Development and implementation of an operating room scheduling tool: An action research study

Development and implementation of an operating room scheduling tool: An action research study

FILIPPO VISINTIN*
filippo.visintin@unifi.it
University of Florence
Dipartimento di Ingegneria Industriale
IBIS Lab.
PO Box 50134, Viale Morgagni 40/44, Florence, Italy

PAOLA CAPPANERA
paola.cappanera@unifi.it
University of Florence
Dipartimento di Ingegneria dell'Informazione
IBIS Lab.
PO Box 50139, Via Santa Marta 3, Florence, Italy

CARLO BANDITORI
carlo.banditori@unifi.it
University of Florence
Dipartimento di Ingegneria Industriale
IBIS Lab.
PO Box 50134, Viale Morgagni 40/44, Florence, Italy

PAMELA DANESE
pamela.danese@unipd.it
University of Padova
Department of Management and Engineering
PO Box 36100, Stradella San Nicola, 3, Vicenza, Italy

DOI: https://doi.org/10.1080/09537287.2017.1310328

Keywords: healthcare, operating room, master surgical scheduling, action research, optimisation.

**Corresponding author*

Abstract

This study examines the master surgical scheduling (MSS) problem. Although based on real data, most studies presenting mathematical models to support MSS do not address the practical problems arising from model utilisation, and do not provide evidence of model implementation. In this study, we report the results of an action research project whose aim was to develop and implement an operating room scheduler in a children's hospital. The findings offer insights into features such that make an MSS optimisation model and scheduler effective and easy to implement, and shed light on those actions facilitating the introduction and use of a scheduler embedding the developed model. Specifically, this study suggests that creating an effective and easy to use operating room scheduler requires clustering patients in homogeneous surgery groups and developing a flexible tool that allows: scheduling surgery groups instead of actual patients, easily adding/removing constraints, changing the objective function(s), and adjusting the planning horizon. In addition, it posits that gaining the commitment of top management by showing credible preliminary results, inferring stakeholder preferences by letting them comment on tentative schedules, introducing changes gradually and involving staff at lower levels of the hospital hierarchy can significantly facilitate the scheduler development and implementation.

1. Introduction

The operating theatre (OT) is considered to be the 'engine that drives the hospital' (Beliën et al., 2006, p. 343), typically accounting for almost 70% of all hospital admissions (Denton et al., 2007). OT activities have a significant influence on those of other departments and, consequently, on hospital performance as a whole (Cardoen et al., 2010; Lin et al., 2013). The OT is also one of the costliest functional areas of a hospital (Denton et al., 2007; May et al., 2011). In the literature, there is unanimous consensus that OT performance depends strongly on how surgical activities are scheduled (Litvak and Long, 2000; Guinet and Chaabane, 2003; Su et al., 2011). There is also consensus that the scheduling process comprises three sequential stages (Beliën and Demeulemeester, 2007): (i) case-mix planning, (ii) master surgical scheduling (MSS) and (iii) patient selection and sequencing. The aim of the first stage is to establish the total amount of operating room (OR) time to be assigned to each surgical specialty, usually on a yearly basis. The second stage involves determining the specialty (or specialties) to assign to each OR for each day of the planning horizon (e.g. two weeks or one month) and, sometimes, the number and types of surgery to be performed each day. Finally, the third stage involves selecting and sequencing patients to undergo surgery in the coming week.

This study focuses on MSS stage. In general, solving an MSS problem is noticeably complex, requiring consideration of the following: (i) many different types of cases, characterised by different priority levels and requiring different procedures; (ii) many different types of resources, such as ORs, OR personnel (e.g. surgeons, anaesthetists and nurses), surgical and electromedical equipment, postsurgical resources (e.g. intensive care units, postsurgical units); (iii) the randomness associated with patients' arrival, duration of surgery and patients' length of stay (LoS) (May et al., 2000); and (iv) the conflicting priorities and preferences of scheduling process stakeholders (Cappanera et al., 2016b, Rich and Piercy, 2013; Glouberman and

Mintzberg, 2001). Coupled with its significant economic and social impact, the complexity of the scheduling process has stimulated intensive research activity in recent years and has encouraged the development of supporting models (principally optimisation models), as reported in the recent and comprehensive literature review by Samudra et al. (2016). However, despite the ever-increasing number of models proposed, studies reporting the results achieved with the implementation of these models are very scarce. This, in turn, limits understanding of whether these models work in practice and of problems that may be encountered during their implementation.

The present study addresses this gap. Specifically, we report the results of an action research project whose aim was to develop an MSS optimisation model and operating room scheduler and to implement them in a real context—the Meyer University Children's Hospital in Florence (hereafter Meyer Hospital), one of Europe's most renowned children's hospitals. As it is typical in action research, this study has two general objectives: (i) to solve a problem of practical relevance (in this case, to smooth the scheduling process and improve OT performance at Meyer Hospital) and (ii) contributing to MSS theory. With regard to the first objective, this study highlights the numerous benefits of using a scheduler embedding an MSS optimisation model. As to the second objective, the study provides insights into features such as objective function(s), constraints and parameters that make an MSS optimisation model and scheduler effective and easy to implement. In addition, this study analyses actions that facilitate the introduction and use in a real setting of a scheduler embedding the developed model. This experience-driven evidence can assist researchers and practitioners in the development and introduction of tools supporting MSS.

The remainder of the paper is organised as follows. Following a review of the literature in Section 2, Section 3 details the research methodology. Sections 4 and 5 describe the steps of the action research project, and Sections 6 and 7 respectively present the results of the project and the implications for practice and theory. Finally, Section 8 presents conclusions and limitations and indicates a direction for future research.

2. Literature review

The healthcare literature is characterised by a large number of quantitative models advanced to support a wide range of processes but with little evidence of their implementation. In their literature review, Brailsford et al. (2009) found that only 5.3% of the reviewed models were actually used in practice. Two often-cited factors (Wilson, 1981; Harper and Pitt, 2004; Brailsford et al., 2009; Van Oostrum et al., 2010; Brailsford and Vissers, 2011; Mahdavi et al., 2013; Virtue et al., 2013) hamper the implementation of models in healthcare organisations.

- *Limited involvement of the stakeholders in model development*. Limited interaction with stakeholders from the early stages of the model development process can lead to poor understanding of the multifaceted and wicked nature of healthcare problems. This, in turn, can lead the modeller to make unrealistic assumptions and/or to neglect relevant constraints and objectives (Eldabi, 2009). This limited involvement also makes it harder for stakeholders to understand the benefits associated with the model implementation, both for themselves and for the organisation as a whole (Wilson, 1981) and/or to develop realistic expectations about project timescale (Eldabi, 2009).

- *Excessive complexity of the model*. Over the years, more and more complex models have been developed. These can require significant time for formulation, which may not be compatible with the needs of healthcare organisations (Proudlove et al., 2007). Additionally, because these complex models require expert end-users to be understood and used, they can be perceived by hospital personnel as inadequate to address their problems (Eldabi, 2009).

Narrowing the scope of the analysis to studies that *specifically address the MSS stage*, Table 1 presents the most relevant studies of MSS-related models. In line with the classification scheme proposed by Brailsford et al. (2009), we distinguish between studies in which models are suggested (i.e. developed without interaction with a client organisation), conceptualised (i.e. developed through interaction with a client organisation) or implemented (i.e. used by a client organisation).

Suggested	Conceptualised	Implemented
Vissers et al. (2005)	Santibanez et al. (2007)	Blake et al. (2002)
Kharraja et al. (2006)	Testi et al. (2007)	
Tànfani and Testi (2010)	Van Oostrum et al. (2008)	
Ma and Demeulemeester (2013)	Zhang et al. (2008)	
Agnetis et al. (2014)	Beliën et al. (2009)	
Fugener et al. (2014)	Chow et al. (2011)	
Yahia et al. (2014)	Agnetis et al. (2012)	
	Carter and Ketabi (2012)	
	Mannino et al. (2012)	
	Banditori et al. (2013)	
	Cappanera et al. (2014)	
	Visintin et al. (2014)	
	Aringhieri et al. (2015)	
	Jebali and Diabat (2015)	
	Cappanera et al. (2016a)	
	Cappanera et al. (2016b)	

Table 1. MSS models: Degree of implementation

All the studies listed in Table 1 present quantitative model(s) and assess the benefits that hospitals could (in theory) achieve by adopting them. As can be noticed, only one study (Blake et al., 2002) reports the results of the model implementation. This does not exclude that the other models have been implemented at a later stage, but simply these studies do not focus on model implementation. This is confirmed by Cardoen et al. (2010), who observe how the surgical scheduling literature lacks contributions addressing implementation aspects and argue that knowledge about the possible causes that could lead to failure or success may be of a great value for the research community, thereby encouraging scholars to share their implementation experiences.

In this regard Blake et al. (2002) stress that stakeholders will be willing to use an MSS optimisation model only if the schedule it produces changes little from week to week. In practice, this allows for better coordination of specialists' activities inside and outside the OT as well as for better predictability of the process and of patient flows. Blake et al. (2002) also warn that stakeholders may find it difficult to identify attributes that characterise a good schedule, making it difficult to identify a suitable objective function. In addition, they argue that arriving at a satisfactory schedule is likely to involve several trial-and-error iterations. For this reason, they advocate development of a user-friendly graphical user interface (GUI) that allows users to add, delete or modify constraints included in the model, to simulate a variety of scenarios and to select those solutions best fitted to their needs. In that study, for instance, stakeholders wanted the schedule to be 'simple', but only through trial and error was it discovered that 'simple' meant that no surgical session would be allocated to more than two specialties in any calendar month.

A further significant issue relates to constraint management. Van Oostrum et al. (2010) suggested avoiding use of too many constraints, especially those based on specialists' personal preferences. On the other hand, they argued that solutions that fail to leave some autonomy to specialists, such as the possibility to select the patients to operate in given sessions, are unlikely to be implemented. On the other hand, they warn that affording too much autonomy to specialists can lead to conflicts when downstream resources (such as general wards) are shared, as specialists may be tempted to secure an undue proportion of shared capacity.

Van Oostrum et al. (2010) also provided insights into definitions of the planning horizon. They argued that a longer planning horizon increases utilisation by making it easier to assign patients to appropriate slots. Against this, they also note that a long planning horizon may cause unacceptable waiting times when cases requiring early surgery are postponed at the end of the planning horizon.

Insight about how an OR scheduler should be developed can be found also in the study of Dios et al. (2015). Despite adopting a fundamentally different scheduling approach (i.e. an "open scheduling approach", where surgery can be assigned to any OR and at any hour of a working shift and does not have to comply with a pre-defined master surgical schedule) and considering no downstream resources, the authors provide suggestions that are valid in general. First, they emphasize that the hospital management's objectives are likely to change over time and, consequently, they advocate the development of tools with a modular architecture allowing the model to be easily updated and/or changed without affecting other modules (i.e. the database and the GUI). Second, consistently with Blake et al., 2002 they stress that one of the greatest value of a scheduling tool is that it allows to perform informed scenario analysis prior to implementation.

