APPOINTMENT SCHEDULING FOR A COMPUTED TOMOGRAPHY FACILITY FOR DIFFERENT PATIENT CLASSES USING SIMULATION

F. Boenzi[(a)], G. Mummolo[(a)], J.E. Rooda[(b)]

[(a)]DM3 - Facoltà di Ingegneria del Politecnico di Bari (Italy)
[(b)]Department of Mechanical Engineering - Eindhoven University of Technology (The Netherlands)

[(a)]boenzi@poliba.it, [(a)]mummolo@poliba.it, [(b)]j.e.rooda@tue.nl

ABSTRACT
In the present paper, a discrete event simulation model of a CT facility within a hospital is presented. The examination facility has to serve different patient classes with different priorities. At the strategic level, outpatient daily access is filtered out by means of the adopted appointment schedule (AS) scheme, whereas, at the tactical level, the decision about which patient to examine is taken by the established priority rule. Being the two levels inter-related, a comprehensive model of the examination process can help analyzing the different patient flows from a global point of view, taking into consideration equipment utilization and patient service performance, both in terms of waiting time at the facility and appointment interval for outpatients. Despite the model has been developed for a specific case-study, it is flexible and different data and settings could be easily implemented. Furthermore, some general considerations are drawn.

Keywords: simulation, health care, outpatient scheduling, CT examinations

1. INTRODUCTION

Complex diagnostic services, as in the case of CT (Computer Tomography) scans or MRI (Magnetic Resonance Imaging), require expensive equipment and very specialized human resources, making their full utilization an unavoidable necessity. As a order of magnitude, the cost of an MRI machine ranges between 1 and 2 million euros, depending on the magnetic field intensity, along with huge costs for building and preparing the space it will occupy. The cost of CT equipment is similar, essentially depending on the number of slices the machine is capable of producing for image computation. Generally, in hospitals, these types of resources are utilized for serving at least both classes of patients: inpatients (patients at the hospital wards) and outpatients, so that "customers" compete for accessing them in short periods of time. The idea behind is that making the resource shared is beneficial for reducing its idle time and achieving better utilization. In outpatient clinics, managers have looked to the popular policy of "Open Access" ("do today's demand today") as a solution for avoiding wasted capacity due to no-shows. Alternative booking techniques, based on short booking window and on the optimal policy from a Markov Decision Process, can perform even better in terms of smoothing out the demand and reducing peak work-load, as illustrated by (Patrick 2012). In hospitals, being only outpatient access planned in advance, in the Appointment Schedule (AS) definition phase, i.e. defining the number of service slots per time session and number of appointments at the beginning of each slot, the scheduler has to take into account allowance for the random (internal) demand component. In daily operations, priority rules are usually implemented (i.e. decision about which class of patient should be served first when both, random and planned demand, are present at the facility). When the diagnostic service facility is also open to patients from the Emergency Department (ED), they generally must be served as soon as possible, possibly on their arrival, unless the resource is already busy. Similarly, priority of outpatients over inpatients is justified by the simple consideration that, in any case, inpatients have to wait in their wards, whereas, for outpatients, excessive waiting time determines overall negative service perception, as pointed out by Sickinger and Kolisch (2009). Occasionally, outpatient prolonged waiting time could lead to over timing the appointment admittance time span and to service denial or incurring additional costs. On the other hand, it should also be noted that too long inpatient waiting time, due to examination postponing, could lead to un-necessary longer stays in hospitals and increasing cost for bed occupancy (Green et al. 2006).

2. LITERATURE SURVEY

The first research papers on outpatient appointment schedule date back to the '50s (Bailey and Welch 1952). Since then, many authors have dealt with this subject in many different settings; an overview can be found in Cayirli and Veral (2003). More specifically, as regards to a CT scan facility, open to the above mentioned patient flows, Kolisch and Sickinger (2008) propose a mathematical model, involving a Markov Decision Process. In their model, it's possible to distinguish two levels: an upper level which we could consider as "strategic", regarding the outpatient Appointment

Schedule (AS) and a lower level, which we could regard as "tactical", involving decisions to undertake on the run at a discrete number of time-points. These decision levels were yet characterized by (Green et al. 2006), who proposed a finite-horizon dynamic control model for the two capacity management tasks (appointment scheduling and real-time capacity allocation), highlighting their interrelation, and applied it to an MRI hospital facility. At the decision level, the choice at the beginning of each time-slot is about serving waiting scheduled outpatients and/or inpatients, stated that if an emergency patient were generated in the previous time-slot, he must always be served. Stochasticity in the system is introduced by diverse probabilities associated to the different classes of patients: no-shows for outpatients and random arrival of an inpatient and/or an ED patient in the previous slot (the limitation to one is among the model assumptions). Instead, service time, equal to the slot time, is assumed to be deterministic. The underlying Markov Decision Process is determined by the established decision policy, which should aim at maximizing an expected total reward function. In general the latter consists of the sum, over the entire appointment time span, of a linear combination of served inpatients' and outpatients' reward, waiting costs and penalty costs for service denial at the end of the shift (ED patients excluded). The optimal policy can be found by the "backward induction algorithm" (Puterman 2005). However, since the authors observe that "the acceptance for computer-based decision rules in medical environments is low", they investigate the performance of "simple decision rules which can be applied manually". The examined rules are LCA (Linear Capacity Allocation) introduced by (Green et al. 2006), FCFS (First-Come-First-Served) and Random, which are compared in combinations of three different scenarios (generated varying the problem parameters) and three different AS schemes from literature (2BEG, Block and Threshold). Even though the LCA rule performs better, the authors underline the importance of the fairness of the rule for the reduction of the perceived waiting time by the patients. FCFS, contrary to LCA, is a very simple and fair decision rule regarding its inter-class selection behavior, so that it can be considered as a "fair heuristic".

