RAIRO-Oper. Res. 53 (2019) 1453–1474
https://doi.org/10.1051/ro/2018079

RAIRO Operations Research
www.rairo-ro.org

A CUBIC CHROMOSOME REPRESENTATION FOR PATIENT SCHEDULING IN THE EMERGENCY DEPARTMENT

Sarah Ben Othman[1], Faten Ajmi[1,*], Hayfa Zgaya[2] and Slim Hammadi[1]

Abstract. In healthcare institution management, hospital flow control and the prediction of overcrowding are major issues. The objective of the present study is to develop a dynamic scheduling protocol that minimizes interference between scheduled and unscheduled patients arriving at the emergency department (ED) while taking account of disturbances that occur in the ED on a daily basis. The ultimate goal is to improve the quality of care and reduce waiting times via a two-phase scheduling approach. In the first phase, we used a genetic algorithm (based on a three-dimensional cubic chromosome) to manage scheduled patients. In the second phase, we took account of the dynamic, uncertain nature of the ED environment (the arrival of unscheduled patients) by continuously updating the schedule.

Mathematics Subject Classification. 90C27.

Received January 31, 2017. Accepted September 8, 2018.

1. Introduction

Controlling hospital flows and anticipating overcrowding phenomena are major challenges in the management of healthcare production systems. Due to fluctuations in patient flow, healthcare stakeholders have to manage congestion and peaks of activity. Long patient waiting times now constitute a key problem in healthcare institutions in general and emergency departments (EDs) in particular. In France, this is often because the arrival of unscheduled patients at the ED interferes with the treatment of scheduled patients (*i.e.* patients to whom a scheduled consultation time has been already given and who are being treated or are in the waiting room). These unscheduled, real-time perturbations in the ED mean that rescheduling is then required. However, EDs lack procedures and tools for decision-making and appropriate rescheduling.

The present study was performed in the ED at Lille University Medical Centre (Lille, France), which is particularly concerned with the issue of scheduling. As elsewhere in France, many patients wait in the ED for hours as many as 10 h, in some cases – before seeing a physician. These delays can even endanger the patient's life. The problem of long waiting times highlights the need to review the ED management process and implement measures to improve the quality of patient care. In the present study, we focused on optimizing the care process. We had noticed that the unscheduled arrival of patients particularly those requiring emergency

Keywords. Emergency department, dynamic scheduling, scheduled and unscheduled patients, genetic algorithm, three-dimensional cubic algorithm.

[1] Ecole Centrale de Lille, France.
[2] EA2694 University of Lille, France.
*Corresponding author: faten.ajmi@centralelille.fr

treatment – perturbs the treatment process in the ED. If the ED is overcrowded, the arrival of unscheduled patients may interrupt the treatment of scheduled patients and/or require rescheduling around the more urgent cases. We have developed a novel, dynamic approach to patient scheduling based on two complementary processes. The first step concerns the management of scheduled patients in the ED, and is based on a genetic algorithm (GA) with a three-dimensional, cubic chromosome. The second phase involves updating the schedule after the arrival of an unscheduled patient, while taking account of staff availability and skills. The priorities here are to save patients' lives, minimize the overall waiting time for both scheduled and unscheduled patients, and optimize resource use. This approach has proved its effectiveness in improving healthcare processes. It optimizes patient treatment while taking account of the various perturbations that can occur in the ED. Performance indicators (such as total workload of medical staff, overall patient waiting time and response time for healthcare tasks) are then generated and analyzed as a guide to the effectiveness of patient management.

2. State of the art

2.1. Optimization of resource allocation and patient scheduling in healthcare organizations

Many studies have focusing on helping health system managers to make decisions and then evaluate their choices' impacts on the system's efficiency and effectiveness [1, 2]. Managers have to make the best possible decisions when faced with the constraints imposed by the environment within which they operate. Furthermore, managers must optimize cost and performance. To this end, optimization systems [3–5] have been used to evaluate alternatives [6, 7]. The effectiveness of optimization systems is often measured in terms of the cost and the quality of services (a reduction in waiting times, the avoidance of a lack of resources, etc.).

Scheduling problems in health facilities are usually linked to the services required by various categories of users. Each service involves several types of resource (*e.g.* physicians, beds, instruments, etc.), each of which has its own costs. Hence, a range of different data must be gathered, and resources must be assigned to care tasks [8]. In this context, the notion of scheduling in healthcare organizations is becoming increasingly complex: (i) staff should have the diverse skills required to meet the patients' needs, (ii) it is not possible to predict a patient's pathway into a healthcare organizations because factors such as the pathology and the institution's management approach are involved, and (iii) the hospital environment is highly stochastic, which complicates resource planning.

2.1.1. Allocation of resources

A general problem in healthcare is the allocation of scarce medical resources (such as operating theatres or medical staff) so that waiting times are as short as possible. A major difficulty lies in the fact that this distribution must be implemented several months in advance – even when the exact number of patients for each specialty remains uncertain. Another problem arises for cyclical schedules, where the allocation is defined over a short period (a week, for example) and then repeated over the time horizon. In most cases, however, demand varies from week to week: even when the exact demand for each week is known in advance, the weekly schedule cannot be adapted accordingly.

2.1.2. Resource optimization

Mathematical optimization is increasingly relevant in healthcare management. As Belien [9] pointed out: "In the near future of public health, resources will become insufficient. Therefore, we need to find effective ways to plan, prioritize and make decisions". The hospital administration's main task is therefore to efficiently distribute the available medical services and resources. A wide variety of assignment and scheduling problems can arise [10, 11]. Resource allocation is directly linked to a planning problem which consists in establishing the sequence for patient admission. As a general rule, patients requiring specific therapy are first placed on a waiting list and then admitted to hospital. Performance indicators related to the length of these lists are used to determine effectiveness. Long queues are to be avoided, as they represent an enormous cost to the healthcare system [12, 13]. The cost of queuing is a design parameter that must be established by the hospital board for each specialty. In a typical case, the cost is represented by a convex function in which marginal costs increase as the tail lengthens. Shorter lists are obviously preferred, although it is usually impossible (and sometimes perhaps not even desirable) to avoid a certain degree of queueing. Indeed, the absence of a queue for certain specialties might reveal the inefficient allocation of certain scarce resources. Hence, the basic scheduling problem in healthcare is the allocation of resources to medical specialties so as to minimize queuing costs. Clearly, the attribution process must be determined in advance, and may involve negotiations. Thus, resources are allocated at the beginning of a time horizon which can be quite long, ranging from a few months to several years. The number of patients for each specialty is therefore estimated in advance, and the actual number may differ considerably from the initial estimate. Furthermore, schedules are often created with reference to the planning horizon (*e.g.* one to four weeks), and then repeated cyclically. Actual demand may vary from one period to another, even when it is known in advance. A schedule must ensure that queues are as short as possible when the demand for care is maximal (relative to the selected schedule).

2.1.3. Staff assignment

Staff assignment is defined as an optimized construction process for the execution of care tasks. It is generally necessary to assign appropriately qualified staff to specific tasks, in order to meet the service's organization demands while complying with work regulations and seeking to satisfy individual preferences. This method has been adapted and applied to different fields, such as transport systems, healthcare systems, manufacturing, emergency services, and public services. Jaumard *et al.* [14] presented a generalized linear programming model (based on the branch and bound algorithm) for the assignment of nurses with different skills. The main problem is to find a set of individual schedules that satisfy demand-side constraints while minimizing wage costs and maximizing nursing preferences and quality of care. Millar and Kiragu [15] used a network model for the cyclical and non-cyclical planning of nurse schedules, in which the network's nodes represented a feasible model of work-stretch and off-stretch patterns. The resulting problem was essentially a model of the shortest path with lateral constraints. According to Blôchliger [16], construction of a practical model must provide a detailed analysis and a description of the basic elements. Ernst *et al.* [17] provided a detailed review of applications, models and algorithms for staff assignment, including the assignment of medical residents in hospitals. Musa and Saxena [18] focused on a single-phase algorithm that took account of scheduling policies for hospital nurses and their preferences for the weekend. Arthur and Ravindran [19] were the first to use this method with the following four objectives: taking account of staff preferences, and decreasing the number of staff, the minimum number of employees, and staff dissatisfaction. In the first phase of their approach, a goal-based programming model was used to assign days-on and days-off to nurses over the two-week planning horizon. The second phase dealt with specific changes to nurse assignment *via* a heuristic procedure. Lastly, Bard and Purnomo [20] developed a dual heuristic to solve nurses' cyclical preference schedules.

