Operations Research Frameworks for Improving Make-Ahead Drug Policies at Outpatient Chemotherapy Infusion Centers

by

Donald B. Richardson

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Industrial and Operations Engineering)
in The University of Michigan
2019

Doctoral Committee:

Professor Amy E.M. Cohn, Chair
Associate Professor Seth D. Guikema
Associate Professor Mariel S. Lavieri
Professor Lisa A. Prosser

Donald B. Richardson

donalric@umich.edu

ORCID: 0000-0002-5094-7548

Dedication

I dedicate this dissertation to the Meyerhoff Scholar program, the late Lamont Toliver, Keith Harmon, Mitsue Wiggs, Michael Goodwyn, Taifa Simpson, Jackie King, Ernestine Baker, Alicia Hall, Sharon Johnson, and all other M's. Each of you gave/give so much to this program which provided me with another support network and the skills to navigate so many challenges grad school threw at me. Thank you for teaching me what a PhD meant and for setting me up to pursue one.

Acknowledgments

"If I have seen further, it is by standing upon the shoulders of giants"–*Sir Isaac Newton*. From my committee, friends, family, and mentors, I am forever grateful for all the giants, who without, I would not have gotten to this point. First I want to start with my advisor and mentor Amy. Thank you so much for taking a chance on me as a Summer Research Opportunity Program (SROP) student who didn't even know how to formulate an LP model before starting my first summer at the Center for Healthcare Engineering and Patient Safety. Even with such a large group, you were so personable and continue to have all of your students' best interests at heart first. You have provided me with so many opportunities from presenting at major conference across the world (literally) to introducing me to such an amazing network of colleagues and friends along the way. Your lessons translate into my research, writing, communication, and project management skills. For this, I cannot be more thankful.

I want to thank each of my committee members for their guidance these last couple of years. Seth, thank you for working with me even after your course ended to work on the predictive modeling work. Mariel, thank for your continued guidance with my dissertation work as well as everything you have done through the IOE

diversity initiative. Lisa, thank you for meeting with me on several occasions and providing such valuable insights from the public health perspective.

Next I want to thank my mom and my dad, both continue to show their love and support as I went through this process as they have done through my life. I also want to thank my extended family who always support and celebrate each academic milestone with me..

Josi, thank you for going through this PhD journey with me and being the best partner a man could ask for these last 4 years. Your love, dissertation proofreading, and constant support goes above and beyond what I could even imagine. I can't wait to walk across the stage together next May.

Next thank you to all of the students both past and present that supported my work. Francisco, thank you for all of our brainstorming sessions with coding and just random "the next big thing" ideas. Matt See, thank you for your willingness to learn and now provide extremely valuable data analysis for this work. It has been amazing to see your growth as you soon will graduate and pursue a PhD of your own. The rest of the chemo team, Ajaay, Hwon, Matt R., Pamela, Jeremy, Elizabeth, Hass, Sarah, Allison, and anyone else who worked with us on this project. You all helped set up such am important problem that we address in the dissertation, thank you.

Thank you to all of the faculty and staff who have touched me along the way. Gisselle, thank you for running SROP and helping me navigate UM's PhD program. You continue to support and love all of your SROP students which I always felt. Gene, thank you for your constant wit and being the backbone that keeps CHEPS running. I want to thank all of the IOE staff again that went above and beyond

to support me during my time here, especially Tina, Matt, Wanda, Rebekah, and Gwen,

I also would like to thank all of our collaborators of the work from the University of Michigan Rogel Cancer Center whom without these projects would not have been possible. I want to thank Carolina Typaldos for always keeping us on track to accomplish the goals on the hospital end but also with my PhD work. Additionally I want to thank Dr. Jerome Winegarden, an oncologist with IHA Hematology Oncology Consultants, for his graciousness and time for providing such valuable insights regarding the complexities of chemotherapy treatment and management of such drugs in the healthcare setting.

I want to thank all of my friends both pre-grad school and those I have made along the way. Each of you have kept me grounded and provided so many outlets to help keep balance in my life. Lastly, I want to thank the Seth Bonder Foundation for their support. Merrill Bonder continues to shower us with support as she attends every conference and cheers us on. I would also like to thank Rackham for supporting me through the Rackham Merit Fellowship.

Table of Contents

List of Figures

Figure

List of Tables

Table

List of Appendices

Appendix

Abstract

Outpatient chemotherapy infusion is one of the most common forms of treatment used to cure, control, and ease symptoms of cancer. Patients who require outpatient chemotherapy infusion undergo lengthy and physically demanding infusion sessions over the course of their treatment. While the frequency and duration of visits vary by patient, drug, and cancer type, most patients will require several treatments over the course of months or years to complete their regimen/treat their disease. Receiving infusion is just one part of the complex treatment process. Patients may have their blood work done, wait for the results to process, visit their oncologist, wait on their order to be placed by their oncologist and prepared by the pharmacy, and then have the infusion administered by infusion clinic staff. Each step introduces randomness which can lead to propagated delays. These delays negatively affect patients as well as clinical operation cost and staff workload.

We focus on optimizing drug preparation at the pharmacy to reduce patient delays. Drugs can be prepared the morning before patients arrive to prevent the patient from waiting the additional time needed to prepare their prescribed drugs in addition to any other wait time incurred during peak pharmacy hours. However, patients scheduled for outpatient chemotherapy infusion sometimes may need to cancel at the last minute even after arriving for their appointment (i.e. patient may

be deemed too ill to receive treatment). This results in the health system incurring waste cost if the drug was made ahead since the drugs are patient specific and have a short shelf life. Infusion centers must implement policies to balance this potential waste cost with the time savings for their patients and staff. In support of this effort, this dissertation focuses on methods and strategies to improve the process flow of chemotherapy infusion outpatients by optimizing pharmacy make-ahead policies.

We propose using three different methods which build upon each other. First we develop a predictive model which utilizes patient-specific data to estimate the probability that a patient will defer or not show for treatment on a given day. Generally, the ability to generate high-quality predictions of patient deferrals can be highly valuable in managing clinical operations, such as scheduling patients, determining which drugs to make before patients arrive, and establishing the proper staffing for a given day. We also introduce how the patient-specific probability of deferral can help determine a "general rule of thumb" policy for what should be made ahead on a given day.

Next we utilize these probabilities in two integer programming models. These multi-criteria optimization models prioritize which and how many drugs to make ahead given a fixed window of time. This is done with the dual objectives of reducing the expected waste cost as well as the expected value of reduced patient waiting time.

Lastly, we utilize simulation to better quantify the impact of our proposed policies. We show that making chemotherapy drugs ahead of an infusion appointment not only benefit the patient they are prescribed for but also subsequent patients due to the decrease load (i.e., reduced blocking) on the pharmacy system as a whole.

Each method utilizes electronic medical record data from the University of Michigan Rogel Cancer Center (UMRCC) but may be generalized to any cancer center infusion clinic.

Chapter 1

Introduction

1.1 Motivation

Cancer is the second leading cause of death in the United States (U.S.), with the number of estimated annual cases increasing from 1.3 million in 2005 to 1.8 million in 2019 [1]. With most patients undergoing a combination of chemotherapy infusion (i.e., intravenously-administered medication), radiation therapy, and surgery, over half of cancer patients in the U.S. will require some form of chemotherapy. Additionally, cancer treatments continue to advance and increase in complexity. This results in an increase in the frequency of patient infusion visits for a given treatment regimen as well as an increase in cancer survivors. Many of these cancer survivors need continued visits to help in the prevention and detection of new or recurrent cancers. They also may need treatment for secondary health problems caused by the chemotherapy. Aside from complexity, as the FDA approves more chemotherapy drugs with increased survival rates, even more patients are opting into chemotherapy [2]. Consequently, this increased demand of patients at outpatient

chemotherapy infusion centers leads to increased patient waiting time, patient dissatisfaction, and overworked staff. Furthermore, patient and nurse safety becomes a concern as demand increases. The risk of the hazardous chemotherapy agents spilling on patients or nurses increases when the nurses are overworked [3], [4], [5]. Therefore, any process improvement on the system has the potential for multifaceted benefits [6].

1.2 Chemotherapy Infusion Drugs

Often, chemotherapy is a drug mixed in solution that is used to treat one's disease. Most patients receive chemotherapy treatment in conjunction with radiation therapy or surgery which focus on removing cancer cells from certain areas in the body. There are over 100 chemotherapy drugs approved for use in the U.S. including some oral drugs [7]. Broadly, these drugs target cells at various stages of the new cell replication process or on a specific and unique molecular target [8]. Although cancer cells grow more rapidly than most healthy cells, chemotherapy drugs may damage both cancer and healthy cells that rapidly divide. The balance between treating a patient's cancer and managing the side effects of such treatment is principal to the role of an oncologist.

While the primary goal of using chemotherapy infusion is to kill cancer cells, there are two main reasons an oncologist prescribes chemotherapy. (1) Curative intent chemotherapy can be given alone as in the case when the cancer is particularly sensitive to the effects of chemotherapy (i.e. Lymphoma or Luekemia) or in combination with surgery or radiation for those that are not. It is also used before

surgery or radiation therapy to shrink a tumor (i.e., neoadjuvant therapy), after surgery to help kill any remnant cancer cells (i.e., adjuvant therapy), or in conjunction with other treatments if a patients' cancer returns. [9] (2) If the cancer is too far progressed or no longer seen as medically curable, then the goal may focus on palliating the disease. Chemotherapy is used to shrink tumors or stop the cancer from spreading which can increase a patient's quality and length of life. There are cases where the cancer will never go away (e.g., low grade lymphoma) but can be managed through treatment similar to other chronic diseases such as diabetes. Oncologists also might recommend chemotherapy to ease the symptoms caused by the cancer. Again, the goal of the chemotherapy in this case is to improve the quality of life for a patient (e.g. shrink a tumor that may be pressing against a nerve).

Due to the complexity of treatment and nature of chemotherapy drugs, many factors are considered when determining what type of drug a patient receives, the frequency, and the dose. Drug choice is mainly driven by the type of cancer, stage/progression, patient's age, patient's current health state, and the patient's previous cancer treatment history. While doses are typically based on the patient's body weight or body surface area, some must be adjusted to better accommodate patients with other health complications or to reduce interference with other ongoing medications or treatments. The chemotherapy dose must stay within the narrow margin of being safe yet effective. Similarly, the frequency at which these drugs are given is determined based on multiple factors. The main goal is to allow sufficient time for normal cells to recover from drug side effects. The frequency magnitude can vary between days and weeks.

1.2.1 Make-ahead Drugs

Most pharmacies cannot stockpile prepared chemotherapy agents due to their limited viability. Some drugs must be administered immediately while others can last in storage for up to 24 hours. All chemotherapy drugs are assigned a ***hang-by time*** which indicates how many hours the infusion center has until they must start administering the drug (i.e., similar to a sell-by date rather than an expiration date).

In this dissertation, we show chemotherapy drugs can be prepared (i.e., compounded) ahead of time to reduce the time a patient waits for their drug to be made similar to [10]. The *compounding process* consists of combining the chemotherapy drug agent with a solution (such as saline or dextrose). Some drugs also require additional time for the agent to settle in the solution. Depending on the drug, this compounding process can take up to an hour to complete for a single dose after all safety protocols are followed. However, given the inherent nature of the sick patient population served, there is a risk of making a last minute cancellation due to not being well enough for treatment after arriving for their appointment [11]. Consequently, if their drug was made ahead, the infusion center could incur a waste cost. However, if there are multiple patients scheduled to receive the same dose of the drug then the subsequent patient would be able to utilize it if the intended patient defers and if the drug remains viable. Infusion centers must implement policies that balance this potential waste cost with the time savings for their patients and staff.

1.2.2 Chemotherapy Infusion Drug Cost

When discussing the cost of chemotherapy infusion drugs, there are multiple perspectives to take and factors to consider. This includes what the hospital pays to stock the drug, what is billed to insurance companies, and what the patients get billed. However this doesn't include the administration cost, any facility fees, or cost associated with compounding the drug (i.e., technician labor cost). This research defines any cost as the amount of money that the hospital would have paid for the drug had it been used. While the hospital buys the stock drugs in standard vial or unit sizes, once compounded the cost range for a single dose of a drug is between \$10-\$25,000 (Note: These cost are based on infusion pharmacy data from the University of Michigan Rogel Cancer Center). With such high cost, hospital administrators are reluctant to approve making all patient drugs ahead of time due to the fear of waste caused by last minute cancellation out of the patients control. The hospital cannot bill a patient or their insurance if they are not well enough to receive treatment.

1.3 University of Michigan Rogel Cancer Center (UMRCC)

At UMRCC, chemotherapy infusion outpatient visits consist of getting blood work done in the phlebotomy lab, seeing their oncologist, waiting for the pharmacy to prepare their chemotherapy drug, and receiving their infusion. However, recurring patients do not always have a clinic visit as seen in Figure 1.1. Although patients are not directly involved with the pharmacy process, the drug preparation phase of their

visit is a key contributor to the overall time a patient spends in the infusion center. While a patients drug is bering prepared, they are either waiting in the waiting room or infusion area being prepared for treatment. Therefore a shorter turnaround-time (TAT) in the pharmacy means a shorter overall time in the system for the patient.

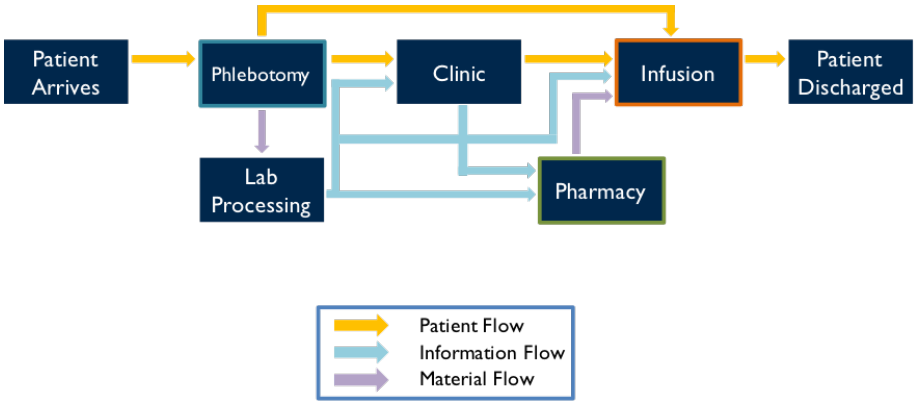


Figure 1.1: Flow of outpatients and their information based on observations and interviews conducted with the medical staff at UMRCC

The patient's orders are only sent to the pharmacy once the patient finishes their clinic visit or checks into the infusion area. The goal is to keep the TAT under 1 hour for each patient. However, through observation and historical data, we see that the TAT can be as long as 2 hours.

As seen in Figure 1.2, drug orders must go through a series of safety checks, compounding, and labeling to ensure the safety of patients and healthcare providers. Our focus is to improve the drug TAT in the pharmacy and in turn reduce the overall time in system (i.e., patient's length of stay at the infusion center) for the patients.

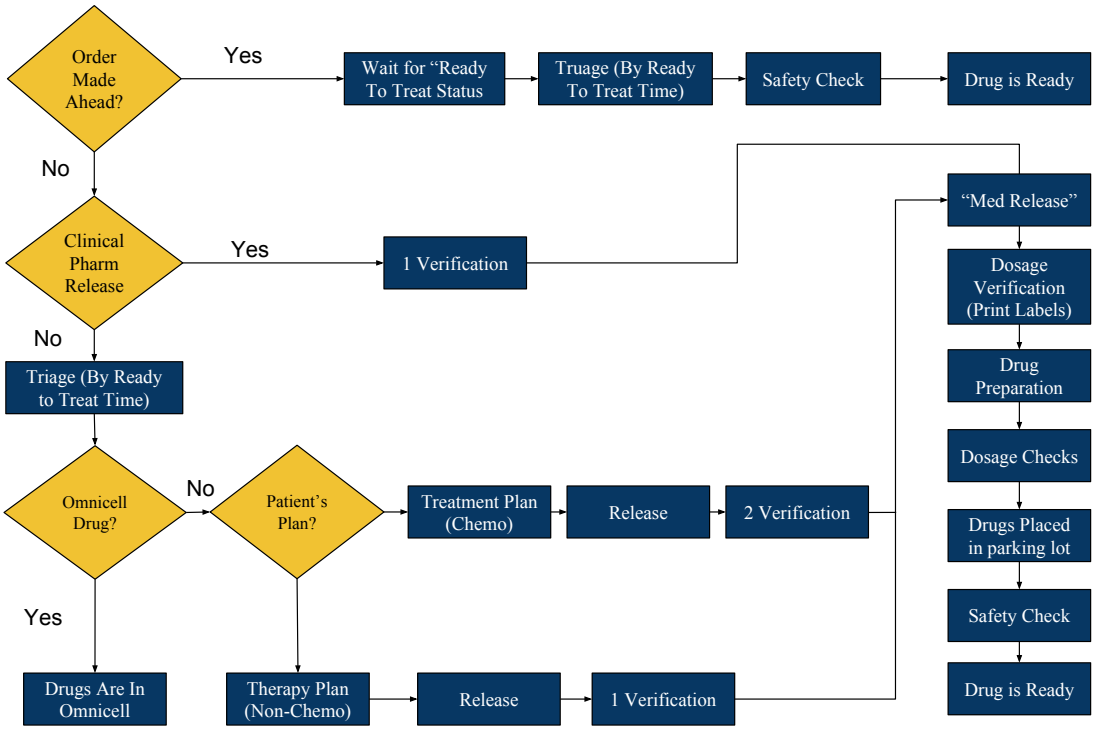


Figure 1.2: Drug order process flow based on observations and interviews conducted with the pharmacists at UMRCC

1.4 Dissertation Summary

In this dissertation, we focus on improving outpatient infusion center pharmacies' efficiency by proposing methods to generate and evaluate make-ahead chemotherapy drug policies (i.e., pre-mix policies) to ensure the drugs are available as soon as the patient is deemed ready to receive them. However when deciding which drugs to pre-mix, the decision maker must evaluate the trade-off between the potential

time saved by making the drugs and the risk of waste if the patient defers (i.e., last minute cancellation due to provider deeming patient not well enough or through self election). We discuss three methods which work in conjunction with each other as seen in Figure 2.2.

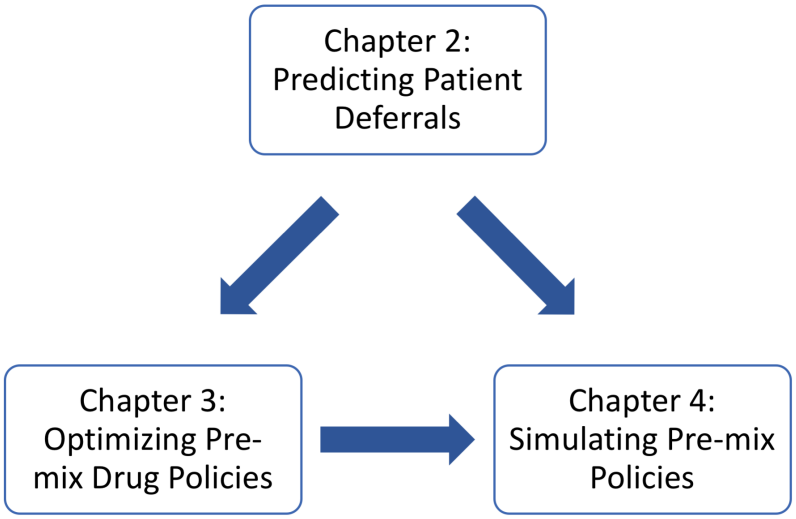


Figure 1.3: Work flow for all methods discussed in this dissertation

In Chapter 2 we discuss reasons for last minute cancellations and then develop a predictive model to better determine the chance a patient will defer their treatment. These high-quality predictions directly feed into determining expected time savings and expected waste cost if a particular drug is pre-mixed. The predictions can also be utilized when scheduling patients, establishing proper staffing both in the pharmacy and infusion area, and our focus of determining which drugs to pre-mix.

In Chapter 3 we utilize these probabilities in a binary integer programming model to determine the most advantageous set of drugs to pre-mix. Given a fixed window of time, our model evaluates a multi-criteria objective by maximizing expected dollar

value of reduced wait while minimizing the expected waste cost. This pre-mix window can either be early in the morning, before patients arrive, or late in the evening to prepare drugs for the next day.

Lastly in Chapter 4 we use discrete-event simulation to evaluate and compare various make-ahead policies and their broader effects on the infusion center as a whole. Here we show that pre-mixing chemotherapy drugs not only benefit the patient they are prescribed but for subsequent patients as well, due to decreasing the pharmacy work load during peak demand times. We evaluate randomness associated with patient arrivals and the various process steps during the compounding process.

These tools for pre-mixing chemotherapy drugs at the UMRCC can be beneficial in regard to decreasing patient wait times and saving nurse overtime cost. This dissertations suggest that our approach will not only help reduce patient waiting times at the UMRCC and other Cancer Centers, but may stand as the foundation for the development of decision support software to help other pharmacies organize drug preparation. Also, this approach may be relevant in other areas of healthcare where there are interactive stages with unpredictability. An example might be an emergency department starting to prepare an inpatient bed prior to a patient being admitted.

1.5 Key Contributions

In Chapter 2 we develop the first prediction model to determine the probability of a patient deferral (i.e., last minute cancellation) in an outpatient chemotherapy infusion setting. All previous efforts focused on no-show predictions. We

also demonstrate a comparative procedure as well as a unique temporal data cross-validation to compare multiple models and determine which is best for our data set. We discuss these steps which can be applied to any datasets looking to predict patient deferrals.

In Chapter 3 we develop an optimization model focusing on pre-mixing chemotherapy drugs. We discuss two variants of our problem and illustrate how to formulate a deterministic IP model where the decision variables are both interdependent and time dependent.

In Chapter 4 we incorporate our predictions from Chapter 2 in a discrete-event simulation to better represent the effects of the drug make-ahead process while considering the chance of patient deferring treatment. This tool can evaluate any pre-mix/make-ahead chemotherapy drug policy at an outpatient infusion center.

We also show while these methods can stand independently, they work in conjunction to provide more realistic and informative solutions.

Chapter 2

Predicting Patient Deferrals

2.1 Introduction

Patients scheduled for outpatient infusion sometimes may be deferred for treatment (i.e. last minute cancellation) after arriving for their appointment. This can be due to a secondary illness, not meeting required treatment parameters (e.g. white blood cell counts), or other medical complications. In most cases an oncologist or nurse order a deferral when the patient has not recovered sufficiently from the prior chemotherapy. After a clinic visit, the oncologist might suggest a treatment plan change as well which would result in a wasted drug if it was made ahead. This might results from a patient progressing on current treatment and another treatment is recommended but can't be given that day. While unlikely, a patient may self elect to defer treatment due to social support or other personal reasons. This includes arranging a ride home, absence of planned family support, etc.(possibly look up a connection to mental health here. Additionally a patient may or may not have progressed on current treatment and decides not to continue but to receive palliative

care.The ability to generate high-quality predictions of patient deferrals can be highly valuable in managing clinical operations, such as scheduling patients, determining which drugs to make before patients arrive, and establishing the proper staffing for a given day.

