2013

Scheduling in Mohs Micrographic Surgery Clinics

Stephen Vincent Steidle
Purdue University, ssteidle@gmail.com

Graduate School ETD Form 9
(Revised 12/07)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Stephen V Steidle

Entitled Scheduling in Mohs Micrographic Surgery Clinics

For the degree of Master of Science in Industrial Engineering

Is approved by the final examining committee:

Seokcheon Lee
Chair

Mark Lawley

Yuehwern Yih

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): Seokcheon Lee

Mark Lawley

Approved by: Ahbi Deshmukh 3/19/2013
Head of the Graduate Program Date

SCHEDULING IN MOHS MICROGRAPHIC SURGERY CLINICS

A Thesis

Submitted to the Faculty

of

Purdue University

by

Stephen V. Steidle

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science in Industrial Engineering

May 2013

Purdue University

West Lafayette, Indiana

To all those who have helped my get where I am today, including my professors, friends, and family.

ACKNOWLEDGMENTS

I would first like to thank my major professors, Dr. Mark Lawley and Dr. Seokcheon Lee for their involvement in this project. Their guidance and support have been invaluable.

I would also like to thank Haolin Feng for his direction and help in developing and executing the sequential scheduling policy and its associated derivations.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

MMS	Mohs Micrographic Surgery
NMSC	Nonmelanoma skin cancer
BCC	Basal cell carcinoma
SCC	Squamous cell carcinoma
k	Number of patients scheduled
p_{ns}	Patient no-show probability
p_r	Patient re-entrant probability
β	Patient re-entrant probability shape coeffecient
r	Revenue from single service
c_w	Hourly waiting cost
c_O	Hourly overtime cost
S_i	Patient schedule for slot i
X_i	Number of scheduled patients who arrive for slot i
Z_i	Number of patients who re-enter the system in slot i following treatment in slot $i-1$
L_i	Number of patients treated in slot i provided the queue is long enough
Y_i	Number of patients who overflow from slot i into slot $i-1$
T_i	Number of patients treated in slot i

ABSTRACT

Steidle, Stephen V. M.S.I.E., Purdue University, May 2013. Scheduling in Mohs Micrographic Surgery Clinics. Major Professors: Mark Lawley and Seokcheon Lee.

Mohs Micrographic Surgery (MMS) is a surgical method used for the excision of aggressive skin cancers in areas of high cosmetic importance, such as the face and hands. The practice has been gaining popularity worldwide for its low recurrence rates and cosmetic results. Current clinics though are plagued by extreme wait times and an overall poor patient experience. In this paper we look to explore this problem by applying systems engineering principles including optimization and scheduling with the goal of improving the patient experience. Currently, little literature exists exploring the difficulties associated with scheduling for MMS clinics which primarily revolve around patient recirculation for an unknown number of repetitions with little predictive ability. By developing a simulation model depicting current clinic operations, we have explored the current practice of clinics through several important performance measures while being able to determine an optimal number of patients to be scheduled. We have also explored the impact of changing re-entrant probability on the nature of the patient schedules. We have developed a set of qualitative scheduling constraints for on-the-fly physician application and a sequential scheduling policy to produce myopically optimal patient schedules for maximizing the patient experience.

1. INTRODUCTION

Skin cancer affects millions of Americans each year, with over 3.5 million cases being diagnosed in at least 2 million patients every year [1]. Of these occurrences, non-melanoma skin cancers (NMSC) are the most significant, including basal cell carcinoma (BCC) and squamous cell carcinoma (SCC), each accounting for an estimated 2.8 million and 700,000 cases a year, respectively [2]. Mohs Micrographic Surgery (MMS) is a technique used in the treatment of NMSCs. MMS has several advantages over other treatments, including being the leader in cost-effectiveness [3], low recurrence rates [4], and less permanent cosmetic damage [5]. However, one of the well-known disadvantages to MMS is long time the procedure can take to complete, often resulting in patients spending entire days at the clinic [7].