In successive work (Sickinger and Kolisch 2009), the authors focus their attention on the "strategic" level, on the basis of the results obtained optimally solving the associated stochastic dynamic program. They carry out an empirical study on a two CT scans examination service with 8 slot available, under three increasing system utilization levels (number of scheduled patients equal to 4,8,12 respectively). They compare the values of the objective function resulting from a proposed Generalized Bailey-Welch (GBW) schedule, a Neighborhood Search (NS) heuristic and the optimal scheduled (obtained by full enumeration). The authors find that the GBW rule and the NS heuristic generate optimal or near-optimal solutions for low and medium utilization, whereas, for high utilization, the GBW rule, contrary to the NS heuristic, does not provide optimal solutions any more. They also analyze the impact of the function parameters on the results (adopting NS schedule as reference) and highlight the gap with GBW. In particular, in cost structures characterized by relevant penalties in case of denial of service for outpatients, the GBW rule becomes the best choice. Anyway, it's observable that, in case of high utilization, nothing assures that all the outpatients will be served within the service time period and for this reason the authors themselves recommend the calculation of a scheduled optimal number.

The problem arising from the co-existence of random urgent patients aside the scheduled ones is often faced by leaving some slots "open" to accommodate urgency. Taking this into account and treating the position of a couple of open slots as a decision variable, Klassen and Rohleder (1996) carry out a full factorial ANOVA analysis on a simulation model (in SIMAN IV simulation language) of a family medicine clinic. The authors adopt 10 known pre-defined AS rules as second decision variable and take into account two environment factors involving probabilistic considerations (3 possible mean values and 5 levels of percentages of clients with "low" standard deviation of lognormal service time distributions). They analyze the system performance in terms of WIT (sum of expected total clients' waiting time and expected total server idle time costs) and other secondary measures. In their model there is not a decision process and a decision policy (tactical level) because of the presence of only two patient classes (scheduled and urgent calls) and of the assumption that the clinic could accept at most two urgent patients per session (number equal to the open slots). According to their findings, simple rules like 2BEG (Bailey's rule, with 2 clients in the first slot) and 4BEG (4 clients at the beginning) perform worse than rules which take into account client's classification into two possible service time variance groups (low and high), when assigning them to the slots. In particular, the LVBEG rule (low variance clients at the beginning) proves to be the best rule, also for its equanimity in balancing clients' time and server time. However, its practical implementation requires the availability, at the clinic, of recorded information about clients' past service times and more attention by the receptionist. The simplest rule FCFA (first-call-first-appointment) is, in all the examined cases, in the group of the best rules and should be preferred for its simplicity. In successive work (Rohleder and Klassen 2002), the authors modify and expand their model, addressing the issues of rolling-appointment horizon and variable demand load, using simulation. They carry out a full factorial analysis considering six demand patterns, six overloading rules and three rule delay periods. Results are summarized in a matrix that outlines good managerial choices for each scenario.

Kaandorop and Koole (2007) consider the problem of optimal outpatient appointment schedule, in which only this class of patients is present. The objective function to be minimized is a linear combination of mean waiting time, mean doctor's idle time and mean tardiness (time exceeding the given session period). They prove that the proposed local search algorithm converges to the global optimum under the problem stated conditions; moreover, they highlight that "for certain parameters value the Bailey-Welch rule is indeed optimal". It should be noted that service time is assumed to be exponentially distributed, which is quite uncommon in healthcare services.

As seen in literature, several cases of appointment service systems, accounting or not for additional random demand, exist, which makes very difficult drawing out general rules, easily understandable by healthcare operators. Simulation in this field of study is deemed to be a very flexible tool, especially for the capability of including particular singular features of real-case systems, typical of the healthcare sector. In the present paper, a discrete event simulation model of one CT examination server within a urban hospital is considered.

The remainder of the work is organized as follows: in Section 3, the process of CT examination demand generation is illustrated, in Section 4 the model is presented, in Section 5 simulation results are illustrated and commented, finally Section 6 presents the conclusions.

3. CT PROCESS DESCRIPTION

The CT scan is located at the radiology department of a large community hospital in Basilicata region (Italy); a schematic layout of the department is reported in Figure 1. In the blank rooms, other type of equipment (traditional x-ray technique machines and ultra-sound scanners) or technical rooms are present, and, on the left part of the figure, the proximity with the ED department, located on the same floor, is depicted.

