An optimization based on simulation approach to the patient admission scheduling problem using a linear programing algorithm

CrossMark

C. Granja [a,c,d], B. Almada-Lobo [b], F. Janela [d], J. Seabra [d], A. Mendes [a,*]

[a] LEPABE, Department of Chemical Engineering, Faculty of Engineering, University of Porto, Portugal
[b] INESC-TEC, Faculty of Engineering, University of Porto, Portugal
[c] Norwegian Centre for Integrated Care and Telemedicine, University Hospital of North Norway, Norway
[d] Siemens S.A. Healthcare Sector, Portugal

ARTICLE INFO

ABSTRACT

Background: As patient's length of stay in waiting lists increases, governments are looking for strategies to control the problem. Agreements were created with private providers to diminish the workload in the public sector. However, the growth of the private sector is not following the demand for care. Given this context, new management strategies have to be considered in order to minimize patient length of stay in waiting lists while reducing the costs and increasing (or at least maintaining) the quality of care.
Method: Appointment scheduling systems are today known to be proficient in the optimization of health care services. Their utilization is focused on increasing the usage of human resources, medical equipment and reducing the patient waiting times. In this paper, a simulation-based optimization approach to the Patient Admission Scheduling Problem is presented. Modeling tools and simulation techniques are used in the optimization of a diagnostic imaging department.
Results: The proposed techniques have demonstrated to be effective in the evaluation of diagnostic imaging workflows. A simulated annealing algorithm was used to optimize the patient admission sequence towards minimizing the total completion and total waiting of patients. The obtained results showed average reductions of 5% on the total completion and 38% on the patients' total waiting time.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

It is common-knowledge that in the last decade there has been an increasing demand for health care services. Among others, this arises from an ageing population and an increasing awareness for preventative care. On the other hand, the world economic situation is leading to a reorganization of healthcare systems. At a macro level, what is noticeable is a centralization and decrease in the number of public healthcare providers [1]. In practice, and mainly in the department of medical imaging, this reflects an inability of healthcare providers to respond to an ever increasing demand. As the patient's length of stay in waiting lists increases, governments are looking for strategies to control the problem. In Portugal, to diminish the workload in the public sector, conventions were created with private providers in which patients can access examinations without losing the benefits of the National Health Service

(SNS, from the Portuguese Serviço Nacional de Saúde). However, the growth of the private sector is not following the demand for care. In this context, new management strategies have to be considered to minimize the patient length of stay in waiting lists, while reducing the costs and increasing the quality of care.

The study of workflows attempts to understand the process and remove components, without added value, which delay the service and introduce complexity that ultimately may result in errors. The opportunities that this type of study has in medical imaging are significant, given that the majority of workflows in medicine have yet to be consistently described [2], which is also the case for medical imaging.

Appointment scheduling systems are today known to be proficient in the optimization of health care services. Their utilization is focused on increasing the usage of human resources, medical equipment and reducing the patient waiting times. This paper considers the appointment scheduling problem in an imaging clinic. The considered imaging clinic provides services on Computed Tomography (CT), MRI, Radiology (RX), Orthopantomography (OT), Densitometry (PX), Mammography (MG) and Ultrasound (US). Appointments are requested by the patient either by phone

* Corresponding author. Address: LEPAE, Department of Chemical Engineering, Faculty of Engineering, University of Porto, Rua Roberto Frias S/N, 4200-465 Porto, Portugal. Fax: +351 22 508 14 49.
  E-mail address: mendes@fe.up.pt (A. Mendes).

or in person. At the moment of request the patient appointment is scheduled for a day and time according to the timetable availability and patient's preference. The size of the time block, i.e. processing time, is defined according to the modality. The aim was to improve the patients' admission scheduling in order to minimize the patients' waiting time, and increase patient throughput. Thus, the problem studied herein considered elective patients and was modeled as a static, multi-stage/multi-server system, with processing times estimated according to the type of clinical examination and defined as a probability distribution. A discrete-event simulation model integrated with an optimization technique was used to minimize the patients' waiting time and increase patient throughput.

This paper is divided in eight sections. In the first section, the problem object of study is presented. In the second section, the problem is defined and the state of the art is reported. In the third section, the case study is described. In the fourth section, the resources' and tasks' characterization parameters are presented. In the fifth section, the modeling approach is explained and the simulation technique described. In the sixth section, the optimization algorithm with which the results were obtained is presented. In the seventh section the optimization results are interpreted and presented. Section eight draws conclusions on the results and foresees future work.

2. Problem definition and literature review

2.1. Environmental factors

The first reference to the Admission Scheduling (AS) problem is from Bailey [3] and dates from 1952. In this work, the problem was addressed as a queuing system with the objective of minimizing patients' waiting time. Bailey concluded that the best solution for the problem was to schedule patients in regular intervals, equal to the average consultation time. The AS problem definition was first addressed by Cayirli and Veral [4] and is briefly described hereafter. AS problems consider the finding of an appointment schedule in a healthcare environment for which an objective function, considering one or multiple performance measures, is optimized. If AS is considered as a queuing system, these performance measures have an essential role as, in queuing systems, issues are often related to the patient. The problem definition is primarily divided into two main categories: static and dynamic. It is considered a static problem when all the decisions are made *a priori*. This means that the proposed appointment system does not consider the system current state. Static problems are the most common type of problems in healthcare environment [3,5–7].

In contrast to this, in a dynamic problem the appointment system is reviewed based on its current state [8–11]. In both categories, problems are further defined according to environmental factors defined in the following sub-sections, namely: number of stages – number of tasks required to complete the process, number of servers – number of available resources with competences to complete the same task, patient tardiness and processing times.

2.1.1. Number of stages

(i) *Single-stage:* system where patients queue for a single stage. The majority of the studies in literature represent single-stage problems. Schemes, illustrating single-stage systems, are provided in Fig. 1(a) and (b).
(ii) *Multi-stage:* system where patients queue for multiple stages, such as registration, examination and checkout. Schemes illustrating multi-stage systems are provided in Fig. 2(a) and (b). A multi-stage system was considered by
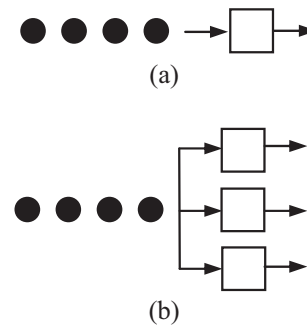

Fig. 1. Queuing system: (a) single-stage/single-server system; (b) single-stage/multi-server system.

Garg et al. [5]. In this work, a hospital scenario was studied and the patient pathway was modeled to consider different phases of care, such as acute, treatment and rehabilitation. In the workflow modeled by Granja et al. [12], three stages were considered in the patient's pathway in radiology: admission, examination and billing. Chen et al. [13] considered four stages in their analysis of surgery admission. Connelly and Bair [14] explored the potential of discrete event simulation in the operations analysis in an emergency department. In their work, a multi-stage system was considered as each patient was modeled as a set of instructions that define a series of individual activities that must be completed in a given order before the patient leaves de emergency department.

2.1.2. Number of servers

(i) *Single-server:* appointment system regards a specific server. Schemes illustrating single-server systems are provided in Fig. 1(a) and Fig. 2(a). They are the most predominant in literature [15,16]. Even if not considered the best option to define the problem, this choice is related to the human background of the problem. The doctor–patient relation is highly considered in quality measures. Always being sent to the same doctor is highly valued by the patient. Therefore, most models consider independent queues for each doctor.
(ii) *Multi-server:* more than one server with the same capabilities is considered in the appointment system [17–20]. The scheduling algorithm decides to which server each patient is queuing. Schemes illustrating multi-server systems are provided in Fig. 1(b) and Fig. 2(b). A multi-server system was considered by Hoot et al. [21]. In this work, the emer-
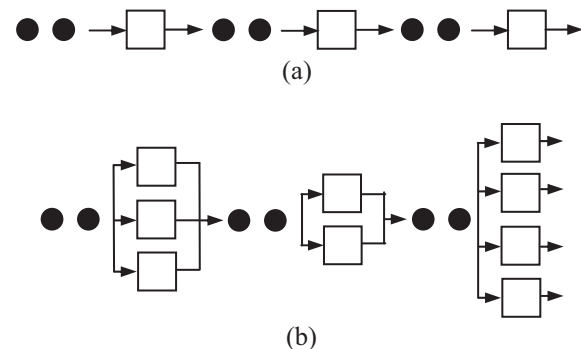

Fig. 2. Queuing system: (a) multi-stage/single-server system; (b) multi-stage/multi-server system.

gency department crowding was studied. Discrete event simulation was applied to real-time forecasting of emergency department crowding, being number of beds used to measure the occupancy level of the emergency department.