3. Formulation of the problem

Emergency services are permanently confronted by interference between the care of scheduled patients and the arrival of unscheduled patients (particularly those requiring urgent treatment). At present, there is no satisfactory solution to this problem.

The term "emergency" covers two distinct phenomena: recurring flows and sanitary crises. Firstly, recurring flows may be seasonal but the average short- or medium-term trends are known (*i.e.* per month or per year). However, even when the flows are known, the establishment of an efficient, effective, short-term management structure is a major challenge for healthcare production systems. Secondly, flows due to sanitary crises (flu epidemics, heat waves, cold waves, etc.) cannot be foreseen in terms of their magnitude and nature.

In the present study, we considered that a given patient's treatment can be splittable or non-splittable. In fact, a patient's treatment can be interrupted in order to deal with a patient requiring treatment more urgently. A patient may be treated at different times in different places.

We next introduce the mathematical model used to formulate the problem, and then assess the set of solutions obtained with our approach.

3.1. Parameters

NP: a set of N patients to be treated, $\text{NP} = \{P_1, P_2, \ldots, P_N\}$
MS: a set of M medical staff members, $\text{MS} = \{m_1, m_2, \ldots, m_M\}$.
N_s: the number of scheduled patients in the ED.
N_{ns}: the expected number of unscheduled patients.
k: the medical staff member index m_k.
P_j^s: the subset of patients corresponding to (splittable) treatments.
P_j^{ns}: the subset of patients corresponding to (non-splittable) treatments.
w_{kl}: the number of patients managed in common by medical staff members m_k and m_l.
W_k: the workload of medical staff member m_k.
W: the workload of all the medical staff members, $W = \sum_{k=1}^{M} W_k$
$C_{j,k}$: the skill of the medical staff member m_k needed to treat patient j.
tar_j: the arrival time of patient j.
c_j: the theoretical completion time for patient $j's$ treatment.
d_j: the theoretical due date for patient $j's$ treatment.
S: the set of sites at Lille University Medical Centre (radiology facility, MRI facility, central laboratory, etc.); $n_s = |S|$, the number of sites in this set.
R: the set of all available healthcare rooms.
S_r^R: the healthcare room's capacity, $r \in R$.
H: The treatment time horizon, which starts at time D_H and ends at the time F_H. The horizon is divided into several periods p whose lengths are not necessarily the same.

3.2. Decision variables

X_{jpr}: Boolean, 1 if a healthcare treatment or a portion of it corresponding to patient j is placed at period p in room r, and 0 if not.
A_{jpr}: an integer representing the number of patients P_j, $\leq N$ having splittable healthcare treatments placed at period p in room r.
C_{lk}^{xy}: Boolean, set to 1 if medical staff members m_l and m_k, managing patients in common, are placed at two periods x and y back-to-back during the same day, and set to 0 if not.
C_{lk}^{SP}: Boolean, set to 1 if medical staff members m_l and m_k, managing patients in common, are placed in two different sites during two periods with a gap between them, and set to 0 if not
U_{jpc}: Boolean, set to 1 if one or more medical staff members are assigned at period p to patient P_j in the corridor c, and set to 0 if not
C_{lk}^{T}: Boolean, set to 1 if medical staff members m_l and m_k are managing patients in common and are placed in healthcare rooms located at different sites during two periods with a gap between them, and set to 0 if not. For this, travel T is necessary.

3.3. Institutional parameters

w^T: penalty weighting for patient travel between different sites within the healthcare organization.
w_c: penalty weighting for using the Emergency Department (ED) corridor c.
BC_p^r: penalty weighting for exceeding the capacity of room r at the period p.
MS_p^k: penalty weighting for exceeding medical staff member $m_k's$ workload during period p.
G^{PS}: gap of the treatment period spread penalty.

3.4. The objective function

Minimize

$$C(w) + C(w^T) + C(w_c) + C(\mathrm{BC}_p^r) + C(\mathrm{MS}_p^k) + C(G^{\mathrm{PS}}) \tag{3.1}$$

where $C(w)$: the cost generated by the waiting times of both scheduled and unscheduled patients in the ED. W is calculated as follows:
$W = Min \sum_{j=1}^{N_s} \sum_{k=1,k\neq j}^{N_{ns}} (W_{s,j} + W_{ns,k})$ where $W_{s,j}$ is the scheduled patients' waiting time and $W_{ns,k}$ is the unscheduled patients' waiting time.
$C(w^T)$: the cost generated by patient travel between the different sites within the healthcare organization.
$C(w_c)$: the cost generated by treating a patient in the corridor of the ED.
$C(\mathrm{BC}_p^r)$: the cost generated by exceeding the capacity of room r at the period p.
$C(\mathrm{MS}_p^k)$: the cost generated by exceeding medical staff member $m_k{}'s$ workload at the period p.
$C(G^{\mathrm{PS}})$: the cost generated by the gap of the treatment period spread penalty.

The objective function is a sum of penalty terms. Each of the terms refers to a specific, flexible constraint (see Sect. 3.6).

3.5. Strong constraints

The following strong constraints influence the solution's feasibility.

SSPT_r^p: the sum of splittable patient treatments (or portions of treatments) allocated to treatment room r at period p should not exceed the treatment room's capacity:

$$\forall r \in R, \forall p \in \mathbb{R}, \sum_{p_j \in P_j^s} A_{jpr} \leq S_r^R. \tag{3.2}$$

Linking the variables X_{jpr} and A_{jpr} related to splittable treatments:

$$\forall P_j \in P_j^s, \forall r \in R, \forall p \begin{cases} A_{jpr} \leq \mathrm{Card}(P_j^s) * X_{jpr} \\ A_{jpr} \geq X_{jpr} \end{cases}. \tag{3.3}$$

The two parts of the above equations are required to check whether $X_{jpr} = 1, A_{jpr} \neq 0$.
SPS_r^p: the sum of patients with splittable treatments should be equal to: $\mathrm{Card}(P_j^s)$,

$$\forall P_j \in P_j^s, \sum_{r\in R} \sum_{p \in H} A_{jpr} = \mathrm{Card}(P_j^s). \tag{3.4}$$

NSPT_r^p: a non-splittable patient treatment should be assigned to a single treatment room:

$$\forall P_j \in P_j^{ns}, \sum_{p\in H} \sum_{r \in R} X_{jpr} = 1. \tag{3.5}$$

RP_r^{pq}: a room cannot be used by two patients in two overlapping periods p and q:

$$X_{j_1pr} + X_{j_2qr} \leq 1, \forall P_{j_1}, P_{j_2} \in \mathrm{NP}, \forall r \in R. \tag{3.6}$$

MSP_r^{pq}: two medical staff members managing patients in common cannot be allocated during the same period or during two overlapping periods p and q:

$$X_{j_1p} + X_{j_2q} \leq 1, \forall P_{j_1}, P_{j_2} \in \mathrm{NP}. \tag{3.7}$$

3.6. Flexible constraints

The solution's quality is determined by the following flexible constraints.

CPP_{kl}^{sp}: whenever two medical staff members managing patients in common are placed on different sites in two consecutive periods, a patient travel penalty is applied:

$$C^{\mathrm{SP}} = w^T \sum_{m_l, m_k \in \mathrm{MS}} C_{lk}^{\mathrm{SP}}. \tag{3.8}$$

MPC_p: whenever a medical staff member is allocated to treat one or more patients in the corridor at period p, a corridor penalty is applied:

$$U^c = w_c \sum_{p_j \in \mathrm{NP}, p \in H} U_{jpc}. \tag{3.9}$$

Cap_p^r: whenever at least two patients are treated at the same period p in the same room r specially in the overcrowding situation, a capacity penalty is applied:

$$\mathrm{BC}^r = \mathrm{BC}_p^r \sum_{p \in \mathrm{NP}, r \in R, p \in H} U_{jpr}. \tag{3.10}$$

In the following section, we solve the above-described problem while meeting the different constraints. To optimize the solution, we decided to adopt an aggregative approach without seeking to apply appropriate weightings. In real-life healthcare situations, it is very difficult to define suitable weights for these criteria. The present study assessed the results of simulations that generated some of these criteria separately or (in some cases) together.

4. The rolling-horizon approach to scheduling

4.1. The scheduling environment

Figure 1 shows the scheduling environment with three kinds of patients: urgent patients (UPs), scheduled patients (SPs) and non-scheduled patients (NSPs).