At the University of Michigan Rogel Cancer Center (UMRCC), more than 50,000 infusion treatments are performed annually with an increase in patient volume of 5% per year [12]. At every step during a patient's visit (blood draw, clinic, pharmacy, infusion, etc.), there is variability, which can cause disruptions that lead to patient delays and overtime for the infusion nurses.

While these disruptions can originate from many sources, one of the biggest impacts on infusion clinics is when patients no-show or defer their treatment. Unlike other outpatient services, which can have up to a 41% no-show rate [13], patient no-shows are typically not as large of an issue regarding chemotherapy infusion appointments. At UMRCC, for example, only approximately 2% of infusion appointments in 2015 resulted in no-shows; anecdotally, it is thought that patients generally understand the importance of their chemotherapy treatment and will do everything in their power to make sure they arrive for their appointment. On the other hand, although patients show for their appointments, once they arrive they may be deemed too ill for treatment. Additionally, patients may make the decision during their pre-infusion clinic visit to transition to palliative care. Similarly, their oncologist may change their treatment plans based on their response to prior treatments. We refer to these situations as patient treatment deferrals, which are a common source of last minute cancellations. At UMRCC, nearly 47% of patients have at least one

of their infusions deferred over the course of their treatment.

Current prediction models in outpatient care primarily focus on determining patient no-shows to help determine more advantageous over-booking policies [14] [15] [16]. Most authors had an objective to improve and/or implement over booking policies/appointment rescheduling. Alaeddini, A, et. al. developed a hybrid model for this exact purpose using a test case from the Veteran Affairs (VA) hospital in Detroit, MI. They presented first a multinomial logistic regression which fed their Bayesian inference model to predicted no-show, cancellation, and completed appointment probabilities. (It was not made clear if a same-day cancelation was classified as a no-show since in most cases this appointment would not have been filled.) They also made comparisons with models such as decision tree, Bayes update, Bayes net, Neural network, boosting, etc. However, they did not use the same holdout for all compared models. (Some held out 2/3 while other only 1/3). They also did not indicate a repeated holdout was used to check the variance of the presented models performance [17]. Daggy, J., et. al. also worked with the VA hospital system where they specifically focused on predicting a patient no-show model. They had three years of data for model fitting and testing. They fit a binary logistical regression model without performance comparisons to other classification models. They also had the objective to improve scheduling policies and reduce idle time for physicians due to patient no-shows [18]. Huang, Y. and Hanauer, D.A. also implemented a binary logistical regression model to improve overbooking policies at the University of Michigan Hospital. Similar to others, they utilize both scheduling and demographic data to help predict patient no-show probabilities, [19]. Binary

logistical regression models seem to be the most popular in the literature but are not commonly compared to other methods to prove they have the best predictive performance.

Since patient no-shows are not as prominent in outpatient chemotherapy infusion, our binary prediction models will focus on predicting patient deferrals. Fuentes, S., & Frdin, J. E. also showed that patient deferrals make up the majority of patient cancellations at an outpatient infusion center. They conducted a study to determine the size of this cancellation problem at the Department of Oncology at Karolinska University hospital in Stockholm. They analyzed the frequency of late cancellations as well as the reasons for them. Patients were organized into three categories of treatment indications: curative treatment, adjuvant therapy, and palliative therapy. The cause of late cancellations were classified in five categories: hematological toxicity, other toxicity, patient choice, administrative reasons, and deteriorated performance status. Of the 2948 bookings (1460 patients) during the four-week period, 383 (13%) were cancelled late. For palliative, there were 16% late cancellations; for adjuvant, 9%; for curative, 8%. Patient deterioration was the dominant reason for late cancellations of palliative treatment, whereas hematological toxicity was the dominant reason for late cancellations of adjuvant treatment. Patient deterioration is the dominating cause, overall (51% of total cancellation) [11]. In our model we consider these deterioration and hematological toxicity cases as patient deferrals. This can result from not meeting required lab parameters to receive treatment or being deemed too ill by a nurse or oncologist. These reasons are also supported by the expert opinions of the oncologist at UMRCC.

The presented models utilize demographic, treatment, and patient scheduling data to determine the probability that a patient will not show for or defer treatment on a given day. This chapter focuses on the model selection/comparison process. This information can then be utilized when scheduling patients, determining which drugs to make-ahead, and determining the proper staffing for a given day. In the following section, we introduce all prediction models explored including all performance comparisons. We follow similar procedures found in the literature for binary classification model comparison [20] [21].

The remainder of this chapter focuses on an outpatient chemotherapy clinic where we want to predict the patient deferrals (i.e. the patient arrives for their appointment but is unfit to receive treatment) and no shows. Also, we are more interested in these predicted probabilities to help determine which drugs we should make-ahead of time rather than over booking polices for patients in infusion.

2.2 Methods

2.2.1 Data Description

Our patient data was provided by UMRCC. Each patient encounter (both completed and canceled) is stored in the University Michigan Health Systems' (UMHS) version of Epic, an electronic medical record called MiChart. We used all available outpatient chemotherapy infusion visits from January 2, 2015 - December 31, 2015 (N=28919) with 3,522 total unique patients being seen (some of them for multiple treatments). This excludes 1,266 encounters due to missing data from one

patient not providing their zip code and the rest not having a treatment protocol available. We attribute this from our data coming from two different systems and not having a one to one mapping of treatment protocols and patient scheduling information. We did not find any commonalities between patients with missing data regarding their demographics or their scheduling information discussed in this section.

The scheduling data includes the patients completion status (completed, cancel, no-show), age, sex (Male, Female), scheduled appointment length, date appointment made, date appointment completed, and date appointment cancelled. We then were able to append calculated columns for the total number of previous visits, total number of previous cancelations, numbers of days since the last cancelation, and number of days since the last visit. Then based on each patient's Medical Record Number (MRN), we were able to include additional demographics such as marital status, race, ethnicity, zip code, and the last recorded BMI. Table 2.1 displays the summary statistics for all continuous variables.

Covariate	Description	Mean	St. Dev.	Min	Max
Length	Scheduled infusion appointment length in minutes	195.6	133.2	30	780
Age	Patient's age in years	59	13.6	16	95
Total Previous Cancellation	Number of cancellation since the patient's last completed visit	0.8	1.4	0	21
Days since Last Cancellation	Number of days since the patient's last cancelled visit	27.3	59.5	0	504
Total Previous Visits	Number of the patient's previously completed visits	8.4	9.2	0	83
Days since Last Visit	Number of days since the patient's last completed visit	15.3	23.3	0	448
BMI (kg/m^2)	Patient's last recorded BMI	28	6.8	12.7	78.7

Table 2.1: Summary statistics of continuous predictors

Factor	Description	Levels
Status (Response)	Appointment completed or deferred/no-show	2
Sex	Either male or female	2
Race	Patient's Race: White or Caucasian (WC), Black or African American (BA), Asian, American Indian and Alaska Native (AA)/Native Hawaiian and Other Pacific Islander (NO), BA and other, Multi-racial and WC, WC/BA, and Multi-racial	8
Ethnicity	Patients Ethnicity: Non-Hispanic, Hispanic, Patient Refused, Unknown	4
Marital Status	Patient's Marital Status: Single, Married, Legally Separated, Divorced, Widowed, Unknown, Significant other, Other	8
Protocol	The various treatment protocol a patient is prescribed by an oncologist. This consists of the type of chemotherapy drug, solution, and additional treatment regimen notes	51
Region	Region of the U.S. that the patient's permanent address is in.	10

Table 2.2: Summary of factors used in our models

Due to the large number of unique zip codes, we decided to only incorporate the region for our patients into our models (i.e. the first digit of the zip code). We also considered using the county code (i.e. first 3 digits) but again encountered over 67 counties, and this exceeds the number of factor levels some statistical methods (e.g. Random Forest) can handle. In Appendix A.1 we illustrate the region breakdown for our dataset. Table 2.2 provides details for all of the factors considered. We note that while we only considered 51 patient treatment protocols, there are over 300 unique treatment protocols (many are research drugs and/or different variants in dosing for common drugs). In order to reduce this factor to a reasonable amount of categories, we only considered levels with at least 150 patient encounters in 2015 with that protocol name and grouped the rest in an "other" category. We note that with additional clinical guidance to create groupings of treatment protocols, we can further reduce the number of treatment protocol factors considered in our prediction model which might increase the variables' significance.

2.2.2 Tested Models

We present the results of seven different types of binary classification models to predict the probability of a patients appointment status: completed (1) or deferred/no-show (0) from the UMRCC patient data. These models were chosen based on popularity for binary classification problems and general reported performance. Each model will be described in the subsequent sections and compared against the Null model which assigned the empirical probability of deferral to each patient visit. Due to the temporal nature of the data, we had to carefully design

our holdout analysis (i.e. determine which data to train our model with and which to reserve for validating the model). Specifically we cannot train our model on a patient's appointment which occurred later than the appointment we plan to predict. We first sorted the appointments by date. Then we held out the last 1000 entries for final validation (Dec 18, 2015-Dec 31, 2015). In the holdout loop, 90 consecutive days were randomly selected to train. The next consecutive 30 days were used for testing the prediction. All models were tested with the original training sets as well as an oversampled dataset to help compensate for our unbalanced classification problem (i.e. many more completed appointments than treatment deferral/no-shows). We note other data re-sampling methods (e.g synthetic data generation, stratified sampling) exist but can be difficult to implement with such a high number of levels for the categorical variable such as in our problem. Due to the large numbers of categories for protocol names, we were unable to utilize under sampling and a combination of both under and over sampling effectively.

All models were fitted in R version 3.0.2 run through RStudio software Version 0.99.903 using a 2.5 GHz Intel Core i7 processor with 16 GB of RAM. Total computational time for all models in the holdout loop was approximately six hours. (We summarize the models below with more details in the supplemental material.)

2.3 Results and Discussion

2.3.1 Model Comparison

We tested two different groups of variables with all of the models. M1 included protocol ID while M2 did not. Since protocol ID has such a large number of levels, it significantly increases the solve time of our prediction models. Table 2.3 consists of all tested models in our holdout analysis. Our goal is to determined the simplest yet most accurate prediction model in this study.

Models Name	**Variable Selection**	**Model Formula**
Null	None	None
GLM 1	Full	M1
GLM 2	Step	M1
CART 1	Full	M1
CART 2	Var. Importance	M1
CART 3	Full	M2
Bag 1	Full	M1
Bag 2	Var. Importance	M1
Bag 3	Full	M2
RF 1	Full	M1
RF 2	Var. Importance	M1
RF 3	Full	M2
MARS 1	Full	M1
MARS 2	Full	M2
BART 1	Full	M1
BART 2	Full	M2

Table 2.3: All tested models: Full indicates all variables were selected from Table 2.2 and 2.1 in the main text, Step indicates a stepwise selection based on AIC reduction was used in R, Variable importance indicates that variable selection was completed by removing all variables with an overall importance less than or equal to 0, M1 includes all variables while M2 includes all variables except protocol ID

For our model comparison, we used AUC from our receiver operation

characteristic (ROC) curves, Brier score, and F1 score as our out-of-sample prediction error measure. All of these methods are commonly used for predictive modeling in the healthcare setting as seen in [20], [21], and [22]. We note that AUC is generally considered a more accurate measure of prediction than simply computing prediction accuracy [23]. However, it has also been shown that AUC can be less precise compared to actual population metrics when dealing with large samples or unbalanced samples [24]. Therefore we consider the Brier score since it directly measure the accuracy of the predicted probabilities. We also measure the F1 score since it considers both the precision (i.e. positive predictive value) and recall (i.e. sensitivity)

After completing our 50-fold repeated holdout, we noticed oversampling caused extremely poor performance in our full Bagged CART and oversample Neural Net models (which performed worse than the Null model with all three error measures). Similarly the model performance for oversampling with GLM 2, GLM 3, Bag 2, Bag 3, RF 1,RF 2, Bart 1, and Bart 2 were very close to the Null model. We later saw the standard sample RF 1 and RF 2 performed poorly as both Brier and F1 scores. These were excluded from the rest of our analysis. We then present the out of sample predictive AUC, Brier score, and F1 score for the rest of the considered models in Figure 2.3.1. Through observation it was clear CART 2 and 3 had very high variance as well as poor prediction performance. We then observe BART, MARS, RF, and Bag 3 models seem to outperform all other tested models. In ordered to determine if their higher performance is significant, we conducted a pairwise

t-test with a null hypothesis that the AUC of the two models are significantly different. Since multiple hypothesis test are being conducted, it is more appropriate to use Bonferroni correction to help avoid false claims of a model being significantly different.

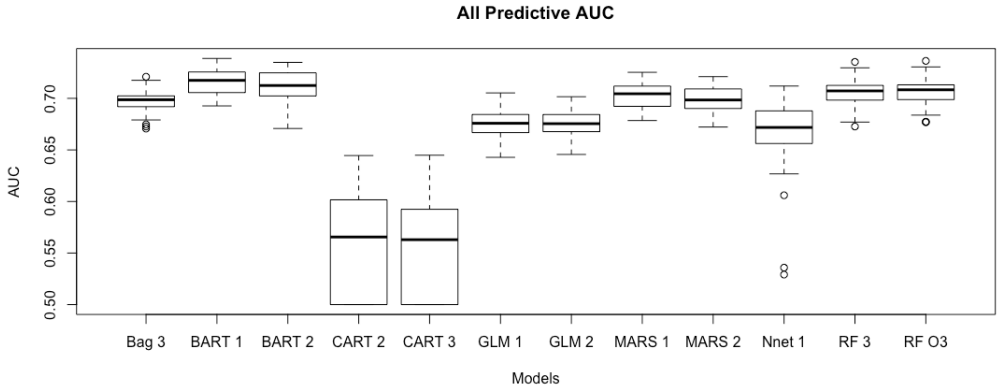


Figure 2.1: Box plot representation of predictive AUC from 50-fold repeated hold out trials

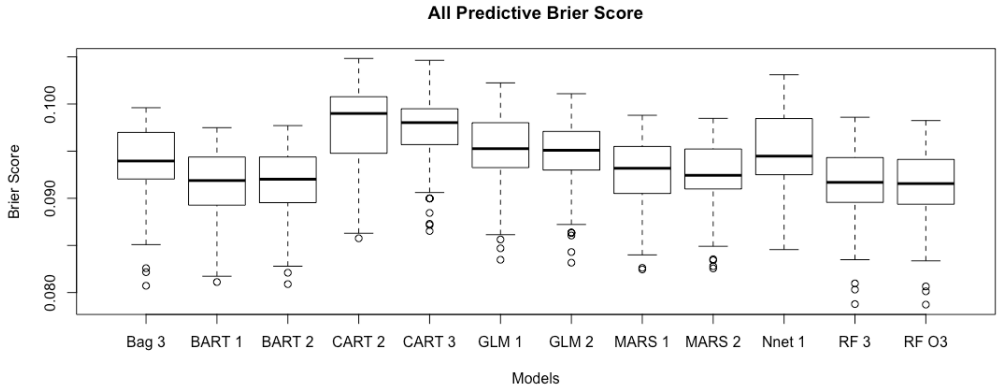


Figure 2.2: Box plot representation of predictive Brier score from 50-fold repeated hold out trials

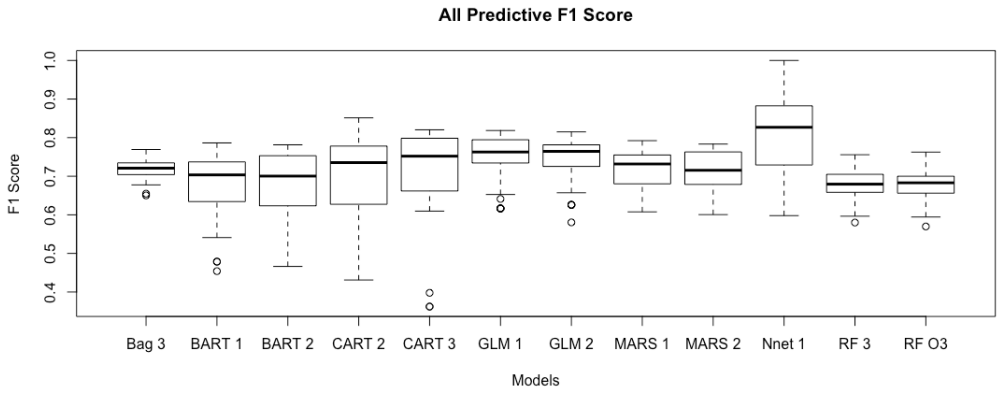


Figure 2.3: Box plot representation of predictive F1 score from 50-fold repeated hold out trials

	GLM 1	GLM 2	CART 2	CART 3	Bag 3	RF 3	RF O3	MARS 1	MARS 2	Nnet 1	Bart 1
GLM 2	1	-	-	-	-	-	-	-	-	-	-
CART 2	0.07634	0.03886	-	-	-	-	-	-	-	-	-
Cart 3	0.14185	0.07446	1	-	-	-	-	-	-	-	-
Bag 3	1	1	0.00149	0.00326	-	-	-	-	-	-	-
RF 3	0.03543	0.06990	2.40E-09	7.63E-09	0.89933	-	-	-	-	-	-
RF O3	0.02141	0.04321	1.00E-09	3.25E-09	0.61289	1	-	-	-	-	-
MARS 1	1	1	2.46E-06	6.58E-06	1	1	1	-	-	-	-
MARS 2	0.92440	1	1.06E-06	2.90E-06	1	1	1	1	-	-	-
Nnet 1	1	1	0.29674	0.51577	1	0.00725	0.00417	0.43645	0.27391	-	-
Bart 1	0.03346	0.06618	2.17E-09	6.92E-09	0.86112	1	1	1	1	0.00681	-
Bart 2	0.06792	0.12996	7.58E-09	2.35E-08	1	1	1	1	1	0.01485	1

Table 2.4: Pairwise t-test utilizing a Bonferroni correction term comparing out of sample predictive Brier Scores

Our Bonferroni test confirms both the GLM and CART models are outperformed by all other models. We then see the Nnet model is also outperformed by our top models (BART 1, MARS 1, RF 3, and RF O3) based on an α level of .05 with all three error measures (Note: RF O3 is the RF 3 model with oversampling). We also notice there is no significant difference between our top models and our Bag 3 model with either error measure. However, Bag 3 performed slightly worse in the

	GLM 1	GLM 2	CART 2	CART 3	Bag 3	RF 3	RF O3	MARS 1	MARS 2	Nnet 1	Bart 1
GLM 2	1	-	-	-	-	-	-	-	-	-	-
CART 2	0.5786	1	-	-	-	-	-	-	-	-	-
Cart 3	0.3183	0.81705	1	-	-	-	-	-	-	-	-
Bag 3	1	1	1	1	-	-	-	-	-	-	-
RF 3	0.0000	0.00015	1	1	0.24380	-	-	-	-	-	-
RF O3	0.0000	9.74E-05	1	1	0.18498	1	-	-	-	-	-
MARS 1	1	1	1	1	1	0.20795	0.15718	-	-	-	-
MARS 2	0.4999	1	1	1	1	0.90635	0.71124	1	-	-	-
Nnet 1	0.0005	9.34E-05	6.75E-09	3.84E-09	5.14E-09	1.13E-17	5.79E+18	6.90E-09	3.49E-10	-	-
Bart 1	0.0000	0.00018	1	1	0.28228	1	1	0.24127	1	1.62E-17	-
Bart 2	0.0002	0.00102	1	1	0.84991	1	1	0.73844	1	2.73E-16	1

Table 2.5: Pairwise t-test utilizing a Bonferroni correction term comparing out of sample predictive AUC

	GLM 1	GLM 2	CART 2	CART 3	Bag 3	RF 3	RF O3	MARS 1	MARS 2	Nnet 1	Bart 1
GLM 2	1	-	-	-	-	-	-	-	-	-	-
CART 2	0.5786	1	-	-	-	-	-	-	-	-	-
Cart 3	0.3183	0.81705	1	-	-	-	-	-	-	-	-
Bag 3	1	1	1	1	-	-	-	-	-	-	-
RF 3	0.0000	0.00015	1	1	0.24380	-	-	-	-	-	-
RF O3	0.0000	9.74E-05	1	1	0.18498	1	-	-	-	-	-
MARS 1	1	1	1	1	1	0.20795	0.15718	-	-	-	-
MARS 2	0.4999	1	1	1	1	0.90635	0.71124	1	-	-	-
Nnet 1	0.0005	9.34E-05	6.75E-09	3.84E-09	5.14E-09	1.13E-17	5.79E+18	6.90E-09	3.49E-10	-	-
Bart 1	0.0000	0.00018	1	1	0.28228	1	1	0.24127	1	1.62E-17	-
Bart 2	0.0002	0.00102	1	1	0.84991	1	1	0.73844	1	2.73E-16	1

Table 2.6: Pairwise t-test utilizing a Bonferroni correction term comparing out of sample predictive F1 scores

Brier score measure than the rest of our top models. We then tested these models on our validation set to see which was better suited for prediction at UMRCC. These prediction models output the probability a patient will defer or not show for treatment on a given day. One application we have applied this to is in the pharmacy at the UMRCC. If pharmacist are working to make ahead certain drugs, information on patient deferrals ahead of time is crucial. When testing our models on the validation set, we tested various decision thresholds (i.e. anyone below the threshold would be assigned a value of 0 for treatment deferral/no show and anyone above assigned a value of 1 for completed appointment).

2.3.2 Selected Model Discussion

Model	Brier score	AUC	F1 score
MARS 1	.11	.67	.88
BART 1	.1	.71	.78
RF 3	.1	.69	.77
RF O3	.11	.66	.72

Table 2.7: Final Model Performance Measures

Considering the results from our final comparisons seen in Table 2.7, BART 1 was chosen as our final model based on slightly outperforming MARS 1 in both Brier score and AUC as well as RF 3 and RF O3 in AUC and F1 score. Our prediction model outputs a probability that a patient will complete their appointment. We then can set a threshold to help us determine what the predicted response will be for a given patient visit. Any predicted probability above the threshold will be classified as a completed appointment and anything below as a deferred/no-show appointment. For example, if we set a threshold at probability .75, we were able to correctly predict 21% of deferrals/no shows and 93% of completed appointments with an overall prediction accuracy of 84%. If we move the threshold from probability .75 to .8 we correctly predicted 34% of deferrals/no shows and 79% of completed appointments with an overall prediction accuracy of 82%. Depending on the applications, the decision maker may favor predicting the positive outcome more accurately than the negative, which in our case would favor the original threshold. Alternatively if predicting the negative outcome accurately is more crucial than our increased threshold is favored. Even if a decision maker is more favorable to a negative or positive response accuracy,

use of optimization techniques to determine the optimal decision threshold has been shown beneficial as done in [25]. In Figure 2.4 illustrates all tested thresholds.

