A Physicians Planning Framework for Polyclinics Under Uncertainty

Mohammad Tohidi[a,*], Masoumeh Kazemi Zanjani[a], Ivan Contreras[a]

[a] *Concordia University and Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT), Montreal, Canada H3G 1M8*

Abstract

The physicians planning procedure in polyclinics is associated with clinics' schedules and requirements, and consists of three levels: clinics scheduling and capacity planning, physicians scheduling, and physicians rescheduling. In this paper we present a bi-level planning framework in which clinic scheduling and capacity planning are developed in the first level, and physicians scheduling and rescheduling in the second level. The first level is modeled as an adjustable robust scheduling problem in order to immune the generated schedules against uncertainty in demand. We develop an implementor/adversary algorithm for solving the robust optimization problem. To cope with variability in patients' treatment times, we formulate the second level as a two-stage stochastic model. Decisions regarding physicians initial schedules are taken in the first stage, and adjustments are considered as the second stage recourse actions. We use a sample average approximation scheme for solving the resulting stochastic optimization problem. We apply our approach to data obtained from a university health center in Montreal, Canada. Using Monte-Carlo simulation, we demonstrate that the schedules generated by our approach are superior to the ones developed by a single level deterministic model.

Keywords: OR in health services, Physician scheduling, Robust optimization, Stochastic programming, Monte-Carlo simulation

1. Introduction

Manpower planning in service industries follows a three-level procedure: planning, scheduling and allocation (Abernathy et al., 1973; Campbell, 2011; Bard & Purnomo, 2005b). The first level conducts the operating policies of service centers. The second level specifies the personnel working shifts over the planning horizon. The last level corresponds to allocating the workers within the

*Corresponding author

Email addresses: m_tohidi@encs.concordia.ca (Mohammad Tohidi), kazemi@encs.concordia.ca (Masoumeh Kazemi Zanjani), icontrer@encs.concordia.ca (Ivan Contreras)

constraints established by previous levels, and making required adjustments. This paper introduces a planning framework for physicians in polyclinics. Polyclinics consist of clusters of clinics that share certain resources such as examination rooms. In these centers, interdisciplinary clinics collaborate in the assessment and treatment of patients, which requires them to be scheduled simultaneously for a certain number of shifts. In this context, we are looking for a weekly schedule for physicians affiliated with different clinics such that the patients scheduled during different clinic sessions (shifts) can be assessed while respecting administrative constraints, resource capacity limitations, and physicians' preferences. It is thus evident that clinics' and physicians' schedules are closely interrelated.

In polyclinic physician planning, the first decision level is to determine clinics' schedules along with a capacity plan (i.e., the number of patients to admit, working hours for physicians and the number of required examination rooms). We denote this as the *clinic scheduling and capacity planning problem* (CSCPP). In this level, clinic schedules are determined according to a weekly demand forecast. Nevertheless, over a long term planning horizon (e.g., one year) the weekly forecast can fluctuate within a given uncertain interval. Hence, hospital's management seeks a robust strategic planning tool to maximize the number of patients who can be served over the week in the presence of worst-case (extremely high) demand scenarios, and to specify the optimal personnel and resource requirements in every planned shift. The aforementioned clinics' schedule along with the capacity plan are expected to be effective over a period of six months to a year.

The second level of the planning process involves the tactical assignment of physicians to the established shifts within the requirements of the shifts and physicians preferences. Generally speaking, physicians scheduling problems consist of constructing work schedules for physicians in a planning horizon such that at each shift there are enough physicians to satisfy demand (Gendreau et al., 2006). Similarly, in polyclinics, physicians are usually scheduled on a weekly-basis once the number of patients scheduled for every shift, physicians availability, and preferences are known. It is worth noting that the outputs of the strategic CSCPP, i.e. number of patients to be served, examination rooms, and physicians' working hours, act as an input to the second decision level. We refer to this problem as the *physician scheduling problem* (PSP). Hospital administration can employ the PSP as a tactical planning tool to obtain physicians' work schedules at the beginning of each week such that all scheduled patients can be assessed, while taking into account capacity constraints and physicians' preferences.

Uncertain events disturb planned operations in health services; physicians' schedules are also

affected by the uncertainty in patients' treatment times. More precisely, if the actual treatment times are longer than the estimated ones, the scheduled physicians might not be sufficient to serve all patients during the regular shifts. In this case, either patients must be rescheduled on another session or extra resources such as on-call doctors or overtime shifts might be deployed. On the contrary, if treatment times are shorter that the estimated ones, physicians would be idle which is not desirable for the administration either. That leads to the third level of the physicians planning procedure, which is to make adjustments to the planned schedules (i.e. prior to each shift). In this operational planning level, proper corrective (recourse) actions must be foreseen in order to minimize the average cost of schedules in the presence of uncertainty in treatment times.

Inspired by a real case study in an ambulatory cancer treatment polyclinic, we present a bi-level physician planning framework under uncertainty. This framework takes into account the uncertainty in the number of arriving patients in the first level, and variability in patients' treatment times in the second level. The decision levels in our approach are similar to the ones introduced by Zaerpour et al. (2017) and follow the three-stage manpower planning process commonly used in service centers. Figure 1 summarizes the physicians planning procedure in polyclinics, and displays our planning framework. At the first level, clinics' weekly sessions and their associated capacities are determined on a long-term cyclic weekly basis. At this level, the problem is formulated as an adjustable robust optimization model denoted as the *robust clinic scheduling and capacity planning problem* (R-CSCPP). It allows hospital managers to plan with inaccurate information regarding the number of arriving patients to each clinic. At the second level, we integrate the PSP with physicians rescheduling decisions as a two-stage stochastic model and denote it as the *stochastic physician scheduling problem* (SPSP).

The main contributions of our paper are the following. The first one is to propose a comprehensive physician planning framework for polyclinics including strategic, tactical and operational levels while incorporating the uncertainty in demand treatment times. The second contribution revolves around formulating the polyclinic capacity planning problem as an adjustable robust optimization model that provides work schedules of clinics under demand uncertainty. We also develop an *implementer-adversary* (I/A) exact algorithm to solve the aforementioned model. The third contribution is to introduce the physician scheduling problem under treatment time uncertainty, which is stated as a two-stage integer stochastic program. This model is unique in terms of first-stage and corrective (recourse) actions. More precisely, while the first-stage decisions focus on the assignment of physicians to regular or on-call duties in different shifts, the second stage (recourse)

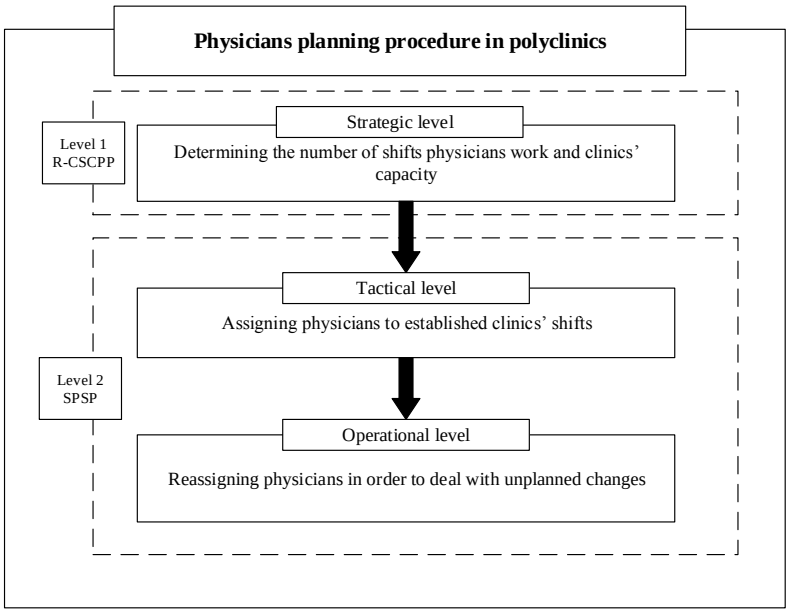

Figure 1: Physicians planning procedure in polyclinics

actions encompass calling on-call doctors and assessing patients during extended shifts in cases where treatment times are stretched-out. The *sample average approximation* (SAA) scheme is used to obtain feasible solutions with small statistical optimality gaps. Finally, the fourth contribution is to develop a Monte-Carlo simulation to demonstrate the importance and potential cost-savings of our proposed planning framework using real data from a polyclinic in a university health center in Montreal, Canada.

The remainder of this paper is structured as follows. Section 2 gives a concise review of the related literature on personnel scheduling problems, with an emphasis on applications in health care. In Section 3, we first present the CSCPP and its robust counterpart and describe the I/A algorithm used for solving it. In Section 4, we introduce the deterministic PSP and extend it to the two-stage SPSP. We then describe the SAA scheme used for solving the stochastic model. In Section 5, using the data provided by the university health center we evaluate the performance of the proposed solution methodologies and compare the generated schedules with the results of a deterministic model that integrates clinics and physicians scheduling into a single level procedure. Finally, concluding remarks are given in Section 6.

2. Literature Review

Workforce allocation and personnel scheduling problems have been widely studied in the literature. Numerous applications arising in different service industries have been considered such as telephone operators, flight crews, bus drivers, and physicians and nurses. We refer to Ernst et al. (2004), Burke et al. (2004), Van den Bergh et al. (2013), and Erhard et al. (2018), for review papers on these classes of problems from an operational research perspective.

Beaulieu et al. (2000) propose the first mixed-integer program (MIP) to model a PSP arising in the emergency room (ER) of a major hospital in Montreal. They divide scheduling constraints into two main categories: *compulsory* and *flexible*. Rousseau et al. (2002) argue that a combination of constraint programming with local search can be a promising generic method to a wide variety of PSPs. Carter & Lapierre (2001) study physicians of ERs at six major hospitals in Canada and propose a tabu search algorithm for assigning physicians to work shifts. Brunner (2010); Brunner et al. (2009, 2010); Stolletz & Brunner (2012) study the problem of flexible scheduling of physicians in which the shifts have variable starting time and duration. Gunawan & Lau (2013) and Van Huele & Vanhoucke (2014) integrate scheduling physicians with other types of decision problems (i.e. duty and surgery scheduling) and consider resource capacity constraints (e.g. available operation rooms, available recovery beds). Tohidi et al. (2017) propose an integrated clinic and physician scheduling problem in polyclinics in a deterministic context.