4.1.1. Assumptions

- A medical staff member is present in the scheduling horizon in the ED. The number of scheduled patients in a scheduling horizon is N_s, whereas the expected number of unscheduled patients is N_{ns}. All the unscheduled patients arrive randomly at the ED; on arrival, they must be assigned with a theoretical scheduled consultation time.
- In France, EDs never close. Each arriving patient j should be registered at the reception desk at time tar_j. None of the patients who arrive at the ED are rejected, and all patients should be treated during the current scheduling horizon or the next scheduling horizon.
- Each patient corresponds to a set of healthcare operations to be executed in a parallel or in a sequential manner by one or more medical staff members (staff physicians, nurses, interns, etc.).
- Medical staff members are organized into teams. Each team contains at least one physician. Some teams contain additional staff members (nurses, paediatricians, etc.), depending on the patient's pathology.
- The scheduling horizon H starts at time D_H and ends at the time F_H. In the present study, we consider that the duration of the scheduling horizon is 4 h.
- The scheduling horizon is divided into several periods whose durations are not necessarily equivalent. If two periods have the same duration, the number and the duration of slots in each period may differ. In general, a period contains multiple slots. A slot is allocated to a scheduled patient. Each period contains at least one slot. A slot's scheduled consultation time is given by the start time of the period to which it belongs. Hence, if two or more slots are included in a period, the scheduled patients assigned to the same slot have the same scheduled consultation time.

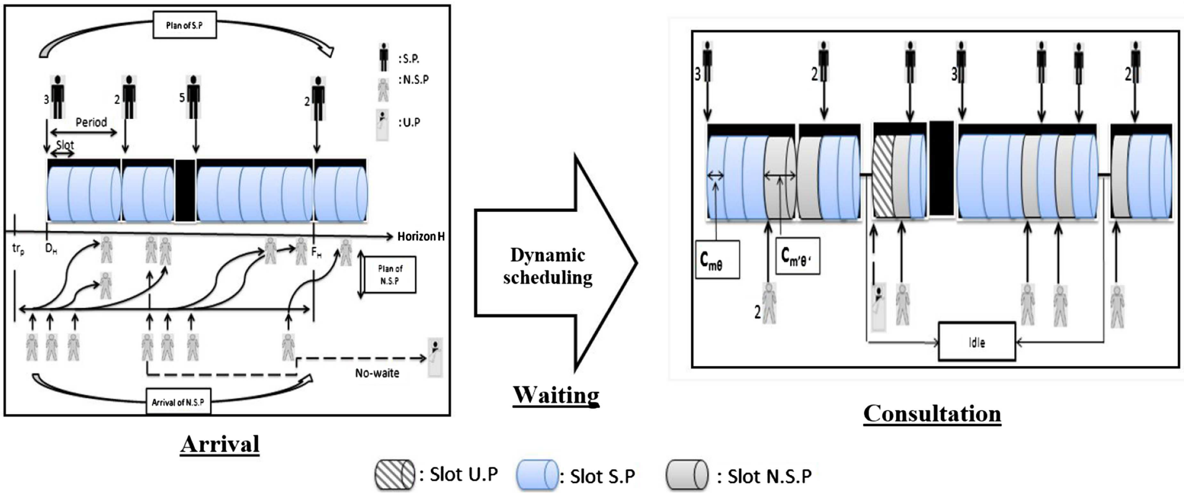

Figure 1. The scheduling environment.

- According to the stochastic behaviour of the medical staff's consultation time, let $C^{j}_{m,\theta}$ be the average consultation time of the medical staff member m having the skill θ to treat patient j.
- When the medical staff member becomes available, the waiting patient with the earliest scheduled consultation time is called. If the waiting room is full and it is not possible to call all the patients during the same scheduling horizon, the remaining patients and the new arrivals will receive a scheduled consultation time in the next scheduling horizon.
- In the ED, the most urgent cases are given the highest priority. Hence, when patients requiring urgent treatment arrive, the current scheduling can be interrupted and rescheduling is required because these patients should not have to wait for a consultation.

4.1.2. Performance measures

Let the waiting time for a scheduled patient j ($W_{s,j}$) be the sum of the patients' waiting times between the registration and the theoretically scheduled consultation time W_{ar}, and the waiting time before the first consultation W_{fc}, where:

$$W_{s,j} = W_{ar,j} + W_{fc,j} \tag{4.1}$$
$$W_{ar,j} = \max(0, t_{s,j} - t_{ar,j}) \tag{4.2}$$
$$W_{fc,j} = \max(0, t_{fc,j} - t_{s,j}) \tag{4.3}$$

where $t_{ar,j}, t_{s,j}$ and $t_{fc,j}$ are respectively the arrival time, the theoretically scheduled consultation time and the first consultation time for the patient j.

The waiting time for an unscheduled patient P_k ($W_{ns,k}$) is the sum of the patients' waiting time between the registration and the theoretically assigned scheduled consultation time $W_{ar,k}$, and the waiting time before the first consultation $W_{fc,k}$, where:

$$W_{ns,k} = W_{ar,k} + W_{fc,k} \tag{4.4}$$
$$W_{ar,k} = \max(0, t_{s,k} - t_{ar,k}) \tag{4.5}$$
$$W_{fc,k} = \max(0, t_{fc,k} - t_{s,k}). \tag{4.6}$$

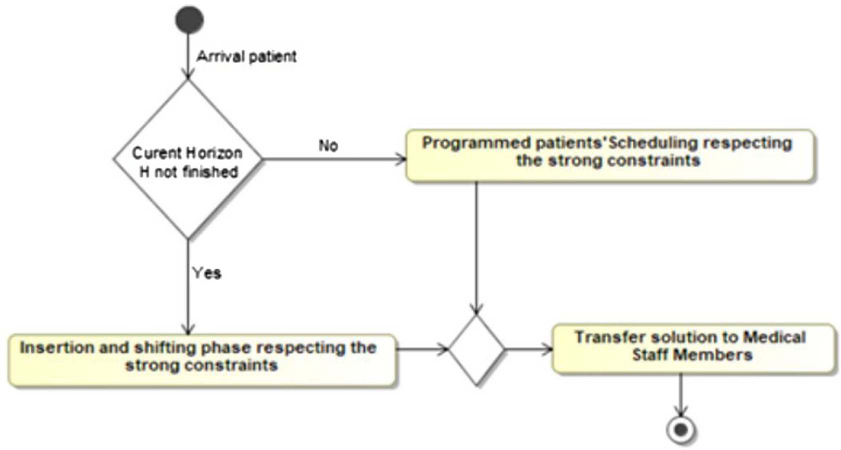

Figure 2. The scheduling approach.

The two equations (4.1) and (4.4) are mathematically equivalent but semantically different. In fact, the scheduling method developed in the present study assigns a theoretical scheduled consultation time to each unscheduled patient at his/her time of arrival and then guides him/her to the waiting room at the scheduled consultation time. On the basis of the scheduled consultation time, W_{ar} is calculated for each registered patient. The objective is to comfort patients and reduce their stress by keeping them informed of their waiting time prior to the first consultation. If the first consultation time is equal to the scheduled consultation time, then the waiting times $W_{fc,j}$ and $W_{fc,k}$ are equal to 0. In the event of perturbations (overcrowding, a lack of medical staff, worsening in the patient's health status, etc.), the consultation times are rescheduled, the first consultation time increases, and so the patient's waiting time lengthens.

Most of patients prefer early scheduled consultation times – especially when they arrive. To satisfy these preferences, the waiting time (based on the arrival time W_{ar}) should be reduced by assigning available medical staff with appropriate skills as quickly as possible.

5. The scheduling procedure

The sequential treatment of patients is dynamically scheduled, which requires the real-time generation of activity plans for each medical staff member. A multidisciplinary medical team is formed, and a treatment role is assigned to each team member. The schedule is updated whenever the patient input stream changes.

Our approach is based on an offline phase and an online phase.

The offline phase consists in scheduling the scheduled patients who arrive at the ED. A GA is applied in this phase, and the scheduled patients correspond to the constraints of the GA. The online phase takes account of a dynamic feature; the arrival of unscheduled patients at the ED. These patients are treated with regard to their pathologies and their emergency status. Hence, the online phase uses a shifting and insertion method based on the notion of periods and horizons (Fig. 2).