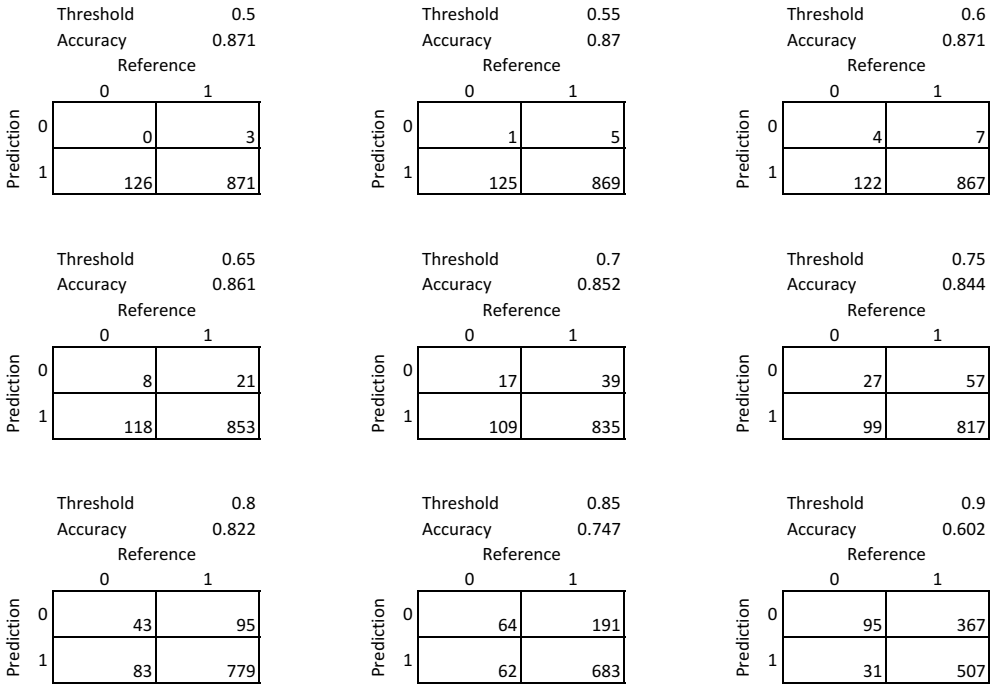


Figure 2.4: Confusion matrix for various thresholds for the top performing BART model

2.4 Variable Significance

For purposes of interpretability, we first utilized the logistic regression model to determine the significance of each variable in predicting patient deferral/no-show status. Appointment Length, Previous number of visits, and Protocol ID have the strongest positive influence on the response while previous number of cancelation and number of days since last visit have the most significant negative effect on the

response.

Next we tested the marginal effect of each covariate using the BART model. Figure 2.5 consist of the partial dependence plots for all covariates and just a few examples for our factors. Each x-axis represent the values taken by the variable being tested and the y-axis represents either the probits or logits for each response. (The higher the probit value the higher the probability of having a completed appointment. This also holds for the logit scale in (i)) The blue lines illustrate the 95% confidence interval for each dependence plot.

Based on the partial dependence plots, we see the number of previous cancellations (c) and the number of days since the last visit (f) have the strongest negative influence on the response (i.e. the higher the values the lower the probability of completion). We also see that new patients and patients with at least 15 previous visits have a higher probability of completing their appointment (e). We also give examples of significant levels based on the logistic model for marital status and race in (h) and (i). Lastly we note there does not seem to be a strong influence with age (b) but there is a strong positive marginal effect with the days since last cancelation on the probability of completing an appointment (d). Both appointment length (a) and last recorded BMI (g) have a slight positive effect on appointment completion at higher values.

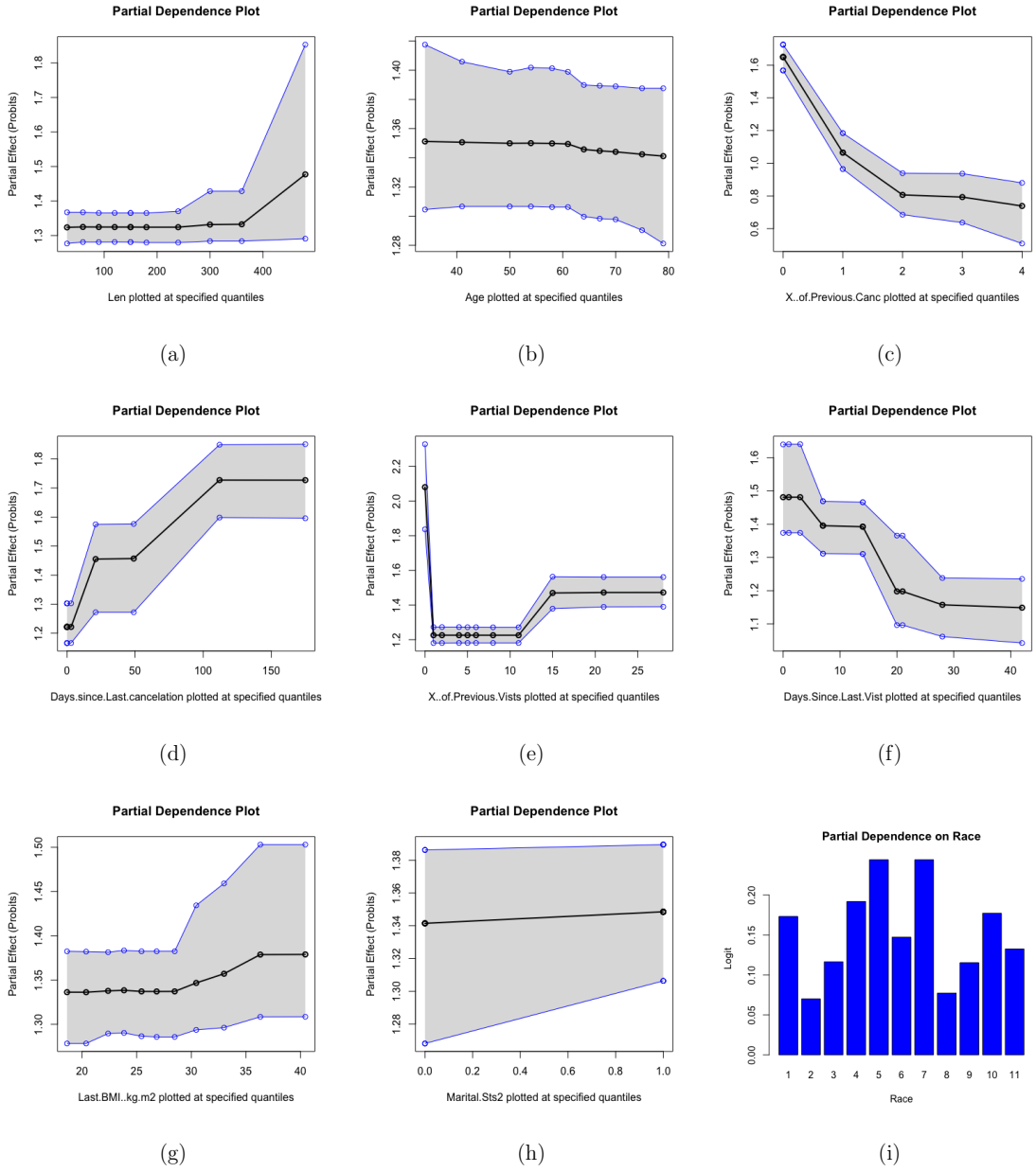


Figure 2.5: Partial dependence plots for all covariates and some example levels of the most influential factors used in our BART model

2.5 Conclusion

This chapter highlights the potential benefits of using statistical models for predicting patient deferrals in an outpatient infusion center. It not only presents a variety of statistical prediction models, but also provides tools for comparing and then selecting the most appropriate of these models for a specific clinical environment. We found the BART model to have the best predictive performance, based on Brier score, AUC, and F1 score, for the UMRCC outpatient infusion center. However, this may not be true in other hospital settings. Our results suggest that, at least for our data set, ensemble methods capable of handling non-linear data with interaction between variables (e.g., BART and RF) provide stronger predictive accuracy than linear models such a logistic regression. We hypothesize that this would generalize to data from other cancer treatment centers. There are also trade-offs between predictive accuracy and interpretability with some models being easier draw inferences about variable importance and influence from than others but at the expense of predictive accuracy. Because a cancer center is most likely to be focused on predictive accuracy in practice, that has been our focus in this paper. However, each cancer center must strike their own balance between interpretability and predictive accuracy. Similar comparison procedures should be followed to determine the best prediction model for each setting. We recognize with guidance from clinical collaborators our mo With proper implementation, these models can enable clinicians and clinical managers to achieve the in-practice benefits of deferral and no-show predictions.

Chapter 3

Optimal Schedule for Pre-mixing Drugs

3.1 Introduction

In this chapter we propose to use Integer programming (IP) techniques to determine the optimal set of drugs to make-ahead at an outpatient chemotherapy infusion center given a finite window of time. IP models are widely used to address scheduling and care delivery problems in healthcare such as scheduling elective surgery patients, clinical providers, or pharmacy orders [26], [27], [28]. There are also more specific research efforts focusing on scheduling of chemotherapy infusion patients as discussed in [29], [30], and [31]. These methods ranged from complex stochastic optimization models to more tractable and interpretable IP models in excel. However, our specific problem domain regarding make-ahead drugs in a pharmacy, specifically chemotherapy, is much more scarce in the literature. Currently this problem has been addressed with queuing/simulation techniques and "Lean" methods but not optimization [10] [32]. Another case study also The models in the literature also disregard the predictability of patient deferrals and in turn miss the

opportunity to utilize the probability of wasting a drug when determining which drugs would be most advantageous to make ahead and when we should make them. More broadly, pharmacy preparation improvements have been studied with batched pharmacy orders [33] [34]. However, due to the highly personalized nature of chemotherapy treatment, these methods are not generalizable to the outpatient chemotherapy infusion center pharmacy domain.

In this chapter we focus on developing two tractable IP models that determine the optimal number drugs to pre-mix. Given a finite amount of time for pre-mix, we considered how to best use that time by maximizing the difference between the value of expected reduced wait time and expected waste cost. This is done by determining a monetary value for patient wait time as discussed in Section 3.2. We utilize patients' probability of treatment deferral determined in Chapter 2 to define our expected waste cost and use the complement (i.e., the probability of completing an appointment) to determine the expected reduced wait time. The first model discussed in Section 3.3 assumes that all drugs being considered will last the entire day since most drugs being considered for pre-mix have a shelf life/hang-by time of at least 8 hours. Our second model discussed in Section 3.4, relaxes this assumption to incorporate all drugs with any hang-by time. So depending on the specific time a drug is made may not lend it viable to all the patients scheduled to receive it that day. We note if multiple patients are scheduled to receive the same pre-mixed dose and drug, then the drug labeling can be updated and given to any of the patients.

Both models consider the schedule of patients on a given day along with their respective probability of treatment deferral. The model then balances the

trade-off between potential reduced wait time and generated waste cost resulting from pre-mixing a chemotherapy drug. In Section 3.2, we discuss how we directly compared patient time with the potential waste cost associated with pre-mixing a drug. After evaluating this trade-off, this model generates a schedule for pre-mixing chemotherapy drugs to be utilized by a pharmacist.

3.2 Reduced Wait Time

One major challenge is how to define the expected value of reduced wait time gained by pre-mixing a particular drug. Setting a dollar value for the reduced wait time gained by pre-mixing a drug is a very controversial topic since we are trading off the value of a patient's time with a potential cost the hospital could incur. Some studies define wait time cost as the average amount a person could have earned working as their wait time cost [35]. Others define wait time cost by the hedonic value of life as a set standard for patient wait time [36]. In our model we assume a patient's wait time is worth the average wage in the Detroit metro area, as reported by the U.S. Bureau of Labor and Statistics, they could have made if working, which in 2017 was $25.22 [37]. Another aspect to consider when defining the value of reduced wait time gained by pre-mixing a drug is the extent to which the opportunity cost is the same for all infusion patients. Should sicker patients or patients with more severe cancers have a higher cost associated with their wait time? Should the time a patients spends away from their loved ones or home be considered? A range of medical, interpersonal, and contextual factors exists in patients' lives that are difficult to quantify, but one must consider their role in defining the value of reduced

wait time when pre-mixing a drug. We anticipate defining the cost for wait time to continue to be a controversial subject so we plan to run a sensitivity analysis on this wait time cost and leave it as an input for the decision maker. We also emphasize the dollar value associated with reduced wait time is a constant value input to the optimization models. This value can easily be scaled nonlinearly if for example the decision maker found a subgroup of patients' time was worth more then the general population or if the decision maker considered 1 hour of reduced wait worth 2 times the cost of 15 min of reduced wait time.

Additionally, our model takes a conservative approach when calculated the expected wait time saved by pre-mixing a dose of a drug. While we only consider the time it would have taken to mix a drug as the amount of reduced observed patient waiting time, these time saving can also reduce the overall load on the pharmacy system as a whole. (i.e., benefitting patients whose drugs were not pre-mixed). This is discussed more when evaluating our simulation model in Chapter 4.

3.3 Chemotherapy Pre-mix Integer Program (CPIP)

As illustrated in Figure 1.2 from Chapter 1, the drug mixing process is complex. However most steps take a trivial amount a time with the majority of the drug processing time coming from the order verification steps taken before the drug is made and the actual compounding time. We simplify our optimization model to focus on these two steps which must still occur in sequence. For each drug, we input the average verification time and compounding process time as input parameters. While there is some variability here depending on the technician mixing the drug,

it is not significant enough to implement a more complex stochastic model. Our objective is to maximize the difference between the expected dollar value of reduced wait time defined in Section 3.2 and the expected waste cost for mixing a drug given a fixed window of time for pre-mix in the morning. We assume the following:

- The drug mixing times are deterministic
- We can predict the patient probability of deferral from a BART model as seen in Chapter 2
- Pharmacy tasks can be reduced to two key steps (1) Drug Verification (2) Drug Compounding (i.e., mixing)
- Drugs have a hang-by time lasting the entire day regardless of when the drug was pre-mixed that morning

3.3.1 CPIP Formulation

We now present the IP model formulation for the CPIP model. We note the use of time discretization in this model to allow us to use a simpler discrete model rather than continuous time model. In the computational experiments, we further explore at what granularity we should discretize time to balance solution quality with model performance. The more granular the discretization, the more exact the solution. Conversely, increased granularity also increases our problem size thus leading to much longer solve times.

Sets

- D: set of drugs d (e.g. 50 mg of Taxotere) Note: Each drug d represents a set of patients scheduled to receive it.

- S: set of stages s (stage 1 verification; stage 2 drug compounding)

Parameters

- p_d^s: the time duration for pre-mixing drug d in stage s $\forall d \in D, s \in S$

- Δ_d: the product of the reward/value of reduced wait time for pre-mixing drug $d \in D$ and the value per unit of reduced wait time (i.e., $(p_d^1 + p_d^2)$*patient time value)

- T: the total time units for the pre-mix period

- $\bar{T}_s^d = T - p_d^s + 1$: Most constraints only need to consider the last time slot a drug can start being pre-made. We use $\bar{T}$ to help simplify this notation for all $\forall d \in D, s \in S$.

- c_d: the cost of drug d $\forall d \in D$

- m_d: the number of doses needed for drug d (i.e., the number of patients scheduled to receive drug d) $\forall d \in D$

- C_s: total number of drugs you can mix at a time in stage s $\forall s \in S$

- $P_d(n)$: the probability of wasting the n^{th} dose of drug d, formally defined in Section 3.3.2 $\forall d \in D, n = 1...m_d$

- M: very large number

- $E_n^d = \Delta_d(1 - P_d(n))$: The expected value of reduced saving when you pre-mix dose n of drug d $\forall d \in D, n = 1...m_d$

- $E_n^d = c_d P_d(n)$: The expected waste cost when you pre-mix dose n of drug d $\forall d \in D, n = 1...m_d$

Variables

- $x_{nts}^d = \begin{cases} 1 & \text{if we start pre-mixing dose } n \in \{1...m_d\} \text{ of drug } d \in D \text{ at time } t \in \{1...T\} \\ & \text{for stage } s \in \{1,2\} \\ 0 & \text{o.w.} \end{cases}$

- $y_n^d = \begin{cases} 1 & \text{if we don't pre-mix dose } n \in \{1...m_d\} \text{ of drug } d \in D \\ 0 & \text{o.w.} \end{cases}$

- $z_{nts}^d = \begin{cases} 1 & \text{if we are pre-mixing dose } n \in \{1...m_d\} \text{ of drug } d \in D \text{ at time } t \in \{1...T\} \\ & \text{for stage } s \in \{1,2\} \\ 0 & \text{o.w.} \end{cases}$

Our objective is to

$$\text{maximize} \sum_{d \in D} \sum_{n=1}^{m_d} \sum_{t=1}^{\bar{T}_1^d} (E_n^d[Reduced\ Wait] - E_n^d[Waste\ Cost]) * x_{nt1}^d \tag{3.1}$$

Note we only need to consider a single stage when evaluating the objective since we constrain our model to processing a drug through both phases if it is pre-mixed.

Constraints

1. If you mix a dose of a drug, it must be done in both phases and can only be started in each phase once for that dose n. If a particular dose n of drug d is not pre-mixed then $y = 1$ forcing all associated x's to 0.

   $\sum_{t=1}^{\bar{T}_s^d} x_{nts}^d + y_n^d = 1 \quad \forall d \in D, n = 1...m_d, s \in S$

2. If a dose of a drug is made then it must be processed a total of p_d^s time units in each stage s. If it is not made then $y = 1$ forcing all associated x's to 0.

   $\sum_{t=1}^{T} z_{nts}^d + p_d^s * y_n^d = p_d^s \quad \forall d \in D, n = 1...m_d, s \in S$

3. We must make the n^{th} dose of drug d before the $(n+1)^{th}$ dose of drug d. If the $(n+1)^{th}$ dose is not made, then the right hand side of our constraint takes on the value of big M resulting in no restrictions on when we pre-mix the previous dose. If a subsequent dose is made, $y_{n'}^d = 0$ forcing the start of the subsequent dose to be no earlier then the start of the previous dose.

   $\sum_{t=1}^{\bar{T}_s^d} t x_{nt1}^d \leq \sum_{t'}^{\bar{T}_s^d} t' x_{n't1}^d + M * y_{n'}^d \quad \forall d \in D, n = 1...(m_d - 1)$ where $n' = n + 1$

4. If you don't make the n^{th} dose you can't make any subsequent doses (note: $n' = n + 1$).

   $y_n^d \leq y_{n'}^d \quad \forall d \in D, n = 1...(m_d - 1)$ where $n' = n + 1$

5. We can only make C drugs at a time at each stage s. Stage 1 capacity is driven by the number of verification pharmacists on hand while stage 2 is limited by the number of pharmacy technicians and/or drug mixing hoods on hand.

$$\sum_d \sum_n z^d_{nts} \leq C_s \quad \forall t = 1...T, s \in S$$

6. Preemptions are not allowed once a drug starts the mixing process. We define the index i to count from 0 to one less than the processing time for the given drug d at stage s. This constrains our z variables to consecutively be 1 until a drug is completed given we decide to pre-mix the drug starting at time t.

$$x^d_{nts} \leq z^d_{n(t+j)s} \quad \forall d \in D, n = 1...m_d, s \in S, t = 1...\bar{T^d_s}, j = 0...(p^s_d - 1)$$

7. This constraint ensures each drug must complete Stage 1 before Stage 2. Since we define our time unit at the beginning of the interval, the sum of when we start a drug in Stage 1 plus the processing time for that drug in Stage 1 must be less than or equal to when we start the drug in Stage 2.

$$\sum_{t=1}^{\bar{T^d_s}} (t + p^1_d) x^d_{nt1} \leq \sum_t t x^d_{nt2} \quad \forall d \in D, n = 1...m_d$$

3.3.2 Probability of Wasting a Drug: Homogeneous Patient Group

Utilizing the probability of wasting a drug has been shown to benefit pharmacy operations in [38]. Now that we have formally defined our IP we must discuss how the probability of wasting the n^{th} dose of drug d, $P_d(n)$, is formulated. We note these probabilities are assumed to be independent between drugs and patient probability of deferrals are also assumed independent from each other. Therefore our formulation discussed in this section applies to each drug individually.

Wasting the n^{th} dose

We begin by assuming there is only one drug to be made, with m patients scheduled to receive this drug. The first dose of the drug that is prepared will be administered to the first patient of the day who does not defer, the second prepared dose to the second patient who does not defer, and so forth. This is based on the notion if a patient defers, then their drug can be used by any subsequent patient since it is the same drug ordered for both patients.

Now suppose each patient has a probability p of being deferred or some other reason not receiving their scheduled treatment. Given that the first dose made is the first dose used, this dose will only be wasted if *all* patients defer. Therefore our probability of wasting the first dose is given by

$$P(1) = (p)^m \quad \text{which can also be written as} \tag{3.2}$$

$$P(1) = \binom{m}{m-0} p^m (1-p)^0 \tag{3.3}$$

i.e., all patient must defer.

Now consider the case where we waste the second dose made. This can happen either if we waste both the first and second dose (which happens if all patients defer) or we use the first dose but waste the second dose (which happens if exactly one

patient receives treatment and all others defer). This is given by the following:

$$p(2) = (p)^m + \binom{m}{m-1} p^{m-1}(1-p) \tag{3.4}$$

Using the same logic we can compute the probability that the n^{th} prepared dose is wasted as

$$P(n) = \sum_{i=1}^{n} \binom{m}{m-(i-1)} p^{m-(i-1)}(1-p)^{i-1}. \tag{3.5}$$

3.3.3 Probability of Wasting a Drug: Heterogenous Patient Group

Now we relax the assumption that all patients have the same probability of deferral. We define the set S to contain the integers from 1 to m representing all patients scheduled to receive the same drug on a given day. For example, given m total patients scheduled for the day, $S = 1, 2, 3, ..., m$.

Similar to the homogeneous case, the probability of wasting the first dose given m total patients is

$$P(1) = \prod_{i \in S} p_i, \tag{3.6}$$

the product of the probabilities of each patient deferring

The probability of wasting the second dose is given by

$$P(2) = \prod_{i \in S} p_i + \sum_{i \in S} \left[(1 - p_i) \prod_{j \in S \setminus i} (p_j) \right] \tag{3.7}$$

We waste the second dose if either everyone defers (i.e. the probability of wasting the first dose) or if exactly one patient is well enough for treatment but the rest defer resulting in only the first pre-made dose being utilized.

The probability of wasting the third dose is given by

$$P(3) = \prod_{i \in S} p_i + \sum_{i \in S} \left[(1 - p_i) \prod_{j \in S \setminus i} (p_j) \right] + \sum_{i} \sum_{j \in S \setminus i} \left[(1 - p_i)(1 - p_j) \prod_{k \in S \setminus \{i,j\}} p_k \right] \tag{3.8}$$

We waste the third dose if all patients defer,(i.e., the probability of wasting the first dose) or only one patient is well enough for treatment (i.e., the probability of wasting the second dose), or exactly two patients are well enough and the rest defer treatment (resulting in the first two doses being utilized but the third being wasted.)

Summarizing, we see that:

A. Probability of wasting the 1^{st} dose — all patients must defer treatment

B. Probability of wasting the 2^{nd} dose — all patient must defer or only one is well enough for treatment (i.e., Case A + Prob(one patient is well enough))

C. Probability of wasting the 3^{rd} dose — all patient must defer, only one patient is well enough for treatment, or only 2 patients are well enough for treatment

(i.e., Case B + Prob(two patients are well enough))

Extending this logic, we observe, recursively, that the probability of wasting the n^{th} dose = Prob(wasting $(n-1)^{th}$ dose) + Prob($n-1$ patients are well enough for treatment).