Some studies focus on a single level manpower planning (e.g., staffing or tactical planning), and consider the variability in daily demand or personnel's capacity. Bard & Purnomo (2004, 2005a,b) present a reactive and real-time scheduling approach to adjust midterm schedules on a daily-basis in response to fluctuations in demand. They apply their method as a recourse decision-making tool for the assignment of nurses to work shifts in a hospital in the U.S. Hur et al. (2004) also investigate real-time work schedule adjustment decisions, and provide heuristics to solve the considered problem. Easton & Goodale (2005) analyze different strategies (e.g. cross-trained workers, overtime, call-in employee, and temporary workers) for scheduling personnel with unplanned absenteeism in a real-time decision-making framework. Campbell (1999) shows cross-trained workers are beneficial to service centers with multiple departments where the demand is subject to uncertainty. Wright & Mahar (2013) investigate on centrally scheduling cross-trained nurses across multiple departments in two hospitals in the U.S. They determine the likelihood of violating the minimum nurse-to-patient ratios through queueing methods.

Other studies aim at integrating different levels of manpower planning, and analyzing the de-

cisions to be taken at each level. Abernathy et al. (1973) introduce a comprehensive three-level manpower planning procedure, which includes policy, staffing, and scheduling decisions, for service industries with demand fluctuations. They present an iterative solution methodology as well as a chance constraint-based method to solve this problem. Venkataraman & Brusco (1996) propose an integrated staffing and scheduling action plan that allows recursion between the staffing and scheduling models and thus enables management to rapidly evaluate the impact of both staffing and scheduling policies. Wright et al. (2006) merge nurse staffing and scheduling by developing a model that incorporates nurse-to-patient ratios and controls the amount of work given to each nurse. Wright & Bretthauer (2010) study two-phase scheduling and rescheduling of cross-trained nurses in a hospital in the U.S. The authors propose a procedure that assigns the nurses to shifts over a mid-term planning horizon in the first phase. In the second phase, the schedules generated by the first phase are adjusted according to the demand at the beginning of each shift. Maenhout & Vanhoucke (2013) integrate staffing and scheduling decisions in a nurse scheduling problem and demonstrate that the schedules generated by the integrated approach are preferable in terms of cost and personnel satisfaction. Ingels & Maenhout (2015) propose a two-phase framework for tackling personnel scheduling problems under demand and employees' availability uncertainty. In the first phase, they assign employees to shifts as regular and reserved workforce. In the second phase, they simulate the random processes, and if needed, optimize the utilization of reserved workforce via an optimization model. They suggest a sequential solution method where the output of the first phase is inserted to the operational phase.

A handful of studies integrate some stages of manpower planning into a two-stage stochastic programming model. Campbell (2011) analyzes staffing and scheduling of cross-trained workers in a service industry with multiple departments and formulates the problem as a two-stage stochastic program. The first stage decisions are to assign workers to work tours, and the recourse actions are to assign cross-trained workers. Bard et al. (2007) tackle the workforce planning at one of the U.S. post service distribution centers by using stochastic programming. The staffing decisions regarding full time and part time workers are made in the first stage, and the final allocations of workers to the shifts and decisions regarding the assignment of additional resources are made during the second stage. Punnakitikashem et al. (2008) formulate the workload assignment of nurses at a hospital in the U.S. as a two-stage stochastic program. The authors consider the first stage decisions to be the assignment of patients to nurses. The recourse actions are considered as the amount of direct or indirect care performed by each nurse. Punnakitikashem et al. (2013) extend the problem of

(Punnakitikashem et al., 2008) by considering the nurse staffing problem along with the workload assignment problem. The authors formulate a two-stage stochastic program in which the staffing decisions are made in the first stage, and in the second stage the duty assignments are executed with respect to outcomes of uncertain parameters. Zhu & Sherali (2009) formulate a multi-category workforce planning with recruitment capacity constraints under demand uncertainty as a two-stage stochastic program. The staffing and allocation decisions are considered as first stage decisions, and the workload assignment decision are considered as recourse actions. Kim & Mehrotra (2015) formulate the scheduling and rescheduling of nurses in polyclinics as a two-stage stochastic program, and propose a formulation which defines the convex hull of the second-stage MIP.

The above literature review indicates the paucity of research on incorporating uncertainty into physician scheduling problems. While the uncertainty has been only considered in some stages of workforce planning (e.g., strategic or tactical level) in service industries, to the best of our knowledge, no prior contribution exists that investigates the integration of the three stages of manpower planning under uncertainty as a unique decision framework. This article, thus, aims to fill this void in the literature given the fact that adopting a deterministic approach could significantly affect the economic viability of the schedule.

3. The Clinic Scheduling and Capacity Planning Problem

In this section we first introduce the deterministic CSCPP. We then incorporate the uncertainty in clinics' demand and present a robust optimization formulation and an exact solution algorithm for solving this problem.

3.1. Problem Definition and Formulation

Let C, I, J, and K denote the sets of clinics, physicians, days and shifts per day, respectively. The set of subsets of interdisciplinary clinics is denoted by T. Each clinic in subset $t \in T$ must be scheduled simultaneously with other clinics in the same subset for at least F_t shifts. The number of patients that can be assessed by a physician in clinic $c \in C$ in one shift is denoted as VPS_c. Let $\hat{H}_i$ and $\hat{Y}_c$ denote the minimum number of shifts that must be assigned to physician $i \in I$ and clinic $c \in C$, respectively. For every $c \in C$, PU_c is a measure of fairness that limits the total difference among the physicians in clinic c in terms of number of assigned shifts. For each $i \in I$, f_{1i} represents the cost of assigning physician i to a shift. Let f_3 be the cost of clinics' staff (non-physicians) per shift. Let R_c be the number of rooms required for non-physician staff in clinic $c \in C$ in a shift. We

also assume that every physician needs one examination room when he/she works. Let f_2 denote the cost of using a room and NWV_c the number of patients per week visiting clinic $c \in C$. We assume that service ability of the polyclinic may not fully satisfy the demand. Therefore, when the demand is more than the planned capacity, patients are not admitted and must be transferred to other hospitals, which incurs a cost of f_4. The CSCPP consists of determining: (i) the set of shifts that each clinic is open during the week, (ii) a tentative work schedule for physicians affiliated with each clinic, (iii) the capacity of each clinic and the number of patients that can be assessed in each shift, and (iv) the number of rooms allocated to each shift, such that the demand, restrictions on interdisciplinary clinics, and physicians requirements are satisfied. The goal is to minimize the total cost of resources (physicians, rooms and clinics' staff) and the cost of rejected patients.

For each $j \in J$ and $k \in K$, we define clinic assignment variables y_{cjk} equal to 1 if and only if clinic $c \in C$ is assigned to shift k on day j. Similarly, for every $j \in J, k \in K$, we define physician assignment variables x_{ijk} equal to 1 if and only if physician $i \in I$ is assigned to day j, shift k. For all $t \in T$, $j \in J$, $k \in K$, g_{tjk} is an auxiliary binary decision variable that equals to 1 if and only if all clinics in subset t are assigned to shift k on day j. For each clinic $c \in C$, nd_c is an integer variable equal to the number of rejected patients. For every physician $i \in I$, h_i is an integer variable representing the total number of shifts assigned to physician i. Finally, r is an integer variable denoting number of required rooms during the week. Using these sets of variables, the CSCPP can be formulated as follows:

$$\text{minimize} \sum_{i \in I} f_{1i} h_i + f_2 r + \sum_{c \in C} \sum_{j \in J} \sum_{k \in K} f_3 y_{cjk} + \sum_{c \in C} f_4 nd_c \tag{1}$$

$$\text{subject to} \sum_{i \in I_c} \sum_{j \in J} \sum_{k \in K} VPS_c x_{ijk} + nd_c \geq NWV_c \qquad c \in C \tag{2}$$

$$\sum_{i \in I} x_{ijk} + \sum_{c \in C} R_c y_{cjk} \leq r \qquad j \in J, k \in K \tag{3}$$

$$\sum_{i \in I_c} x_{ijk} \geq y_{cjk} \qquad c \in C, j \in J, k \in K \tag{4}$$

$$\sum_{c \in C | c \in t} y_{cjk} \geq |t| g_{tjk} \qquad t \in T, j \in J, k \in K \tag{5}$$

$$\sum_{j \in J} \sum_{k \in K} g_{tjk} \geq F_t \qquad t \in T \tag{6}$$

$$\sum_{j \in J} \sum_{k \in K} x_{ijk} \leq h_i \qquad i \in I \tag{7}$$

$$\sum_{j \in J} \sum_{k \in K} x_{ijk} \geq \hat{H}_i \qquad i \in I \tag{8}$$

$$\sum_{j \in J} \sum_{k \in K} y_{cjk} \geq \hat{Y}_c \qquad c \in C \tag{9}$$

$$h_i - h_{i'} \leq PU_c \qquad c \in C, i, i' \in I_c \tag{10}$$

$$\sum_{k \in K} x_{ijk} \leq 1 \qquad i \in I, j \in J \tag{11}$$

$$x_{ijk},\ y_{cjk},\ g_{tjk} \in \{0, 1\} \tag{12}$$

$$h_i,\ r,\ nd_c \in \mathbb{Z}^+. \tag{13}$$

The first three terms of the objective function are the total resource cost (i.e. cost of physicians, rooms, and clinics) and the last term is the total cost of rejected patients. Constraints (2) ensure that each clinic demand (i.e. visiting patients) is either served by physicians or rejected. Constraints (3) specify the number of required rooms in each shift. Constraints (4) link the assignment of physicians with corresponding clinics. Constraints (5) and (6) guarantee that every clinic in each subset of interdisciplinary clinics is assigned with other clinics in the subset for the predefined number of shifts. Constraints (7) record the total number of shifts that each physician works. Constraints (8) ensure that physicians are assigned to their minimum required shifts. Similarly, constraints (9) assure the minimum number of shifts that each clinic must be open. Constraints (10) restrict the difference in the number of assigned shifts between any two physicians of a clinic. Constraints (11) prevent physicians from working more than one shift per day. Finally, constraints (12) and (13) are the standard integrality and non-negativity constraints. It should be noted that the physicians' schedule that is provided by the CSCPP model can be considered as a rough estimate of the actual shift assignment that will be provided by the second-level SPSP model.