5.1. The offline phase: scheduling with a GA

The treatment plan is generated by applying a dynamic, responsive GA. The algorithm is designed to (i) optimize the assignment of patients to medical staff with the skills needed to treat them, and (ii) minimize patient waiting times and overall costs while maintaining the quality of care. The scheduling algorithm selects the appropriate medical staff member for the treatment of a given patient, according to the staff's availability and skills. An emergency alert resulting from the need for a medical staff member triggers an updating process by

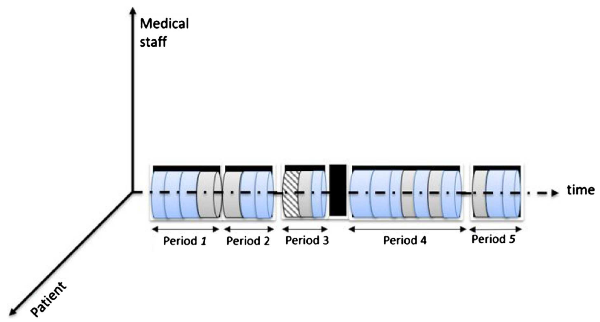

FIGURE 3. A representation of a cubic chromosome (a patient/medical staff/time cube).

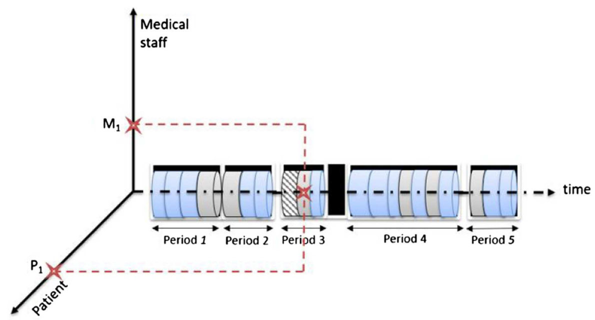

FIGURE 4. A single cubic assignment.

the scheduling algorithm. This situation may lead to the interruption of a patient's treatment and the initiation of treatment of a more urgent case or a shift in one or more treatment processes to make space for unscheduled patients requiring urgent treatment.

The goal is to minimize the overall waiting time that the patient spends in the ED and the costs described in the third section of this article.

The purpose of the GA is to provide an approximate solution to the optimization problem, insofar as an exact method cannot solve the problem within a reasonable time. The potential solution(s) provided by the GA necessarily requires the involvement of the different medical staff members in the ED. In the following section, we will describe our patient scheduling approach in detail.

5.1.1. Definition of the chromosome

We chose to use a three-dimensional cube chromosome with the following three axes: "medical staff", "patients", and "time" (Fig. 3). The time axis is divided into intervals of different sizes. The scheduling horizon is divided into several periods that do not necessarily have the same duration. If two periods have the same duration, the number and the duration of slots in each period may differ. In general, a period contains more than one slot.

In view of the division of the time axis into many slots, each medical staff member is assigned to a patient in a specific slot from a specific period. For example, Figure 4 shows a medical staff member M_1 treating patient P_1 in the second slot of period 3. The Figure 5 shows the sequential assignment of medical staff members M_{y1} and M_{y2}, to patients P_{z1} and P_{z2} in periods 3 and 4, respectively. The Figure 6 shows a multiskill cubic assignment, which is made possible by our choice of the type of chromosome. Here, patient P_{Z1} needs two different skill sets for his/her treatment, medical staff members M_{y1} and M_{y3} are assigned to the same period 3 and the same

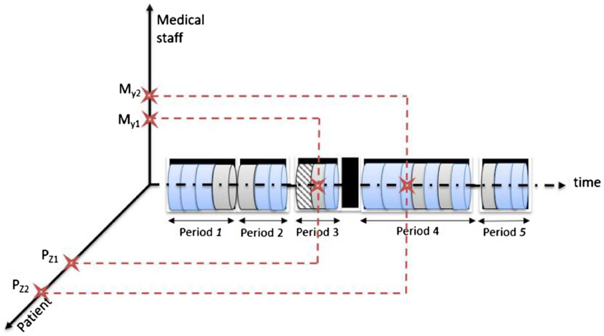

FIGURE 5. A sequential cubic assignment.

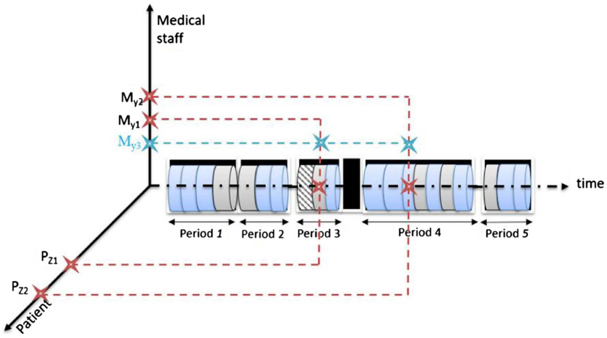

FIGURE 6. A multiskill cubic assignment.

Algorithm 1. Construction of the $IniPopL$.

input : Fixed NL, the size of the chromosome population $InitPopL$
Fixed NR, the size of the chromosome population $IniPopR$
output: $IniPop$L of NL chromosomes generated by applying the list algorithm
$IniPopR$ of NR chromosomes generated by applying a random process

```
Begin
i ← 0;
IniPopL = ∅;
while i ≤ NL do
    Find a cubic chromosome i as a feasible (or suboptimal) solution of a single objective optimization model by
    applying a list algorithm;
end
IniPopL = IniPopL ∪ {chromosome i};
i ← i + 1;
end
```

slot. Patient P_{Z2} needs a pair of different skills, and medical staff members M_{y2} and M_{y3} are assigned to the same period 4 and the same slot.

5.1.2. The initial chromosome population

The first step is the formation of an initial population as the starting point for execution of the algorithm.

Algorithm 2. The cubic GA approach.

input : $InitPopL$, $InitPopR$, N is the global size of initial population
output: a set of N good scheduling solutions

```
Begin
Construction of IniPopL: find NL feasible cubic chromosomes
Construction of IniPopR: find NR = N − NL partial feasible solutions at random
Merging of IniPopL and IniPopR (N cubic chromosomes)
while (stop criterion are not reached) do
    - Evaluate individuals
    - Select 2 parents P1 and P2 at random
    - Apply a controlled crossover algorithm with a probability p_c,
    in order to obtain offspring1 and offspring2
    - Apply a controlled mutation algorithm with a probability p_m
    - Select N new individuals and build a new population
    - Update the stopping criterion
end
end
```

Patient \ Time	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	75-80	80-85	85-90	90- 95
Patient 1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Patient 2	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
Patient 3	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0
Patient 4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1

FIGURE 7. Chromosome A.

Patient \ Time	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	75-80	80-85	85-90	90- 95
Patient 1	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0
Patient 2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
Patient 3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Patient 4	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0

FIGURE 8. Chromosome B.

We used two methods to build the initial population:

- The first method consists in recovering the initial population solutions ($IniPopL$) generated by a list algorithm with dynamic priority rules.
- The second method consists in generating initial population solutions at random ($IniPopR$) but which are viable because they comply with the strong constraints.

The details of the GA used in the present study are as follows: (see Algorithms 1 and 2).

5.1.3. The controlled crossover schema

This used in order to move the start time of the patient treatment process forward or backward for a given medical staff member. It does not change the assignment of patients, *i.e.* which medical staff member treats which patient). Only the "time" and "patient" axes are considered.

Time / Patient	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	75-80	80-85	85-90	90-95
Patient 1	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0
Patient 2	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
Patient 3	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0
Patient 4	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0

FIGURE 9. Offspring Chromosome 1.

Time / Patient	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	75-80	80-85	85-90	90- 95
Patient 1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Patient 2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
Patient 3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Patient 4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1

FIGURE 10. Offspring Chromosome 2.

Example

The time axis is divided into 5 min intervals. Each slot in the time axis is a Boolean equal to 1 if the patient is assigned to the slot, or 0 if not. (see Figs. 7 and 8).

If the mask is 0110, then the chromosome resulting from the crossover will be as follows: (see Figs. 9 and 10)

The crossover yields two viable offspring chromosomes, so no correction is needed. The viability is checked on the "medical staff" axis. In fact, we need two different medical staff members in the slots [10–15], [55–60] and [60–65] on offspring chromosome 2 because the same medical staff member cannot treat two different patients at the same time (a constraint related to the equation (3.7)). Furthermore, patient 2 and patient 4 in offspring chromosome 2 must be treated by two different medical staff members in the slot [80–95].

This phenomenon shows the value of using a three-dimensional cubic chromosome to check compliance with strong constraints.

5.1.4. The controlled mutation schema

The mutation is a partially random operation that enables us to modify the solutions and move towards an optimum or perhaps move out of a local optimum. In our case, the mutation modifies the Booleans present in our chromosomes. Not all chromosomes are mutated; the probability of mutation is <1. If the chromosome is selected, it will then go through the slots (according to the "medical staff", "patient to treat", "time" axes) and change their values in accordance with simple rules.