In a standard MMS clinic setting, the physician begins by identifying the visible margins of the tumor in question. After identification, the physician applies local anesthesia and makes his first incision, cutting at between 30° and 45°. The sample is then sent to a pathologist for analysis. Typically when examining the tissue margins for signs of cancer, pathologists use a method known as the bread-loaf technique. This method reviews less than 1% of the overall margins [6]. However, in MMS nearly 100% of the margin is examined, enabling the pathologist to identify clear margins much more reliably [7]. However, the trade off for 100% margin clearance, as previously mentioned, is the long and tedious process that is MMS [8]. Usually, multiple excisions are required to achieve clear margins. The typical method of patient flow through this system is as follows. A patient sees the physician for initial tumor identification and first round of excision. Following the first excision, the patient returns to the waiting room to wait for the pathology report to be completed. The returning pathology report will have one of two possible results: clear or unclear margins. Clear margins

mean that the margins of the excised area are cancer free and the patient requires no more excisions. Unclear margins indicate that the margins still contained cancer cells and addition excision is required. If the patient has clear margins, then he or she returns to the physician for wound repair. This rotation through physician and pathology service continues until the patient's margins are determined to be clear and wound repair has been conducted. This process is demonstrated in Figure 1.1 (Reprinted with permission. Copyright 2013 University of Wisconsin Hospitals and Clinics Authority. [9]).

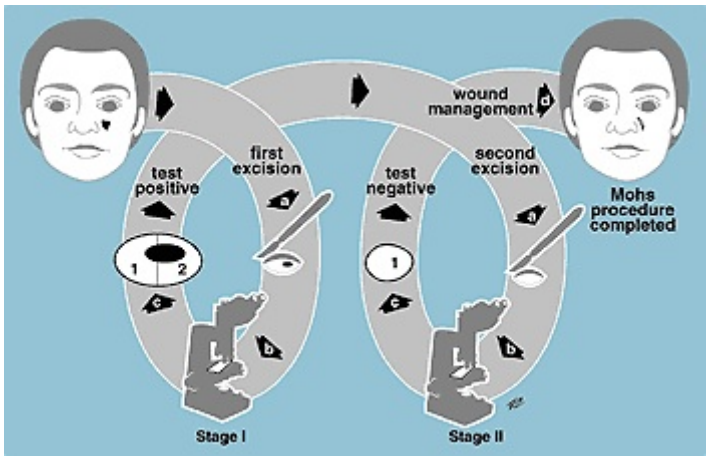

Figure 1.1.: The Mohs Surgery Process [9]

1.1 Objectives

When considering this problem from a scheduling perspective, it is the rotational aspect of the patient's progress through the system that defines our problem. Several examples in literature exist examining scheduling problems featuring recirculation [10, 11], and scheduling in healthcare scenarios is also a very heavily researched area [12,13] . However, what differentiates our problem from other examples in literature is the stochastic nature of the amount of recirculation. When scheduling patients into this system, it is difficult, if not impossible, to know precisely how many repetitions a patient will require to complete treatment. This detrimental characteristic

of our system can lead to poorly structured schedules causing excessive bottlenecking and patient wait time. Patients waiting to see the physician a second time are delayed due to the availability of the pathologist or the physician is seeing another patient. We hope to develop a multi-faceted approach to this problem. With the little attention paid to MMS in scheduling and simulation literature, our first objective is to lay a foundation by developing a simulation model of an MMS clinic. The development and analysis of this model should lead to an increased understanding of how this system operates. Our second objective is to develop a system of scheduling constraints and heuristics based on patient characteristics that would enable a physician or individual responsible for scheduling able to make informed decisions about what types of patients to schedule. Finally, our last objective is to develop an algorithmic scheduling policy based on an objective function that takes into account not only the clinic's costs and revenues, but also the patient experience. By combining these three objectives together, we can fulfill our overall goal of providing tools to improve clinic operations while simultaneously improving the patient experience.

1.2 Organization

Chapter 1 outlines what MMS is and how it is used, as well as the objectives of this research and why it is important to MMS clinics.

Chapter 2 presents a literature review of research into MMS as well as important peripheral papers. It begins with a review of non-melanoma skin cancers and their risk factors and symptoms. Next is a review of MMS and its implications. This is followed by an analysis of the current literature directly related to MMS. Finally, a review of relevant clinical scheduling policies is conducted.

Chapter 3 presents a simulation model of an MMS clinic. All relevant parameters and aspects of the clinic are discussed. Additionally, an objective function is developed for

optimizing clinic function in relation to operations and the patient experience. This function is then analyzed and its behaviors under various conditions are discussed.