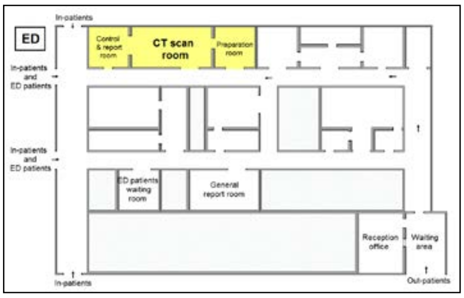


Figure 1: Radiology Department Layout

The diagnostic examination facility has to serve different patient classes with different priorities; specifically, in descending priority order: a) ED patients, b) urgent in-patients, c) outpatients and d) non-urgent in-patients; each of them, except for scheduled patients, is characterized by a random arrival process. These patients flows are coexisting during the reception opening time (from 8 a.m. to 8 p.m., from Monday to Friday), whereas access for urgent cases of type a) and b) is possible at any time. On their arrival, patients c) and d) are checked-in and their data put in the Radiology department Information system (RIS) by reception staff, whereas for patients a) and b), the examination requests are generally sent in electronic form, which makes check-in possible also when the reception is closed. The list of waiting patients builds up the work-list for the technologist in the CT control room. Once registered, outpatients, stationing in a dedicated area, wait for call and generally reach the examination room autonomously, following signposts; sometimes, they are accompanied by relatives or by department attendants. In-patients and ED patients are always accompanied by attendants or by one of the technologists themselves, because generally they are on wheelchairs or wheel-beds. A relevant difference between outpatients and other types of patients, in addition to diverse priority, is that, for CT scan examination which require intravenous administration of a contrast medium by a nurse, patients have to pass first through a preparation room, whereas hospitalized and ED patients generally can access the CT room directly. In the remainder of the paper, the preparation phase for outpatients has not been taken into account for total process time quantification (as if it were part of waiting time) because the two processes (examination and preparation), taking place in two different rooms, can be regarded as independent and parallel processes (i.e. a patient can undergo a CT scan, while the next one is being prepared). The two activities can still, in rare cases, overlap. This happens when preparation time takes too long (especially with elderly persons) and, meanwhile, the CT room has turned free or in the case when, at the end of the examination, a patient pleads indisposition and has to be monitored by the nurse, who, therefore, can't take care of the next one, causing delays.

As regards the process of examination generation, for inpatients (urgent and not), examination requests are generated by doctors in the various hospital departments and ED patients requests are generated, when needed, after their arrival at the ED. For out-patients, the requests are generated by family or speciality doctors. After that, possible points of access to health-care services are by phone-call to a unified regional call-centre or by taking the paper request to a "CUP" (unified booking centre) office. In any case, information about the next available appointment slot in a health-care structure able to dispense the requested service are shared in real-time. Requests are added so forth and build up waiting lists. For outpatients, a random generation process is adopted at the origin of the demand, whereas their daily access to the examination process is filtered out through the AS scheme established by the department director. Inpatient and ED patient access requests can be considered random generated and flow to the daily work-list; anyway, at the "tactical" level, the decision about which patient to examine is set by the mentioned priority rule. The described process and its integration with the examination process are illustrated in Figure 2, in which the time-span within the two vertical dashed lines represents the time elapsed from the input of the request into the "CUP" system to the appointment day.

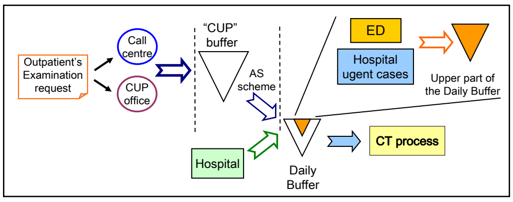


Figure 2: Examination Generation process and buffering

The aim of the present paper is analyzing the performance of the system altogether in terms of waiting time for the different patient classes and machine utilization by means of simulation. Differently from existing literature, which generally focuses on patient waiting time during the appointment session, for outpatients it's very important to know also the "long" waiting time (usually measured in weeks) to the appointment day. Moreover, according to recent regional legislation, outpatients have to be differentiated at the origin, in order to account for particular urgent cases and their waiting time must comply with specified limits. Three classes of outpatients have been characterized depending on this time limit: appointment within 10 days, denoted as "B"; within 30 days, denoted as "D"; without particular time constraint, denoted as "N". Correspondingly, three different booking lists have been set up at "CUP" and at the radiology department. Outpatient waiting time in the department on the appointment day could be considered a secondary measure of performance, whereas it remains very important for emergency cases. Of course, the two aspects are strongly inter-related because, at the strategic level (AS), scheduling a bigger number of outpatients can shorten their waiting queue and "long" waiting time, but, on the other hand, can congest the system, increase unacceptably waiting time for the random low-priority component and eventually lead to lateness of the examination session (overtime), post-poning or cancellation of scheduled patients. The simulation model aims at offering a global view of the system performance when the different patient flows are coexisting.