2.1.3. Patient tardiness

(i) *Delays:* difference between the scheduled time for the appointment and the time of the patient's arrival, is denominated by patient tardiness. Alexopoulos et al. [22] modeled patient tardiness by an asymmetric Johnson distribution. More recent studies in literature handle this uncertainty by considering it while modeling the patient arrival, instead of estimating patient's tardiness. Williams et al. [23] defined the patient arrival as a dynamic random Poisson process; the patient arrival process is described by an exponential distribution with time varying parameters.

(ii) *No-shows:* no-shows and late-cancelations, hereafter referred to as no-shows, are a major concern for healthcare managers. Liu et al. [17], defined no-shows as a binomial random variable with two parameters: (a) the number of appointments for the day and (b) the probability that the patient is a no-show. The no-show probability is estimated based on the time elapsed between the day the appointment was requested and the day the appointment was scheduled to. The influence of appointment date on no-shows was studied by Alaeddini et al. [18]. However, patient related variables were also considered. A hybrid probabilistic model was presented using patient social and demographic data, individual attendance records and appointment date. Lin et al. [19] studied the use of overbooking to decrease the impact of no-shows in admission scheduling. In this work, the authors presented a model for sequential assignment of heterogeneous no-show patients to time slots with the possibility of slot overbooking and in-clinic patient waiting.

(iii) *Walk-ins:* patients that come into the system without an appointment. Very few AS models consider walk-ins. Walk-ins are generally defined as probability related to a percentage of the scheduled appointments [20,24,25].

2.1.4. Processing times

(i) *Processing time:* was defined by Bailey [3] in the particular case of outpatient scheduling, as the sum of all the times the patient is with the doctor preventing him from attending to other patients. Extrapolating to other clinical environments, the processing time represents the sum of all the times the patient is using clinical resources in activities that add value to the process, or activities which the process completion depends on. Meaning that, in systems that do not refer to outpatient scheduling, i.e. patients that are hospitalized, the patient interacts with and uses other clinical resources, such as nurses, technicians, beds and rooms.

Mazier et al. [9] used real system data on admission and planned discharge. These authors used a random variable to correct the data and define the processing times distribution. The random variable was estimated using empirical data for the day of discharge. A heterogeneous stochastic model was used by Adan et al. [24]. The processing time was defined as a probability estimated according to patient group. Groups were defined by patient's age and empirical data regarding the use of resources.

The problem studied herein considered elective patients and was modeled as static, multi-stage/multi-server system and processing times were estimated according to the type of clinical examination and defined as a probability distribution.

2.2. Decision level in production control

A framework for production control in healthcare, describing the different decision levels in production planning, was introduced by Groot et al. [25] and further developed by Vissers et al. [26]. In the framework by Vissers et al. five hierarchical levels were proposed for resource capacity planning that are, from the highest to the lowest level: strategic planning, patient volumes planning and control, resources planning and control, patient group planning and control and patient planning and control. The hierarchical levels are defined as follows:

(i) *Strategic planning:* is the highest level of the framework. At this level, decisions encompass two to five years. They refer to top management and define the future direction of the provider.

(ii) *Patient volumes planning and control:* the target number of patients and service levels are set for each patient group and the production volume is settled. These decisions have a range from one to two years.

(iii) *Resources planning and control:* is devoted to the allocation of resources to the patient group, such as specialties or departments. It is at this level that the batching rules are set by patient group. At this level the decisions have an effect on the system for a period of three months to one year.

(iv) *Patient group planning and control:* refers to the availability of the resources required to perform the activities defined in the Patient Planning and Control level. At this level the decisions have a time range from weeks to three months.

(v) *Patient planning and control:* is devoted to the scheduling of daily activities required for patient planning. The decisions have a timeframe from one day to one week and refer to the point in time when waiting patients are admitted. That is, the definition of admission and discharge dates.

Although this framework treated healthcare providers as independent business units it is limited to resource capacity planning and does not consider online decision making.

Hans et al. [27] presented a two-dimensional framework for healthcare planning and control which considers four management areas and four hierarchical levels. The managerial areas are: medical planning, resource capacity planning, materials planning and financial planning. As for the hierarchical levels considered, these are: strategic, tactical, offline operational and online operational.

Considering the scope of the work presented herein, the literature review on the state of the art, given in this section, will only consider the *resource capacity planning* managerial area. This area is defined in the healthcare planning and control framework presented by Hans et al. [27]. Moreover, the identified research works are categorized according to the hierarchical levels described in the same framework.

In what concerns the four hierarchical levels proposed by Hans et al. [27], the *strategic level* has a similar scope as in the framework defined by Vissers [26]. This level addresses structural decisions with a long-term effect on the organization. Diwas and Terwiesch [15] studied the effect that the focus at firm level, operating unit level and process flow level, has on the operational performance in an AS problem describing cardiac care delivery in hospitals. Firm level denotes the percentage of hospital patients that are admitted for cardiac care. Operating unit level refers to the percentage of patients that require the same type of care delivery from the cardiology department. Process flow level, as indicated by the name, refers to the patient pathway. At this level, the process was modeled as a single-stage/single-server system. The effect of focus on performance was measured by the outcome of patient treatment. The authors were able to demonstrate that an increase of 10% at

firm level was associated with a reduction of 3.88% in the patient length of stay; the same increase at operating unit level implies a reduction in length of stay of 4.92%. Lee and Kwak [16] developed an enterprise resource planning (ERP) for healthcare using a multi-criteria decision-making model. The model was developed in the scope of the implementation of a new hospital providing patient-oriented healthcare services. At process level the system was modeled as a single-stage/single-server and the global work objectives were the maximization of service quality and minimization of cost for the organization, through an efficient resource allocation. Using goal programming the authors achieved satisfying solutions for the problem that allowed the increase of patient throughput. Mullen [28] demonstrated the importance of waiting lists prioritization in healthcare organizations, according to political and clinical objectives. This author investigated the importance of prioritization rules for patients' clinical aspects, such as pain and severity, and social factors, such as age, employment status and number of dependents.

Tactical level addresses the processes organization and execution. At this level the decisions have an intermediate planning horizon. An AS problem in a surgery department aiming to minimize the patients' length of stay in waiting list was presented by Arenas et al. [29]. The problem was defined as a single-stage/multi-server system. The number of requests and processing time were estimated using empirical data and described by the average. Using goal programming the authors were able to present a solution to minimize the patients' length of stay to a maximum of six months but no solution was found to minimize it to a maximum of four months. Gallivan et al. [30] studied the effect of the variability of patient length of stay on the capacity requirements in cardiac surgery. The number of service requests was estimated based on empirical data and described by the total average, not considering daily or weekly patterns. The effect of the variability of length of stay was determined by investigating the bed occupancy in the intensive care unit. The results of this study revealed that not considering this variability could lead to an operational overload of 41% of the operational time, indicating that the variability in length of stay is determinant in hospital activities. AS for elective patients in a hospital was presented by Nunes et al. [31]. The AS problem was modeled as a single-stage/multi-server and aimed at the prevention of resources idleness and excessive usage. Adan and Vissers [32] studied the AS problem at an orthopedics department in a general hospital. The goal was to find a feasible admission profile while setting targets for patient throughput and resource utilization. Good results were obtained when maximum weight was given to operating room use. However, because only elective patients were considered, the optimization level achieved was small when compared to the real scenario. Holm et al. [33] used a discrete event simulation model to tackle the AS problem through the optimization of bed allocation among wards, aiming the minimization of hospital overcrowding. The model does not consider the patients' flow among the wards and each ward has an unlimited number of beds. Thus, the model is classified as a single-stage/multi-server being both elective and walk-in patients considered. Patients' length of stay in the ward is defined by probability distribution and the arrival of walk-ins is defined by a Poisson distribution. In the proposed model, each bed allocation scenario is simulated generating a matrix *M*, denoting the utility of increasing the number of beds, which is then fed to an optimization algorithm that yields the optimum allocation scenario. Bair et al. [34] approached the AS problem in an emergency department (ED) through the optimization of ED bed allocation as well as patients transfer to inpatient beds. The problem was modeled as a multi-stage/multi-server system and both patient inter-arrival time and processing time were defined by a probability distribution. The authors demonstrated that by increasing the number of inpatient beds available it is possible to decrease both ED overcrowding and the number of patients that leave the ED without being seen. Patrick et al. [8] formulated the AS as a capacity allocation problem of appointment slots for Computed Tomography (CT) demand in a hospital. Three priority classes were considered for the patient that constrained the length of stay in waiting list to 7, 14 and 28 days, respectively. Given the complexity of the multi-priority patient scheduling problem, linear approximations for the value functions were used to develop an approximate dynamic programming method. The authors demonstrated that, contrary to current practice, small amounts of flexibility have a significant positive impact in the scheduling of the high-priority patients.