Chapter 4 delineates a method of generating random patients for scheduling in a clinic and presents several sets of constraints and heuristics that could be used in an MMS clinic. Each of these is identified in detail and explained. These heuristics are then analyzed using a simulation model similar to that in the previous chapter. The results of these simulation runs are then presented and analyzed.

Chapter 5 describes a sequential scheduling policy as an alternative to the heuristic method presented in the previous chapter. An objective function similar to that of Chapter 3 is utilized to schedule patients according to the policy presented. Initial experimental results of this model are presented along side an analysis of its behavior.

Chapter 6 discusses the overall conclusions of this research as well as describing the important contributions made by it. Future areas of research are also presented.

2. LITERATURE REVIEW

2.1 Overview of Non-melanoma Skin Cancers

Skin cancer is defined as a cluster of abnormal skin cells that grow uncontrollably. Skin cancers are caused when DNA and RNA in skin are damaged by exposure to ultraviolet radiation and these damages go unrepaired. These cancers can be broadly grouped into two categories: melanoma and non-melanoma. Non-melanoma skin cancers (NMSCs) include two primary subcategories, namely basal cell carcinomas and squamous cell carcinomas. An additional wide variety of subtypes exists under the umbrella of NMSC, but it is typically used to identify those two primary subtypes. The various types of carcinomas are typically identified by the layer of the epidermis in which they arise. Basal cell carcinomas are identified by the fact that they are found in the topmost level of the skin, the epidermis, and squamous cell carcinomas begin in the middle layer.

2.1.1 Risk Factors

The primary cause for all skin cancers is identified as over-exposure to ultraviolet (UV) radiation. This is typically encountered through prolonged exposure to sunlight or tanning beds. Additionally though, there are a combination of personal and environmental risk factors that can increase an individuals susceptibility to NMSC. The major personal factor that puts individuals at an increased risk for NMSC is a susceptibility to UV radiation. Characteristics indicative of this typically include light skin, hair or eye color [14]. Individuals who possess these characteristics and who are over-exposed to UV radiation have an increased risk of developing NMSC at some point during their lives. Environmental risk factors are primarily linked to

geographic location. Areas nearer to the equator which typically have thinner ozones and therefore increased penetration by UV radiation typically see increased rates of NMSC than areas further from the equator [15].

2.1.2 Symptoms

Basal cell carcinomas and squamous cell carcinomas vary in their physical presentations. Basal cell carcinoma presentations are typically divided into the following clinical subtypes, but often include a mix of multiple [16].

- Classic rodent ulcer - a lesion on the surface of the skin with necrotic tissue in the center
- Nodular or cystic - a small, translucent nodule on the skin through which exposed blood vessels can be seen
- Superficial - an erythematous patch of skin, often difficult to differentiate from other conditions such as eczema
- Morphoeic - a waxy scar made up of white sclerotic plaque
- Pigmented basal cell carcinoma - similar in appearance to a nodular or cystic BCC, but with dramatically increased melanin levels

Squamous cell carcinomas are not as varied as basal cell carcinomas, but unlike BCCs, they present with precursor lesions. As such, these precursor lesions are used to identify an SCC. There are three primary types of precursor lesions with squamous cell carcinomas [16].

- Actinic keratosis - a flat, crusty or scaled area that develops into a wart-like surface
- Squamous-cell carcinoma in situ (also known as Bowen's disease) - a clearly defined erythematous plaque that increases in size and features a scaled or crusted surface

- Keratoacanthoma - a symmetrical dome-shaped skin inflammation with keratin scales on its peak

While all three of these have been identified as precursor lesions for squamous cell carcinomas, developing any of these does not guarantee a diagnosis of SCC.

2.2 Overview of Mohs Micrographic Surgery

Mohs Micrographic Surgery is a procedure that has been used for the past 70 years to treat skin cancer. Treatment methods started in 1938 by Frederic E. Mohs with the utilization of a zinc chloride paste to fix the tumor in situ, so that as each layer of the tumor was excised, the margins could be evaluated. However, this fixed method had substantial drawbacks, including only being able to perform one layer of excision per day and necrosis caused by the zinc chloride paste. In 1953, Mohs performed several layers without the utilization of the chloride paste and saw success in freezing the horizontal layers. This began the modern fresh tissue technique, which enabled the surgeon to perform multiple layers as well as wound repair in a single day [17]. Initially, only BCC and SCC were treated using Mohs surgery. However, in the last 40 years, this procedure has gained significantly in not only popularity but also application [5]. Beyond simply BCC and SCC, today MMS is used to treat other tumor types such as verrucous carcinoma, extramammarry Paget disease, and microcystic adnexal carcinoma [17].

