Health Care Manag Sci
DOI 10.1007/s10729-011-9153-z

Managing magnetic resonance imaging machines: support tools for scheduling and planning

Adam P. Carpenter · Lawrence M. Leemis · Alan S. Papir · David J. Phillips · Grace S. Phillips

Received: 16 July 2010 / Accepted: 27 February 2011
© Springer Science+Business Media, LLC 2011

Abstract We devise models and algorithms to estimate the impact of current and future patient demand for examinations on Magnetic Resonance Imaging (MRI) machines at a hospital radiology department. Our work helps improve scheduling decisions and supports MRI machine personnel and equipment planning decisions. Of particular novelty is our use of scheduling algorithms to compute the competing objectives of maximizing examination throughput and patient-magnet utilization. Using our algorithms retrospectively can help (1) assess prior scheduling decisions, (2) identify potential areas of efficiency improvement and (3) identify difficult examination types. Using a year of patient data and several years of MRI utilization data, we construct a simulation model to forecast MRI machine demand under a variety of scenarios. Under our predicted demand model, the throughput calculated by our algorithms acts as an estimate of the overtime MRI time required, and thus, can be used to help predict the impact of different trends in examination demand and to support MRI machine staffing and equipment planning.

Keywords Simulation · Forecasting · Decision support systems · Radiology · Scheduling · Resource planning · Multicriteria optimization · Linear programming · Combinatorial optimization

This paper was supported in part by NSF grants DMS-0703532, DUE-0849955, and a NASA/VSGC New Investigator grant.

A. P. Carpenter · L. M. Leemis · A. S. Papir · D. J. Phillips (✉)
Mathematics Department, The College of William & Mary, Williamsburg, VA, USA
e-mail: phillips@math.wm.edu

A. P. Carpenter
e-mail: atom.carpenter@gmail.com

L. M. Leemis
e-mail: leemis@math.wm.edu

A. S. Papir
e-mail: apapir@gmail.com

G. S. Phillips
Seattle Chilldren's Hospital, Seattle, WA, USA
e-mail: grace.phillips@seattlechildrens.org

1 Introduction

In this paper, we devise models and algorithms to support decision makers at Seattle Children's Hospital (SCH) with the resource allocation and planning problems faced by the Radiology Department concerning MRI examinations. SCH serves pediatric inpatients and outpatients with over 50 specialty clinics. MRI is an imaging modality which uses magnetic fields to help diagnose a wide variety of conditions including tumors, infections, cardiac conditions and musculoskeletal disorders. MRI machines are differentiated by magnet strength, measured in Tesla, and we subsequently refer to MRI machines as *magnets*. SCH owns one 1.5 Tesla (1.5 T) and one 3 Tesla (3 T) magnet, which are typical diagnostic strength magnets. Since stronger magnets generate finer detail but also more artifacts, some examination types are best done on the 1.5 T magnet, others on the 3 T magnet, and still others can be done on either. In addition, each examination type consists of a certain number of mandatory *sequences* and an additional number of discretionary sequences. Typically, multiple images comprise each sequence.

Published online: 29 March 2011

SCH performs MRI examinations on inpatients and outpatients, who might be in an urgent care situation and require an expedited MRI examination.

A critical resource planning question at SCH has been when to increase magnet resources, due to both increasing patient needs, and the high expense of the magnets and support staff. In addition, an important goal is to increase magnet utilization during the normal working day (7 AM–7 PM at SCH) while reducing the overtime pay required to meet the patient demand.

Based on interviews with multiple doctors, technologists, and schedulers at SCH we learned the following about the current situation:

- The current process for scheduling MRI patients was effective and flexible, and would be difficult to change due to both medical and logistical challenges. In particular, SCH uses process-oriented improvement strategies and a flexible inpatient scheduling policy that allows sensible rescheduling of patients as needed throughout a given day.
- There was considerable skepticism that mathematical modeling could address the interest of **both** the patients and SCH with respect to scheduling due to the complexity and uniqueness of individual medical conditions. In addition, there was a reluctance to use reimbursement rates as a metric in determining magnet utilization.
- Although automated data collection of magnet utilization and examination type is already in place, the information was not being used to assess magnet utilization or to identify possible areas of scheduling improvement.

Therefore, we provide SCH with a *descriptive* analysis. In particular, we use optimization models and simulation techniques to identify possible areas of improvement in the current scheduling practices and also to predict the impact of patient examination trends on magnet utilization.

In order to balance both the concerns of the hospital and the patient, we modeled the scheduling problem faced by SCH as having two competing objectives, magnet overtime versus a fairness metric, which we call *fairtime*, that measures the magnet time patients receive and is unbiased by the particular examination type that the patient requires. We devise an algorithm to compute tradeoff curves between the two competing objectives. Since there are several other objectives of importance (e.g., revenue), our model and algorithm are meant to be used in conjunction with other information that the decision maker possesses.

Our contributions are as follows:

(1) Design a decision support tool to help assess current scheduling practices and to identify potential areas of improvement. In particular, our tool is meant to work in conjunction with the process-oriented improvement system in place.
(2) Design a simulation based on current data that, in conjunction with the decision support tool, can be used to perform scenario analysis of varying MRI examination demands. The results of the analysis can then be used to aid in decisions about both personnel and magnet planning.

We provide theoretical and practical evidence that our algorithm performs well. Our algorithm is provably efficient, but is an approximation of the true optimal solution. Since our problem is computationally difficult to approximate, we derive *resource augmentation* guarantees in order to measure the theoretical performance of our algorithm. Practical experience from both real and simulated data support that our algorithm finds solutions that are close to optimality.

We ran the algorithm on a calendar year of patient data from SCH. The data includes demographic information about the patients, examination start times, MRI examination type & duration, inpatient versus outpatient status and magnet type.[1] Based on an analysis of the patient data we designed a Monte Carlo simulation to work with the algorithm to predict solutions to the stochastic variant of the scheduling problem. The results of this stochastic optimization problem are then used as predictions for the impact of scenarios, which then lend support for resource planning decisions.

We emphasize that our decision support tool is not meant to make scheduling or purchasing decisions. Instead, the tool is designed to provide support for decision makers who wish to identify potential areas of improvement in scheduling or analyze predicted outcomes from possible future trends in MRI examination demand.

The literature on medical appointment scheduling is extensive and we describe only a small sample of work in this area. Further references can be found in the surveys by Cayirli and Veral [2] and Gupta and Denton [7]. The work of Vermuelen et al. [15] also uses a multicriteria objective where minimum resource allocation metrics similar to our own fairness metric are used. In their work, they are considering

[1] Our data was collected and analyzed with the approval of the SCH Institutional Review Board.

problems across hospital departments, so their metrics are not directly applicable to our problems. Work by Green et al. [6] also analyze the problem of scheduling MRI magnets. They focus on the two related problems of serving waiting patients and outpatient scheduling. They show that a particular dynamic policy is optimal under some structural assumptions and provide outpatient scheduling heuristics that perform well in computations. Our work departs from theirs in several respects. Their methods seek to maximize profit whereas our methods seek to simultaneously maximize both regular-hour utilization and a patient fairness metric. They also categorize patients into three categories, outpatients, inpatients, and emergency, whereas we use these categories in addition to the over 100 examination types performed on the MRI magnets. In particular, their model assumes that patients all require the same magnet time (one appointment slot) to obtain their analytical results whereas our model takes into account the substantial variability in magnet time. They do provide a simulation analysis of their online policy performance against different examination times. They also deal with the more difficult problem of determining an optimal online policy whereas we focus on an offline optimal schedule for evaluative purposes. In part, the latter difference highlights their focus on a prescriptive model versus our focus on a descriptive model. Because our goal was an offline algorithm to evaluate a schedule and predict impacts of resource allocation decisions, we do not incorporate wait lists (an online policy concern) into the model.

Also relevant to our problem is the work of Patrick and Puterman [11], who use scheduling techniques to determine a scheduling policy for outpatient CT-scanner examinations in order to minimize the amount of overtime used, which is equivalent to our utilization objective. Our work differs from Patrick and Puterman as they focus solely on developing models and algorithms to improve resource utilization, whereas our models and algorithms help to improve resource planning in addition to resource utilization. In addition, our methods also consider a second objective which measures the time allocated to each patient. Kolisch and Sickinger [9] also consider CT-scanner examination scheduling where they use a Markov Decision Process model to analyze priority policies in order to maximize a revenue-based objective which takes into account the same patient groups as in [6]. The same authors also show that a generalization of the classical Bailey–Welch priority rule [1, 17] works well when scheduling two CT-scanners [13]. Vermeulen et al. [16] describe a CT-scanner examination scheduling algorithm that adapts to current and future demands within five patient groups through the use of heuristics and simulation.