Operational level also addresses process execution. However, at this planning level decisions have a short-term horizon. Given the stochastic nature of healthcare processes and considering the low level decisions involved in operational planning, this level contains two sub categories: *offline* and *online*.

Offline operational level, the level approached in the work presented herein, is devoted to operational decision making that involves advanced planning of the processes, whereas online operational level refers to decisions that take into consideration the current status of the processes. Demmester et al. [6] described the AS problem as an offline single-stage/multi-server system. By means of a hybrid tabu search algorithm, the authors studied the daily assignment of patients to beds in a hospital with the goal of minimizing patient waiting. The proposed algorithm revealed to be very effective in finding a feasible solution when compared to commercial integer programming solvers. However, the single-stage modeling of the problem used limits the proposed system applicability. Single-stage models do not allow the algorithm to have full knowledge of the patient pathway. Therefore, it cannot be applied to problems, as the case-study considered herein that considers more than one type of resources, since it is unable to handle the interaction in between them. Su and Shih [35] used a case-study of an outpatient urology clinic to investigate and demonstrate the benefits of applying simulation models to offline AS problems. A multi-stage/single server system was considered, with both elective and walk-in patients. Patients' processing times were determined from empirical data and fitted to a probability distribution, which goodness-of-fit was proved by a Kolmogorov-Simirnov test [36]. The authors sequentially numbered admission slots and setup four scenarios considering different admission policies to infer from its implications on patients' throughput and waiting time. In the outpatient clinic setting considered, patients' appointments are scheduled to a specific physician. Thus, the proposed method is not suitable for AS problems where patients queue for multiple server, i.e. more than one resource with the same capabilities is considered for allocation by the appointment system. Pham and Klinkert [37] proposed a mixed integer programming (MIP) model for the offline AS problem in a surgery department, considering both elective and emergency patients. Surgery processes were modeled as multi-stage/multi-server system with time constraints on the maximum waiting between stages. The authors introduced the multi-mode concept to refer to the multiple groups of resources, selected from the resources pool, available to execute each activity. Each mode was defined by a set of resources and characterized by the time interval of availability. Considering that the authors used MIP, and have identified as future work the development of a heuristic solution, one may conclude that they were not able to overcome the problem of excessive calculation time inherent to this type of algorithms. Cardoen et al. [38] determined the offline optimal sequence for the daily schedule of patients to operating rooms in an ambulatory surgical center. Using a MIP method the authors simultaneously optimized six objectives by incorporating them into a single multi-objective function with

weighting factors. The method revealed several limitations concerning the non-interchangeable time slots, resource capacity, and multi-tasking. Moreover, deterministic processing times were used for surgeries which introduces a huge error in the description of the actual events. Lamiri et al. [2] developed a stochastic algorithm combining Monte Carlo Simulation and MIP for scheduling patients to operating rooms. Their objective was to minimize operating room overtime costs, considering both elective patients, who may be delayed in a time horizon, and emergency patients, which have to be admitted on the day of arrival. The authors demonstrated that stochastic optimization models achieve considerably good results, when compared to deterministic models. However, solutions were achieved by solving random instances of the problem. Thus, it is difficult for the authors to know whether the algorithm is converging to an optimum value.

As an example of the *online operational level* the reader is referred to Li et al. [39]. The authors proposed a method for patient scheduling aimed at minimizing patient waiting time and resolving waiting room overcrowding. The model considered patient priorities in real-time and its implementation allowed a significant reduction of patient waiting times.

In this paper, the AS problem is addressed at an offline operational level. The proposed method aims to minimize the patients' waiting time and maximize the patients' throughput by minimizing the processes' total completion time.

3. Optimum workflow: case-study

To unequivocally understand the optimum operation of clinical workflows defined by national and international guidelines [40,41], the processes from all the modalities existing in the studied clinic were theoretically modeled. The case of Magnetic Resonance Imaging (MRI) is briefly presented hereafter.

In MRI, and diagnostic imaging in general, implemented workflows are notoriously inefficient. These workflows require professionals to perform a heightened number of tasks, since the patient is scheduled until the image report is ready for delivery. To evaluate and understand the workflow in an MRI examination it is necessary to understand the main types of examinations and their utility, as well as the technical protocols involved in each of them. Technical protocols mainly influence the time required for image acquisition during an examination. It should also be understood for each type of examination: (a) how the members of clinical staff are involved, e.g. exams that require contrast administration imply additional human resources, and (b) which are the technical requirements, e.g. not all equipment is able to perform a given set of protocols.

To understand and outline the capabilities and responsibilities of each human resource, national and international guidelines on the execution and interpretation of MRI examinations were analyzed. The identified capabilities by human resource category are as follows:

Radiologist: the radiologist's capabilities may be divided in two main sub-categories: executive and cognitive [42]. Executive capabilities may be described, defined, taught and subjected to execution patterns, which makes them transferable to other members of staff with appropriate training, e.g. radiology technician. During examination, the radiologist performs a supervision role, unless the examination requires advanced anatomical knowledge. Cognitive capabilities are those in which the radiologist cannot be replaced and that add value to the process such as examination reporting. Therefore, the radiologist should be available to perform them as much as possible.

Radiologist technician: it is competence of the technician the preparation and positioning of the patient for examination, as well as their surveillance during the examination. Technicians also plan, program and execute the procedures necessary for the examination. They may also write a preliminary report describing the observations, to aid the radiologist on reporting. Daily quality control assessment, such as equipment calibration, is also the responsibility of the technician [42,43].

Nurse: the nurse is responsible, in cooperation with the technician, for preparing and positioning the patient. The nurses also have to prepare the intravenous line for contrast administration, detect eventual contraindications and monitor the patient during injection [43].

In the general case of an imaging clinic, the following physical resources were identified: front office, waiting room, control room, exam room, dressing room, recovery room, processing room, and radiologist room. For a detailed explanation on the specifications of physical resources, please consult the following references [44,45].

Parallel to the patient workflow there is an information and/or resource flow, hereafter denoted by information flow, independent from the patient workflow. The information workflow includes tasks that despite the patient not being aware of, are of extreme relevance to accomplish the examination request. This includes tasks such as: (a) the patient registry on the IT system; (b) the analysis of the patient and the examination information in order to define the examination sequences; (c) the registry of examination information on the modality; (d) the equipment preparation for examination; (e) the images processing; and (f) the images reporting. Together, the patient and the information workflow compose the clinic workflow, since the patient requests the examination until the final report is delivered.

Considering the specification on the resources, given herein, the guidelines and the best practices [46,48–56], all the tasks required to accomplish an examination were identified, characterized, and the optimum workflow defined, as illustrated in Fig. 3.