The reasoning behind the utilization of MMS is that in most physical presentations of skin cancer, the tumors grow contiguously but unpredictably. The external presentation and margins of the tumor may have little to do with the actual subcutaneous spread of the tumor. Typically, long finger-like extensions grow outward from the central mass of the tumor. In order to avoid recurrence, which can often be attributed to incomplete excision [18], Mohs attempts to completely excise all cancer cells. MMS is able to identify these extensions though the examination of horizontal cross sections removed layer by layer by the physician. It is

this type of analysis, in contrast to the standard "bread-loaf" technique used in traditional tumor excision methodologies, that allows such accurate identification of the tumors actual margins [7]. The difference between these two techniques is demonstrated in Figures 2.1 and 2.2 (Images reprinted with permission from Medscape Reference (`http://emedicine.medscape.com/`), 2013, available at: `http://emedicine.medscape.com/article/1125510-overview`).

It is this technique that allows MMS to feature substantially lower recurrence rates than other types of surgical excisions [4]. As such, MMS "remains the gold standard for the surgical management of basal cell and squamous cell carcinomas" [7]. Not only do these techniques offer much lower recurrence rates, but they also enable the physician to conserve as much tissue as possible during the resection and eventual wound closing. Because the physician is able to accurately identify the location of any residual extensions, nodules, or tumor, areas that are unaffected are able to be left intact [8]. This permits maximum effectiveness while minimizing cosmetic damage. Since a substantial portion of BCC and SCC occur in areas of high cosmetic significance, such as the face and hands, the use of Mohs is even more justified [19].

Much research has also been done into the cost-effectiveness of Mohs surgery [19–23]. The problem with much of this literature is that MMS is performed under various conditions by various types of professionals, leading to difficulty in comparisons. For example, the paper by Bialy only studies MMS procedures performed by otolaryngologic (ENT) surgeons [20] while the analysis by Smeets is from a training hospital [19]. With the wide variety of conditions under which Mohs is performed, a consistent cost analysis can be difficult to obtain. A review published in 2009 reviews all of these and many more in an attempt to determine the cost-effectiveness of Mohs surgery [23]. It is their finding that when coupled with the decreased recurrence rates and decreased cosmetic damage, Mohs is in fact a cost- effective treatment. In fact, it is on average less than the overall cost of traditional excision, especially in light of the recurrence rates for such procedures and the additional cost incurred when undergoing a second procedure.