4. MODELLING

4.1. Patient data

The employed data come from two different sources: RIS data reporting the CT examinations performed at the radiology department for each patient class in the period October 2012 – May 2013 and data of examination requests (classified as outpatient N, D or B) arrived at the "CUP" office over the same period. For the codification of the various types of examinations into the model, a national codification system, set up in (VV.AA. 2006), has been partially used. The latter system, also for accounting purposes, consider a single body-part CT scan as a coded examination. In reality, the majority part of patients undergo some typical CT sequences (as in the case of chest-abdomen, brain-chest-abdomen or "total-body" scan); for these examinations, aggregating the sequence into a whole has led to create appropriate additional new codes. RIS data for outpatients have not been used, since their access to the service is regulated by means of scheduling. For inpatients and ED patients RIS data have been filtered with the check-in time, excluding the examinations not within the reception opening days and hours of the day. The resulting examination mix is reported in Figure 3. Average inter-arrival time and throughput values of examination requests (outpatients) and of patients (ED and inpatients) are reported in Table 1 (averaged over the net reception worked hours in the observed period).

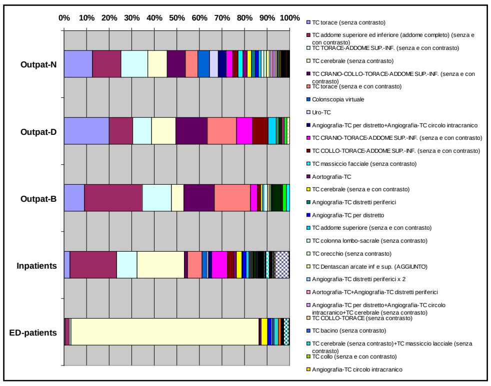


Figure 3: Patient Examination Mix

Inpatients and urgent inpatients have not been differentiated with regard to the examination mix (i.e. any type of examination could be requested with urgency); their relative proportions in current data are 78.8% and 21.2% respectively.

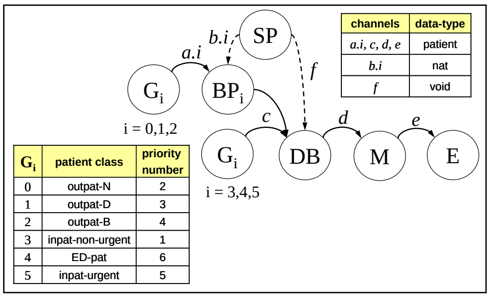
Table 1: Patient Flow Data

	Outp-N	Outp-D	Outp-B
t_a (min/pat)	105.67	1276.55	798.99
Thrput (pat/h)	0.57	0.047	0.075
	Inpat	**ED-pat**	**Urg-Inpat**
t_a (min/pat)	93.80	175.45	348.15
Thrput (pat/h)	0.64	0.34	0.17

4.2. Discrete-event simulation model

The model is a stochastic discrete-event simulation model built by means of the process algebra language Chi 1.0 (Hofkamp and Rooda 2007). With Chi, a symbolic representation of a system is translated into a model, consisting of parallel processes, which communicate, in a synchronous way, one with each other via channels. Data exchanged on channels can represent physical entities (e.g. patients) or information contents (e.g. signals, data, etc.). Among the principal advantages of the language are its capability of preserving formalism and it's ease of comprehension and transparency even to non-expert people. Process based language Chi has been used in manufacturing modeling as well as in the health-care sector. For example in (Jansen et al. 2012), an aggregate model of an MRI department is developed employing effective process time (EPT) concepts.

In the present paper, instead of aggregating examination data, their differentiation is maintained on the basis of the examination types because of the necessity, for future model developments, of characterizing some particular types, for which separated outpatient booking lists are currently adopted in the department (in addition to the three mentioned waiting lists). In Figure 4, the Chi model of the CT diagnostic service is depicted, along with two tables: one summarizing the correspondence between generator processes and classes of patients and priority rank; the other reporting channels and exchanged types of data. The arrowed lines represent the patient flows and the dashed lines represent exchanged data.



channels	data-type
a.i, c, d, e	patient
b.i	nat
f	void

G_i	patient class	priority number
0	outpat-N	2
1	outpat-D	3
2	outpat-B	4
3	inpat-non-urgent	1
4	ED-pat	6
5	inpat-urgent	5

Figure 4: Model and Processes

The model consists of six patient generator processes G_i, three buffer processes BP_i, the scheduling process *SP*, the daily buffer process *DB*, the examination process *M* and the exit process *E*.

The generators G_i $(i=0,1,2)$ represent the arrival processes of examination requests arriving at "CUP", for the three types of waiting queues (N, D and B). Examination requests (patients in the model) are then put in queue in the buffer processes BP_i, as happens at the booking office. Patient data-type contains information regarding the code of the requested examination, chosen randomly on the basis of the found mix (see previous section) and the assigned priority. The arrival process is modeled as a homogeneous Poisson process (HPP), with mean inter-arrival time according to Table 1; this assumption of the model comes from the hypothesis that the served population remains constant and there is not a seasonal component. The generators G_i $(i=3,4,5)$ represent generators of the random component of the demand, represented by inpatients, ED patients and urgent inpatients. For each of these generators also, a HPP is assumed. This hypothesis is indeed an approximation, because variability in the course of the day and eventually from a day of the week to another day occurs. Generators of inpatients and urgent inpatients utilize the same examination mix, but have different mean inter-arrival times, respecting the proportions found in current data.

