 1, & \text{if a time extension has been used on session} i \\ 0, & \text{otherwise} \end{cases}$$

Recall that according to design consideration no. 11 (p. 5), the duration of a scheduled extension may not exceed 10 min. The purpose of this variable is to indicate whether or not the scheduling stays within the regular session hours.

4.1.3 Constraints

The constraints on the ILP model are the following:

1. Interventions cannot be scheduled for operating rooms or on sessions for which they are not feasible. The number M_1 must be equal to or greater than the maximum number of operations that are feasible on a working day.

$$\sum_{p,doc1,doc2} t_i^{p,doc1,doc2,pab} \leq M_1 \cdot cor_i^{pab}, \quad \forall pab, i \tag{3}$$

   In this inequality the coefficient M_1 is equal to the ratio of the duration of the longest session and the duration of the shortest intervention. In the cases dealt with for this study, $M_1 = 20$ given that the sessions were 5 h long and the shortest operation was estimated at 15 min.
2. Interventions cannot be scheduled for times when there are no doctors available to perform them. The number M_2 must be greater than the maximum number of operations a doctor can carry out on a working day.

$$\sum_{p,doc2,pab} t_i^{p,doc1,doc2,pab} \leq M_2 \cdot d1_i^{doc1}, \quad \forall doc1, i \tag{4}$$

$$\sum_{p,doc1,pab} t_i^{p,doc1,doc2,pab} \leq M_2 \cdot d2_i^{doc2}, \quad \forall doc2, i \tag{5}$$

   In this inequality the coefficient M_2 takes the same value as coefficient $M_1 = 20$ in constraint set (3) above.
3. The assignment of the same doctor as principal and secondary doctor for a given intervention should be avoided.

$$\sum_{pab,i} Coin_{doc1,doc2} \cdot t_i^{p,doc1,doc2,pab} \leq 0, \quad \forall p, doc1, doc2 \tag{6}$$

4. The scheduling of interventions at times when patients are not available should be avoided.

$$\sum_{doc1,doc2,pab} t_i^{p,doc1,doc2,pab} \leq disp_i^p, \quad \forall i, p \tag{7}$$

5. No patient can be operated on more than once over the defined time horizon.

$$\sum_{i,doc1,doc2,pab} t_i^{p,doc1,doc2,pab} \leq 1, \quad \forall p \tag{8}$$

6. If a doctor has been preassigned as principal doctor to an intervention, he or she and no one else must perform it. The parameter $Asig_p^{doc1}$ indicates whether a preassignment exists for a given patient, and if so, identifies the preassigned doctor.

$$\sum_{i,doc2,pab} t_i^{p,doc1,doc2,pab} \geq Asig_p^{doc1}, \quad \forall p, doc1 \tag{9}$$

7. Interventions must not be scheduled for operating rooms that do not have the required characteristics.

$$\sum_{i,doc1,doc2} t_i^{p,doc1,doc2,pab} \leq f_p^{pab}, \quad \forall p, pab \tag{10}$$

8. The schedules for each day must not exceed the maximum time plus the permitted time extension ST_i. If they nevertheless do, the variable x_i is 1.

$$\sum_{p,doc1,doc2,pab} Dura_p \cdot t_i^{p,doc1,doc2,pab} \leq Dur_i + ST_i \cdot x_i, \quad \forall i \tag{11}$$

$$\sum_{p,doc1,doc2,pab} Dura_p \cdot t_i^{p,doc1,doc2,pab} \geq Dur_i \cdot x_i, \quad \forall i \tag{12}$$

9. Special patients must be scheduled as the first patient in the morning, implying that on any given morning no more than one such patient may be assigned. The parameter MN_i indicates the session (morning or afternoon), and since its value is either 1 or 0, no more than one special patient can be scheduled.

$$\sum_{p,doc1,doc2,pab} ESP_p \cdot t_i^{p,doc1,doc2,pab} \leq MN_i, \quad \forall i \tag{13}$$

10. Nature of the variables.

$$x_i \in \{0,1\}\, \forall i$$
$$t_i^{p,doc1,doc2,pab} \in \{0,1\}\, \forall p, doc1, doc2, pab, i$$

4.1.4 Objective function

The objective function of the ILP model incorporates 3 criteria that are set out below.

1. *Compliance with patient priority* During preliminary investigations before the models were developed, discussions were held with doctors at the hospital regarding operating room assignment criteria. The lexicographic rule that best approximates their wishes, already cited here above, is that “operating on a given patient is preferred to operating on

all other lower-priority patients combined". This concept is incorporated into the decision process in the following form:

$$\lambda_F \cdot \sum_{i,p,doc1,doc2,pab} Pri_p \cdot t_i^{p,doc1,doc2,pab} \tag{14}$$

where the value of Pri_p is greater for higher-priority patients. As was also noted earlier, if the rule is strictly observed this value will grow exponentially with the number of patients. λ_F is a weighting factor that sets the importance to be attached to this criterion in the objective function.

