

Chemotherapy Operations Planning and Scheduling

Ayten Turkcan

Northeastern University, Mechanical and Industrial Engineering, Boston, MA 02115
a.turkcan@neu.edu

Bo Zeng

University of South Florida, Industrial and Management Systems Engineering, Tampa, FL 33620
bzeng@eng.usf.edu

Mark Lawley

Purdue University, Weldon School of Biomedical Engineering, West Lafayette, IN 47907
malawley@purdue.edu

November 21, 2011

Chemotherapy operations planning and scheduling in oncology clinics is a complex problem due to several factors such as the cyclic nature of chemotherapy treatment plans, the high variability in resource requirements (treatment time, nurse time, pharmacy time) and the multiple clinic resources involved. Treatment plans are made by oncologists for each patient according to existing chemotherapy protocols or clinical trials. It is important to strictly adhere to the patient's optimal treatment plan to achieve the best health outcomes. However, it is typically difficult to attain strict adherence for every patient due to side effects of chemotherapy drugs and limited resources in the clinics. In this study, our aim is to develop operations planning and scheduling methods for chemotherapy patients with the objective of minimizing the deviation from optimal treatment plans due to limited availability of clinic resources (beds/chairs, nurses, pharmacists). Mathematical programming models are developed to solve the chemotherapy operations planning and scheduling problems. A two-stage rolling horizon approach is used to solve these problems sequentially. Real-size problems are solved to demonstrate the effectiveness of the proposed algorithms in terms of solution quality and computational times.

Key words: Chemotherapy, planning, scheduling, oncology, acuity, resource allocation

1. Introduction

Cancer is the second most common cause of death in the U.S., accounting for 26% of all deaths [2]. The National Cancer Institute estimates that approximately 11.4 million Americans with a history of cancer were alive in January 2006. Some of these individuals were cancer-free, while others still had evidence of cancer and may have been undergoing treatment [2]. The demand for oncology services is projected to increase from 41 million in 2005 to 61 million in 2020 due to the aging population, the age-sensitive nature of cancer, and the increase in cancer survivors [19, 41].

Chemotherapy is one of the most commonly used cancer treatment therapies, along with surgery

and radiotherapy. It is a systemic treatment that uses drugs to kill cancer cells. Sophisticated treatment methods and improved management of side effects are increasing the demand for chemotherapy, and oncology clinics are experiencing ever higher workloads that can result in laboratory, pharmacy, and chemotherapy administration delays [1, 18, 22, 30]. Reducing waiting times for the first visit and waiting times in the clinic for chemotherapy administration are among the highest priorities for quality improvement in outpatient cancer treatment facilities [21].

Studies identified appointment scheduling that does not include clinic resources (nurse staffing and chair availability), and nursing care requirements, as the main cause of delays and unbalanced workload (Gruber et al. [23], Chabot and Fox [9]). Most previous studies propose using scheduling templates/rules based on nursing or pharmacy times (Diedrich and Plank [17], Hawley and Carter [24], Langhorn and Morrison [29]). Scheduling decisions are made in an ad-hoc manner according to physician and scheduler experiences and patient preferences. To the best of our knowledge, there is no study that proposes optimization methods to schedule chemotherapy treatments optimizing several objectives such as minimizing treatment delay, patient waiting time and staff overtime, and maximizing staff utilization. In this paper, we develop planning and scheduling methods for chemotherapy patients with the objective of decreasing patient waiting time and maximizing adherence to treatment plans, while considering the limited availability of clinic resources (beds/chairs, nurses, pharmacists). This study differs from previous studies in that it develops and uses optimization methods rather than scheduling templates and ad-hoc rules.

We believe the contributions of this study are as follows:

1. We propose an integer programming model to solve the planning problem for chemotherapy patients in infusion clinics. The objective of the planning problem is to minimize unnecessary treatment delays due to limited resources.
2. We propose a heuristic and an integer programming model to solve the scheduling problem with the objective of minimizing overtime. The proposed scheduling method finds appointment times and assigns patients to nurses and chairs simultaneously.
3. To the best of our knowledge, the proposed method is the first optimization-based method that considers patient acuities to determine nurse assignments.

The remainder of the paper is structured as follows. The chemotherapy planning and scheduling problem characteristics are explained in detail in Section 2. Section 3 presents a mathematical programming model, and section 4 presents a two-stage algorithm to solve planning and scheduling problems sequentially. Section 5 presents a variety of computational studies to illustrate the effectiveness of the proposed algorithms, and Section 6 provides some concluding remarks and discusses future work.

2. Problem characteristics

We define the planning and scheduling problems as the allocation of patient treatments to clinic days subject to available resource capacities, and setting of appointment times and allocation of chemotherapy patients to nurses and chairs/beds on each day, respectively. Planning and scheduling problems are complicated due to sheer volume of patients needing treatment, the cyclic nature of their treatment plans, the high variability in resources required, and the complexity of chemotherapy administration [3]. Chemotherapy regimens, which vary widely in the length of treatment, amount of direct nursing care required, multiple day treatments, and support therapies required, are the main inputs to the planning and scheduling problem [17, 3]. Section 2.1 provides a brief explanation of chemotherapy and the cyclic nature of chemotherapy treatment plans. Section 2.2 presents the oncology clinic environment and patient treatment process, which is important in determining appointment durations. Section 2.3 discusses the complicating role of patient acuity in planning and scheduling patient treatment, emphasizing the importance of intensity/acuity tools for better allocation of resources. Section 2.4 reviews existing planning and scheduling studies.

2.1. Chemotherapy treatment plans

Chemotherapy is a systemic treatment that uses drugs to treat cancer patients. The main aims of the chemotherapy treatment are: i) to stop or slow tumor growth, ii) to control or prevent the spread of cancer cells, and iii) to relieve cancer symptoms such as pain (palliative chemotherapy). Chemotherapy drugs affect cancer cells by altering cellular activity during one or more phases of the cell cycle (see Figure 1.a). The treatment decision depends on the stage of the disease, expected survival rate, recurrence risk, and the patient's health condition. A cancer's stage is based on the primary tumor size and whether it has spread to other areas of the body. If cancer cells are present only in the layer of cells where they developed and have not spread, the cancer is referred to as *in situ*. If cancer cells invade neighboring tissues and spread to other parts of the body through the blood and lymph systems, the tumor is said to be invasive/malignant (see Figure 1.b). Cell-cycle phase specific chemotherapy drugs are most effective against cells that are rapidly dividing (especially when tumor size is small) [4]. Cell-cycle phase non-specific chemotherapy drugs affect cells in all phases of the cell cycle and are most effective against slow dividing cells (especially when tumor size is large) [4].

Chemotherapy drugs affect not only cancer cells, but also rapidly dividing normal cells. Therefore, chemotherapy treatments are given in cycles with intervening periods of rest that allow the body to recover before the next treatment is given [4]. Chemotherapy protocols show the types of drugs, doses, and schedule of drugs based on the type of cancer, stage of cancer, and other specifics about

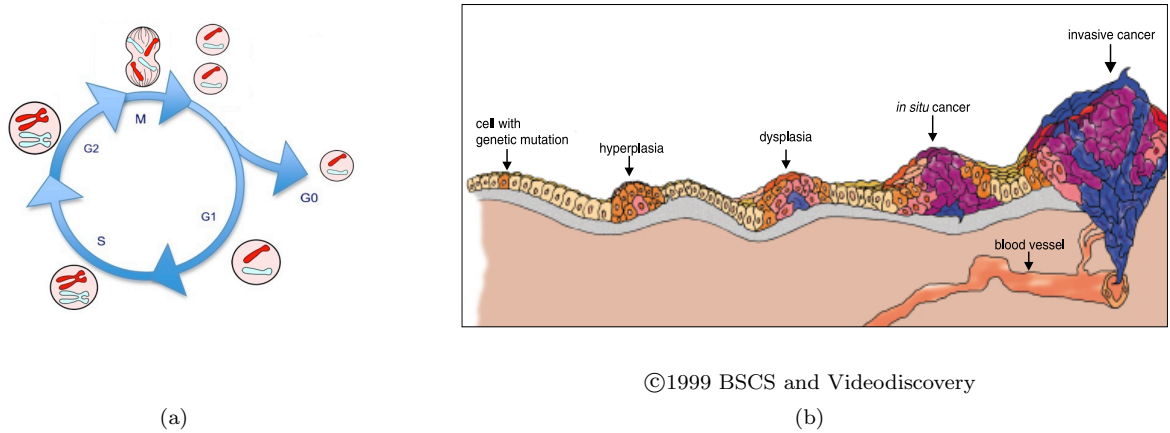


Figure 1 (a) Cell cycle includes five phases (G_0 , G_1 , S , G_2 , M) during which the cell grows, replicates, divides and rests (adapted from [32, 34]), (b) Stages of tumor growth [6]

the person's cancer (a comprehensive list of protocols can be found on the National Comprehensive Cancer Network (NCCN) website [33]). For example, Figure 2 shows the treatment timeline of a chemotherapy protocol to treat patients with Wilms tumor (a type of kidney tumor occurring in children) [26]. The cycle length is 21 days and the cycle is repeated 7 times. Different drugs are given on different days of the treatment, and lab tests are performed every week. Typically, the oncologist sees the patient at the beginning of each cycle, checks the lab results, and decides if the patient can start the cycle. If the patient has not recovered sufficiently from the previous cycle, the treatment might be delayed or the dosage might be reduced.

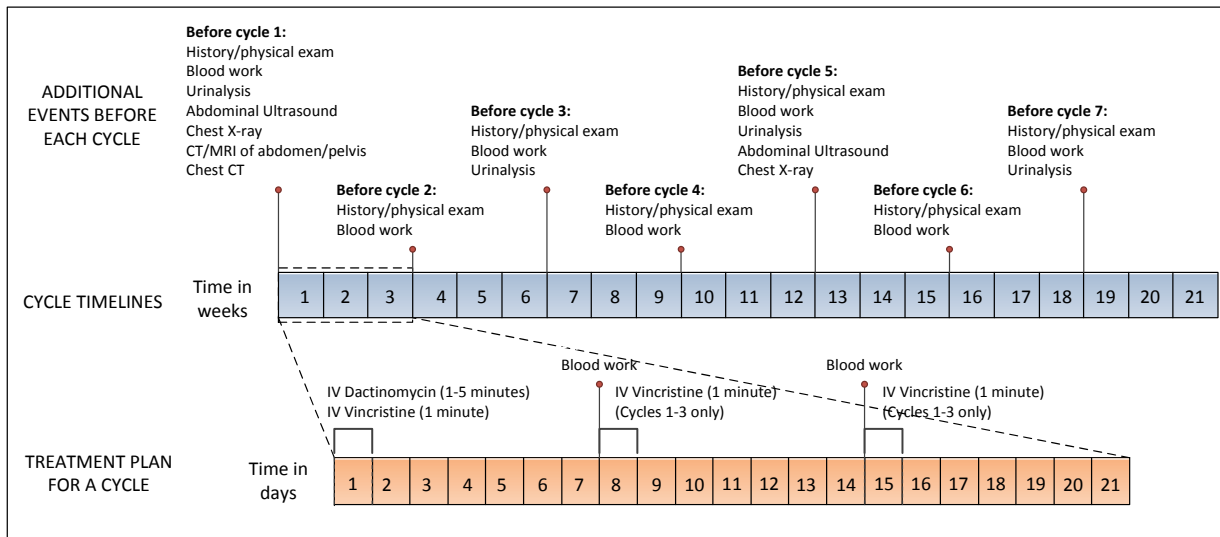


Figure 2 Treatment plan for treating a cancer patient with Wilms tumor (adapted from [26])

It is very important to adhere to the patient's treatment plan to achieve the best results, since

delaying the treatment decreases its effectiveness due to reduced dose intensity. Many studies show the correlation between low dose intensity and poor health outcomes (i.e. decreased tumor growth control, poorer quality of life, and shortened overall survival) [5, 8, 10, 37, 45]. We consider the minimization of treatment delays in the planning problem to achieve better health outcomes.