We describe a mixed-integer programming formulation and our algorithm in Section 2. In Section 3, we describe our data and simulation for generating predicted future data. In Section 4, we provide results from our algorithm when used on the real and simulated data from Section 3. We discuss our results in Section 5.

2 Calculating a fair resource plan

In this section, we describe an algorithm that determines the trade-off between additional overtime magnet hours required and allocating each patient as much time as possible within the schedule. We use this algorithm to both evaluate past scheduling decisions and to construct forecasts by scheduling simulated data. Because we use simulation to handle uncertainty, we assume all the input data to the algorithm is deterministic.

Mathematical formulation We now describe a mathematical programming formulation of our problem. We make the following assumptions.

1. The schedulers at SCH currently allocate examinations into 45-min time slots. However, the technologists use such schedules as guidelines and make decisions throughout the day to accommodate patients' actual varying examination times. Therefore, the schedule used at SCH is actually flexible and the optimization tool only schedules to the day, rather than in 45-min time slots. For non-emergent patients, we allow the optimizer to reschedule the appointment to any day within two days of their simulated scheduled appointment. We assume emergent patients must be seen on the day they arrive into the system.
2. We assume that the magnet strength (1.5 T vs. 3 T) scheduled for the patient is fixed because such a decision also involves diagnostic, as well as logistic, determinations.
3. For a given MRI examination on a given magnet, there are also a number of discretionary sequences that can be performed in addition to those required by the examination type. Therefore, we assume that each exam has a magnet-specific minimum amount of operation time required and a magnet-specific maximum amount of operation time that could also be used.

Minimizing magnet operation-time involves determining the operation-time allocated for a specific exam conducted on a specific magnet. If all SCH wished to do was maximize the throughput of examinations, the scheduling algorithm should allocate every patient the minimum magnet time required. However, such an allocation ignores the diagnostic value of the discretionary sequences, and so we also create a second objective in order to maximize the minimum normalized time each exam-magnet combination receives. Let $\mathcal{T}$ denote the set of examinations. For a given examination $j \in \mathcal{T}$, let ℓ_j and u_j denote the minimum and maximum of magnet time required, respectively, and p_j be the decision variable indicating the amount of magnet time allocated. Thus, $\ell_j \le p_j \le u_j$. We define

$$\varphi_j = \frac{p_j - \ell_j}{u_j - \ell_j} \tag{1}$$

as the *fairtime of examination* j and define *fairtime* as the minimum fairtime over all examinations, i.e., $\varphi = \min_j \varphi_j$. The use of a minimum resource allocation to patients is not new, e.g., see [15] where the minimum service level for each hospital department is used as an objective when scheduling block appointments for patients.

Below is an exact formulation to schedule examinations on one magnet while simultaneously minimizing the maximum extra magnet time required over all days and maximizing the fairtime.

$$\begin{aligned}
&\text{min/max} \quad \{F, \varphi\} \\
&\text{subject to} \quad \sum_t x_{jt} = 1 && \forall j && \text{(a)} \\
&\qquad \sum_j p_j x_{jt} \le b + F_t, && \forall t && \text{(b)} \\
&\qquad F_t \le F, && \forall t && \text{(c)} \\
&\qquad \ell_j \le p_j \le u_j && \forall j && \text{(d)} \\
&\qquad \varphi_j = \tfrac{1}{u_j - \ell_j}(p_j - \ell_j) && \forall j && \text{(e)} \\
&\qquad \varphi_j \ge \varphi && \forall j && \text{(f)} \\
&\qquad x_{jt} = 0 && t \notin \{r_j, \ldots, d_j\} && \text{(g)} \\
&\qquad F, F_t, p_j \ge 0, x_{jt} \in \{0, 1\}, \forall j, t,
\end{aligned} \tag{2}$$

where the problem parameters are

- r_j, d_j the earliest and latest day examination j can be conducted,
- b the minutes each magnet is available for each day,

and the decision variables are

- p_j the minutes allocated to examination j,
- x_{jt} indicates that examination j is on day t,
- F_t the extra minutes allocated to the magnet on day t,
- F the maximum amount of minutes allocated over all days.

Constraints (a) ensure that each examination is conducted. Constraints (b) ensure that the magnet is allocated enough minutes on a given day t to conduct all examinations. Constraints (c) ensure that the maximum F_t is being minimized. Constraints (d) enforce the upper and lower bound on the examination time allocated to examination j. Constraints (e) calculate the fairtime for examination j. Constraints (f) ensure the minimum fairtime is allocated and constraints (g) ensure that examinations are only scheduled on feasible days. Note that Eq. 2 is a *quadratic* mixed-integer bicriteria program and is generally difficult to solve as modelled. However, the following proposition simplifies the problem. In particular, the following proposition shows that if an optimal solution exists for Eq. 2, then an optimal solution exists where $\overline{\varphi}_j$ are all the same.

Proposition 1 *For every feasible solution* $(x_{jt}, p_j, \varphi_j, \varphi, F_t, F)$ *to Eq. 2 there exists another feasible solution with equal fairtime and the same objective value.*

Proof Let $(x_{jt}, p_j, \varphi_j, \varphi, F_t, F)$ denote a feasible solution vector to Eq. 2. The solution, $(x_{jt}, \overline{p}_j, \overline{\varphi}_j, \varphi, F_t, F)$ is also feasible where

$$\overline{p}_j = \ell_j + \varphi(u_j - \ell_j), \forall j,$$

and

$$\overline{\varphi}_j = \varphi, \forall j.$$

Note that the objective values are identical for both solutions. Therefore, we need only show that the constraints are satisfied for the new solution. Constraints (a), (c), (d), (e) and (f) are all satisfied by definition because the variables in the new feasible solution have not changed on these constraints. To see that constraint (b) is satisfied, note that $\overline{\varphi}_j = \varphi \le \varphi_j$ for all j, which implies $\overline{p}_j \le p_j$. Because $x_{jt} \ge 0$,

$$\sum_j \overline{p}_j x_{jt} \le \sum_j p_j x_{jt} \le b_t + F_t, \forall t.$$

Thus, all constraints are satisfied. □

So, to make the problem linear, we can instead consider the following two families of (linear) integer programs (IP).

$$
\begin{aligned}
G(\varphi) = \min\ & F \\
\text{subject to}\quad & \sum_t x_{jt} = 1 && \forall j \quad (a_\varphi) \\
& \sum_j p_{jt} x_{jt} \le b + F_t, && \forall t \quad (b_\varphi) \\
& F_t \le F, && \forall t \\
& x_{jt} \in \{0, 1\} && \forall j, t,
\end{aligned} \tag{3}
$$

where for all j and t,

$$
p_{jt} = \begin{cases} \ell_j + \varphi(u_j - \ell_j), & t \in \{r_j, \ldots, d_j\}, \\ M, & \text{otherwise.} \end{cases}
$$

Our approach The optimization problems (Eq. 3) are *NP-Hard*, which indicates that it is unlikely for efficient, exact algorithms to exist (see, e.g., [5]). Moreover, commercial solvers could not solve the mixed-integer programs to a reasonable accuracy even after several hours of computation. To develop effective methods to solve our problems, we consider techniques from *approximation algorithms*. For $\beta \ge 1$, a β-approximation algorithm for a minimization problem with optimal objective v^* efficiently finds a solution with objective value $\overline{v}$ such that $\overline{v} \le \beta v^*$. However, because the optimization problems we consider can have $v^* = 0$, finding a β-approximation algorithm is as difficult as finding an exact algorithm. We address this differently by employing *resource augmentation* techniques [8]. For $\gamma \ge 1$, we define a γ-speed scheduling algorithm to be one which finds a solution with objective $\overline{v} \le v^*$ assuming all process times were scaled by γ, e.g., if p_{ij} represent the processing time required by a job j on machine i, then a γ-speed algorithm sets $p_{ij} \leftarrow \frac{p_{ij}}{\gamma}$, i.e., all machines were "sped up" by a factor of γ. Such an approach is natural in our situation because we are interested in minimizing the amount of overtime magnet time required. Note that 1-approximation algorithms find the exact optimal solution, as do 1-speed algorithms.