2. *Penalty for time extensions* The term x_i is equal to 1 if a time extension is used in session i. It is included in the objective function to impose a penalty for the use of extensions. Its relative weight can be modelled via a parameter λ_H that remains constant for the entire scheduling period as follows:

$$\lambda_H \cdot \sum_i x_i \tag{15}$$

3. *Reward for scheduling urgent patients early in the week* The idea behind this term is to reward operating on higher-priority patients early in the scheduling period. It appears in the objective function in the following form:

$$\lambda_S \cdot \sum_{i,p,doc1,doc2,pab} \delta_i^p \cdot t_i^{p,doc1,doc2,pab} \tag{16}$$

where parameter λ_S models the relative weight to be given to this criterion. The value of δ_i^p is determined as

$$\delta_i^p = M - (p-1) - (i-1) \tag{17}$$

where M must be greater than the maximum number of patients plus the number of sessions (recall that patients are ordered by decreasing priority).

Thus, the complete objective function of the ILP model is written as follows:

$$\begin{aligned} \max \quad & \lambda_F \cdot \sum_{i,p,doc1,doc2,pab} Pri_p \cdot t_i^{p,doc1,doc2,pab} - \lambda_H \cdot \sum_i x_i \\ & + \lambda_S \cdot \sum_{i,p,doc1,doc2,pab} \delta_i^p \cdot t_i^{p,doc1,doc2,pab} \end{aligned} \tag{18}$$

The patient priority weighting factor λ_F is set to 1 in all cases for simplicity. The values for λ_H and λ_S were chosen solely as a function of the priority value Pri_p of the highest priority patient, without regard for patient numbers.

To determine appropriate values for these parameters, we performed a sensitivity analysis on various different possibilities. As an example, consider an instance similar to our real case but with a reduced waiting list of 50 patients. The solution generated by the model always assigns the first 20 patients but never the 21st, the next assignment varying with the particular combination of λ_H and λ_S values as shown in Table 1. To satisfy the compliance with patient priority criterion stated above, the appropriate parameter combinations are those for which the next assignment is the closest one beyond the 21st patient. As can be seen, combinations in the upper right-hand entries of the table all result in the assignment of the 22nd patient, the closest one possible, and are therefore the preferred parameter values.

Recall that according to the definition of the problem, the following are preferred: (1) high values for λ_H, to avoid as much as possible the use of scheduled extensions; (2) low

Table 1 First patient assigned after 20th patient according to priorities in ILP1 solution for different values of λ_H and λ_S (instance with 50 patients)

λ_S	λ_H									
	Pri_1	Pri_5	Pri_{10}	Pri_{20}	Pri_{24}	Pri_{25}	Pri_{30}	Pri_{40}	Pri_{50}	Pri_{75}
Pri_{50}	31	31	31	31	31	22	22	22	22	22
Pri_{45}	31	31	31	31	31	22	22	22	22	22
Pri_{40}	31	31	31	31	31	22	22	22	22	22
Pri_{37}	34	34	34	34	34	22	22	22	22	22
Pri_{35}	34	34	34	34	34	22	22	22	22	22
Pri_{25}	34	34	34	34	34	37	37	37	37	37

penalties, in order to schedule higher priority patients earlier given that the urgency for a patient is already incorporated in Pri_p. This last objective is included only to obtain a better schedule given the patient assignment.

The results obtained with other instances were similar. The general rule decided upon for the parameters is the following:

$$\lambda_H = Pri_{\lfloor N/2 \rfloor} \quad \text{and} \quad \lambda_S = Pri_{\lfloor 3N/4 \rfloor} \tag{19}$$

where N is the number of patients on the waiting list. For the example just considered above, the rule generates $\lambda_H = Pri_{25}$ and $\lambda_S = Pri_{37}$.

The magnitudes of some of the terms in the objective function seem at first glance not to be comparable. When tests were run leaving only the patient priority criterion in the objective function and incorporating the other two criteria as constraints with an adjustable parameter, the results turned out to be very similar to those obtained with the version presented above. Due to the construction of λ_H and λ_S, it was possible to make the objective function terms more comparable.

4.2 Variant of the integer linear programming model (ILP2)

We now present integer linear programming model ILP2, a variant of ILP1 in which the treatment of patient priority is modified. In this version, Pri_p is replaced by new weights calculated for patient assignment that retain the property of being greater for higher-priority patients. The motive is to avoid the problem noted above of the exponential growth of Pri_p values as patient priority increases. The weight function is written as follows:

$$w_p = \alpha^{CAT_p} \cdot (1/Q_p) \tag{20}$$

where $CAT_p \in \{1, 2, 3, 4, 5\}$ and depends on both the patient's waiting time and his or her assigned category, the latter determined by the diagnosis and the seriousness of the medical condition on a scale of decreasing importance from A to E. The actual values of CAT_p for each waiting time and category are decided by the doctors and set out here in Table 2.

As for α, it is the scale factor of CAT_p. After various tests were carried out, it was decided to use the values $\alpha = 2$ and $\alpha = 5$.