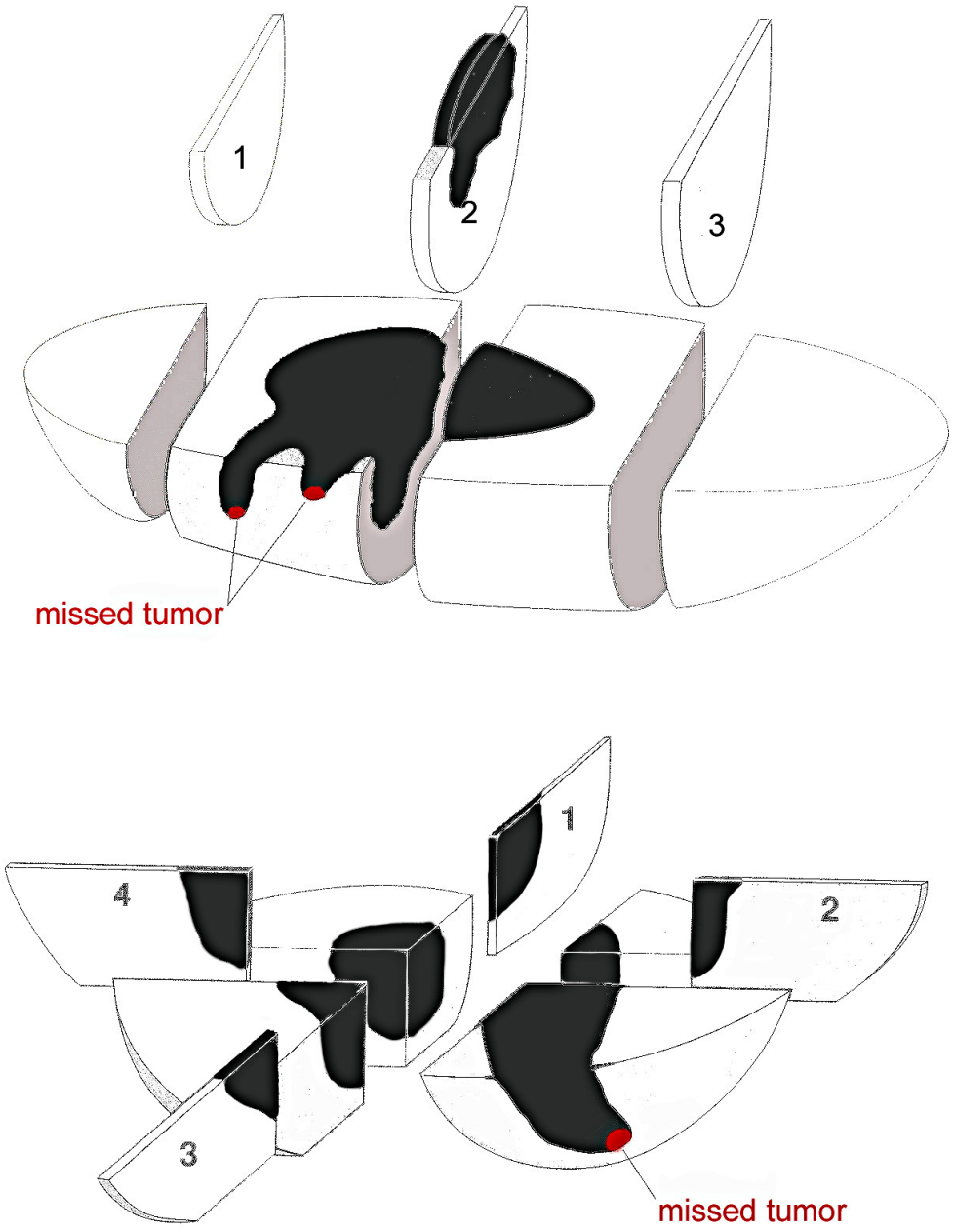

Figure 2.1.: The Bread-Loaf Technique [17]

The most consistently cited drawback to Mohs surgery is the tediousness of the process, for both the patient and the physician [24–27]. The duration of the procedure as well as the repeated excisions can be difficult for a patient. This becomes even more