3.1. Case study description

This article approaches an AS problem in a private imaging clinic. Since the clinic is not integrated in a hospital, no service requests, besides the scheduled patients, are delivered by the clinic. In the field work, carried out in the beginning of the project, a non-significant number of patient delays was observed. This is mainly due to the social environment in which the clinic is integrated. As the clinic is situated in a rural area in the north of Portugal, and serves a vast area, due to social factors the patients tend to arrive early for the appointment, often several hours before the scheduled time. Thus, it was not possible to characterize this environmental factor.

The considered clinic provides services on Computed Tomography (CT), MRI, Radiology (RX), Orthopantomography (OT), Densitometry (PX), Mammography (MG) and Ultrasound (US). It has 8 daily working hours and is composed of human, physical and technical resources, such as: Physicians; Technicians; Assistants; Administrative; Waiting room; Changing Room; Control Rooms; Examination rooms; Modalities.

CT and MRI modalities use the same human and physical resources. The call for examinations is made by the same assistant, who directs the patients to the dressing rooms. There are four dressing rooms shared by both modalities. RX/OT and MG/PX share the technician that executes the examination.

In the considered case, appointments are requested by the patient either by phone or in person. At the moment of request the patient appointment is scheduled for a day and time according to the timetable availability and patient's preference. The size of the time block, i.e. processing time, is defined according to the modality. At the day of the appointment the patient refers to the
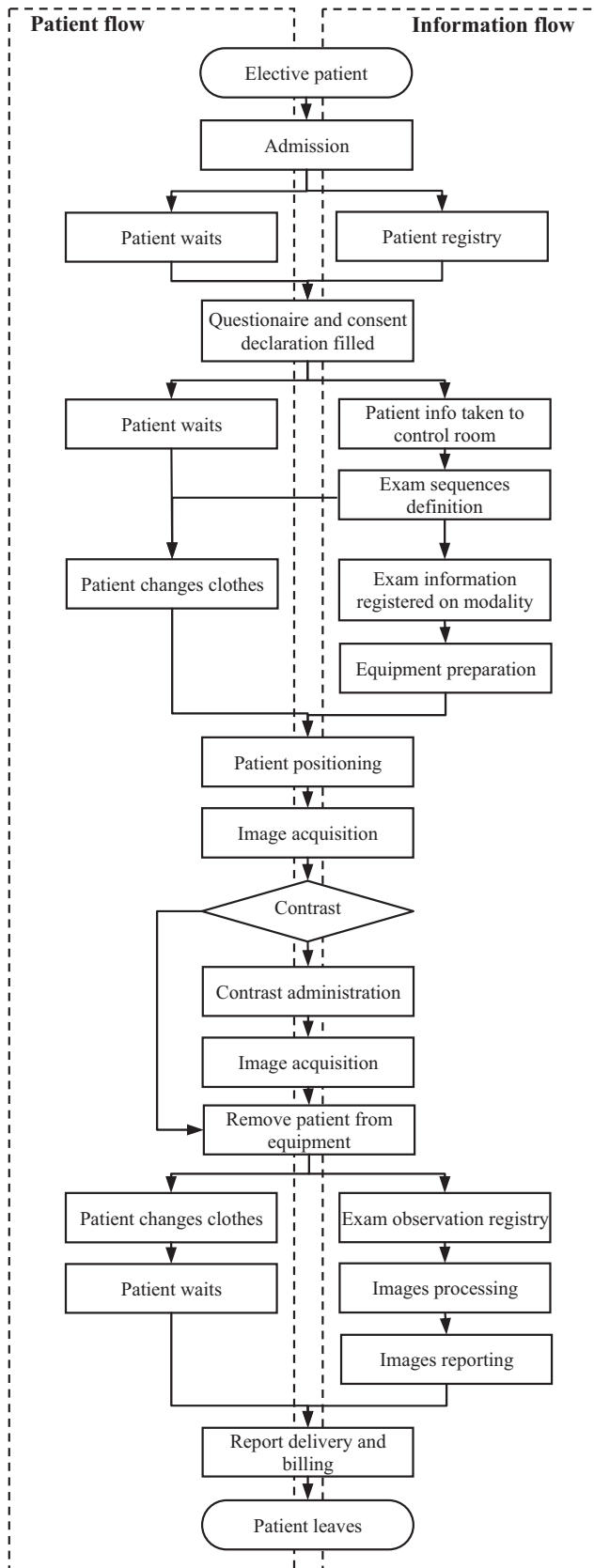
Fig. 3. Proposed scheme of the optimum clinic workflow.

reschedule made by the technicians, based on their empirical knowledge on the type of examination. Depending on the number of patients waiting and the modality required to fulfill the examination, more than one patient may be called at the same time. In these cases, if the patients are scheduled either to CT, MRI or RX then they are directed to dressing rooms. Otherwise, only one patient enters directly to the examination room and the remainders are directed to the interior waiting area. After discharge, patients are directed to billing.

The proposed AS problem was defined as a multi-stage/multi-server system. Patient tardiness, i.e. delays, no-shows and walk-ins, was not considered, given that in the case presented herein patient earliness was assumed to be more frequent. The system was modeled based on a new concept of task. Here, task is defined as all activities using the same set of resources during the same time range. Tasks are characterized by their space and time dependency of the resources and their sequence in time. Thus, for a task to be fully defined the previous task has also to be defined in time, either by its ending time or operation mode, i.e. if waiting is or not allowed.

The processes implemented at the studied clinic were modeled by *in loco* observation and characterized according to the parameters described in the following sections. The parameters were defined based on the optimum workflow, to ensure the process models accuracy. These parameters allow describing of the real case, while maintaining the modeling flexibility and robustness.

4. Resources and tasks characterization

No distinction was made between human, physical or technical resources while defining the characterization parameters of the modeling approach. Furthermore, given that both the patient workflow and the information flow are considered in the characterized processes, for modeling purposes patients are considered as resources.

Resources are defined by the activities they are able to perform instead of their type, (e.g. physician, nurse, or technician). This ensures that during simulation all the resources executing the activity in the real case are considered for allocation. This allows the simulation tool to describe situations such as the technician referring to the waiting room to call the patient for examination when the assistant is not available. These situations occur in the real case even though it is not of the resource responsibility.

During the fieldwork at the clinic it was noticed that resources often operate in multi-tasking mode. This is particularly relevant in cases where, during the examination of a patient, the technician processes the examination images from another patient, since this last task is extremely time consuming when compared to the duration of an examination. In this sense, resources are characterized according to their capacity when performing an activity and their availability is defined in time slots that are used by the activities.

Furthermore, it was observed that physicians have preferences when choosing the technician that executes certain examinations. To be able to consider these preferences resources are classified as belonging to a team or not, when performing specific activities. This feature is also used to tackle the common situation of dressing rooms being chosen for patient use according to the examination modality.

Each resource *m* belonging to the set of resources M_m is characterized by the parameters described in Table 1.

The *in loco* observation of the clinical workflows of the studied clinic led to the conclusion that the process could not be modeled as a single stage. Considering a single-stage system means that all the resources needed to perform the process are used in

clinic for admission. Subsequent to admission, the patient waits to be called for examination. The service is provided to the patient when the necessary resources are available and according to the

Table 1
Problem characterization parameters.

Resource parameters	
A_m	Activity performed by resource m. Each resource is able to perform one or more activities (index a)
$C_{m,a}$	Number of time slots of resource m used to perform activity a
$G_{m,a}$	Defines if the resource m operates in group the activity a. Assumes the value o if activity a is not performed in group by resource m
Task parameters	
$\mu_{i,j}$	Average duration of task i from process j
$\sigma_{i,j}$	Standard deviation of the duration of task i from process j
$\varepsilon_{i,j}$	Operation mode of task i from process j. Assumes the value w or o whether waiting is allowed or not, respectively
$w_{i,j}$	Maximum waiting time allowed for task i from process j
$p_{i,j}$	Processing time of task i from process j
Task resources parameters	
$\eta_{m,i,j}$	Occupation mode of resource m in task i from process j; takes the value of zero waiting (zw) or non-zero waiting (nzw) as the occupation mode is, respectively, continuous and the task $T_{i,j}$ must start immediately after task $T_{i-1,j}$ finishes for the resource m, or discontinuous and task $T_{i,j}$ does not have starting time constraints related to resource m
$R_{m,i,j}$	Constraint of resource m in task i from process j; constraints the choices of resource m to perform task i from process j to the one chosen to perform a previous task in process j
Task precedence relations	
$TB_{i,j}$	Set of tasks that have to be finished before task i from process j starts
$TA_{i,j}$	Set of tasks to be initiated as task i from process j finishes

continuous mode throughout the process, which does not describe the reality.