To generalize this, suppose we have m patients scheduled to receive the same drug on a given day. If we waste dose n then at most $n-1$ people scheduled to receive that drug can be well enough to receive treatment that day in order for the n^{th} dose to be wasted. Conversely, $m-(n-1) = m-n+1$ patients must defer to waste the n^{th} dose of a drug. Considering this, we define s_i^n as the i^{th} subset of S which contains $n-1$ patients who complete their appointment with the remaining $m-(n-1)$ patients deferring treatment that day. Since we want to consider the combinations of patients that need to defer out of the total m patients, we recognize that $i = 1...\binom{m}{n-1}$. (Ex. If we were considering the probability of wasting the third dose with four total patients scheduled to receive that drug then $|s_i^3| = 3-1 = 2\ \forall i$. $s_1^3 = \{1,2\}, s_2^3 = \{1,3\}, ...$). We note counting the negative cases of deferral is equivalent to counting the combinations of the positive cases (i.e., $\binom{m}{m-n+1} = \binom{m}{n-1}$).

Now we can generalize our probability for wasting dose n of our drug with the

following recursive function:

$$P(1) = \prod_{j \in S} p_j \tag{3.9}$$

$$P(n) = P(n-1) + \sum_{i=1}^{\binom{m}{m-n+1}} \Big[\prod_{j \in s_i^n} p_j \Big] \Big[\prod_{k \in S \setminus s_i^n} (1 - p_k) \Big] \tag{3.10}$$

3.4 CPIP-Hang-by Time (CPIP-HT)

In CPIP from Section 3.3, we made the simplifying assumption that all drugs lasted the entire day (i.e. had a hang-by time of at least 8-12 hours). Therefore no matter what time they were pre-mixed in the morning, they would be viable for all patients scheduled to receive that drug on a given day. In reality, some drugs have shorter hang-by times (ranging from one to six hours). Therefore you must consider when patients are scheduled for treatment when deciding which drug you decide to pre-mix. Specifically, the probability of a particular dose of a drug being used (and, conversely, of being wasted) depends on both the time that the dose is mixed and the time of the patient appointments. Even more challenging, as we see in the following example, it also depends on the time that other doses of the same drug are mixed. Thus, we expand the original CPIP model to incorporate this additional problem characteristic in CPIP-HT.

Example: Suppose we have two patients scheduled to receive the same drug, with probabilities of deferral p_1 and p_2, respectively. In Figure 3.1a we illustrate a timeline representing both the window of time to pre-mix and the scheduled time for each patient (indicated with their probability of deferral). Figure 3.1b illustrates

how pre-mixing a dose of in the beginning of our window results in that dose being viable only for the first patient. This results in the probability of wasting the first dose of the drug as p_1. However in Figure 3.1c we show if the dose is made at the end of the period it is viable for both patients. The probability of wasting the first dose of the drug now becomes $p_1 * p_2$. This shows that the probability of wasting the dose depends on the time it was made.

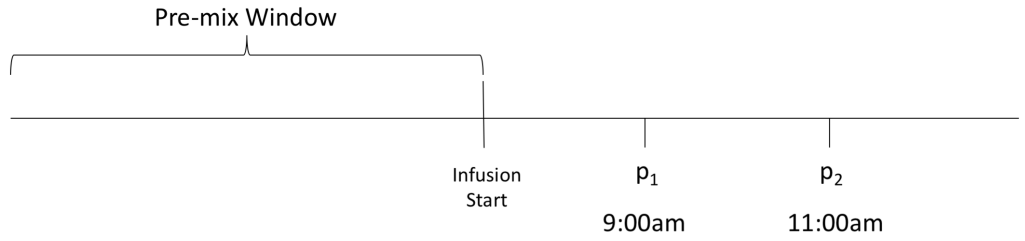
Now suppose we pre-mix two doses of this drug. In Figure 3.1d we consider the case where both doses are viable for all patients hence the probability of wasting the first and second dose follows the same formulation from Section 3.3.2 as

$$P(1) = p_1 * p_2 \tag{3.11}$$

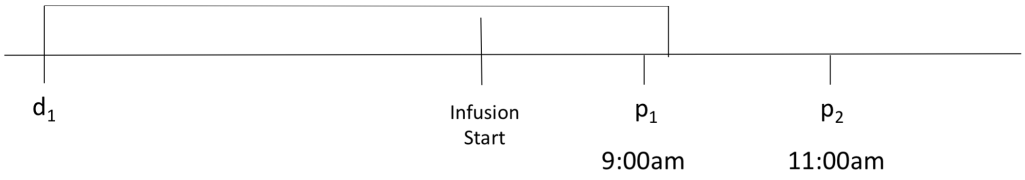
$$P(2) = P(1) + (1 - p_1) * p_2 + (1 - p_2)p_1. \tag{3.12}$$

However if we make the first dose earlier in our pre-mix window, as we see in Figure 3.1e, the first dose is now only viable for the first patient while the second is viable for both. Therefore, the probability of wasting the first dose is simply p_1, the probability of wasting the second dose is given by

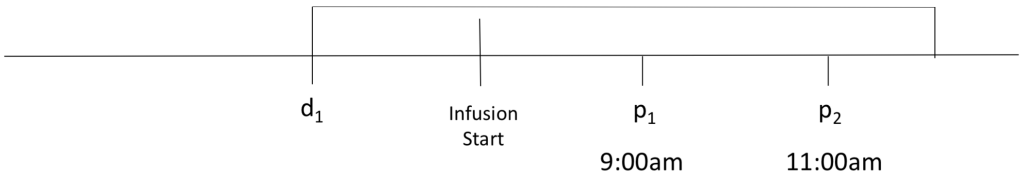
$$P(2) = p_1 * p_2 + (1 - p_1)p_2 = p_2. \tag{3.13}$$


Pre-mix Window
Infusion Start
p_1
p_2
9:00am
11:00am

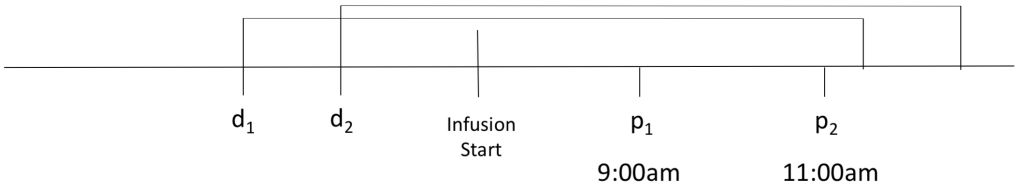
(a)


d_1
Infusion Start
p_1
p_2
9:00am
11:00am

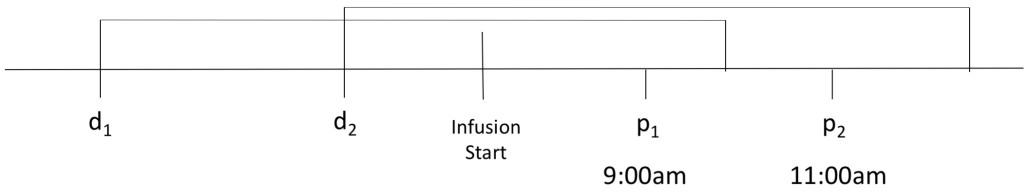
(b)


d_1
Infusion Start
p_1
p_2
9:00am
11:00am

(c)


d_1
d_2
Infusion Start
p_1
p_2
9:00am
11:00am

(d)



(e)

Figure 3.1: Example of eligibility windows: (a) Example of a timeline to represent pre-mix window and two patients scheduled for the same drug (b) Here we show a single dose of a drug pre-mixed and being eligible for a single patient (c) Here is the same drug from the previous example but made at a later time to represent it being eligible for both patients (d) Now we introduce two drugs both being eligible for both patients (e) Illustrates two doses of the same drug being pre-mixed again but the first being eligible for one patient and the second for both

This shows that the probability of wasting the second dose depends not only on the time it was made, but also on the time that the first dose was made

3.4.1 Patient Eligibility Vectors

To determine the potential waste cost (and, in turn, the objective value) associated with a particular dose of a given drug, we need to know when *all* the doses of that drug were made. To accommodate this, first consider replacing the simple variable x^d_{nts}, which determines whether the n^{th} dose of drug d was made at time t in stage s, with a complex variable $y^{ds}_{t_1,t_2...t_m}$ which determines whether you made a particular set of doses for a given drug d at respective times $t_1, t_2, ...t_m$ in stage s. (Note some t_i's are null if we do not pre-mix dose i.) This variable allows us to fully capture the probability of deferral for each dose, and therefore the expected cost of the associated drug. Unfortunately, this comes at the cost of an exponentially large number of variables.

Consider, however, the following example. Suppose we have three patients all scheduled to receive the same drug. Depending on when you pre-mix a dose you then could have from zero (i.e., you decided not to make the drug) to three total patients eligible to receive the drug based on when they are scheduled for treatment and when the drug will surpass its hang-by time and expire.

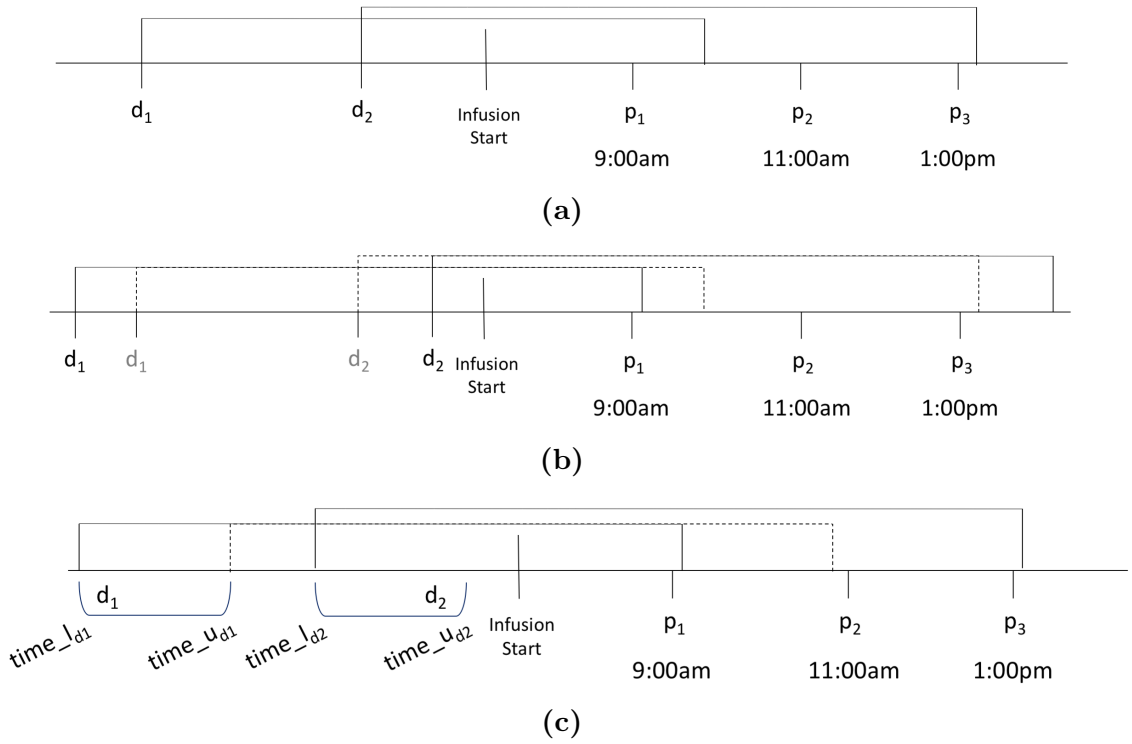


Figure 3.2: Example of eligibility windows: (a) Given 3 patients while pre-mixing two drugs (b) Pre-mixing the same two drugs but a different times while keeping the same outcome (c) The window of time where we can pre-mix these drugs to get the same outcome

In Figure 3.2a we consider the case where we make 2 doses at times such that:

- for d_1 - 1 patient is eligible

- d_2 - 3 patients are eligible
- d_3 - 0 patients are eligible (not mixed)

Then in Figure 3.2b we consider making dose d_1 earlier in our pre-mix window and and d_2 later but–observe that we still get the same outcome for patient eligibility! Therefore in both cases, the expected waste cost will be the same. Further, as we see in Figure 3.2c, *any* set of doses pre-mixed in the indicated windows of time will yield the same probabilities of waste. We denote the bounds on these time windows as $time\_u_n$ and $time\_l_n$ for the upper and lower bound on when we can pre-mix the n^{th} dose of a drug which is further discussed in Section 3.4.4. No matter when doses d_1 and d_2 were pre-mixed in this window, they would have the same associated objective value for all combinations of our decision variable $y^{ds}_{t_1 t_2 \ldots t_m}$. This examples illustrates that we focus solely be on the *number* of patients eligible to receive a drug (i.e. within the hang-by time window for each drug) rather than the time each dose was made. We exploit this fact to reduce the size of CPIP-HT. Specifically, we define a *patient eligibility vector* e^d_i as the i^{th} collection of how many patients are eligible for each dose of drug d along with the associated pre-mix time window $[time\_l_n, time\_u_n]$ required for the vector to be feasible. Based on this definition, we now present the CPIP-HT formulation.

3.4.2 CPIP-HT Formulation

Now that we have defined our patient eligibility vectors and the associated pre-mix time window imposed by them, we present CPIP-HT. Our model follows the same introduction as CPIP from Section 3.3 but relaxes that assumption all

drugs have a hang-by time lasting the entire day. To determine our patient eligibility vectors, we also assume patients arrive on time for their scheduled infusion appointment (Note: the decision maker can incorporate a patient arrival tolerance for a more conservative approach that account for tardiness)

We then define the full model as follows (Note: all new sets and parameters are in bold):

Sets

- D: set of drugs d (e.g. 50 mg of Taxotere) Note: Drugs are defined by hang-by time and the set of patients to receive that drug, along with their scheduled time of appointments.
- S: set of stages $s \in \{1, 2\}$
- E_d: Set of eligibility vectors for all $d \in D$. The size of this set depends on the number of doses scheduled for a given drug (eligibility vectors are further discussed in Section 3.4.1)

Parameters We define all of our parameters for CPIP-HT as follows:

- p_d^s: the time duration for processing drug $d \in D$ in stage $s \in S$
- Δ_d: the value for time savings of drug $d \in D$ (i.e., $(p_d^1 + p_d^2) *$patient time value)
- T: the total time units for the pre-mix period
- $\bar{T}_s^d = T - p_d^s + 1 \ \forall d \in D, s \in S$: Most constraints only need to consider the last time slot a drug can start being pre-made. We use $\bar{T}$ to help simplify this notation.

- c_d: the cost of one dose of drug $d \in D$
- m_d: the number of doses needed for drug $d \in D$ based on the number of scheduled patients
- C_s: total number of drugs you can process at a time at stage $s \in S$
- δ_i^d: total number of doses being mixed of drug $d \in D$ if we choose eligibility vector $i \in E$
- $time\_u_{ni}^d$: upper bound time unit on when dose $n = 1...m_d$ of drug $d \in D$ can be made given eligibility vector $i \in E$ is selected (Defined in Section 3.4.4)
- $time\_l_{ni}^d$: lower bound time unit on when dose $n = 1...m_d$ of drug $d \in D$ can be made given eligibility vector $i \in E$ is selected (Defined in Section 3.4.4)
- M: very large number
- $E_i^d[Waste\ Cost]$: the expected wast cost associated with selecting eligibility vector $i \in E$ for drug $d \in D$ (Defined in Section 3.4.3)
- $E_i^d[Reduced\ Wait]$: the expected reduced wait time associated with selecting eligibility vector $i \in E$ for drug $d \in D$(Defined in Section 3.4.3)

We then want to

$$\text{Maximize} \sum_{d \in D} \sum_{i \in E} (E_i^d[Reduced\ Wait] - E_i^d[Waste\ Cost]) * a_i^d \tag{3.14}$$

Variables

We introduce a new variable a which allows us to select a patient eligibility vector for each drug being considered for pre-mix.

- $a_i^d = \begin{cases} 1 & \text{if we select patient eligibility vector } i \in E \text{ for drug } d \in D \\ 0 & \text{o.w.} \end{cases}$

All variables from the CPIP model in Section 3.3 still hold.

Constraints

First we define our new constraints for CPIP-HT as follows:

1. You must select a exactly one patient eligibility vector $i \in E$ for a given drug $d \in D$.

   $\sum_{i \in E} a_i^d = 1 \quad \forall d \in D$

2. Here we ensure that the selected eligibility vector a_i^d of drug $d \in D$ has the corresponding start pre-mixing time slot assigned for all doses δ_i^d determined by the eligibility vector $i \in E$. For example, if we selected an eligibility vector which indicates four of five doses of a drug are pre-mixed, we must have x_{t1}^d, x_{t2}^d, x_{t3}^d, and x_{t4}^d set to 1 in exactly one of our t time slots while $x_{t5}^d = 0$ for all t.

   $\sum_{n=1}^{m_d} \sum_{t=1}^{\bar{T}} x_{nt}^d = \sum_i \delta_i^d a_i^d \quad \forall d \in D$

3. The n^{th} dose of drug d must be started in phase 2 before the upper time limit to ensure it wont expire before all eligible patients are scheduled to receive that

drug from our selected eligibility vector i. We further discuss how this time bound is calculated in Section 3.4.4.

$$\sum_{t=1}^{\bar{T}} tx_{nt2}^{d} \leq \sum_{i \in E} time\_u_{ni}^{d} a_i^d \quad \forall d \in D, n = 1...m_d$$

4. Drugs must be started in phase 2 after the lower time limit to ensure it wont expire before all eligible patients are scheduled to receive that drug from our selected eligibility vector i. We further discuss how this time bound is calculated in Section 3.4.4.

$$\sum_{t=1}^{\bar{T}} tx_{nt2}^{d} \geq \sum_{i \in E} time\_l_{ni}^{d} a_i^d \quad \forall d \in D, n = 1...m_d$$

Once a vector is selected we still utilize the CPIP constraints from Section 3.3.1

Next we further define some parameters from Section 3.4.2.

3.4.3 Probability of Waste, Expected Waste, and Expected Savings Definitions

The CPIP-HT model described in Section 3.4.2 requires us to calculate $P(n, i)$, the probability of wasting the n^{th} dose of a drug given an associated patient eligibility vector i. These probabilities, which also depend on the number of patients eligible for doses 1 through $n-1$, are in turn used to compute the expected waste cost and value of reduced wait time.

To derive $P(n, i)$, we begin by considering the probability of wasting the first dose of a drug, given patient eligibility vector e^i. Note that this dose will be used so long as at least one patient within the hang-by time window of that dose is well enough for treatment; conversely, it is only wasted if *all* of the patients eligible for that dose

defer treatment. Recall that, given a patient eligibility vector $e^i = [w_1, w_2, ..., w_m]$ where w_n is the number of patients eligible to receive the n^{th} dose of that drug. We also note that any patient in the window of eligibility for dose n is also in the window for eligibility for all subsequent doses.

We first consider the probability of $P(1, i)$, i.e., wasting the first dose made under eligibility vector e^i. We can compute P(1,i) as the product of the probabilities of deferrals of all patients from 1 to w_1,

$$\prod_{k \in \{1,...,w_1\}} p_k. \tag{3.15}$$

We next consider the probability $P(2, i)$ of wasting the second dose given patient eligibility vector i. This can happen in one of two ways – either *none* of the patients in $\{1...w_2\}$ (i.e., patients eligible to receive the second dose) receive treatment or *only one* patient from $\{1...w_2\}$ receives treatment, and this patient must also be eligible to receive the first dose and therefore be in set $\{1...w_1\}$ (i.e., there is only one patient being treated, and they receive the first dose). This can be expressed as the product of the probabilities of deferrals of all patients from 1 to w_2,

$$\prod_{k \in \{1,...,w_2\}} p_k \tag{3.16}$$

plus the second term is comprised of all combinations of the probability one patient is well enough for treatment that is also eligible for the first dose *times* the probability

the rest of the patients eligible for the second dose defer treatment

$$+ \sum_{k \in \{1 \ldots w_1\}} (1 - p_k) \prod_{j \in \{1 \ldots w_2\}/k} p_j \tag{3.17}$$

Similarly, the probability $P(3, i)$ of wasting the third dose given patient eligibility vector i can occur in three ways–all patients eligible to receive the third dose defer:

$$\prod_{k \in \{1, \ldots, w_3\}} p_k, \tag{3.18}$$

all combinations of only one patient receives treatment and that patient is also eligible to receive the second dose of the drug (which includes all patients eligible to receive the first dose of the drug):

$$\sum_{k \in \{1 \ldots w_2\}} (1 - p_k) \prod_{j \in \{1 \ldots w_3\}/k} p_j, \tag{3.19}$$

or all combinations of all but two patient receive treatment where both of these patients are eligible for the second dose and at least one is eligible for the first dose:

$$\sum_{k \in \{1 \ldots w_2\}} \sum_{l \in \{1 \ldots w_1\}} (1 - p_l) \prod_{j \in \{1 \ldots w_3\}/k,l} p_j. \tag{3.20}$$

To define this in a general sense, we recognize that if there are $k < n$ patients who receive treatment and the n^{th} dose is wasted, then at least k patients must be eligible for the dose $n - 1$, at least $k - 1$ of these must be eligible for dose $n - 2$,..., and at least 1 of these patients must be eligible for dose $n - k$. We define the set s^i_{nkj}

as the j^{th} subset of patients associated with patient eligibility vector i such that:

- $\geq k$ of the patients are in $\{1...w_{n-1}\}$
- $\geq k-1$ of the patients are in $\{1...w_{n-2}\}$
- $\geq k-2$ of the patients are in $\{1...w_{n-3}\}$

$$\vdots$$

- ≥ 1 of the patients are in $\{1...w_{n-k}\}$

We then define S^i_{nk} as a set containing all j subsets s^i_{nkj} where $|S^i_{nk}| = \binom{w_k}{k}$ for all patient eligibility vectors i and $k < n$. We now can generalize $P(n,i)$ as follows:
for $n = 1$

$$P(n,i) = \prod_{l \in \{1...w^i_n\}} p_l \tag{3.21}$$

for $n > 1$

$$P(n,i) = \prod_{l \in \{1...w^i_n\}} p_l \tag{3.22}$$

$$+ \sum_{k=1}^{n-1} \Big(\sum_{j=1}^{|S^i_{nk}|} \prod_{q \in s^i_{nkj}} (1-p_q) \prod_{r \in \{1...w^i_n\}/s^i_{nkj}} p_r \Big) \tag{3.23}$$

Expected Waste and Savings

Now that we have defined the probability of wasting the n^{th} dose of a drug given the patient eligibility vector e^i and m_d total patients, we can expand this to consider

multiple drugs and define our expected waste and value of reduced wait per unit of time as

$$E_i^d[Waste\ Cost] = c_d * \sum_{n=1}^{m_d} P_d(n,i) \tag{3.24}$$

$$E_i^d[Reduced\ Wait] = \Delta_d * \sum_{n=1}^{m_d} \Big(1 - P_d(n,i)\Big) \quad \text{where } \Delta_d \text{ is defined in Section 3.4.2} \tag{3.25}$$

where $(1 - P_d(n,i)) = 0$ and $P_d(n,i) = 0$ for all doses n not pre-mixed given patient eligibility vector e^i. We recall Δ_d as the dollar value place on a unit of reduced patient waiting time. This value is a function of how long the drug takes to mix which is the minimum time a patient would need to wait for the drug. We also recall m_d as the total number of doses (patients) scheduled for a given drug d.