2.2. Clinic environment and patient flow in oncology and infusion clinics

In the last two decades, chemotherapy administration has shifted from the inpatient setting to the outpatient setting due to sophisticated delivery methods, new oral preparations of drugs, and improved management of side-effects, enabling patients to tolerate their treatments without being hospitalized. Cancer care is provided in a wide range of settings from solo practices to large academic medical centers [7]. Published studies provide information about the size of the cancer centers and infusion clinics [9, 11, 16, 24]. Table 1 shows the number of patients treated per day in different infusion clinics. The number of patients seen per day shows significant variability among different settings. Table 2 shows the number of new patients seen per week in different settings. Hawley and Carter [24] also presented the number of patients assigned to each nurse per day, which is 6–8 patients.

Study	Setting	Number of clinics	Number of patients
Collingwood [11]	Mountain States Tumor Institute	5 cancer centers	10–85 patients/day
Chabot and Fox [9]	Tower Hematology Oncology Medical Group	1 clinic	50–60 patients/day
DeLisle [16]	A large, private oncology practice	11 clinics	10–70 patients/day
Raad et al. [14]	Oncology centers in Australia	6 oncology centers	6–26 patients/day

Table 1 Number of patients seen per day in different infusion clinics

Study	Setting	Clinic size	Number of new patients
Hawley and Carter [24]	Cleveland Clinic Cancer Center at Hillcrest Hospital	24 chairs, 3 private rooms	30 new patients/week
Sadki et al. [38, 39]	ICL ambulatory care unit in France	18 beds	21 new patients/week

Table 2 Number of new patients seen per week in different clinics

Figure 3 presents the typical flow of chemotherapy patients in oncology and infusion clinics. All patients arrive to the clinic by appointment. The patient service assistant (PSA) registers the patient. The medical assistant (MA) prepares patient charts and takes vitals. The phlebotomy staff draws blood for laboratory tests. If the patient has a port (device inserted under the skin of the patient by surgical procedure to facilitate the blood drawing process), these tasks are performed by a nurse. The patient waits in the waiting room until the laboratory results become available.

Sometimes, the patient has his/her blood drawn and analyzed in another clinic prior to the appointment. In that case, there will be no phlebotomy work in the clinic. If the patient has to see the oncologist, all these tasks are performed in the oncology clinic and the oncologist sees the patient. If the patient's health is suitable for the treatment, he/she is sent to the infusion clinic for a same-day treatment or an appointment is scheduled for chemotherapy administration at a later date. If the patient has not totally recovered, the oncologist may delay the treatment until the patient becomes ready for the treatment. When the patient comes to the infusion clinic, he/she waits for an available chair/bed and nurse. When these resources become available, the chemotherapy nurse takes the patient to the chair/bed. The pharmacy staff prepares the chemotherapy drugs. The portering staff transports the drugs from pharmacy to the clinic. Chemotherapy nurses administer the chemotherapy to the patient. Sometimes, the patient does not have to see the oncologist. In that case, the patient goes directly to the infusion clinic for treatment. Sadki et al. [38, 39], Baldwin [3], van Lent et al. [42], Sepulveda et al. [40] provide similar patient flow diagrams.

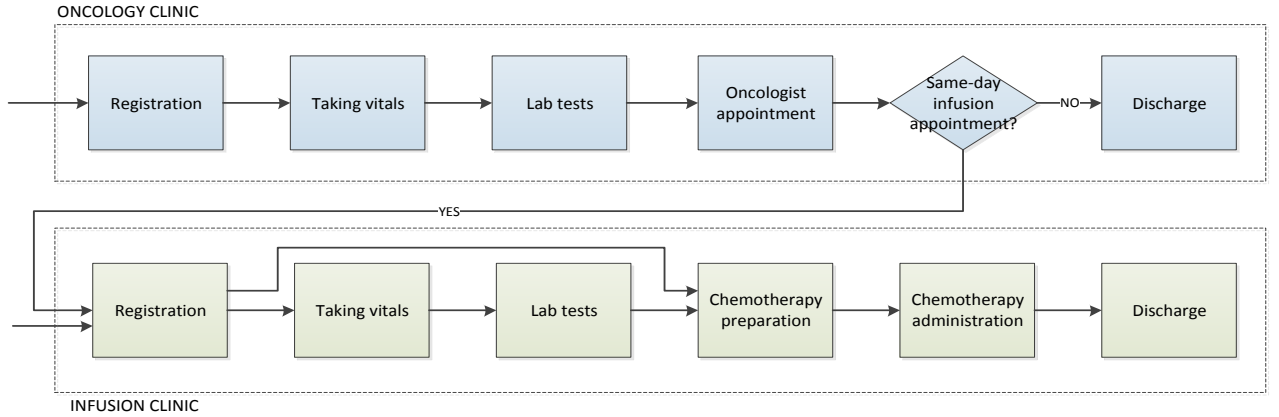


Figure 3 Patient Flow

Infusion clinics consider one or more of these processes in determining appointment durations for patient scheduling. Chemotherapy administration is the main component of the appointment duration. The lab and pharmacy times might be included in the scheduled duration depending on the current practice in the clinic. For example, if chemotherapy is prepared after the patient is taken to the infusion chair/bed, then the pharmacy time is included in the scheduled duration. If the patient needs labs and if the clinic is not using next-day chemotherapy scheduling (where labs are performed and the patient is seen by the oncologist on one day, and chemotherapy is administered on the next day [18]), the scheduler might adjust the scheduled appointment duration

to give enough time for lab tests. If the patient has to see an oncologist on the same day, there should be enough time between oncologist and infusion appointments. If the infusion clinic is small and there is not enough ancillary service staff, then the vitals are taken by the nurses. If there is no pharmacist, the chemotherapy is prepared by the infusion nurse. Since there is a variety of clinic practices, we included all processes in the patient flow. In our study, we assume that the scheduled duration includes all processes that are performed when the patient is in the chair or an infusion nurse is performing the task even if the patient is not in the chair.

2.3. Chemotherapy nursing and high variability in resource requirements

Chemotherapy chairs/beds and nurses are two key resources that should be considered while planning and scheduling patients. The availability of these resources determines the capacity of the clinic. The clinic space is the fixed capacity and nurses determine the flexible capacity since staffing levels can be adjusted by clinic managers. Even though the clinics might have limited space to handle the patient volume, nurse staffing is a more crucial problem in oncology settings. According to a survey of chemotherapy infusion clinics (ambulatory, outpatient, and physician practices) conducted by the Oncology Nursing Society (ONS), 47% of respondents reported open positions, which reflects the significance of the nursing shortage [25]. In two other surveys, 59% and more than 80% of oncology nurses perceived staffing as inadequate [7, 28]. Different methods are used to increase staffing levels such as hiring less experienced nurses [25], asking nurses to work overtime [7, 25], hiring an increased number of unlicensed ancillary personnel [25], and using agency nurses, internal float pools, and nurses from other departments [7]. Since chemotherapy nurses are specially trained, pulling nurses from other departments is not always possible, which makes the resource allocation decisions even more important.

Chemotherapy nurses perform several tasks such as administering chemotherapy, managing side-effects, stabilizing patients during an emergency, documenting important information in patient charts, providing counseling to patients and family members, and triaging patient questions and problems [35]. There is high variability in nurse workflow due to hundreds of cancer-specific protocols that require different infusion methods and treatment durations (ranging from 15 minutes to more than 8 hours). There is also high variability in treatment duration for each regimen due to patient specific factors such as difficult vein access, risk for side effects, and change in dosage. The variability in treatment durations and nurse workflow complicates the planning and scheduling process [14, 15, 38, 39].

In literature, there are studies that propose patient intensity/acuity tools in ambulatory oncology settings [9, 12, 16, 24, 27, 31] to establish appropriate staffing levels [20, 24, 44] and improve

scheduling with better resource allocation [9, 13, 27]. Chabot and Fox [9] develop a patient-classification system that represents patient care and staffing needs. Acuity levels are assigned to each regimen based on the number of agents, pre-medications, complexity of administration and assessments required. Hawley and Carter [24] use total treatment time, time with patient and/or family members, blood draws and any additional nursing needs assessed by the nurse at the time of the treatment to determine the acuity level. In this study, we propose acuity-based planning and scheduling methods for better allocation of nurses.

2.4. Literature on chemotherapy planning and scheduling

Sadki et al. [38, 39] is the only study that considers the chemotherapy planning problem in literature. They propose an integer programming model to solve the planning problem with the objective of balancing the workload throughout the week. We note that the planning problem may not always be able to generate a feasible daily schedule due to acuity levels and available nurse capacity. Therefore, the impact of planning can be measured only when the scheduling problem is completely solved. We consider both planning and scheduling problems to generate feasible daily schedules considering both chair and nursing times. The minimization of treatment delay is the main objective for the planning problem in our study.

There are a few studies about chemotherapy appointment scheduling in nursing literature [9, 17, 18, 23, 24, 29]. Langhorn and Morrison [29], Diedrich and Plank [17], and Hawley and Carter [24] propose scheduling templates/rules based on nursing or pharmacy times. These templates show nurses, chairs, and appointment slots on a spreadsheet. Scheduling rules include the maximum number of patients that can be scheduled at any time slot, earliest and latest appointment times that treatments can be scheduled, and calculation of appointment durations based on procedures to be performed. The scheduling decisions themselves are made in an ad-hoc manner according to physician and scheduler experiences and patient preferences. Chabot and Fox [9] discussed the scheduling system changes resulting from their previously mentioned patient-classification system. These included representing nurse schedules in the schedule log, assigning patients to nurses, and better coordination of oncologist appointments and infusion appointments. A three-year implementation of the new acuity system and scheduling guidelines resulted in more patients treated, more balanced workload throughout the day, reduced overtime, and increased staff and patient satisfaction. Dobish [18] proposes a next-day chemotherapy schedule, with laboratory and physician appointments on one day and chemotherapy administration on the next day. The implementation of this next-day approach resulted in improved efficiencies for pharmacy and nursing and reduced in-clinic waiting times for patients. After the change, the pharmacy was able

to prepare 95% of the orders on time for the patients seen the day before. The percentage was only 44% for the patients seen and treated on the same day. Gruber et al. [23] changed nurse working hours and established scheduling procedures to improve on-time starts (from 11% to 94%). We note that none of these authors use formal methods to optimize an objective function subject to constraints. The fact that they could achieve significant improvements without rigorous modeling reflects the degree of inefficiency in current practice.

3. Problem definition

In this section, we first introduce the notation and define the planning and scheduling problem with its underlying assumptions. Then, we propose an integer programming model (IP_1) to solve the planning problem, which assigns new patients' treatments to days without changing the plans of existing patients. The objective is minimizing the treatment delays. The planning problem provides an input (the set of patients assigned to each day) to the daily scheduling problem. An integer programming model (IP_2) is proposed to solve the scheduling problem that considers both chair and nurse availabilities. The complexity of the proposed model is reduced by considering only the nurse availabilities in a revised model, IP_3 . A heuristic ($ALTT$) is proposed to provide schedules in short computation times. The proposed scheduling methods find appointment times and assign patients to chairs and/or nurses.