Because solving Eq. 3 is NP-Hard, we *relax* the integrality constraints $x_{jt} \in \{0, 1\}$ to $0 \le x_{jt} \le 1$ to create a solvable model. We attempted to solve Eq. 3 using CPLEX 11 on two dual-core Opteron 2.6 Ghz processors with 8 GB of memory but did not close the integrality gap to within 10% even after 24 h of computation. Once the LP relaxation of $G(\varphi)$ is solved we can then solve the LP relaxation of the following IP.

$$
\begin{aligned}
H(\varphi) = \min\ & \sum_{j,t} t x_{jt} \\
\text{subject to}\quad & \sum_t x_{jt} = 1 && \forall j \quad (a_\varphi) \\
& \sum_j p_{jt} x_{jt} \le b + G(\varphi), && \forall t \quad (b_\varphi) \\
& x_{jt} \in \{0, 1\} && \forall j, t.
\end{aligned} \tag{4}
$$

We denote $G^{LP}(\varphi)$ and $H^{LP}(\varphi)$ as the LP relaxations of $G(\varphi)$ and $H(\varphi)$ respectively. Note that $G^{LP}(\varphi) \le G(\varphi)$ and $H^{LP}(\varphi) \le H(\varphi)$. Also, all optimal solutions to $G(\varphi)$ are also feasible to $H(\varphi)$ and all optimal solutions to $G^{LP}(\varphi)$ are feasible to $H^{LP}(\varphi)$. Also, in $H^{LP}(\varphi)$, the objective is equivalent to minimizing the average *flowtime* for all examination, i.e., the average number of days from the release day to the day when the examination is completed. For a given φ, holding all p_{jt} to the "minimum" φ for permissible days is without loss due to Proposition 1. Note that solving for $G(\varphi)$ is the computational bottleneck of the FindTradeoff algorithm (see Algorithm 1). By considering each day as "machine," Eq. 3 is actually a problem of minimizing the maximum tardiness where all jobs have an identical due date of b and the machines are unrelated. However, such a scheduling problem is at least as hard as finding the minimum *makespan*[2] on identical parallel machines, which is NP-Hard [4]. Such scheduling problems are also known to be difficult for integer programming software to solve. Moreover, for our problem, if b is set to the optimal makespan i.e., the minimum $\max_t \sum_j p_j x_{jt}$ possible for any schedule, then the optimum maximum tardiness is zero. Thus, no β-approximation algorithm exists if $P \ne NP$ because any β-approximate solution is identical to the true optimum. Thus, we focus on finding a γ-speed scheduling algorithm. We use an algorithm of Shmoys and Tardos [12] that converts solutions of the LP relaxation of $G(\varphi)$ into integral solutions by solving a minimum cost matching problem on an associated bipartite network. The key theorem regarding the objective guarantee is as follows.

Theorem 2 [12] *A solution,* $\mathbf{x}$, *to* $H^{LP}(\varphi)$ *can be converted into an integral solution,* $\overline{\mathbf{x}}$, *to* $H(\varphi)$ *with* $\sum_{jt} t\overline{x}_{jt} = H^{LP}(\varphi)$ *and constraint* (b_φ) *of Eq. 4 satisfied so that*

$$
\sum_j \overline{x}_{jt} \le b + G^{LP}(\varphi) + p_{\max},
$$

where $p_{\max} = \max_{j,t}\{p_{jt} : t \in \{r_j, \ldots, d_j\}\}$.

[2]The makespan is the maximum completion time over all jobs.

Algorithm 1 The FindTradeoff algorithm

FindTradeoff
Inputs: $\ell_j, u_j, b_t, \epsilon$
Outputs: $k = \frac{1}{\epsilon} + 1$ pairs, $\{(F^{(i)}, \varphi^{(i)})\}$
Set $\varphi \leftarrow 0, \quad i \leftarrow 1$
while $(\varphi \leq 1)$
  do
    Set $p_j \leftarrow \ell_j + \varphi(u_j - \ell_j)$
    Solve $G(\varphi)$ and set $(F^{(i)}, \varphi^{(i)}) \leftarrow (G(\varphi), \varphi)$
    Set $\varphi \leftarrow \varphi + \epsilon, \quad i \leftarrow i + 1$
  end
return $\{(F^{(i)}, \varphi^{(i)})\}$

As a corollary, we can state the following.

Corollary 1 *There is an algorithm that can find an integral solution* $\overline{\mathbf{x}}$ *to* $G(\varphi)$ *with* $\sum_{j,t} t\overline{x}_{jt} \leq H^{LP}(\varphi)$ *and objective* $\overline{F}$ *so that* $\overline{F} \leq G(\varphi) + p_{\max}$ *where* $p_{\max} = \max_{j,t}\{p_{jt} : t \in \{r_j, \ldots, d_j\}\}$. *Moreover, such a solution can be found in time proportionate to solving the* G^{LP} *and* H^{LP}.

Proof Suppose that $\mathbf{x}$ is a solution to $G^{LP}(\varphi)$. Then $\mathbf{x}$ is feasible to $H^{LP}(\varphi)$ and we can find the optimal solution, $\mathbf{x}^*$, to $H^{LP}(\varphi)$. We can then find an integral solution $\overline{\mathbf{x}}$ to $H(\varphi)$ by Theorem 2 with

$$\sum_j \overline{x}_{jt} \leq b + G^{LP}(\varphi) + p_{\max}.$$

Clearly, $(\overline{\mathbf{x}}, \overline{F})$ is feasible to $G(\varphi)$ with $\overline{F} \leq G^{LP}(\varphi) + p_{\max}$. □

With respect to our problem, the maximum examination length is known to be no more than some small fraction of the daily magnetic availability. We denote this fraction by α. Then we can state the following theorem as a consequence of Theorem 2.

Theorem 3 *Suppose there is some* $\alpha \in (0, 1)$ *so that* $p_{\max} \leq \alpha b$. *There is an* $(\alpha + 1)$*-speed algorithm to find integral solutions to* $G(\varphi)$.

Proof Let φ be given. Using Corollary 1, solving $G^{LP}(\varphi)$, $H^{LP}(\varphi)$ and applying Shmoys–Tardos will generate a solution $(\overline{\mathbf{x}}, \overline{F}, \overline{F}_t)$ to $G(\varphi)$. Let $\overline{C}$ denote the makespan of the schedule generated by the Shmoys–Tardos algorithm and C^* denote the optimal makespan, i.e., $\overline{C} = \max_t\{b + \overline{F}_t\} = b + \overline{F}$ and $C^* = b + F^*$. Then, by Corollary 1, we have that

$$\overline{C} = b + \overline{F} \leq b + F^* + \alpha b = C^* + \alpha b.$$

Now consider the same solution if each machine's "speed" was increased by a factor of $(\alpha + 1)$, which is equivalent to setting all process times to $\frac{1}{\alpha+1} p_{jt}$. Thus, at this speed, the approximate solution has completion time

$$\overline{C}^S = \frac{1}{1+\alpha}\overline{C} \leq \frac{1}{1+\alpha}(C^* + \alpha b).$$

For a real number $a \in \mathbb{R}$, let $(a)^+ = \max\{0, a\}$. Then the objective on the faster machines is

$$\overline{F}^S = \left(\frac{1}{1+\alpha}\overline{C} - b\right)^+ \leq \left(\frac{1}{1+\alpha}(C^* + \alpha b) - b\right)^+$$
$$= \left(\frac{1}{1+\alpha}(C^* - b)\right)^+ \leq (C^* - b)^+ = F^*. \qquad \square$$

So, in our implementation of FindTradeoff, we solve $G(\varphi)$ approximately. An interesting feature of our method is that we also return a solution with optimal average *flowtime*. In particular, we have the following corollary.

Corollary 2 *The algorithm* FindTradeoff, *when implemented via Theorem* 3, *is a* $(\alpha + 1)$*-speed algorithm to the bicriteria problem of* $G(\varphi)$ *where both* F *and average flowtime are minimized.*

We note that we could use Theorem 2 on the first LP solution. In Section 4 we examine the benefit of using the second over the first.

3 Simulating schedules

In this section, we describe our simulation model used to generate inputs for FindTradeoff. Since we seek to simulate the MRI examination demand at SCH, we first conduct an analysis of our data sets. In this study, we use both MRI patient data (the *patient dataset*) and weekday examination counts (the *trend dataset*) from SCH radiology department. We describe and analyze the patient dataset in Sections 3.1 and 3.2 and the trend dataset in Section 3.3 below. We describe our simulation model in Section 3.4.

Our data analysis and simulation model was written in S-Plus [14] and performed on two dual-core Operon 2.6 GHz processors with 8 GB of RAM.

3.1 Collecting the patient data set

The patient data set is composed of 5,917 MRI examinations conducted from July 1, 2007 to June 30, 2008 at SCH and consists of two basic types of examinations. Not included in the data set were 17 records for reasons described below. Patients who made examination appointment times through the schedulers at SCH are referred to as *scheduled patients*. Scheduled patients' appointment times were recorded in the schedulers' computer system. Patients who did not schedule examination appointment times and were added on an ad hoc basis to the preexisting schedule by the technologists are referred to as *unscheduled patients.* Typically, unscheduled patients had symptoms requiring urgent exams.

Depending on their condition, previous examination history, and diagnosis, patients received varying amounts of pre-examination preparation and were examined on either the 1.5 T or 3 T magnets. Some exams require that a patient receive intravenous contrast and/or anesthesia in order to receive a diagnostic quality examination. In addition, preparation time for both the patient and magnet was necessary prior to each exam. The amount of patient and magnet preparation time required prior to an exam was not recorded, although the need for intravenous contrast and anesthesia was.