Process *SP* represents the scheduling process for authorizing outpatient buffers BP_i to release fixed numbers of patients to process *DB* in the course of the day, according to the established AS. This information is communicated to each BP_i via channels *b.i* of integer numbers. The probability of no-shows is not included in the current model and therefore the scheduled number always corresponds to the number of released patients, unless the buffer becomes empty. *SP* assumes also the function of a cyclic clock in the model, because it takes care of the passing time at disposal to complete the daily schedule. At the end of the day, a signal is sent via channel *f* to process *DB*. Generators G_i $(i=3,4,5)$ are not "filtered" by a scheduling process, but are directly linked to the *DB* process.

Process *DB* simulates the daily buffer of patients to be examined each day, sorted according to their priority number. This buffer is indirectly related to the *SP* process, since its filling up follows a cyclic behavior, on the basis of the AS; in addition, it is also subject to "disturbances" due to the unplanned arrival (according to HPPs) of the other types of patients. At the end of the day, a signal is received by *SP* via channel *f* and the remaining numbers of patients are monitored, in order to calculate average values. The most critical event which could happen is that in this buffer there are still outpatients; this means that the random arrival of patients with greater priority has prevailed, impeding the completion of the daily schedule. According to the hospital staff, this event is rarely possible, but, in any case, scheduled outpatients must be examined that day. The same could happen for inpatients, who are normally examined only during the reception opening time. However, to a limited extent, this is not a serious problem, considering that inpatients are in wards at the hospital. Patients remaining in the daily work-list are normally examined in overtime and this doesn't have consequences on the next appointment sessions. Therefore, in the model, buffer *DB* is emptied at the end of each day.

Process *M* represents the examination process, in terms of CT room occupation time and is modeled as a time delay for the patient. For some examination codes, collection of empirical data has led to determine maximum likelihood estimated parameters of *gamma* probability distributions, with acceptable results of goodness-of-fit tests at the significance level 0.05 (Boenzi et al. 2012). For other codes, for which commonly used PDFs don't fit satisfactorily, empirical distributions are implemented. For all the other CT examination codes with a scarce number of observations, process mean values μ are assumed on the basis of technologists' esteems. *Gamma* PDFs are then adopted, assuming a worst-case approach with regard to the highest variability for process time. Among the observed examination types, the maximum found coefficient of variation $c=\sigma/\mu=\sqrt{1/\alpha}=0.55$ is selected and it is employed to calculate a common shape parameter $\alpha=3.3$. Then, different scale parameters β are calculated as μ/α.

In the model, at last, patients are sent to the exit process *E*, in which indicators regarding patient flow (data deriving from time-stamps at each stage) are calculated.

5. SIMULATION RESULTS

All the following simulation results are calculated as average values of five independent simulation runs, time-terminated at 500.000 minutes, time limit which corresponds to approximately 3 years of continuous reception time operation. In all the simulations, the system starts in empty conditions, i.e. preceding patient queues are disregarded.

5.1. Current system

The AS currently adopted at the radiology department is depicted in Figure 5.

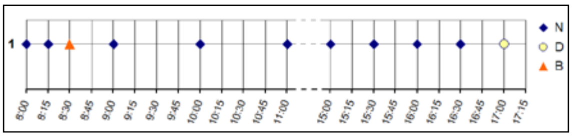


Figure 5: Current AS for CT examinations

The daily schedule (Monday – Friday) follows the reported scheme, except for Wednesday, when only the first part of the schedule (till the dashed lines) is adopted, because in the afternoon a special health monitor program is in place. Consequently, in the simulation model, a generic cycle of the scheduling process *SP* (see Section 4.2) can be either 720 minutes long for a full day (8 – 20) and 420 minutes for half a day (8 – 15). A complete weekly cycle comprises four full days and one half-day, i.e. 3,300 minutes in total. It can be observed that, in the first part of the morning shift, an approach similar to the Bailey-Welch rule is adopted, placing three patients, with only 15 minutes time-elapse, in the first operative hour. This can account for possible no-shows (not modeled), but inevitably increases patient waiting time. It can also be observed that, in the morning, the examination slot duration is set to one hour, whereas in the afternoon shift it is reduced to half an hour. This scheduling decision is due to the assumption that the random demand, especially by ED patients, is more intense in the morning. The model presented in this paper, however, doesn't take this aspect into consideration. Outpatient admission is distinguished substantially between two blocks and the last appointment in the morning is at 11 a.m. The clearance time between the two blocks is devised, according to the department staff, just in order to process the main part of inpatient examinations. Comparing the weekly accepted number of outpatients with the observed average weekly demand (Table 2), it's possible to observe that, with the current schedule, some extra capacity is employed. This is clearly an effective strategy to reduce waiting queues.