Table 3 provides notation that will be used throughout the paper. We assume that the treatment plan (cycle length, C_i , and the number of times a cycle will be repeated, F_i) is known for each patient. This is a realistic assumption because chemotherapy treatments are planned by oncologists based on the chemotherapy protocols. If the oncologist changes the treatment plan for an existing patient, the patient can be considered as a new patient and appointments can be scheduled based on the new treatment plan. We assume that the length of treatment on each day, r_{id} , is also known. In a real clinic, the appointment durations should be estimated for each treatment regimen based on the infusion durations and additional tasks that should be performed when the patient is in the clinic (i.e., blood draws, laboratory tests, drug preparations, etc.). Since it would be difficult for the scheduler to calculate the appointment durations, the clinic staff should provide these estimates to the scheduler. The effect of treatment delays changes according to cancer type, stage of the disease and the patient's health. We use priorities (w_i^d) for each patient i to incorporate the effect of treatment delays. As w_i^d increases, the negative effect of treatment delay on patient's health increases.

We assume there are new patients waiting for their treatment to start (P^N) and existing patients who are already in their treatment cycles (P^E). The planning problem is to assign the cycle of

P^N, P^E	Set of new and existing patients ($i \in P^N \cup P^E$ and $P^N \cap P^E = \emptyset$)
P^R	Set of patients who are referred to receive chemotherapy in the last planning period ($P^R \subset P^N$)
P^O	Set of patients whose treatment cycles have terminated in the last planning period ($P^O \subset P^E$)
P^{NE}	Set of patients whose treatments started in the last planning period ($P^{NE} \subset P^N$)
PR_t	Total treatment time of the patients assigned to day t in previous planning horizon
PB_t	Total acuity of the patients assigned to day t in previous planning horizon
C_i	Cycle length of the treatment for patient i (in days) ($d = 1 \dots C_i$)
F_i	Number of cycles that will be repeated for patient i ($f = 1 \dots F_i$)
K	Number of chairs ($k = 1 \dots K$)
T	Length of planning horizon (days) ($t = 1 \dots T$)
Δ	Re-planning frequency (days) ($\Delta < T$)
N_t	Number of nurses on day t ($j = 1 \dots N_t$)
U	Target nurse utilization
S	Number of slots on each day ($s = 1 \dots S$)
H_t	Number of regular working hours on day t
w^o	Cost of overtime
w^u	Cost of idle time
w_i^d	Effect of treatment delay on patient i
r_{id}	Treatment length on day d of each cycle for patient i ($d = 1 \dots C_i$)
a_{id}	Acuity level on day d of each cycle for patient i ($d = 1 \dots C_i$)
est_i	Earliest treatment start day for patient i
A^{max}	Maximum acuity level a nurse can handle at any time
X_{it}	Binary variable, 1 if the treatment of patient i starts on day t , 0 otherwise
R_{it}	Treatment time required for patient i on day t
A_{it}	Acuity level of patient i on day t per time slot ($A_{it} = \{1, 2, 3, \dots\}$)
B_{it}	Total acuity of patient i on day t , i.e. $R_{it}A_{it}$
P_t	Set of patients who have treatment on day t , i.e. $P_t = \{i : R_{it} > 0\}$
G_t^o	Over utilization on day t
G_t^u	Under utilization on day t
C_t^{max}	Completion time of all treatment on day t
Y_{ijkst}	Binary variable, 1 if the treatment of patient i is started by nurse j on chair k at time slot s on day t , 0 otherwise
M_{jt}	Completion time of all treatments assigned to nurse j on day t

Table 3 Notation

treatments to a sequence of days for each patient. The scheduling problem is finding appointment times for all patients on their assigned days. The planning problem is solved for a number of weeks or months. The planning horizon (T) is divided into days, and days are divided into smaller time slots (S) to find the appointment days and times for patients.

The oncology clinic has limited resources. The number of chairs/beds (K) and other equipment determine the fixed capacity of the clinics. The clinic staff (nurses, pharmacists, etc.) who provide care determine the flexible capacity. Thus, the capacity can be increased by increasing the number of nurses (N_t) or working overtime. However, increasing the number of nurses and working overtime will increase the cost, since staff are paid more than the regular rate when they work after clinic hours. When the clinic resources are not fully utilized, the idle time can be thought of as lost

capacity and reduced access to care. We assume that the number of nurses and their normal working hours (H_t) are given for each day in the planning horizon.

As discussed in Section 2.3, the workflow of chemotherapy nurses is complicated. Considering only the total number of patients assigned to each nurse does not reflect the actual workload of the nurses throughout the day. If patients with high acuity levels are assigned to the same nurse, there may be delays in treatment and the probability of making errors increases because of heavy workload. We consider acuity levels (a_{id}) for different treatment types to achieve a well-balanced workload for nurses. The acuity levels can be determined by nurse input or by performing time studies and analyzing the time spent for all treatment-related tasks. Even though there are studies that develop patient intensity/acuity tools, none of them provide a comprehensive list of acuity levels for treatment regimens. In this study, we assign acuity levels to the regimens according to the number of agents, pre-medications and complexity of administration as in Chabot and Fox [9]. The acuity levels we assign might not be exact. Our aim is not to develop an acuity system, but to propose an acuity-based scheduling method and compare it with current practice that does not consider acuity levels. Chabot and Fox [9] propose a scheduling method that assigns patients to nurses according to the acuity level. In their study, they assume that all nursing time is required at the beginning of the treatment. However, even though the nurse starts the treatment for each patient, he/she also has to monitor the patient throughout the treatment, may need to perform additional tasks, and must end the treatment. The nursing time may not be as intense as it is at the beginning of the treatment, but assigning a large number of patients to a nurse at later times might increase the risk of errors due to higher workload. In current practice, some clinics assign a fixed number of chairs to each nurse to have a balanced workload. However, a simple count of number of patients assigned is not a good representation of the actual workload assigned to a nurse [15]. Thus, we use an acuity level which represents the amount of nursing time per slot. We assume an upper bound (A^{max}) on the amount of acuity that can be assigned to a single nurse, thus the maximum number of patients that can be assigned depends on the acuity mix. We assume that a nurse can start at most one treatment per slot since he/she must carefully assess the patient, check dosages, and start the treatment.

We propose a two-stage algorithm to solve the chemotherapy operations planning and scheduling problems sequentially. At Stage 1, the planning problem is solved to find the treatment start days of the patients (X_{it}), which is the first day of the first cycle. Since the treatment plan for each patient is known, the treatment days ($t' = t + f \cdot C_i + d - 1$), resource requirements (R_{it}) and acuity levels (A_{it}) are also determined at this stage. At Stage 2, the daily scheduling problem is solved

for the given set of patients assigned to the same day. Patients are assigned to chairs, nurses, and appointment times. Figure 4 illustrates the two-stage algorithm to solve planning and scheduling problems.

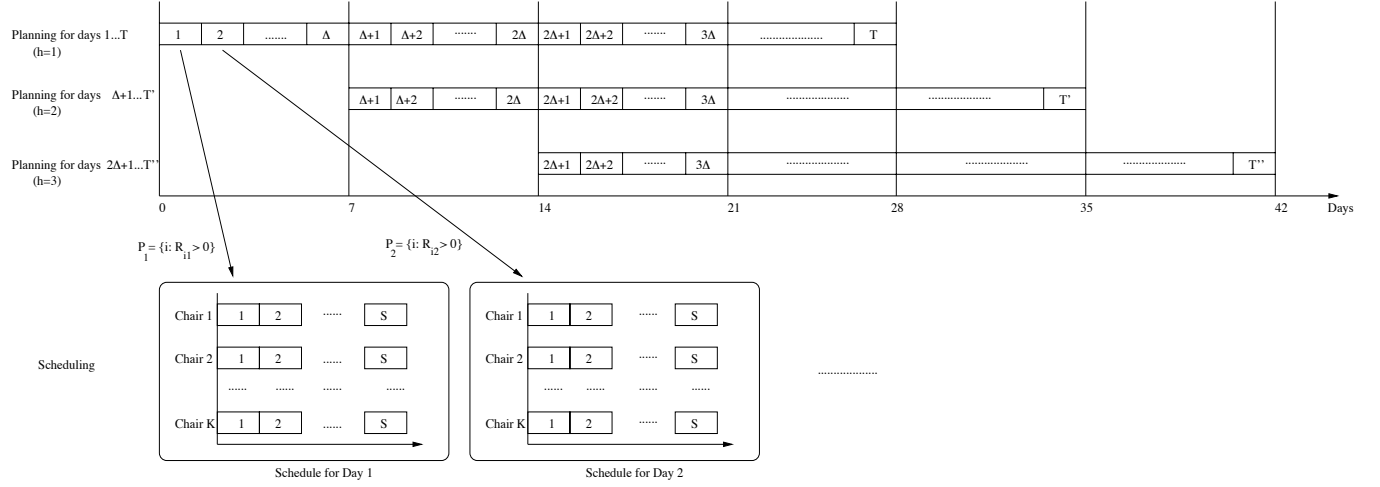


Figure 4 Chemotherapy operations planning and scheduling

3.1. Planning

The planning problem is as follows:

$$(IP_1) \quad \min \sum_{t=1}^T (w^o G_t^o + w^u G_t^u) + \sum_{i \in P^N} \sum_{t=1}^T w_i^d (t - est_i) X_{it} \quad (1)$$

$$st \quad \sum_{t=1}^T X_{it} \leq 1 \quad \forall i \in P^N \quad (2)$$

$$R_{it} = \sum_{f=1}^{F_i} \sum_{d=1}^{C_i} r_{id} X_{i,t-(f-1)C_i-d+1} \quad \forall i \in P^N, t = 1 \dots T \quad (3)$$

$$A_{it} = \sum_{f=1}^{F_i} \sum_{d=1}^{C_i} a_{id} X_{i,t-(f-1)C_i-d+1} \quad \forall i \in P^N, t = 1 \dots T \quad (4)$$

$$B_{it} = \sum_{f=1}^{F_i} \sum_{d=1}^{C_i} r_{id} a_{id} X_{i,t-(f-1)C_i-d+1} \quad \forall i \in P^N, t = 1 \dots T \quad (5)$$

$$G_t^o - G_t^u = \sum_{i \in P^N} R_{it} + PR_t - K \cdot H_t \quad t = 1 \dots T \quad (6)$$

$$\sum_{i \in P^N} B_{it} + PB_t \leq UN_t H_t A^{max} \quad t = 1 \dots T \quad (7)$$

$$X_{it} \in \{0, 1\} \quad \forall i \in P^N, t = 1 \dots T \quad (8)$$

The objective is to minimize total staff overtime and idle time, and total treatment delay. The first term in the objective function is total overtime and idle time cost of clinic staff. The second term is the total weighted treatment delay, which is calculated according to the earliest treatment start time (est_i) and planned treatment start time ($\sum_t t X_{it}$). The earliest treatment start time is determined by the oncologist and depends on the patient's health status, the treatment plan and other treatments (surgery, radiotherapy) that need to be coordinated with chemotherapy. The delays are multiplied by w_i^d to incorporate the differences among treatments in terms of their effect on health outcomes. The patient's treatment will start on at most one of the days in the planning horizon, which is guaranteed by constraint (2). If the planning horizon is not long enough and the number of patients is high, then the treatment may not start in the planning horizon. The resource requirement (treatment time, nurse time, pharmacy time) and acuity level per unit time for each patient on a given day after the treatment start (a patient's acuity level is positive only on appointment days, zero otherwise) are calculated in constraints (3) and (4). Total acuity level of

patient i on a day is calculated by constraint (5). For example, for patient i^* with cycle length of 21 days ($C_{i^*} = 21$), treatment on days 1 and 3 of each cycle (treatment length on day 1 is $r_{i^*1} = 90$, acuity level on day 1 is $a_{i^*1} = 2$, treatment length on day 3 is $r_{i^*3} = 60$, acuity level on day 3 is $a_{i^*3} = 1$, treatment length and acuity levels on days 2, 4–21 are zero), and two cycles ($F_{i^*} = 2$), constraints (3)–(5) would be as follows:

$$\begin{aligned}
R_{i^*1} &= 90X_{i^*t} & t &= 1, 2 \\
R_{i^*t} &= 60X_{i^*,t-2} + 90X_{i^*t} & t &= 3 \dots 21 \\
R_{i^*t} &= 90X_{i^*t-21} + 60X_{i^*t-2} + 90X_{i^*t} & t &= 22, 23 \\
R_{i^*t} &= 60X_{i^*,t-23} + 90X_{i^*,t-21} + 60X_{i^*t-2} + 90X_{i^*t} & t &= 24 \dots T \\
\\
A_{i^*1} &= 2X_{i^*t} & t &= 1, 2 \\
A_{i^*t} &= 1X_{i^*,t-2} + 2X_{i^*t} & t &= 3 \dots 21 \\
A_{i^*t} &= 2X_{i^*t-21} + 1X_{i^*t-2} + 2X_{i^*t} & t &= 22, 23 \\
A_{i^*t} &= 1X_{i^*,t-23} + 2X_{i^*,t-21} + 1X_{i^*t-2} + 2X_{i^*t} & t &= 24 \dots T \\
\\
B_{i^*1} &= 180X_{i^*t} & t &= 1, 2 \\
B_{i^*t} &= 60X_{i^*,t-2} + 180X_{i^*t} & t &= 3 \dots 21 \\
B_{i^*t} &= 180X_{i^*t-21} + 60X_{i^*t-2} + 180X_{i^*t} & t &= 22, 23 \\
B_{i^*t} &= 60X_{i^*,t-23} + 180X_{i^*,t-21} + 60X_{i^*t-2} + 180X_{i^*t} & t &= 24 \dots T
\end{aligned}$$