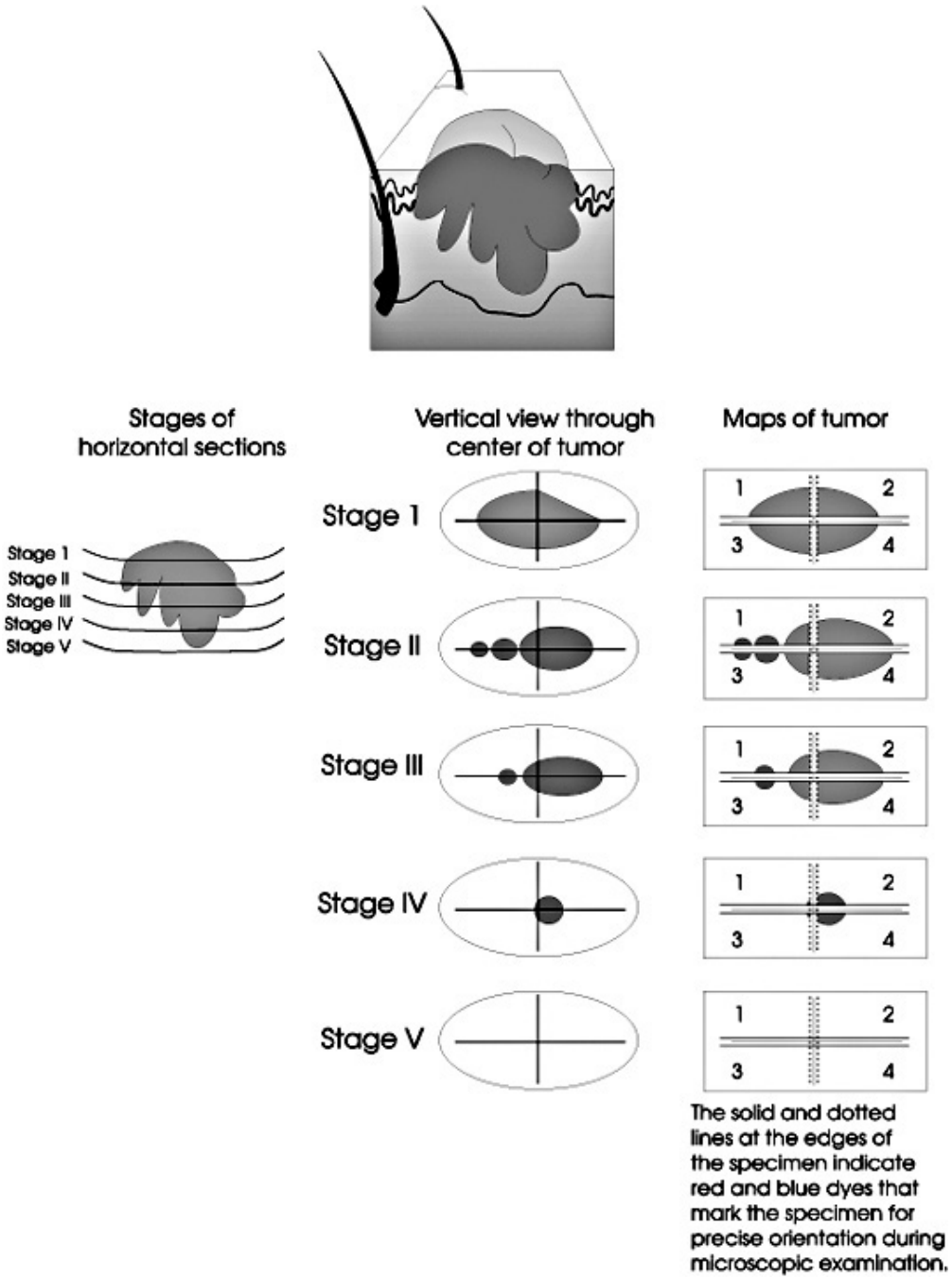

Figure 2.2.: The Mohs Technique [17]

problematic when the patient suffers from a debilitating condition such as dementia, poor sphincter control or spinal arthritis. Some also deem the necessity of removing all cancer cells to avoid recurrence questionable and cite the fact that even with this,

recurrence does still occur in some Mohs cases. [4, 26]. While the procedure may be long and difficult, most physicians still identify it as the best treatment option available for patients [5, 23].

In summation, Mohs Micrographic Surgery has been identified as the best treatment of choice when considering basal cell and squamous cell carcinomas. This procedure has dramatically decreased recurrence rates while enabling the physician to conserve as much tissue as possible. This minimizes the cosmetic harm caused by the procedure, which is exceptionally valuable since many BCC and SCC occur on the face. Not only is this the best treatment option available, but it has also been found to not only be cost-effective but also less costly than traditional excision methods. The primary drawback seen in Mohs is the tediousness and the length of the procedure which can be burdensome on the patient.

2.3 Current Mohs Micrographic Surgery Literature

Little literature exists today studying the operational aspects of Mohs Micrographic Surgery. Literature that does exist about Mohs is more focused on cost-effectiveness, recurrence rates, and improvements to the procedure. Considering the interesting nature of the operation of an MMS clinic, we find this lack of literature surprising. To improve the current status of MMS clinics, we began by reviewing any mention of Mohs, heuristics, or scheduling and only a single paper was found, *The Webb and Rivera (WAR) Score* (from now on referred to as *WAR Score*) [28]. The stated goal of this paper is to obtain an easily used preoperative tool that would enable a physician to ascertain the difficulty of and time required to complete an MMS case. The data for this study was obtained through a questionnaire which inquired about four preoperative characteristics (original lesion size biopsied, recurrent or not, located on nose, eyelid, ear, or lip, and whether or not it is of an aggressive subtype) and two post operative questions (number of stages required and time from first cut to final suture or repair). Based on the answers to these questions, a statistical

analysis was completed to identify which characteristics were statistically significant predictors of both the number layers and amount of time required to complete the procedure. These predictors were then ranked and scored based on how significant each was. Based on these rankings and scores, the WAR score was developed, with higher scores being associated with more complex and longer duration surgeries. The researchers found that the WAR score was significantly correlated to both the number of layers required and the overall surgery time. The most difficult aspect of this problem was the lack of ability to accurately predict the difficulty of the procedure. As mentioned previously, carcinomas such as those treated with Mohs surgery often feature extensions that can infiltrate far beyond the physical presentation of the tumor. As such, the number of layers required can be very difficult to ascertain [29].