3.4.4 Calculating Pre-mix Windows

To determine the time window when each dose must be pre-mixed given patient eligibility vector i, we first define some additional terminology and notation:

- T: the total time units we have to pre-mix drugs (e.g., discretizing the time units by 5 minute increments with a two-hour to pre-mix period yields $T = 24$)
- $prep_d$: total time is takes to prepare drug $d \in D$ (i.e., the processing time in both stages $p_d^1 + p_d^2$
- $hangby_d$: hang-by time for drug $d \in D$ (i.e., how long the pharmacy has to administer the drug after it is mixed) in time units

- $sch\_time_d^n$: the number of time units from the end of the pre-mix window until the patient scheduled to receive the n^{th} dose of drug $d \in D$ is ready for treatment for $n = 1...m_d$ (e.g., given a patient scheduled to receive the first dose of drug d at noon with infusion appointments starting at 8:00am, then $sch\_time_d^1 = 4$ hours $*60$ min/hour $*\frac{1}{5}$ units/min $= 48$)

- t_{nd}: the time we start mixing the n^{th} dose of drug $d \in D$ for $n = 1...m_d$

- *storage time*: the time a drug is stored after being pre-mixed before the infusion appointments start. (e.g., if a drug is finished being mixed 1 hour into the two-hour pre-mix period, the storage time will be 1 hour as well.)

- e^{di}: patient eligibility vector $i \in E_d$ given drug $d \in D$

- $time\_u_{ni}^d$: upper bound time unit on when dose $n = 1...m_d$ of drug $d \in D$ can be made given eligibility vector $i \in E$ is selected

- $time\_l_{ni}^d$: lower bound time unit on when dose $n = 1...m_d$ of drug $d \in D$ can be made given eligibility vector $i \in E$ is selected

Using this notation, for a given dose n of drug d we then calculate:

1. Complete preparation time of each dose

2. Storage time of each pre-mixed dose

3. Determine how much life a drug has left based on when it was pre-mixed before patients arrived

4. Upper and lower time limits on when a dose can be pre-mixed given patient eligibility vector $i \in E_d \; \forall d \in D$

Suppose we start preparing dose n of a given drug d at t_{nd}. Then we will complete preparing the dose at time $t_{nd} + prep_d - 1$. (Note we define our time units at the start of the unit. So a drug that takes 4 units starting at t=1 will be mixed until t=4). The *storage time* will then be

$$T - (t_{nd} + prep_d - 1) \quad \forall d \in D, n = 1...m_d \tag{3.26}$$

Given $hangby_d$ as the time until drug d must be administered, the latest time a patient can be scheduled and still receive it is the hang-by time minus the storage time (e.g., if a drug has an 8 hour hang-by time and is finished pre-mixing 30 minutes into the 2 hour pre-mix window, then there will be an adjusted hang-by time set to 6.5 hours). This is expressed as

$$hangby_d - (T - (t_{nd} + prep_d - 1)) \quad \forall d \in D, n = 1...m_d. \tag{3.27}$$

Now we can define our pre-mix windows introduced in Section 3.4.1. Referring to Figure 3.2c we see that a drug must be pre-mixed late enough that is it viable for the correct number of patients associated with eligibility vector i, but it also must be mixed early enough to not be eligible for any later patients. In our example with three patients, to determine the time windows for the patient eligibility vector $e^{di} = [1\ 3\ 0]$ we need to ensure the first dose is made late enough to be eligible for the first patient but early enough that it is not eligible for the second or third patient.

We define the upper and lower pre-mix window limits for all doses in patient eligibility vector e^{di} as follows

$$time\_l_n^d = sch\_time^d_{n=e_j^{di}} - hangby_d + T - prep_d + 1 \quad \forall d \in D, j, n = 1...m_d, i \in E \tag{3.28}$$

$$time\_u_n^d = sch\_time^d_{n=e_{j+1}^{di}} - 1 - hangby_d + T - prep_d + 1 \quad \forall d \in D, j, n = 1...m_d, i \in E \tag{3.29}$$

where e_j^{di} and e_{j+1}^{di} are the j and $j+1$ element in e^{di} (e.g., in our example $e_1^{di} = 1$ and $e_2^{di} = 3$) (Note: for any $time\_u_n^d > T$ we determine that vector as infeasible)

3.4.5 Generating Patient Eligibility Vectors

Recall the i^{th} *patient eligibility vector* e^{di} (also referred to as just an eligibility vector) associated with a drug d has one element j for each dose n of drug d (i.e., $|e^{di}| = m_d$). e^{di} specifies the number of patients eligible to receive dose n (with $e_j^{di} = 0$ if dose j was not pre-mixed). Associated with that, as described in Section 3.4.4, is a time window $[time\_l_n^d, time\_u_n^d]$ such that, if dose n is pre-mixed at any time within that time window, the number of eligible patient will be exactly e_n^{di}. Also recall that by definition, we make the first dose of a drug before the second dose and so forth. Similarly, if we don't make the n^{th} dose then we don't make any future $(n+1)^{th}$ doses either for all n. Finally, because the doses are all pre-mixed prior to the arrival of the first patient, then $e_{j+1}^{di} \geq e_j^{di}$, except for doses that are not pre-mixed (i.e., any patient eligible to receive the n^{th} dose will also be eligible for the $(n+1)^{th}$ dose.

Consider, for example, eligibility vector [1 3 4 0]. The cardinality of this vector is

four, meaning we have four patients receiving this drug and therefore four potential doses to pre-mix. $e_1^{di} = 1$, meaning that the first dose must be mixed late enough to still be viable for the first patient, but early enough that is expires before the treatment time of the second patient. $e_2^{di} = 3$, meaning that the second dose must be pre-mixed late enough to still be viable for the first three patients, but early enough that it expires before the treatment time of the fourth patient. $e_3^{di} = 3$, meaning that the third dose must be mixed late enough to be viable for all four patients. $e_4^{di} = 0$, meaning the fourth dose is not pre-mixed in this vector i. Given an eligibility vector, we then can determine the expected waste cost and expected reduced wait time associated with it as described earlier.

To determine the total number of eligibility vectors for each drug, we rely on a data structure called a *multiset* [39]. Formally a multiset A is defined as

$$A = \{a_1^{m(a_1)}, ..., a_n^{m(a_1)}\} \tag{3.30}$$

where a_n is the n^{th} unique element in the set and m is a function expressing the multiplicity of that element. One example of a multiset is expressing the prime factors of a number. For example $100 = 2^2 5^2$ which results in the multiset $\{2, 2, 5, 5\}$. In this example $a_1 = 2, a_2 = 5, m(a_1) = 2$, and $m(a_2) = 2$. Unlike a set, in which each element must be unique, a multi-set can repeat elements. However, as in a set, the order does not matter, i.e. $\{3, 4, 4\}$ and $\{4, 3, 4\}$ are the same multiset. Thus we can think of our eligibility vectors as multi-sets, and use the associated properties to count the number of potential eligibility vectors.

When counting a multiset, we define q as the cardinality of an individual multiset

and define r as the cardinality of the finite set that all of the elements of our multiset are taken from (i.e., the number of options for each element in the multiset). Therefore, the number of multisets of cardinality q with elements taken from a finite set r is given by

$$g(q,r) = \binom{q+r-1}{q}. \tag{3.31}$$

[40] In our problem context, q is the size of the eligibility vector m, (i.e., the number of doses to potentially be pre-mixed). Each element of the vector (the number of patients eligible for that dose) can be from 0 to m, with 0 representing the decision to not pre-mix. Thus $r = m+1$ resulting in us counting the number of combinations of a pre-mix vector as

$$g(m, m+1) = \binom{2m}{m}. \tag{3.32}$$

Given that Equation 3.32 is the total number of eligibility vectors for a given drug, we can then sum over all drugs d to get the total number of eligibility vectors (and associated binary decision variables) which is given as

$$\sum_{d \in D} \binom{2m_d}{m_d} \tag{3.33}$$

Table 3.1 illustrates how these values grow. However, we can also exclude eligibility vectors for which the j^{th} element of the vector is less than j since they are sub optimal (e.g., [1 1 0 0] would be suboptimal since there is no value in pre-mixing two drugs that will only be viable for the same single patient). Table 3.1 also illustrates the saving we get after pruning the sub-optimal vectors.

Number of Drugs	# of Doses for Each Drug	Number of Pre-mix Vectors Upper Bound	Number of Pre-mix Vectors Pruned	Eliminated Vectors
200	1	400	400	0
100	2	600	500	100
50	4	3500	2450	1050
25	8	321750	7940	313810
20	10	3695120	2225320	1469800

Table 3.1: Growth of the upper bounds on the number of pre-mix vectors given the total number of drugs and doses on a given day, we emphasizes this table starts with the most likely case to very rare instances of our problem

3.5 Computational Experiments

In this section we analyze the performance, tractability, and solution quality of our IP models from Section 3.3 and Section 3.4. We seek to answer the following questions:

- How large is our problem (i.e., number of variables and constraints)?
- What is the computational time to generate the model inputs and solve the model?
- How granular do we need to discretize time to still have a quality solution?
- How sensitive are the solutions based on the decision makers' value of wait time?

To run these experiments and address these questions, we use a Windows server with two Intel Xeon E5-2620 8-core processors running at 2.10-GHz with hyper-threading

and 128 GBs of RAM. The model was run with the IBM ILOG CPLEX Optimizer (version 12.8.0) Python API package.

3.5.1 Input Data Sources

We utilize data from the pharmacy's system, DoseEdge, to determine mixing time ranges and drug order quantity/frequency, as well as MiChart data to generate a sample patient schedule used in our experiments. Our collaborators from UMRCC also provided us all drug costs to the pharmacy, drug hang-by times, as well as the pharmacy capacity (e.g., number of compounding hoods and maximum number of pharmacists scheduled at one time). While the probability of deferral for patients varies in our experiments, all cases utilized data further discussed in Chapter 2.

3.5.2 Design of Experiments

To address these questions, we test three scenarios:

1. at most 1 dose of a drug is scheduled
2. 2-5 doses of each drug are scheduled
3. 10 doses of each drug are scheduled with 2 doses of one additional drug

We use a drug order schedule based on a single day with 122 patients scheduled for chemotherapy infusion treatment with a fixed two hour pre-mix window. All experiments have a fixed capacity based on three pharmacists and six technicians on the schedule during the pre-mix period (i.e., there is capacity to verify three drugs and compound six drugs at a time) to mimic the actual pre-mix window staffing

schedule. We also incorporate variable mixing times for each drug and verification in all of our experiments since some orders are more complex than others. In the first set of experiments we discretized time into five, two, of one minute increments. For example, given a drug that takes 23 minutes to mix, using five minute granularity would result in the drug reserving 25 minutes to mix, two minutes granularity would be 24 minutes and one minute granularity would be the true time of 23 minutes (i.e., we round up to the nearest discretized minute). While this drug's mixing time varies by the granularity of the discretization, pre-mixing this drug in *all* time varied cases results in 23 minutes of saved wait time.

For all three scenarios in each time discretization, 10 instances of the problem were generated and solved. These instances varied only by which drugs were randomly generated to be scheduled for use on the given day. We report the median, maximum, and minimum result of those 10 instances for each scenario. In Section 3.5.3 we address our problem size and performance questions. In Section 3.5.4, we determine the best time discretization granularity level. Lastly, Section 3.5.5 addresses how sensitive our model is to the decision makers value on reduced patient wait time. We note all models were solved with a two hour solve time limit as well as a 1% relative optimality gap limit.

3.5.3 Models Performance and Tractability

Table 3.2 highlights how much larger CPIP-HT is compared to the CPIP model in all discretized time cases. Additionally, as we increase the granularity of our time discretization for both models the problem increases in size. However, referring to

Table 3.3 we see the solve time between CPIP and CPIP-HT is not significantly different until we get to the one minute granularity case where CPIP-HT solves most instances of the problem significantly faster. However, having to generate such a large number of input parameters (i.e. patient eligibility vectors) in the CPIP-HT model, the load time is significantly longer for scenario 3 resulting in CPIP-HT actually performing slower in most instances based on total load plus solve time. We note both models follow the trend that increasing the granularity of the time discretization increases the size of the model resulting in longer load and solve times.

Model	Case	Scenario	Number of Variables	Number of Constraints
CPIP	5 min	1	10876	20688
		2	10905	21076
		3	10906	21642
	2 min	1	26806	113939
		2	26913	115085
		3	26875	118970
	1 min	1	53379	437824
		2	53580	442860
		3	53514	457218
CPIP-HT	5 min	1	11118	20930
		2	11351	21207
		3	1377780	21668
	2 min	1	27048	114181
		2	27358	115191
		3	1393749	118996
	1 min	1	53621	438066
		2	54061	442966
		3	1420388	457244

Table 3.2: Total number of variables and constraints for all scenarios of all time discretization cases for the CPIP model and the CPIP-HT model

			Load Time (sec)			Solve Time (sec)		
Model	Case	Scenario	Median	Min	Max	Median	Min	Max
CPIP	5 min	1	2.0	1.9	2.0	5.7	4.3	9.5
		2	1.8	1.7	1.9	12.0	3.4	26.4
		3	1.7	1.6	1.7	7.7	4.4	14.4
	2 min	1	6.3	6.1	6.6	31.9	22.1	254.1
		2	6.3	6.0	6.6	432.0	40.5	1290.0
		3	6.4	6.1	6.6	54.3	29.7	722.9
	1 min	1	21.1	20.0	22.0	2029.7	261.4	7200
		2	21.1	20.5	22.6	6982.7	2323.2	7200
		3	21.8	20.7	22.5	1568.6	275.8	6514.8
CPIP-HT	5 min	1	1.7	1.6	5.1	5.9	4.1	11.6
		2	1.6	1.4	1.8	10.1	5.1	18.1
		3	6632.4	6607.6	6638.0	85.1	77.2	93.6
	2 min	1	6.3	5.8	10.0	54.3	24.1	593.3
		2	6.3	6.0	6.5	446.4	53.0	1270.1
		3	7097.8	7069.5	7102.6	126.4	92.3	161.3
	1 min	1	20.4	19.4	25.5	3238.4	206.2	4933.1
		2	20.3	19.3	21.5	3200.3	340.6	7200
		3	6628.4	6597.6	6643.1	308.6	145.7	1681.0

Table 3.3: Summary table of model load and solve time for all instances in all scenarios for our variable drug cases

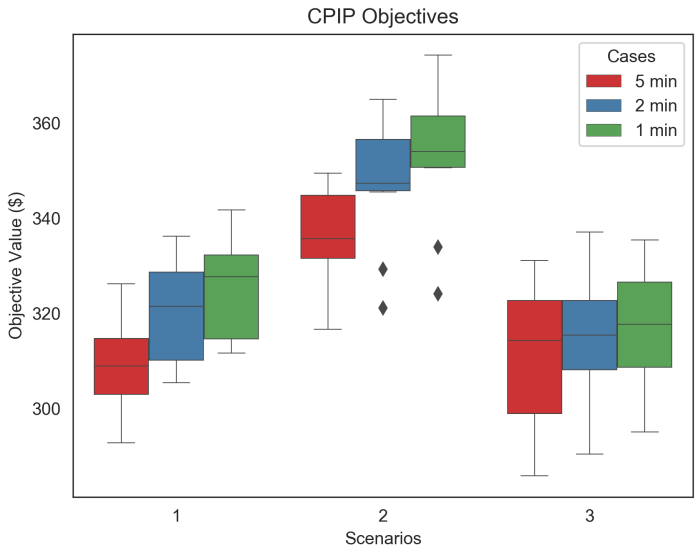
CPIP-HT's superior solve time in scenario 3 is partly attributed to the model pre-solve. Referring back to our time bounds in Equation 3.3, there are some instances of the dataset which cause the upper and lower time bound to be outside of our pre-mix window. This results in an infeasible patient eligibility vector which can quickly be discarded during pre-solve. For example, suppose we have an eligibility vector $[1, 3, 3]$ associated with a drug with a three hour hang-by time. The time

bounds on this vector would be feasible if one patient was scheduled right when infusion opened then the other three receiving the same drug were scheduled an hour later. However, if these patients were scheduled hours apart, it would not be possible to pre-mix three doses of the same drug and have the first dose only be viable for the first patient and the next two doses viable for all patients.

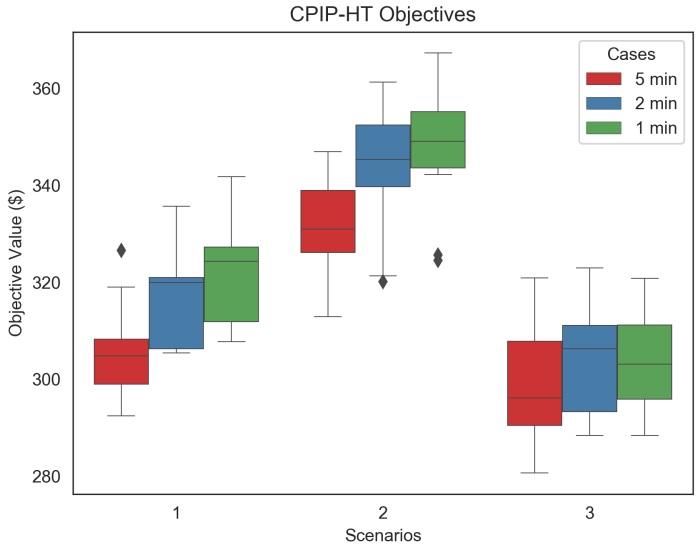
3.5.4 Solution Characteristics

Next we compare the objective values between all scenarios of the five, two, and one minute discretization cases with both CPIP and CPIP-HT models. In Figure 3.3a and 3.3b, we notice the objective values increase between scenarios 1 and 2. However, the objective value in scenario 3 decreases back to the objective value of scenario 1 or lower. While about 46-54 drugs were pre-mixed in both models for all granularity levels, only 36-38 drugs were pre-mixed in scenario 3. Although fewer drugs were pre-mixed, the pre-mixed drugs had the highest median preparation time across all instances of the problem resulting in a higher expected saved waiting time.

While not significant, there is a difference between the CPIP and CPIP-HT objective values since the later incorporates hang-by time reducing the number of feasible solutions in our CPIP-HT model.



(a) Objective values across all instances for each scenario in the CPIP Model



(b) Objective values across all instances for each scenario in the CPIP-HT Model

Figure 3.3: Objective Value for all instances and scenarios

From these results we observe there is no significant difference between the solution quality of the 2 min time granularity case with this 1 min case. However, there is a significant increase in both load and solve time when moving from 2 minute to 1 minute granularity. Therefore we continue using discretized time units of 2 minutes for the rest of our analysis.

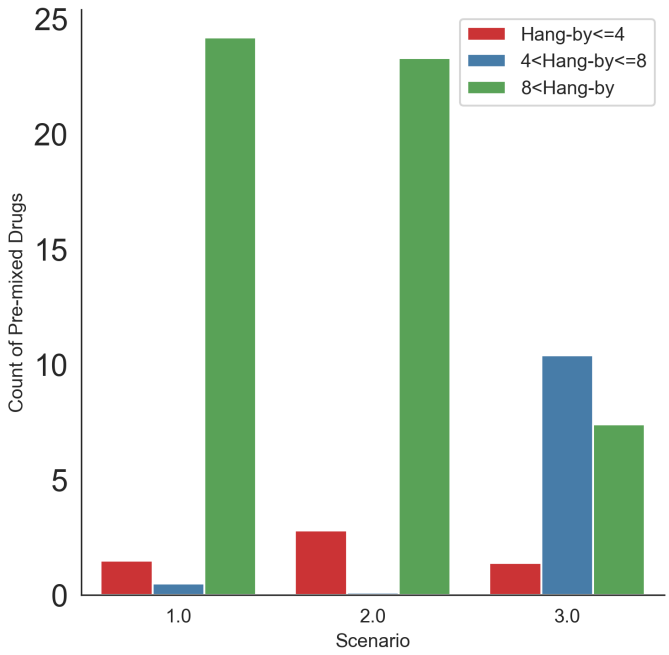


Figure 3.4: Number of drugs pre-mixed by hang-by time in the 2 min variable time drug case

3.5.5 Wait Time Value Test

Next we show how varying the value of patient wait will affect the model solution. We increment our reward by \$10 from \$0-\$100 then by increments of \$100 from \$100-\$1000 and observed how the number of pre-mixed drugs and their

characteristics changed as we increase this value. We note since our models reached optimality well within our two hour limit, we reduced the solve time limit to 1 hour to run all 10 instances for each reward value in a reasonable amount of time. As mentioned in the previous section, analysis was conducted discretizing time units to two minutes.

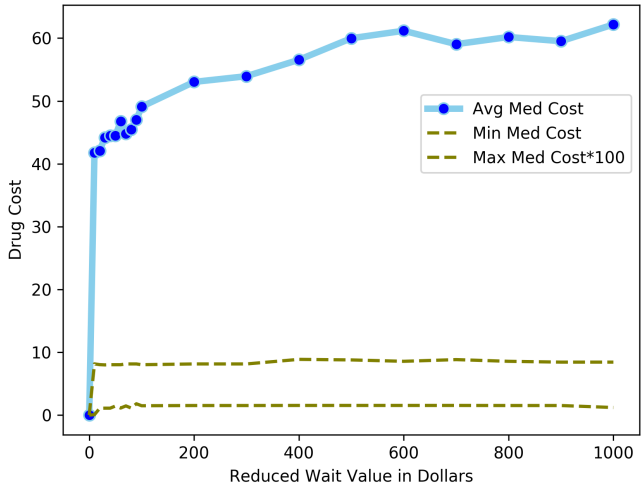


Figure 3.5: Average, minimum, and maximum median drug cost for all pre-mixed drugs as we increase the value of patient wait time (Note the max med cost is scaled by 100)

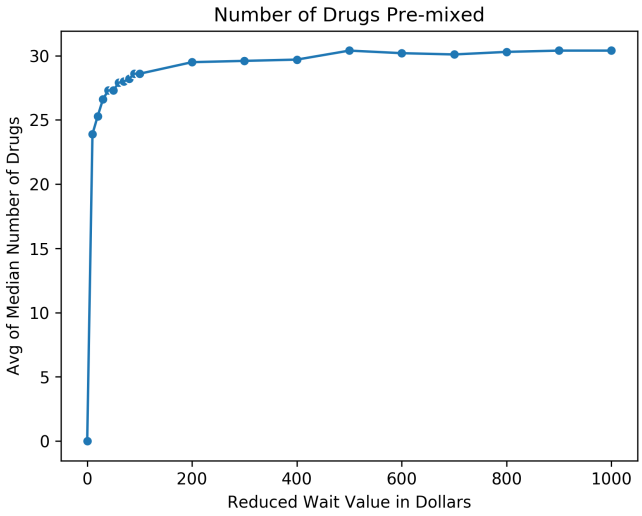


Figure 3.6: Total number of drugs pre-mixed as we increase the value of reduced wait time

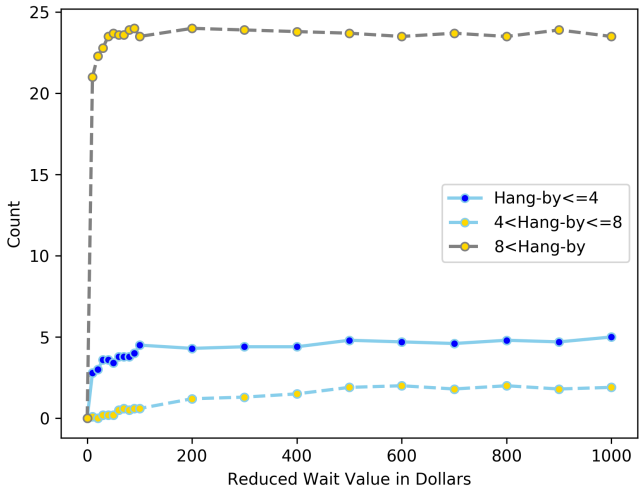


Figure 3.7: Total number of drugs pre-mixed grouped by hang-by time

We notice in each figure after our reward (i.e., valued on a patients reduced wait

time) reaches about $200, our solutions stabilize. While in Figures 3.5 and 3.6 we do see a significant increase in the median cost of all drugs and the median number of drugs pre-mixed as we increase the reward value, respectively. However in Figure 3.7, there is not as significant change in the hang-by time distribution of the drugs.