Table 2: Outpatient Examination Demand

Average weekly demand (over 35 weeks)		**Current weekly schedule**	**Extra-capacity**
Outp-N	31.23	41	1.31
Outp-D	2.59	4	1.54
Outp-B	4.13	5	1.21

Simulation results are reported in Tables 3 and 4. The Appointment Interval (A.I.) represents the time-span for obtaining an appointment at the facility (measured in effective days, including Saturday and Sunday) and waiting time (w-time) values, instead, refer to waiting at the department. Utilization of the system is 0.498.

Table 3: Outpatient Flow Performance

	Outp-N	**Outp-D**	**Outp-B**
Avg Thrput (pat/h)	0.572	0.048	0.077
Avg A.I. (days)	**0.31**	**1.79**	**2.96**
Min A.I. (days)	0.0001	0.0094	0.0138
Max A.I. (days)	2.06	10.39	15.61
Avg w-time (min)	**9.50**	**8.08**	**12.01**
Min w-time (min)	0	0	0
Max w-time (min)	182.11	93.97	114.10

Table 4: Hospital Patient Flow Performance

	Inpat	**ED-pat**	**Urg-Inpat**
Avg Thrput (pat/h)	0.63	0.34	0.17
Avg w-time (min)	**17.51**	**6.64**	**7.14**
Min w-time (min)	0	0	0
Max w-time (min)	273.75	116.33	116.14

Similar results are obtained assuming that, in the model, the examination process start is systematically 40 minutes delayed, from 8 a.m. to 8.40 a.m. This is a pessimistic but realistic assumption, because, occasionally, due to organizational reasons and limitation of personnel resources, it could happen that the CT facility is not fully operational at the start of the shift. Waiting time results are reported in Table 5, in which an increase for low-priority patients can be observed and the minimum waiting time for B-type outpatients is ten minutes, as expected. The above assumption will be held also in the following.

Table 5: Patient Flow Performance at the department with delayed CT room availability

	Outp-N	**Outp-D**	**Outp-B**
Avg w-time (min)	**29.29**	**7.91**	**15.28**
Min w-time (min)	0	0	**10**
Max w-time (min)	219.02	77.36	101.78
	Inpat	**ED-pat**	**Urg-Inpat**
Avg w-time (min)	**28.05**	**7.60**	**8.71**
Min w-time (min)	0	0	0
Max w-time (min)	297.88	112.24	137.17

Since the system starts in empty conditions and an extra-capacity is put at disposal, outpatients are characterized by very brief average appointment time-spans, also with regard to maximum values. In these conditions, if the examination demand remained the same, waiting lists would be progressively emptied and the strategic objectives, regarding reduced appointment time-span, met. Over-sizing of the current AS is also testified by Table 6, reporting the time average buffer size and the average number, at the end of a generic cycle, of additional patients which could have entered the system if the buffer could have released the requested number.

Table 6: Outpatient Buffers and additional potential patients

	Outp-N	Outp-D	Outp-B
Avg Buffer size (patients)	1.83	0.93	2.17
Avg additional patients (patients/cycle)	1.95	0.28	0.17

5.2. New schedule analysis

In this and in the following sections, some scenario hypothesis are formulated and simulation results are illustrated and compared. The first hypothesis to be investigated assumes an increment of outpatients exactly equal to the current weekly scheduled number. Results are reported in Table 7. Utilization of the CT room (0.557) is only marginally improved, but the average A.I. for outpatients of type B (and, much worse, its maximum value) is above the acceptable limit and tends to increase in time, i.e. the system doesn't reach a stationary condition. This trend is also confirmed by the buffer time-history reported in Figure 6 (reporting the results of five simulation runs) and is due to the variability of the arrival process.

Table 7: Outpatient Flow Performance under the hypothesized increment and the current AS

	Outp-N	Outp-D	Outp-B
Avg Thrput (pat/h)	0.737	0.0701	0.0874
Avg A.I. (days)	**4.08**	**28.79**	**21.78**
Min A.I. (days)	0.0042	2.20	0.77
Max A.I. (days)	**15.35**	**57.98**	**56.47**
Avg w-time (min)	30.09	9.41	14.92
Min w-time (min)	0	0	**10**
Max w-time (min)	219.71	111.16	93.09

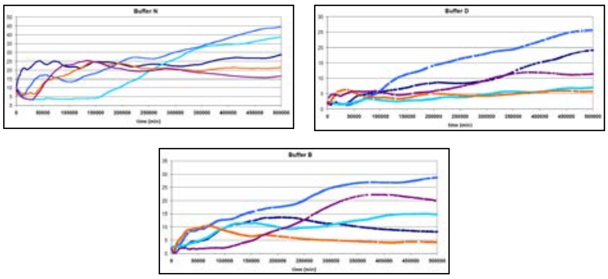


Figure 6: Outpatient Buffers time-history

In order to accommodate the increased examination demand, a new AS, reported in Figure 7, is proposed and tested. On its basis, the weekly admitted outpatient numbers increase according to Table 8.