Constraint (6) is used to calculate the overtime and idle time. Since the appointment times are not determined in the planning part, the overtime and idle time are approximate values based on total available capacity ($K \cdot H_t$). Constraint (7) is used to control the total acuity of the patients assigned to nurses on each day. The maximum total acuity that can be assigned to nurses during normal working hours is $N_t H_t A^{max}$. However, this is an over-estimate, because it does not consider that a nurse can start at most one treatment at any slot. Therefore, we multiply the maximum total acuity level by U . The value of U should be selected carefully in order not to overload nurses. Constraints (6) and (7) are necessary to make a reasonable number of assignments to days based on available capacity.

3.2. Scheduling

The proposed model finds the treatment days, resource requirements, and acuity levels on each treatment day for new patients. The resource requirements and acuity levels for existing patients were already calculated in previous time periods and are used as inputs in the proposed model to calculate the overtime and idle time. The resource requirements (R_{it}) and acuity levels (A_{it}) for all patients ($i \in P^N \cup P^E$) are used as inputs to the scheduling problem. We proposed two integer programming models and a heuristic. The first integer programming model considers both chair and nurse availabilities. The second integer programming model considers just the nurse availabilities

assuming that nurses are the limited resources. The heuristic, which sorts the patients according to treatment duration and assigns them to chairs and nurses one at a time, is proposed to find schedules in small computation times.

3.2.1. Integer programming models The integer programming model considering both chair and nurse availabilities is as follows:

$$(IP_2) \quad \min \sum_{t=1}^T C_t^{max} \quad (9)$$

$$st \quad \sum_{j=1}^{N_t} \sum_{k=1}^K \sum_{s=1}^{S-R_{it}+1} Y_{ijkst} = 1 \quad \forall i \in P_t, t = 1 \cdots T \quad (10)$$

$$\sum_{i \in P_t} \sum_{j=1}^{N_t} \sum_{u=\max\{s-R_{it}+1, 1\}}^{\min\{S-R_{it}+1, s\}} Y_{ijkut} \leq 1 \quad k = 1 \cdots K, s = 1 \cdots S, t = 1 \cdots T \quad (11)$$

$$\sum_{i \in P_t} \sum_{k=1}^K \sum_{u=\max\{s-R_{it}+1, 1\}}^{\min\{S-R_{it}+1, s\}} A_{it} Y_{ijkut} \leq A^{max} \quad j = 1 \cdots N_t, s = 1 \cdots S, t = 1 \cdots T \quad (12)$$

$$\sum_{i \in P_t} \sum_{k=1}^K Y_{ijkst} \leq 1 \quad j = 1 \cdots N_t, s = 1 \cdots S, t = 1 \cdots T \quad (13)$$

$$\sum_{k=1}^K \sum_{s=1}^{S-R_{it}+1} Y_{ijkst} (s + R_{it} - 1) \leq M_{jt} \quad \forall i \in P_t, j = 1 \cdots N_t, t = 1 \cdots T \quad (14)$$

$$M_{jt} \leq C_t^{max} \quad j = 1 \cdots N_t, t = 1 \cdots T \quad (15)$$

$$Y_{ijkst} \in \{0, 1\} \quad \forall i \in P_t, j = 1 \cdots N_t, k = 1 \cdots K, s = 1 \cdots S, t = 1 \cdots T \quad (16)$$

The objective is to minimize the total completion time of all treatments on day t . Each patient who has treatment on day t ($i \in P_t$) is assigned to a nurse, a chair, and a time slot by constraint (10). Constraint (11) ensures that at most one patient is assigned to a chair. The total acuity level assigned to a nurse cannot exceed the maximum acuity level, which is controlled by constraint (12). Constraint (13) ensures that a nurse can start at most one treatment per slot. The total completion time of the treatments assigned to a nurse on a given day is calculated in constraint (14). The total completion time of all treatments for each day is calculated in constraint (15). Constraint (16) is the integrality constraint. The proposed model IP_2 can be decomposed into smaller subproblems for each day t . The subproblems can be solved independently to find the appointment times, nurses and chairs on each day.

The proposed integer programming models can be solved optimally. However, the length of the planning horizon, number of time slots on each day, and number of patients, nurses and chairs affect the computational complexity of the problem. The first model (IP_1) has $|P^N| \times T$ binary variables, $3|P^N| \times T + T$ continuous variables, and $|P^N| + 3|P^N| \times T \times \max_i \{C_i F_i\} + 2T$ constraints. The second model (IP_2) has $|P_t| \times N_t \times K \times S$ binary variables, and $|P_t| + K \times S + 2N_t \times S + |P_t| \times N_t + N_t$ constraints for a given day t . Please note that the second model is solved for all days in the planning horizon. If the number of slots, patients, chairs, and nurses are 40, 50, 20, and 7, respectively, then there will be 280,000 binary variables and 8,687 constraints in IP_2 for each day.

One way to reduce the computational complexity of IP_2 is considering a single resource rather than considering both resources (chairs and nurses). For example, if the nurses' time is the limiting resource, as appears to be the case in most of the clinics, the chair assignment can be removed from the model. The index k for chairs and constraint (11) are removed from the model to have a smaller size problem with $|P_t| \times N_t \times S$ binary variables, and $|P_t| + 2N_t \times S + |P_t| \times N_t + N_t$ constraints for a given day t . The patients should be assigned to chairs after the following integer programming model (IP_3), which is used to assign patients to nurses and to find appointment times, is solved.

$$(IP_3) \quad \min \sum_{t=1}^T C_t^{max} \quad (9)$$

$$st \quad \sum_{j=1}^{N_t} \sum_{s=1}^{S-R_{it}+1} Y_{ijst} = 1 \quad \forall i \in P_t, t = 1 \cdots T \quad (10')$$

$$\sum_{i \in P_t} \sum_{u=\max\{s-R_{it}+1, 1\}}^{\min\{S-R_{it}+1, s\}} A_{it} Y_{ijut} \leq A^{max} \quad j = 1 \cdots N_t, s = 1 \cdots S, t = 1 \cdots T \quad (12')$$

$$\sum_{i \in P_t} Y_{ijst} \leq 1 \quad j = 1 \cdots N_t, s = 1 \cdots S, t = 1 \cdots T \quad (13')$$

$$\sum_{s=1}^{S-R_{it}+1} Y_{ijst}(s + R_{it} - 1) \leq M_{jt} \quad \forall i \in P_t, j = 1 \cdots N_t, t = 1 \cdots T \quad (14')$$

$$M_{jt} \leq C_t^{max} \quad j = 1 \cdots N_t, t = 1 \cdots T \quad (15')$$

$$Y_{ijst} \in \{0, 1\} \quad \forall i \in P_t, j = 1 \cdots N_t, s = 1 \cdots S, t = 1 \cdots T \quad (16')$$

3.2.2. Heuristic We propose a heuristic ($ALTT_t$) to find appointment schedules in short computation times. The basic steps of the algorithm is as follows:

At Step 1, $NurseAvail_{jst}$, $ChairAvail_{kst}$, and $TreatStart_{jst}$ are initialized as zero. If

Algorithm 1 *ALTT_t*: Longest treatment time first rule incorporating acuity levels

```

1:  $NurseAvail_{jst} = 0, ChairAvail_{kst} = 0, TreatStart_{jst} = 0, M_{jt} = 0$ 
2: Sort all patients according to their treatment times, i.e.  $R_{it} \geq R_{i+1,t}$ 
3: for all  $i = 1$  to  $I$  do
4:   for all  $s = 1$  to  $S, j = 1$  to  $J$  and  $k = 1$  to  $K$  do
5:     Assign patient  $i$  to nurse  $j$ , chair  $k$  and slot  $s$  temporarily
6:     if  $NurseAvail_{j,s',t} + A_{it} \leq A^{max}, TreatStart_{j,s,t} < 1$  and  $ChairAvail_{j,s',t} < 1$  for all  $k, j, s$  and  $s'$  where
        $s' \geq s$  and  $s' < s + R_{it}$  and given  $t$  then
7:       Go to Step 9
8:     end if
9:      $(s, j, k)$  satisfies the constraints (6), (7) and (8). Go to Step 11.
10:   end for
11:    $(s^*, j^*, k^*) = (s, j, k)$  and  $Y_{i,j^*,k^*,s^*,t} = 1$ 
12:    $TreatStart_{j^*,s^*,t} = 1$ 
13:   for all  $s' \geq s^*$  and  $s' \leq s^* + R_{it} - 1$  do
14:      $NurseAvail_{j^*,s',t} = NurseAvail_{j^*,s^*,t} + A_{it}$ 
15:      $ChairAvail_{j^*,s',t} = 1$ 
16:   end for
17:   if  $s^* + R_{it} - 1 > M_{j^*,t}$  then
18:      $M_{j^*,t} = s^* + R_{it} - 1$ 
19:   end if
20:   if  $M_{j^*,t} > C_t^{max}$  then
21:      $C_t^{max} = M_{j^*,t}$ 
22:   end if
23: end for

```

$NurseAvail_{jst}$ is zero, that means nurse j is available at slot s on day t . If $ChairAvail_{kst}$ is zero, that means chair/bed k is available at time slot s on day t . When $TreatStart_{jst}$ is zero, nurse j can start a treatment at slot s on day t . The completion time of all patients assigned to nurse j on day t (M_{jt}) is also initialized to zero at Step 1. At Step 2, the patients are sorted in non-increasing order of their treatment times (R_{it}). Each patient in the sorted list is assigned to a nurse, a chair/bed and a time slot (j, k, s) temporarily at step 5. At Step 6, it is checked whether constraints (6), (7) and (8) are satisfied. If all constraints are satisfied, then the patient is assigned to the corresponding nurse, chair/bed and slot (Step 11). At Steps 12–16, $TreatStart_{jst}$, $NurseAvail_{jst}$, and $ChairAvail_{kst}$ are updated. At Steps 17–19, the completion time of all patients assigned to a nurse on day t (M_{jt}) is updated. At Steps 20–22, the completion time of all treatments on day t is calculated.

The proposed algorithm is similar to longest processing time (LPT) rule, which is used to minimize makespan in parallel machine scheduling [36]. The LPT rule is modified to control the total

acuity level and number of treatment starts assigned to each nurse. The number of slots is chosen as large as possible so that all patients can be assigned to nurses, chairs/beds and time slots. The proposed algorithm aims to minimize the completion time of all patients.