The present modeling approach considers each task as a stand-alone process, i.e. performed by the same set of resources in the same time space. The use of this concept makes necessary to establish precedence relations between tasks in order to sequence them in time and to prevent tasks from starting before previous tasks are completed. Additionally, for a task to be fully defined in time it is also necessary to either characterize its processing time or its operation mode (i.e. defining whether waiting time is allowed or not). However, in diagnostic imaging waiting time limits have to be imposed to tasks in processes that involve the administration of contrast material to the patient. For the same reason it is necessary to impose that examination tasks are non-preemptive, i.e. when the examination starts the process cannot be interrupted. The non-preemption constraint is imposed to the patient resource and not to the tasks. This ensures that the second image acquisition task is performed on time, independent of the technician being available or not; since the process of bio-distribution of the contrast material takes 20 min and the technician might not be available.

In the healthcare environment the doctor/patient relationship is very important. Patients feel reassured and consider a parameter of quality is met when the same resources assist them during their stay at the clinic. However, this does not apply to all the resources that intervene in the process; as an example, it is not relevant that the assistant is the same throughout the process, as it is not important to the patient or to the clinic quality. To describe this preference, the decision of the resource needed to perform an activity in a given task; is constrained to the previously used resource (e.g. the same nurse, the same dressing room).

Each task i in the set of tasks T_i ($i = 1, \ldots, n_j$) from process j is characterized by the parameters shown in Table 1. The task resources m and task precedence relations are characterized by the parameters presented in Table 1.

According to the aforementioned characterization parameters for both resources and tasks, each process at the clinic was modeled.

5. Methodology

The modeling methodology was implemented in three stages: workflow definition, data collection and simulation and optimization via simulated annealing. First, it was accomplished through a holistic overview of the clinic workflows to identify possible tasks that were not taken into account in the previous definition of an optimum workflow. The data collected at this stage established the principles for the elaboration of a detailed checklist, generic to all providers, that systematizes the relevant data collection. In the following, the used modeling methodology is exemplified for the MRI modality.

5.1. Data collection

To better understand the resources utilization at the studied clinic, 966 medical exams were investigated and characterized. For a period of four months, data was collected by the observation of the daily work. The workload characterization data is given in Table 2 and Fig. 4.

Data in Fig. 4 evidences a characteristic pattern concerning the frequency of a given type of exam. From a total number of forty-one different types of medical exams, six represent 76.4% of the total workload (cervical, lumbar, knee, brain, breast and shoulder). This led to the restriction of the exams observed, during the second phase to the six stated above, given the amount of observation time required to have a sample that would be statistically representative.

In the second stage, the major sources of disturbances to the normal workflow were investigated. In this sense, it was studied the influence on the clinical workflow of events/features, such as: (a) resource sharing between modalities – several times this leads to idle processes due to the assistant's participation in the ultrasound processes, which require his/her attention for long periods of time; and (b) rescheduling of patients by the technicians, based on their empirical knowledge, it is not always effective and in the majority of times leads to longer patient waiting and extensive working hours. The standard checklist, developed during the first stage, made it possible to recreate the idiosyncratic reality of the studied clinic.

At a third and last stage, the processing time was determined for every previously identified task. The collected data was subjected to an outlier analysis. These points were not included in the calculation of the parameters required for processing time definition, considering that their numerical value was very distinct from the rest of the sample. Since these points were part of the collected data, the cause of variability was investigated.

In the task of equipment preparation one outlier was removed from the sample. The cause for its value being extremely high was due to a patient undergoing two examinations without leaving the examination room. This resulted in the technician preparing both coils needed for the examinations, doubling the time required for the task. In this case the influencing variable was the number of examinations performed on the same patient.

Table 2
Information regarding the MRI workload on the studied clinic during the observation period.

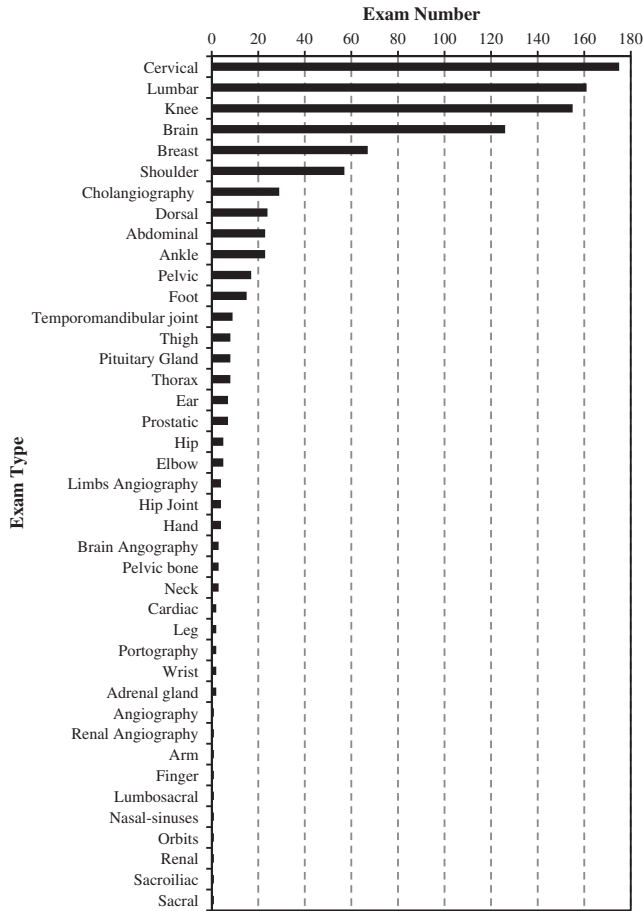|	
Working hours per day, h	8
Number of working days	82
Total number of MRI exams	966
Average number of MRI exams per day	12
Minimum number of MRI exams per day	6
Maximum number of MRI exams per day	18
MRI exams requiring contrast, %	75.9

Fig. 4. MRI exams frequency by examination type.

The outliers in the patient positioning were all associated to the breast examination. During this examination the contrast injection is automatic and, in these cases, an intravenous line has to be prepared for the injector, which does not happen in any other type of examination. In this case the influencing variable was the examination type.

The statistical analysis on the clinic workflow main tasks is given in Table 3. Considering that the type and number of sequences differ by examination type, the data relating to the image acquisition task was segmented by examination type, see Table 4.

5.2. Simulation

The proposed simulation model is based on the dynamic allocation of resources, taking into consideration the stochastic nature of

Table 3
Statistical analysis of the main tasks of the clinic workflow.

Parameter	Task		
	1st cloth change (min)	Prepare equipment (min)	Patient positioning (min)
Average	2.30	0.62	1.96
Median	1.93	0.53	1.79
Standard deviation	0.88	0.29	0.65
	Contrast (min)	Remove patient (min)	2nd cloth change (min)
Average	2.07	1.34	1.95
Median	1.92	1.17	1.83
Standard deviation	0.72	0.65	0.88

Table 4
Statistical analysis of the image acquisition task by examination type.