Although the above-mentioned studies provide useful insights about the implementation process of OR scheduling or MSS models, their implementation still represents a major organisational challenge, as several implementation aspects remain underexplored. Specifically, the literature lacks studies *based on experience-driven evidence* that link the features of a model (e.g. objective function(s), constraints, planning horizon) to issues that might arise if the model were embedded in a scheduler and used in a real context. In addition, there is a lack of contributions that discuss how the introduction and use of such a scheduler should be managed to avoid or address these problems.

The present study aims to bridge this gap in the literature analysing the following research questions.

RQ1: What features make an MSS optimisation model and scheduler embedding the model effective and easy-to-use?

RQ2: What actions facilitate scheduler implementation?

3. Methodology

This study adopts an action research (AR) approach (Coughlan and Coghlan, 2002). Commencing in 2011, the study involves a multi-year research collaboration between the IBIS Lab, a laboratory of the Department of Industrial Engineering of the University of Florence, and the Meyer Hospital.

Meyer Hospital's top management commissioned this project with the aim of improving OT performance and increasing personnel competence in healthcare operations management (OM). Specifically, in 2011, the surgical scheduling process was the hospital's greatest concern. The hospital Planning Department lacked the tools and skills to manage OR scheduling; decisions about resource utilisation were based on intuition and frequently influenced by undue pressure from surgeons. This resulted in an unsatisfactory throughput, limited and unbalanced utilisation of resources (ORs, beds) and frequent problems of overcrowding. The hospital therefore needed a tool to automate and support the MSS process. It was eventually agreed to develop an optimisation model tailored to the hospital's needs, to embed this in an OR scheduler and to train personnel in its use. The collaboration agreement allowed us, as researchers, to freely access requisite data; to work with hospital personnel to understand scheduling problems, patients' requirements and operators' needs; and to verify the applicability of solutions and evaluate subsequent implementation problems. This close working relationship with hospital personnel helped to develop awareness over time of how their decisions could impact significantly on OR performance (e.g. bed management, resource optimisation, waiting lists).

AR methodology is particularly appropriate for present purposes because it is useful for the study of implementation processes and, specifically, for developing a fuller understanding of problems arising, solutions adopted and factors facilitating or hindering implementation (Coughlan and Coghlan, 2002, Farooq and O'Brien, 2015). By participating in and analysing the implementation of an MSS model and scheduler over time, it was possible to examine the transferability of the model developed, along with problems that characterised its implementation and solutions adopted. On that basis, the reflection process typical of this methodology enabled us developing generalisable findings relating to features of the MSS model that make it effective and easy to use and actions facilitating scheduler introduction and use, so addressing the literature gaps highlighted in the previous section.

The study comprised one pre-step (analysis of the context) and four cyclical steps: diagnosing, planning action, taking action, and evaluating action (Coughlan and Coghlan, 2002). To ensure that all stakeholders in the surgical planning process were represented and that all relevant data could be easily accessed, the AR team included both researchers from the IBIS Lab and hospital personnel, specifically:

- the General Director and the Medical Director, who commissioned the project and played an active part in defining objectives and revising deliverables;
- the OT manager, who was the responsible for the assignment of OR sessions to the surgical specialties;
- the bed manager, who was responsible for the allocation of patients to the postsurgical units;

- the Head of the Planning Department. As a nurse, in addition to being involved in actual planning of surgeries (i.e. assignment of patients to surgical sessions, in accordance with the MSS), she served as an interface between nurses working in the Planning Department (six people, each specialising in one or more surgical specialties), the OT manager and the bed manager and
- a member of the IT Support staff who could access the hospital information systems.

On occasion, other hospital employees, including nurses, surgeons and Planning Department personnel, were involved in the project. Stakeholders' participation in the AR project was fundamental in order to make their needs and preferences explicit and implement an actionable solution (Oral, 2012).

After a couple of plenary meetings with hospital top management to scope the project, the team commenced the *analysis of the context.* This aimed to understand the particular characteristics of the hospital (such as mission, services, resources, patient types, particularities of children's hospitals) by examining internal documents and conducting interviews with top management and other personnel. The *diagnosis* phase examined the surgical planning and scheduling process, beginning with extensive data collection. Data were gathered by a range of methods that included direct observation, interviews, querying databases and analysis of internal documents. Direct observations encompassed both the activities of the Planning Department and activities performed in the OT. The interviews extended to all hospital members of the AR team, Planning Department personnel and some OT personnel, including surgeons, anaesthetists and nurses. Databases were queried to retrieve surgical records as well as data pertaining to admissions and length of stay. Finally, a significant number of non-digitised documents such as surgery request forms (see Appendix A) were read and digitised. For three months, one IBIS Lab researcher worked inside the hospital for three days each week, taking an active part in all stages of the surgical scheduling process and carefully mapping the process using the BPMN formalism (Business Process Model and Notation) (White, 2004). Based on the identification of several criticalities (see Section 4) in the *planning action* phase, the team agreed that

- the MSS process needed to be reengineered and
- the reengineered process should use a scheduler to automate and optimise scheduling activities.

The *taking action phase*, then, involved (i) reengineering of the surgical scheduling process, (ii) development of a model to optimise and automate scheduling activities, (iii) development of a user-friendly scheduler embedding the model and (iv) implementation of the scheduler. Each action involved all members of the team, and the outputs of each action were systematically revised, with stakeholder involvement. The main outcome of reengineering the surgical process was the decision to implement a surgery groups-based scheduling approach (Van Oostrum et al., 2010; Banditori et al., 2013). This entailed a thorough analysis of surgical records as well as intense interaction with surgeons to form the surgery groups. Development of the optimisation model was by far the most complex and time-consuming step. The model as eventually implemented was the result of a cyclic process of formulation, in which several model versions were developed, real data were input and solved and the relevant solutions were critically discussed with the team.

To facilitate discussion, team members were provided with printed copies of several simulated schedules, along with a table reporting: (i) the total number of scheduled patients; (ii) the daily utilisation of ORs and bed units; and (iii) the surgery mix.

For each revision cycle, feedback from AR team members was used to improve the model. Following three iterations (and two models discharged), we arrived at a version of the model whose solutions were judged to be implementable and satisfactory. The technical features of the models can be found in Banditori et al. (2013), Cappanera et al. (2014), Cappanera et al. (2016a, 2016b). Compared to these studies, which focus on the mathematical models development, this AR study addresses the practical problems arising from model implementation. By working closely with hospital personnel in order to continuously adjusting the model, it was possible to progressively identify those features that made it effective and implementable.

To facilitate utilisation of the optimisation model, we created a scheduler embedding the model. The scheduler implementation process included installation of the scheduler in the hospital IT infrastructure, as well as end-user training. Following implementation, members of the AR team used the scheduler for several months (from November 2013 to July 2014). In particular, the scheduler was (and still is) used to (i) create the schedule at the beginning of the time horizon; (ii) perform ongoing adjustments to the schedule as needed.

Throughout the process of taking action, feedback from stakeholders and end users was leveraged to gradually improve the optimisation model, the scheduler and the overall scheduling and planning process. Specifically, the *evaluation phase* involved periodic meetings in which the AR team analysed the value of the KPIs reported in Table 2 and commented on the quality of the schedule produced.

Throughput indicators	**Throughput [Surgeries]**
	Average weekly number of surgeries [surgeries/week]
	Average weekly number of surgeries per available OR hour [surgeries/hours]
	Average weekly number of day surgeries
	Average weekly number of inpatient surgeries
	Percentage of day surgeries [%]
	Percentage of inpatient surgeries [%]
Resource utilisation indicators	**OR utilisation**
	Average weekly OR utilisation [hours/week]
	Average OR utilisation rate [%]
	Bed utilisation (weekends included)
	Average inpatient bed utilisation rate [%]
	Average day surgery bed utilisation rate [%]
	Inpatient surgery beds unit
	Beds available [beds]
	Mean bed utilisation [beds]
	St Dev bed utilisation [beds]
	Max bed utilisation [beds]
	Overbooking rate [%]
	Day surgery beds unit
	Beds available
	Mean bed utilisation
	St Dev bed utilisation
	Max bed utilisation
	Overbooking Rate [%]
Mix complexity	**Surgical time inpatient surgeries**
	Mean surgical time [min]
	LoS inpatient surgeries
	Mean LoS [days]
	Surgical time day surgeries [min]
	Mean surgical time [min]
	LoS day surgeries
	Mean LoS [days]

Table 2. KPIs used in the evaluation phase

Prior to release of the scheduler, the evaluation was based on simulated results; subsequently, it was based on real data. Once the scheduler was released, end users were also included in the evaluation. In particular, they were asked to use the scheduler under the supervision of a developer and to provide feedback. Direct observation and feedback from end users enabled bug fixing and improvement of the software.

Overall, this AR project required significant efforts by the Meyer Hospital and IBIS Lab, in terms of hours dedicated to the above-mentioned activities. Specifically, top managers (Medical and General Directors) were involved every 3 months in formal project review meetings lasting at least 1 hour each. Middle managers (Bed manager, OR manager and Head of the Planning Department) were involved in meetings taking place (approximately) on a monthly basis. Quite often, the Medical Director took part in these meetings as well. Three researchers from the IBIS Lab were involved in this project: one worked inside the hospital for at least 3 days per week; a second researcher was mainly involved in model(s) development and testing, while a further one acted as a project coordinator.

4. Context and issues to be addressed

4.1 The surgical planning and scheduling process at the Meyer Hospital

The Meyer Hospital is one of Europe's most renowned children's hospitals. Its activities include medical and surgical services, health education, health promotion and disease prevention, teaching, research and community health. The Meyer Hospital has 247 beds; in 2014, it cared for 33,413 inpatients and completed approximately 7,000 surgeries.

The Hospital's surgical facilities include seven ORs; five of these are partially interchangeable and host 15 surgical specialties: urology, otorhinolaryngology, paediatric surgery, neonatal surgery, ophthalmology, orthopaedic surgery, gynaecology and obstetrics, trauma centre, hand and microsurgery, oral and maxillofacial surgery, orthopaedic oncology, cardiothoracic surgery, gastroenterology, burns and plastic surgery. The remaining two ORs are dedicated almost entirely to specific surgical specialties or treatments (haemodynamics and bronchial endoscopy). These also accommodate emergencies and urgent cases, which represent a small percentage (10–15%) of treated cases as is common in children's hospitals. Meyer Hospital allocates 42 beds to elective patients. Beds are organised into three physically distinct units, one of which accommodates day surgery patients, who do not require an overnight stay. The other two units accommodate patients with a longer expected LoS—that is, inpatient surgeries that occupy a bed for more than one day.