A magnet is considered to be in operation from the time the first image is acquired (magnet on-time) until the time the last image is acquired (magnet off-time). Both magnet on-time and off-time were automatically recorded by the computers used to operate the magnets. The magnet operation-time was calculated as the interval between magnet off-time and magnet on-time. During the examination, the technologist recorded additional information including the exact exam type performed. After the completion of an exam, a radiologist interpreted the results and dictated a report. A radiologist and an undergraduate assistant combined these separate computer records and encrypted any personal patient information as per IRB specifications.

Throughout the course of the year-long period, the Radiology Department performed 5,934 examinations, of which 17 were omitted and 5,917 included in our study. The 17 patient records omitted, lacked crucial data fields (nine records), appeared to have incorrect data fields (six records), or corresponded to patients too old to be considered representative of the pediatric population (two records). A complete list of the 17 omitted records with justifications is available from the authors upon request.

3.2 Data analysis of the patient data set

The initial analysis of the patient data set yields several trends in patient arrivals and magnet operation-times.

1. There is often a strong preference to perform certain examination types on a specific magnet strength (1.5 T vs. 3 T).
2. Magnet operation-time distributions differ depending on examination type, magnet type, and scheduled status.
3. The volume of unscheduled patients examined varies throughout the day, exhibiting similar behavior, but at different magnitudes, for each weekday and magnet type.

SCH Radiology Department performed 112 different types of MRI examinations during the study period. The proportion of male to female patients was roughly equal in the 5,917 examination records (51.7% male). Unscheduled patients accounted for a significant fraction of the examinations performed (19.7%). The 1.5 T magnet was utilized approximately 50% more than the 3 T magnet with respect to both the number of examinations (59.0% on the 1.5 T) and the amount of magnet operation-time (60.8% on the 1.5 T). The vast majority of MRI examinations conducted were on children (99.9% under age 18). The ages of patients in the data ranged from one day to 25 years, with a median age of 9.58 years.

For most examination types, the data indicates a strong preference for a particular magnet. However a magnet preference that is apparent for scheduled examinations was not apparent among unscheduled examinations.[3]

The amount of magnet operation-time required for an examination is dependent upon the type of examination, the magnet on which the examination is performed, and the scheduling status of the examination. This implies that the median and variance of magnet operation-time differs not only between examination types, but also between magnet type and scheduled status. These trends are graphically represented in Fig. 1. Thus, the distribution of magnet operation-times for a single examination type may exhibit different behavior depending on scheduled status and magnet type.

To further analyze the magnet operation-time, we performed a statistical analysis using a Cox Proportional Hazards (CPH) model [3]. We examined a large

[3] We suspect magnet availability is a mitigating factor.

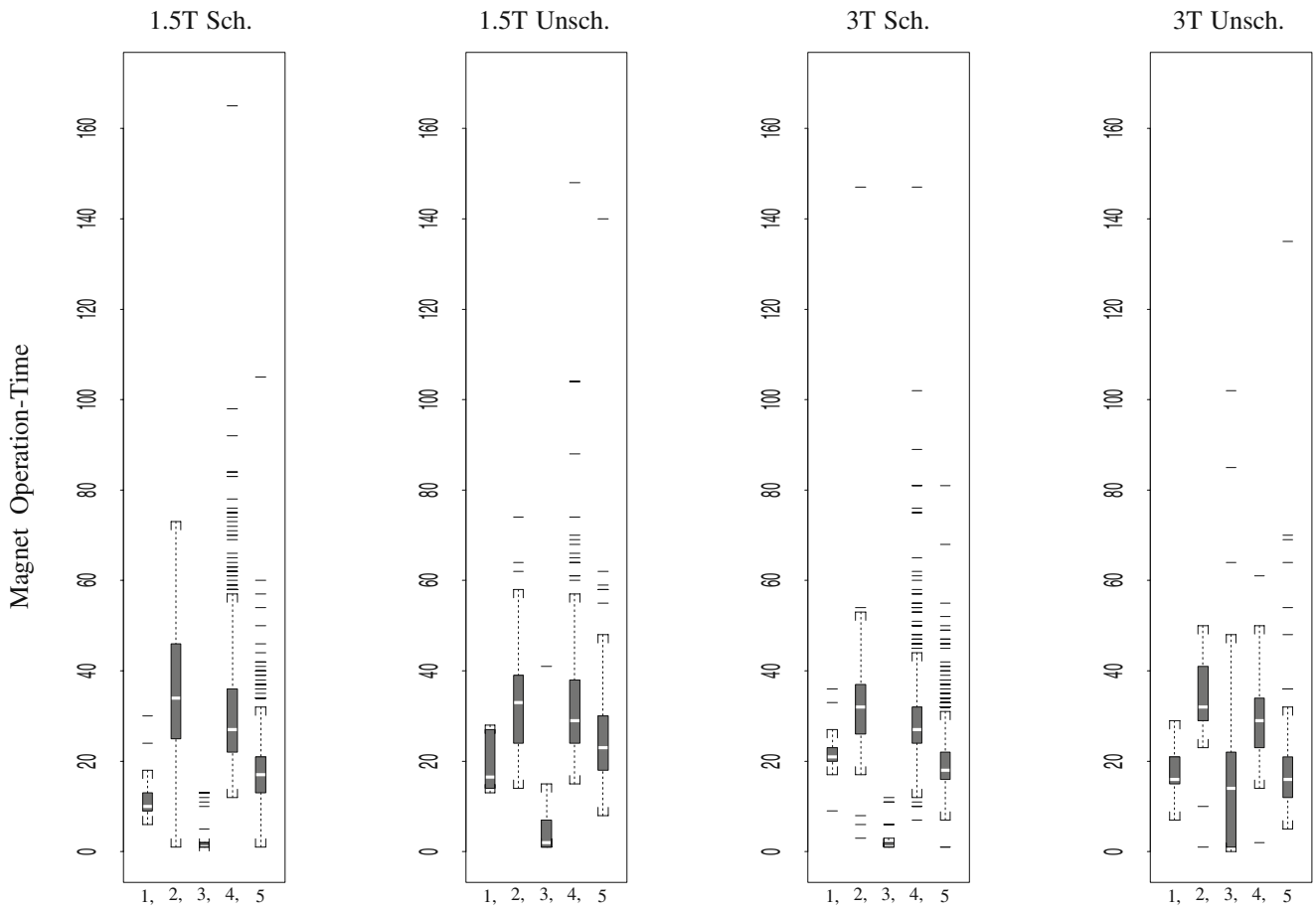

Fig. 1 Magnet operation-times by magnet and examination type. *Shaded boxes* represent the middle 50% of magnet operation-times, with the *white line* demarcating the median. The whiskers extend out from the 25th and 75th percentiles to the most extreme value still within 1.5 times the range of the middle 50% of the data. The examination types plotted above represent all types with at least 50 examinations in an individual category. The single zero magnet operation-time present is due to rounding. The examination types are listed as follows: *1* abdomen without contrast; *2* angio head without contrast; *3* brain limited; *4* brain with and without contrast; *5* brain without contrast

subset of the patient data set (3,901 records) corresponding to the five most common examination types, so that the effects of examination type would be accounted for with a sufficient number of observations. Table 1 gives the results of our CPH analysis. In a CPH model, a log-linear function is used to estimate the covariates effect on the examination time. A positive coefficient implies that as a single covariate changes (and all others are held constant) from a lower to a higher level, the magnet operation-time decreases on average. Similarly, a negative coefficient implies that examinations should take longer on average moving from a lower to a higher level for an individual covariate. The results in Table 1 are listed in decreasing order of significance with covariates satisfying a $p < .05$ level of significance, above the horizontal rule. The results of the CPH analysis indicate that examination type plays a significant role in determining magnet operation-time, as does contrast, inpatient status, magnet type and scheduled status.

Using the magnet on-times as estimates of daily arrival rates,[4] we found that the average arrival rates of unscheduled patients vary depending on the day of the week, with Mondays and Fridays having the most unscheduled arrivals to both magnets. Also, as shown in Fig. 2, the average arrival rates for the 1.5 T magnet were consistently higher than for the 3 T magnet, both across individual weekdays and in aggregate. Despite having the most unscheduled cases on Mondays and Fridays (see Fig. 3), the Radiology

[4]Using magnet on-times as an estimate of arrival rates does not take reneging into account. However, based on interviews with technologists and doctors at SCH, the occurrence of reneging was insignificant.