Parameter	Exam requiring contrast administration			
	Cervical (min)	Lumbar (min)	Brain (min)	Breast (min)
Average	20.40	26.49	26.70	25.90
Median	19.50	25.50	25.17	26.63
Standard Deviation	4.46	6.04	8.23	4.88
Parameter	Exams without contrast administration			
	Cervical (min)	Lumbar (min)	Knee (min)	Shoulder (min)
Average	18.00	15.71	15.78	16.28
Median	18.29	15.62	15.58	15.83
Standard Deviation	3.79	1.25	1.33	2.79

tasks processing times. The scheduling system, in which the simulation is based, considers frequently occurring situations in healthcare that are seldom studied, such as activities performed simultaneously and remotely, technical patient waiting, multi-activities of resources and their interdependency. In order to guarantee the scheduling results reproducibility, for a given set of patients, the events mentioned before are described by modeling parameters. By not letting the simulation technique decide on these situations the model guarantees that they are solved following the real events, and that the results obtained for the same set of patients in the same sequence are the same. This allows the scheduling system to have a high degree of accuracy. The scheduling system only has to handle the time related constraints. In this sense, the decision variables are:

$t_{0_{i,j}}$ Task i from process j starting time.
$t_{f_{i,j}}$ Task i from process j ending time, defined as:

$$t_{f_{i,j}} = t_{0_{i,j}} + p_{i,j} + \Delta t_{w_{i,j}}, \tag{1}$$

where

$\Delta t_{w_{i,j}}$ Waiting time of task i from process j.
The constraints reads:

$$\varepsilon_{i,j} = \begin{cases} w \to t_{0_{i,j}} \geqslant t_{f_{i-1\,j}} \\ o \to t_{0_{i,j}} = t_{f_{i-1\,j}} \end{cases} \tag{2}$$

$$\eta_{i,j} = \begin{cases} zw \to \Delta t_{w_{i,j}} = 0 \\ nzw \to \Delta t_{w_{i,j}} = t_{0_{i+1\,j}} - (t_{0_{i,j}} + p_{i,j}) \end{cases} \tag{3}$$

For each task, $p_{i,j}$ is defined by a normal distribution, based on $\mu_{i,j}$ and $\sigma^2_{i,j}$ specified in task parameters. The definition of $p_{i,j}$ is bounded as follows:
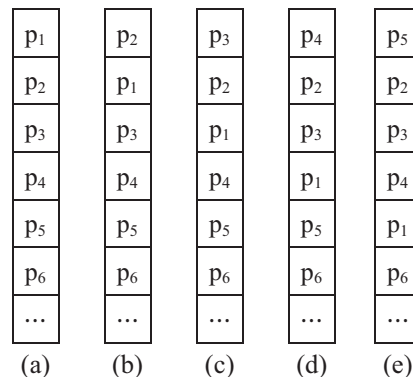

Fig. 5. Illustration of the swap relation in the search space S_1. (a) Incumbent solution; (b) $N_1 \langle p_1, p_2 \rangle$; (c) $N_2 \langle p_1, p_3 \rangle$; (d) $N_3 \langle p_1, p_4 \rangle$; (e) $N_4 \langle p_1, p_5 \rangle$.

$$p_{i,j} > 0, \quad \forall\, i \in T_j, j \in J \tag{4}$$

$$\mu_{i,j} - 3\sigma_{i,j}^2 \leqslant p_{i,j} \leqslant \mu_{i,j} + 3\sigma_{i,j}^2, \quad \forall\, i \in T_j, j \in J \tag{5}$$

Considering the magnitude of processing time for some tasks, constraints (4) ensure that only valid values are used for simulation. Moreover, constraints (5) impose a three sigma precision to $p_{i,j}$, centered in $\mu_{i,j}$. This implies that the defined $p_{i,j}$ matches 99.7% of the observed value.

6. Optimization via simulated annealing

Simulated annealing (SA) metaheuristic was originally proposed by Kirkpatrick et al. [46] and Černý [47]. The algorithm is

Fig. 6. Flowchart of the simulated annealing algorithm where, N represents the iteration number and N_{max} the maximum number of iterations allowed per cycle.

based on the thermodynamic principles of the physical annealing process. In the annealing process, a solid is heated to a temperature above its melting point and then cooled slowly. The structural proprieties of the solid depend on the cooling rate. If the cooling process is fast the solid crystals will be small and with imperfections. The analogy between the SA and the physical process is as follows:

Objective function, $F_{obj} \rightarrow$ System energy, E_i.
Initial solution, $S_0 \rightarrow$ Initial state.
Accepted solution, $S_m \rightarrow$ System state.
Neighbor solution, $S \rightarrow$ Change of state.
Control parameter, $T \rightarrow$ Temperature.
Best solution, $S_{beat} \rightarrow$ Solidification state.

One of the benefits of a SA metaheuristic is that it can avoid local minima by accepting a worse solution, under given conditions. In the SA version used herein, shown in Fig. 6 as a flowchart, the Boltzmann distribution (see Eq. (6)) was considered to define the probability of acceptance of worse solutions and the cooling rate was defined as being geometrical, Eq. (7).

$$p = \frac{\exp(-E_{i+1}/kT)}{\exp(-E_{i+1}/kT) + \exp(-E_i/kT)} \approx \exp(-\Delta E/kT) \tag{6}$$

$$T_{i+1} = \beta \cdot T_i \tag{7}$$

The algorithm of Fig. 6 has five parameters: T_0, the system initial temperature; T_{min}, stopping criteria; N, the number of neighbors evaluated at each temperature; β, the cooling rate; and k, the Boltzmann constant. An additional stopping criterion is used to evaluate if the system is *frozen*, based on the evolution of the objective function in a given interval of iterations.

In SA original configuration [46,47], the search space was defined by the random generation of neighbors. In the work presented herein, SA searches on the space of patient sequences, representing the order by which elective patients are admitted in a working day, with relations in between neighborhoods, denoted as S. The first space investigated, denoted as S_0, has only one solution that represents the original patient sequence at the studied provider. In this way, the simulation results for S_0 were used to
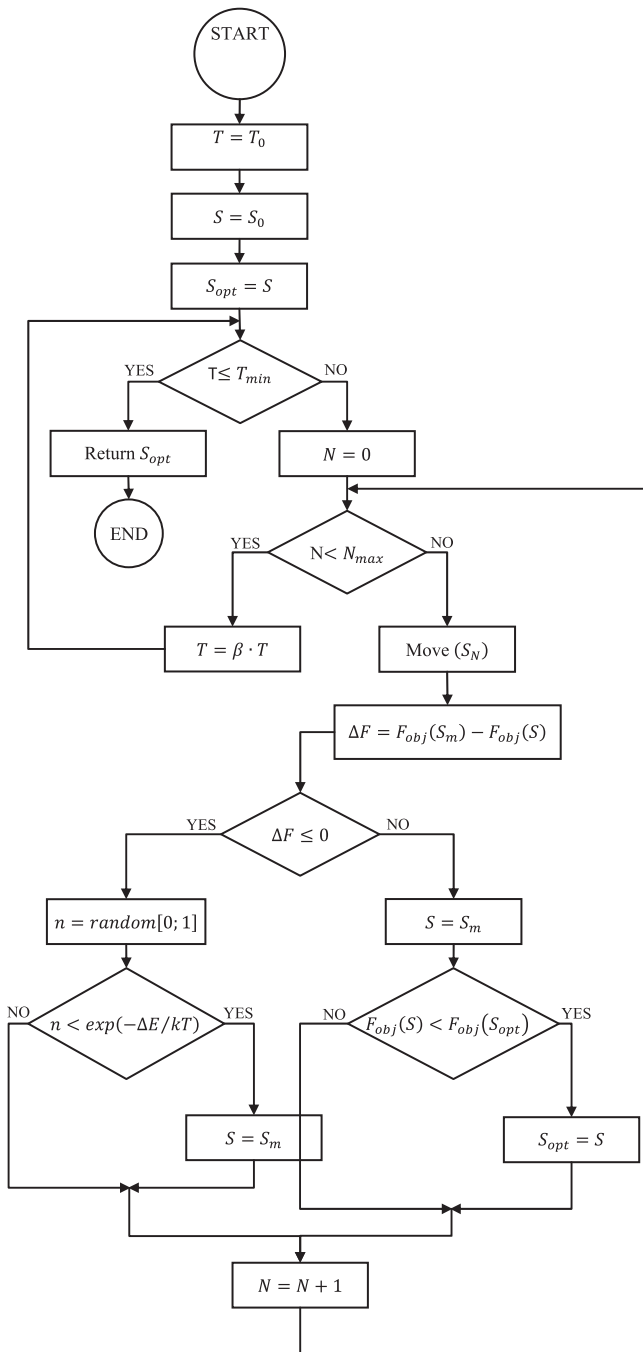
Table 5
Statistical analysis on the main tasks of the clinic workflow.