3.5.6 Summary

In Section 3.5, we test the CPIP and CPIP-HT model on 10 instances of three scenarios with various time discretization levels of our problem. We determined discretizing to 2 minutes is sufficient to maintain a good tradeoff between quality solutions and model run-time performance. We also demonstrate in some cases the CPIP-HT model reaches a more accurate solution in less computational time than the CPIP model. This can be due to many things such as the CPIP-HT model is able to eliminate many variables in the pre-solve due to some eligibility vectors being infeasible. Additionally, since our decision variable defines a series of doses to pre-mix rather than single doses, this potentially eliminates some fractional solutions which would lead to infeasible integer solutions. We also demonstrated how our decision maker might evaluate their value on patient wait time and how that affects the CPIP-HT model solutions.

3.6 Conclusion

In this chapter, we developed two IP models to help determine the optimal set of drugs an outpatient chemotherapy infusion center pharmacy should pre-mix, given a finite amount of time. By pre-mixing chemotherapy drugs before patients arrive, the

infusion center reduces the amount of time patients need to wait for their drugs to be prepared and, in turn, decreases the load on the system during peak hours in the pharmacy. Additionally, our model minimizes the projected waste cost associated with pre-mixing this set of drugs.

To do so, we first incorporated patients' probability of treatment deferrals from Chapter 2 to determine the probability of wasting a dose of a drug given it was pre-mixed with a pre-determined number of patients eligible to receive it. We developed two probability functions. The first assumes all drugs remain viable the entire day after they are mixed. The second version relaxes this assumption by addressing the time dependence between doses when considering hang-by time (i.e., the probability of wasting the n^{th} dose of a drug depends on how many patients were eligible to receive that dose as well as when all prior doses were made). Both of these functions were directly incorporated in our CPIP and CPIP-HT models, respectively, to determine the expected waste cost and value of reduced wait time associated with pre-mixing a set of doses of various drugs.

In our computational experiments, we compared various scenarios between the CPIP and CPIP-HT models. While CPIP-HT run time was significantly higher, it still remains tractable and is the more accurate model since it incorporates drug hang-by time. Through our experiments, we show an optimal set of drugs for pre-mix can be found in a timely manner while minimizing waste cost and maximizing the value of reduce wait time of a patient. We also empirically showed the point at which a decision maker's value on wait time no longer affects the IP solution.

As future work, we look to expand how we define the value of reduced wait

time. Currently all doses of the same drug have the same value of reduced wait as it is only dependent on the time it takes to prepare the drug in the pharmacy. However in Chapter 4, we show that pre-mixing drugs for patients strategically spaced throughout the schedule is more beneficial then pre-mixing for the first few patients in the morning. To address this, we suggest introducing a time-dependent reward. Consequently, the value of reduced wait for a dose of a drug being pre-mixed for a patient scheduled during peak hours would be higher than another dose of the same drug being pre-mixed for a patient schedule during off-peak hours.

Chapter 4

Modeling the Impact of Pre-mix Policies through Discrete-Event Simulation of the Pharmacy

4.1 Introduction

Healthcare operations, care delivery, and treatment decision making are just a few examples where data-driven modeling has greatly impacted performance and grown in popularity as data access has improved [41]. Discrete-event simulation stands as one of the most tractable methods practiced to improve overall healthcare system performance [42], [43], [44], [45]. This increasing trend of performance improvement models exists for cancer treatment centers as well [46], [47].

Cancer patients who require chemotherapy infusion undergo exhausting and lengthy infusion sessions over the course of their treatment. These session lengths can increase even more during peak demand hours due to delays in the pharmacy. Through observations of work flow and interviews with pharmacists at UMRCC, we determined peak drug demand hours on certain days of the week. During

these peak hours, the pharmacy can get backed up to the point of taking up to two hours to get a drug out to a patient. Therefore, one major opportunity to reduce patient delays is by optimizing drug preparation at the pharmacy. Drugs can be prepared (i.e., compounded) ahead of time to prevent patients from waiting as their chemotherapy drug(s) is/are compounded as done in [10]. However, patients scheduled for outpatient chemotherapy infusion may be deferred for treatment (i.e., last minute cancellation due to not meeting medical parameters or personal reasons such as a family member not being available to support the patient) after arriving for their appointment [11]. Consequently, the infusion center may incur a waste cost if a drug is made ahead and the patient defers treatment. Infusion centers must implement policies, determining which drugs to make before patients arrive given a fixed window of time, to balance this potential waste cost with the time savings for their patients and staff. In support of this effort, we developed a discrete-event simulation, using UMRCC data, that has widespread applicability to evaluate the effectiveness of pre-mixing chemotherapy drugs as well as the drug planning process in general. This simulation allows us to take multiple sets of drugs determined for pre-mix, along with the patient schedule for the day, and simulate the performance (i.e., time in system and staff utilization) of these sets (i.e., compare the pre-mix policies). Through our computational experiments, we test proposed methods to create pre-mix policies. For example, if a drug is below a certain cost threshold and one or more patients are scheduled to receive it below a certain probability of deferral, then we can pre-mix the drug. We incorporate patient deferral probabilities from the prediction model in Chapter 2. The simulation allows us to test various

make-ahead policies for mixing drugs throughout the day. Improvements made at the pharmacy may reduce patient delays and nurse overtime in the infusion area. We propose various policies ranging in risk tolerance. More conservative policies mimic the current state of pre-mix which are solely cost based. Conversely, less conservative policies consider pre-mixing higher-cost drugs if the probability of wasting them is low enough. Although all policies save time, the purpose of the simulation is to determine how much time and whose time (patient or pharmacist/tech).

As previously mentioned we use a discrete-event simulation model to predict the effectiveness of various make-ahead drug policies and validate this model using a case study based on data from the UMRCC. The expected wait time for premixing a drug is determined by the known preparation time of the drug. In 3, we introduced the formulation for an expected saved wait time for a patient if their drug is pre-mixed. This expectation is based on an assumed distribution for the mixing time of that patients drug. This also assumes that the pharmacy is running smoothly and never gets backed up during the day. Through observations and interviews with pharmacists at UMRCC, we know there are peak drug demand hours on certain days of the week. During these peak hours, the pharmacy can get backed up to the point it takes up to two hours to get a drug out to a patient. We then utilize the discrete event simulation to relax this assumption and to better predict the affect of pre-mixing on the entire system rather than for the specified patient.

Next we will use the simulation to test various policies for mixing drugs throughout the day. Discrete simulation will also let us test multiple objectives such as nurse utilization, chair utilization, patient waiting time, and unit closing

time as done in [48]. Instead of just considering the drug cost for various policies as done in [10], we also introduce how to incorporate the patient deferral probabilities from our prediction model in Chapter 2. For example, if a drug is below a certain cost and two or more patients are scheduled to receive it below a certain probability of deferral then we can premix the drug. We then want to explore the pharmacy interaction with other areas in the cancer center (i.e. model the entire network). Improvements made at the pharmacy may reduce patient delays and nurse overtime in the infusion area. Discrete simulation has been used in other outpatients cancer centers with similar objectives [46], [47]. The rest of this section is structured as follows: Section 4.2 defines the problem we are solving, Section 4.3 describes the simulation modeling construction and all assumptions, Section 4.4 presents our computational experiments, and Section 4.4.4 provides discussion and conclusion.

4.2 Problem Description

The drug mixing process at the pharmacy consists of a series of order verifications, the actual compounding of the drug, and safety checks to ensure safe delivery to patients. The process we model begins when a drug order has been received by the pharmacy (i.e., the patient has arrived at infusion and is ready to be treated). Figure 4.1 illustrates the various steps taken and various checks needed to complete a chemotherapy drug order. If a drug is pre-mixed, all steps are performed in advance except for the final safety check. The drug mixing process is carried out by various pharmacist and technical staff. Pharmacists conduct both the order verification and safety check of all drugs. They are assigned to one of these two tasks for the first half

of the day then switch. Pharmacy technical staff serve one of two roles: compounding drugs under the hood or printing all labels and collecting supplies.

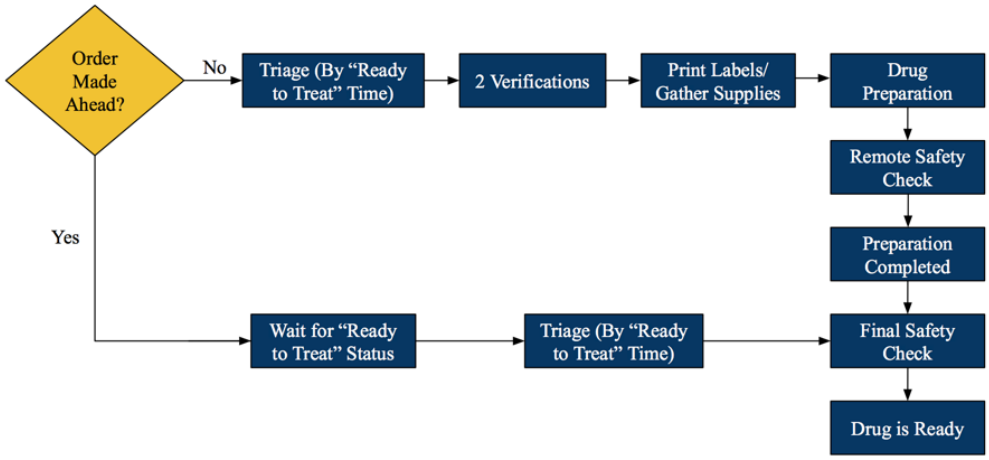


Figure 4.1: Pharmacy process flow for chemotherapy infusion drug orders

For simulation purposes, we simplified this process flow into five main steps.

1. Check if the drug was made ahead (pre-mixed), if yes skip to Step 5 otherwise continue to Step 2.
2. Complete two drug verifications; must be done by two different pharmacists
3. Print labels and collect supplies for the drug order
4. Compound the drug and perform the first safety check. If order is mixed incorrectly, the same tech must remake the order (i.e. repeat Step 4)
5. Perform the safety check and deliver the completed order.

Arrivals are determined based on the patients' appointment time plus some random deviation where a negative deviation means the patient was early and

positive means they were late. Each time a new order arrives, the orders are sorted first by arrival time then by appointment time (i.e., if two orders arrive at the same time before the pharmacy opens, the order with the earlier appointment time will be processed first). Orders are then released to follow the process in Figure 4.1. Once a drug is verified twice (Step 2) and all supplies have been gathered (Step 3), it then can proceed to the mixing hood. After the drug is mixed, the remote safety check is performed by the pharmacist. Before compounding the drugs, the mixing tech must take a photo of all materials and another photo after the drug has been fully compounded. The safety check pharmacist must review these images before the mixing process can be marked as complete. However, during this check, the mixing tech cannot conduct any other work. Since the historical data is only captured when a drug entered and exited the mixing hood, we included this initial safety check in the mixing time for our drugs. We note this safety check is performed in addition to the final safety check/drug delivery in step 5 of our simulation. Refer to Appendix C for more detailed simulation logic.

4.3 Simulation Model

Our simulation was built using the *SimPy* module in Python version 3.6. The model is initiated with a set of orders scheduled for the day, with some being flagged as pre-mixed, each of which has a scheduled arrival time for the corresponding patient. The actual arrival time of orders, the length of each service, and the potential for the compounding service to fail are all stochastic. Due to the interest of our collaborators, variable staffing was outside the scope of this project. Our metrics

are the time in system for each order (i.e., perceived patient wait time) and the utilization of all resources (i.e., pharmacist and techs at the pharmacy).

4.3.1 Input Parameters

Our simulation input parameters are determined using a combination of observations, historical data, and expert pharmacist opinion. Using the *Scipy* module in Python, we fit all distributions used for arrival time and service time estimation as seen in Table 4.1.

Process	Distribution	Description
Patient Arrivals	JohnsonSU (-0.428,1.41,-2.767,45.511)	Negative values=early arrival Positive=late arrival
First Verification	Triangular (1,2,15)	Expert Opinion in min
Second Verification	Triangular (1,2,5)	Expert Opinion in min
Print Labels/Kit	Triangular (1,3,5)	Expert Opinion in min
Drug Mix Time	Beta (1.461, 1376723443.471, 1.019, 7036129537.303)	Historical Data
Safety Check	Pearson3(2.509, 3.583, 3.240)	Historical Data

Table 4.1: All input distribution and parameters used in our simulation model

While the arrivals are appointment driven, most patients' actual arrival time will deviate from their scheduled appointment time due to the stochastic nature of any previous appointments as well as general tardiness or earliness. An additional delay can occur since a pharmacy order is only initiated once a patient has arrived and checked-in for their infusion appointment. We approach modeling the arrival process by first determining the distribution of the deviation from scheduled appointment time to actual arrival time (check-in time) similar to [49] [50]. Figure 4.2 presents a histogram of the arrival deviation data along with the JohnsonSU distribution used for arrival estimations in the simulation similar to [44].

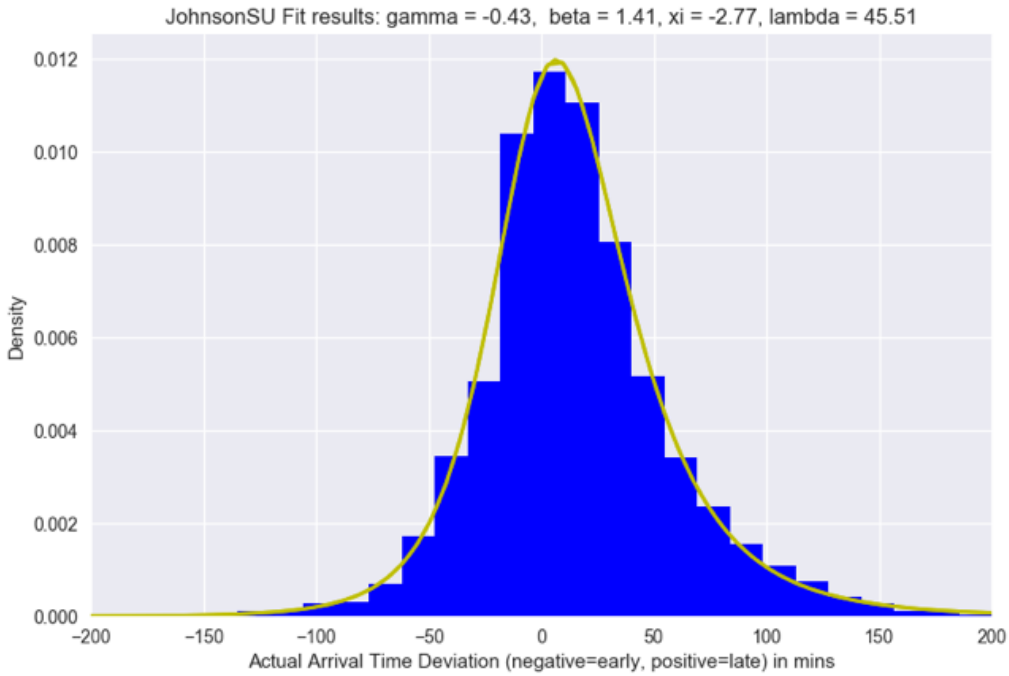


Figure 4.2: JohnsonSU Arrival Deviation Distribution Fit based on historical data taken from UMRCC MiChart (Epic Product) Medical Records from 2016

The model also incorporates patients deferrals/no-shows. From the pharmacy perspective, this will only affect the metrics if a patient's drug is pre-mixed and they have a same day deferral/no-show. To determine this probability, we trained a Bayesian Additive Regression Tree (BART) model with demographic (e.g., race, gender, age, sex, and marital status), scheduling history (i.e., number of previous cancellations, previous appointments, and past treatment protocols), and medical treatment patient data to make the treatment deferral/no-show predictions for each patient. This model was chosen based on a comparative procedure, discussed in Chapter 2, which tested multiple binary decision models to determine the model with the best out-of-sample prediction based on AUC, Briar Score, and F-1 Score.

With a decision threshold of probability .75, we correctly predicted 93% of completed appointments and 21% of deferrals/no-shows with an overall accuracy of 84%. We note that while a false negative is not ideal, there is no waste cost associated with it (i.e., the drug would be mixed as planned.) However, a false positive could result in a drug being wasted if pre-mixed. We emphasize that the UMRCC system was used as an example case study. The presented computational experiments utilized historical data from 2016 and expert opinion from the UMRCC pharmacy to estimate model inputs and validity. We note our modeling approach can be replicated for other facilities with minor modifications to the pharmacy process flow as well as the appropriate data sources. After observing the simulated no pre-mix case discussed in Section 4.4, it was determined a reasonable estimate if the pharmacy did not pre-mix based on the observed data. We highlight that we focus on comparing various make-ahead policies as a proof of concept.

4.3.2 Modeling Conditions

The infusion center pharmacy opens at 6:00 am to complete all pre-mix drug orders. We assume that all pre-mixed orders will not expire before they are administered if the patient appointment is at noon or before (most drugs have an 8-12 hour lifetime). At 7:30am they finish pre-mixing and begin making orders as patients arrive (earliest appointment time is 7:30am). We assume if a drug is pre-mix that it will be completed in this window of time. This is when our simulation starts and runs until all patients are served each day. The simulation only considers a single arrival stream of patients based on their appointment time in infusion. There are

some cases when a same day appointment can occur but the majority of appoints will be scheduled ahead of time. The simulation then models a single drug order for each patient with a drug compounding probability of failure of 5%. Through observations and interviews with pharmacists, we found this to be a reasonable assumption. There is also a chance for pre-mixed orders to be wasted if a patient defers or doesn't show. We assume constant patient volumes for each iteration of the simulation buy vary them by day of week to best represent the cancer centers current schedule.

4.4 Computational Experiments

We present three tractable and easily implementable pre-mix policies. Then we compare their performance with the baseline scenario (i.e., no pre-mix policy). Table 4.2 and 4.4 highlight our performance metrics (total time in system and staff utilization) compared across scenarios. Since staffing is out of the scope of this paper, we keep a constant schedule across all scenarios based on the current staffing schedule at UMRCC. We simulate a week (i.e., Monday through Friday) with a constant amount of pre-mixed drugs each day since the pharmacy has the same limited time to pre-mix drugs each day. This week was simulated 20 times to insure our average drug order time in system metric stayed within an error of 4 minutes for all days and scenarios based on a 95% confidence level. We first ran 10 iterations of our simulation to determine our sample mean and standard deviation of the average drug order time in system. We then utilized the t-distribution to find the number of iterations required to ensure our error or confidence interval half with was at most 4 minutes.

4.4.1 Scenario 1

In the first scenario, we only consider pre-mixing drugs for the first 20 (this number is adjustable and dependent on the time allotted for pre-mix) patients who have a probability of deferral/no-show of 0.1 or lower. This threshold is used as an example; it is ultimately determined by the decision maker, depending on their risk tolerance. The probability of deferral/no-show for patients scheduled to receive the drug orders were determined using a BART model with the input data mentioned in Section 4.3.1.

4.4.2 Scenario 2

For our second scenario, we disregard the probability of deferral/no-show and focus solely on patient appointment times when deciding to pre-mix. Since the pharmacy schedule is heaviest before noon, we focus on pre-mixing only for patients with scheduled appointments from 8am-12pm . This also ensures that all pre-mixed drugs won't expire before the patients' appointments. Next, we determine the proportion of appointments in each hour block from the first appointment until noon. This proportion is used to determine the number of pre-mixed drugs in a respective hour block (e.g., suppose from 8-9am there are 20 patients out of a total of 100 that morning, then given we can only pre-mix 20 drugs, the first 5 drugs in that time window will be pre-mixed).

4.4.3 Scenario 3

In our last scenario, we combine ideas from both Scenarios 1 and 2. Using the proportions for each hour block determined in Scenario 2, we then only assign the allocated number of pre-mixed drugs to patients that fall below the probability of deferral/no-show threshold from Scenario 1. This policy should incorporate the benefits of pre-mixing throughout the first half of the day while being more risk adverse in regards to wasting drugs. This also may spread out the appointments that have pre-mix drugs (assuming the first five patient in our previous example do not all fall under the probability of deferral/no-show threshold).

4.4.4 Results/Discussion

Referring to Table 4.2 and 4.3, it is clear that pre-mixing chemotherapy drugs has a significant impact on a drug orders average time in the system and the percentage of patients waiting past the goal TAT of 1 hour as Scenarios 1, 2, and 3 all outperform the No Pre-mix case. On Day 3 we see that Scenario 1 out performs Scenario 2. This is an example where being more risk-seeking by not considering the probability of deferral/no-show had less reward (i.e., drugs may have been pre-mixed for patients who actually deferred treatment resulting in no benefit for pre-mixing) for the patients. However, all other days we see Scenario 2 dominates 1. This results from Scenario 1 only mixing the first 20 drugs that fit the criteria. This reduces the potential propagated time savings compared to spreading the pre-mixed drugs throughout the schedule. While not always significant, we see that the average time in system for Scenario 3 and the average number of wasted drugs is lower than

Scenarios 1 and 2. We hypothesized this to be the case since Scenario 3 is a more efficient r̈ule of thumbp̈olicy by utilizing the more conservative approach of Scenario 1 but also by lightning the pharmacy load throughout the morning instead of just in the first couple of hours similar to Scenario 2. While the average percentage of drug order TATs greater than 1 hour in Scenario 3 slightly increases for days 1,4, and 5, it is not by a significant amount compared to the improvements on the heavier day 3 with 153 drug orders compared to 121, 116, and 129 from scenarios 1, 4, 5 respectively.