Finally, $Q_p \in \{1, 2, \ldots, 10\}$ is a proportion of the duration of the operation to be performed on patient p. This time period is discretized in integers of 1–10, where 1 is assigned to the longest procedure and 10 to the shortest. The idea behind this coefficient, based on

Table 2 Variant ILP model categories by patient waiting time

	<1 week	<1 month	<3 months	<6 months	≥6 months
A	3	4	5	5	5
B	2	3	4	5	5
C	1	2	3	4	5
D	1	1	2	3	4
E	1	1	1	2	3

suggestions by the doctors, is that if the value of α^{CAT_p} turns out to be the same for more than one intervention, greater priority is assigned to the longer ones.

With this alternative version of patient priority, the objective function of the variant ILP model is

$$\max \quad \lambda_F \cdot \sum_{i,p,doc1,doc2,pab} w_p \cdot t_i^{p,doc1,doc2,pab} - \lambda_H \cdot \sum_i x_i$$
$$+ \lambda_S \cdot \sum_{i,p,doc1,doc2,pab} \delta_i^p \cdot t_i^{p,doc1,doc2,pab} \tag{21}$$

In this implementation it was decided for simplicity to set $\lambda_F = 1$. After a number of tests, the values chosen for the other parameters were

$$\lambda_H = \frac{\max(w_p)}{2} \text{ and } \lambda_S = \frac{1}{\lambda_H} \tag{22}$$

4.3 IP feasibility model algorithm

The underlying approach of the algorithm based on a feasibility model, hereafter simply "feasibility model algorithm" (IPFA), is to divide the problem into two parts. The first part solves the assignment problem of deciding which patients should be operated on over the 1-week scheduling period while the second part solves the timetable problem of determining when (that is, on what day) their operations should be carried out. The two parts are described below.

1. *Assignment problem* The "who to operate on" problem is solved by a binary tree algorithm that runs a feasibility IP model for each patient in the order of priority to determine whether he or she can be assigned or not. This implies that the model is executed n times. A flow diagram of the algorithm is shown in Fig. 1.
   The IP model itself is just an adaptation of the ILP1 model discussed above. For each patient p, the patients not chosen in the previous steps are eliminated and constraint 8 is modified to become
   $$\sum_{i,doc1,doc2,pab} t_i^{p',doc1,doc2,pab} = 1, \tag{8'}$$
   for each patient $p' < p$ who has been previously chosen. This forces the procedure to add feasible patients as it progresses while maintaining the assignment of those with greater priority who have already been chosen.
   The objective function consists simply in maximizing the sum of the variables corresponding to the current patient. If this maximum value is 1, the current patient is included.

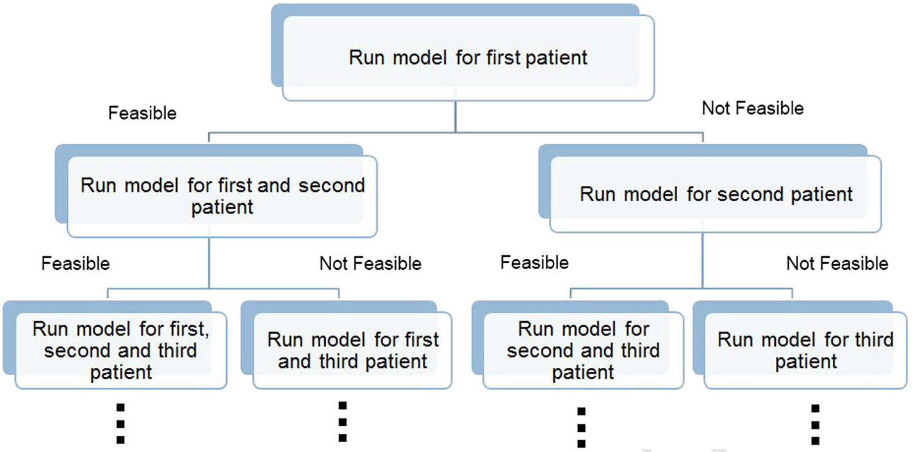


Fig. 1 Search tree for the assignment problem

When the algorithm terminates, the complete list of patients that can be feasibly assigned while satisfying the priority criterion will have been determined. Since the feasibility search proceeds in the same order as the priority criterion results, the patient assignment strictly satisfies that criterion. In other words, an assigned patient will not conflict with one having higher priority because the fact that the latter was not assigned previously means it was not feasible to do so and not that the method simply chose a patient having lower priority.

2. *Timetable problem* Once the patient assignment problem has been solved, the problem of when to operate can be dealt with by an ILP model adapted from ILP1, the first of our four methods. Since the list of patients to be operated on has already been decided, this model will include no terms for patient priority. The objective function will then take the following form:

$$\max \quad -\lambda_H \cdot \sum_i x_i + \lambda_S \cdot \sum_{i,doc1,doc2,pab,p\in A} \delta_i^p \cdot t_i^{p,doc1,doc2,pab} \tag{23}$$

where A is the set of patients assigned by the binary tree algorithm in the assignment problem. For simplicity, the value of λ_S was set to 1. The value of λ_H was chosen so as to be equal to the largest value that can be taken by the other term in the objective function. The formula for the term is

$$\lambda_H = \sum_{i,p\in A} \delta_i^p \tag{24}$$

This value depends on $|A|$, the number of patients assigned in the "Assignment problem", and the number of sessions. Note that since the assignment problem has already been solved, the $t_i^{p,doc1,doc2,pab}$ variables have only to be considered for patients $p \in A$. The constraints for this model are the same as the ones in ILP1 except for (8), which is simply changed an equality so that the scheduling of the patients chosen in the Assignment problem stage is ensured. Thus,

$$\sum_{i,doc1,doc2,pab} t_i^{p,doc1,doc2,pab} = 1, \quad \forall p \in A.$$

Finally, the assignment obtained upon including the last patient in the "Assignment problem" can be used as an initial solution to accelerate the solution of the "Timetable problem".