Hospital waiting lists for patients needing surgery are populated on the basis of surgery request forms, which are completed by surgeons following patient visits. For each case, the form clearly indicates (i) the surgical specialty, (ii) the diagnosis, (iii) the procedure that the patient needs to undergo and (iv) a priority class. Priority class determines Maximum Time Before Treatment, expressed in days —in other words, the case's due date. The four priority classes (A, B, C, D) are assigned waiting times of 30, 60, 180 and 360 days, respectively. In some cases, patients are not assigned to a class and are recorded as class D. In addition, the form indicates (v) expected duration of the procedure (surgery duration) and (vi) expected LoS. There are three possible time ranges for procedure duration: less than one hour (short duration), between one and two hours (medium duration) and more than two hours (long duration). With respect to expected LoS, the form indicates whether the patient requires day surgery or inpatient surgery. Finally, the form includes (vi) the patient's contact information.

At the beginning of this study, the Meyer Hospital's scheduling process was managed by the typical three-stage approach (Beliën and Demeulemeester, 2007) described in the Introduction. Specifically, in the case mix planning stage, a certain number of OR hours were assigned to each specialty. This occurred once a year and entailed negotiation between hospital management and the surgeons responsible for the various specialties. To date, this phase has not undergone significant changes.

In the MSS stage, OR and bed managers produced the so-called 'allocation grid'—a timetable indicating the surgical specialties operating in each OR for each day and session of the planning horizon (one month). In addition, the grid provided a rough indication of the number of day surgeries and inpatient surgeries to be performed in each session. The grid was produced on a monthly basis, taking into consideration the number of hours to be allocated to each specialty (in accordance with the case mix planning output), availability of anaesthetists, nurses and electromedical equipment and, if possible, the surgeon's preferences and needs. In general, the allocation grid did not undergo any major variation during the year, as keeping the grid constant made it easier for surgeons to coordinate their activities within and outside the hospital. Nevertheless, minor changes usually arose each month.

Finally, in the patient selection and sequencing phase, the Planning Department personnel called some patients on the waiting list and assigned them to one of the sessions in the allocation grid, with at least one week's notice. To create a list of patients suitable for scheduling, Planning Department staff regularly called patients on the waiting list and arranged a pre-admission assessment with the anaesthetist. If the anaesthetist's feedback was positive, the patient could be operated on within six months, although not necessarily by the surgeon who prescribed the procedure.

Patients were selected from the ready-to-be-operated list and scheduled to ensure that:

(i) the sum of expected durations of surgeries assigned to each session did not exceed the duration of the session itself;

(ii) the expected number of hospitalised patients for each day of the planning horizon did not exceed the number of expected available beds (bed availability was controlled using a table on a large whiteboard, where each row represented a day and each column represented a specific bed; each time a new patient was scheduled, her/his name was inserted in the column corresponding to the bed she/he was expected to occupy, for a number of bed-day slots corresponding to his/her expected LoS);

(iii) the final schedule presented an approximate fixed number of short-, medium- and long-lasting surgeries to avoid leaving too many long-lasting surgeries on waiting lists, which would have made the scheduling process more complex in subsequent weeks or months and

(iv) when possible, patients with closer due dates were given higher priority.

The Planning Department was also asked to perform ongoing schedule adjustments to manage:

- *no-shows* (patients who become unavailable - typically because they are ill - a few days before the scheduled surgery);
- *urgent elective cases* (elective patients whose clinical condition has deteriorated rapidly and who need to be scheduled ahead of less urgent patients already on the list);
- *bed shortages* (caused by surgical patients with a longer than expected LoS or by non-surgical patients temporarily accommodated in surgical wards).

The feasibility of each adjustment, in terms of OR time and bed availability, was assessed by Planning Department personnel, based on experience and rudimentary tools such as the whiteboard mentioned above. Sporadically, surgeons who sought to disrupt the schedule were asked to estimate Surgical Time (ST) and LoS for the patients to be scheduled.

4.2 Open issues

Surgery request forms and anaesthetist assessments (pre-admission data, see Appendix A) were handwritten, making it impossible to subject these to detailed quantitative analysis. However, it was self-evident from the huge piles of paper on the Planning Department nurses' desk and shelves that waiting lists were growing rapidly, leading to high and rapidly increasing waiting times.

…every month the number of incoming surgery request forms we process is higher than the number of surgeries that we can schedule. [In addition,] we frequently come across surgery request forms for patients whose due date has already expired. [Nurse working in the Planning Department]

According to the Planning Department nurses, this was due to a chronic shortage of beds and a high level of uncertainty about expected ST and LoS, which led them to be conservative both when producing the initial schedule and in managing ongoing adjustments.

These whiteboards are always full of patients! [Bed Manager]

The estimates in the allocation grid are too inaccurate. The allocation grid indicates the specialty associated with each session and gives us a rough idea of the number of inpatient and day surgeries to be scheduled in each session. As we are not doctors and we have only a vague idea of the ST and LoS associated with the cases in the waiting list, if we feel that by scheduling the requested number of surgeries we can cause OR overtime and/or bed overbooking, then we schedule less surgeries than we should. [Nurse in the Planning Department]

It is very difficult to cope with short notice cancellations. Every time a patient needs to be replaced, we are supposed to replace her/him with a patient that has already been visited by the anaesthetist and with a similar LoS and ST. As it is not easy to identify a patient needing the same type of surgery as the cancelled one, we tend to replace cancelled surgeries with simpler and less resource-consuming surgeries… [Nurse in the Planning Department]'

Sometimes surgeons ask us to schedule specific patients in specific sessions. However, we have no clear idea whether this would lead to bed overbooking. This can lead us to cancel more than the necessary number of already scheduled surgeries. [Nurse in the Planning Department]

The analysis of LoS and ST of processed patients (post-admission data, see Appendix A), confirmed that while the ORs were underutilised (< 70%), bed utilisation was indeed very unbalanced across weekdays and very high on Thursdays and Fridays (close to 100%). OR and bed managers were not surprised by these results.

We suffer from a lack of beds. Ward overcrowding is the norm, especially on Thursdays and Fridays. I spend most of my time fighting against bed shortages [...] We often need to keep the day surgery unit open at night and/or to postpone surgeries already scheduled. [Bed Manager]

ORs are systematically underutilised but we are not understaffed. This can only be a consequence of poor planning. [OR Manager]

In summary, after direct observation of the surgical scheduling process, the interviews and data analysis revealed the following.

- In the MSS phase, specialties were assigned to OR sessions, also providing an indication of the number of day and inpatient surgeries to be performed in each session. However, creation of the allocation grid focused only on ORs—that is, it took no account of the effect of these decisions on bed utilisation. The actual availability of beds was considered only during patient selection and sequencing, but this cascade approach was suboptimal, as the unavailability of beds did not allow the expected number of surgeries to be scheduled, causing OR underutilisation, especially on Thursdays and Fridays.
- The cascade approach helped planners to cope with the complexity of simultaneously considering impacts on both OR and bed utilisations in a given schedule. However, without the aid of a scheduler, it was impossible to arrange surgeries one month in advance so as to ensure optimal utilisation of both resources.
- A scheduler would not in itself have solved the problem. In fact, as the estimates of ST and LoS in the surgery request forms were very inaccurate, a mathematical model using this data would probably have also produced misleading results.
- As well as making the task of creating the initial allocation grid more complex, the uncertainty about ST and LoS also made the management of short-notice cancellations and last-minute surgeons' requests more complex.
- There was no scarcity of anaesthetists, nurses and electromedical equipment, and their availability did not represent a constraint for OT performance.

5. The new scheduling process: MSS model and scheduler adopted

5.1 Creation of surgery groups and patient classification

To redesign the scheduling process, the AR team agreed to focus on three critical resources—surgical teams, ORs and beds—and to classify patients into homogeneous *surgery groups*. Patients in the same surgery group were expected to need a surgical team of the same specialty and to have similar ST and LoS. For this reason, it was decided that, in addition to assigning specialties to different OR sessions, the new MSS process would also have to determine the number of surgeries to be performed in each OR session and the surgery group these surgeries should belong to. Determining the number of patients to be scheduled and the respective surgery groups rather than scheduling actual patients would have two main advantages. First, it would facilitate the control of bed capacity utilisation at the beginning of the scheduling process—that is, before selecting patients for surgery. Second, it would make it easier to perform ongoing adjustments to the schedule. For example, replacement of an already scheduled patient would simply require the selection of another patient from the waiting list of the same surgery group.

To create surgery groups, data relating to elective surgeries performed in the triennium (2009–2011) were extracted from the Meyer Hospital databases (see Appendix A), and ST and LoS were determined for each patient. Each surgery was characterised by two sets of ICD9-CM codes.[1] The first set contained up to six codes identifying diagnoses, and the second set contained up to six codes identifying surgical procedures performed. In addition, the surgical record indicated the name and specialty of the surgeon. In practice, cases with the same diagnosis and procedure codes were sometimes carried out by surgeons from different specialties (e.g. urology and paediatric surgery). Additional special codes were added to distinguish surgeries sharing the same ICD9-CM code but characterised by significantly different ST and/or LoS. For example, bilateral surgeries (i.e. surgeries performed in a single procedure on both eyes, both hands etc.) bore the same code as their unilateral counterparts but often took much longer.

Since both ST and LoS tend to be distributed according to a lognormal distribution (May et al., 2000; Stepaniak et al., 2009; Marazzi et al., 1998; Carter and Ketabi, 2012; Cappanera et al., 2014), we filtered outliers for each possible combination of diagnosis codes, procedure codes and specialty (hereafter, *code combination*), using the adjusted boxplot method, which has proved effective for skewed distributions (Hubert and Vandervieren, 2008). With this cleaned dataset, we calculated several descriptive statistics (mean, median, standard deviation, quartiles) for both ST and LoS, for each code combination.

The clusterisation of code combinations followed a very simple and pragmatic procedure. For ST, discrete time intervals were identified, each of 30 minutes' duration (1–30, 30–60, 60–90 etc.). Similarly, for LoS, we identified discrete time intervals corresponding to days (0, 1, 2 etc.). Then the surgery group name was created by combining an abbreviation of the specialty name (e.g. URO = Urology, ORL = otorhinolaryngology) with the upper bound of the ST time interval (30, 60, 90 etc.) and the LoS (0, 1, 2, 3 etc.). Finally, code combinations were assigned to surgery groups, based on the mean value of ST and LoS—and so, for example, the group URO-90-2 represented all the code combinations of the urology specialty for which the mean ST value fell within the interval (60–90] and the (rounded) mean LoS value was 2. The mean ST value of all the cases belonging to each surgery group was calculated (e.g. for URO-90-2, mean (ST) was 85 mins), and mean ST values and rounded mean LoS values were subsequently used as input data for the optimisation model. Using the mean values of ST and LoS in this way is conservative, as it allows overestimation of the duration of most scheduled surgeries. Because ST and LoS are lognormally distributed, their median is smaller than their mean. Application of this procedure led to the creation of almost 150 surgery groups, which were revised and

[1] The International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) is the official system for assigning codes to diagnoses and procedures associated with hospital utilisation in the United States and other countries, including Italy (see http://www.cdc.gov/nchs/icd/icd9cm.htm).

validated by surgeons from the respective surgical specialties and subsequently used to characterise all patients in the waiting list.