Table 1 The results of the Cox Proportional Hazards model analysis of examination lengths on a large subsection of data (3,901 examinations)

Covariate (low value, high value)	coef	Exp (coef)	SE (coef)	*z*-score	*p*-value
Brain without contrast	−0.57106	0.565	0.0361	−15.827	0.0e+00
Spine total without contrast	−0.39959	0.671	0.0436	−9.174	0.0e+00
Brain with and without contrast	−0.45485	0.635	0.1071	−4.247	2.2e–05
Contrast dye status (used, not)	0.37663	1.457	0.1050	3.589	3.3e–04
Outpatient status (out, in)	−0.08983	0.914	0.0293	−3.063	2.2e–03
Magnet type (1.5 T, 3 T)	0.04757	1.049	0.0182	2.618	8.8e–03
Spine total with and without contrast	−0.26772	0.765	0.1111	−2.409	1.6e–02
Scheduled status (scheduled, unscheduled)	−0.06495	0.937	0.0285	−2.280	2.3e–02
Repeat examination status (repeat, new)	−0.02975	0.971	0.0175	−1.699	8.9e–02
Priority status (priority, not)	−0.04411	0.957	0.0375	−1.175	2.4e–01
Gender (male, female)	−0.01368	0.986	0.0162	−0.842	4.0e–01
Anesthesia status (used, not used)	−0.01978	0.980	0.0244	−0.811	4.2e–01
Age (years)	0.00046	1.000	0.0040	0.114	9.1e–01

For each binary, non-numerical covariate, we list our designations of the 'high' and 'low' settings. The entries are sorted in decreasing level of significance. The excluded examination was MR Angio Head without Contrast

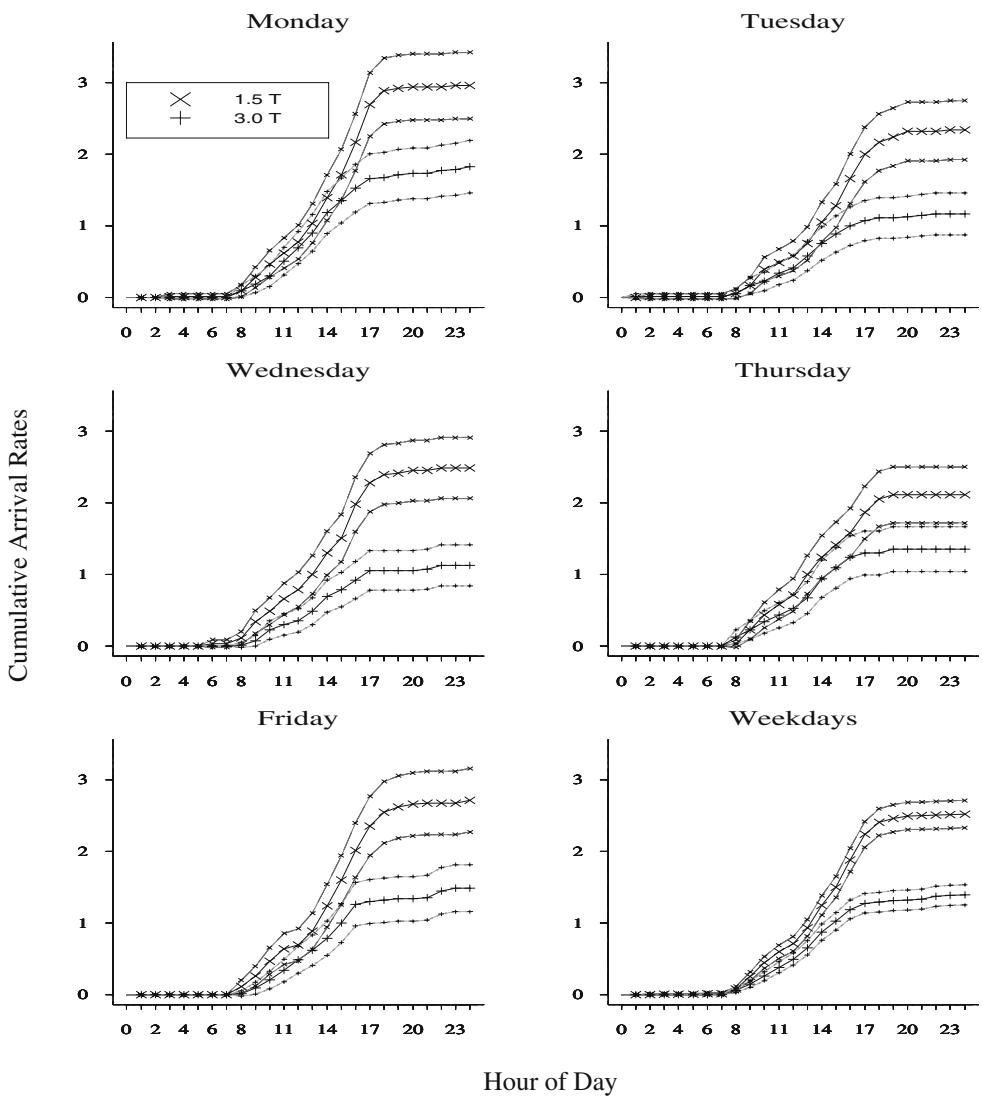

Fig. 2 Magnet-specific cumulative arrival rates. The cumulative arrival rates to each magnet, by day, with linear interpolation, with associated 95% confidence intervals. The vertical axes represent the number of patients arriving on average while the horizontal axes represent the time of day, from midnight to midnight. The larger marks represent the average arrival rates while the smaller marks represent the bounds of the 95% confidence intervals

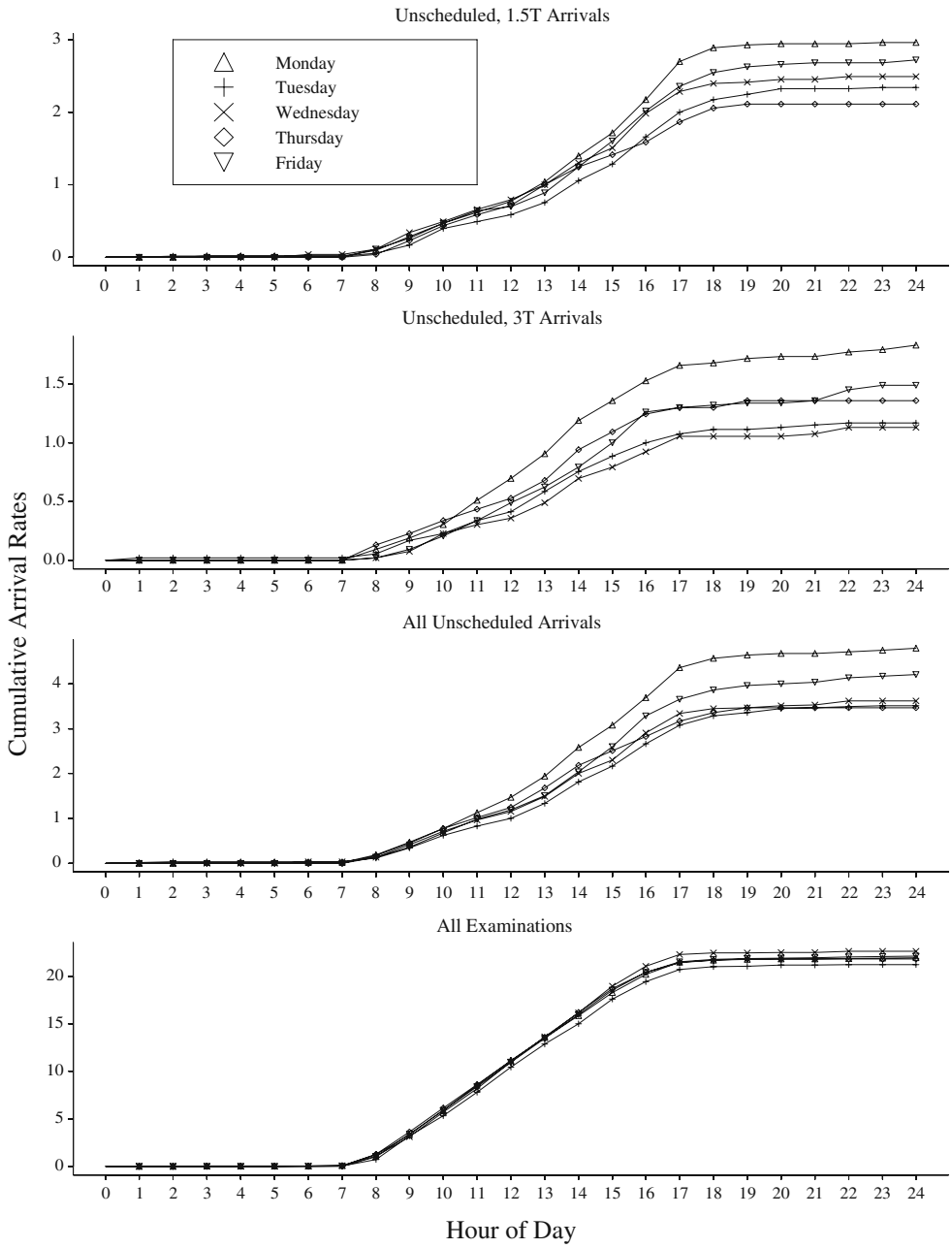

Fig. 3 The cumulative arrival rates by day. The first two plots show the cumulative arrival rates of unscheduled examinations for the 1.5 T magnet and the 3 T magnet, respectively. The third plot shows the cumulative arrival rates of unscheduled examinations for both magnets, and the fourth plot displays how the average total number of examinations increases over the course of a day

Department performed the highest number of examinations on Wednesdays. This implies that Wednesdays also have the most scheduled examinations on average. The data also suggests a increasing trend across the year in the average number of examinations performed daily. We note that the trend is supported by regression analyses on the trend data set, as described below in Section 3.3 (see Fig. 4).