3.2. A Robust Formulation for the CSCPP

We recall that in practice, the weekly demand NWV_c used in CSCPP is usually not known in advance. At the strategic decision-making level, where it is not possible to have access to accurate information on the number of arriving patients to clinics, estimating the probability distribution of the demand is not straightforward. Therefore, we follow a more realistic approach in which weekly demands NWV_c are modeled as an interval of uncertainty. By estimating the upper and lower

bound of the demand for each clinic, the uncertain set of demand for clinic $c \in C$ is defined as:

$$\Omega_c = \left\{ NWV_c \in \mathbb{Z}^+ : \underline{NWV_c} \leq NWV_c \leq \overline{NWV_c} \right\},$$

where $\mathbb{Z}^+$ is the set of positive integers and $\overline{NWV_c}$, $\underline{NWV_c} \in \mathbb{Z}^+$. In practice, it is very unlikely that all uncertain parameters simultaneously achieve their upper bounds. Hence, our goal is to formulate the robust counterpart of the CSCPP, denoted as R-CSCPP, such that the plan's cost is as small as possible when the demand takes its worst case scenario within a certain level of conservatism. To control the degree of conservatism of the solution, in a similar fashion to the approach suggested by Bertsimas & Sim (2004), we use a budget of uncertainty (Γ). Even though each clinic's demand may vary within the corresponding interval, we restrict the overall demand of all clinics to Γ by adding the following constraint:

$$\sum_{c \in C} NWV_c \leq \Gamma.$$

Let $NWV = (NWV_1, \ldots, NWV_{|C|})$ denote the vector of demands for each of the $|C|$ clinics. Then, the demand uncertainty set Ω is stated as follows:

$$\Omega = \left\{ NWV : NWV_c \in \Omega_c, \forall c \in C, \sum_{c \in C} NWV_c \leq \Gamma \right\},$$

where we denote each possible outcome of the uncertainty set by $\omega \in \Omega$. The R-CSCPP is formulated as follows:

$$\text{minimize } \eta$$

$$\text{subject to (3)–(13)}$$

$$\eta \geq \sum_{i \in I} f_{1i} h_i + f_2 r + \sum_{c \in C} \sum_{j \in J} \sum_{k \in K} f_3 y_{cjk} + \sum_{c \in C} f_4 nd_c^{\omega} \quad \omega \in \Omega \tag{14}$$

$$\sum_{i \in I_c} \sum_{j \in J} \sum_{k \in K} VPS_c x_{ijk} + nd_c^{\omega} \geq NWV_c^{\omega} \quad c \in C,\ \omega \in \Omega \tag{15}$$

$$\eta,\ nd_c^{\omega} \in \mathbb{R}^+. \tag{16}$$

The R-CSCPP seeks a clinic schedule and capacity plan such that the cost is minimized under worst-case demand scenarios within the uncertainty set Ω. Constraints (14) formulate the total

cost of each scenario, in which η is a decision variable that captures the maximum cost caused by the worst case scenario. Constraints (15) are the demand coverage constraint under different scenarios (ω) from the uncertainty set Ω. Observe that although Ω is a finite set, its size grows exponentially in C. The R-CSCPP thus involves a huge number of constraints (14) and (15), so many that we cannot add them explicitly while using a general purpose solver. In the next section we describe a cutting plane method that is used to efficiently solve this problem.

3.3. An Implementor/Adversary Algorithm for the R-CSCPP

We use the so-called *implementor/adversary* algorithm introduced by Bienstock (2007) to solve the R-CSCPP. It has been successfully applied to solve various robust optimization problems (see for instance, Holte & Mannino, 2013; Tang & Wang, 2015). This algorithm decomposes the original problem into two simpler ones: the *implementor* and the *adversary* sub-problem. The main idea of the method is to initially consider a small subset of scenarios (and consequently a sub-set of constraints and variables) in the implementor problem, and to generate new ones through the adversary problem if needed. It can also be interpreted as if the implementor problem tries to find a solution that minimizes the cost, but the adversary problem generates a scenario that makes the solution infeasible, if any.

Let $\tilde{\Omega} \subset \Omega$ be a subset of the uncertainty set. The restricted implementor problem (I-CSCPP) is a relaxation of the R-CSCPP problem that considers a subset constraints (14) and (15), the ones associated with $\tilde{\Omega}$. Let $\tilde{x} = (\tilde{h_i}, \tilde{r}, \tilde{x_{ijk}}, \tilde{y_{cjk}})$ be the optimal solution to the restricted implementor problem and $\eta_{\tilde{x}}$ be the corresponding objective function value, which provides a valid lower bound (L) on the optimal solution value of R-CSCPP. For any $\omega \in \Omega$, let $f(\tilde{x}, \omega)$ be the associated cost of $\tilde{x}$ and $n\tilde{d}_c^{\omega}$ be the number of rejected patients in scenario ω. If $f(\tilde{x}, \omega) \leq \eta_{\tilde{x}}$, and $n\tilde{d}_c^{\omega}$ ensures the feasibility of constraints (15) for all $\omega \in \Omega$, then $\tilde{x}$ is the optimal solution to the original R-CSCPP with optimal solution value $\eta_{\tilde{x}}$. Otherwise, there exists at least one scenario $\omega \in \Omega \setminus \tilde{\Omega}$ that leads to a violated inequality of the type (14) or (15).

In order to prove optimality, or to determine which scenario leads to violated inequalities, we solve the adversary problem (i.e. separation problem) to find a scenario ω^* that maximizes the cost $f(\tilde{x}, \omega)$. In particular, for a given solution $\tilde{x} = (\tilde{h_i}, \tilde{r}, \tilde{x_{ijk}}, \tilde{y_{cjk}})$ of the restricted implementor problem with scenario set $\tilde{\Omega}$, the adversary problem of the R-CSCPP can be stated as the following

MIP:

$$\text{maximize} \sum_{c \in C} f_4 nd_c + \sum_{i \in I} f_{1i} \tilde{h}_i + f_2 \tilde{r} + \sum_{c \in C} \sum_{j \in J} \sum_{k \in K} f_3 \tilde{y_{cjk}} \tag{17}$$

$$\text{subject to} \quad \underline{NWV_c} \leq nwv_c \leq \overline{NWV_c} \qquad c \in C \tag{18}$$

$$\sum_{c \in C} nwv_c \leq \Gamma \tag{19}$$

$$nd_c \leq M z_c + nwv_c - \sum_{i \in I_c} \sum_{j \in J} \sum_{k \in K} VPS_c \tilde{x_{ijk}} \qquad c \in C \tag{20}$$

$$nd_c \leq M (1 - z_c) \qquad c \in C \tag{21}$$

$$z_c \in \{0, 1\} \tag{22}$$

$$nd_c \in \mathbb{R}^+ \tag{23}$$

$$nwv_c \in \mathbb{Z}^+, \tag{24}$$

where nwv_c are integer variables representing the demand for clinic $c \in C$, and M is a large constant. Constraints (18) and (19) ensure that for each $c \in C$, the clinic demand value nwv_c is within the uncertainty set Ω. For each $c \in C$, z_c are binary decision variables equal to 1 if the demand of clinic c is more than the capacity allocated according to schedule $\tilde{x}$ (i.e., $nwv_c - \sum_{i \in I_c} \sum_{j \in J} \sum_{k \in K} VPS_c \tilde{x_{ijk}} < 0$), and 0 otherwise. Constraints (20) and (21) ensure that for each $c \in C$, nd_c represents the number of rejected patients associated with current assignments of physicians $\tilde{x_{ijk}}$.

The optimal solution value of the adversary problem, denoted by $f(\tilde{x}, \omega^*)$, provides a valid upper bound (U) on the optimal solution value of R-CSCPP. If there is a new scenario ω^* that leads to a violated inequality, it is added to $\tilde{\Omega}$, and the associated implementor problem is resolved once again. This procedure is implemented in an iterative manner until the gap between the upper and lower bounds reaches a predefined threshold value ϵ. Our implementor/adversary algorithm is summarized in Algorithm 1.

4. The Physician Scheduling Problem

In this section, we first provide a formal definition and an MIP formulation for the deterministic variant of the PSP. We then show how we incorporate the uncertainty in patient treatment times into the PSP and formulate it as a two-stage integer stochastic program. Finally, we present an

Algorithm 1 Implementor/adversary algorithm for R-CSCPP

Initialization: $\tilde{\Omega} = \emptyset$, $L = -\infty$, $U = +\infty$
while $U - L > \epsilon$ **do**
  Solve the restricted implementer problem with $\tilde{\Omega}$ to obtain $\tilde{x}$ and L.
  $L \leftarrow \eta_{\tilde{x}}$
  Solve the adversary problem (17) - (24) to obtain ω^*
  if $U > f(\tilde{x}, \omega^*)$ **then**
    $U \leftarrow f(\tilde{x}, \omega^*)$
  end if
  $\tilde{\Omega} \leftarrow \tilde{\Omega} \cup \omega*$
end while

SAA scheme to generate feasible solutions and obtain a statistical estimation of their optimality gap.