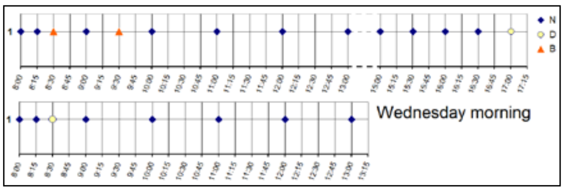


Figure 7: New Proposed AS

Table 8: Outpatient Examination Demand Increment

Hypothesized outpatient weekly demand		New weekly schedule	Extra-capacity
Outp-N	41	51	1.24
Outp-D	4	5	1.25
Outp-B	5	8	1.6

Simulation results are reported in Table 9, showing that offering extra-capacity assures achieving the stated appointment-time objectives, also with regard to maximum waiting time.

Table 9: Outpatient Flow Performance under the hypothesized increment and the new AS

	Outp-N	Outp-D	Outp-B
Avg Thrput (pat/h)	0.741	0.0747	0.0912
Avg A.I. (days)	**0.28**	**2.56**	**1.05**
Min A.I. (days)	0.000084	0.00926	0.00159
Max A.I. (days)	1.68	11.44	4.84
Avg w-time (min)	26.40	9.47	15.11
Min w-time (min)	0	0	0
Max w-time (min)	242.22	105.37	122.76

5.3. Hypothesized scenarios

Maintaining the illustrated hypothesis of increased outpatient demand and employing the newly proposed AS, additional "what-if" scenarios take into consideration the eventuality of rising up of the random urgent patient component (ED patients and urgent inpatients), who are assigned greater priority ranks with respect to outpatients. The scenarios comprise the following: a_1) increasing ED demand 50%; a_2) increasing ED demand 100% (doubling the current figure); b) changing the percentage of urgent inpatients from the current figure (21.2%) to three different levels: b_1) 50%, b_2) 63.3% and b_3) 75%, constant in time. The second level has been calculated in such a way that the summed throughputs of current ED patients and urgent inpatients is equal to the summed throughputs of current urgent inpatients and doubled ED patients, i.e. the urgent throughput is the same for a_2) and b_2). Scenarios a_1) and a_2), even if clearly over-estimated, could be the consequence, for example, of the closure of one or more neighboring EDs. As regards to scenarios b), they come from the consideration that, realistically, inpatient whole throughput can't increase, because it is linked to the hospital bed capacity. Instead, its urgent component could increase, considering that hospitalized patients, as society, are an aging population and that employing the form of urgent examination request could be increasingly utilized by doctors at wards, in order to shorten their dismissal. As also pointed out in Section 4.2, the major impact that urgent patients can have on the appointment schedule is not succeeding in examining all the planned outpatients. In order to monitor this, the remaining number of patients in the daily work-list, in the course of a simulation run, is summed up and average values over the total number of daily cycles are calculated. Therefore the average value can also be regarded as the probability of finding a certain type of patient remaining in the work-list, at the end of a generic daily schedule. Comparative results are reported in Figure 8. Urgent patients do not represent a serious concern because access for them is granted at any time: therefore they can always be present in the work-list. Instead, non-urgent inpatients should be preferentially completed during the reception opening time. It can be observed that in scenario a_2), due to their low priority and the increased number of high priority patients, inpatients have to wait and therefore it's more likely to find any of them in list at the end of a day. In scenarios b), this effect is mitigated, also compared to initial system conditions, because quantitatively their presence is reduced. For outpatients, the probability of over-timing is, in all the examined cases, very low. It can be observed that the impact on over-timing in scenario b_2) is more severe, compared to scenario a_2), even if the urgent throughput is the same. This can be explained considering the different examination mix of inpatients and ED patients and the greater variability for the first ones, with longer time examinations. In general, the impact of increasing urgent inpatient percentage is greater than increasing ED throughput, even if in the first case there is not a net increment of patients.

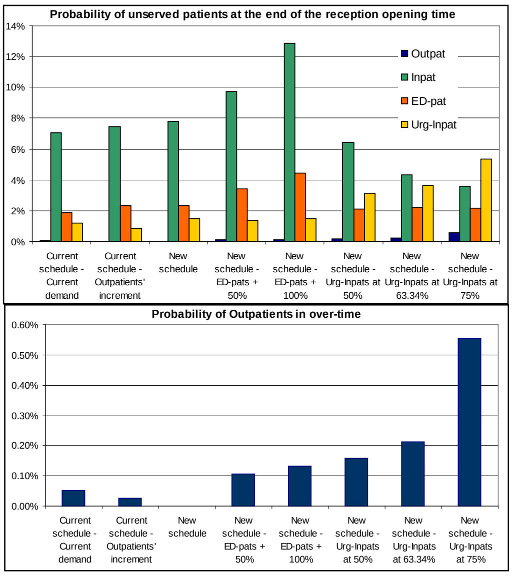
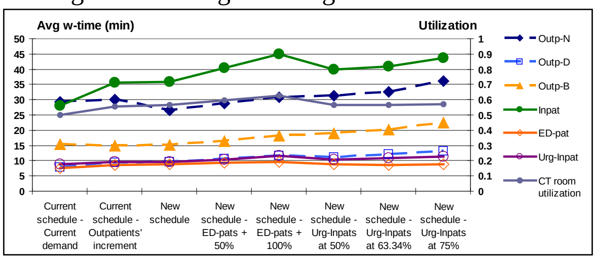