3.3 The trend data set analysis

The trend data set consisted of the total number of examinations per weekday dating from January 1, 2003 to January 31, 2009. As of January 1, 2003, SCH operated a single 1.5 T magnet. On October 14, 2006 a 3 T magnet was installed and the 1.5 T magnet was replaced with a newer 1.5 T magnet. We removed 71 weekdays from our data set on which the MRI magnets were not fully operational, such as holidays, staff training days, or days on which maintenance or installation occurred. A detailed description of the days we removed are available on request.

In order to determine the trend in MRI examinations, we regressed the number of daily MRI examinations on the weekdays elapsed in the trend data set. Since the number and type of available magnets

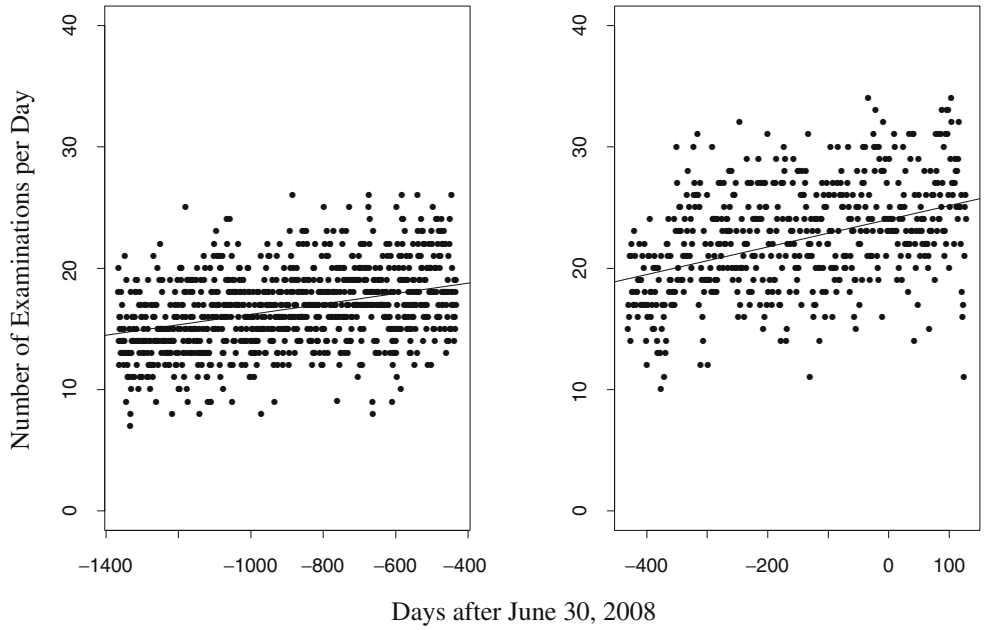

Fig. 4 The single magnet data and regression line (*left*) and the two-magnet data and regression line (*right*). The horizontal axis measures the days in our trend data set with day 0 representing June 30, 2008, the last day of the patient data set. The single magnet plot ends the day of the installation of the 3 T magnet, and the two-magnet plot begins the day after the installation

changed, we performed two separate linear regressions. One regression was performed on the weekdays from January 1, 2003 to October 14, 2006, and the other was performed on weekdays from October 14, 2006 until January 31, 2009. The slope found on the first regression is 0.0048 ($p < 0.001$) and on the second regression is 0.0114 ($p < 0.001$). The variance in the MRI examinations explained by the weekdays elapsed is similar for both regressions ($r^2 = 0.1565$ for the first, $r^2 = 0.1738$ for the second), although the overall variance is high.

3.4 The simulation model

In order to simulate the number of patients requiring MRI examinations, we consider the arrival process of unscheduled patients and the distribution of scheduled patients. We assume that the unscheduled and scheduled patients are distributed independently and model each population separately. Scheduled patients call the scheduler's office to obtain an appointment for their MRI examination. The scheduler then allocates a time in the schedule that is mutually acceptable to the hospital and the patient. Appointments such as these may be scheduled months in advance. Examinations performed on unscheduled patients, in contrast, are much harder to plan for. Unscheduled patients may arrive to the Radiology Department from other departments within SCH or from outside clinics or hospitals with little to no warning, and may require immediate attention. Therefore, we also model the priority status for patients which indicates how soon the patient must undergo an MRI examination.

Modeling the unscheduled patient arrival process relied heavily on the use of magnet on-times to approximate arrival times. As shown in Fig. 3, unscheduled patients do not arrive at a constant rate. Arrivals are infrequent during the night and early morning hours, increase dramatically during the afternoon, and subsequently decrease to nighttime levels. Arrival rates also differ by day of the week and magnet type (Figs. 2 and 3). Since SCH only records magnet on-time, and no more than one examination can be performed at a time on a given magnet, it is nearly impossible to determine the frequency with which groups of two or more simultaneous patients arrive from our data. Anecdotal evidence suggests that group arrivals are infrequent, and so we assume that patients arrive singly. Given this information, we model the arrival of unscheduled patients as a non-homogeneous Poisson process. A Poisson process assumes that the chance of an arrival in a given time interval only depends on the average arrival rate and length of the time interval. A non-homogeneous Poisson process has time-dependent arrival rates. Average arrival rates were calculated for each hour of each weekday. For more information about non-homogeneous Poisson processes, including applications, see [10].

Once the arrival time of an unscheduled patient had been generated, we independently and randomly sample examinations from the data recorded from July 1, 2007 to June 30, 2008 to determine the examination type, priority status, contrast and anesthesia requirements and magnet operation time. This method is a simple way to incorporate the underlying correlations

of exam traits to be represented in our simulated data. By doing so, we are assuming that the examination type required by unscheduled arrivals are independent of time of day and day of week. Such an assumption was reasonable in light of the small proportion of unscheduled arrivals.

We took a different approach in modeling the distribution of patients requiring scheduled examinations. We do not have information regarding when patients first contacted the scheduling office. Instead, we use the magnet on-times and examination dates of scheduled examinations in order to model the distribution of MRI examinations amongst the scheduled patients. We generate the examination information for a day by randomly selecting a weekday from the patient data set, using the scheduled examinations from that day as the examinations requiring scheduling. We also use the actual magnet operation-time as simulated operation time in order to account for codependencies in magnet operation-time and the examination type.

Since the simulation is used to model future patient load of the MRI magnets, we use the slope of the second regression as the trend in increased MRI examinations. Since schedulers can, ostensibly, control the number of scheduled MRI examinations, we assume that the increase in examinations was all in unscheduled patients.[5] Thus, each day of the simulation, the arrival rate of unscheduled patients was increased by the slope of the second regression (i.e., 0.0114).

Depending on what date we began our simulation sampling, we scaled the arrival rate by this rate.

For the results reported, the simulation begins the day after the end of our study period, July 1, 2008. For each ten-year analysis, we generated enough replications so that the relative sample error was below 3%. Each replication consisted of ten years of patient data. The number of replications required to reduce the standard error below 3% varied but was always fewer 100 replications. Therefore, we generated 100 replications for each analysis.

4 Results

Our tool can be used to help identify potential improvements in scheduling and/or resource allocation as well as resource planning. In Section 4.1, we discuss how the former can be done by running the scheduling algorithm on the real data and making a comparison to the actual schedule. In Section 4.2, we describe how the algorithm and simulation models can be used to help forecast and plan for additional magnet resources. Since the utilization is higher on the 1.5 T magnet, we present results only for this magnet. Our methods can be used identically on the 3 T magnet.

For our aggregate measures, we focused on yearly magnet overtime minutes. We define magnet overtime as the amount of magnet time that occurred outside normal operating hours, which were from 7 AM to 7 PM. When measuring the magnet overtime minutes for a year, applying a statistic (e.g., average, median, etc.) without removing the days where no magnet overtime occurred artificially skews the statistic based on the incidence of magnet overtime. For example, suppose 2009 had two days with magnet overtime of 50 min in each day, and in 2010 there were 20 days with magnet overtime of 50 min in each day. The ordinary average overtime in 2009 (assuming 250 workdays) is 0.2 min and 4 min in 2010. However, this masks the actual change: the conditional average magnet overtime in both years was 50 min whereas the incidence sharply rose from 2 days out of 250 to 20 days out of 250. Therefore, we examined the incidence and the quantity separately. Our chief metrics were to examine the incidence of days with magnet overtime and the conditional magnet overtime minutes, i.e., the average magnet overtime minutes over days when an overtime event occurred.