The planning and scheduling problems are solved for the patients who need treatment for several weeks. The schedules cannot be fixed for the whole planning horizon. New patients with different acuity levels and treatment plans are referred to receive chemotherapy treatment and they should be added to the existing schedule. The planning problem should be solved frequently to minimize treatment delays for new patients. We use a rolling horizon approach to solve planning and scheduling problems sequentially. The proposed method is explained in detail in the following section.

4. Rolling horizon methodology

We propose a rolling horizon approach to solve planning and scheduling problems with the objective of minimizing treatment delays for new patients. We solve the planning problem every Δ days for a planning horizon of T days (where $\Delta \leq T$). The planning problem finds the treatment start days for new patients for the first Δ days starting from current time (t^c). Once the first treatment day of a new patient is identified, the remaining treatment days are fixed in the planning horizon according to the chemotherapy cycle. (As noted below, the planning horizon T must be sufficiently long to cover the treatment cycles for all patients referred up to t^c .) No new treatment start is planned for time periods $t^c + \Delta + 1, \dots, T$.

The detailed appointment schedule is generated only for the first Δ days. After the schedule for Δ days is executed, the planning and scheduling problems are solved for time periods $t^c + \Delta + 1, \dots, t^c + \Delta + T$. The planning horizon T must be updated every time the planning problem is solved since the set of new patients will change. The patients whose treatments have started in the last Δ days, P^{NE} , are removed from the set of new patients, P^E , and added to the set of existing patients, P^E . Patients referred to receive chemotherapy in the last Δ days, P^R , are added to the new patient set, P^E , and patients whose treatment cycles finished in the last Δ days, P^O , are removed from the set the existing patient set, P^E .

The planning horizon T is calculated according to the total treatment length of new patients and the treatment completion time of existing patients, i.e. $T = \max\{t^c + \Delta + \max_{i \in P^N}\{C_i F_i\}, \max\{t : R_{it} > 0, i \in P^E\}\}$. The first term is the maximum treatment completion time for the new patients. The second term is the treatment completion time for all existing patients. The treatment length of a new patient is calculated by multiplying the cycle length (C_i) with number of cycles (F_i).

The treatment completion time of an existing patient i is the last time period that the patient is treated and is calculated as $\max\{t : R_{it} > 0\}$.

The following algorithm gives the basic steps of a rolling horizon approach:

Algorithm 2 Rolling horizon algorithm

- 1: Initialize $t^c = 0$.
 - 2: Calculate the planning horizon for the given set of existing (P^E) and new patients (P^N).
 $T = \max\{t^c + \Delta + \max_{i \in P^N}\{C_i F_i\}, \max\{t : R_{it} > 0, i \in P^E\}\}$.
 - 3: Solve IP_1 for all patients in $P^E \cup P^N$ and the planning horizon of $[t^c + 1, T]$.
 - 4: **for all** $t = t^c + 1, t^c + 2, \dots, t^c + \Delta$ **do**
 - 5: Determine the set of patients who have treatment on day t .
 $P_t = \{i : R_{it} > 0 \text{ and } i \in P^E \cup P^N\}$.
 - 6: Solve IP_2 (or IP_3) for day t .
 - 7: **end for**
 - 8: After Δ days, find sets P^{NE} , P^O and P^R .
 $P^{NE} = \{i : X_{it} = 1, t \in [t^c + 1, t^c + \Delta], i \in P^N\}$
 $P^O = \{i : \max\{t : R_{it} > 0\} \in [t^c + 1, t^c + \Delta]\}$
 $P^R = \{i : est_i \in [t^c + 1, t^c + \Delta]\}$
 - 9: Update P^E and P^N .
 $P^E = P^E \cup P^{NE} \setminus P^O$
 $P^N = P^N \cup P^R \setminus P^{NE}$
 - 10: Increase t^c by Δ and go to Step 2.
-

At Step 1, the current time is initialized to zero. The set of existing patients (P^E) and new patients (P^N) are assumed to be known at time zero. At Step 2, the planning horizon T is calculated according to the treatment completion times of existing patients and maximum completion time of new patients. The planning problem IP_1 is solved at Step 3. The treatment start times are found for the time interval $[t^c + 1, t^c + \Delta]$. The scheduling problem is solved for each day $t \in [t^c + 1, t^c + \Delta]$ at Step 6. According to the solution of the planning problem, patients whose treatments have started (P^{NE}) and whose treatment cycles have terminated (P^O) during time interval $[t^c + 1, t^c + \Delta]$ are found at Step 8. The set of patients who are referred to receive chemotherapy are also found at this step. At Step 9, the sets of existing and new patients are updated. The schedules for time periods $t^c + 1 \dots t^c + \Delta$ are executed and the current time is updated at Step 10. Steps 2–10 are repeated every Δ days.

We present a numerical example to clarify the basic steps of the rolling horizon approach. Assume that there are two nurses ($N_t = 2$) and five chairs ($K = 5$) in an infusion clinic. Four hours is used as the normal working hours. The total working hours is 20 hours ($KH_t = 5 \times 4 = 20$ hours =

1200 minutes). Δ is chosen as seven days, which corresponds to a week. At the beginning of week 11, the set of existing patients is $P^E = \{1, 2, \dots, 80\}$ and the set of new patients is $P^N = \{81, 82, \dots, 96\}$. The planning problem is solved to assign new patients to days 71–75 (Monday through Friday) at week 11. The clinic is closed on weekends, therefore no patient is scheduled on days 76–77. Table 4 shows existing and new patients assigned to each day after the planning problem is solved. The bold numbers represent the new patients whose treatment started at week 11. Patients who have multiple day treatments can easily be seen in the table. For example, patient 14 has treatments on days 71, 72, and 73. The total workload (total treatment duration) on each day is also shown in the same table.

	Day, t	$P_t = \{i : R_{it} > 0, i \in P^E\} \cup \{i : X_{it} = 1, i \in P^N\}$																Total workload (minutes) (Existing + New)
During week 11 (Steps 3 and 5)	71	1	4	5	7	8	10	11	12	13	14	15	16	17	20			1185 (1155 + 30)
						37						59	60		62			
													93					
	72	23	34	35	36		10			13	14							1185 (1185 + 0)
				70			38	40	43	45		59	60		62			
	73						10				14							1155 (675 + 480)
											57	59		61				
							85							95	96			
	74						10					57	59					1155 (555 + 600)
		66		74	84	85			89		91	92		95	96			
	75						10											1005 (300 + 705)
							85	86	87		90	57	59			95	96	

Table 4 Numerical example - Planning

During week 11, the treatments of patients 84, 85, 86, 87, 89, 90, 91, 92, 93, 95, 96 (P^{NE}) are started, and treatment of patients 4, 8, 10, 35, 62 (P^O) are terminated. Ten new patients ($P^R = \{97, 98, \dots, 106\}$) are referred to receive treatment. At the beginning of week 12, sets of existing and new patients are updated as follows:

$$P^E = P^E \cup P^{NE} \setminus P^O = \{1, 2, \dots, 80\} \cup \{84, 85, 86, 87, 89, 90, 91, 92, 93, 95, 96\} \setminus \{4, 8, 10, 35, 62\}$$

$$P^N = P^N \cup P^R \setminus P^{NE} = \{81, 82, \dots, 96\} \cup \{97, 98, \dots, 106\} \setminus \{84, 85, 86, 87, 89, 90, 91, 92, 93, 96, 96\}$$

After the patients who have treatments on days 71–75 are determined, appointment scheduling problem is solved for each day. An example schedule for day 74 can be seen in Figure 5. Different colors show different nurses; patients 10, 59, 66, 74, 84, 92 are assigned to nurse 1, and patients 57, 85, 89, 91, 95, 96 are assigned to nurse 2. The numbers in the boxes show the patient number and the acuity level. For example, patient 96 with acuity level 2 (96, 2) is assigned to chair 1. Patients

92 and 96 are scheduled to arrive at time zero. Patients 74 and 89 are scheduled to arrive at time 15. They cannot be scheduled to arrive at time zero, because a nurse can start the treatment of only one patient. The completion time is 270 minutes, which corresponds to 30 minutes of overtime. Even though the total treatment duration assigned to day 74 is less than the total available nursing time ($1155 < 1200$), the actual schedule has overtime because of nurse capacity.

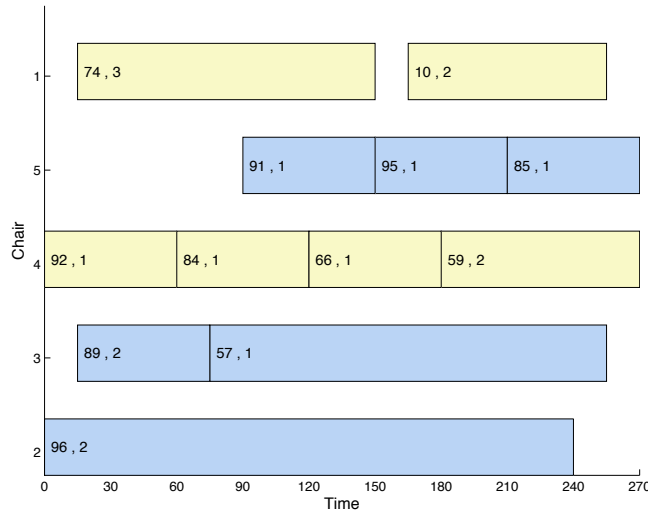


Figure 5 Example schedule for day 74

5. Comparison of proposed method with current practice

The aim of this section is to compare the proposed rolling horizon approach with current practice. In the following sections, we first explain the planning/scheduling method used in current practice and then use randomly generated patient mixes to compare the proposed rolling horizon approach with current practice.

5.1. Current practice

In current practice, after the oncologist sees the patient, the oncology nurse sends the scheduling request to the scheduler. The scheduler selects the patient whose appointment request is made first and schedules all treatments in one treatment cycle. The patient's treatments are scheduled to the first available time slot according to patient and nurse preferences. Even though the chair availability is considered, the acuity levels and nurse availabilities are not considered by the scheduler due to either no expertise in infusion nursing or no software capability to consider chair and nurse availabilities and treatment acuity levels at the same time. Figure 6 shows the planning/scheduling process performed by the scheduler.

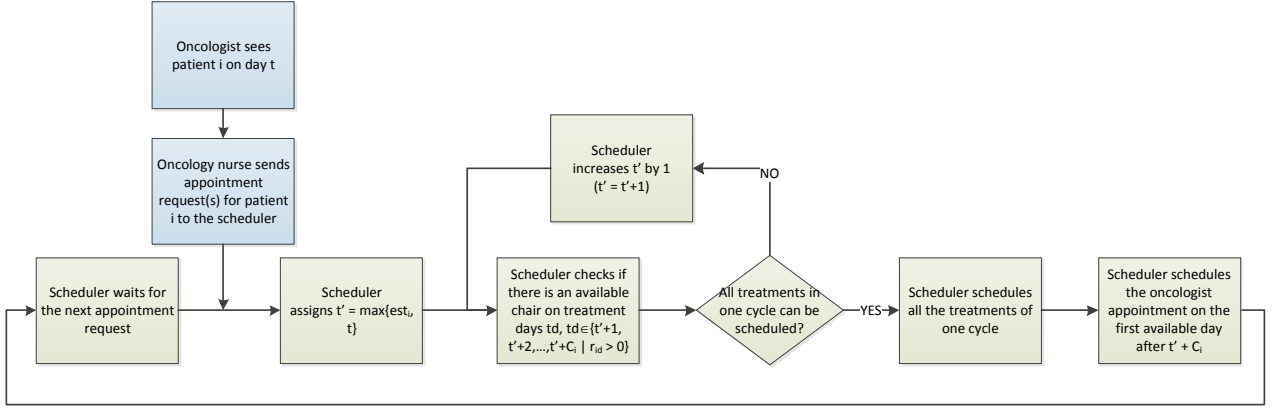


Figure 6 Planning and scheduling algorithm in current practice

After the schedules are made, the charge nurse in the infusion clinic assigns nurses to patients based on acuity levels, nurse skills, nurse working hours, and patient preferences. Nurse assignment is performed the day before the clinic session starts.