4.1. Problem Definition and Formulation

Consider the sets and parameters previously defined for the CSCPP. We denote as $(\tilde{h_i}, \tilde{r}, \tilde{x_{ijk}}, \tilde{y_{cjk}})$ the optimal solution of the R-CSCPP, which acts as an input to the PSP. For each $j \in J$ and $k \in K$, let D_{cjk} denote the number of patients visiting clinic $c \in C$, and R_{jk} the number of available examination rooms. For every physician $i \in I$, H_i is the maximum number of on-duty shifts. The aforementioned parameters are hence calculated as follows:

$$D_{cjk} = \sum_{i \in I_c} VPS_{cjk} \tilde{x_{ijk}} \quad c \in C, j \in J, k \in K \tag{25}$$

$$R_{jk} = \tilde{r} - \sum_{c \in C} R_c \tilde{y_{cjk}} \quad j \in J, k \in K \tag{26}$$

$$H_i = \tilde{h_i} \quad i \in I. \tag{27}$$

Let NJ_i be the set of days that physician $i \in I$ is not available to work. TK_i are the percentage of total working shifts that physician $i \in I$ must be assigned to the shifts that he/she prefers to work in. In order to reflect physicians preferences for working in a specific shift, for every $j \in J$ and $k \in K$, we define PR_{ijk} equal to 1, if physician $i \in I$ prefers to work on day j, shift k, 0 otherwise. The PSP consists of assigning physicians to the shifts on a weekly basis, such that the demand per shift, examination room capacity, and physicians preferences are satisfied while the total cost of physicians is minimized. Physician assignment variables x_{ijk} are defined as in the CSCPP. The PSP can be formulated as the following MIP:

$$\text{minimize} \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} f_{1i} x_{ijk} \tag{28}$$

$$\text{subject to} \sum_{i \in I_c} VPS_{cjk} x_{ijk} = D_{cjk} \qquad c \in C,\ j \in J,\ k \in K \tag{29}$$

$$\sum_{i \in I} x_{ijk} \leq R_{jk} \qquad j \in J,\ k \in K \tag{30}$$

$$\sum_{j \in J} \sum_{k \in K} (x_{ijk} - x_{i'jk}) \leq PU_c \qquad c \in C,\ i, i' \in I_c \tag{31}$$

$$\sum_{k \in K} x_{ijk} = 0 \qquad i \in I,\ j \in NJ_i \tag{32}$$

$$\sum_{j \in J} \sum_{k \in K} x_{ijk} \leq H_i \qquad i \in I \tag{33}$$

$$\sum_{j \in J} \sum_{k \in K} (TK_i x_{ijk} - PR_{ijk} x_{ijk}) \leq 0 \qquad i \in I \tag{34}$$

$$\sum_{k \in K} x_{ijk} \leq 1 \qquad i \in I,\ j \in J \tag{35}$$

$$x_{ijk} \in \{0, 1\}. \tag{36}$$

The objective function represents the total cost of physicians. Constraints (29) ensure that adequate number of physicians are assigned to each shift to visit the scheduled patients. Constraints (30) restrict the number of physicians in each shift to the number of available examination rooms. Constraints (31) restrict the difference in the number of assigned shifts between any two physicians of each clinic. Constraints (32) prevent assigning physicians to the shifts that they are not available to work. Constraints (33) limit the number of shifts that a physician can be on-duty to their corresponding maximum workload. Constraints (34) ensure that a certain percentage of physicians' workload are assigned according to their preferences. Similarly, constraints (35) prohibit physicians from working more than one shift per day. Lastly, constraints (36) are the standard integrality constraints.

4.2. Dealing with Patient Treatment Time Uncertainty

We now consider the number of patients that a physician can assess per shift (VPS_{cjk}) to be an uncertain parameter. In particular, we assume that these parameters can be represented with a finite set of scenarios with known probability. This is a realistic assumption due to the fact that only few classes of patients are visited in the polyclinic under discussion for which the probability distribution of treatment times can be realistically estimated. We denote Ξ as the set of random scenarios, where $\Xi = \{\xi_1, \ldots, \xi_{|\Xi|}\}$, and P^{ξ} the probability of scenario $\xi \in \Xi$. In this context, we assume that the physicians schedule must be determined at the beginning of the

planning horizon (week) without full knowledge of the list of patients that must be served during each shift. In other words, the treatment times are not revealed to the decision maker at this stage. The patients' treatment times, on the contrary, will be finalized only few hours before the beginning of a shift. At this point, the decision maker would have a clearer knowledge on the treatment times by referring to the initial clinical assessments, which are done prior to the patients' appointments with the doctors. Given the fact that physicians have been already assigned to shifts, depending on the patients mix, it is possible that all patients could not be assessed during the regular duration of a shift. In order to avoid this situation, three types of corrective (recourse) actions with associated costs are considered: *i)* assigning a certain number of physicians as on-calls, *ii)* serving patients during extended (over-time) shifts, and *iii)* requesting additional examination rooms. It should be noted that the third recourse action is a consequence of the first one in the sense that additional physicians require more resources. On the other hand, the initial schedule might lead to an idle time for physicians due to low treatment times required for patients scheduled during a shift. The under-utilization of physicians would also incur a cost for hospital's management.

Given the above assumptions, the stochastic PSP can be reformulated as a multi-stage stochastic program with recourse (MSP) where each stage denotes a shift during the planning horizon. The first stage decisions (those made at the beginning of the week) revolve around assigning physicians to either regular or on-call duties over all shifts during the planning horizon. On the contrary, calling on-call physicians to work during a shift, assigning physicians to extended (over-time) shifts, and requesting extra examination rooms are the decisions that are made as corrective actions at the beginning of each shift (stage) for different scenarios in terms of patient treatment times. Nonetheless, it is noteworthy that the aforementioned recourse decisions made in each stage (shift) do not affect the decision in proceeding stages (shifts) which is a common feature in the majority of personnel scheduling problems (see e.g., (Bard et al., 2007)). In other words, the non-anticipativity condition (NAC) should only be satisfied for the first-stage decisions. Consequently, the aforementioned MSP model can be transformed into a two-stage integer stochastic problem with recourse, denoted by SPSP.

Consider the sets and parameters defined for the PSP. Similar to assigning on-duty physicians, for each $i \in I$, $j \in J$, $k \in K$, we define x'_{ijk} equal to 1, if and only if physician i is assigned on day j, shift k, as on-call. Let f_5 be the cost of assigning a physician as on-call per shift. Let VPS^{ξ}_{cjk} denote the number of patients that can be visited in clinic $c \in C$ during shift $j \in J$ by physician $k \in K$ under scenario $\xi \in \Xi$. As mentioned, if the number of on-duty physicians is not enough for

a certain scenario during a shift, on-call physicians are called to work, and additional examination rooms are prepared to accommodate them. For each $i \in I$, $j \in J$, $k \in K$, $\xi \in \Xi$, we define s^{ξ}_{ijk} equal to 1, if and only if on-call physician i is called to work on day j, shift k, under scenario ξ. Also, for each $j \in J$, $k \in K$, $\xi \in \Xi$, the decision variable er^{ξ}_{jk} is the number of additional examination rooms prepared for shift k on day j, under scenario ξ. f_6 and f_7 represent the costs for calling on-call physicians to work and preparing additional examination room, respectively.

For each $c \in C$, $j \in J$, $k \in K$, $\xi \in \Xi$, the recourse decision ot^{ξ}_{cjk} represents the number of patients visiting clinic c in shift k on day j who are processed during overtime shift, under scenario ξ, and f_8 is the corresponding cost. To measure the under-utilization of on-duty physicians under some scenarios, we define the recourse decision oa^{ξ}_{cjk} for each $c \in C$, $j \in J$, $k \in K$, $\xi \in \Xi$ that represents the amount of time physicians in clinic c are idle in shift k on day j, under scenario ξ. f_9 is the cost of idle physicians. The objective of the SPSP is to minimize the cost of on-duty and on-call physicians, plus the expected cost of recourse decisions over all treatment time scenarios. The SPSP can be formulated as follows:

$$\text{minimize} \sum_{i\in I}\sum_{j\in J}\sum_{k\in K}(f_{1i}x_{ijk} + f_5 x'_{ijk}) + \sum_{\xi\in\Xi} P^{\xi}\Bigg(\sum_{i\in I}\sum_{j\in J}\sum_{k\in K} f_6 s^{\xi}_{ijk} + \sum_{j\in J}\sum_{k\in K} f_7 er^{\xi}_{jk} + \sum_{c\in C}\sum_{j\in J}\sum_{k\in K}(f_8 ot^{\xi}_{cjk} + f_9 oa^{\xi}_{cjk})\Bigg) \tag{37}$$

subject to(33) – (36)

$$\sum_{i\in I_c} VPS^{\xi}_{cjk}(x_{ijk} + s^{\xi}_{ijk}) + ot^{\xi}_{cjk} - VPS^{\xi}_{cjk} oa^{\xi}_{cjk} = D_{cjk} \qquad c \in C,\ j \in J, k \in K, \xi \in \Xi \tag{38}$$

$$\sum_{i\in I}(x_{ijk} + s^{\xi}_{ijk}) - er^{\xi}_{jk} \le R_{jk} \qquad j \in J, k \in K, \xi \in \Xi \tag{39}$$

$$\sum_{j\in J}\sum_{k\in K}(x_{ijk} + x'_{ijk} - x_{i'jk} - x'_{i'jk}) \le PU_c \quad c \in C, i, i' \in I_c \tag{40}$$

$$\sum_{k\in K}(x_{ijk} + x'_{ijk}) \le 0 \qquad i \in I, j \in NJ_i \tag{41}$$

$$x_{ijk} + x'_{ijk} \le 1 \qquad i \in I, j \in J, k \in K \tag{42}$$

$$\sum_{k\in K}(x_{ijk} + s^{\xi}_{ijk}) \le 1 \qquad i \in I, j \in J, \xi \in \Xi \tag{43}$$

$$s^{\xi}_{ijk} \le x'_{ijk} \qquad i \in I,\ j \in J, k \in K, \xi \in \Xi \tag{44}$$

$$x'_{ijk},\ s^{\xi}_{ijk} \in \{0,1\} \tag{45}$$

$$ot^{\xi}_{cjk},\ er^{\xi}_{jk} \in \mathbb{Z}^{+} \tag{46}$$

$$oa^{\xi}_{cjk} \in \mathbb{R}^{+}. \tag{47}$$

The first two terms of the objective are the total cost of assigning physicians as on-duty and on-call, respectively. The last four terms are the expected cost of calling on-call physicians to work, preparing additional examination rooms, processing patients in overtime shifts, and physicians' idle time over all scenarios. Constraints (38) specify the number of patients processed by on-duty or on-call physicians, the number of patients assessed during overtime, and the amount of physicians' idle time during each shift for all scenarios. Constraints (39) are the examination rooms capacity constraints. Constraints (40) ensure fair assignments of on-duty and on-call shifts among the physicians of each clinic. Constraints (41) forbid the on-call or on-duty assignments to the shifts that physicians are not available to work in. Constraints (42) prevent simultaneous assignments of physicians as on-call and on-duty to any given shift. Constraints (43) guarantee that physicians work no more than one shift per day. Constraints (44) assure that physicians can be called to work in a shift only if they have been already assigned as on-call to that shift. Finally, constraints (45) $-$ (47) are the standard integrality and non-negativity constraints.