Drug Order Time in System and Wasted Drug Results

	Metrics	Scenarios			
Day of Week		No Pre-mix	1	2	3
1	Average	52.79	30.7	26.64	26.17
	C.I.	(49.51, 56.07)	(29.53, 31.87)	(26.23, 27.05)	(25.65, 26.68)
2	Average	85.63	46.6	41.73	38.19
	C.I.	(80.79, 90.46)	(44.43, 48.8)	(39.21, 44.25)	(36.19, 40.19)
3	Average	58.04	35.44	37.69	27.47
	C.I.	(54.74, 61.34)	(33.65, 37.22)	(34.87, 40.51)	(26.59, 28.35)
4	Average	38.1	24.78	22.82	22.43
	C.I.	(36.3, 39.89)	(24.18, 25.37)	(22.35, 23.33)	(22.06, 22.81)
5	Average	47.86	28.32	25.73	25.7
	C.I.	(44.95, 50.78)	(27.53, 29.09)	(25.27, 26.19)	(24.71, 26.7)
Avg # of drugs wasted per day		0	2.81	3.13	2.32

Table 4.2: Contains the comparison of time in system for the various scenarios in minutes

Percent of Patient's Drug Orders with TAT >1 Hour

		Metrics	Scenarios			
Days	# of Orders		No Pre-mix	1	2	3
1	121	Average	34.38	7.23	1.53	1.94
		(Min, Max)	(20.66, 52.07)	(2.48, 19.01)	(0, 4.96)	(0, 3.3)
2	168	Average	77.65	50.57	35.71	35.45
		(Min, Max)	(71.43, 82.74)	(36.31, 66.67)	(16.07, 61.9)	(16.67, 52.38)
3	153	Average	59.41	19.31	29.29	5.78
		(Min, Max)	(33.33, 76.47)	(5.88, 36.6)	(13.68, 53.57)	(.65, 14.37)
4	116	Average	27.63	5.56	1.34	2.07
		(Min, Max)	(17.24, 37.93)	(1.72, 16.38)	(0, 3.45)	(.86, 4.31)
5	129	Average	43.76	7.13	3.06	4.26
		(Min, Max)	(31.78, 61.24)	(2.33, 22.48)	(0, 10.08)	(0, 12.4)

Table 4.3: Average, maximum, and minimum percentage of patient who wait past the 1-hour TAT after running 20 iterations of the simulation

Looking at Day 2 (highest drug order demand day of the week) in Table 4.4, we also notice a significant effect on staff utilization. For example, both Verification Pharmacists and Printing Technicians have a very high utilization in the No Pre-mix case. By simply pre-mixing 20 drugs before the patient rush, we are able to decrease their utilization by almost 10% and 5% respectively. This supports the idea that pre-mixing chemotherapy drugs will benefit both the patients and the staff to better ensure safe delivery of such high hazardous drugs. Since we assume the same distribution for mixing all chemotherapy drugs and that arrivals are appointment

driven, the utilization across all pre-mix policies is relatively the same. We note that our utilization calculation for the Safety Check Pharmacist are an extreme under-estimate since we were only able to capture the start time for compounding a drug to the completed time. As seen in Figure 1.2, this also incorporated the Remote Safety check which is performed by the Safety Pharmacist. It is here where almost all of the 5% of failed drug orders are determined and cycled back for re-compounding. Refer to Appendix C for utilization results on all other days.

Day 2 Utilization

Resource	Scenarios											
	No Pre-mix			1			2			3		
	Avg.	CI-	CI+	Avg.	CI-	CI+	Avg.	CI-	CI+	Avg.	CI-	CI+
Verification Pharm	0.89	0.8	0.99	0.82	0.77	0.87	0.81	0.77	0.85	0.81	0.77	0.86
Print Tech	0.71	0.56	0.85	0.65	0.62	0.67	0.65	0.61	0.69	0.65	0.62	0.68
Mix Tech	0.3	0.28	0.32	0.28	0.27	0.29	0.26	0.25	0.27	0.27	0.26	0.28
Safety Pharm	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table 4.4: Utilization comparison between policies on Day 2 of our simulation as percentages

4.5 Conclusion

Using discrete event simulation, we evaluated various pre-mix policies to determine which most benefited both the individual patients and the entire outpatient infusion system. We used estimated probabilities of deferral/no-show from our BART model to develop these policies for chemotherapy drugs at an outpatient chemotherapy infusion center pharmacy. These experiments were not meant to determine the best approach for UMRCC but to demonstrate the idea of

pre-mix as a technique that can be utilized by other institutions. The discussed model serves multiple purposes both for our current and future work. While we can test current r̈ule of thumbp̈olices as done in this paper, we also can evaluate optimization models that determine, within a fix window of time, what set of drugs are optimal to pre-mix to minimize the expected time in system as well as expected waste cost. Our immediate next steps will incorporate the drug cost into the simulation model as well as introduce additional patient arrival streams from those with a clinic visit. Future extensions will also allow for more dynamic policies to be tested that provide unattainable improvements from a static model. This work provides an invaluable tool to both engineers and medical professional working to reduce patient waiting time in an outpatient chemotherapy infusion center by helping insure the safety of the patients and improves their overall satisfaction.

Chapter 5

Conclusion

5.1 Conclusion/Future Work

In this dissertation, we proposed three methods which, in conjunction, generate and evaluate make-ahead chemotherapy drug policies (i.e., pre-mix policies) at an outpatient chemotherapy infusion center.

In Chapter 2, we developed a predictive model to determine the probability a patient will defer their treatment on a given day. We described the comparative procedure used during the modeling selection process and emphasize this selection process must be repeated if applied in another hospital setting. We found a patient's previous number of cancelations and the number of days since the patient's last visit had the most positive influence on predicting a treatment deferral. Conversely, a patient's appointment length, previous number of infusion center visits, and which treatment regimen they were on had the strongest negative influence on predicting a treatment deferral (i.e., higher the value the greater chance a patient will not defer). Future research directions include exploring additional data re-sampling methods

(e.g., synthetic data generation, stratified sampling) and additional classification models such as Naive Bayes and support vector machines. We also suggest to explore additional forms of feature selection that are not limited to stepwise selection and variable importance, as done in this dissertation.

In Chapter 3, we utilized the patient probability of deferral from Chapter 2 in our IP models to determine the optimal set of drugs to pre-mix given a fixed window of time. We introduce two versions of the problem with their respective model formulations. The CPIP Model assumes all drugs pre-made will last (no expiration) for the entire day. While this model provides a fast and valuable solution, we found its limitations include not considering drugs with varied expiration or hang-by times affected the accuracy of our results. We then introduced the CPIP-HT model which incorporated drug hang-by time in determining the optimal set of drug to pre-mix and when to mix them. This model also allows for further extensions of considering pre-mixing drugs for multiple days given the capacity and required expirations dates. The CPIP-HT formulation addresses the time dependence and interdependences between pre-mixing multiple doses of the same drug (i.e., when the n^{th} dose is pre-mixed will affect the probability of waste of any subsequent dose). We show our new model remains tractable in many realistic instances of the problem.

We emphasize that our objective function from both models captures a conservative estimate on how much wait time is saved for patients prescribed the pre-mixed drugs. In practice, these time savings would propagate through the system benefiting many patients. Utilizing the simulation model from Chapter 4, we evaluated our CPIP-HT relying on the same parameters from our computational

experiments using a two minute discretization for our time units from Section 3.5.2.

Table 5.1 and 5.2 display the results of our implementation. In Table 5.1 we reported the average between 20 iterations of the total time in system for each drug order. The CPIP-HT model suggested pre-mixing from 25-27 drugs each day, resulting in an estimated 40% reduction in our total time in system metric on average. This time directly reflects the amount of time a patient must wait (either in the waiting room or the infusion chair) for their drug. Additionally we measure the number of patients who waited longer than 1 hour for their drug order in Table 5.2. At the time of our study, this was the goal turn-around time for the pharmacy. Our simulation estimates on average a 70% decrease in this metric. In Chapter 4, showed how strategically selecting which drugs to pre-mix based on when their prescribed patients with a low probability of deferral are scheduled to receive their drug proves to benefit patient time savings even more than our model. This led us to a potential next step of introducing time depended rewards for pre-mixing chemotherapy drugs in the CPIP-HT model.

Average Patient Observed Drug Order Time in System (min)			
	Metrics	Scenario	
Days		No Pre-mix	CPIP-HT
1	Average	52.79	28.39
	CI	(49.51, 56.07)	(27.04, 29.74)
2	Average	85.63	51.88
	CI	(80.79, 90.46)	(49.85, 53.91)
3	Average	58.04	29.42
	CI	(54.74, 61.34)	(28.24, 30.61)
4	Average	38.10	24.07
	CI	(36.3, 39.89)	(23.26, 24.87)
5	Average	47.86	37.63
	CI	(44.95, 50.78)	(36.14, 39.12)

Table 5.1: Comparison of simulated results of using the CPIP-HT model to the not pre-mixing at the UMRCC for a single week

Average Percent of Patients Drug Orders Past 1-Hour TAT Goal				
		Metrics	Scenarios	
Days	# of Orders		No Pre-mix	CPIP-HT
1	121	Average	34.4	6.7
		(Min, Max)	(20.66, 52.07)	(0, 14)
2	168	Average	77.6	41.0
		(Min, Max)	(71.43, 82.74)	(31.5, 53.8)
3	153	Average	59.4	6.4
		(Min, Max)	(33.33, 76.47)	(1.3, 13.7)
4	116	Average	27.6	2.7
		(Min, Max)	(17.24, 37.93)	(.86, 5.2)
5	129	Average	43.8	25.0
		(Min, Max)	(31.78, 61.24)	(14, 33.3)

Table 5.2: Percentage of patient that waiting longer than the 1 hour goal TAT. We compare the results from the no pre-mix case to using the CPIP-HT recommendations

Based on these results, we determine one potential next steps of defining a time-dependent parameter representing the value on reduced wait time when a drug is pre-mixed. This allows the decision maker to weight pre-mixed drugs for patients scheduled during peak hours higher than patients schedule at less congested times of the day to potentially better decrease the load on the system during that time. Additional next steps include performing a sensitivity analysis on all input parameters to gain further insights on how our model outputs might change with another patient population.

In Chapter 4 we introduced a pre-mix policy evaluation tool which utilizes discrete-event simulation methods as well as the probabilities of patients' deferrals

determined in Chapter 2. We demonstrated the benefits of pre-mixing chemotherapy drugs by simulating make-ahead "rule-of-thumb" policies. We also used this tool to support the idea that pre-mixing a drug not only benefits the prescribed patient, but also benefits other subsequent patients as we lighten the load on the system during infusion peak hours. Lastly, we compared the results from a realistic instance of the CPIP-HT model from Chapter 3 with the no pre-mix base case. Potential next steps involve exploring simulation optimization methods to generate more robust pre-mix policies that can be utilized throughout the day rather than just a fixed period of time as done in Chapter 3.

Bibliography

Bibliography

[1] A. C. S. ACS, "Cancer Facts & Figures 2019," p. 76.

[2] L. Levit, M. Patlak, *et al.*, *Ensuring Quality Cancer Care through the Oncology Workforce: Sustaining Care in the 21st Century: Workshop Summary*. National Academies Press.

[3] D. Schrijvers, "Extravasation: A dreaded complication of chemotherapy," vol. 14, pp. iii26–iii30.

[4] B. Y. He, K. Mendelsohn-Victor, M. C. McCullagh, and C. R. Friese, "Personal Protective Equipment Use and Hazardous Drug Spills among Ambulatory Oncology Nurses: Results from a Mailed Survey," in *Oncology Nursing Forum*, vol. 44, p. 60, NIH Public Access.

[5] J. M. Harrison, B. Lavasseur, P. J. Stella, P. T. Adams, L. Swafford, J. Lewis, K. Mendelsohn-Victor, and C. R. Friese, "Factors associated with toxicity-related service use among community oncology patients.,"

[6] G. Lam, O. Jouini, and J. Stal-Le Cardinal, "Outpatient chemotherapy planning: A literature review with insights from a case study," vol. 6, no. 3, pp. 127–139.

[7] N. N. C. Institute, "A to Z List of Cancer Drugs."

[8] V. V. Padma, "An overview of targeted cancer therapy," vol. 5, no. 4.

[9] A. C. S. ACS, "How Is Chemotherapy Used to Treat Cancer?."

[10] I. H. Masselink, T. L. van der Mijden, N. Litvak, and P. T. Vanberkel, "Preparation of chemotherapy drugs: Planning policy for reduced waiting times," vol. 40, no. 2, pp. 181–187.

[11] S. Fuentes and J.-E. Frdin, "Why is intravenous chemotherapy cancelled and how often. Could it be prevented? A prospective analysis of all planned and given intravenous anti-tumor treatments at the Department of Oncology, Karolinska University Hospital, Stockholm during one month," vol. 54, no. 7, pp. 1056–1062.

[12] Admin, "Facts and Figures,"

[13] N. L. Lacy, A. Paulman, M. D. Reuter, and B. Lovejoy, "Why we don't come: Patient perceptions on no-shows," vol. 2, no. 6, pp. 541–545.

[14] K. Muthuraman and M. Lawley, "A stochastic overbooking model for outpatient clinical scheduling with no-shows," vol. 40, no. 9, pp. 820–837.

[15] C. Zacharias and M. Pinedo, "Appointment Scheduling with No-Shows and Overbooking," vol. 23, no. 5, pp. 788–801.

[16] L. R. LaGanga and S. R. Lawrence, "Clinic overbooking to improve patient access and increase provider productivity," vol. 38, no. 2, pp. 251–276.

[17] A. Alaeddini, K. Yang, P. Reeves, and C. K. Reddy, "A hybrid prediction model for no-shows and cancellations of outpatient appointments," vol. 5, no. 1, pp. 14–32.

[18] J. Daggy, M. Lawley, D. Willis, D. Thayer, C. Suelzer, P.-C. DeLaurentis, A. Turkcan, S. Chakraborty, and L. Sands, "Using no-show modeling to improve clinic performance," vol. 16, no. 4, pp. 246–259.

[19] Y. Huang, D. Hanauer, *et al.*, "Patient no-show predictive model development using multiple data sources for an effective overbooking approach," vol. 5, no. 3, pp. 836–860.

[20] J. Futoma, J. Morris, and J. Lucas, "A comparison of models for predicting early hospital readmissions," vol. 56, pp. 229–238.

[21] C. G. Walsh, J. D. Ribeiro, and J. C. Franklin, "Predicting risk of suicide attempts over time through machine learning," vol. 5, no. 3, pp. 457–469.

[22] T. Razzaghi, O. Roderick, I. Safro, and N. Marko, "Fast imbalanced classification of healthcare data with missing values," in *Information Fusion (Fusion), 2015 18th International Conference On*, pp. 774–781, IEEE.

[23] J. Huang and C. X. Ling, "Using AUC and accuracy in evaluating learning algorithms," vol. 17, no. 3, pp. 299–310.

[24] B. Hanczar, J. Hua, C. Sima, J. Weinstein, M. Bittner, and E. R. Dougherty, "Small-sample precision of ROC-related estimates," vol. 26, no. 6, pp. 822–830.

[25] G.-G. P. Garcia, M. S. Lavieri, M. A. McCrea, T. W. McAllister, and S. P. Broglio, "Data-driven Stochastic Optimization Approaches to Determine Decision Thresholds for Risk Estimation Models," p. 51.

[26] A. Huggins, D. Claudio, E. Prez, and S. Marcos, "Improving Resource Utilization in a Cancer Clinic: An Optimization Model," p. 9.

[27] A. Rais and A. Viana, "Operations Research in Healthcare: A survey," vol. 18, no. 1, pp. 1–31.

[28] J. T. Blake and J. Donald, "Mount Sinai Hospital Uses Integer Programming to Allocate Operating Room Time," vol. 32, no. 2, pp. 63–73.

[29] P. Santibez, R. Aristizabal, M. L. Puterman, V. S. Chow, W. Huang, C. Kollmannsberger, T. Nordin, N. Runzer, and S. Tyldesley, "Operations Research Methods Improve Chemotherapy Patient Appointment Scheduling," vol. 38, no. 12, pp. 541–AP2.

[30] J. Castaing, A. Cohn, B. T. Denton, and A. Weizer, "A stochastic programming approach to reduce patient wait times and overtime in an outpatient infusion center," vol. 6, no. 3, pp. 111–125.

[31] Y.-L. Huang, A. H. Bryce, T. Culbertson, S. L. Connor, S. A. Looker, K. M. Altman, J. G. Collins, W. Stellner, R. R. McWilliams, A. Moreno-Aspitia, S. Ailawadhi, and R. A. Mesa, "Alternative Outpatient Chemotherapy Scheduling Method to Improve Patient Service Quality and Nurse Satisfaction," vol. 14, no. 2, pp. e82–e91.

[32] T. I. P. Soh, Y. S. Tan, Z. Hairom, M. Ibrahim, Y. Yao, Y. P. Wong, S. W. Lim, S. E. Lim, B. C. Goh, and C. S. Tan, "Improving Wait Times for Elective Chemotherapy Through Pre-Preparation: A Quality-Improvement Project at the National University Cancer Institute of Singapore," vol. 11, no. 1, pp. e89–e94.

[33] G. Dobson, D. Tilson, and V. Tilson, "Optimizing the timing and number of batches for compounded sterile products in an in-hospital pharmacy," vol. 76, pp. 53–62.

[34] A. C. Svirsko, "Multiobjective and Robust Optimization in Pharmacy Delivery and Emergency Department Nurse Staffing."

[35] K. N. Ray, A. V. Chari, J. Engberg, M. Bertolet, and A. Mehrotra, "Opportunity costs of ambulatory medical care in the United States.," vol. 21, no. 8, pp. 567–574.

[36] K. R. Yabroff, W. W. Davis, E. B. Lamont, A. Fahey, M. Topor, M. L. Brown, and J. L. Warren, "Patient time costs associated with cancer care," vol. 99, no. 1, pp. 14–23.

[37] U. B. of Labor Statistics, "Occupational Employment and Wages in Detroit-Livonia-Dearborn - May 2014 : Midwest Information Office : U.S. Bureau of Labor Statistics,"

[38] V. Tilson, G. Dobson, C. E. Haas, and D. Tilson, "Mathematical Modeling to Reduce Waste of Compounded Sterile Products in Hospital Pharmacies," vol. 49, no. 7, pp. 616–627.

[39] A. Syropoulos, "Mathematics of Multisets," in *Multiset Processing* (C. S. Calude, G. Pun, G. Rozenberg, and A. Salomaa, eds.), vol. 2235, pp. 347–358, Springer Berlin Heidelberg.

[40] E. W. Weisstein, "Multiset."

[41] D. Thara, B. Premasudha, V. R. Ram, and R. Suma, "Impact of big data in healthcare: A survey," in *Contemporary Computing and Informatics (IC3I), 2016 2nd International Conference On*, pp. 729–735, IEEE.

[42] J. Barjis, "Healthcare simulation and its potential areas and future trends," vol. 2, no. 5, pp. 1–6.

[43] K. Katsaliaki and N. Mustafee, "Applications of simulation within the healthcare context," vol. 62, no. 8, pp. 1431–1451.

[44] T. R. Rohleder, P. Lewkonia, D. P. Bischak, P. Duffy, and R. Hendijani, "Using simulation modeling to improve patient flow at an outpatient orthopedic clinic," vol. 14, no. 2, pp. 135–145.

[45] M. D. Seminelli, J. W. Wilson, and B. M. McConnell, "Implementing discrete event simulation to improve optometry clinic operations," in *Proceedings of the 2016 Winter Simulation Conference*, pp. 2157–2168, IEEE Press.

[46] J. C. Woodall, T. Gosselin, A. Boswell, M. Murr, and B. T. Denton, "Improving patient access to chemotherapy treatment at Duke Cancer Institute," vol. 43, no. 5, pp. 449–461.

[47] B. Liang, A. Turkcan, M. E. Ceyhan, and K. Stuart, "Improvement of chemotherapy patient flow and scheduling in an outpatient oncology clinic," vol. 53, no. 24, pp. 7177–7190.

[48] F. F. Baesler and J. A. Seplveda, "Healthcare II: Multi-objective simulation optimization for a cancer treatment center," in *Proceedings of the 33nd Conference on Winter Simulation*, pp. 1405–1411, IEEE Computer Society.

[49] K. Chung and D. Min, "Staffing a service system with appointment-based customer arrivals," vol. 65, no. 10, pp. 1533–1543.

[50] O. Jouini and S. Benjaafar, "Queueing Systems with Appointment-Driven Arrivals, Non-Punctual Customers, and No-Shows," p. 41.

[51] S. Schulze, "U.S. ZIP Codes: Free ZIP code map and zip code lookup."

Appendices

Appendix A

Chapter 2 Appendix

A.1 Region Analysis

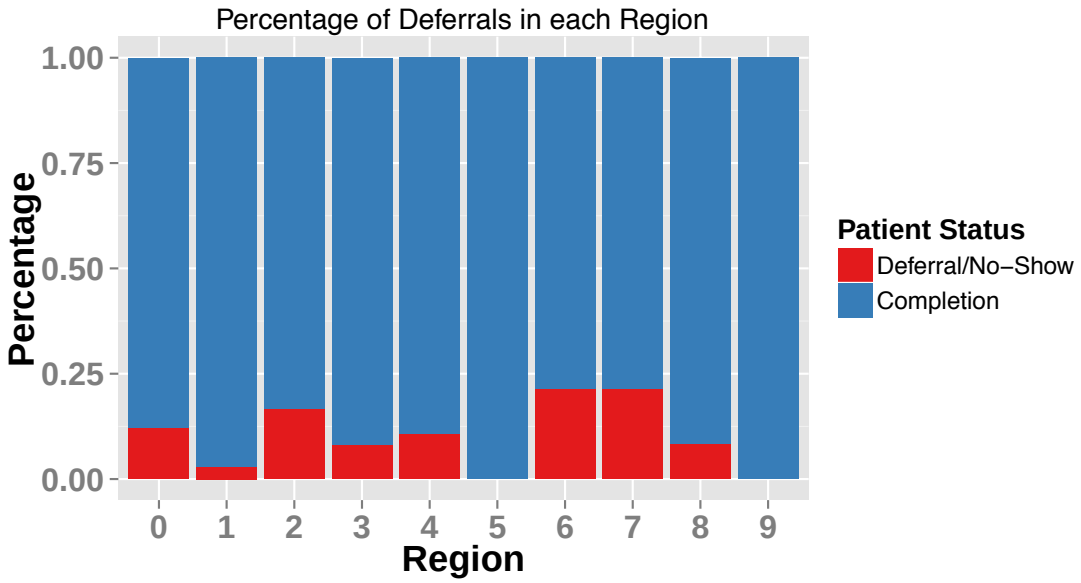


Figure A.1: Here we illustrate the proportion of patients that deferred vs completing their appointment broken down by each region. The total number of patient encounters in each region are as follows: Region 1 – 35 , Region 2 – 30, Region 3 – 98, Region 4 – 27491, Region 5 – 31, Region 6 – 42, Region 7 – 42, Region 8 – 95, Region 9 – 8

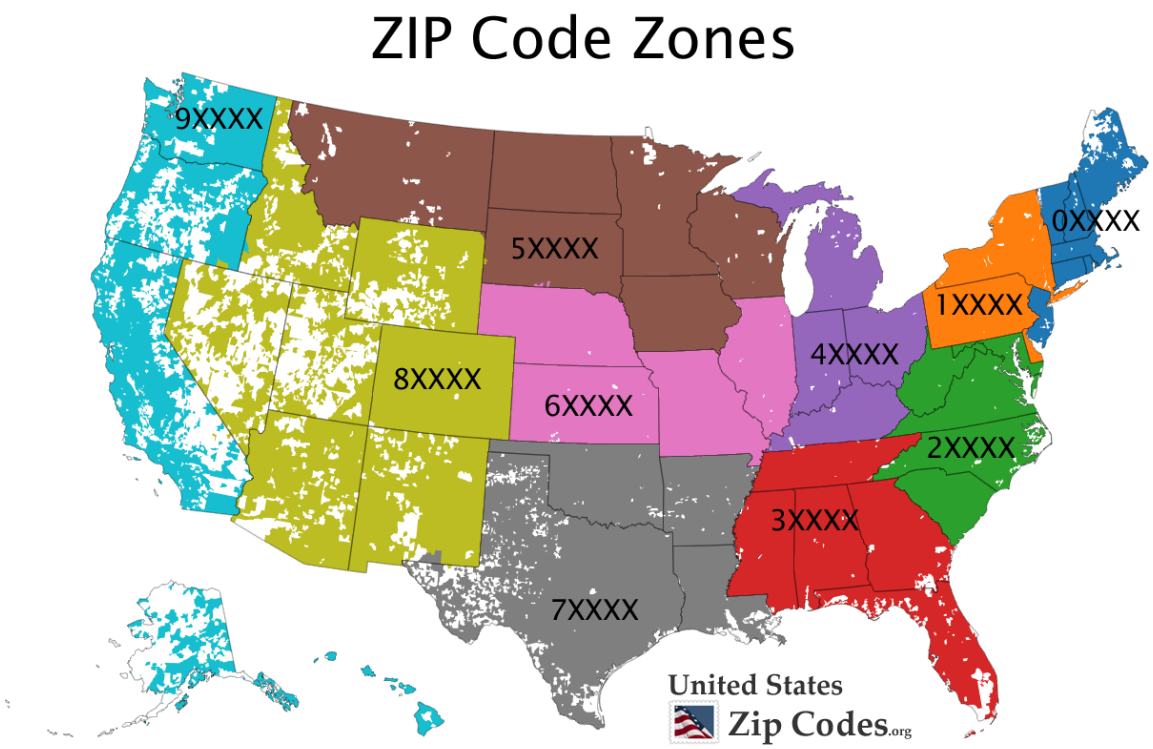


Figure A.2: Here we have a map of the United States displaying the location of each region (i.e., the first digit of a zip code) taken from [51]

A.2 Modeling Process

In Figure A.3, we walk through the predictive modeling process. We note that some of these steps are dependent on the type of response variable (e.g., classification or continuous).