Figure 8: Average Number of patients remaining in the Daily Work-list per cycle

In Figure 9, average waiting time values at the department, along with CT room utilization, are summarized. It can be noticed that, in general, waiting time is acceptable, even though, especially for outpatients of type N, it could be improved changing the AS and taking into account, in simulation, the possibility of no-shows. Inpatients are, in all the examined cases, the most penalized service users and the impact of their lowest priority is particularly evident on the maximum waiting time (not reported). In turn, this could eventually lead to the decision of postponing them to the next daily cycle, incurring costs for additional hospitalization. In all cases, utilization is low, ranging from around 0.5 to 0.62.

Figure 9: Average Waiting Time and Utilization



6. CONCLUSIONS

In the present paper a discrete-event simulation model of a hospital CT facility has been presented. The aim of the model is giving a global view of the problem of coexisting admitted classes of patients, in terms of local performance (waiting time at the department for the different patient classes and CT equipment utilization) and long-term performance, represented by the average appointment interval for outpatients. The last is determined by the AS policy and, as illustrated by means of a case-study, the two aspects are inter-related. Current situation and some hypothesized scenarios, employing a different AS, have been illustrated and qualitatively compared. Even though the obtained simulation results refer to a particular case-study, two general recommendations can be drawn: 1) in schedule planning, setting up extra-capacity with respect to the current average external examination demand, instead of offering a capacity strictly equal to it, prevents the making up of appointment queues, because of the random nature of the process. Therefore, external demand should be periodically monitored by the “CUP” staff and eventually determine the AS redesign.

2) It should be avoided a too rigid application of the priority rule for non-urgent inpatients (lowest priority) in order to prevent examination postponing. Therefore an alert system for excessive waiting time should be implemented, permitting in some cases overriding the rule at the expense of outpatients. This could in turn cause the increase of average outpatient waiting time, but, as illustrated, over-timing is a rare event.

Future work comprises model validation, for which data collecting is in progress, and finding strategies for AS improvement.

REFERENCES

Bailey, N.T.J., Welch, J.D., 1952. Appointment systems in hospital outpatient departments. *Lancet* 259: 1105–1108.

Boenzi, F., Mummolo, G., Rooda, J.E., 2012. Analysis of a Diagnostic Radiology Department with different patient flows using different data sources. *Proceedings of the International Workshop on Applied Modelling and Simulation*, 114–124. September 24-25, 2012, Rome, Italy.

Cayirli, T., Veral, E., 2003. Outpatient scheduling in health care: a review of literature. *Prod Oper Manag* 12: 519–549.

Green, L.V., Savin, S.V., Wang, B., 2006. Managing patient service in a diagnostic medical facility. *Oper Res* 54(1): 11–25.

Hofkamp, A.T., Rooda, J.E., 2007. C*hi 1.0 Reference Manual*. Eindhoven University of Technology. Available from: http://seweb.se.wtb.tue.nl/sewiki/_media/chi/chi10_refman1689.pdf [Accessed 15 July 2013].

Jansen, F. J. A., Etman, L. F. P., Rooda, J. E., Adan, I. J. B. F., 2012. Aggregate simulation modeling of an MRI department using effective process times. *Proceedings of the 2012 Winter Simulation Conference*, 919–930. December 9-12, 2012, Berlin, Germany.

Kaandorop, G., Koole, G., 2007. Optimal outpatient appointment scheduling. *Health Care Manag Sci* 10: 217–229.
Klassen, K.J., Rohleder, T.R., 1996. Scheduling outpatient appointments in a dynamic environment. *Journal of Operations Management* 14: 83–101.
Kolisch, R., Sickinger, S., 2008. Providing radiology health care services to stochastic demand of different customer classes. *OR Spectrum* 30(2): 375–395.
Patrick, J., 2012. A Markov decision model for determining optimal outpatient scheduling. *Health Care Manag Sci* 15: 91–102.
Puterman, M.L., 2005. *Markov decision processes: discrete stochastic dynamic programming*. New York. Wiley.
Rohleder, T., Klassen, K.J., 2002. Rolling Horizon Appointment Scheduling: A Simulation Study. *Health Care Manag Sci* 5: 201–209.
Sickinger, S., Kolisch, R., 2009. The performance of a generalized Bailey–Welch rule for outpatient appointment scheduling under inpatient and emergency demand. *Health Care Manag Sci* 12: 408–419.
VV.AA., 2006. *Metodologia di determinazione dei volumi di attività e della produttività dei medici radiologi - Nomenclatore SIRM-SNR delle prestazioni radiologiche.* OMICRON Editrice Genova. Available from: http://www.sirm.org/index.php/documenti/doc_download/35-nomenclatore-e-calcolo-volumi-di-attivita [Accessed 15 July 2013].