5.2 Model(s) development

The AR team agreed that model solutions would have to meet the following criteria:

Efficiency: The model should allow scheduling of a large number of surgeries and should lead to high utilisation of ORs and beds. Increasing throughput was considered necessary to reduce waiting times and to increase revenues.

Long-term orientation: The model should always schedule a certain proportion of complex surgery groups (characterised by long ST and LoS) to avoid any excess of these in the waiting list.

Robustness: The model solution should be robust in coping with variations of LoS and ST. In other words, even if a certain surgery and/or LoS lasted longer than expected, the solution should minimise overtime and/or overbooking cancellations.

Equity: The model should take into consideration the number of patients in each surgery group and their due dates, prioritising surgery groups with a higher number of patients with closer due dates.

In line with several previous works in the MSS literature (see e.g. Cardoen et al., 2010), it was decided to use integer/mixed-integer programming (MIP). In addition, as a deterministic environment would not have allowed us to test the robustness of the proposed models, it was decided - as in Ewen and Mönch (2014) - to create a simulation model for that purpose. A testing environment integrating optimisation and simulation models allowed the AR team to assess the effects of certain solutions before actually implementing them and to provide valuable feedback for successive refinements of the model.

It is worth to point out that, given the stochastic nature of ST and LoS, stochastic programming could have been a promising methodology to employ. However, as suggested by Saccani et al. (2016), such a modelling approach can cause two problems: (i) the formulation of stochastic optimisation models can be very complex and time consuming (which is a big issue in time-constrained project like ours), (ii) when applied to real-life data, solving these models can be extremely computationally demanding. This often leads to the formulation of simplified assumptions that can undermine the credibility of the results and, consequently, the chances that the model will be eventually used. This is confirmed by the fact that there is no evidence in the literature of stochastic models supporting the MSS problem being actually employed in hospitals (Samudra et al., 2016). For this reason, we decided to opt for a sequential optimisation and simulation approach.

In the testing environment, the optimisation model was coded using AMPL (Fourer et al., 1993), with direct input of hospital data, and solved using the IBM-ILOG CPLEX solver. The model solutions were subsequently simulated using Rockwell Arena, and the relevant results were analysed using MS Excel. Visual Basic was used to integrate all software and to automate most of these tasks. The testing environment was used extensively in developing each model and to evaluate the implications of implementation. Once a model was developed, stakeholder feedback was used to create the following model. In total, three different models were created and tested before a reasonably stable version was embedded in the scheduler.

All the models returned a solution – the schedule - indicating for each OR and session of the planning horizon: (i) the number of patients to be scheduled and (ii) the surgery group these patients must belong to. The models are characterised by an objective function, which includes the criterion/criteria to be optimised, the variables relating to decisions to be taken, and a set of constraints describing requirements that the model solution must respect. Constraints were of two types: utilisation of critical resources (ORs, beds, surgical teams) and quality requirements, relating mainly to patients' due date and the scheduled case mix. Despite sharing a similar structure, the models were characterised by different objective functions and constraints.

Model 1

At the beginning of the project, management's top priority was to increase OT efficiency. Consequently, the objective function of Model 1 was a hierarchical function whose most important term related to maximisation of the number of scheduled patients; other terms were added to penalise the value of the objective function whenever quality criteria were not met. Equity was controlled using both hard[2] and soft constraints (referred to as *due-date constraints*). Hard due-date constraints obliged the model to schedule all patients whose due date expired within the planning horizon. Soft due-date constraints penalised the objective function when patients with approaching (but not expiring in the planning horizon) due-dates were not scheduled. Hence, while hard constraints guaranteed that all patients with due date expiring in the planning horizon were scheduled, soft constraints allowed to prioritize, if possible, those patients whose due date was closer but not expiring in the planning horizon.

The long-term orientation of the solution was ensured by hard constraints (called *mix constraints*), bounding the scheduled percentages of short, medium and long ST surgeries. Finally, robustness against the uncertainty of ST and LoS was achieved by use of *resource slacks*—that is, by instantiating the model with reduced availability of daily OR time and beds, respectively, to create buffers to absorb peaks in daily utilisation of OR and beds that may follow schedule execution. The size of these resource slacks was determined by means of a combined optimisation-simulation approach. This is an iterative approach in which the (deterministic) optimisation model is instantiated using many combinations of resource slacks. The solution corresponding to each combination is subsequently simulated in a stochastic environment, where ST and LoS are free to vary in accordance with suitable probability distributions, causing overtimes and overbooking cancellations. We found that respective resource slack values of 10% and 12% for ORs and beds ensured efficient and robust scheduling.

[2] In mathematical optimisation, constraints can be either hard or soft. Hard constraints set conditions for the variables that must be strictly satisfied. Soft constraints include some variable values that are penalised in the objective function if and to the extent that the conditions for the variables are not satisfied.

Following discussion with the AR team, it was concluded that the quality of this model's solutions was unsatisfactory. The emergence of several interesting and unexpected issues suggested a need to modify the model. First, daily utilisations of beds and ORs were judged to be too unbalanced (in some cases, OR session utilisation was less than 50%). Second, the schedule contained too many day surgeries. In addition, we noticed that mix constraints and due-date constraints caused problems. Mix constraints had an undue negative impact on throughput, as they prevented scheduling of additional short surgeries (even if the final solution provided for enough space to accommodate them). Moreover, they allowed control of the mix only with regard to ST and not to LoS; for that reason, the model scheduled day surgeries instead of inpatient surgeries with the same ST whenever this was possible. Finally, in several instances, the joint presence of hard mix and due-date constraints made it impossible to find a feasible solution. This latter issue was particularly critical; as users without an operational research background may be unable to detect the constraint or combination of constraints that creates infeasibility, a model with too many (conflicting) hard constraints was considered too difficult to be managed and implemented.

Model 2

One of the most significant changes in moving from Model 1 to Model 2 concerned the objective function. In Model 2, the new objective function still comprised hierarchical terms and still included maximisation of the throughput term. However, two terms were added to achieve more balanced utilisation of beds and ORs. Balancing utilisation of critical resources in this way achieved a fairer distribution of workload across surgeons and nursing staff and ensured higher robustness. In fact, if resources are properly balanced and average resource utilisation is not too high, some idle resources should always be available to absorb any unexpected peaks caused by ST and LoS variability (Belien et al., 2009), even without resorting to resource slacks. In addition, *mix constraints* were modified in Model 2. First, constraints were added to control the mix in terms of LoS (day surgeries vs. inpatient surgeries), so preventing excessive scheduling of day surgeries. Second, the constraints controlling the ST mix were relaxed by switching from three ranges (short, medium and long ST) to two ranges (short and long ST). In so doing, it became possible to avoid the negative impact of these constraints on throughput while making an infeasible solution less probable. Additionally, simulation was used to test schedule robustness. The simulation results showed that efficiency and robustness could be achieved by balancing resource utilisation. Specifically, we found that a utilisation target of 85% was adequate to guarantee both efficiency and robustness (which is consistent with the results for Model 1). While the results for Model 2 were judged more satisfactory by the AR team, they still presented some issues. In particular, hard constraints still occasionally led to infeasibilities that would be too difficult for hospital personnel to manage. Moreover, the final solution was unexpectedly characterised by an excessive number of short-stay inpatient surgeries (i.e. surgeries with LoS of two days), again undermining the long-term orientation of the solution.

Model 3

To avoid infeasibilities caused by hard constraints in managing case mix and due dates, it was decided to convert these constraints into objective function terms. The objective function of Model 3, then, comprises five terms:

1. maximisation of the number of scheduled surgeries (efficiency);
2. balancing of daily utilisations of ORs (robustness against ST variability);
3. balancing of daily utilisations of beds (robustness against LoS variability);
4. patient due date fulfilment (equity) and
5. matching desired mix of surgeries (long-term orientation).

A goal programming approach was adopted to deal with this multi-criteria problem, which allows a weight to be assigned to each term or criterion according to the preferences of the decision maker. The flexibility of this approach was greatly appreciated by stakeholders, and the solutions produced by the model were judged satisfactory by the AR team. On that basis, Model 3 was integrated in the scheduler.

5.3 Scheduler

The model had to be embedded in a user-friendly scheduler, which allowed unskilled end users to perform three main tasks:

1. setting model parameters and finding the optimal schedule;
2. visualising the schedule along with the values of a few relevant performance indicators (number of scheduled surgeries, daily OR and bed utilisation, etc.) and
3. assessing the impact and feasibility in terms of OR and bed utilisation of changing *ex-post* some of the scheduled surgery groups, due to last-minute schedule changes.

The scheduler embeds an optimisation model, which is coded using the open language Python-Pyomo (Hart et al., 2012), and uses the open source COIN-OR CBC solver (Forrest and Lougee-Heimer, 2005). The software GUI comprised two spreadsheets, named *Input* and *Output*. The Input spreadsheet contains a form in which the end user can insert the following data.

- Length of the planning horizon
- For each day, OR and session of the planning horizon:
    - Session duration
    - Session type (inpatient, day surgery, mixed)
    - Specialty operating in that session
- Number of available beds for each day of the planning horizon and each unit (day surgery and inpatient surgery bed units)
- Target daily utilisation for OR and beds
- Weight associated with each term of the objective function.

Together with the waiting lists stored in the hospital database, these data are imported, elaborated and input to the optimisation model via Python. The solution produced by the solver is then elaborated, formatted and visualised on the Output worksheet, which also contains several forms. One of these reports the solution—that is, the surgery groups to be scheduled for each day, session and OR. The other forms report (see also Figures 1 and 2 in Section 6):

- the daily OR and bed utilisation associated with the solution;
- the input value of OR time and beds on which the solution was based and
- other summary performance (number of surgeries planned, surgery group etc.).

The cells of these forms can be easily edited by the end user to manage ongoing schedule adjustments, including:

- addition, substitution or cancellation of one or more of the scheduled surgery groups and
- modification of the input capacity (in terms of beds and/or OR time).

Each time a cell is edited, it turns yellow to warn the end user that the system has yet to check whether or not the performed changes are feasible. Once the operator has performed the desired changes, the check can be triggered by pressing a button. The check process verifies whether:

- the addition, substitution or cancellation of one or more of the scheduled surgery groups is compatible with the available resources and
- it is still possible to implement the initial schedule, based on the modified capacity.