4.4 Constructive algorithm (CA)

Another approach is to develop a constructive algorithm that finds feasible solutions based on assignment rules defined by the hospital. The solutions can then be evaluated to determine which is the most suitable. This idea is captured by the algorithm set out below, which generates partial patient assignments constituting feasible schedules but which could potentially include more patients. Thus, the patients are visited by the CA in order of priority and at each iteration, all non-dominated partial feasible schedules generated up to that point considering the current patient plus all those with greater priority are stored in the stack. The algorithm's steps are as follows:

1. *Preprocess* This step generates lists of sessions in which a patient can be operated on. Taken into consideration are the operation duration time (which must not extend beyond the length of the session), the availability of a doctor who can perform the operation, and whether the operation is "special" (in which case, as noted earlier, it cannot be performed in the afternoon).
2. *Construction of feasible subassignments* In this step a method is used to construct feasible combinations of the assignments determined in the previous step. The criteria for these combinations are the feasibility of each added patient, that no more than two special operations can be performed in a single morning session and that the corresponding operation times cannot add up to more than the accumulated times of the corresponding sessions.
   The technique consists in generating arrays using a two-input stack to test all possible combinations of feasible patient assignments for satisfaction of the above criteria. Thus, an array is removed from the stack and the feasibility of adding a feasible assignment from the next patient to it is tested. If a feasible combination results, that assignment is added to the array which is then returned to the stack. If, however, the combination is not feasible, another assignment from the next patient is tested. By construction, each of the arrays in the stack is a feasible operation schedule. When this procedure terminates, the stack will have combinations of patients assigned to sessions with the maximum number of feasible patients while also complying with their priority levels.
3. *Dominance* The results of the previous step may provide more than one feasible solution. In this step, the solutions are evaluated on the twin criteria that the younger is the patient, the greater is the priority for operating in the morning, and more urgent operations are scheduled where possible early in the week. The result of this step is a single feasible solution that is better than the other combinations. Thus, we say that the *worst* solution is *dominated* and is eliminated from the stack, leaving only the dominant partial solution.
4. *Choosing the solution* Finally, the best solution in the stack is chosen by comparing the solutions' individual objective function values for the ILP1 model (18).

5 Results

In this section we compare the results of the various methods presented above and contrast them with a real-world situation. Each method's performance can be evaluated in different

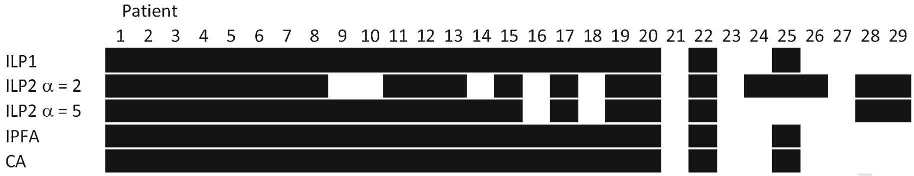


Fig. 2 Example of patients selected by the different methods

scenarios according to the quality of the solution delivered. The quality indicators defined for this purpose are:

1. *Execution time* A routine implementation of a system with an operating room scheduling method such as those developed here must be able to solve the scheduling problem within a time limit ensuring it would be of practical use in the context of the intended application. Since in the present case it would be used at meetings of doctors to define operating schedules, the time limit would have to be no longer than 10 or 15 min.
2. *Patient priority compliance* The specific priority levels established for each patient on the waiting list previous to the running of the model or algorithm must be complied with in the assignments. To compare priority compliance we graph them in Fig. 2 and contrast the sequences of assigned and non-assigned patients.
3. *Operating room capacity utilization* According to the Chilean Ministry of Health, the percentage capacity utilization of an operating room is defined by the following formula:

$$\text{Percentage utilization} = \frac{h_O + h_P}{h_D} \tag{25}$$

   where h_O is the total monthly hours of utilization, h_P the monthly preparatory hours and h_D the total available monthly hours.

For each test scenario the different indicator values for the four methods are calculated and then compared, thus determining the relative quality of each solution. The test scenarios are described below.

5.1 Test scenarios and comparisons

Since the particular characteristics of the different surgical specialties and their corresponding waiting lists will vary from hospital to hospital, the proposed methods were developed to handle a range of scenarios based on actual data supplied by a specific institution. To test and compare the results of the methods we therefore defined a set of such scenarios incorporating variations in three different key characteristics.