The slots are changed at random. Each slot has a predetermined probability of being mutated. If the selected slot is to be changed from 1 to 0, there are no additional conditions; only one patient is treated at a given time by a medical staff member. If the selected slot is to be changed from 0 to 1, we have to check that the medical staff member has the requisite skills and is available to treat the patient in the slot. If the condition is checked, the slot is mutated.

This first phase of the mutation can thus remove patients from a medical staff member or assign them to him/her if he has the needed skills and is available. However, the durations of the patient treatment processes may be inaccurate, and the treatment is divided into several slots. This mutation is controlled by the medical staff axis.

	0-5 min	5-10 min	10-15 min	15-20 min
Patient 1 (treatment duration: 10 minutes)	0	1	1	0
Patient 2 (treatment duration: 3 minutes)	1	0	0	0
Patient 3 (treatment duration: 5 minutes)	1	0	0	0

Mutation: changes are in shown grey

	0-5 min	5-10 min	10-15 min	15-20 min
Patient 1 (treatment duration: 10 minutes)	0	0	1	1
Patient 2 (treatment duration: 3 minutes)	0	1	0	0
Patient 3 (treatment duration: 5 minutes)	1	0	0	0

FIGURE 11. Mutation.

In order to comply with the viability of the final set of generated solutions (resulting from the application of the GA with controlled crossover and mutation operators), we set a number of constraints to be complied with by these operators. These constraints guide us in the search for the optimal solution and accelerate the convergence.

5.1.5. Selection

After crossover, our population increases as the offspring chromosomes join the parent chromosomes. It is then necessary to select the chromosomes that will be part of the new population before rescheduling.

We first evaluated our set of solutions by calculating the value of the objective function (see Sect. 3.3). We calculated its strength of each solution and normalized it as a percentage of the total strength. Selecting only the strongest solutions would not guarantee a great diversity in our solutions, and selecting solutions at random would perhaps remove good solutions. We decided to select a percentage of the best solutions, and then select those that remain on the roulette wheel. The probability of selection corresponds to the normalized strength. This ensures the selection of varied, strong solutions.

5.2. The online phase: real-time rescheduling

This phase deals with the interference between scheduled and unscheduled patients arriving at the ED, which prompts real-time rescheduling. The goal is to reduce the waiting time of both scheduled and unscheduled patients. The process looks at whether an unscheduled patient can be inserted into the schedule generated by the offline phase without affecting his/her neighbouring patients. To this end, the process first seeks medical staff members who have the appropriate skills for treating the patient to be inserted.

To take account of interference between scheduled and unscheduled patients, we need to consider the inter-period waiting time in each horizon.

This work assumes that the scheduling horizon which represents consultation time window is divided into several periods as already mentioned above. A consultation for an unscheduled patient is scheduled in the first empty slot in the period, as shown in Figure 12.

In principle, the start time of the first empty slot gives the patient's theoretical scheduled consultation time. The maximum acceptable number of unscheduled patients in the period is difficult to estimate because slots in the same period can have different lengths. For non-urgent patients, the real-time rescheduling is performed by the Algorithms 3 and 4.

The end of the consultation window $t_{\max,H}$ is used as the scheduled consultation time for patients who are not included in any of the periods in the horizon H.

Algorithm 3. Search First Free Slot (Period p, Horizon H, Patient w, Deb_p, Fin_p).

input : $\mathbf{Nbs_p}$: Number of slots in the period p.
TAB[$\mathbf{1} \dots$ **Nbsp**] a table contains the length of each slot.
FREE[$\mathbf{1} \dots$ **Nbsp**] a table contains 1 or 0.
$FREE(i) = 0$ if the slot i is free, otherwise the slot i is full.
output: T_w the start time of the first free slot

Begin Calculate $HW_{p,h}$;
if $(HW_{p,h} + C^w_{m,\theta} < HW_{p,h,max})$ **then**
    Assign-patient w to the period p;
end
$T_w = Deb_p$: the slot's start time
for $i = 1$ *to* Nbs_p **do**
    if $FREE[i] == 0$ **then**
        return T_w
    else
        $T_w = T_w + TAB[i]$
    end
end
$T_w = Fin_p$// the period p is overloaded
return T_w
end Search First Free Slot

Algorithm 4. Scheduling new arrivals ($Patient\ w, Time\ t, Horizon\ H$).

input : $\mathbf{Deb_H}$: the Start time of the horizon H represents the start time of the consultation;
We consider that the arrival of unforeseen patient follows poisson distribution
t: Current arrival time of patients
$\mathbf{Fin_H}$: The end of the horizon H;
$\mathbf{NB_{p,H}}$: Number of periods in the horizon H;
$\mathbf{TAB_H}$: $Table[1 \dots NB_{p,H}]$ contains the different lengths of each period.
output: T_w the start time of the first consultation specifying the horizon, period and slot.

Begin
if $Urgent\_patient$ **then**
    No-wait-consultation
end
if $t \leq Deb_H$ **then**
    $p = 1$ (the first period in the horizon)
end
else
    for $(i = 1$ *to* $NB_{p,H})$ **do**
        if $t \leq Deb_H + TAB_H[i]$ **then**
            $p = i$;
            Save the start of the period Deb_p
            Save the end of the period Fin_p
        end
    end
end
T_w = Search First Free Slot ($p, Horizon\ H, Patient\ w, Deb_p, Fin_p$)
end Scheduling new arrivals

The present approach assumes that each period p has its own length Δp, and that the start time of each period is x minutes behind the start time of its first free slot Figure 12. The maximum workload level per period

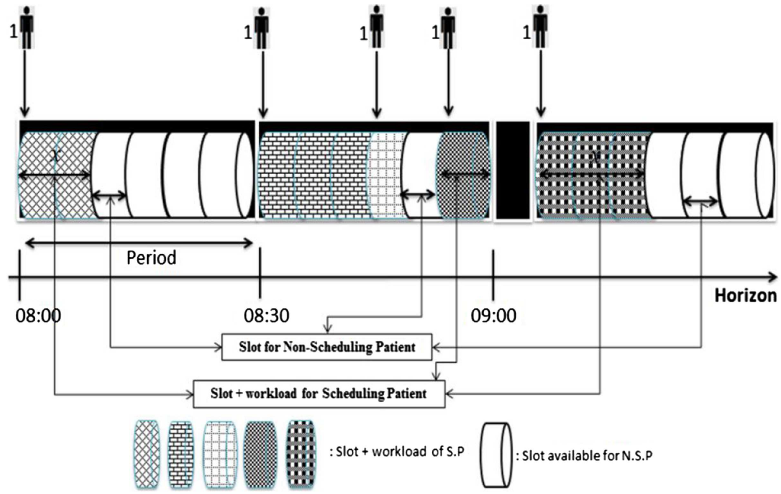

FIGURE 12. An example in which an unscheduled patient's consultation time is determined on the basis of free slots.

is $\mathrm{HW}_{p,H,\max}$. For example, if there are three scheduled patients j and two unscheduled patients k in period p, then $\mathrm{HW}_{p,H}$ is given by $\mathrm{HW}_{p,H} = 3 * C^{j}_{m,\Theta} + 2 * C^{k}_{m',\Theta'}$

6. Simulations and results

Prior to our simulations, we collected data in the ED at Jeanne de Flandre Hospital (part of Lille University Medical Centre). This ED receives about 24 000 visits per year (an average of 458 per week and 66 per day), of which 20% take place in a short-stay hospitalization unit (SSHU) and 80% take place in an outpatient unit. It has 10 beds in the SSHU, 10 consultation boxes in the outpatient unit, a suturing room, a plaster room, an emergency room, and two waiting rooms. In the event of overcrowding, vacant beds in the SSHU can be transformed into consultation boxes.

In the present section, we describe the effectiveness and efficiency of our approach to scheduling. We first describe the real data collected in the ED. Next, we generated realistic random instances of the real data and studied dynamic, rolling-horizon scheduling in more detail. With a view to investigating the interactions between the objective functions and determining how the patients' waiting times affect costs, we carried out several different computational experiments.

6.1. Description of the data

We analyzed a sample of data collected over a period of almost three years, from January 2011 to November 2013 (Fig. 13).

Figures 14–16, respectively show the numbers of patients per month, per week, and per day for the three years of the study period.