A.2.1 Data Exploration

In healthcare, data exploration can be one of the most lengthy processes. Due to the system complexity, some data is not always correctly formatted or uniform.

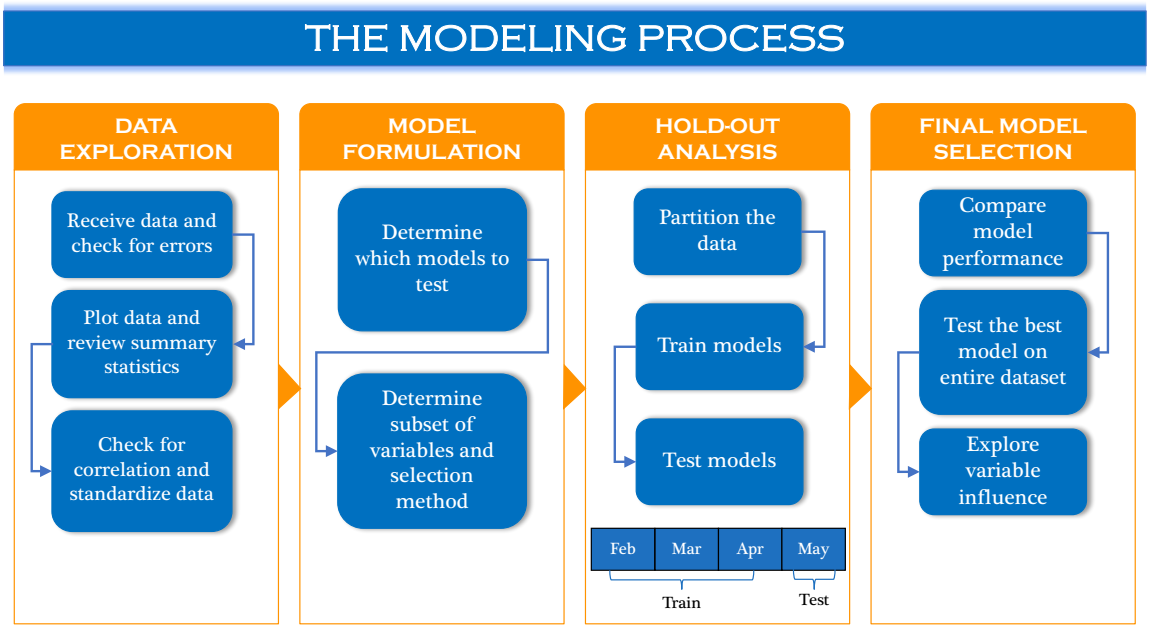


Figure A.3: An overview of the modeling process

Common problems include getting data with missing values, inconsistent naming conventions, or errors occurring from manual data entry. Using tools such as **R** or **Excel** are extremely useful when initially cleaning the data. When working with a large dataset, simply removing the entire entry or row with an error can suffice. In some cases this error may be consistent in a particular variable. Removing the variable completely can help solve the issue but then there is a risk that a significant predictor is left out.

Next, it is always good to plot the data and review the summary statistics. This can be used to help catch any errors. For example, if one of the predictors is scheduled appointment length for infusion and the observed mean is 5 minutes, then something may have gone wrong while collecting/cleaning the data since appointment lengths

have a much higher average. Plots can also help to better understand the data and the relationships between various predictors.

While not as important as using tree based methods, checking for correlation/collinearity in your data is important when trying to fit a model. For example, if the model includes both the weight and BMI of a patient, some correlation issues may arise since the two predictors are closely related. If using a linear model then this can lead to non-unique beta values. This will make it difficult to make any inferences from the fitted models. The data can also be standardized to help interpret the influence of a particular variable directly from the beta values. This is especially useful if the predictors are on drastically different scales.

A.2.2 Model Formulation

The application and type of response you are looking to predict will help determine which models you should test on your data. In our case we are dealing with a binary classification model. This leads us to explore methods such as logistic regression, classification trees, and other ensemble methods (e.g. Random Forrest and Bagged Classification Trees).

Next, you can select your variables. If you have a small number of predictors, it is best to group various subsets and test them all in the hold-out analysis. This way you ensure yourself to find the optimal variable set. However, most problems have a large number of predictors which lead us to conduct variable selection directly in the hold-out since it is not practical to test all subsets of variables. Depending on the model, R has many packages to help with this. Either utilizing variable importance

or simply performing a step wise selection.

A.2.3 Hold-out Analysis

If your goal is to select a model with the best predictive performance, then the hold-out design is one of the most important steps. A poorly designed hold-out can lead to a biased model that may perform well predicting on your current data set but poorly on a future data set. First, as done in this paper, you can immediately hold out a portion of the data for final testing. Next, it is good to split your data into a training set and a testing set. All models are then fitted to the training set and use the predictors from the testing set to make a prediction. These predictions are then compared to the actual response from the testing set to determine the prediction error. These two sets should be randomly selected to prevent a false sense of high predictive performance. It is typical to run this holdout at least 30 times to be able to measure the variability in the prediction error for various models.

When dealing with temporal data, you must take care in making sure you do not train a model on future events and use it to predict past events. For example, training on a patients 10th visit to predict their 3rd visit could result in a false sense of superior predictive performance when in reality this would never happen. To overcome this with our patient appointment data, we created smaller training and testing sets. The model was then trained on a random 90 consecutive day period and tested on the following 30 day period.

A.2.4 Final Model Selection

Finally, you must compare the prediction error and model fitting error for all tested models and subsets of variables. When comparing multiple models, it is also important to validate any significant differences between the predictor errors by using measures such as the Bonferroni pairwise t-test. Once we select a final model, we then can test it on the initially held out dataset to determine how well our model would perform in practice. It is also good practice to explore the influence of each predictor on the response (i.e. generate partial dependence plots).

Appendix B

Chapter 3 Appendix

Drug Name	Hang-By
Abatacept	12 hours
Ado-trastuzumab emtansine	Immediate
Agalsidase beta	12 hours
Aldesleukin	12 hours
Alemtuzumab	6 hours
Alemtuzumab	4 hours
Alglucerase	12 hours
Alglucosidase alfa	12 hours
Alpha1-Proteinase Inhibitor (human)	Immediate
Alpha1-Proteinase Inhibitor (human)	Immediate
Antithymocyte globulin (Equine)	12 hours
Arsenic	12 hours
Asparaginase Erwinia chrysanthemi	3 hours
Atezolizumab	4 hours

Azacitadine	Immediate
Bacillus Calmette-Guerin	Immediate
Belatacept	3 hours
Belimumab	7 hours
Belinostat	12 hours
Bendamustine	5 hours
Bevacizumab	7 hours
Bleomycin	12 hours
Bortezomib	7 hours
Brentuximab Vedotin	12 hours
Busulfan	6 hours
Cabazitaxel	7 hours
Carboplatin	7 hours
Carfilzomib	3 hours
Carmustine	6 hours
Certolizumab pegol	Immediate
Cetuximab	6 hours
Cisplatin	12 hours
Cladribine	12 hours
Clofarabine	12 hours
Cyclophosphamide	12 hours
Cytarabine	12 hours
Cytarabine Intrathecal	7 hours
Cytarabine Liposome Intrathecal	3 hours
Cytomegalovirus Immune Globulin	6 hours
Dacarbazine	7 hours
Dactinomycin	12 hours
Daratumumab	8 hours

Daunorubicin	12 hours
Daunorubicin citrate (liposomal)	Immediate
Decitabine	Immediate
Degarelix	Immediate
Denileukin diftitox	5 hours
Dexrazoxane	Immediate
Docetaxel	3 hours
Doxorubicin	12 hours
Doxorubicin liposomal	12 hours
Eculizumab	12 hours
Elotuzumab	5 hours
Epirubicin	12 hours
Eribulin mesylate	3 hours
Etoposide	12 hours
Floxuridine	12 hours
Fludarabine	12 hours
5-Fluorouracil	12 hours
Fosaprepitant	12 hours
Gemcitabine	12 hours
Golimumab	3 hours
Hemin	3 hours
Idarubicin	12 hours
Ifosfamide	12 hours
Imiglucerase	Ref: 12 hours
Immune Globulin	12 hours
Immune Globulin 10% liquid	12 hours
Immune Globulin	12 hours
Infliximab	12 hours

Ipilimumab	12 hours
Irinotecan	12 hours
Ixabepilone	3 hours
Laronidase	1 hour
Leucovorin calcium	12 hours
Levoleucovorin	11 hours
Mechlorethamine	Immediate
Melphalan	3 hours
Mesna	12 hours
Methotrexate	12 hours
Methotrexate Intrathecal	7 hours
Methylene Blue	12 hours
Mitomycin	12 hours
Mitoxantrone	12 hours
Natalizumab	7 hours
Nelarabine	6 hours
Nivolumab	3 hours
Obinutuzumab	12 hours
Ofatumumab	12 hours
Olaratumab	3 hours
Oxaliplatin	4 hours
Paclitaxel, albumin-bound	7 hours
Paclitaxel	12 hours
Pamidronate	12 hours
Panitumumab	4 hours
Pegaspargase	12 hours
Pegloticase	3 hours
Pembrolizumab	3 hours

Pemetrexed	12 hours
Pentamidine	12 hours
Pentostatin	7 hours
Pertuzumab	Immediate
Pralatrexate	12 hours
Ramucirumab	3 hours
Rasburicase	12 hours
Rituximab	12 hours
Romidepsin	12 hours
Romiplostim	12 hours
Sargramostim	12 hours
Sebelipase	Immediate
Streptozocin	11 hours
Talimogene laherparepvec	Immediate
Temozolomide	12 hours
Temsirolimus	5 hours
Teniposide	12 hours
Thiotepa Intrathecal	Immediate
Thiotepa	Immediate
Tocilizumab	12 hours
Topotecan IV	12 hours
Topotecan Intrathecal	3 hours
Trabectedin	6 hours
Trastuzumab IV	12 hours
Trastuzumab Intrathecal	7 hours
Ustekinumab	3 hours
Vedolizumab	3 hours
Velaglucerase alfa	12 hours

Vinblastine	12 hours
Vincristine	12 hours
Vinorelbine	12 hours
Zoledronic acid	12 hours
Ziv-aflibercept	3 hours

Table B.1: List of Chemotherapy drugs used by the UMRCC's Pharmacy and their hang-by and expiration times

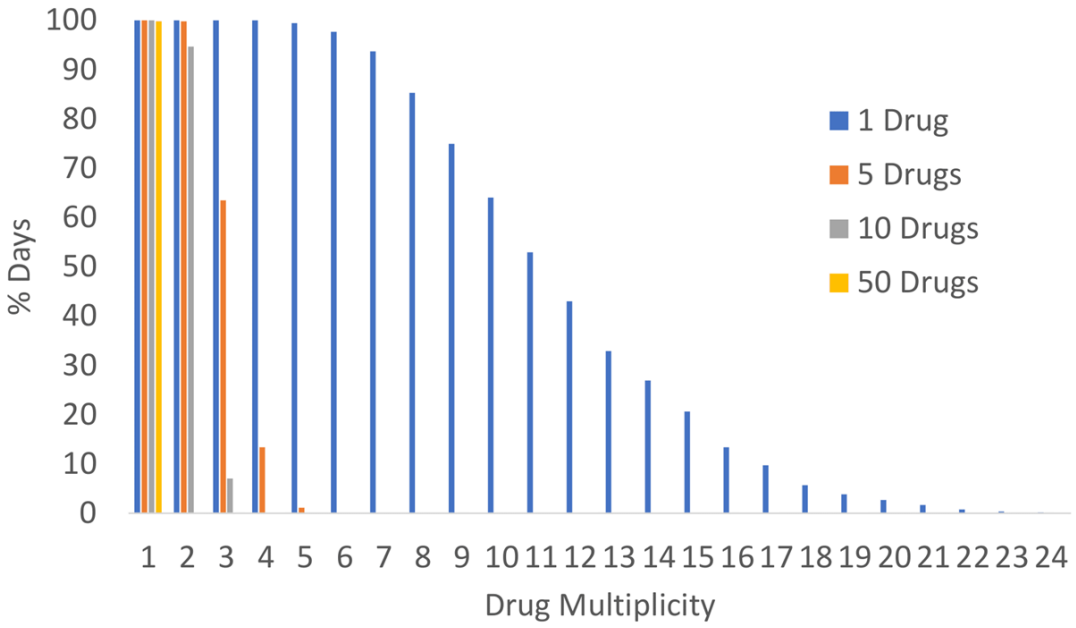


Figure B.1: Drug multiplicity count for all drugs at specific doses that go through the UMRCC pharmacy between March 2015 and March 2017.

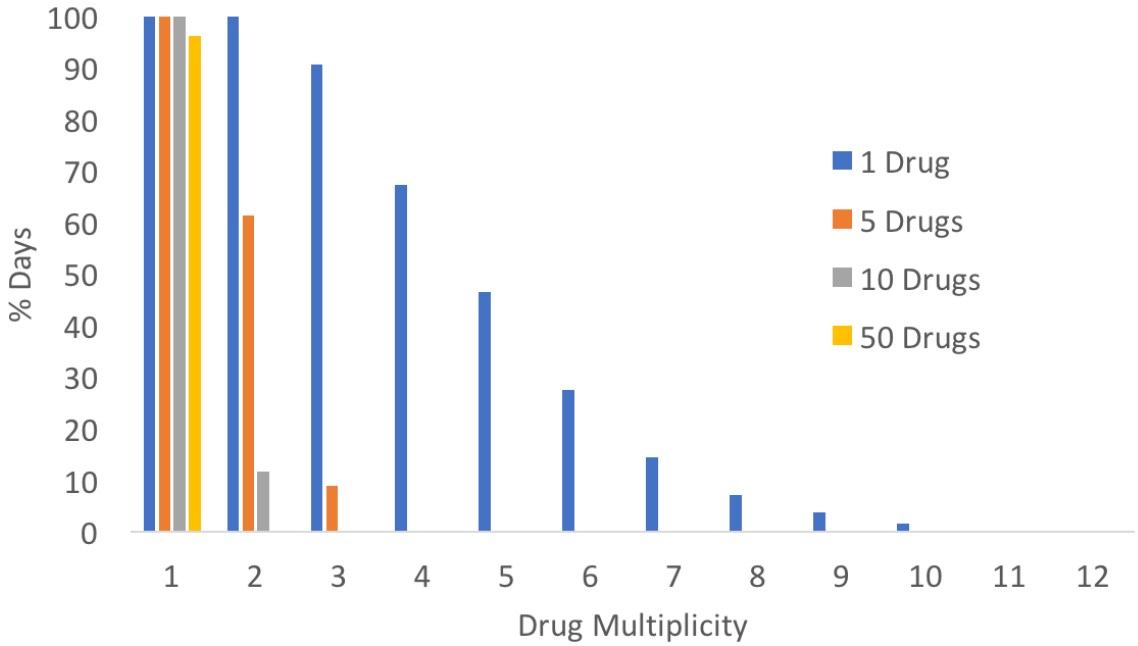


Figure B.**2:** Drug multiplicity count for all drugs classified as "high hazard" at specific doses that go through the UMRCC pharmacy between March 2015 and March 2017.

Appendix C

Chapter 4 Appendix

C.1 Simulation logic

This simulation consists of ten queues:

- The order arrival queue (Aq) which contains one entry per patient pharmacy order and is sorted by actual arrival time of the order into the pharmacy queue
- The order verification queue (Vq) is initially empty but later contains all orders waiting for verification. These orders are sorted by scheduled appointment time
- The Available-Pharmacists-to-Verify Queue (VP_q) contains one entry for each pharmacist available for the verification task, sorted by time that they become available (i.e. when their shift starts or when they have completed a task)
- The print labels queue (L_q) follows a First in First Out (FIFO) policy. This queue contains an entry for each order after it finishes two verifications

- The Available-Tech-to-Print Queue (PT_q) contains one entry for the single print tech working during that shift
- The mixing drug queue (Mq) follows a FIFO policy. This queue contains an entry for each order after finishing printing the labels
- The Available-Tech-to-Mix-Drug Queue (MT_q) contains one entry for each mixing tech available for this task, sorted by the time they become available
- The safety check and sorting queue (Sq) also follows a FIFO policy. This queue contains an entry for each order after drug mixing is complete
- The Available-Pharmacist-to-Safety-Check- Queue (PS_q) contains one entry for contains one entry for each safety check pharmacist tech available for this task, sorted by the time they become available
- The Event queue is used to sequentially order and trace all events in the simulation

These queues are initialized as follows:

- For each drug order, we generate the actual arrival time from a perturbation of the scheduled arrival time (i.e. some patients arrive before and some after the scheduled appointment time). Each order is then placed into A_q,
- All order with an actual arrival time before the start-of-operations are placed into V_q by order of their appointment time

- VP_q, PT_q, MT_q, and PS_q are then generated based on the staff schedule for all pharmacy techs and pharmacist.
- If a drug is pre-mixed, skip to step 6 below

Next we continue to loop through the following steps until we work through the entire arrival queue:

1. Any drug order with an actual arrival time at or before the simulation clock time will then move into the verification queue.

   This queue is then sorted by scheduled appointment time

2. Verification pharmacist then take the order from the top of the queue to complete service. After the first verification is complete, the order is then placed back in the verification queue with a higher priority than the orders who have not received a first verification yet.

3. Next the order must go through a second verification done by a different pharmacist than their first verification. After the second verification, the drug order is then sent to the print labels/drug kit queue.

4. Print techs than take orders in their queue at a first come first serve basis. After service is complete, the order along with all supplies are placed in the drug mixing queue.

5. Mixing tech again grab the orders by FIFO and start working on the drug. There is a chance the tech makes a mistake which is caught by one of the

safety checks. If this is the case, the tech must re mix the entire drug. After mixing service is successfully completed. The drug is placed into the sort/safety check queue

6. Another group of pharmacist will pull from the sort/safety check queue by FIFO. Once this service is complete the drug is ready for the patient.

Next we present the additional staff utilization tables

C.2 Pharmacy Staff Utilization Tables

Utilization Day 1

Resource	Scenarios											
	No Pre-mix			1			2			3		
	Avg.	CI-	CI+	Avg.	CI-	CI+	Avg.	CI-	CI+	Avg.	CI-	CI+
Verification Pharm	0.65	0.61	0.69	0.54	0.51	0.58	0.54	0.52	0.57	0.54	0.52	0.55
Print Tech	0.52	0.49	0.54	0.43	0.41	0.45	0.43	0.41	0.45	0.43	0.40	0.46
Mix Tech	0.21	0.20	0.21	0.17	0.16	0.18	0.17	0.16	0.17	0.16	0.16	0.17
Saftey Pharm	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table C.1: Utilization comparison between policies on day 1 of our simulation as percentages

Utilization Day 3

Resource	Scenarios											
	No Pre-mix			1			2			3		
	Avg.	CI-	CI+	Avg.	CI-	CI+	Avg.	CI-	CI+	Avg.	CI-	CI+
Verification Pharm	0.87	0.83	0.90	0.76	0.72	0.80	0.75	0.70	0.80	0.75	0.72	0.79
Print Tech	0.69	0.66	0.72	0.60	0.57	0.63	0.60	0.57	0.63	0.60	0.58	0.63
Mix Tech	0.28	0.27	0.29	0.25	0.23	0.26	0.23	0.23	0.24	0.24	0.23	0.25
Saftey Pharm	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table C.2: Utilization comparison between policies on day 3 of our simulation as percentages

Utilization Day 4

Resource	Scenarios											
	No Pre-mix			1			2			3		
	Avg.	CI-	CI+	Avg.	CI-	CI+	Avg.	CI-	CI+	Avg.	CI-	CI+
Verification Pharm	0.56	0.51	0.61	0.46	0.42	0.51	0.46	0.42	0.51	0.46	0.41	0.51
Print Tech	0.41	0.34	0.47	0.34	0.28	0.39	0.33	0.28	0.39	0.33	0.28	0.38
Mix Tech	0.19	0.18	0.20	0.16	0.15	0.16	0.15	0.14	0.16	0.15	0.14	0.15
Saftey Pharm	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01

Table C.3: Utilization comparison between policies on day 4 of our simulation as percentages

Utilization Day 5

Resource	Scenarios											
	No Pre-mix			1			2			3		
	Avg.	CI-	CI+	Avg.	CI-	CI+	Avg.	CI-	CI+	Avg.	CI-	CI+
Verification Pharm	0.69	0.64	0.73	0.58	0.55	0.62	0.58	0.55	0.62	0.58	0.54	0.62
Print Tech	0.55	0.52	0.58	0.47	0.44	0.50	0.46	0.44	0.49	0.47	0.45	0.49
Mix Tech	0.22	0.22	0.23	0.20	0.19	0.21	0.19	0.18	0.20	0.19	0.19	0.20
Saftey Pharm	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table C.4: Utilization comparison between policies on day 5 of our simulation as percentages