Modality	Exam type	Load
MRI	Cervical	1
	Cervical with contrast	2
	Lumbar	1
	Lumbar with contrast	2
	Brain with contrast	2
	Breast with contrast	1
	Knee	2
	Shoulder	1
CT	Brain	9
	Brain with contrast	1
	Lumbar	6
	Thorax	3
	Abdominal	3
	Cervical	3
	Peri-nasal sinus	2
MG	Mammography	20
PX	Densitometry	10
OT	Orthopantomography	4
CR	Thorax	6
	Lumbar	5
	Cervical	4
	Knee	4
	Foot	3
	Full body	2

Table 6
Simulation results for the space S_0 and optimum results of the simulated annealing algorithm.

	Run 1			Run 2			Run 3		
	S_0 (min)	Optimum (min)	$\left|\frac{Optimum}{S_0} - 1\right|$ (%)	S_0 min	Optimum (min)	$\left|\frac{Optimum}{S_0} - 1\right|$ (%)	S_0 (min)	Optimum (min)	$\left|\frac{Optimum}{S_0} - 1\right|$ (%)
Total completion	28,459	27,030	5.0	27,439	26,157	4.7	26,779	25,245	5.7
Total waiting	269,727	166,962	38.1	146,191	91,112	37.7	143,852	89,664	37.7
F_{obj}	149,093	96,996	34.9	86,815	58,634	32.5	85,316	57,455	32.7

check the simulator accuracy on reproducing the original scheduling. Each new solution generated in each neighborhood S was simulated and the total completion (TC), which denotes the final processing time of the last task to end, and the sum of patients' waiting time, or total waiting (TW), were used in the evaluation of the objective function, according to Eq. (8).

$$F_{obj} = 0.5TC + 0.5TW \qquad (8)$$

The random generation of patient sequences, in the search for new solutions, means that there is no control in the search method. The new solution may or not be close to the previous and the search procedure may overcome the optimum solution, or take a longer time to find it.

Considering that the clinic processes share one or more resources, slight changes in the patient sequence have a large impact in the objective function. In this sense, a patient swap move was considered to generate neighbor solutions. In the sequence generation, a pair of patients $\langle p_1, p_2 \rangle$ is swapped for the full set of exams they are scheduled to. The swap is restricted and does not occur if the set of exams to be performed by both patients is the same. In each search space one of the patients was kept unchanged in the $\langle p_1, p_2 \rangle$ pair selected for swap to generate a new neighbor, as shown in Fig. 5.

7. Results and discussion

Experiments were conducted in order to determine the optimum patient admission sequence that minimizes total completion and total waiting.

Total completion was used as a measure of the system efficiency. It was considered that, a sequence of patient admissions was more efficient if the production ends earlier. Total waiting was used as a measure of production quality. TW, as it is a performance indicator that directly affects the patient, was used to measure production quality. It was considered that a sequence of patient admissions that, on average, imposes less waiting time to the patient denotes better production quality.

The developed work was applied in the operational analysis of the imaging clinic described in Section 3.1, with the exception of the US modality. Simulation inputs are based on empirical data and are representative of the workload in a typical day at the studied clinic. The set of exams investigated is presented in Table 5.

The results obtained for the space S_0 were compared to the reference of 8 h of work per day, denoting a relative error in simulation of 1.2%. The results for S_0 evidenced that the number of administrative staff is a limiting variable of the system. The bottleneck mostly influences the MG modality since the patient length of stay, in most cases, is shorter than the time given between admissions, turning the technician idle. The situation in which the idleness does not occur is when the patient is scheduled for MG and OT. Since the technician executing both modalities is the same, the patient performs the processes in continuum. The best result was obtained in the trial in which the optimum solution was found at S_{11}. In the sequence evaluated as the optimum, the MRI patients were admitted by examination type. In practice this means that less room preparation is needed, since the coils necessary for the

exams are the same, diminishing the processing time and, consequently, the patient wait (see Table 6).

The same variable, examination type, also influences the CR modality. In the optimum sequence, all the patients scheduled for the thorax examination and three of the patients scheduled for the lumbar examination were admitted consecutively. This also diminishes the processing time since the equipment existing at the studied clinic had a rotary table that had to be set according to the examination type. In the mentioned type of examinations, the table is set to the vertical position in all of them, which eliminates the machine setup time from these processes, except for the first to be admitted.

Resource sharing was also identified as a limiting variable. The CR and the MG modalities share the equipment in which the images are digitalized. Several times this equipment was required by both modalities at the same time, or was in use when required, causing idleness of technicians and increased patient waiting, since the patient cannot be discharged until the image quality was evaluated.

8. Conclusions

It was demonstrated that the application of the developed tools could be used for the evaluation of patient admissions. The gathering of simulation with optimization techniques was demonstrated to be very effective in the determination of system bottlenecks as well as their evaluation by identifying their cause. A drawback of simulation tools, based on linear programming, is that this approach attempts to solve the AS problem as a queuing system, by generating an empirical fit to the system behavior defined in the process model. Such limitation was minimized by the modelling methodology, the system module that defines the system behavior. Here, each resource is characterized according to its competences. Thus, the simulation tool allocates a resource to a task according to the activities they are able to perform, therefore, being flexible when generating a solution and more truthful to the actual events. Additionally, a sequence generation algorithm, that allows to explore different system settings when searching for a solution, was embedded in the developed system.

An imaging clinic was studied in terms of efficiency and production quality, denoted by total completion time and total patient waiting, respectively.

The patient sequence of the simulated optimum solutions was similar to the patient call order imposed by the technicians. If during patient admission the type of examination is considered, and as a result, patients with similar exams are admitted in continuum, this would have a positive impact in the patient waiting time. On the other hand, if the technician overrides the patient sequence on the day of the examination, to improve the internal sequence of examinations, these results on an extension of patient waiting time, assuming that he/she arrived on time. Therefore, it is concluded that the type of examination should be considered in patient admission.

The process limitations caused by the sharing of the resource used for image development can only be resolved by the acquisition of additional development equipment, or by updating the ana-

log equipment that is used for image acquisition, to digital equipment that does not require image development.

In future work it would be interesting to test other optimization tools that guarantee the finding of a global optimum. It would also be important to consider in the objective function the cost of resources (human, technical and physical).

References

[1] Direcção de Serviços de Epidemiologia e Estatísticas de Saúde – Divisão de Estatísticas da Saúde, "Centros de Saúde e Hospitais – Recursos e Produção do SNS 2007". Direcção Geral de Saúde, Lisboa; 2008.

[2] Lamiri M, Xie X, Dolgui A, Grimaud F. A stochastic model for operating room planning with elective and emergency demand for surgery. Eur J Oper Res 2008;185(3):1026–37.

[3] Bailey NT. A study of queues and appointment systems in hospital out-patient departments, with special reference to waiting-times. J Roy Stat Soc Ser B Met 1952;14(2):185–99.

[4] Cayirli T, Veral E. Outpatient scheduling in health care: a review of literature. Prod Oper Manage 2003;12(4):519–49.

[5] Garg L, McClean S, Meenan B, Millard P. A non-homogeneous discrete time Markov model for admission scheduling and resource planning in a cost or capacity constrained healthcare system. Healthc Manage Sci 2010;13(2):155–69.

[6] Demeester P, Souffriau W, De Causmaecker P, Vanden Berghe G. A hybrid tabu search algorithm for automatically assigning patients to beds. Artif Intell Med 2010;48(1):61–70.

[7] Ceschia S, Schaerf A. Multi-neighborhood local search for the patient admission problem. Hybrid Metaheuristics 2009;5818:156–70.

[8] Patrick J, Puterman ML, Queyranne M. Dynamic multipriority patient scheduling for a diagnostic resource. Oper Res 2008;56(6):1507–25.