The need for a method of predicting the amount of subcutaneous spread of the neoplasm is obvious. Two studies were found examining this and attempting to develop methods to accomplish it. The primary source of literature in the area comes from two sources by one individual, R. Sonia Batra, MD. Her two papers in this area, *Predictors of extensive subclinical spread in nonmelanoma skin cancer treated with Mohs micrographic surgery* (from now on referred to as *Predictors*) [30] and *A Risk Scale for Predicting Extensive Subclinical Spread of Nonmelanoma Skin Cancer* (from now on referred to as *Risk Scale*) [31] form the basis of the patient distribution data used in this paper. *Predictors* [30] focuses primarily on identifying which patient characteristics are indicative of a patient having extensive subclinical spread. Batra uses five categories of characteristics: age, sex, carcinoma location, carcinoma size, and carcinoma classification. A multivariate analysis is then conducted to identify the most significant predictors, including single characteristics and combinations of them. Her paper *Risk Scale* [31] conducts a similar experiment but offers an additional point system for determining what level of risk a patient is at for extensive subclinical spread. Using the predictors established in the previous paper, Batra assigns each one a point value based on how strong of an indicator of extensive subclinical spread it is. One characteristic about both of these papers to note is that Batra cites as a primary

weakness of both papers the fact that patients used in these studies were not general NMSC patients, but ones whose carcinomas has been identified as particularly suited to MMS. This actually causes the data used to be more suited to our uses, as we are not considering the general population of NMSC sufferers, but only those who receive MMS.

2.4 Scheduling for Outpatient Clinics

In order to tackle the problem of scheduling the clinics, a wide variety of healthcare scheduling literature was reviewed. Gupta [13] and Cayirli's [12] reviews provided excellent foundations for this work, and several examples of healthcare heuristics were found [32, 33] . However, little of the literature existing in the field is strictly applicable to this problem, due to the aforementioned re-entrant problem. Through our initial analysis of these works, it was decided to pursue a specific variety of scheduling procedure: a sequential, myopic policy that accounted for no-show probability. A review of sequential scheduling methodologies yielded several results [34, 35]. Overbooking models also yielded a pair of results [35, 36]. Turkcan's 2011 paper [34] presents a method for developing sequential clinical schedules utilizing specific service criteria, such as maximum waiting and overtime. Muthuraman's 2008 paper [36] delineates of substantial model for overcoming patient no-show through overbooking. A combination of these two can be used to produce a model for our clinics, after the incorporation of the re-entrant problem.

3. SIMULATION AND NUMERICAL ANALYSIS OF CLINIC OPERATIONS

3.1 Simulation Model Development

Our first objective is to model how MMS clinics are currently operated, in order to establish a baseline to which our other proposed solutions can be compared. The model of current practice is based on the following assumptions.

1. Patients arrive individually in 30 minute intervals with k patients scheduled [13].
2. All patients have a no-show probability of $p_{ns} = 0.1$ [36].
3. The probability of re-entrance into the queue, p_r, is a function of the number of repetitions through the queue, n, and a shape coefficient, β of the form

$$p_r(n, \beta) = e^{-\beta(n-1)}. \tag{3.1}$$

4. System has two stages, each with a single server [17].
5. Patients are punctual (arrive at start of appointment) [34].

The $p_r(n, \beta)$ function was defined in this manner to meet certain requirements for the re-entrant probability. These requirements follow.

1. Patients must enter the system at least two times: excision and wound repair [17].
2. The probability of a patient re-entering the queue should decrease with the number of entrances.
3. We must be able to account for varying patient characteristics and physician behaviors.

The first requirement arises simply from the way patients flow in the system. A patient cannot enter the system once and leave. This would be the same as a patient receiving one cut from the physician and leaving. This is guaranteed by the $(n-1)$ expression in the exponent. On the first visit, $p_r(1,\beta) = 1$ regardless of the β value. Our second requirement corresponds to the removal of additional layers. The more times a patient re-enters the system, the more layers they will have removed. As there is a limit amount of layers a physician can reasonably perform, usually between two and six [30], we can assume that with increasing re-entrances, the probability of re-entrance declines. This is confirmed by the negative value of the exponent. Finally, different patients have different needs and different physicians have different practices. In order to account for these, we wish to be able to control the shape of the curve. We do this through β. Now, we can state these mathematically as follows.