5.2. Experimental settings

The main input to the planning and scheduling problem is patient mix, that is, the number of patients, their prescribed treatment regimens, and their acuity levels. The cancer type and the treatment regimen are determined randomly. The cancer type is generated according to the expected number of new cases [2], and percentage of patients receiving chemotherapy [43]. We considered four cancer types with highest incidence rates (lung, breast, colorectal, and prostate). Table 5 shows the percentage of expected new cases over all new cases for each cancer type. The percentage of the patients receiving chemotherapy is based on the values for 2002 presented by Warren et al. [43], who used SEER-Medicare data for 306,709 persons aged 65 and older and diagnosed with breast, lung, colorectal, or prostate cancer.

When new patients are generated, the percentages in the last column are used. Thus, the probability that the patient will have lung cancer is 0.4184. The probabilities for breast, prostate and colorectal are 0.2540, 0.0717 and 0.2560, respectively. The treatment protocol of the patient is determined according to the cancer type. Fifty-nine treatment protocols obtained from NCCN website [33] are used in our computations. Each protocol has equal probability of being selected. The cycle length, number of cycles, treatment days, and treatment times for each treatment day change between 7-28 days, 1-6 cycles, 1-5 days, and 30-480 minutes, respectively.

The second important parameter is the acuity level of each chemotherapy regimen. None of the existing studies provide a comprehensive list of acuity levels for treatment regimens. In our

Cancer type	Percentage of new cases [2]	Percentage of patients receiving chemotherapy [43]	Number of regimens	Average treatment duration per day (min, max)	Percentage of patients in the patient mix ($14.83 \times 35.3 + 13.13 \times 24.2 + 13.00 \times 6.9 + 10.93 \times 29.3 = 1250$)
Lung	14.83%	35.3%	11	74 (15,240)	$14.83 \times 35.3 / 1250 = 41.84\%$
Breast	13.13%	24.2%	39	63 (15,240)	$13.13 \times 24.2 / 1250 = 25.40\%$
Prostate	13.00%	6.9%	4	98 (60,180)	$13.00 \times 6.9 / 1250 = 7.17\%$
Colorectal	10.93%	29.3%	5	137 (75,255)	$10.93 \times 29.3 / 1250 = 25.60\%$

Table 5 Patient mix

experiment, we used three acuity levels as in Chabot and Fox [9]. According to their patient classification system [9], we assigned acuity levels to the regimens according to the number of agents, pre-medications and complexity of administration. As the number of drugs and treatment duration increase, the acuity level increases.

The other parameters are related to the clinic environment. We consider an infusion clinic with 20 chairs and 7 nurses. The Poisson distribution with mean five is used to generate the number of new patients per day. The planning model assigns approximately 53.5 patients to each day (as explained in Section 5.3). A nurse sees approximately 7–8 patients per day. The problem size is comparable to the ones reported in the literature (see Table 1 in Section 2.2). The normal working hours are 8:00am – 4:00pm ($H_t = 480$ minutes). If the total completion time of all treatments exceeds eight hours, overtime cost is incurred. The slot length is chosen as 30 minutes and number of slots for normal working hours is 16. The clinic is assumed closed on weekends.

5.3. Planning

The proposed planning and scheduling methods and the current practice are coded in C++. The callable libraries of Cplex 12.0 are used to solve the integer programming models. All computations are performed in a personal computer with 2 GHz CPU and 2 GB memory. In our computations, we solve the planning problem for one year with a non-empty schedule at the beginning. The performance measures are number of patients assigned to each day, daily workload in terms of total treatment duration, total acuity violation, and total treatment delay.

The average number of patients assigned to each day is 53.5 for both current practice and proposed planning method. The standard deviation is 7.3 for current practice and 8.6 for the proposed method. The F-test to test the equality of variances shows that variances are not equal (p-value = 0.01). The number of patients assigned to each day can be a misleading performance measure due to high variability in treatment durations. Figure 7 shows the distribution of daily workload (total treatment duration in terms of number of slots) for both methods. The average workload per day is 185 slots (92.5 hours) for both methods. Since the fixed capacity is 320 slots (8 hours * 20 chairs = 160 chair hours = 320 slots), the chairs are not fully utilized. This is due

to patient mix with high acuities. The standard deviations of daily workload are 26.4 and 35.8 for current practice and the proposed method, respectively. F-test with a p-value of less than 0.001 shows that the proposed method gives daily plans with higher variances.

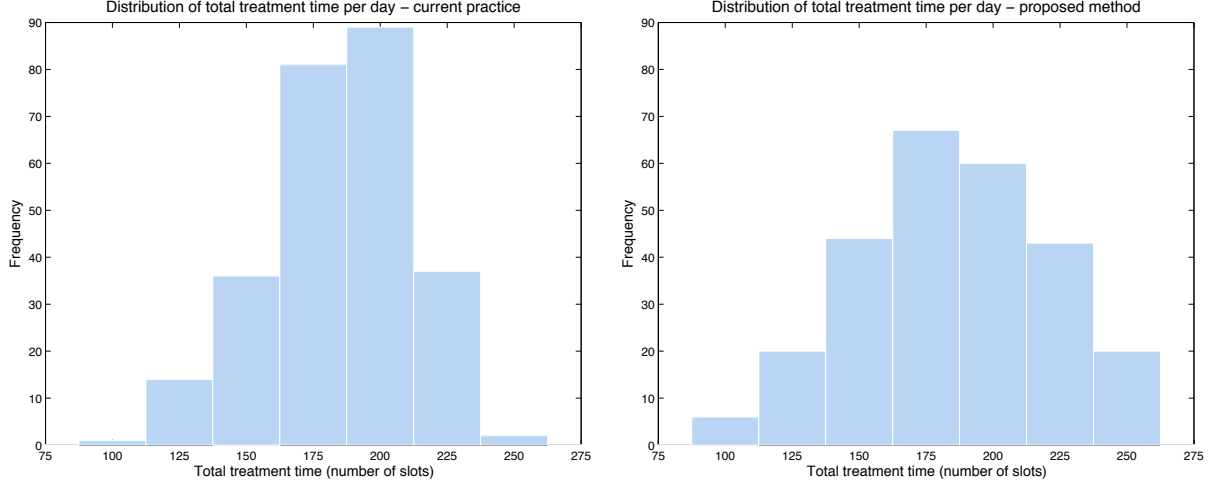


Figure 7 Distribution of daily workload (total treatment duration)

The next performance measure is total acuity violation. Current practice considers only chair availability while scheduling patients. Therefore, the generated schedules might not be feasible due to limited nurse availability. If the total acuity level assigned to a slot exceeds the available nurse capacity, either additional nurses are required to treat the patients on time or the patients have to wait until the nurses become available. Figure 8.(a) shows the distribution of total acuity violation per day for current practice. The acuity violation is calculated as $\sum_s \max(0, TotalAcuity_{st} - N_t A^{max})$. The average acuity violation per day is 94.6. Since a nurse can handle a total acuity of at most 64 ($A^{max} \times \text{Number of slots} = 4 \times 16$) in 8 hours, 1.5 additional nurse FTE is required to handle the workload. If the salary of an oncology nurse is \$40,000 per year, then the annual cost of additional nurses would be \$60,000 ($1.5 FTE \times 40,000 / FTE$). Even though total treatment time is much less than the available chair capacity, the acuity violation shows that the nurse capacity is exceeded by current practice. In contrast, the proposed method does not lead to acuity violation, since this is stated explicitly as a constraint.

Another important point that should be noted is that the workload throughout the day generated by the current practice shows high variability. While there are time slots with high acuity, there are underutilized time slots as well. In order to see if the underutilized time slots can be enough to accommodate the workload of overloaded slots, we should look at the difference between the total acuity violation and underutilization. Figure 8.(b) shows the histogram of these differences.

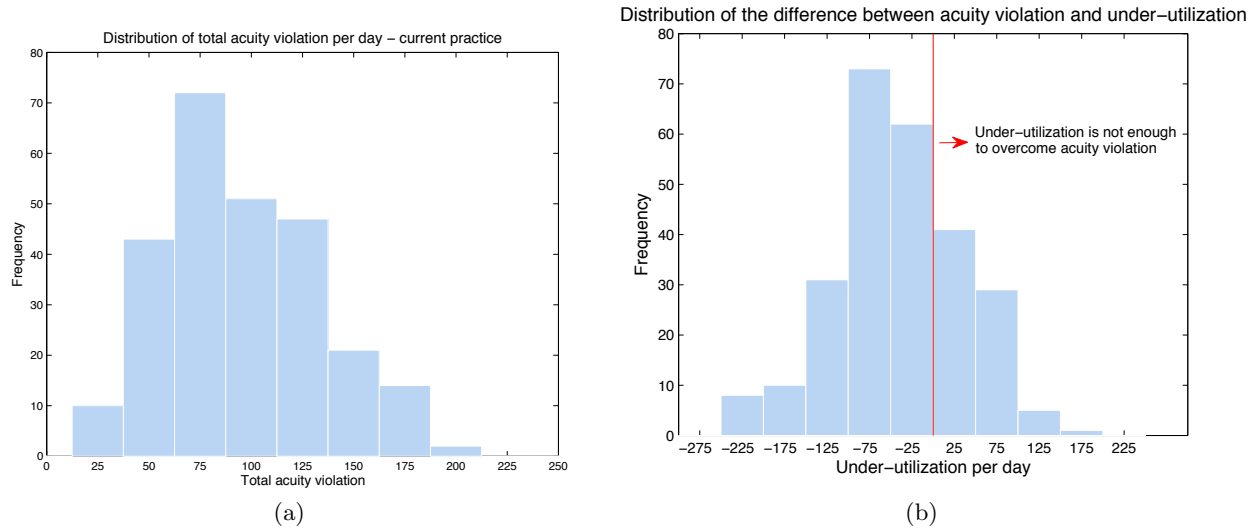


Figure 8 (a) Distribution of total acuity violation, (b) Distribution of difference between total acuity violation and under-utilization per day

In 76 days out of 260 days (29%), the underutilized time slots are not enough to accommodate the workload in overloaded time slots. This means current practice cannot generate feasible schedules (schedules free of acuity violation) unless additional resources and/or overtime are used. If no additional resources are used, the schedules generated by current practice will cause additional patient waiting times or safety problems due to heavy workload assigned to nurses. Since nurses have to perform several tasks such as assessment, education, IV access, and monitoring, the patients will have to wait while the nurse is completing the tasks for other patients. The heavy workload will increase the nurse burnout, which may increase the risk of making safety errors.

The objective of the planning model is to minimize the total treatment delay. Figure 9 shows the histograms of treatment delays per day. The treatment delay for a day is calculated for all patients planned on that day. The treatment delay for new patients is calculated as $(est_i - tX_{it})$, where est_i is the earliest start time for patient i and tX_{it} is the assigned treatment start time. The treatment delay for the existing patients is calculated as $(tX_{it} + (f - 1)C_i - tX_{it}^f)$, where $(tX_{it} + (f - 1)C_i)$ is the desired treatment start time of cycle f and tX_{it}^f is the actual start time of cycle f . The histograms show that proposed method can start patient treatments at their earliest start time on 133 days (51% of days). The current practice can do that for only 103 days (40% of days). The treatment delays range between 0 and 19 days for the proposed method, and between 0 and 45 days for the current practice. The average treatment delay is 4.0 days with current practice and 3.1 days for the proposed method. The standard deviation is 7.7 and 4.1 for current practice and the proposed method, respectively. Although the difference between the mean treatment delays is

not statistically significant (the p-value of paired t-test is 0.14), the variance of treatment delays is higher for the current practice (the p-value of the F-test is < 0.001). We believe the main reason for high variability in treatment delays among days is that the proposed method plans the treatments for all cycles at the beginning, which does not cause any other delay during the treatment. The proposed method also plans several patients simultaneously, which gives a smaller treatment delay compared to planning of patient treatments one by one.