If yes, the yellow cells turn green; otherwise, they turn red. In this latter case, the system also highlights in red those cells corresponding to the day, OR sessions and/or bed units lacking the required capacity to accommodate the proposed change. In this way, the operator can immediately understand the impact of an adjustment, visually monitor expected utilisation of resources throughout the planning period and predict days when criticalities are more likely to occur.

5.4 Feedback from scheduler use and improvements

Once developed, the scheduler was transferred to the Meyer Hospital, and the first schedule actually implemented by the hospital was produced on 25 October 2013, using a preliminary (beta) version of the scheduler. The schedule covered the four-week period from 4 November to 1 December.

The first three schedules were entirely produced and managed by the IBIS Lab researchers. The ex-post changes were performed at the presence of the bed manager and/or Head of the Planning Department and discussed with them. The five subsequent schedules were again produced by the developers, but ex-post changes were entirely managed by the Bed Manager. Afterwards, the scheduling process was executed by hospital personnel, initially under the supervision of the developers (for two planning periods) and then autonomously.

We observed application of the scheduler for 15 consecutive planning periods (to July 2014).

During this scheduler implementation process, feedback from stakeholders' and developers' use-experiences, as well as from bed managers and the Head of the Planning Department, was leveraged to gradually improve the optimisation model, scheduler and overall scheduling process. More specifically, during implementation, the optimisation model was improved as follows.

1. Several hard constraints were added to accommodate specific requests. Each time a schedule was created, surgeons made very specific requests concerning the number of cases and/or types of surgery (inpatient surgeries vs. day surgeries) and/or surgery groups to be scheduled in specific sessions or weeks of the planning horizon (typically one month).
2. An earliest programmable date constraint was also added. For certain cases, surgeons clearly warned against scheduling the surgery until the patients had reached a certain age or clinical condition (e.g. until they had completely recovered from an ongoing disease) by means of a note on the surgery request form.
3. An additional term in the objective function took account of long-waiting patients. A significant number of patients on the waiting list have been diagnosed with diseases that require surgery before a certain age or before puberty. These patients are not assigned a specific due date (Class D patients), and scheduling them is quite complex because the surgery they need is typically characterised by a long ST and leads to a long LoS. As these cases were neither urgent nor easy-to-schedule, they were not prioritised and, in consequence, the surgery groups to which these surgeries belong included a significant number of long-waiting patients. By adding a term in the objective function, it became possible to increase the number of long-waiting scheduled patients, simply by increasing the weight associated with the added term.

The changes introduced in the optimisation model required us to modify the scheduler accordingly, with the following improvements for the end-user.

1. *Managing added mix constraints*: As previously noted, hard constraints can lead to model infeasibilities. For operators without an operational research background, it is difficult to set such constraints or to investigate the root causes of model infeasibility. The scheduler was therefore modified by (i) adding a new user form to allow intuitive setting of mix constraints to enable end users to accommodate surgeons' requests and (ii) creating alerts to identify the constraint(s) causing model infeasibility.
2. *More easily setting the weight associated with the added term in the objective function*: This was already possible for other terms.

During the implementation phase, the AR team also decided to modify the scheduling process**.** These adjustments were undertaken to address the recurring problems arising in the first months of implementation, some of which were caused by the newly added scheduler functionalities. More specifically, the following adjustments were made.

1. *The planning horizon was shortened from 4 to 2 weeks.* Although longer planning horizons allowed for better theoretical solutions (e.g. it allowed scheduling approximately 2 additional surgeries per week), changes in external conditions constantly required multiple ex-post adjustments, resulting in implemented solutions that were worse than those obtained with a shorter planning horizon. The most common causes driving modification of the schedule included the following.
    a. *Deferrable elective surgeries*. These are surgeries not classified as emergency or urgent but whose deferral for more than two weeks may jeopardise life or essential function or cause severe pain. Neonatal surgeons, in particular, regularly require scheduling of new-born elective patients who, by definition, were not in the waiting list at the beginning of the time period, with two weeks' notice.
    b. *Bed unavailability caused by other than elective surgeries*. For example, periodically, overcrowding in medical wards meant that the bed manager might decide to use surgical beds to accommodate medical patients as well.

    Shortening the planning horizon allowed most of these situations to be managed without any ex-post adjustments to the initial schedule. As most of these changes were foreseeable with two weeks' notice, it was sufficient to adjust the data to take account of the reduced capacity to determine the optimal solution. In addition, even when one or more adjustments were actually needed, these affected only a few scheduled sessions.

2. *A formal procedure was introduced to process surgeons' case-specific requests.* As well as leading on occasion to model infeasibilities, accommodating surgeons' requests to schedule specific cases in given sessions had two undesirable effects. First, they undermined solution efficiency, reducing the space of admissible solutions. Second, they rendered the scheduling process unfair; by requiring that a certain number of their patients be scheduled in a specific session or week, surgeons *de facto* reserved a number of beds for several days. As beds were shared across specialties, this often prevented scheduling of patients in other specialties. On occasion, entire OR sessions would have to remain empty as a consequence of bed shortages. These requests (e.g. to schedule one or two complex surgeries in a morning session) were sometimes fully justified for clinical reasons (e.g. not performing complex procedures when tired or avoiding return of patients to wards after undergoing critical surgery at night, when there is a smaller number of nurses available). Sometimes, however, surgeons would place these requests just to secure capacity. Although aware of the problems this would create, nurses in the Planning Department were not in a position to reject these requests and always tried to accommodate them. To limit opportunistic behaviours and reduce conflicts, it was decided that every surgeon request had to be authorised (via mail) by the OT manager, a senior anaesthesiologist who had the competence and authority to reject them. By its very existence, this authorisation process

dramatically reduced opportunistic behaviours, and arguments between the OT manager and surgeons were indeed very rare. Obviously, this procedure could not prevent surgeons opportunistically exaggerate the severity of the condition of their patient to assign them a higher priority class when they prescribed the surgery in the first place. However, checking all the requests upon arrival was neither feasible, given the number of request processed every day, nor desirable, since reviewing the requests would have likely created conflicts. We are confident that the emphasis put by top management on the importance of the scheduler, the continuous analyses of KPIs and of anomalies in the schedules produced helped to discourage opportunistic behaviours also when filling the surgery request form.

Finally, throughout the implementation process, the weights associated with different terms of the objective function were adjusted several times as a consequence of a change in top management's priorities. The initial focus on scheduling long-waiting patients shifted to increasing OR utilisation and balancing utilisation of beds and OR. Identifying suitable combinations of weights required several rounds of trial and error. As is well known (Cappanera et al., 2016a, 2016b), in goal programming, it is not easy to predict *ex-ante* the effect of given variations in one or more weights on the model's solution. In this regard, software to facilitate sensitivity analysis has proved to be of paramount importance.

6. Achieved outcomes

This AR project led to full implementation of a surgery groups-based scheduling process supported by scheduling software, and all stakeholders in the MSS process are now familiar with this concept. Inside the hospital, surgeries are often referred by the name of their surgery group. The scheduler is used by the Bed Manager and by the head of the Planning Department to assess the impact of every single scheduling decision. Everyone agrees that the scheduling process is now under control and that a better balanced utilisation of the resources has significantly smoothed operations, both in the OT and in the wards.

Surgeons now complete surgical sessions on time, OR overtime is less likely to occur… [OT Manager]

Bed utilisation can now be controlled and critical situations handled on time, taking more informed decisions. We still suffer from a lack of capacity, but dealing with cancellations and postponements is no longer part of my daily activities… [Bed Manager]

Scheduling patients is now much easier. We still feel that there are not enough beds and OR sessions to cope with the massive amount of surgeries we are supposed to schedule, but now we are aware that what we are doing is ok and that with the resources available, it cannot be much improved… [Head of the Planning Department]

These statements can be better understood looking at Figures 1 and 2 which present screenshots of the scheduler output[3]. The left-hand part of Figure 1 reports the ORs and, for each OR session scheduled, the number of unutilised OR minutes. For example, on 01/04/2014 in OR3-Session 2, 360 minutes out of 360 (the

[3] For explanatory purposes, the example was created setting the beds' target utilisation equal to 100%

entire session) is not utilised. The right-hand part refers to the beds, and reports for each ward the beds that will not be utilised. Figure 2 reports the schedule, relevant to April 1st, 2014, indicating for each session the number of surgery to perform and the surgery group[4]. As can be noticed, OR3/session 2, which is dedicated to Neonatal Surgery, could not be used because of a lack of beds on April 2nd , 2014. To assess the feasibility of a change in the schedule, the end-user can just add/change/remove a surgery group from the form reported in Figure 2 and the values reported in Figure 1 will update automatically.

	OR1		OR2		OR3		OR4		Beds		
Day	SESS1	SESS2	SESS1	SESS2	SESS1	SESS2	SESS1	SESS2	DS	WARD1	WARD2
31/03/2014	160		30		66	137		35	0	9	4
01/04/2014	36		12	57	30	360	22	1	0	6	0
02/04/2014	137		25		304		13		0	0	0
03/04/2014	26		24	268	48	37	25		0	1	0
04/04/2014	133		21	46	40	16	59		0	0	0
05/04/2014										0	3
06/04/2014										2	5
07/04/2014	196		12	4	66	162		120	0	1	0
08/04/2014	15		7	27	108		24	60	0	4	0
09/04/2014	40		0		203		13		0	0	0
10/04/2014	3				30	23	5		0	0	1
11/04/2014	12	45	21	17	302	85			0	0	0

Figure 1. Summary relevant to the period 31/03/2014-11/04/2014 – unutilised resources

01/04/2014	OR1	OR2	OR3	OR4
SESS1	URO-0-90-0	EGDS-1-90-1	CHNEO-0-90-2	OCU-0-30-0
SESS1	URO-0-90-0	EGDS-1-60-1	CHNEO-1-120-6	OCU-0-30-0
SESS1	URO-0-90-0	EGDS-0-60-0	CHNEO-0-120-1	OCU-0-30-0
SESS1	URO-0-90-0	EGDS-0-60-0		OCU-0-30-0
SESS1		EGDS-0-60-0		OCU-0-90-1
SESS1		EGDS-0-60-0		OCU-0-60-0
SESS1				OCU-0-60-0
SESS1				
SESS2		ORL-0-90-1		ORTO-0-60-1
SESS2		ORL-0-90-1		
SESS2		ORL-0-90-1		
SESS2				
SESS2				
SESS2				
SESS2				
SESS2				

Figure 2. OR schedule relevant to 01/04/2014

Surgeons who feared that implementation of the scheduler would reduce their autonomy no longer see it as a threat. Instead, they recognise that the assignment of shared resources (beds) to different surgical specialties is now fairer, as it is informed by shared criteria and accomplished using a mathematical model. In addition, the utilisation of surgery groups still leaves room to select the cases for surgery and to sequence them within the OR session. This is recognised as one of the key factors for successful implementation of an MSS approach, whatever the model used (Van Oostrum et al., 2010).