4.3. The Sample Average Approximation Scheme

The SPSP is a mixed-integer programming model that is hard to solve if the number of scenarios is large. More precisely, considering the variety of patients' treatment times in each clinic, a fairly large number of scenarios can be expected. However, by using Monte-Carlo sampling techniques such as the SAA scheme (Shapiro & Homem-de Mello, 2000; Mak et al., 1999; Kleywegt et al., 2002; Verweij et al., 2003), we can obtain an approximate solution by randomly selecting a subset $\{\xi_1, \ldots, \xi_N\}$ of scenarios from the original uncertainty set. In the SAA method, the expected recourse function is approximated by the sample average function by considering N randomly selected scenarios from the original uncertainty set. The SAA problem is formulated as follows:

$$\begin{aligned}
\text{minimize} & \sum_{i\in I}\sum_{j\in J}\sum_{k\in K}\left(f_{1i}x_{ijk} + f_5 x'_{ijk}\right) \\
& + \frac{1}{N}\sum_{n=1}^{N}\sum_{j\in J}\sum_{k\in K}\left(\left(\sum_{i\in I} f_6 s^n_{ijk} + f_7 er^n_{jk}\right) + \sum_{c\in C}\left(f_8 ot^n_{cjk} + f_9 oa^n_{cjk}\right)\right) \\
\text{subject to} & (33)-(36), (38)-(47).
\end{aligned}$$

The SAA method proceeds by solving the SAA problem repeatedly for M independent samples of size N. Let denote the corresponding optimal solutions by $\tilde{x}_N^1, \ldots, \tilde{x}_N^M$ and their objective values by $z_N^1, \ldots, z_N^M$. Let $\bar{Z}_N$ be the average of the M objective values:

$$\bar{Z}_N = \frac{1}{M} \sum_{m=1}^{M} z_N^m. \tag{48}$$

On one hand, it is well-known that $E[\bar{Z}_N] \leq z^*$ (Mak et al., 1999). Therefore, $\bar{Z}_N$ provides an unbiased estimate for a lower bound on the optimal solution value of the SPSP. On the other hand, for any feasible point $\tilde{x}$, it is obvious that the corresponding objective value of the SPSP for any larger scenario set of size N' provides a statistical upper bound for z^*. Hence, an upper bound can be estimated by:

$$\begin{aligned}
\tilde{Z}_{N'}(\tilde{x}) = \min & \sum_{i\in I}\sum_{j\in J}\sum_{k\in K}(f_{1i}\tilde{x_{ijk}} + f_5\tilde{x_{ijk}}') + \frac{1}{N'}\sum_{n=1}^{N'}\{\sum_{j\in J}\sum_{k\in K}(\sum_{i\in I} f_6 s_{ijk}^n \\
& + f_7 er_{jk}^n) + \sum_{c\in C}\sum_{j\in J}\sum_{k\in K}(f_8 ot_{cjk}^n + f_9 oa_{cjk}^n)\} \\
\text{subject to} & (33) - (47),
\end{aligned}$$

where N' is typically a quite large sample size and independent of N. Then, for any feasible solution $\tilde{x}$ we have that $E[\tilde{Z}_{N'}(\tilde{x})] \geq z^*$. Variances of the lower bound and upper bound estimators are calculated as follows:

$$\begin{aligned}
\tilde{\sigma}^2_{\bar{Z}_N} =& \frac{1}{(M-1)M}\sum_{m=1}^{M}(z_N^m - \bar{Z}_N)^2, \\
\tilde{\sigma}^2_{\tilde{Z}_{N'}(\tilde{x})} =& \frac{1}{(N'-1)N'}\sum_{n=1}^{N'}\Bigg(\sum_{i\in I}\sum_{j\in J}\sum_{k\in K}(f_1\tilde{x_{ijk}} + f_5\tilde{x_{ijk}}') + \sum_{j\in J}\sum_{k\in K}(\sum_{i\in I} f_6 s_{ijk}^n \\
& + f_7 er_{jk}^n) + \sum_{c\in C}\sum_{j\in J}\sum_{k\in K}(f_8 ot_{cjk}^n + f_9 oa_{cjk}^n) - \tilde{Z}_{N'}(\tilde{x})\Bigg)^2,
\end{aligned}$$

respectively. According to Verweij et al. (2003), one should take $\tilde{x}^*$ as the best solution among $\tilde{x}_N^1, \ldots, \tilde{x}_N^M$ candidate solutions, that is:

$$\tilde{x}^* \in \arg\min \tilde{Z}_{N'}(\tilde{x}) : \tilde{x} \in \{\tilde{x}_N^1, \ldots, \tilde{x}_N^M\}. \tag{49}$$

The quality of the solution $\tilde{x}^*$ can be stated by an optimality gap estimate, computed as follows:

$$Gap = \frac{\tilde{Z}_{N'}(\tilde{x}^*) - \bar{Z}_N}{\tilde{Z}_{N'}(\tilde{x}^*)} \times 100. \tag{50}$$

5. Computational results

In this section, we present the results of computational experiments obtained using data from a real case study in a polyclinic in Montreal, Canada. The computational experiments consist of three parts. First, we evaluate the convergence and solution time of the I/A algorithm on the R-CSCPP for the real-case problem. We then investigate on the performance of the SAA method implemented for the SPSP. Using a Monte Carlo simulation, we also compare the cost of schedules obtained from our framework with the cost of the expected value problem (EVP), where random parameters are substituted by their mean values. All experiments were run on an HP server with 20 Intel(R) Xeon(R) CPU E5-2687W v3 @ 3.10GHz processors and 512 GB RAM under Linux environment. All formulations and algorithms were coded in C++, and the associated MIPs were solved using *Concert Technology* of CPLEX 12.7.0.

The studied polyclinic operates five days a week and two shifts every day. Table 1 shows the information of the polyclinic, in which the first two columns are the clinics' identification number and their discipline. The third column is the number of physicians in each clinic. In the fourth and fifth columns, the upper bound and lower bound on the number of visiting patients to each clinic obtained from the historical data are given. Finally, the last three columns contain the minimum, average, and maximum number of patients that can be assessed by one physician in one shift, respectively.

Table 1: Polyclinic information

ID	Clinic	# Physicians	Demand		# Patients/Physician/Shift		
			$\underline{NWV_c}$	$\overline{NWV_c}$	$\underline{VPS_c}$	VPS_c	$\overline{VPS_c}$
1	Breast	19	90	270	2	4	13
2	Urology	15	115	345	3	7	21
3	Hematology	21	95	285	2	4	12
4	Gynecology	7	66	197	7	10	20
5	Hepatology	10	110	330	5	7	12
6	Lung	19	85	255	2	4	10
7	Musculoskeletal	6	35	105	2	5	18
8	Melanoma	9	200	600	12	18	24
9	Upper GI	10	10	30	1	2	4
10	Pain	7	8	23	1	2	3
11	Cancer Rehab.	6	13	38	1	2	3
12	Colorectal	12	75	225	2	4	8
13	Brain	6	15	45	1	2	12

Table 2 contains the information about the interdisciplinary clinics. As the table shows, there

are clusters of interdisciplinary clinics that must be scheduled for a certain number of shifts simultaneously. The first two columns show the clusters' identification number and the interdisciplinary clinics in each cluster, and the third column is the number of shifts that they must be scheduled simultaneously during a week.

Table 2: Interdisciplinary clinics

Clusters	Interdisciplinary groups	Frequency
1	(6) & (10)	4
2	(6) & (11)	4
3	(12) & (5)	4
4	(2) & (10)	3
5	(6) & (8)	1

Without loss of generality, we assume rational values for each cost category according to its priority level. Based on the results of a survey conducted at the hospital, service-level-related costs have the highest priority for the administration, followed by the cost of physicians that are higher than the cost of the other resources. The following cost values are provided in Tables 3 and 4 for the R-CSCPP and the SPSP, respectively. As can be seen in these tables, we assume the costs of rejecting and rescheduling a patient (i.e. f_4 and f_8) are significantly higher than the costs of resources and physicians, as the polyclinic administration would like to process as many patients as possible during regular shifts. Moreover, for the sake of simplicity, we assume equal costs for all physicians.

Table 3: Cost values in the R-CSCPP

Parameter	Value
f_1	5
f_2	1
f_3	1
f_4	50

Table 4: Cost values in the SPSP

Parameter	Value
f_1	5
f_5	2.5
f_6	10
f_7	7.5
f_8	50
f_9	15

5.1. Performance of the I/A on the R-CSCPP

We now present the results of the experiments for evaluating the performance of the I/A algorithm. In order to evaluate the behavior and convergence of the I/A algorithm, the R-CSCPP is solved with different values of the budget of uncertainty Γ within its range. Theoretically, Γ can

range in the interval of $[\sum_{c \in C} \underline{NWV_c}, \sum_{c \in C} \overline{NWV_c}] = [917\ ,\ 2748]$. However, in our case, since a physician assesses multiple patients in a shift (i.e. VPS_c), and the demand constraint is not an equality constraint, the number of processed patients do not equal Γ. Consequently, $\Gamma = 1667$ represents the case in which all clinics' demand are simultaneously at their upper bounds.