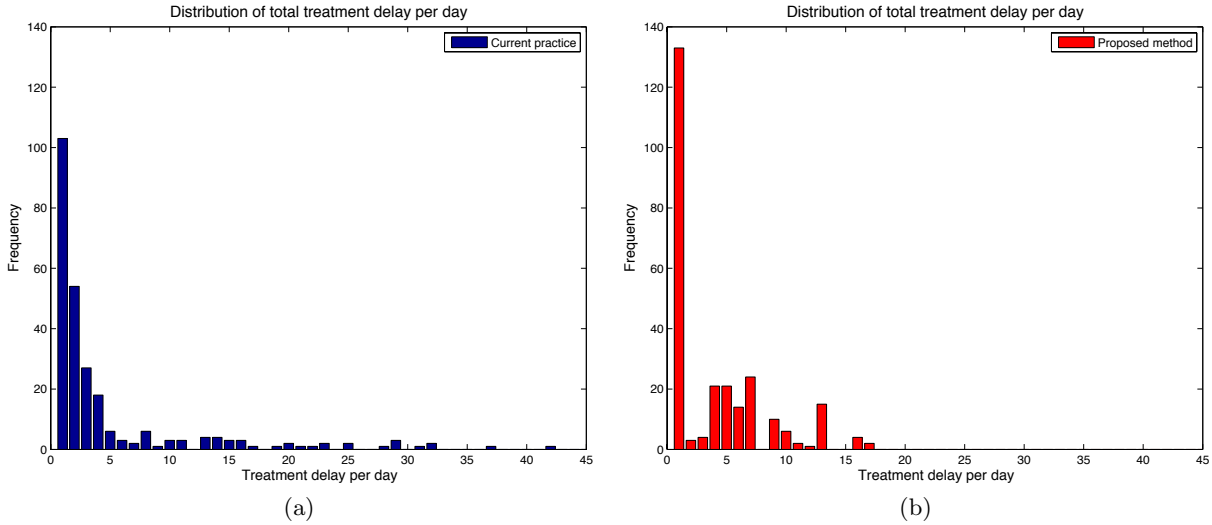


Figure 9 Distribution of treatment delay per day for current practice and the proposed method

Performance measure		Current practice	Proposed method	F-test (or t-test)	Significant difference?
Number of patients seen per day	Mean	53.5	53.5		
	Std. dev.	7.3	8.6	$p < 0.01$	Yes
Total workload per day (hours)	Mean	92.5	92.5		
	Std. dev.	26.4	35.8	$p < 0.001$	Yes
Acuity violation per day	Mean	94.6	No violation		
Acuity violation > underutilization	Mean	76 days	No violation		
Treatment delay (days)	Mean	4.0	3.1	$p = 0.14$	No
	Range	0–45	0–19		
	Std. dev.	7.7	4.1	$p < 0.001$	Yes

Table 6 Comparison of current practice and the proposed planning method: Summary of the results

To summarize these findings, the proposed method generates plans with greater variance in total daily treatment time than does current practice. This is due to the fact that current practice does not consider limits on daily acuity, whereas the proposed method generates plans with no acuity violation. Thus, while total treatment time from day to day is more consistent under current practice, nurses are more often overloaded with too much acuity assignment. That is, they are handling too many complicated tasks at the same time, which has implications for patient safety. In fact, protecting nurses and patients from acuity overload seems to come at the price of more

variation in total daily treatment time. On a final note, the proposed method provides smaller average treatment delay as well as smaller variance and maximum value of treatment delay, which bodes well for healthier patient outcomes.

5.4. Scheduling

It is important to verify that the generated plans can be used to find feasible daily schedules. We solve the scheduling problem using the proposed heuristic and integer programming model, IP_3 . We compare the daily schedules generated by the proposed integer programming model and current practice. The performance measures used for comparison are chair utilization, total acuity per slot, completion times (C_{max}), and computation times.

The scheduling problem is solved for 50 days. Figure 10 shows the number of chairs occupied and total acuity per slot for current practice and the proposed integer programming method, IP_3 . Since both methods consider chair availability, the number of patients at each slot does not exceed the total number of chairs. However, as is clear, the proposed method provides much better control over nurse workload throughout the day by guaranteeing that assigned acuity level does not exceed the required threshold. Under current practice, the charge nurse must either request additional nurses to treat all patients on time, or the patients should wait until the nurses become available to safely deliver the chemotherapy.

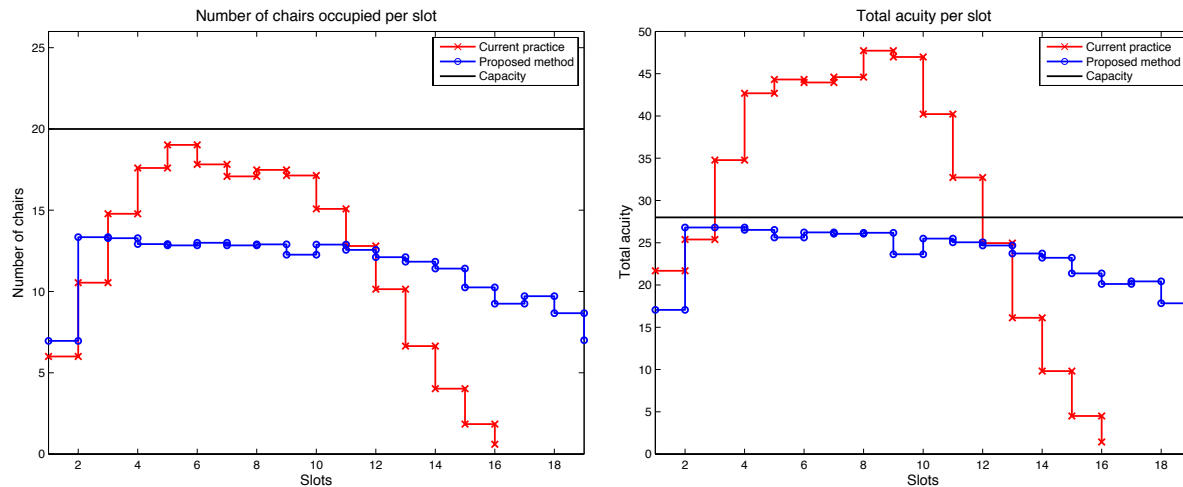


Figure 10 Chair and nurse utilization per slot (average of 50 schedules)

We compare the proposed integer programming model and the heuristic in terms of treatment completion times for each day. Figure 11 shows the histogram of the difference between completion times found by $ALTT$ and IP_3 . The proposed integer programming model gives the same results as the proposed heuristic for 14 days (28%), and gives better results for 36 days (72%). Paired

t-test (with a p-value < 0.001) shows that the difference between the methods is significant. The average difference between the completion times is 0.97 slots, which corresponds to 29.1 minutes. Since there are seven nurses in the clinic, the improvement in total nurse time is 3.4 hours. If the demand for chemotherapy treatment is high, more patients can be treated in that 3.4 hours. If one more treatment (with a treatment duration of less than 3.4 hours) can be added to each day, 260 more treatments can be performed in a year, which improves the access to care.

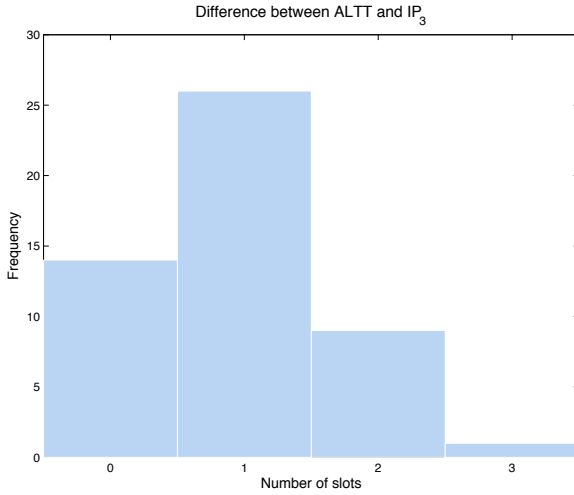


Figure 11 Difference between the completion times found by $ALTT$ and IP_3

Since IP_3 considers just nurse availability, we check the schedules in terms of chair feasibility. The maximum number of chairs is 18, which means all schedules are chair feasible. This may not be the case in other infusion clinics where there is limited space for the chairs. In that case, IP_2 that considers both chair and nurse availability should be solved or clinic managers should reduce the number of nurses and solve IP_3 again.

When we look at the computation times, IP_3 finds an optimal schedule in 13 minutes on the average. The proposed heuristic finds solutions in less than one second. To give an idea about the computation times of IP_2 , we solved five problems with a time limit of one hour. IP_2 could not find any optimal solution within the one-hour time limit. The percentage gap between the lower bound and the best feasible solution was 40% at the end of one hour. All these problems were solved optimally by IP_3 within 20 minutes, and the average computation time was 7.5 minutes. If the computation times are important or an optimization software is not available, then proposed heuristic that considers both chair and nurse availabilities can be used to find better schedules than current practice.

5.5. Recommendations for implementing optimization and heuristic methods

Due to high variability in clinic practices, we would like to make some recommendations to clinic managers about when and how the proposed methods can be used.

1. If there is a patient waiting list and reducing the waiting times for treatment is important for the clinic, then the proposed optimization-based planning method can reduce the treatment delays by planning several patient treatments simultaneously rather than planning each patient's treatment one by one. (In our experiments, the reduction in treatment delay was on average one day. This improvement was achieved by just considering on the average five patients (which is equal to the number of new patient arrivals per day) simultaneously instead of one patient at a time, and planning all treatments instead of only one cycle at a time.)

2. The proposed rolling horizon approach is flexible in the sense that it can be solved as frequently as needed (every day or every week). That means, if an urgent patient arrives, the treatment plan can be prepared without waiting for the next planning horizon.

3. The importance of treatment delays (w_{id}) can be determined by oncologists. The stage of the disease, invasiveness of the tumor, and patients health condition are the factors that might be important in determining w_{id} .

4. The decision to use one of the scheduling methods (IP_2 , IP_3 , or $ALTT$) should be based on the patient mix and available computational resources. If the patients have high acuities, IP_3 should be used because nurses would be the limiting resource. If the patients have low acuities, then IP_2 that considers both chair and nurse availabilities should be used. If the computation time is important and/or an optimization package is not available, then $ALTT$ can be used.

5. Time studies or nurse experience can be used to determine acuity levels for each treatment. Treatment acuity levels can be adjusted after the first treatment according to difficulty and workload the nurse experiences. The charge nurses who do the nurse assignments can determine the maximum acuity level a nurse can handle safely at any given time.

6. The computational results show that there is high variability in treatment completion times among days. In order to reduce to variability in completion times, we recommend the clinic managers to adjust the number of nurses by solving the scheduling problem with different number of nurses to have similar completion times among different days. Another way is to add workload balancing as an another objective to the planning model. We would like to refer the reader to the studies by Sadki et al. [38, 39] how workload balancing can reduce the variability between days. We are currently developing planning models to find a more balanced workload (in terms of treatment duration and total acuity) among days while keeping the treatment delays low.

7. If staggered nurse schedules are used, the proposed scheduling methods can easily handle nurse working hours with additional constraints.