1. *Number of patients* This characteristic refers to the number of patients on the waiting list for a given specialty to be scheduled. In a 1-week period, a specialty with 10 available sessions can operate on approximately 25 patients. For testing purposes the patients must be at least double this number so that different assignment alternatives can be considered. The numbers of patients used in the test scenarios were 50, 100 and 200.
2. *Operation duration as a function of session duration* Operation duration times vary from specialty to specialty. Information on operation times were obtained from historical data provided by the hospital. Four different values of operation duration as a percentage of session duration were tested: 12.5, 25, 37.5 and 50 %.

Table 3 Average execution times in minutes by number of patients and model or algorithm

No. of patients	ILP1	ILP2	IPFA	CA
50	0:07	0:08	6:01	0:04
100	0:15	0:15	23:31	0:04
200	0:28	0:29	97:18	0:07

ILP1 refers to the integer linear programming model, ILP2 to the ILP variant model, IPFA to the algorithm using an IP model for feasibility checking and CA to the constructive algorithm

3. *Number of available sessions and/or operating rooms per patient* To ensure the problem is both non-trivial and realistic, a patient must be schedulable for more than one session and/or operating room. The estimates of the number of *a priori* feasible sessions per patient per week were made using three different values for this characteristic: 2.5, 3.5 and 4.5.

As regards the third characteristic, for all instances the conditions at the hospital in our case study were maintained. Thus, the specialties had 2 available operating rooms for 10 sessions distributed in 2 sessions per day across the 5-day week.

All in all, the variations described above define 36 test scenarios, one of them real and the others derived from it, with 3 different numbers of patients, 4 operation duration times, 1 option for the number of available sessions and 3 different numbers of available sessions per patient.

5.2 Results obtained

For all comparison purposes the models and algorithms were run on the same computer, powered by an AMD Phenom II X4 965 3.4GHz processor with 8 GB of RAM. The constructive algorithm was written in Java using NetBeans 6.9.1. The maximum available memory for running the algorithm was set at 6.5 GB. The other 3 methods were modelled in GAMS 23.5 and solved using CPLEX 12.2.

5.2.1 Execution times

The average execution times for the 3 different numbers of patients in the test scenarios are summarized in Table 3 by method.

Note first of all that the average execution times in the table relate to solved scenarios, which were 100 % of all cases with the exception of the constructive algorithm, where the proportion was 60 % (for the remaining 40 % Java stopped the algorithm due to RAM assignment problems).

As regards the actual results, the execution times for ILP1 and ILP2 were quite similar, in both cases depending on the number of patients. The constructive algorithm generally delivered very good run times regardless of the number of patients in every case where it could find a solution.

The feasibility model algorithm, on the hand, took significantly longer than the others to reach a solution, but by construction its solution was the best one from the standpoint of compliance with the priority order of the assigned patients. It delivered the same solution as the constructive algorithm, but as already explained the latter did not execute to completion in every scenario, failing to do so in particular for interventions with operation duration times much shorter than session durations and cases where the number of available sessions per patient was close to the number of sessions per week. For the instance sizes tested, the feasibility model algorithm did not create any memory problems.

5.2.2 *Patient priority compliance*

There is no simple way of representing patient priority compliance, but one alternative is to indicate which patients were assigned under each model or algorithm. This is done in Fig. 2, which depicts the scenario with 50 patients, average operation times 37.5 % of session durations and an *a priori* average of 3.5 available sessions per patient. The patients are ordered in the figure by priority, starting with patient 1, the highest priority, at the far left.

For ILP2, different values of α were studied, two of which ($\alpha = 2$ and $\alpha = 5$) are shown in Fig. 2. As can also be seen, the constructive and feasibility model algorithms and ILP1 model all generated the same results as regards priority compliance. The results for the three methods were better in all scenarios than those delivered by ILP2.

5.2.3 *Operating room capacity utilization*

OR use in percentage terms for the solutions obtained with the different methods in all evaluated instances is shown in Table 4.

The test scenarios demonstrated that when operation duration times are very short relative to session durations (12.5 %) and the number of patients to be scheduled is low, the maximum patient assignment does not cover a large proportion of the available days. As a result, in such cases the operating room capacity utilization rate is low. If the operation times are close to one-half of the session times, the number of combinations that produce good capacity utilization rates declines. For scenarios approximating real ones, all of the models and algorithms achieved capacity utilization rates of 95 %.

5.3 Results obtained in real-world cases

Real data generated by manual methods were studied for general surgery operating rooms nos. 3 and 4 at Luis Calvo Mackenna Hospital in the Chilean capital of Santiago during the second week of August 2009. For reasons of confidentiality the patients are identified only by number. The total operating times in minutes shown in Fig. 3 refer to minutes per indicated session.

The real case schedules generated by the constructive algorithm, ILP1 model and ILP2 for the same week are shown in Fig. 4. All three formulations delivered the same solution.