[4] Please notice that some sessions are totally (e.g. OR1/Sess2) or partially (e.g. OR4/Sess2) used for urgent cases or for specialties (e.g. Neuro Surgery) that are not scheduled using the model.

The general and medical directors, whose commitment was fundamental to overcoming initial resistance to change within the hospital, were satisfied with the outcomes achieved (see next paragraph) and decided to extend the collaboration with the IBIS Lab to other functional areas such as the Department of Diagnostic Imaging and the Emergency Department.

6.1 Before and after implementation comparison

In this section, we compare the performance in the period January 2013-October 2013 (before tool implementation), and in the period November 2013-July 2014 (after tool implementation). Before presenting the results in details, it is important to point out that due to some external factors, in the two periods we registered a significant change both in the demand for surgeries and in the hospital available capacity. More precisely, total demand for surgeries increased by an average of 12.73 surgeries per week (+10.31%) during the period November 2013–July 2014 due to an increase in demand for day surgeries (+18.51%). In that same period, the weekly number of available OR hours decreased from 201.9 to 176 hours per week (-12.8%). This followed a budget cut arising from the financial crisis faced at that time by the Regional Health System, which forced the hospital to reduce the number of surgical sessions. (In the region, elective surgeries absorb 20% of the Health System's total resources (Nuti et al., 2012)).

Tables 3 and 4 report the values, before and after implementation, of the KPIs presented in Section 3 and a measure of their (absolute and relative) variation. Since our data were not normally distributed, to statistically assess the significance of the difference between before and after performance we couldn't compare the KPIs values – that, except for the percentage of day/inpatient surgeries and for the overbooking rate, are mean values - using a parametric test (e.g. a t-test). Instead, we compared:

- the *weekly number of surgeries/day surgeries/inpatient surgeries,* the *weekly number of surgeries per available OR hour,* the *weekly OR/bed utilisation,* the *OR/bed utilisation rates,* the *LoS* and the *ST,* using the independent 2-group Mann-Whitney U Tests,
- the *percentage of day/inpatient surgeries* and the *bed overbooking rates* using the χ-squared test for the proportions.

Significant before-after differences are identified in the last column of Tables 3 and 4.

Table 3 shows that, despite the mentioned adverse external factors, thanks to the scheduler, neither the weekly number of surgeries nor the weekly OR utilisation decreased significantly (-0.052% and -1.64% on average, respectively) despite the 12.8% OR capacity reduction.

	Before	After	Variation	Sig
Throughput [Surgeries]				
Average weekly number of surgeries [surgeries/week]	96.6	96.1	-0.05 (-0.052%)	n.s.
Average weekly number of surgeries per available OR hour [surgeries/hours]	0.48	0.54	+28.8%	***
Average weekly number of day surgeries	46.46	53.34	+6.88 (+14.80%)	***
Average weekly number of inpatient surgeries	50.14	42,76	-7.38 (-14.71%)	***
Percentage of day surgeries [%]	48.1	55.5	+7.4 (+15.38%)	***
Percentage of inpatient surgeries [%]	51.9	44.5	-7.4% (-14,26%)	***
OR utilisation				
Average weekly OR utilisation [hours/week]	140	137.7	-2.3 (-1.64 %)	n.s.
Average OR utilisation rate [%]	69.3	78.1	+8.8 (12.7%)	***
Bed utilisation (weekends included)				
Average inpatient bed utilisation rate [%]	74.6	72.5	-2.1 (-2.81%)	n.s.
Average day surgery bed utilisation rate [%]	66.42	77.14	+10.72 (+16.14%)	***
Inpatient bed overbooking rate %	9.23%	1.07%	-8.16 (-88.4%)	***
Day surgery bed overbooking rate %	6.65%	2.2%	-4.45 (-66.9%)	n.s.
n.s. p>0.05, *** p<0.001				

Table 3. Before and after comparison: resource utilisation

This is because the OR utilisation rate increased by 12.7% and the average number of surgeries executed per available OR hour by 28.8%. The utilisation rate for day surgery beds increased by 16.14% while the inpatient bed utilisation rate decreased slightly (-2.81 %). As can be noticed, the utilisation rates (especially for inpatients beds) both before and after the scheduler implementation are not extremely high. This is mostly due to patient-driven cancellations/postponements of surgeries, typical of children's hospitals, which have a tremendous impact on the bed saturation. Kids fall ill very easily and a common cold can determine the postponement of a surgery. Despite the use of surgery groups allows better managing patient-driven cancellations, it is not always possible to replace cancelled surgeries with others with the same expected LoS and ST. In fact, quite often they are replaced with surgeries with shorter LoS, which hampers the bed utilisation. Moreover, if a patient falls ill the day of the surgery or the day before, it is not possible to arrange a replacement. Indeed, one of the most significant benefits achieved with the scheduler, was to reduce the hospital-driven cancellations due to bed overbooking while not decreasing the resource utilisation. The overbooking rate reported in Table 3 indicates, as a percentage, the proportion of days when the number of occupied beds was higher than the capacity reserved for elective cases. As can be observed, this significantly decreased for the inpatients bed (-88.4%) while for the day surgery beds the reduction (-66.9%) was not statistically significant. Finally, it should be noted that a management decision to treat certain cases as day surgeries that had previously been classified as inpatient surgeries led to a significant increase in the percentage of day surgeries performed.

Overall, it can be argued that the project led to an overall increase in OR efficiency. As we did not wish to increase OR efficiency merely by scheduling simpler surgeries, it is important to demonstrate that the complexity of surgeries performed did not decrease significantly. To prove that, consider the mean value ST and LoS of surgeries performed as proxies of their complexity. In practice, it is more difficult to slot longer surgeries into surgical sessions. Second, it is more complex to schedule patients who will occupy beds for several days rather than for one or a few days. Finally, surgeons may need to perform a long and difficult procedure as the first surgery of the morning session and/or may not want any subsequent surgeries to be scheduled. It follows that scheduling long surgeries is likely to reduce the model's degrees of freedom.

Table 4 shows that, for inpatient surgeries, both ST and LoS increased (+6% and +7.41%, respectively). For day surgeries, however, ST and LoS did not change significantly in absolute value (+0.09 min and -0.03 days, respectively). The average LoS of the day surgeries is larger than zero, implying that certain patients were kept under observation for one or more days rather than being discharged on the same day as expected. These data confirm that the previously mentioned gain in efficiency was not linked to any diminished complexity of the scheduled surgeries, confirming the long-term orientation of the schedules produced.

	Before	**After**	**Variation**	**sig**
Surgical time inpatient surgeries				
Mean surgical time [min]	97.96	103.94	+5.98 (+6%)	*
LoS inpatient surgeries				
Mean LoS [day]	2.7	2.9	+0.2 (+7.41%)	***
Surgical time day surgeries [min]				
Mean surgical time [min]	52.28	52.37	+0.09 (+0.17%)	n.s.
LoS day surgeries				
Mean LoS [day]	0.05	0.02	-0.03 (-60%)	***
n.s. p>0.05; * p<0.05; *** p<0.001				

Table 4. Before and after comparison: complexity of surgeries

Another objective of this study was to balance utilisation of resources (principally beds) to achieve robustness. In fact, it has been shown (Cappanera et al., 2014) that by balancing bed daily utilisation, it is possible to absorb the unexpected peaks caused by LoS variability and so prevent overbooking cancellations. To assess whether such a result was achieved, Table 5 presents relevant descriptive statistics for daily bed utilisation, for both day surgery and inpatient beds.

	BEFORE	**AFTER**	**VARIATION**
Inpatient surgery beds unit			
Beds available [beds]	28	28	0 (0%)
Mean bed utilisation [beds]	20.9	20.3	-0.6 (-0.03%)
St Dev bed utilisation [beds]	6.1	4.3	-1.8 (-29.5%)
Max bed utilisation [beds]	35	30	-5 (14.3%)
Day surgery beds unit			
Beds available [beds]	14	14	0
Mean bed utilisation [beds]	9.3	10.8	+1.5 (16.1%)
St Dev bed utilisation [beds]	3.8	3.1	-0.7 (-18.4%)
Max bed utilisation [beds]	24	16	-8 (-33%)

Table 5. Before and after comparison: bed utilisation descriptive statistics

Looking first at inpatient beds, we can observe a slight decrease (-0.03%) in the mean value of bed utilisation, but there is a substantial decrease of the standard deviation (-29.5%), indicating a considerable improvement in balancing of inpatient beds. The impact of such an improvement on the robustness of the schedule is well captured by the mentioned decrease in the overbooking rate (see Table 4). Moreover, we can observe a substantial decrease (from 35 to 30) in the maximum value for bed utilisation, which is further evidence of better balancing and higher robustness. The results are even better for the day surgery unit, with more effective bed balancing and an increase in the mean value of bed utilisation.

Finally, it is worth noting that these results were achieved by giving higher priority to patients with a closer due date while also reducing the previously mentioned backlog of long-waiting (inpatient) patients. In fact, in the post-implementation period, 390 patients with expired due dates were scheduled, as compared to only 192 before implementation. In other words, the implemented scheduling process was also equitable.

7. Discussion

In this section, we present the lessons learned from this AR project and discuss the contribution of the present study in light of previous research on MSS.

7.1 MSS model and scheduler characteristics

During the AR project, we developed some insights into the features that can make an MSS model and scheduler more effective and easy to use. These can be summarised as follows.