The results for $\Gamma = 1067$ are plotted in Figure 2. The algorithm converges after 173 seconds and 20 iterations. For each iteration, the lower bound returned by the implementor (lower dotted line) and the upper bound returned by the adversary (upper solid line) are plotted. The dash line with circle marks shows the optimality gap (i.e. percent deviation of implementor from the best known solution of the adversary problem) on the secondary axis on the right side of the figure. As can be seen, the algorithm requires few number of iterations to converge, and we observe a similar behavior under different Γ values. Table 5 demonstrates the results of the I/A algorithm with

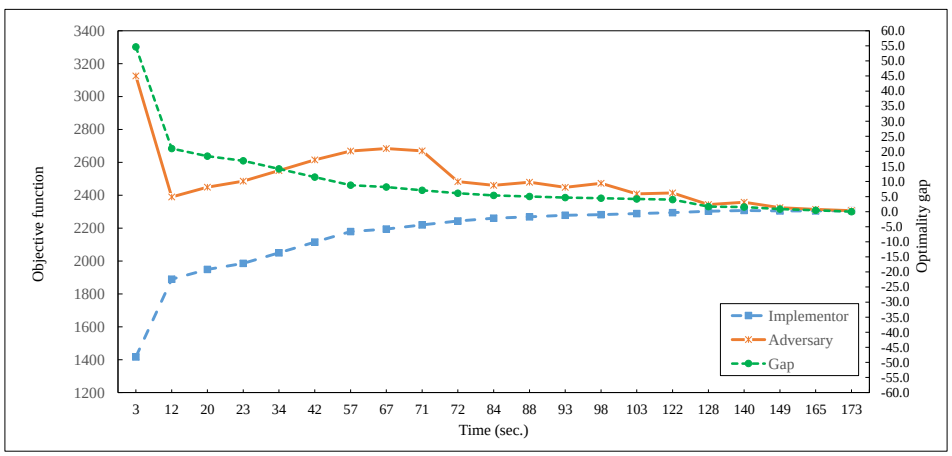

Figure 2: Implementor and adversary $\Gamma = 1067$

different values of Γ. The first column shows the value of Γ selected for the instance. The second column demonstrates the total time in seconds that the algorithm spent in order to converge. The remainder columns give the percentage of the worst case demand served in each clinic, starting from the breast clinic (c_1) to the brain clinic (c_{13}). As can be seen in the table, the algorithm converges in less than 200 seconds for all the instances. In these experiments, in each instance of Γ, we included the scenarios generated by the smaller Γ values in the implementor problem; for example, in the first iteration, the implementor problem of instance $\Gamma = 1017$ includes the scenarios generated by the $\Gamma = 917$ and $\Gamma = 967$. Including those scenarios in the implementor model and warm-starting it improves the convergence time of the algorithm. As the row of $\Gamma = 1667$ shows, 100 percent of

the worst case demand in all clinics is served; therefore, any Γ such that $1667 \leq \Gamma \leq 2780$ will have the same results as $\Gamma = 1667$.

Table 5: Results of the I/A algorithm

Γ	Time	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	c_{13}
917	3	34	34	33	35	34	34	33	35	33	33	35	33	35
967	21	34	34	33	35	34	34	38	35	33	33	35	33	35
1017	107	38	38	38	40	38	39	48	35	40	33	35	39	35
1067	173	66	58	64	75	62	69	90	47	60	50	55	72	48
1117	57	99	90	99	100	94	98	100	65	100	92	95	98	96
1167	31	99	98	99	100	100	100	100	74	100	100	95	98	96
1217	20	99	98	99	100	100	100	100	82	100	100	95	98	96
1267	32	99	98	99	100	100	100	100	91	100	100	95	98	100
1317	35	99	98	99	100	100	100	100	100	100	100	95	98	100
1367	25	100	100	99	100	100	100	100	100	100	100	95	98	100
1417	7	100	100	99	100	100	100	100	100	100	100	95	98	100
1467	6	100	100	99	100	100	100	100	100	100	100	95	98	100
1517	25	100	100	99	100	100	100	100	100	100	100	95	100	100
1567	22	100	100	100	100	100	100	100	100	100	100	95	100	100
1667	5	100	100	100	100	100	100	100	100	100	100	100	100	100

5.2. Performance of the SAA Scheme for the SPSP

We next present the results of the SAA method for the SPSP. The outcomes of the random parameters corresponding to the number of patients assessed by a physician during a shift for each scenario ξ are generated according to a truncated Poisson distribution with mean values equal to VPS_c, $c \in C$.

Table 6 summarizes the computational results of the SAA method on the SPSP. Recall that our bi-level framework starts by solving the R-CSCPP with a certain budget of uncertainty Γ. Then, it proceeds to solving the SPSP. In this table, for each combination of Γ and a number of sample scenarios N, rows $\bar{Z}_N$, $\sigma_{\bar{Z}_N}$, $\tilde{Z}_{N'}(\tilde{x})$, $\sigma_{\tilde{Z}_{N'}(\tilde{x})}$, Gap, and CPU give the average solution of the $M = 40$ sample scenarios solutions, standard deviation of the average solution, the estimated upper bound for the selected solutions among M solutions, the standard deviation of the upper bound, the gap between the average solution and the upper bound, and total CPU time in seconds for solving the problems, respectively. The approximated objective value $\tilde{Z}_{N'}(\tilde{x})$ is computed with sample size $N' = 2,000$.

Table 6: Results of the SAA method with $M = 40$ and $N^{'} = 2,000$

Inst.	Ind.	N							
		1	5	10	20	30	40	50	60
917	$\bar{Z}_N$	3818.62	4447.86	4551.53	4573.13	4614.51	4635.68	4642.20	4647.52
	$\sigma_{\bar{Z}_N}$	138.93	63.35	35.82	17.60	21.28	15.68	13.78	12.44
	$\tilde{Z}_{N'}(\tilde{x})$	11677.80	6053.71	4837.10	4681.55	4677.83	4672.75	4645.12	4642.61
	$\sigma_{\tilde{Z}_{N'}(\tilde{x})}$	50.70	32.98	14.72	13.89	14.09	13.88	13.97	13.57
	Gap	67.30	26.53	5.90	2.32	1.35	0.79	0.06	-0.11
	CPU	274.37	475.33	4162.30	9812.14	11055.88	25280.52	59168.13	84156.51
1067	$\bar{Z}_N$	7521.90	8414.72	8848.81	8793.63	8865.92	8891.89	8890.51	8900.37
	$\sigma_{\bar{Z}_N}$	377.74	113.12	68.65	44.72	35.51	36.58	28.18	21.48
	$\tilde{Z}_{N'}(\tilde{x})$	20857.40	10355.80	9146.95	8948.85	8947.39	8934.56	8903.35	8915.84
	$\sigma_{\tilde{Z}_{N'}(\tilde{x})}$	75.31	47.73	29.85	29.14	29.27	29.05	28.98	28.50
	Gap	63.94	18.74	3.26	1.73	0.91	0.48	0.14	0.17
	CPU	280.00	1058.00	3122.27	4583.59	9672.30	13190.48	28403.95	63560.44
1167	$\bar{Z}_N$	10610.00	11014.70	11637.70	11718.30	11786.20	11801.78	11804.80	11788.80
	$\sigma_{\bar{Z}_N}$	432.94	241.72	103.67	69.92	64.60	51.89	45.56	36.00
	$\tilde{Z}_{N'}(\tilde{x})$	22360.90	13735.20	12100.60	11913.60	11882.50	11836.80	11806.00	11793.80
	$\sigma_{\tilde{Z}_{N'}(\tilde{x})}$	90.22	70.45	44.85	42.49	43.50	43.81	43.36	45.67
	Gap	52.55	19.81	3.83	1.64	0.81	0.30	0.01	0.04
	CPU	269.46	593.03	4371.88	6143.82	9264.97	15477.57	21402.85	40532.60
1467	$\bar{Z}_N$	10472.00	11392.70	12041.50	12088.74	12068.50	12085.90	12158.70	12174.89
	$\sigma_{\bar{Z}_N}$	542.15	202.76	89.53	71.13	53.61	53.88	43.53	41.07
	$\tilde{Z}_{N'}(\tilde{x})$	24354.10	13266.30	12348.50	12234.30	12240.30	12239.90	12197.80	12180.40
	$\sigma_{\tilde{Z}_{N'}(\tilde{x})}$	88.39	61.85	46.17	46.42	45.16	45.20	46.00	45.36
	Gap	57.00	14.12	2.49	1.19	1.40	1.26	0.32	0.05
	CPU	269.65	582.08	5344.20	7482.70	8612.77	11942.59	15724.38	20116.55
1667	$\bar{Z}_N$	10528.70	11466.40	12442.00	12567.60	12579.21	12591.60	12533.60	12610.00
	$\sigma_{\bar{Z}_N}$	400.60	157.42	131.86	62.31	62.10	48.19	43.89	46.25
	$\tilde{Z}_{N'}(\tilde{x})$	23860.80	14012.70	12914.60	12681.90	12683.40	12676.50	12632.00	12635.50
	$\sigma_{\tilde{Z}_{N'}(\tilde{x})}$	90.27	68.92	48.62	47.88	47.50	47.69	48.36	47.51
	Gap	55.87	18.17	3.66	0.90	0.82	0.67	0.78	0.20
	CPU	272.27	541.83	4494.05	6239.83	10121.72	17450.84	24467.91	24806.62

As can be seen from Table 6, in all the instances, a high quality solution can be obtained by using a relatively small sample size. In particular over all the instances, a sample size of 30 scenarios can provide solutions within 2% of optimality gap. It is also noteworthy that the standard deviations of statistical lower and upper bounds are small enough, which suggests that the values of M and N' are adequately large. It is also worth mentioning that the CPU time increases with the expansion of the sample size and it reaches to almost 24 hours for a sample size of 60 scenarios. Observe that the total CPU time for solving all the instances with a sample size of 30 scenarios is no more than three and half hours. Figure 3 visualizes the results of the table and plots the percent deviation gap between the statistical lower bounds and upper bounds and CPU times over different sample sizes. The main axis on the left measures the gap for the solid lines, and the secondary axis on the right shows the CPU time for the series plotted by the dash lines. As the figure shows, the approximated solution of five scenarios sample size can be at most %27 away from its lower bound. As we increase the sample size, both the lower bounds and upper bounds improve (i.e. former

increase and latter decrease) and the gap between them declines. Nevertheless, the improvement in the gap for sample sizes greater than 30 scenarios is not remarkable. We can also observe a significant hike in the CPU time for sample sizes greater than 30 scenarios. Taking into account the percent deviation gap and the CPU time simultaneously makes the sample size of 30 scenarios a reasonable choice for generating good schedules, as both of these indicators are relatively low over all the Γ instances.