The real case corresponds to the scenario with 100 patients, 10 available sessions, average operation times 37.5 % of session durations and an *a priori* average of 2.5 available sessions per patient. The same case was also evaluated assuming 3.5 available sessions per patient. This can be done by relaxing some of the rules on assigning doctors to patients. The results obtained are set out in Table 5, which compares the manual method assignment with those generated by the IP1 model (or the CA and IPFA algorithms) for the real case and the relaxed rule case.

They demonstrate that the use of less restrictive policies has a positive impact on operating room capacity utilization and that the proposed mathematical scheduling methods can improve capacity utilization in real-world situations, the increases in the case studied ranging from 10 to 15 % over the existing rate using manual methods.

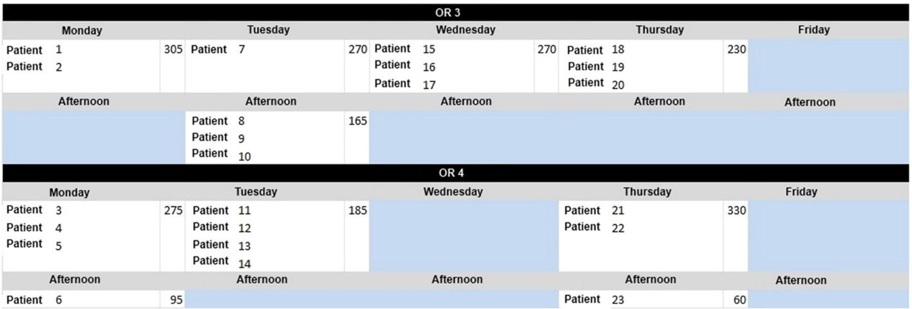
Table 4 Operating room utilization by instance and method

Number of patients	Duration	Sessions available	ILP1	ILP2 $\alpha = 2$	ILP2 $\alpha = 5$	IPFA	CA
50	12.5	2.5	43	43	43	43	
100	12.5	2.5	77	78	78	77	
200	12.5	2.5	98	100	100	100	
50	25	2.5	86	88	85	86	
100	25	2.5	95	96	96	95	
200	25	2.5	62	96	96	62	
50	37.5	2.5	96	99	99	96	96
100	37.5	2.5	95	99	98	95	95
200	37.5	2.5	95	99	99	98	98
50	50	2.5	50	51	51	50	50
100	50	2.5	61	57	57	61	61
200	50	2.5	71	72	74	72	72
50	12.5	3.5	49	50	49	49	
100	12.5	3.5	84	85	84	84	
200	12.5	3.5	99	100	100	100	
50	25	3.5	96	94	94	96	
100	25	3.5	93	96	96	93	
200	25	3.5	94	96	96	95	
50	37.5	3.5	97	97	97	97	
100	37.5	3.5	95	99	98	95	95
200	37.5	3.5	96	99	98	96	96
50	50	3.5	49	52	52	49	49
100	50	3.5	63	57	57	63	63
200	50	3.5	71	74	74	72	72
50	12.5	4.5	51	51	51	51	
100	12.5	4.5	47	48	48	47	
200	12.5	4.5	99	100	100	100	
50	25	4.5	96	96	96	96	
100	25	4.5	74	96	96	74	
200	25	4.5	97	98	98	97	
50	37.5	4.5	99	98	98	99	
100	37.5	4.5	98	99	99	98	
200	37.5	4.5	98	99	99	98	
50	50	4.5	53	53	53	53	
100	50	4.5	66	59	57	66	
200	50	4.5	76	72	71	76	

Note that in some instances ILP2 delivers better utilization levels than the other methods, but in those cases its compliance with patient priority is lower than the others

6 Conclusions

Four alternative mathematical methods were developed for the operating room scheduling problem at public hospitals, all of which delivered good solutions according to criteria set by



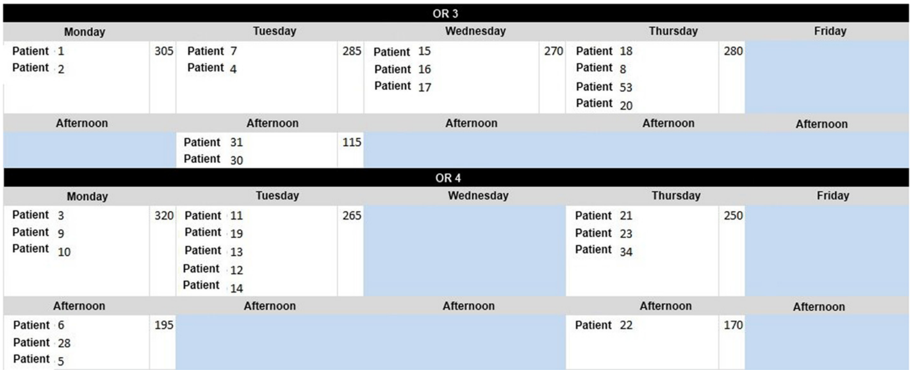
OR 3

	Monday		Tuesday		Wednesday		Thursday		Friday
	Patient 1 Patient 2	305	Patient 7	270	Patient 15 Patient 16 Patient 17	270	Patient 18 Patient 19 Patient 20	230	
	Afternoon		Afternoon		Afternoon		Afternoon		Afternoon
			Patient 8 Patient 9 Patient 10	165					