The monthly variations are almost identical from one year to the next, and depend on certain periods of the year. The autocorrelation function is represented graphically (*i.e.* as a correlogram) in Figure 17. It shows a peak with a shift of 12 intervals, reflecting the correlation of the data for each 12-month period and thus the seasonality for each 12-month period.

The weekly variations in Figure 15 shows troughs for holiday periods and peaks for flu epidemics. Outside the summer holidays, the mean data were very similar from one week to another. In contrast, we observed regular

number	Arrival Date	Exit Date	duration	MD	diagnostique	echographie	scanner	radiologie	biologie	nbreTests	codeN
69453	28/11/2013 08:13:00	28/11/2013 13:18:00	305	3091	Autres pneumopathies bacté	0	0	0	0	0	
69454	28/11/2013 08:19:00	28/11/2013 16:44:00	505	3091	Gale	0	0	1	1	2	
69455	28/11/2013 08:33:00	28/11/2013 08:35:00	2	3092	Intoxication par autres analgé	0	0	0	0	0	
69456	28/11/2013 09:14:00	28/11/2013 12:25:00	191	3091	Fractures multiples des méta	0	0	0	0	0	
69457	28/11/2013 09:27:00	28/11/2013 15:41:00	374	3091	Gastroentérites et colites d¿c	0	0	1	1	2	
69458	28/11/2013 09:31:00	28/11/2013 14:30:00	299	3091	Nausées et vomissements	0	0	0	1	1	
69459	28/11/2013 09:34:00	28/11/2013 14:24:00	290	3091	Morsure de la joue et de la lè	0	0	1	0	1	
69460	28/11/2013 09:34:00	28/11/2013 11:57:00	143	3091	NA	0	0	0	0	0	
69461	28/11/2013 09:40:00	28/11/2013 11:19:00	99	3092	Autres difficultés liées à l'env	0	0	0	0	0	
69462	28/11/2013 09:42:00	28/11/2013 14:57:00	315	3091	Contusion de(s) doigt(s) sans	0	1	0	0	1	
69463	28/11/2013 09:59:00	29/11/2013 09:59:00	1440	3092	Laryngite (aiguë)	0	0	0	0	0	
69464	28/11/2013 10:13:00	28/11/2013 13:02:00	169	3091	Malaise	0	0	0	0	0	
69465	28/11/2013 10:14:00	28/11/2013 11:36:00	82	3092	Diabète sucré insulino-déper	0	0	0	0	0	
69466	28/11/2013 10:21:00	28/11/2013 13:43:00	202	3091	Asthme, sans précision	0	0	0	0	0	
69467	28/11/2013 10:35:00	29/11/2013 10:16:00	1421	3092	Gastroentérites et colites d¿c	0	0	0	0	0	
69468	28/11/2013 10:50:00	28/11/2013 17:27:00	397	3091	Pyélonéphrite (non obstructi	0	0	0	1	1	
69469	28/11/2013 10:52:00	28/11/2013 14:27:00	215	3091	Autres conjonctivites	0	0	0	0	0	
69470	28/11/2013 10:59:00	28/11/2013 15:07:00	248	3091	Inquiétude et préoccupation	0	0	1	1	2	
69471	28/11/2013 11:02:00	28/11/2013 20:14:00	552	3091	Plaie ouverte d'autres parties	0	0	1	0	1	
69472	28/11/2013 11:05:00	28/11/2013 15:21:00	256	3091	Bronchiolite (aiguë), sans pré	0	0	0	0	0	
69473	28/11/2013 11:07:00	28/11/2013 15:54:00	287	3091	Bronchiolite (aiguë), sans pré	0	0	0	0	0	
69474	28/11/2013 11:15:00	28/11/2013 15:05:00	230	3091	Rhinopharyngite (aiguë) [rhu	0	0	1	1	2	
69475	28/11/2013 11:32:00	28/11/2013 23:02:00	690	3092	Convulsions, autres et non pr	0	1	0	1	2	
69476	28/11/2013 11:47:00	28/11/2013 14:57:00	190	3091	Rhinopharyngite (aiguë) [rhu	0	0	0	1	1	
69477	28/11/2013 11:50:00	28/11/2013 15:19:00	209	3091	Hernie inguinale, (unilatérale	0	0	0	0	0	
69478	28/11/2013 11:58:00	28/11/2013 17:54:00	356	3091	Rhinopharyngite (aiguë) [rhu	0	0	1	0	1	
69479	28/11/2013 12:12:00	28/11/2013 17:11:00	299	3091	Douleur aiguë	0	0	1	0	1	
69480	28/11/2013 12:30:00	28/11/2013 16:43:00	253	3091	Néphrite tubulo-interstitielle	0	0	1	0	1	
69481	28/11/2013 12:36:00	28/11/2013 18:28:00	352	3091	Entorse et foulure de la chevi	0	0	0	0	0	
69482	28/11/2013 12:46:00	28/11/2013 15:44:00	178	3091	Malaise	0	0	0	0	0	

Record: No Filter Search

Figure 13. Database from the ED in Jeanne de Flandre Hospital.

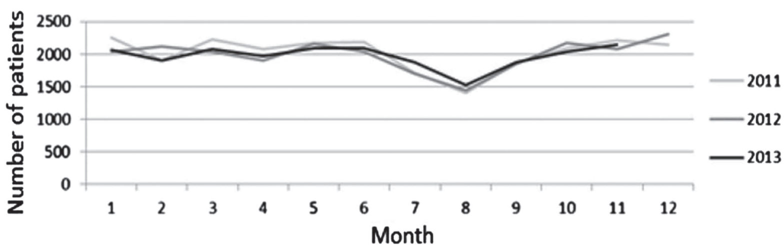

Figure 14. The number of patients per month at the ED in Jeanne de Flandre Hospital.

variations over the seven days of the week; this can be seen as recurring peaks in correlogram with a period of 7 (Figs. 18 and 19).

Figure 16 shows the daily frequencies, where the phenomenon becomes less stable due to the appearance of the variations previously absorbed in the broader views. We noted three very strong peaks (one per year): 119 patients on November 27th 2011, 97 on February 19th 2012, and 100 on February 7th 2013 (compared with a daily of 66 patients).

The value of managing overcrowding is emphasized by refining the time horizon. For effective decision-making, it is best to adopt a time scale that enables patient rescheduling.

Following our observations in the ED and interviews with the medical staff, we noted that waiting times in the ED could be as long as 5 h. The ED at Jeanne de Flandre Hospital did not have a decision support system or information system capable of managing overcrowding. Medical staff members gave the highest consultation priority to the most urgent patients and then to previously scheduled patients. Unscheduled patients in the ED

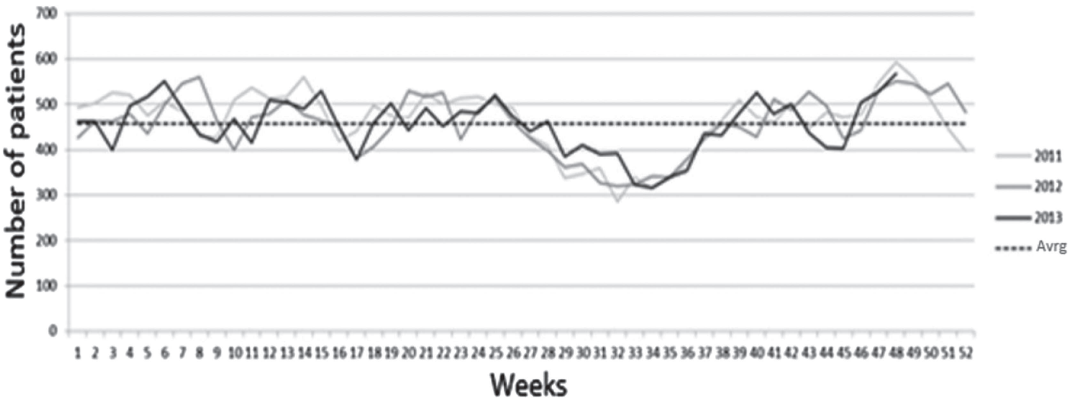

FIGURE 15. The number of patients per week at the ED in Jeanne de Flandre Hospital.