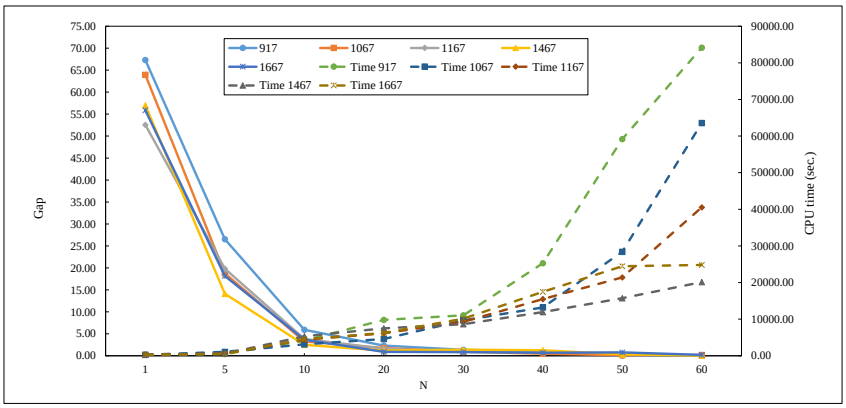

Figure 3: Percent deviation gap and CPU time in convergence of the SAA method

5.3. The Value of Stochastic Schedule

We now demonstrate the superiority of the stochastic solution over the deterministic schedule by comparing the results of the single level deterministic EVP with our bi-level framework through a Monte Carlo simulation. In the absence of uncertainty, the CSCPP and the PSP can be integrated into a single-level problem. However, in our approach, we first solve the R-CSCPP with a certain budget of uncertainty Γ. After that, we plug the obtained clinics and resource plans into the SPSP and develop physicians' work schedules. Throughout this section, we refer to the bi-level framework as *sequential approach*. Figure 4 summarizes the simulation procedure. As can be seen, in the sequential approach, first the R-CSCPP is solved (with a specific value of Γ), then its solution (D_{cjk}, R_{jk}, H_i) is set into the SPSP. After that, the SPSP is solved for a sample size of 30 scenarios, and physicians' work schedules $(\bar{x_{ijk}}, \bar{x_{ijk}}')$ are generated. In order to make the decisions regarding utilizing the available on-call physicians, assigning additional examination rooms in each shift, and also to calculate the total cost of processed patients during over-time shifts as well as physicians' idle time, we formulate a mathematical program (SIM) that receives physicians works schedules and the number of available rooms in different shifts as inputs. In this problem, the number of weekly

arriving patients as well as patients' treatment times can vary simultaneously in each replication of the simulation. Hence, the generated physicians' work schedules are plugged into the SIM, and then it is optimized for a set of $\Phi = \{\phi_1, \ldots, \phi_{100}\}$ replications. In the deterministic approach, physicians' work schedules are developed in one step and plugged into the SIM model.

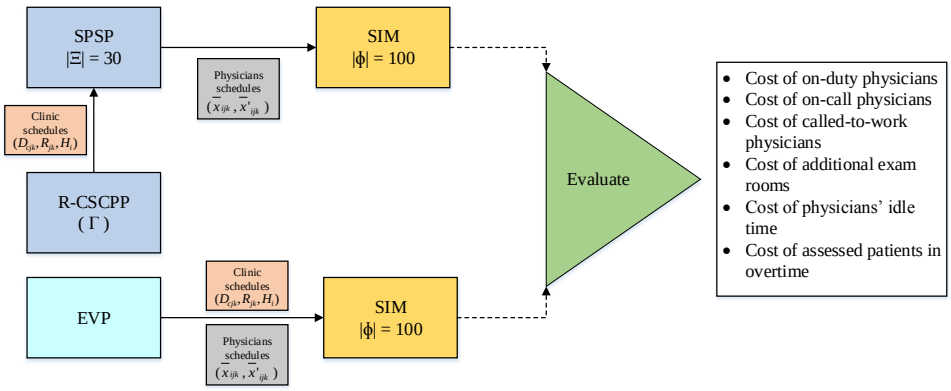

Figure 4: Simulation procedure

For a given replication ϕ, the SIM problem is formulated as follows:

$$\text{minimize} \sum_{i\in I}\sum_{j\in J}\sum_{k\in K}(f_1 \bar{x_{ijk}} + f_5 \bar{x_{ijk}}' + f_6 s_{ijk})$$

$$+ \sum_{j\in J}\sum_{k\in K}(f_7 er_{jk} + \sum_{c\in C}(f_8 ot_{cjk} + f_9 oa_{cjk})) \tag{51}$$

$$\text{subject to} \sum_{i\in I_c} VPS^{\phi}_{cjk}(\bar{x_{ijk}} + s_{ijk}) + \sum_{j\in J}\sum_{k\in K} nd_{cjk}$$

$$- \sum_{j\in J}\sum_{k\in K} VPS^{\phi}_{cjk} oa_{cjk} = NWV^{\phi}_{c} \qquad c \in C \tag{52}$$

$$\sum_{i\in I}(\bar{x_{ijk}} + s_{ijk}) - er_{jk} \leq R_{jk} \qquad j \in J,\ k \in K \tag{53}$$

$$\sum_{k\in K}(\bar{x_{ijk}} + s_{ijk}) \leq 1 \qquad i \in I, j \in J \tag{54}$$

$$s_{ijk} \leq \bar{x_{ijk}}' \qquad i \in I, j \in J, k \in K \tag{55}$$

$$s_{ijk} \in \{0, 1\} \tag{56}$$

$$ot_{cjk}\ er_{jk} \in \mathbb{Z}^+ \tag{57}$$

$$oa_{cjk} \in \mathbb{R}^+ \tag{58}$$

As can be seen in Figure 4, physicians' work schedules $(\bar{x_{ijk}}, \bar{x_{ijk}}')$ as well as the number of available

examination rooms (R_{jk}) are the inputs of the problem. In order to incorporate the randomness in the number of weekly arriving patients, the demand constraint (52) is formulated as weekly demand for each clinic. Values of the random parameters (VPS^{ϕ}_{cjk}, NWV^{ϕ}_{c}) are generated from the distributions provided in Table 7.

Table 7: Distributions of the random parameters

Parameter	Distribution	Mean	Min.	Max.
NWV^{ϕ}_{c}	Uniform	-	$\underline{NWV_c}$	$\overline{NWV_c}$
VPS^{ϕ}_{cjk}	Truncated Poisson	VPS_c	$\underline{VPS_c}$	$\overline{VPS_c}$

Figure 5 displays the simulation results. The vertical axis is the operational cost calculated by the simulation model, and the horizontal axis is replication number. As the figure demonstrates, in 99 out of 100 replications, the cost of the deterministic approach is higher than the cost of the sequential method with $\Gamma \geq 1067$. Throughout the replications, we observe an intense fluctuation in the cost of deterministic approach with a standard deviation of 4858 versus 452 for the sequential approach with $\Gamma = 1067$. Moreover, the worst case cost of the deterministic method can be as high as 30000 (replication 14), whereas the worst case cost of the sequential approach is no more than 15000 (replication 59) for any budget of uncertainty. The significant superiority of the sequential method over the deterministic approach regardless of the uncertainty budget implies the effectiveness of the recourse actions against variability in patients' treatment times in the SPSP. It is also noteworthy that when the value of Γ increases from its lower bound (i.e. 917), the cost decreases remarkably. The decline in the cost by increasing the uncertainty budget confirms the efficacy of the adjustable robustness of the R-CSCPP.

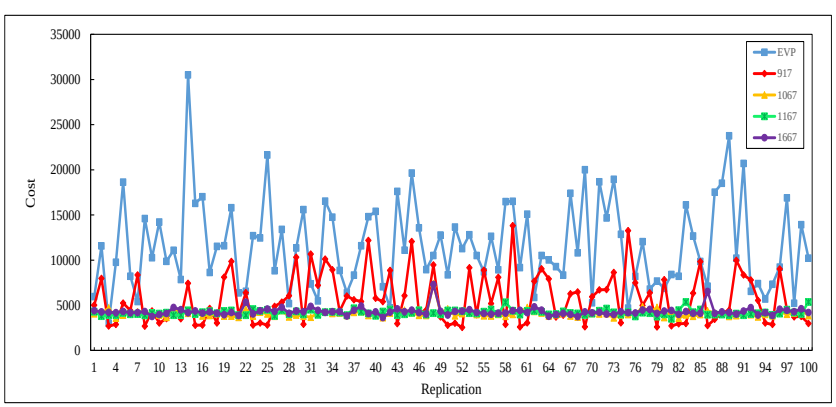

Figure 5: Simulation results

Figure 6 plots the average operational cost of the sequential method with different uncertainty

budgets along with the average cost of the EVP. The cost of the sequential method is plotted with regard to the primary axis on the left side of the graph and the cost of the deterministic approach is measured with respect to the secondary axis on the right side of the graph. The dash lines are percent deviation of the sequential method from the deterministic method.

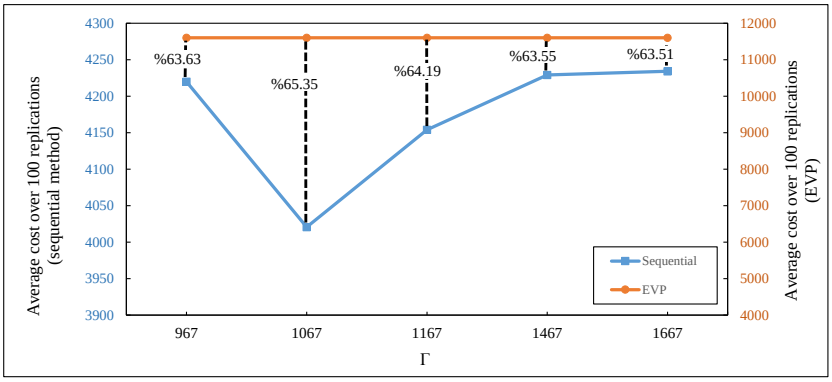

Figure 6: Average cost for different Γ values

As the dash lines demonstrate, the average cost of the sequential method is significantly lower than the average cost of the deterministic approach over all simulation replications. Additionally, the decline in the operational cost with increasing the uncertainty budget is more evident in the plot of the average cost. Moreover, an intriguing observation in this figure is the trend of the average cost throughout the uncertainty budget values; it can be noted that the decrease in the average cost ceases at a certain level of Γ (1067), and it rises from $\Gamma = 1067$ to $\Gamma = 1667$. The turnabout in the average cost is the result of over-protecting the schedules against demand uncertainty. Recall from Section 3 that the R-CSCPP provides the weekly number of patients that must be visited in each clinic that also acts as an input to the SPSP. Under a high budget of uncertainty, more clinic sessions and working hours could be scheduled for physicians to maximize the number of patients that could be scheduled during regular shifts. On the contrary, the actual number of patients who require appointments and hence scheduled during each week (generated in the simulator) might be significantly smaller than the number obtained by the R-CSCPP under a high uncertainty budget. This, in return, would lead to an increased number of idle hours for physicians initially scheduled by the SPSP under the aforementioned uncertainty budget. In other words, assigning extra resources to shifts in order to hedge against the demand uncertainty may lead to an increase in physicians' idle time. It may also increase the number of on-duty physicians which reduces the number of on-call physicians, that results in less flexibility toward patients' treatment times variability. All in

all, it is safe to conclude that choosing extreme values (low or high) for uncertainty budget results in higher costs than those of moderate values.