OR 4

	Monday		Tuesday		Wednesday		Thursday		Friday
	Patient 3 Patient 4 Patient 5	275	Patient 11 Patient 12 Patient 13 Patient 14	185			Patient 21 Patient 22	330	
	Afternoon		Afternoon		Afternoon		Afternoon		Afternoon
	Patient 6	95					Patient 23	60	

Fig. 3 Real case operating room schedule: manual method



OR 3

	Monday		Tuesday		Wednesday		Thursday		Friday
	Patient 1 Patient 2	305	Patient 7 Patient 4	285	Patient 15 Patient 16 Patient 17	270	Patient 18 Patient 8 Patient 53 Patient 20	280	
	Afternoon		Afternoon		Afternoon		Afternoon		Afternoon
			Patient 31 Patient 30	115					

OR 4

	Monday		Tuesday		Wednesday		Thursday		Friday
	Patient 3 Patient 9 Patient 10	320	Patient 11 Patient 19 Patient 13 Patient 12 Patient 14	265			Patient 21 Patient 23 Patient 34	250	
	Afternoon		Afternoon		Afternoon		Afternoon		Afternoon
	Patient 6 Patient 28 Patient 5	195					Patient 22	170	

Fig. 4 Real case operating room schedule: CA, IP1, IPFA

Table 5 Real case models results

	Minutes used	Minutes available	Percent utilization (%)	Improvement (%)
Manual method	2185	2820	77.50	–
Model	2455	2820	87.10	9.60
Model (3.5 sessions)	2624	2820	93.00	15.60

medical personnel. The first method was an integer linear programming model (ILP1), the second a variant on that model (ILP2), the third a constructive algorithm and the fourth an algorithm based on a feasibility model.

The four models and algorithms were tested on 36 scenarios for a public hospital in Chile, one of them real and the others derived from it. The test results were compared for three criteria: execution time, compliance with patient priority levels and operating room capacity utilization. In general terms, the results in the real scenario showed that the methods were able to increase the capacity utilization rates of operating rooms by 10–15 % over the existing rates achieved by manual methods. This points to a significant opportunity for other public hospitals seeking to improve on the schedules obtained using manual approaches. The outcome in any given case will of course depend on the constraints imposed, but the fewer are these restrictions, the greater, obviously, will be the utilization rate.

The execution time results for the four proposed methods' revealed that the constructive and the two ILP models solved the scheduling problem in a matter of seconds whereas the feasibility model algorithm, which must be executed as many times are there are patients, required significantly more time. It is also true, however, that run times for the constructive algorithm might extend beyond what is reasonable in real-world applications if there are many feasible solutions.

As for patient priority compliance, the constructive algorithm delivered the same results as the feasibility model algorithm, which by construction performed best on this criterion. In the case of ILP1, priority compliance was the main problem because the lexicographic rule for modelling it led to exponential growth of the values that define the priority of each patient. This complicated the feasibility of solving the scheduling problem when there were many patients. ILP2, by contrast, attempted to get around this weakness by modelling priority differently, but the result was weaker compliance with the priority levels. Although this could be countered by adjusting certain parameters for each scenario, such a solution would impair the model's responsiveness in day-to-day applications.

In regard to operating room capacity utilization rates, the third test criterion, the best results were generated by ILP2 while those of the other formulations depended strictly on the durations of the assigned patients' operations.

It should be evident from the foregoing that choosing the method which will give the best results is not a one-dimensional decision. Much will depend on establishing a clear idea of the considerations involved in the functioning of the hospital and its surgical specialties as well as the characteristics of the patients and the particular interventions they require. The real scenarios tested in this study demonstrated that a constructive algorithm could be successfully applied, or alternatively an ILP model. Yet both may experience difficulties in certain scenarios, reducing somewhat the robustness of their solutions. If the execution time requirements of the method's intended application permit the use of the feasibility model algorithm, this approach will provide optimal solutions in terms of patient priority under any scenario. The variant ILP model offers reasonable execution times and better operating room utilization rates but since its solutions do not strictly comply with patient priority assignments, its choice would require the approval of hospital personnel.

A key contribution of the present study was the incorporation of the relative patient priority concept into the scheduling methods, as they proved to be fundamental in the modelling of the entire scheduling problem. A survey of the literature revealed that many studies centre the patient assignment decision on the minimization of operation costs without including criteria such as the patient's biomedical condition or waiting list time.

To determine relative patient priority in the real scenarios studied, information from a previous study was used to prioritize patients on the basis of waiting time and biomedical complexity criteria (Barros and Julio 2011). The four methods proposed in this paper are currently being trialled at a children's hospital in Santiago, Chile. The authors have developed and implemented a computer application based on these methods that generates weekly operating room schedules, and tests have so far delivered satisfactory results from both clinical and hospital resource use points of view.