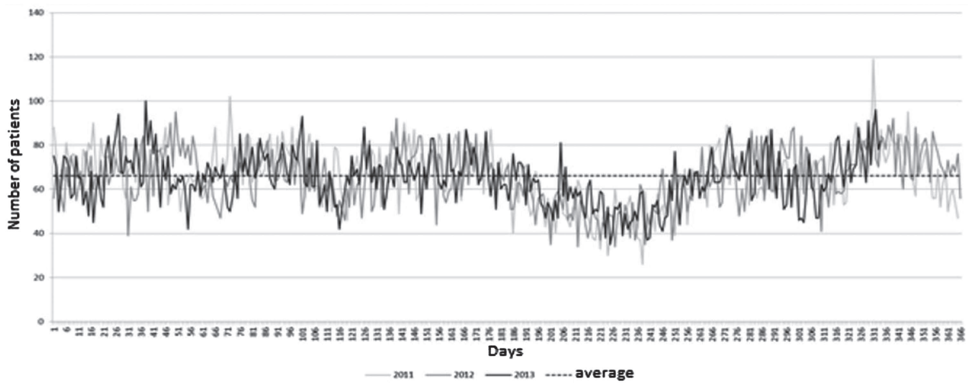

FIGURE 16. The number of patients per day at the ED in Jeanne de Flandre Hospital.

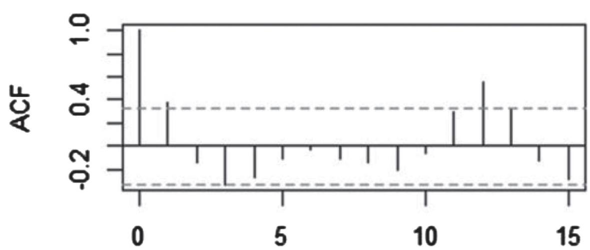

FIGURE 17. Correlogram related to the monthly period. ACF: autocorrelation function.

had to wait in the waiting room and sometimes in corridors without obtaining a scheduled first consultation time, which increased their level of anxiety.

Our approach's level of performance was compared with that of the conventional method used in the ED. A database analysis enabled us to simulate the patients' waiting times, which appeared to be excessive in some cases.

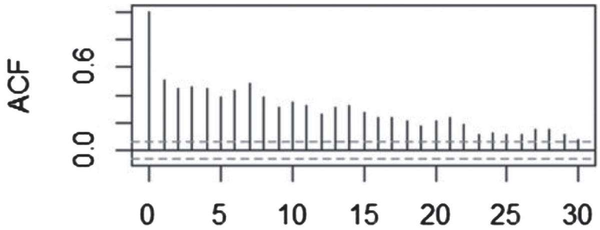

FIGURE 18. Correlogram relating to the daily period. ACF: autocorrelation function.

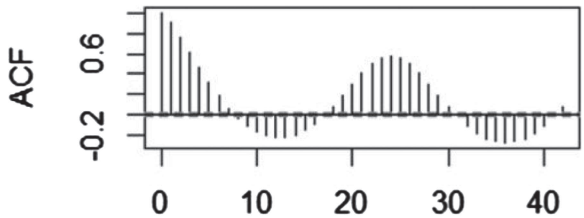

FIGURE 19. Correlogram related to the hourly period. ACF: autocorrelation function.

6.2. Computational results and discussion

As emphasized in Section 4, our two-phase, rolling-horizon patient scheduling is revised whenever unscheduled patients requiring urgent treatment arrive in the ED, in order to optimize the schedule for patients whose waiting times are longer or shorter than expected. The waiting list varies over time as patients arrive and as patients are treated.

We analyzed our computational results in two main steps. Firstly, we investigated the effectiveness of the GA-based algorithm and validated our chromosome model. Secondly, we implemented our approach in a real ED, and evaluated its applicability and performance.

In order to evaluate our approach's level of performance, extend our computational results and generalize our method, we applied our two scheduling phases to solve 10 randomly generated problem instances with different numbers of patients. The results were compared with those obtained in practice (according to the ED database used by the medical staff) and those generated by the list algorithm (implemented in Java). These instances were generated from real data. In the ED, emergencies can be treated with the list algorithm in order to find a quick (but not necessary optimal) solution. The list algorithm's dynamic priority rules mean that it is particularly suited to the scheduling problem. It is flexible and is easily implemented in real time. Our problem is solved by dynamic priority rules: the patients' arrival time and level of urgency. The algorithm maintains a list of all the ready-to-be-scheduled tasks after registration at the reception desk.

Table 1 shows the real ED data related to the test problem scheduling, together with the results obtained with the list algorithm, the GA-based approach, and the practical case. The gap between the solutions (related to the mean total waiting time per patient per instance (a day) is shown in Figure 20.

Table 1 shows that minimization of the waiting time is associated with an increase in the medical staff's workload – especially when using the GA-based approach (Fig. 21). In fact, adjusting the physicians' total idle time minimizes the average total waiting time.

As can also be seen from Table 1, the solutions obtained with the list algorithm for the first seven test instances are close to those obtained by the practical solution. For instances 8, 9 and 10, the list algorithm is markedly better than the practical solution. The gap between the solutions was low and never exceeded 5.41%.

TABLE 1. A comparison of the GA-based approach, the list algorithm and the practical case.

Day	Number of patients		GA based approach		List algorithm		Practical		Number of
	Scheduled patients	Unscheduled patients	$\overline{W}$(min)	Medical staff's workload (%)	$\overline{W}$(min)	Medical staff's workload (%)	$\overline{W}$(min)	Medical staff's workload (%)	no-wait consultations (%)
1	8	48	205.3	80.2	242.2	82	245.72	80	0.55
2	17	50	212.6	71.3	216.4	73.5	218.47	72	0.5
3	20	44	230.5	67	267.8	65	266.39	65	0.48
4	28	31	298.5	88.2	358.8	87.2	359.08	85	0.58
5	6	58	197.5	80.5	215.2	79	215.65	78	0.52
6	12	105	290.2	88.3	303.4	90.2	305.05	89	0.50
7	12	70	222.6	75	230.2	76	231.04	76	0.64
8	14	38	223.5	85	265.4	90	270.18	90	0.68
9	10	29	278.4	93.2	296.2	92	305.26	92	0.58
10	18	24	187.6	83.2	198.5	82.5	209.63	81	0.78

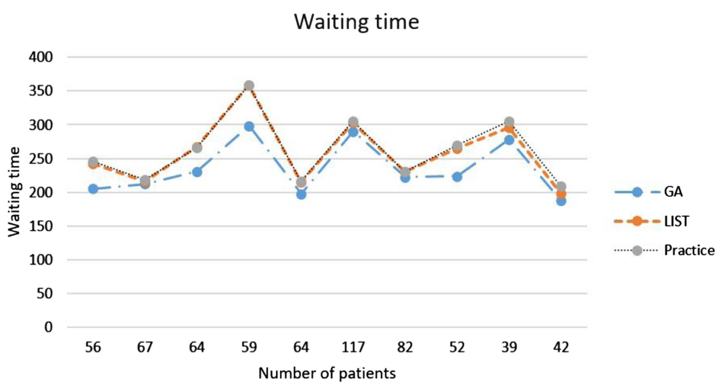

FIGURE 20. The waiting time as a function of the number of patients.

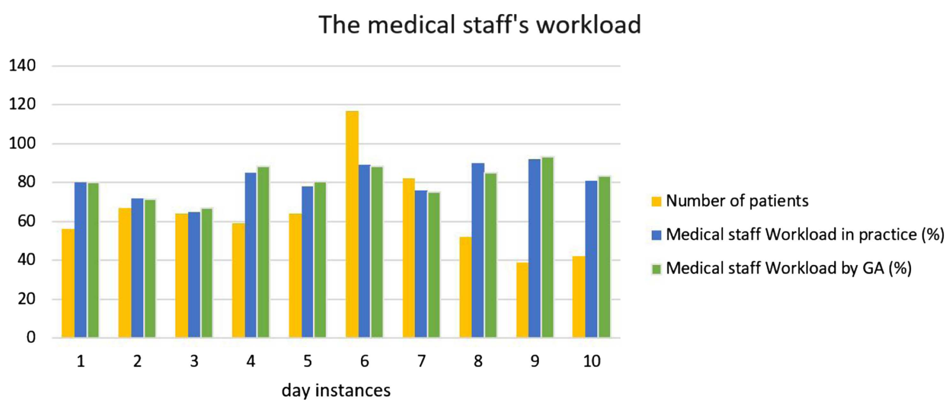

FIGURE 21. The medical staff's workload on day instances.

Table 2. Computation time.