FindTradeoff predicts lower bounds on the amount of overtime minutes for the following reasons.

- FindTradeoff works with magnet ontimes only—thus, there are no setup times.
- FindTradeoff is allowed to schedule all scheduled patients within two days of their originally simulated examination date and all unscheduled patients in the same day as their original simulated examination date.
- FindTradeoff is working with planned examination times, whereas in reality, some examinations can take longer than planned for.

The predictions are therefore conservative predictions of how many magnet overtime minutes occur, based on what the fairtime, φ, is desired.

4.1 Real data

In this section we show how our algorithms and models can be used as a decision support tool to determine time periods where potential efficiency improvements exist. We focus, as previously stated, on the scheduling for

[5] We note that the increase could also be due to other factors such as improved efficiency in conducting examinations or an increased proportion of shorter MRI examinations required.

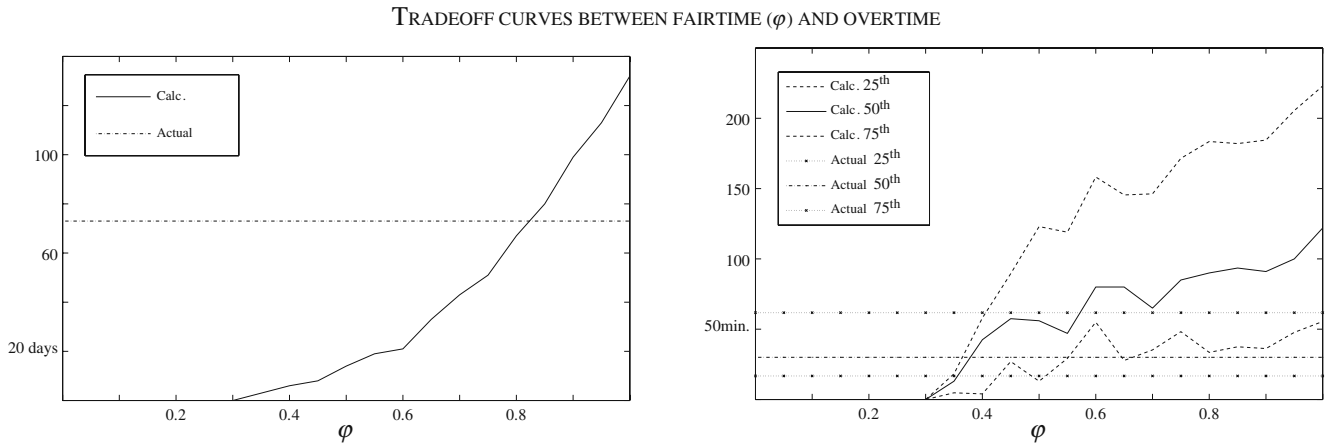

Fig. 5 These results were generated when FINDTRADEOFF was run on the real data for the 1.5 T magnet. The graph on the *left* displays the tradeoff between fairtime (φ) and days with overtime. The graph on the *right* displays the tradeoff between percentage of allocated examination time and the quartiles of the conditional overtime. The *horizontal lines* are the quartiles from the actual overtime that occurred

the 1.5 T magnet. We examine which time periods contained magnet overtime for the real schedule but none or little for the algorithm. These periods are potential areas of inefficiency that can be further examined by the decision maker. By identifying these periods, the aggregate amount of data that the decision maker has to consider is reduced.

The actual overtime can be estimated from the data, by adding up the magnet times that occurred outside of the usual working day (i.e., beyond 7 AM to 7 PM). We omitted weekend hours from our analysis since these appointments were never scheduled. We then ran FINDTRADEOFF on the original data set by treating each of the magnet on times as their originally scheduled times. We allowed the algorithm to reschedule their examinations within two days of their original appointment. This allows the algorithm considerably more flexibility than the actual schedule. The resulting tradeoff curves are depicted in Fig. 5. Both the incidence and conditional overtime are shown as described above. From the tradeoff curves, we can see that using a φ from 0.45 to 0.55 results in an average overtime that is within the same amount of conditional overtime experienced in reality. Based on this observation, we then compared the overtime for each day from the data to overtime the algorithm attained for $\varphi = 0.45$ and $\varphi = 0.55$. The real schedule had 94,465 minutes of magnet ontime versus the algorithm which scheduled 118,294 minutes for $\varphi = 0.45$ and 131,592 minutes for $\varphi = 0.55$. These comparisons are shown in Fig. 6. As can be seen, the period from August 20, 2007 to November 5, 2007 contains days of overtime for the real schedule but none for the algorithm, even as φ was increased.

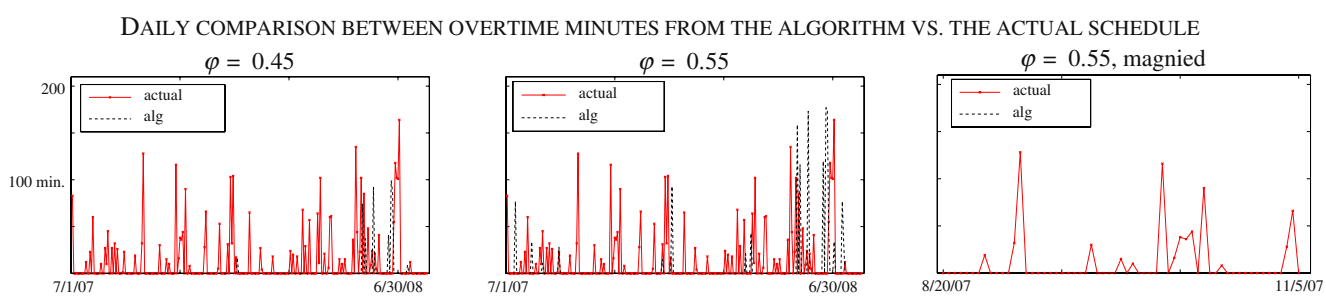

Fig. 6 The first two graphs compare the schedule from FINDTRADEOFF to actual schedule on the 1.5 T magnet when varying fairtime (φ) from 0.45 to 0.55, which increases the total minutes scheduled from 118,294 to 131,592 over one year. The original schedule contained 94,465 minutes of magnet ontime. The comparison between the algorithmically generated schedule and the actual schedule can be used to identify "hotspots" in the scheduling as shown in the third graph, for a two-and-a-half month interval. The vertical axis on all graphs are the same

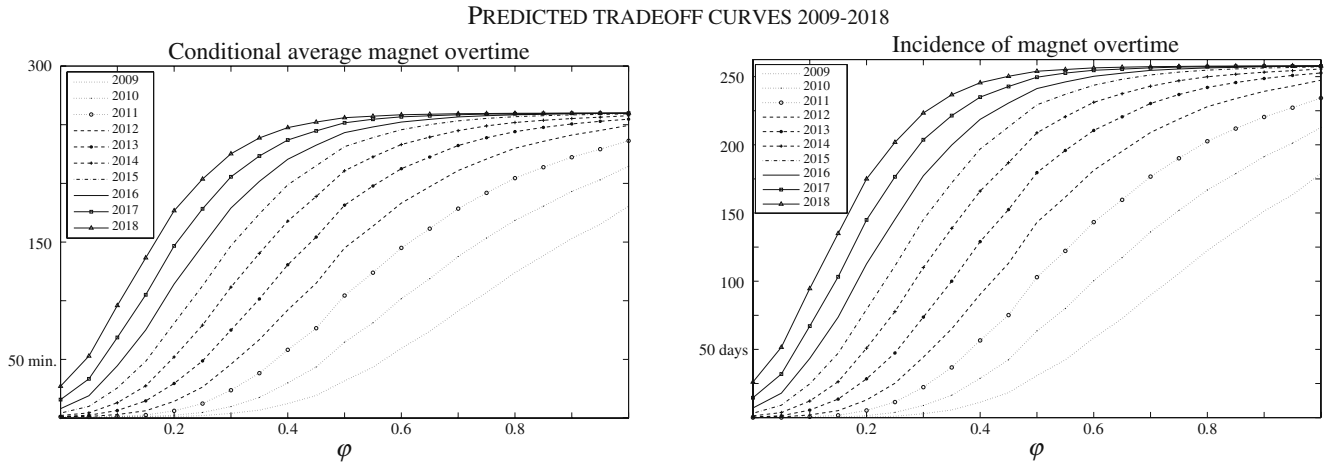

Fig. 7 The two graphs show the tradeoff curves for 2009–2018 on the 1.5 T magnet assuming the trends described by the 2007–2008 dataset. The graph on the *left* represents the average magnet overtime among days that had magnet overtime. The graph on the *right* represents the number of total weekdays with a magnet overtime examination. The total number of weekdays in a year is 252 (excluding holidays)

4.2 Simulated results

In this section, we describe the results of our simulation when combined with FindTradeoff. Moreover, we describe how the results could be used to aid in determining when examination demand will exceed the capacity of magnet and personnel resources. As we previously noted, we analyze data for the 1.5T magnet.