Acknowledgments The authors would like to thank Oscar Barros for providing the inspiration to study the operating room scheduling problem, and Daniel Espinoza and Kenneth Rivkin for their interesting suggestions. They are also grateful to the medical and administrative staff at Luis Calvo Mackenna Hospital for their collaboration. The first author was partly financed by UBACyT Grant No. 20020130100808BA (Argentina), ANPCyT PICT Grant No. 2012-1324 (Argentina) and FONDECyT Grant No. 1140787 (Chile) as well as by the Institute Complex Engineering Systems Institute (Chile).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Barros, O., & Julio, C. (2011). Enterprise and process architecture patterns. *Business Process Management Journal*, *17*, 598–618.

Cardoen, B., Demeulemeester, E., & Beliën, J. (2009). Scheduling surgical cases in a day-care environment: An exact branch-and-price approach. *Computers and Operations Research*, *36*, 2660–2669.

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

Dexter, F., & Macario, A. (2002). Changing allocations of operating room time from a system based on historical utilization to one where the aim is to schedule as many surgical cases as possible. *Anesthesia and Analgesia*, *94*, 1272–1279.

Guerriero, F., & Guido, R. (2011). Operational research in the management of the operating theatre: A survey. *Health Care Management Science*, *14*, 89–114.

Guinet, A., & Chaabane, S. (2003). Operating theatre planning. *International Journal of Production Economics*, *85*, 69–81.

Gupta, D. (2007). Surgical suites' operating management. *Production and Operations Management*, *16*, 689–700.

Hadorn, D., & Holmes, A. (1997). The New Zealand priority criteria project. Part 2: Coronary artery bypass graft surgery. *British Medical Journal*, *314*, 135.

Hilkhuysen, G., Oudhoff, J., Rietberg, M., Wal, G., & Timmermans, D. (2005). Waiting for elective surgery: A qualitative analysis and conceptual framework of the consequences of delay. *Public Health*, *119*, 290–293.

Jebali, A., Alouane, A., & Ladet, P. (2006). Operating rooms scheduling. *International Journal of Production Economics*, *99*, 52–62.

MacCormick, A. D., Collecutt, W., & Parry, B. (2003). Prioritizing patients for elective surgery: A systematic review. *ANZ Journal of Surgery*, *73*, 633–642.

Marques, I., Captivo, M., & Pato, M. V. (2011). An integer programming approach to elective surgery scheduling: Analysis and comparison based on a real case. *OR Spectrum*, *34*, 407–427.

Min, D., & Yih, Y. (2010). An elective surgery scheduling problem considering patient priority. *Computers and Operations Research*, *37*, 1091–1099.

Mullen, P. (2003). Prioritising waiting lists: How and why? *European Journal of Operational Research*, *150*, 32–45.

Noseworthy, T., McGurran, J., & Hadorn, D. (2003). The Steering Committee of the Western Canada Waiting List Project: Waiting for scheduled services in canada: development of priority-setting scoring system. *Journal of Evaluation in Clinical Practice*, *9*, 23–31.

Oudhoff, J., Timmermans, D., Knol, D., Bijnen, A., & Wal, G. (2007). Prioritising patients on surgical waiting lists: A conjoint analysis study on the priority judgements of patients, surgeons, occupational physicians, and general practitioners. *Social Science and Medicine*, *64*, 1863–1875.

Reveco, C., & Weber, R. (2011). Gestión de capacidad en el servicio de urgencia en un hospital público. *Revista Ingeniería de Sistemas*, *25*, 57–75. (in Spanish).

Russell, C., Roberts, M., Williamson, T., McKercher, J., Jolly, S., & McNeil, J. (2003). Clinical categorization for elective surgery in Victoria. *ANZ Journal of Surgery*, *73*, 839–842.

Santibáñez, P., Begen, M., & Atkins, D. (2007). Surgical block scheduling in a system of hospitals: An application to resource and wait list management in a British Columbia health authority. *Health Care Management Science*, *10*, 269–282.

Siciliani, L., & Hurst, J. (2005). Tackling excessive waiting times for elective surgery: A comparative analysis of policies in 12 OECD countries. *Health Policy*, *72*, 201–215.

Testi, A., Tanfani, E., Valente, R., Ansaldo, L., & Torre, C. (2006). Prioritizing surgical waiting lists. *Journal of Evaluation in Clinical Practice*, *14*, 59–64.

Valente, R., Testi, A., Tanfani, E., Fato, M., Porro, I., Santo, M., et al. (2009). A model to prioritize access to elective surgery on the basis of clinical urgency and waiting time. *BMC Health Services Research*, *9*, 1.

Journal: 10479
Article: 2172

Author Query Form

Please ensure you fill out your response to the queries raised below and return this form along with your corrections

Dear Author

During the process of typesetting your article, the following queries have arisen. Please check your typeset proof carefully against the queries listed below and mark the necessary changes either directly on the proof/online grid or in the 'Author's response' area provided below

Query	Details required	Author's response
1.	The affiliation 1 has been split into three different affiliations. Please check and correct if necessary.	
2.	Please check and confirm whether the inserted city name 'Buenos Aires' is correctly identified.	
3.	Please check and confirm the author names and initials for the author "Patricio Wolff" is correctly identified.	
4.	Please check and confirm whether the mail id 'prey@dii.uchile.cl' should appear in the publication.	