6. Conclusion

This paper introduced a framework for planning physicians in polyclinics under uncertainty. The procedure addresses the problem in three levels of strategic, tactical, and operational planning. In the strategic level, we proposed an adjustable robust approach that plans clinics work schedules and assigns required capacity to each shift. In particular, the model provides the decision maker with the option of protecting the plans against the uncertainty in the number of arriving patients to the polyclinic. The robust problem was solved with an implementor/adversary algorithm, which can prove optimality in small CPU times. We combined tactical physicians scheduling with operational rescheduling decisions into a two-stage stochastic program that incorporates the uncertainty in patients' treatment times. In the first stage, physicians are assigned as on-call or on-duty, and in the second stage, the on-call physicians are called to work if needed. Since the variety of patients' treatment times in each clinic results in a fairly large number of scenarios, we applied the sample average approximation scheme to obtain high quality solutions by considering only a sub-set of scenarios. The results from the computational experiments with the data provided by a polyclinic in a university health center in Montreal confirmed the efficiency of the proposed framework. Furthermore, we investigated the impact of including uncertainty in our approach by applying a Monte Carlo simulation. We compared our framework to its deterministic counterpart and demonstrated that the additional capacity included in the plan by the robust model, and the corrective (recourse) decisions implemented by the stochastic model, result in schedules that have significantly lower cost than those generated by the deterministic one. Finally, we analyzed the value of the budget of uncertainty and how it influences the downstream operational cost. The results showed that extreme values cause over or under protection phenomena, which have higher costs as compared to moderate values.

Acknowledgments

This research was partly funded by the Canadian Natural Sciences and Engineering Research Council under grants 418609-2012 and 402043-2011. This support is gratefully acknowledged.

References

Abernathy, W. J., Baloff, N., Hershey, J. C., & Wandel, S. (1973). A three-stage manpower planning and scheduling model—-a service-sector example. *Operations Research*, *21*, 693–711.

Bard, J. F., Morton, D. P., & Wang, Y. M. (2007). Workforce planning at usps mail processing and distribution centers using stochastic optimization. *Annals of Operations Research*, *155*, 51–78.

Bard, J. F., & Purnomo, H. W. (2004). Real-time scheduling for nurses in response to demand fluctuations and personnel shortages. In *Proceedings of the 5th international conference on the practice and theory of automated timetabling* (pp. 67–87).

Bard, J. F., & Purnomo, H. W. (2005a). Hospital-wide reactive scheduling of nurses with preference considerations. *IIE Transactions*, *37*, 589–608.

Bard, J. F., & Purnomo, H. W. (2005b). Short-term nurse scheduling in response to daily fluctuations in supply and demand. *Health Care Management Science*, *8*, 315–324.

Beaulieu, H., Ferland, J. A., Gendron, B., & Michelon, P. (2000). A mathematical programm g approach for scheduling physicians in the emergency room. *Health Care Management Science*, *3*, 193–200.

Van den Bergh, J., Beliën, J., De Bruecker, P., Demeulemeester, E., & De Boeck, L. (2013). Personnel scheduling: A literature review. *European Journal of Operational Research*, *226*, 367–385.

Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations Research*, *52*, 35–53.

Bienstock, D. (2007). Histogram models for robust portfolio optimization. *Journal of Computational Finance*, *11*, 1–64.

Brunner, J. O. (2010). *Flexible shift planning in the service industry: the case of physicians in hospitals* volume 640. Springer Science & Business Media.

Brunner, J. O., Bard, J. F., & Kolisch, R. (2009). Flexible shift scheduling of physicians. *Health Care Management Science*, *12*, 285–305.

Brunner, J. O., Bard, J. F., & Kolisch, R. (2010). Midterm scheduling of physicians with flexible shifts using branch and price. *IIE Transactions*, *43*, 84–109.

Burke, E. K., De Causmaecker, P., Berghe, G. V., & Van Landeghem, H. (2004). The state of the art of nurse rostering. *Journal of Scheduling*, *7*, 441–499.

Campbell, G. M. (1999). Cross-utilization of workers whose capabilities differ. *Management Science*, *45*, 722–732.

Campbell, G. M. (2011). A two-stage stochastic program for scheduling and allocating cross-trained workers. *Journal of the Operational Research Society*, *62*, 1038–1047.

Carter, M. W., & Lapierre, S. D. (2001). Scheduling emergency room physicians. *Health Care Management Science*, *4*, 347–360.

Easton, F. F., & Goodale, J. C. (2005). Schedule recovery: Unplanned absences in service operations. *Decision Sciences*, *36*, 459–488.

Erhard, M., Schoenfelder, J., Fügener, A., & Brunner, J. O. (2018). State of the art in physician scheduling. *European Journal of Operational Research*, *265*, 1–18.

Ernst, A. T., Jiang, H., Krishnamoorthy, M., & Sier, D. (2004). Staff scheduling and rostering: A review of applications, methods and models. *European Journal of Operational Research*, *153*, 3–27.

Gendreau, M., Ferland, J., Gendron, B., Hail, N., Jaumard, B., Lapierre, S., Pesant, G., & Soriano, P. (2006). Physician scheduling in emergency rooms. In *Practice and Theory of Automated Timetabling VI* (pp. 53–66). Springer.

Gunawan, A., & Lau, H. C. (2013). Master physician scheduling problem. *Journal of the Operational Research Society*, *64*, 410–425.

Holte, M., & Mannino, C. (2013). The implementor/adversary algorithm for the cyclic and robust scheduling problem in health-care. *European Journal of Operational Research*, *226*, 551–559.

Hur, D., Mabert, V. A., & Bretthauer, K. M. (2004). Real-time work schedule adjustment decisions: An investigation and evaluation. *Production and Operations Management*, *13*, 322–339.

Ingels, J., & Maenhout, B. (2015). The impact of reserve duties on the robustness of a personnel shift roster: An empirical investigation. *Computers & Operations Research*, *61*, 153–169.

Kim, K., & Mehrotra, S. (2015). A two-stage stochastic integer programming approach to integrated staffing and scheduling with application to nurse management. *Operations Research*, *63*, 1431–1451.

Kleywegt, A. J., Shapiro, A., & Homem-de Mello, T. (2002). The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, *12*, 479–502.

Maenhout, B., & Vanhoucke, M. (2013). An integrated nurse staffing and scheduling analysis for longer-term nursing staff allocation problems. *Omega*, *41*, 485–499.

Mak, W.-K., Morton, D. P., & Wood, R. K. (1999). Monte carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters*, *24*, 47–56.

Punnakitikashem, P., Rosenberber, J. M., & Buckley-Behan, D. F. (2013). A stochastic programming approach for integrated nurse staffing and assignment. *IIE Transactions*, *45*, 1059–1076.

Punnakitikashem, P., Rosenberger, J. M., & Behan, D. B. (2008). Stochastic programming for nurse assignment. *Computational Optimization and Applications*, *40*, 321–349.

Rousseau, L.-M., Pesant, G., & Gendreau, M. (2002). A general approach to the physician rostering problem. *Annals of Operations Research*, *115*, 193–205.

Shapiro, A., & Homem-de Mello, T. (2000). On the rate of convergence of optimal solutions of monte carlo approximations of stochastic programs. *SIAM journal on optimization*, *11*, 70–86.

Stolletz, R., & Brunner, J. O. (2012). Fair optimization of fortnightly physician schedules with flexible shifts. *European Journal of Operational Research*, *219*, 622–629.

Tang, J., & Wang, Y. (2015). An adjustable robust optimisation method for elective and emergency surgery capacity allocation with demand uncertainty. *International Journal of Production Research*, *53*, 7317–7328.

Tohidi, M., Kazemi, M., & Contreras, I. (2017). Integrated physicians and clinic scheduling in polyclinics. URL: `https://users.encs.concordia.ca/ icontrer/web/template.pdf`.

Van Huele, C., & Vanhoucke, M. (2014). Analysis of the integration of the physician rostering problem and the surgery scheduling problem. *Journal of Medical Systems*, *38*, 1–16.

Venkataraman, R., & Brusco, M. (1996). An integrated analysis of nurse staffing and scheduling policies. *Omega*, *24*, 57–71.

Verweij, B., Ahmed, S., Kleywegt, A. J., Nemhauser, G., & Shapiro, A. (2003). The sample average approximation method applied to stochastic routing problems: a computational study. *Computational Optimization and Applications*, *24*, 289–333.

Wright, P. D., & Bretthauer, K. M. (2010). Strategies for addressing the nursing shortage: Coordinated decision making and workforce flexibility. *Decision Sciences*, *41*, 373–401.

Wright, P. D., Bretthauer, K. M., & Côté, M. J. (2006). Reexamining the nurse scheduling problem: Staffing ratios and nursing shortages. *Decision Sciences*, *37*, 39–70.

Wright, P. D., & Mahar, S. (2013). Centralized nurse scheduling to simultaneously improve schedule cost and nurse satisfaction. *Omega*, *41*, 1042–1052.

Zaerpour, F., Bischak, D. P., & Menezes, M. B. (2017). Coordinated lab-clinics: A tactical assignment problem in healthcare. *European Journal of Operational Research*, *263*, 283–294.

Zhu, X., & Sherali, H. D. (2009). Two-stage workforce planning under demand fluctuations and uncertainty. *Journal of the Operational Research Society*, *60*, 94–103.