[9] Mazier A, Xie X, Sarazin M. Scheduling inpatient admission under high demand of emergency patients. In: 2010 IEEE conference on automation science and engineering (CASE), Toronto; 2010.

[10] Helm J, AhmadBeygi S, Van Oyen M. Design and analysis of hospital admission control for operational effectiveness. Prod Oper Manage 2011;20(3):359–74.

[11] Cabrera E, Taboada M, Iglesias M, Epeldee F, Luque E. Simulation optimization for healthcare emergency departments. Proc Comput Sci 2012;9:1464–73.

[12] Granja C, Mendes J, Janela F, Soares J, Mendes A. Optimisation-based on simulation: a diagnostic imaging department case-study. In: Second international conference on information, process, and knowledge management (eKNOW), Saint-Marteen; 2010.

[13] Chen N, Zhan Z, Zhang J, Liu O, Liu H. A genetic algorithm for the optimization of admission scheduling strategy in hospitals. In: 2010 IEEE congress on evolutionary computation (CEC), Barcelona; 2010.

[14] Connelly L, Bair A. Discrete event simulation of emergency department activity: a platform for system-level operations research. Acad Emerg Med 2004;11(11):1177–85.

[15] Diwas SK, Terwiesch C. The effects of focus on performance: evidence from California Hospitals. Manage Sci 2011;57(11):1897–912.

[16] Lee C, Kwak N. Strategic enterprise resource planning in a health-care system using a multicriteria decision-making model. J Med Syst 2011;35(2):265–75.

[17] Liu N, Ziya S, Kulkarni V. Dynamic scheduling of outpatient appointments under patient no-shows and cancellations. Manuf Serv Oper Manage 2010;12(2):347–64.

[18] Alaeddini A, Yang K, Reddy C, Yu S. A probabilistic model for predicting the probability of no-show in hospital appointments. Health Care Manage Sci 2011;14(2):146–57.

[19] Lin J, Muthuraman K, Lawley M. Optimal and approximate algorithms for sequential clinical scheduling with no-shows. IIE Trans Healthc Syst Eng 2011;1(1):20–36.

[20] Sickinger S, Kolisch R. The performance of a generalized Bailey–Welch rule for outpatient appointment scheduling under inpatient and emergency demand. Health Care Manage Sci 2009;12(4):408–19.

[21] Hoot NR, LeBlanc LJ, Jones I, Levin SR, Zhou C, Gadd CS, et al. Real-time evaluation. J Am Med Inform Assoc 2009;16(3):338–45.

[22] Alexopoulos C, Goldsman D, Fontanesi J, Kopald D, Wilson J. Modeling patient arrivals in community clinics. Omega 2008;36(1):33–43.

[23] Williams P, Tai G, Lei Y. Simulation based analysis of patient arrival to health care systems and evaluation of an operations improvement scheme. Ann Oper Res 2010;178(1):263–79.

[24] Adan I, Bekkers J, Dellaert N, Vissers J, Yu X. Patient mix optimisation and stochastic resource requirements: a case study in cardiothoracic surgery planning. Health Care Manage Sci 2009;12(2):129–41.

[25] Groot P, Kremer P, Vries GD. Een raamwerk voor produktiebesturing van ziekenhuizen. Acta Hospitalia 1993;3:45–57.

[26] Vissers J, Bertrand J, De Vries G. A framework for production control in health care organizations. Prod Plan Control 2001;12(6):591–604.

[27] Hans E, Houdenhoven M, Hulshof P. A framework for healthcare planning and control. In: Hall R, editor. Handbook of healthcare system scheduling, vol. 168. US: Springer; 2012. p. 303–20.

[28] Mullen P. Prioritising waiting lists: how and why? Eur J Oper Res 2003;150(1):32–45.

[29] Arenas M, Bilbao A, Caballero R, Gomez T, Rodriguez M, Ruiz F. Analysis via goal programming of the minimum achievable stay in surgical waiting lists. J Oper Res Soc 2002;53(4):387–96.

[30] Gallivan S, Utley M, Treasure T, Valencia O. Booked inpatient admissions and hospital capacity: mathematical modelling study. Brit Med J 2002;324(7332):280–2.

[31] Nunes L, de Carvalho S, Rodrigues R. Markov decision process applied to the control of hospital elective admissions. Artif Intel Med 2009;47(2):159–71.

[32] Adan I, Vissers J. Markov decision process applied to the control of hospital elective admissions. Int J Oper Prod Man 2002;22(4):445–61.

[33] Holm L, Luràs H, Dahl F. Improving hospital bed utilisation through simulation and optimisation: with application to a 40% increase in patient volume in a Norwegian general hospital. Int J Med Inform 2012.

[34] Bair A, Song W, Chen Y, Morris B. The impact of inpatient boarding on ED efficiency: a discrete-event simulation study. J Med Syst 2010;34(5):919–29.

[35] Su S, Shih C. Managing a mixed-registration-type appointment system in outpatient clinics. Int J Med Inform 2003;70(1):31–40.

[36] Ross SM. Introduction to probability and statistics for engineers and scientists. Amsterdam: Elsevier; 2004.

[37] Pham D, Klinkert A. Surgical case scheduling as a generalized job shop scheduling problem. Eur J Oper Res 2008;185(3):1011–25.

[38] Cardoen B, Demeulemeester E, Beliën J. Optimizing a multiple objective surgical case sequencing problem. Int J Prod Econ 2009;119(2):354–66.

[39] Li M, Tsai J, Chen W, Lin H, Pan H, Yang T, et al. Redefining the sonography workflow through the application of a departmental computerized workflow management system. Int J Med Inform 2012.

[40] European Association of Radiology. <http://www.ear-online.org>.

[41] American College of Radiology. <http://www.acr.org>.

[42] Reiner B, Siegel E, Carrino J, Goldburgh M. SCAR Radiologic Technologist Survey: analysis of the impact of digital technologies on productivity. J Digit Imaging 2002;15(3):132–40.

[43] American Radiological Nurses Association. Orientation manual for radiologic and imaging nursing. American Radiological Nurses Association; 2008.

[44] Reiner B, Siegel E. The cutting edge: strategies to enhance radiologist workflow in a filmless/paperless imaging department. J Dig Imaging 2002;15(3):178–1900.

[45] Junk RP, Gilk T. Safety considerations in the design of magnetic resonance imaging. American Institute of Architects; 2007.

[46] Kirkpatrick S, Gelatt Jr C, Vecchi M. Optimization by simulated annealing. Science 1983;220(4598):671–80.

[47] Černý V. Thermodynamical approach to the traveling salesman problem: an efficient simulation algorithm. J Optim Theory Appl 1985;45(1):41–51.

[48] Kanal E, Barkovich A, Bell C, Borgstede J, Bradley Jr W, Froelich J, et al. ACR guidance document for safe MR practices: 20 07. Am J Roentgenol 2007;188(6):1447–74.

[49] Wijewickrama A, Takakuwa S. Outpatient appointment scheduling in a multi facility system. In: Proceedings of the 40th conference on winter simulation; 2008. p. 1563–71.

[50] Cayirli T, Yang KK, Quek SA. A universal appointment rule in the presence of no-shows and walk-ins. Prod Oper Manage 2012.

[51] American College of Radiology. ACR practice guidelines for magnetic resonance imaging. American College of Radiology; 2005.

[52] National Imaging Associates. Diagnostic imaging guidelines. National Imaging Associates Inc.; 2005.

[53] Dreyer KJ, Hirschorn DS, Thrall JH, Mehta A. PACS: a guide to the digital revolution, vol. 48. Verlag: Springer; 2004.

[54] Siegel EL, Reiner B. Work flow redesign: the key to success when using PACS. J Dig Imaging 2003;16(1):164–8.

[55] Ralston M, Coleman R, Beaulieu D, Scrutchfield K, Perkins T. Progress toward paperless radiology in the digital environment: planning, implementation, and benefits. J Dig Imaging 2004;17(2):134–43.

[56] Reiner B, Siegel E, Carrino J. Workflow optimization: current trends and future directions. J Dig Imaging 2002;15(3):141–52.