1. *Scheduling surgery groups*. Several authors (Santibanez et al., 2007; Van Oostrum et al., 2008; Belien et al., 2009) have highlighted the importance of considering downstream resources (such as post-surgical beds) when creating the master surgical schedule. The present study contributes to existing knowledge by suggesting that this can be achieved by clustering patients in homogeneous surgery groups and by scheduling according to those groups. Indeed, by considering downstream resource bottlenecks at an early stage of the surgical scheduling process, we achieved higher throughput and better resource utilisation and balancing. Our clinical observation of the implementation process also revealed that the use of surgery groups-based scheduling made the task much easier for nurses in the Planning Department in terms of (i) the creation of ready-to-be-operated patient lists (see Section 4.1); (ii) patient selection and sequencing and (iii) the management of schedule disruptions. By organising waiting lists by surgery group, nurses could easily identify those groups with a higher number of patients with imminent due-dates and give them higher priority, ensuring that schedules were more equitable. In addition, these nurses were relieved of the burden of monitoring patient due dates while at the same time reducing the number of patients scheduled late. In this regard, it is worth noting that the introduction of an earliest programmable date (see Section 5.4) and allowing patients to enter the waiting list only when they were actually eligible for surgery further helped to filter waiting lists by focusing only on a meaningful subset of patients.
2. *Relying on a flexible model and scheduler*. The needs of surgical process stakeholders inevitably change over time, creating a need for scheduling tools that can be easily reconfigured even by end-users with no operational research background. Most previous studies presented models tailored to well-defined hospital settings but without this dynamic perspective. Others (e.g. Dios et al., 2015) suggest to develop modular schedulers allowing to easily “swap” the model “module” when the

stakeholders' needs change. We suggest that not only the software needs to be flexible to allow changing the model, but also the model needs to be flexible to allow the end user to add-remove constraints and to modify the objective function without the support of any developer. Here, we achieved flexibility in two ways. First, we developed a multi-criteria optimisation model that takes account of a relatively large number of possible constraints and objective functions. The model was designed in such a way that changing the relative importance of objective functions and/or activating or deactivating constraints requires only the setting of a few parameters. Second, we embedded the model in a flexible scheduler, allowing easy setting of model parameters and assessment of the effect of different parameter combinations on the schedule before commencement of the planning horizon. Subsequently, throughout the planning horizon, it allows users to assess and even to *visualise* the effects of changes that might be requested to address contingent problems related to the OR and bed utilisation. By exploring different scenarios, both before and after schedule implementation, end users can develop a greater sense of the consequences of their decisions. As many previous studies do not address implementation issues (see Table 1), they underestimate the need for flexible tools and fail to assess how a lack of user-friendliness can hamper model utilisation in real settings.

3. *Dealing with surgeons' preferences and using hard constraints***.** Granting stakeholders (principally surgeons) the possibility of scheduling specific cases or types of cases in certain sessions of the planning horizon is often the key to their willingness to accept a software-generated schedule (van Oostrum et al., 2010). This can be obtained by using suitable hard constraints. However, such constraints cause two problems. First, they lead to model infeasibilities that end users may not be able to address, making the scheduling tool unusable. Second, they can encourage opportunistic behaviours. Van Oostrum et al. (2010) suggested avoiding the use of soft constraints in place of hard ones, but they offer little insight into how best to manage the trade-off between acceptability of the solution for the stakeholder and usability of the scheduling tool for end users. In this regard, our research suggests that hard constraints should be used to accommodate surgeons' requests, but it is also necessary to provide other stakeholders with the right tools for dealing with them. We suggest first that an *organisational filter* should be set up by appointing a specialist with sufficient authority to evaluate the appropriateness of surgeons' requests and to prevent opportunistic behaviours. Second, we suggest designing the scheduler in such a way that, whenever an infeasibility occurs, the end user can (i) identify the constraints that caused it and (ii) check whether these infeasibilities can be removed by making some adjustments (e.g. by accepting the risk of overbooking beds). As well as facilitating the scheduling process, these alerts can help to increase user knowledge about the process itself.
4. *Adjusting the length of the planning horizon.* In general, longer planning horizons better enable assignment of patients to appropriate slots and support increasing utilisation (Van Oostrum et al., 2010). However, some patients may end up waiting too long if postponed at the end of the planning horizon. The present findings suggest that when resource availability is subject to frequent and unexpected variation, it is preferable to shorten the planning horizon, making adjustments every time

a change occurs and then running the model again. In these cases, longer planning horizons (e.g. one month) are likely to require multiple adjustments that will cascade, influencing each other and leading eventually to a poorly implemented solution.

7.2 Scheduler implementation

During the AR project, certain actions proved fundamental to the successful introduction and use of the scheduler, principally in terms of overcoming stakeholder scepticism about the model and changing their established way of working (Wilson, 1981; Blake et al., 2002; Eldabi, 2009). These can be summarised as follows.

1. *Gaining the commitment of top management by showing credible preliminary results.* The surgical scheduling process strongly influences hospital performance and directly affects the daily routines of many personnel. These routines are often a consequence of years of negotiation, and without a strong commitment from top management, it is almost impossible to change them (Blake et al., 2002). Good preliminary results proved very helpful in increasing the commitment of top management. In this regard, we found it very useful to present the results of simulations in which we assessed, in a stochastic environment, the performance that might be achieved by using schedules produced with the optimisation model.
2. *Inferring stakeholder preferences by enabling them to comment on tentative schedules.* One of the hurdles in creating commitment among stakeholders of a scheduling process is that it is difficult to develop a satisfactory model based only on preliminary interviews (Blake et al., 2002). By creating tentative schedules and asking stakeholders to comment on them, we were able to develop a good understanding of their preferences and needs. As explained in Section 5.2, stakeholder feedback about the schedules produced by the three models (and even after scheduler implementation) enabled us to change the model's objective function and constraints several times. These changes addressed issues not explicitly mentioned at the outset, such as the need to achieve better bed-OR balance, the need to schedule additional short surgeries if there was space, the need to avoid scheduling too many day surgeries and to minimise the occurrence of infeasible solutions.
3. *Introducing changes gradually*. Even with the strong commitment of top management, resistance to change can be strong. For this reason, it is advisable to introduce change gradually, moving to the next step only when most stakeholders clearly understand the benefits, *for the hospital* and *for themselves*, of the proposed change. This aligns with Framinan and Ruiz (2012) and Carvahlo et al. (2014), who suggested that such incremental deployment is advisable. As explained in Section 5.4, the first schedules were developed by the IBIS Lab, and adjustments and *ex-post* changes were managed in collaboration with hospital staff before handing over full responsibility for the process to nurses and bed managers. Based on the AR team feedback, examples of the perceived benefits at individual level were as follows.
    - Top management: higher revenues, reduced costs, higher patient satisfaction;

- Surgeons: less overtime work, more balanced workload across sessions, higher control of the waiting list;
- Bed managers: fewer emergency situations to manage, better balanced utilisation of beds in surgical units across the week, higher control of bed occupancy;
- Planning Department nurses: faster selection and sequencing of patients, easier replacement of unavailable patients, fewer complaints from surgeons and patients' parents;
- Operating theatre manager: better balancing of work across OR sessions, less overtime work, fewer complaints from surgeons.

4. *Involving staff at lower levels of the hospital hierarchy.* Hospitals are hierarchical organisations in which employees other than managers and physicians are seldom considered when making decisions about critical processes such as surgical scheduling (Aas, 1997). However, while managers and physicians play a central role in identifying the desired outcomes of these processes, our findings suggest that they are not fully aware of the constraints and problems affecting process execution. Nursing, IT and administrative staff can provide very useful insights and should be involved in the scheduling tool development process from an early stage.

8. Conclusions and future research

This study sets out to improve the MSS process at a leading European hospital. As an AR project, it contributes to solving a practical problem, as well as to knowledge.

The practical contribution of this study was to improve the Meyer Hospital scheduling process, in terms of efficiency, robustness, long-term orientation and equity. These improvements were achieved by reengineering the hospital's surgical process and by developing and implementing an MSS scheduler to support such a process. The scheduler embeds a novel, purpose-built mixed-integer programming model. In terms of theoretical findings, this research contributes to literature on MSS analysing those features that make an MSS optimisation model and scheduler effective and easy to use, and the actions facilitating scheduler implementation.

Reasoning about the contextual conditions at the Meyer Hospital can help to understand the generalisability of these results and the extent to which they might be transferable to other contexts. This can also help to indicate future research directions. The Meyer Hospital is characterised by (i) frequent need to reschedule surgical activities (e.g. to deal with children getting sick); (ii) a low volume of emergencies and urgencies and (iii) few critical resources (surgical teams, beds and ORs). The findings here are therefore useful for those hospitals where there is critical reliance on a robust surgical scheduling process, with no significant interaction between elective and urgent surgical activities (e.g. because the latter involve dedicated resources, as at the Meyer Hospital). To the best of our knowledge, these characteristics are shared by most children's hospitals, as well as by certain other types of hospital and surgical facility. Future research should consider addressing the MSS issue in other hospitals with different characteristics. Implementation in other settings might require consideration of resources such as intensive care units (that are becoming increasingly more critical (Azadeh et al., 2016)), electromedical devices and anaesthetists, which were not a cause of bottlenecks at the Meyer Hospital but may be critical elsewhere. We expect our own future research to extend the scope of application of our scheduling tools.

Acknowledgements

We are grateful to Meyer University Children's Hospital, and in particular to General Director Dr. Alberto Zanobini, and to the Medical Director Dr. Francesca Bellini for supporting the research project that has inspired this study and for authorizing us to disclose the Meyer Hospital name in the manuscript. We are also grateful to IBIS Lab for authorizing us to disclose the lab name in the manuscript.

References

AAS, I. M. 1997. Organizational change: decentralization in hospitals. The International journal of health planning and management, 12(2), 103-114.

AGNETIS, A., COPPI, A., CORSINI, M., DELLINO, G., MELONI, C. & PRANZO, M. 2012. Long term evaluation of operating theater planning policies. Operations Research for Health Care, 1, 95-104.

AGNETIS, A., COPPI, A., CORSINI, M., DELLINO, G., MELONI, C. & PRANZO, M. 2014. A decomposition approach for the combined master surgical schedule and surgical case assignment problems. Health Care Management Science, 17, 49-59.

ARINGHIERI, R., LANDA, P., SORIANO, P., TÀNFANI, E. & TESTI, A. 2015. A two level metaheuristic for the operating room scheduling and assignment problem. Computers & Operations Research, 54, 21-34.

AZADEH, A., TOHIDI, H., ZARRIN, M., PASHAPOUR, S., & MOGHADDAM, M. 2016. An integrated algorithm for performance optimization of neurosurgical ICUs. Expert Systems with Applications,43, 142-153.

BANDITORI, C., CAPPANERA, P. & VISINTIN, F. 2013. A combined optimization–simulation approach to the master surgical scheduling problem. IMA Journal of Management Mathematics, 24, 155-187.

BELIËN, J., DEMEULEMEESTER, E. & CARDOEN, B. 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., DEMEULEMEESTER, E. & CARDOEN, B. 2009. A decision support system for cyclic master surgery scheduling with multiple objectives. Journal of Scheduling, 12, 147-161.

BLAKE, J. T. & DONALD, J. 2002. Mount Sinai hospital uses integer programming to allocate operating room time. Interfaces, 32, 63-73.

BRAILSFORD, S., HARPER, P., PATEL, B. & PITT, M. 2009. An analysis of the academic literature on simulation and modelling in health care. Journal of Simulation, 3, 130-140.

BRAILSFORD, S. & VISSERS, J. 2011. OR in healthcare: A European perspective. European Journal of Operational Research, 212, 223-234.

CAPPANERA, P., VISINTIN, F. & BANDITORI, C. 2014. Comparing resource balancing criteria in master surgical scheduling: A combined optimisation-simulation approach. International Journal of Production Economics, 158, 179-196.