Day instance	GA-based approach		List algorithm	
	$\overline{W}$(min)	Computation time (s)	$\overline{W}$(min)	Computation time
1	205.3	9.3	242.2	10.2
2	212.6	10.5	216.4	11.4
3	230.5	77.9	267.8	65.7
4	298.5	120.6	358.8	119.3
5	197.5	14.2	215.2	15.2
6	290.2	120.2	303.4	140.2
7	222.6	80.9	230.2	76.9
8	223.5	92.6	265.4	146.7
9	278.4	70.2	296.2	89.3
10	187.6	9.6	198.5	10.4

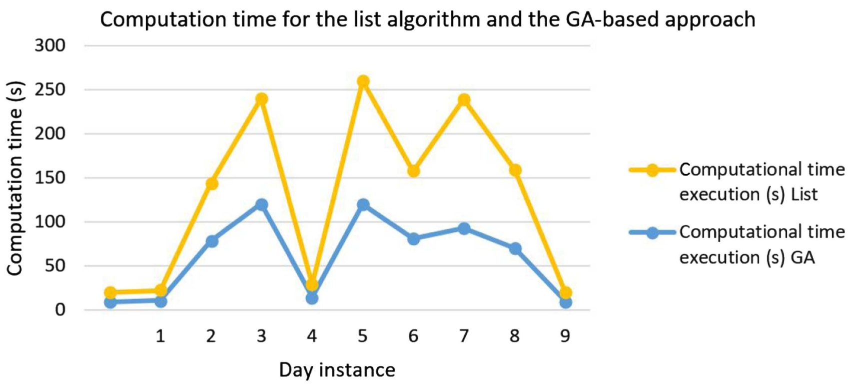

Figure 22. Execution times using the list algorithm and the GA-based approach.

In fact, the list algorithm uses dynamic priority rules to schedule the patients. These rules depend on the care tasks that have yet to be performed. As the tasks do not have the same care pathway, the waiting time can be reduced for some patients (for the same scenario), while the workload of the medical staff remains the same.

Table 2 shows the computation times for the list algorithm and the GA-based approach. In order to avoid the blind aspect of the genetic operators, we designed a controlled genetic crossover and mutation operators for the cubic representation of the chromosome.

Furthermore, we integrated the solutions found by the list algorithm in the initial population into the GA, in order to accelerate convergence on the best solution. As can be seen from Table 2, the GA-based approach performs far better than the list algorithm in terms of the computation time and performs better than both the current system and the list algorithm in terms of the waiting time. The computation time are compared in Figure 22. The GA-based approach's computation time is significantly shorter than that of the other methods.

In Table 2, the difference in execution time between the GA and the list algorithm is less than one minute. If this time difference is sufficient for the GA to generate high-quality solutions minimizing the patients' waiting time, then the solutions are relevant for clinical practice.

As discussed above, the GA-based method can address real-life scheduling problems in EDs, so that the patients' waiting time can be minimized as a function on the urgency of the required treatment. The second phase of our scheduling method reschedules the medical staff's tasks whenever a new patient arrives in the ED; the goal is to minimize the patients' waiting times by optimizing the use of resources (medical staff members) and ensuring that a patient with a more severe condition is prioritized. Hence, to better address scheduling in real-world EDs, the GA-based scheduling approach requires a reliable information system and an adequate amount of training for scheduling staff.

In addition, we have already studied and developed a multi agent system to model the communication and the interaction between the different medical staff member and the software agents [21]. The scheduled and re-scheduled approaches proposed in this paper are integrated in the agent behaviour in order to communicate to the medical staff member, the new care tasks to realize.

7. Conclusion

In the present study, we developed an innovative GA-based approach for scheduling both scheduled and unscheduled patients in an ED. The GA-based method assigns a theoretical consultation time to each patient on arrival. The goal of patient scheduling is to minimize the total waiting time and the overall cost. The GA-based approach grant a higher priority to the most urgent patients, while optimizing the medical staff's workload. The GA's performance has been enhanced by the incorporation of a cubic chromosome representation with novel, controlled genetic operators. In order to demonstrate the superiority of our approach, we applied it to a real ED. Simulations revealed that the GA-based approach improves the performance of patient scheduling in the ED and makes efficient use of the available resources. The computational results of our approach exceeded those of the practical approach. In the future, we intend to improve our approach to multiskilled healthcare task scheduling in the ED by combining a GA with multi-agent systems.

Acknowledgements. This work was supported and funded by the Federative Research Structures Technologies for Health and Drugs (SFR-TSM): http://sfr-tsm.ec-lille.fr/. The three components: CRIStAL CNRS UMR 9189 (http://www.cristal.univ-lille.fr/), the EA2694 of the Public Health Laboratory of Lille University (http://ea2694.univ-lille2.fr) and the adult emergency department of CHRU de Lille are full partners of the SFR-TSM.

References

[1] F. Flessa, Where efficiency saves lives: a linear programme for the optimal allocation of health care resources in developing countries. *Health Care Manage. Sci.* **3** (2000) 249–267.

[2] A.H. van Zon and G.J. Kommer, Patient flows and optimal health-care resource allocation at the macro-level: a dynamic linear programming approach. *Health Care Manage. Sci.* **2** (1999) 87–96.

[3] D. Baubeau, A. Deville and M Joubert, Les passages aux urgences de 1990 1998: une demande croissante de soins non programmés. Études et résultats N° 72 (2000) 1–8.

[4] M.A. Bjorkgren, U. Hakkinen and M. Linna, Measuring efficiency of long-term care units in Finland. *Health Care Manage. Sci.* **4** (2001) 193–200.

[5] S.H. Jacobson, E.C. Sewell, E. Deuson and B.G. Weniger, An integer programming model for vaccine procurement and delivery for childhood immunization: a pilot study. *Health Care Manage. Sci.* **2** (1999) 1–9.

[6] M.J. Ferreira de Oliveira, 3D Visual simulation platform for the project of a new hospital facility. In: Monitoring, Evaluating, Planning Health Services. *Proceedings of the 24th meeting of the ORAHS EURO-WG, World Scientific, Singapore* (1999) 82–95.

[7] J.M.H. Vissers, J.D. Van Der Bij and R.J. Kusters, Towards decision support for waiting lists: an operations management view. *Health Care Manage. Sci.* **4** (2001) 133–142.

[8] D.E. Burke and N. Menachemi, Opening the black box: measuring hospital information technology capability. *Health care Manage. Rev.* **29** (2004) 210–217.

[9] J. Belin, *Exact and heuristic methodologies for scheduling in hospitals: problems, formulations and algorithms*. Ph.D. thesis, Faculty of Business and Economics, Katholieke Universiteit Leuven (2006) 279.

[10] A. Rais and A. Viana, Operations research in healthcare: a survey. *Int. Trans. Oper. Res.* **18** (2011) 1–31.

[11] D. Lee, H. Shin and B.K. Choi, Mediator approach to direct workow simulation. *Simul. Model. Pract. Theory* **18** (2010) 650–662.

[12] J. Patrick, M.L. Puterman and M. Queyranne, Dynamic multipriority patient scheduling for a diagnostic resource. *Oper. Res.* **56** (2008) 1507–1525.

[13] E. Hans, G. Wullink, M.V. Houdenhoven and G. Kazemier, Robust surgery loading. *Eur. J. Oper. Res.* **185** (2008) 1038–1050.

[14] B. Jaumard, F. Semet and T. Vovor, A generalised linear programming model for nurse scheduling. *Eur. J. Oper. Res.* **107** (1998) 1–18.

[15] H.H. Millar and M. Kiragu, Cyclic and non-cyclic scheduling of 12 h shift nurses by network programming. *Eur. J. Oper. Res.* **104** (1998) 582–592.

[16] I. Blhliger, Modeling staff scheduling problems. A tutorial. *Eur. J. Oper. Res.* **158** (2004) 533–542.

[17] A.T. Ernst, H. Jiang, M. Krishnamoorthy and D. Sier, Staff scheduling and rostering: a review of applications, methods and models. *Eur. J. Oper. Res.* **153** (2004) 3–27.

[18] A.A. Musa and U. Saxena, Scheduling nurses using goal-programming techniques. *IIE Trans.* **16** (1984) 216–221.

[19] J.L. Arthur and A. Ravindran, SMultiple objective nurse scheduling. *AIIE Trans.* **13** (1981) 55–60.

[20] J.F. Bard and H.W Purnomo, Cyclic preference scheduling of nurses using a Lagrangian-based heuristic. *J. Schedul.* **10** (2007) 5–23.

[21] S. Ben Othman, H. Zgaya, S. Hammadi, A. Quilliot, A. Martinot and J.M. Renard, Agents endowed with uncertainty management behaviors to solve a multiskill healthcare task scheduling. *J. Biomed. Inf.* **64** (2016) 25–43.