8. The proposed scheduling methods assume that the appointment times and nurse/chair assignments are made once all patients assigned to a day are known. However, it is not practical for the clinic to wait until the last day to schedule all the patients. Therefore, a cut-off point like one week can be determined by the clinic. All patients that are assigned to come on a day can be scheduled one week before the appointment day. If new patients arrive after the schedule is generated, they can be added to the existing schedule by using the proposed scheduling methods.

6. Concluding remarks and future research

In oncology clinics, the scheduling of treatments (surgery, radiotherapy, chemotherapy, etc.) for cancer patients is very important in delivering the right care at the right time. We consider planning and scheduling problems for chemotherapy patients. Chemotherapy drugs are given to patients on several days with the objective of minimizing the number of cancer cells while sparing the normal cells. Adherence to chemotherapy protocols is crucial in achieving the best health outcomes. Our aim is to achieve the best health outcomes by minimizing the treatment delays due to limited resources.

Most previous studies propose using scheduling templates/rules based on nursing or pharmacy times. The scheduling decisions are made in an ad-hoc manner according to physician and scheduler experiences and patient preferences. To the best of our knowledge, there is no study that proposes optimization methods to schedule chemotherapy treatments considering acuity levels and optimizing several objectives such as minimization of treatment delay, minimization of staff overtime and under-utilization, and maximization of staff utilization.

We proposed a two-stage approach to solve the planning and scheduling problems sequentially. Integer programming models are proposed to solve these problems. A rolling horizon approach is used to schedule new patients who are referred to receive chemotherapy. A heuristic and an integer programming model are proposed to reduce the computation times for the scheduling problem. The computational results show that the planning and scheduling problems can be solved in reasonable times. The proposed planning and scheduling models can be used as a decision making tool in determining the optimal staffing levels.

In a real clinic environment, there are many uncertainties such as delays in getting lab results, cancellations, add-on patients, and variability in treatment durations, that affect the daily performance. Delays increase the patient waiting times and affect the clinic flow. Most of the cancellations

occur on the same day after the lab results are performed and the patient is seen by the oncologist. Add-ons might occur within the last few days. One future research area is developing stochastic planning and scheduling methods that consider uncertain treatment durations, cancellations and add-ons. One drawback of the proposed deterministic planning model is the high variability in workload among days. As a future research, the workload balancing objective can also be added to the stochastic planning model to find plans with a more balanced workload among days while keeping the treatment delays low.

There are several clinics where patients are scheduled according to chair availability and nurse staffing/nurse assignments are performed after the schedule is generated. Nurse assignment is either performed by the charge nurse according to the patient acuities, nurse skills, nurse working hours, and patient and nurse preferences, or the nurse assignment is done in real time where the arriving patient is assigned to the nurse who has the least number of patients. Even though we mentioned about nurse assignment when we explained the current practice in Section 5.1, we did not use any staffing method to find feasible nurse assignments for a given schedule. Another future research area is developing nurse assignment/staffing models that incorporate patient acuities, nurse skills, and patient and nurse preferences.

References

- [1] Aboumater H. J., Winner L. E., Davis R. O., Trovitch P. B., Berg M. M., Violette K. M., Messersmith W. A., Maylor K. K., Lehmann C. U. 2008. No time to waste: Decreasing patient wait times for chemotherapy administration using automated prioritization in an oncology pharmacy system. *Am J Manag Care* **14**(5) 309–16.
- [2] American Cancer Society. 2010. *Cancer Facts & Figures*. Atlanta: American Cancer Society.
- [3] Baldwin C. 2006. Cancer clinic resource scheduling: Your resource to productivity. *Hematology and Oncology News* 18–20.
- [4] Barton-Burke M., Wilkes G. M., Ingwerson K. 2001. *Cancer Chemotherapy: A Nursing Approach*. Jones and Bartlett Publishers, Inc., Sudbury, MA.
- [5] Bonadonna G., Valagussa P., Moliterni A. 1995. Adjuvant cyclophosphamide, methotrexate, and fluorouracil in node-positive breast cancer. *New England Journal of Medicine* **332** 901–906.
- [6] BSCS & Videodiscovery. 1999. *Cell Biology and Cancer*. NIH publication No. 99-4646.
- [7] Buerhaus P., Donelan K., DesRoches C., Lamkin L., Mallory G. 2001. State of the oncology nursing workforce: Problems and implications for strengthening the future. *Nursing Economics* **19**(5) 1–11.

- [8] Cairo M. 2000. Dose reductions and delays: limitations of myelosuppressive chemotherapy. *Oncology* **14**(supplement 8) 21–31.
- [9] Chabot G., Fox M. 2005. The creation of a patient-classification system in an outpatient infusion center setting. *Oncology Nursing Forum* **32**(3) 535–538.
- [10] Chang J. 2000. Chemotherapy dose reduction and delay in clinical practice: evaluating the risk to patient outcome in adjuvant chemotherapy for breast cancer. *European Journal of Cancer* **36**(supplement 1) 11–14.
- [11] Collingwood J. 2005. A snapshot of oncology nursing staffing. *Oncology Issues* 41.
- [12] Cusack G., Jones-Welles A., Chisholm L. 2004. Patient intensity in an ambulatory oncology research center: A step forward for the field of ambulatory care. *Nursing Economics* **22**(2) 58–63.
- [13] Cusack G., Jones-Welles A., Chisholm L. 2004. Patient intensity in an ambulatory oncology research center: A step forward for the field of ambulatory care - part III. *Nursing Economics* **22**(4) 193–195.
- [14] de Raad J., van Gool K., Haas M., Haywood P., Faedo M., Gallego G., Pearson S., Ward R. 2010. Nursing takes time: Workload associated with administering cancer protocols. *Clinical Journal of Oncology Nursing* **14**(6) 735–741.
- [15] Delaney G., Jalaludin B., Moylan E., Barton M. 2002. The development of a model of outpatient chemotherapy delivery - Chemotherapy Basic Treatment Equivalent (CBTE). *Clinical Oncology* **14** 406–412.
- [16] DeLisle J. 2009. Designing an acuity tool for an ambulatory oncology setting. *Clinical Journal of Oncology Nursing* **13**(1) 45–50.
- [17] Diedrich J., Plank D. P. 2003. Efficient system to schedule chemotherapy and support therapies for oncology nurses. *Oncology Nursing Society Congress Abstracts*. Abstract 76.
- [18] Dobish R. 2003. Next-day chemotherapy scheduling: a multidisciplinary approach to solving workload issues in a tertiary oncology center. *Journal of Oncological Pharmacy Practice* **9** 37–42.
- [19] Erikson C., Salsberg E., Forte G., Bruinooge S., Goldstein M. 2007. Future supply and demand for oncologists : Challenges to assuring access to oncology services. *Journal Oncology Practice* **3**(2) 79–86.
- [20] Gaits V. A. 2005. Nurse resource allocation in ambulatory cancer centers: Guidelines for clinicians and executives. *Oncology Issues* 36–42.

-
- [21] Gesell S., Gregory N. 2004. Identifying priority actions for improving patient satisfaction with outpatient cancer care. *Journal of Nursing Care Quality* **19** 226–233.
- [22] Grannan M., Roberson J., Fine B., Buswell L., Thompson L. 2002. Working together on “the wait”. *Oncology Nursing Forum* **29**(356) Abstract 106.
- [23] Gruber M., Kane K., Flack L., Abbotoy J., Recchio J., Williamson K., Horan K., McCarthy P. 2003. A perfect day work redesign in a chemotherapy and infusion center. *Oncology Nursing Forum* **30**(4) 567–568.
- [24] Hawley E., Carter N. G. 2009. An acuity rating system for infusion center nurse staffing: The cleveland clinic cancer center at hillcrest hospital experience. *Oncology Issues*. November/December, 34–37.
- [25] Ireland A., DePalma J., Arneson L., Stark L., Williamson J. 2004. The oncology nursing society ambulatory office nurse survey. *Oncology Nursing Forum* **31**(6) E147–E156.
- [26] IU Simon Cancer Center. 2006. *Wilms and other renal tumors*. <http://www.cancer.iu.edu/renaltumors/wilms/timeline/>.
- [27] Jones A., Cusack G., Chisholm L. 2004. Patient intensity in an ambulatory oncology research center: A step forward for the field of ambulatory care - part II. *Nursing Economics* **22**(3) 120–123.
- [28] Lamkin L., Rosiak J., Buerhaus P., Mallory G., Williams M. 2001. Oncology nursing society workforce survey part i: Perceptions of the nursing workforce environment and adequacy of nurse staffing in outpatient and inpatient oncology settings. *Oncology Nursing Forum* **28**(10) 1545–1552.
- [29] Langhorn M., Morrison C. 2001. Redesigning processes in ambulatory chemotherapy: Creating a patient appointment scheduling system: Part II. *Canadian Oncology Nursing Journal* **11**(3) 157–159.
- [30] McNulty E., O’Sullivan M., Smith E. P. 2001. Pre-mixing chemotherapy agents to decrease patient waits in an ambulatory setting. *Oncology Nursing Forum* **28**(329) Abstract 58.
- [31] Moore M., Hastings C. 2006. The evolution of an ambulatory nursing intensity system: measuring nursing workload in a day hospital setting. *Journal of Nursing Administration* **36**(5) 241–248.
- [32] National Center for Biotechnology Information. 2004. *A Science Primer*. http://www.ncbi.nlm.nih.gov/About/primer/genetics_cell.html.
- [33] National Comprehensive Cancer Network. 2009. *NCCN Chemotherapy Order Templates*. <http://www.nccn.org/ordertemplates/default.asp>.

-
- [34] National Institute of General Medical Sciences. 2003. *Findings*. <http://publications.nigms.nih.gov/findings/sept03/mix.html>.
- [35] Oncology Nursing Forum. 2004. *Medicare Reimbursement for Chemotherapy Administration and Related Practice Expenses, Issue Brief*.
- [36] Pinedo M. 2009. *Planning and Scheduling in Manufacturing and Services*. Springer, New York, NY.
- [37] Rosenthal D. 2007. Consequences of mucositis-induced treatment breaks and dose reductions on head and neck cancer treatment outcomes. *Journal of Supportive Oncology* **5**(9) 23–31.
- [38] Sadki A., Xie X., Chauvin F. 2010. Bedload balancing for an oncology ambulatory care unit. *8th International Conference of Modeling and Simulation*.
- [39] Sadki A., Xie X., Chauvin F. 2010. Patient assignment for an oncology outpatient unit. *6th Annual IEEE Conference on Automation Science and Engineering*. 891–896.
- [40] Sepulveda J. A., Thompson W. J., Baesler F. F., Alvarez M. I., Cahoon L. E. 1999. The use of simulation for process improvement in a cancer treatment center. Farrington, P. A., Nemphard, H. B., Sturrock, D. T., Evans, G. W., eds., *Proceedings of the 1999 Winter Simulation Conference*. 1541–1548.
- [41] U.S. Census Bureau. 2000. *Projected Population of the United States, by Age and Sex: 2000 to 2050*. <http://www.census.gov/population/www/projections/usinterimproj/natprojtab02a.pdf> (accessed Feb 2011).
- [42] van Lent W. A., Goedbloed N., van Harten W. 2006. Improving the efficiency of a chemotherapy day unit: Applying a business approach to oncology. *European Journal of Cancer* **45** 800–806.
- [43] Warren J. L., Yabroff K. R., Meekins A., Topor M., Lamont E. B., Brown M. L. 2008. Evaluation of trends in the cost of initial cancer treatment. *Journal of National Cancer Institute* **100** 888–897.
- [44] West S., Sherer M. 2009. The "right" nurse staffing model. *Oncology Issues* 26–30.
- [45] Wood W., Budman D., Korzun A. 1994. Dose and dose intensity of adjuvant chemotherapy for stage II, node-positive breast carcinoma. *New England Journal of Medicine* **330** 1253–1259.