CAPPANERA, P., VISINTIN, F. & BANDITORI, C. 2016a. A goal-programming approach to the master surgical scheduling problem. In Proceedings of the Second International Conference on Health Care Systems Engineering, 27-29 May, Lyon, France.

CAPPANERA, P.; VISINTIN, F. & BANDITORI, C., 2016b. Addressing conflicting stakeholders' priorities in surgical scheduling by goal programming. Flexible Services And Manufacturing Journal, pp. 1-20, DOI: 10.1007/s10696-016-9255-5

CARDOEN, B., DEMEULEMEESTER, E. & BELIËN, J. 2010. Operating room planning and scheduling: A literature review. European Journal of Operational Research, 201, 921-932.

CARTER, M. W. & KETABI, S. 2012. Bed Balancing in Surgical Wards via Block Scheduling. Journal of Minimally Invasive Surgical Sciences, 1, 129-137.

CARVALHO, A. N., SCAVARDA, L. F. & LUSTOSA, L. J. 2014. Implementing finite capacity production scheduling: lessons from a practical case. International Journal of Production Research, 52, 1215-1230.

CHOW, V. S., PUTERMAN, M. L., SALEHIRAD, N., HUANG, W. & ATKINS, D. 2011. Reducing surgical ward congestion through improved surgical scheduling and uncapacitated simulation. Production and Operations Management, 20, 418-430.

COUGHLAN, P. & COGHLAN, D. 2002. Action research for operations management. International Journal of Operations & Production Management, 22, 220-240.

DENTON, B., VIAPIANO, J. & VOGL, A. 2007. Optimization of surgery sequencing and scheduling decisions under uncertainty. Health Care Management Science, 10, 13-24.

DIOS, M, MOLINA, J.M., ANDRADE, J.L., FERNANDEZ-VIAGAS, V., FRAMINAN, J.M., 2015. A Decision Support System for Operating Room Scheduling, Computers & Industrial Engineering, 88, pp. 430-443.

ELDABI, T. 2009. Implementation issues of modeling healthcare problems: misconceptions and lessons. Winter Simulation Conference. Winter Simulation Conference, 1831-1839.

EWEN H., MÖNCH L., 2014. A simulation-based framework to schedule surgeries in an eye hospital, IIE Transactions on Healthcare Systems Engineering, 4 (4), pp. 191-208.

FAROOQ S., O'BRIEN C., 2015. An action research methodology for manufacturing technology selection: a supply chain perspective, Production Planning & Control, 26:6, 467-488

FORREST, J. & LOUGEE-HEIMER, R. CBC user guide.

FOURER, R., GAY, D. & KERNIGHAN, B. 1993. AMPL, Boyd & Fraser.

FRAMINAN, J. M. & RUIZ, R. 2012. Guidelines for the deployment and implementation of manufacturing scheduling systems. International Journal of Production Research, 50, 1799-1812.

FÜGENER, A., HANS, E. W., KOLISCH, R., KORTBEEK, N. & VANBERKEL, P. T. 2014. Master surgery scheduling with consideration of multiple downstream units. European Journal of Operational Research, 239, 227-236.

GLOUBERMAN, S. & MINTZBERG, H. 2001. Managing the care of health and the cure of disease-Part I: Differentiation. Health care management review, 26, 56-69.

HARPER, P. & PITT, M. 2004. On the challenges of healthcare modelling and a proposed project life cycle for successful implementation. Journal of the Operational Research Society, 55, 657-661.

HART, W. E., LAIRD, C., WATSON, J.-P. & WOODRUFF, D. L. 2012. Pyomo–optimization modeling in python, Springer Science & Business Media.

HUBERT, M. & VANDERVIEREN, E. 2008. An adjusted boxplot for skewed distributions. Computational statistics & data analysis, 52, 5186-5201.

KHARRAJA, S., ALBERT, P. & CHAABANE, S. Block scheduling: Toward a master surgical schedule. Service Systems and Service Management, 2006 International Conference on, 2006. IEEE, 429-435.

JEBALI, A., DIABAT, A., 2015. A stochastic model for operating room planning under capacity constraints. International Journal of Production Research, 53(24), pp.7252–7270.

LARSSON, A., 2013. The accuracy of surgery time estimations, Production Planning & Control, 24:10-11, 891-902

LIN, Q. L., LIU, L., LIU, H. C., & WANG, D. J. 2013. Integrating hierarchical balanced scorecard with fuzzy linguistic for evaluating operating room performance in hospitals. Expert Systems with Applications, 40(6), 1917-1924.

MA, G. & DEMEULEMEESTER, E. 2013. A multilevel integrative approach to hospital case mix and capacity planning. Computers & Operations Research, 40, 2198-2207.

MAHDAVI, M., MALMSTRÖM, T., VAN DE KLUNDERT, J., ELKHUIZEN, S. & VISSERS, J. 2013. Generic operational models in health service operations management: A systematic review. Socio-Economic Planning Sciences, 47, 271-280.

MANNINO, C., NILSSEN, E. J. & NORDLANDER, T. E. 2012. A pattern based, robust approach to cyclic master surgery scheduling. Journal of Scheduling, 1-11.

MARAZZI, A., PACCAUD, F., RUFFIEUX, C. & BEGUIN, C. 1998. Fitting the distributions of length of stay by parametric models. Medical care, 36, 915-927.

MAY, J. H., SPANGLER, W. E., STRUM, D. P. & VARGAS, L. G. 2011. The surgical scheduling problem: Current research and future opportunities. Production and Operations Management, 20, 392-405.

MAY, J. H., STRUM, D. P. & VARGAS, L. G. 2000. Fitting the Lognormal Distribution to Surgical Procedure Times*. Decision Sciences, 31, 129-148.

ORAL, M. 2012. Action research contextualizes DEA in a multi-organizational decision-making process. Expert Systems with Applications, 39(7), pp.6503-6513.

PROUDLOVE, N., BLACK, S. & FLETCHER, A. 2007. OR and the challenge to improve the NHS: modelling for insight and improvement in in-patient flows. Journal of the Operational Research Society, 58, 145-158.

RICH, N., PIERCY N., 2013. Losing patients: a systems view on healthcare improvement, Production Planning & Control, 24:10-11

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.

SACCANI, F., VISINTIN, F., MANSINI, R., COLOMBI M., 2016. Improving spare parts management for field services. A model and a case study for the repair kit problem, IMA Journal of Management Mathematics, 28, 185-204.

SAMUDRA, M., VAN RIET, C., DEMEULEMEESTER, E., CARDOEN, B., VANSTEENKISTE, N., RADEMAKERS, F.E., 2016. Scheduling operating rooms : achievements , challenges and pitfalls. Journal of Scheduling, 19(5), pp.493–525.

STEPANIAK, P. S., HEIJ, C., MANNAERTS, G. H., DE QUELERIJ, M. & DE VRIES, G. 2009. Modeling procedure and surgical times for current procedural terminology-anesthesia-surgeon combinations and evaluation in terms of case-duration prediction and operating room efficiency: a multicenter study. Anesthesia & Analgesia, 109, 1232-1245.

SU M.C., LAI S.C., WANG P.C., HSIEH Y.Z., LIN S.C. 2011. A SOMO-based approach to the operating room scheduling problem, Expert Systems with Applications, 38, 15447–15454

TÀNFANI, E. & TESTI, A. 2010. A pre-assignment heuristic algorithm for the Master Surgical Schedule Problem (MSSP). Annals of Operations Research, 178, 105-119.

VAN OOSTRUM, J. M., BREDENHOFF, E. & HANS, E. W. 2010. Suitability and managerial implications of a master surgical scheduling approach. Annals of Operations Research, 178, 91-104.

VAN OOSTRUM, J. M., VAN HOUDENHOVEN, M., HURINK, J., HANS, E., WULLINK, G. & KAZEMIER, G. 2008. A master surgical scheduling approach for cyclic scheduling in operating room departments. OR Spectrum, 30, 355-374.

VIRTUE, A., CHAUSSALET, T. & KELLY, J. 2013. Healthcare planning and its potential role increasing operational efficiency in the health sector: A viewpoint. Journal of Enterprise Information Management, 26, 8-20.

VISINTIN, F., CAPPANERA, P. & BANDITORI, C. 2016. Evaluating the impact of flexible practices on the master surgical scheduling process: an empirical analysis. Flexible Services and Manufacturing Journal vol. 28, pp. 182-205.

VISSERS, J., ADAN, I. & BEKKERS, J. 2005. Patient mix optimization in tactical cardiothoracic surgery planning: a case study. IMA Journal of Management Mathematics, 16, 281-304.

WHITE, S. A. 2004. Introduction to BPMN. IBM Cooperation, 2, 0.

WILSON, J. T. 1981. Implementation of computer simulation projects in health care. Journal of the Operational Research Society, 825-832.

YAHIA, Z., HARRAZ, N. & ELTAWIL, A. B. Building master surgery schedules with leveled bed occupancy and nurse workloads. Industrial Engineering and Engineering Management (IEEM), 2014 IEEE International Conference on, 2014. IEEE, 89-93.

ZHANG, B., MURALI, P., DESSOUKY, M. M. & BELSON, D. 2008. A mixed integer programming approach for allocating operating room capacity. Journal of the Operational Research Society, 60, 663-673.

APPENDIX A

In this section, we provide details about the Meyer Hospital's information systems and the study's data sources.

The data used here were acquired from several different sources. For simplicity, data can be classified as pre-admission and post-admission data. Pre-admission data include the information recorded on surgery request forms and anaesthetist pre-admission reports. In 2011, these documents were all handwritten. Completed surgery request forms were sent or brought directly to the Planning Department by surgeons. Planning Department personnel then placed these in plastic folders and archived them in special binders. Each binder was associated with a different specialty and contained requests relating to all the patients waiting for surgery of that specialty—in other words, a specialty-based waiting list. Similarly, pre-admission reports were handed by anaesthetists to one of the nurses in the Planning Department, who would again place them in the same plastic folder containing the patient's surgery request form. Since 2013, the data contained in surgery request forms and pre-admission reports (which are still handwritten) are input by the Planning Department's nurses, using a waiting list management software. This is subsequently used to retrieve the requisite information to schedule pre-admission tests and surgeries, to update patient status and to control the waiting list. The data stored in the software database are now used to input the scheduler.

Post-admission data were (and still are) collected once the patient is admitted; these can be subdivided into data pertaining to surgical activity and those pertaining to patient stay. The former includes several timestamps of relevance to surgical activities, including induction, surgery execution, operating surgeon, diagnosis and procedure performed. These data are directly input by surgeons in the OR and stored in a standalone database. Data pertaining to patient stay include admission and discharge data and the surgical unit where patients were hospitalised. These data are stored in another database.