We generated simulated data for 2009 through 2018 and created schedules by using FindTradeoff with results shown in Fig. 7. As can be seen in the right graph, given the current trends, the predicted incidence and severity of magnet overtime steeply rises. By determining the desired φ as described in Section 4.1, the decision maker can use the predicted results as a lower bound (see above) on the amount of magnet overtime incidence and minutes. Thus, the model can predict when staffing or equipment changes would be required to avoid exceeding some cutoff for overtime incidence and minutes. For example, if $\varphi = 0.3$ were desired and a threshold of 50 days of overtime were set, then under current trends, FindTradeoff suggests increasing 1.5 T magnet resources in 2012, as can be seen in Fig. 7.

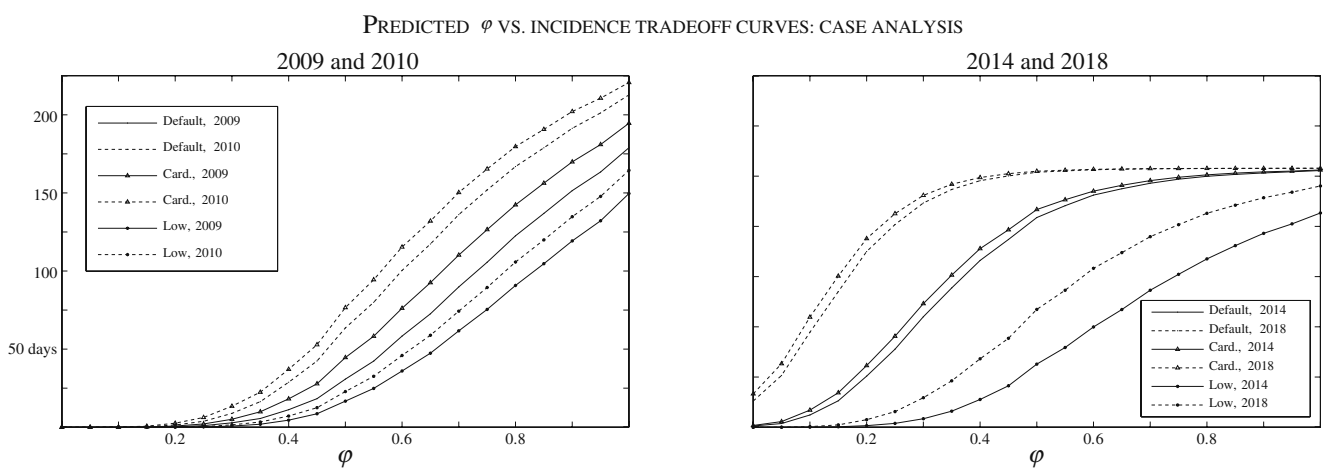

Fig. 8 These two graphs show a case analysis of the predicted tradeoff curves for 2009, 2010, 2014 and 2018 on the 1.5 T magnet between fairtime (φ) and the incidence of magnet overtime days. *Default* is the case where the trends are as predicted by the data. *Card* is the case where the number of cardiac patients increases three-fold. *Low* is the case where the increase in examinations predicted by the data is reduced by a third. The graph on the left depicts the predictions for 2009 and 2010, and the graph on the right predictions for 2014 and 2018. Both graphs have the same *y*-axis. The graphs show the high sensitivity in overtime incidence to examination demand, especially as the forecast horizon is extended

In addition to using the historical trends to generate predictions, we can also perform case analysis. We demonstrate this with two different scenarios. In the first, we simulated increased cardiac examinations by a factor of three. In the second, we decreased the overall trend by a third. Both of these cases are shown in Fig. 8, along with the original simulation for comparison. As shown, the impact of increasing the cardiac examinations has a modest shift on the tradeoff curves, whereas a reduction in the overall trend results in a larger improving shift in the magnet overtime incidence. Additional graphs displaying the results of case analysis are available from the authors upon request.

5 Discussion

The use of technological devices has become ubiquitous in modern medical diagnosis and care. Given the cost and importance of their use, effective allocation of such resources is an important problem facing healthcare providers. In this paper, we have developed a tool that can perform both retrospective analysis of previous decision making as well as a case analysis of the impact of varying trends in MRI examination demand.

Our results have demonstrated the following:

- Optimization models, which are typically used in prescriptive modeling, can be used to great effect in descriptive modeling. In particular, we showed how optimization techniques can be used to evaluate scheduling decisions, purchasing policies, and MRI examination difficulty.
- Approximation algorithm techniques along with Monte Carlo simulation can be used to successfully find tradeoff curves for a bicriteria problem.
- The metric fairtime effectively measures the fairness of magnet time allocation to patients that is unbiased with respect to examination length. This normalization also helps allow for multiple kinds of examination types to be handled within the model.

Our results and methods come with limitations as well. For example, revenue and costs were not addressed by our methods. In addition, given the difficulty in using the past to predict the future, much of the accuracy of our forecasting is dependent on the decision maker having a good understanding of future trends in patient demand and choosing accurate scenarios to examine.

Response from SCH has been positive to our tool. A future study has been planned where the impact of increasing certain examination types will be analyzed using our simulation and algorithm. In addition, since the demand for MRI examinations is rising, interest has been expressed in using the tool to analyze the potential impact of purchasing an additional magnet versus extending magnet operating hours. In particular, the tool has helped demonstrate that extending magnet operational hours is a viable option due to the sensitivity of overtime incidence to demand trends.

Our algorithm and simulation has been a positive step towards developing an evaluation and prediction tool for better resource allocation and planning for medical technology. However, many challenges remain, including extending the tool to other treatment and diagnosis modalities such as CT-scanner scheduling. We are also interested in applying similar optimization and simulation tools to other areas in medicine where descriptive modeling is required.

Acknowledgements We thank the referees and editor for their extremely helpful comments. Also, we thank Eyjo Ásgeirsson and Cliff Stein for many useful discussions and Tom Crockett for his help with the computational experiments. Finally, we thank the technologists and schedulers at SCH for their valuable information and insights.

References

1. Bailey N (1952) A study of queues and appointment systems in hospital out-patient departments, with special reference to waiting-times. J R Stat Soc, B (Methodological) 14:185–199
2. Cayirli T, Veral E (2003) Outpatient scheduling in health care: a review of literature. Prod Oper Manag 12:519–549
3. Cox D (1972) Regression models and life-tables. J R Stat Soc, B (Methodological) 34(2):187–220
4. Garey M, Johnson D (1978) "Strong" NP-completeness results: motivation, examples, and implications. J Assoc Comput Mach 25:499–508
5. Garey M, Johnson D (1979) Computers and intractability: a guide to the theory of *NP*-completeness. W.H. Freeman, New York
6. Green L, Savin S, Wang B (2006) Managing patient service in a diagnostic medical facility. Oper Res 54:11–25
7. Gupta D, Denton B (2008) Appointment scheduling in health care: challenges and opportunities. IIE Trans 40:800–819
8. Kalyanasundaram B, Pruhs K (1991) On-line weighted matching. In: Proceedings of the 2nd ACM-SIAM symposium on discrete algorithms, pp 234–241
9. Kolisch R, Sickinger S (2008) Providing radiology health care services to stochastic demand of different customer classes. OR Spectrum 30:375–395
10. Leemis L, Park S (2005) Discrete event simulation: a first course. Prentice-Hall, Upper Saddle River
11. Patrick J, Puterman ML (2008) Improving resource utilization for diagnostic services through flexible inpatient scheduling: a method for improving resource utilization. J Oper Res Soc 58:235–245
12. Shmoys DB, Tardos E (1993) An approximation algorithm for the generalized assignment problem. Math Program A 62:461–474

13. 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
14. Venables W, Ripley B, Venables W (2002) Modern applied statistics with S-PLUS, 4th edn. Springer Verlag
15. Vermeulen I, Bohte S, Elkhuizen S, Bakker P, La Poutré H (2008) Decentralized online scheduling of combination-appointments in hospitals. In: Proceedings of ICAPS-2008, the international conference on automated planning and scheduling. Sydney, Australia, pp 372–379
16. Vermeulen I, Bohte S, Elkhuizen S, Lameris H, Bakker P, Poutré H (2009) Adaptive resource allocation for efficient patient scheduling. Artif Intell Med 46:67–80
17. Welch J, Bailey NTJ (1952) Appointment systems in hospital outpatient departments. Lancet 259:1105–1108