1. $p_r(1,\beta) = 1 \ \forall \ \beta$
2. For a constant β, as n increases, $p_r(n,\beta)$ decreases
3. For $n \neq 1$, as β decreases, $p_r(n,\beta)$ increases

The first point is essential because we know that patients in this system must visit the physician at least twice. After the first visit the patient must re-enter at least once. This is guaranteed since the probability of re-entrance after the first visit is 1. The second point indicates that the more a patient re-enters the system, the lower his or her probability of re-entrance becomes. This is reasonable due to the fact that the more layers that are taken from the patient, the lower the probability of another layer being needed is. Our final point describes the function of β. For different types of patients or physician practices, the expected number of layers required per patient changes. For example, in an area with a high prevalence of skin cancer, patients in general may require more layers to completely excise the tumor. Alternatively, if a physician is prone to taking much deeper layers than another physician in his area, his patients may in general require less layers. Since a decrease in β corresponds to an increase in $p_r(n,\beta)$, a patient's probability at a given value of n is higher for a

smaller β. This means that while β decreases, the probability of a patient re-entering the system increases. Some examples of $p_r(n, \beta)$ are displayed in Figure 3.1.

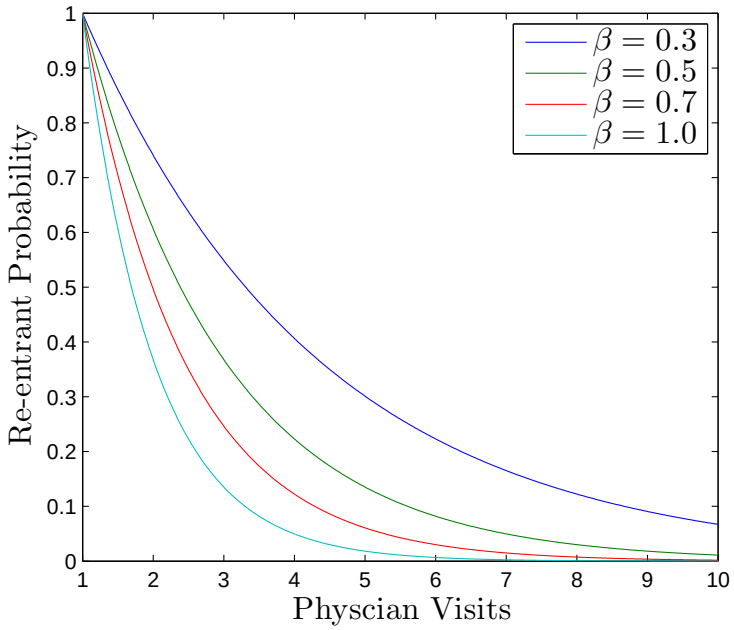

Figure 3.1.: Re-entrant Probability for Various β

Using this set of requirements and assumptions, a simulation model was developed and analyzed for various values of k and β. The service times utilized in the simulation are estimates derived from a wide variety of basic MMS literature, such as patient pamphlets, physician websites, and MMS overviews [37–39]. However, these resources did not provide distributions, but rather a range of maxima and minima. Since the gamma distribution is used to model healthcare service times [12, 35] we decided to approximate a gamma distribution using a triangular distribution. We found various values for service times from these sources and used them to approximate a triangular distribution. A triangular distribution has three parameters, a, b, and c, with a and b as the minimum and maximum values respectively and c as the mode. The parameters we used for the distributions are as follows. For the physician, we used $f_{phys}(a, b, c) = f_{phys}(10, 30, 15)$ and for pathology, we used $f_{path}(a, b, c) = f_{path}(15, 45, 40)$. We were

then able to calculate the variation of the proposed triangular distributions using the mode (c) as an estimate of the mean and the variance of the triangular distribution for the variance of the gamma distribution. This enabled us to calculate the parameters of each gamma distribution, α and β in $\Gamma(\alpha, \beta)$. They took the form of $\Gamma_{phys} = (12.45, 0.83)$ and $\Gamma_{path} = (37.20, 0.93)$. These distributions are seen in Figures 3.2 and 3.3.

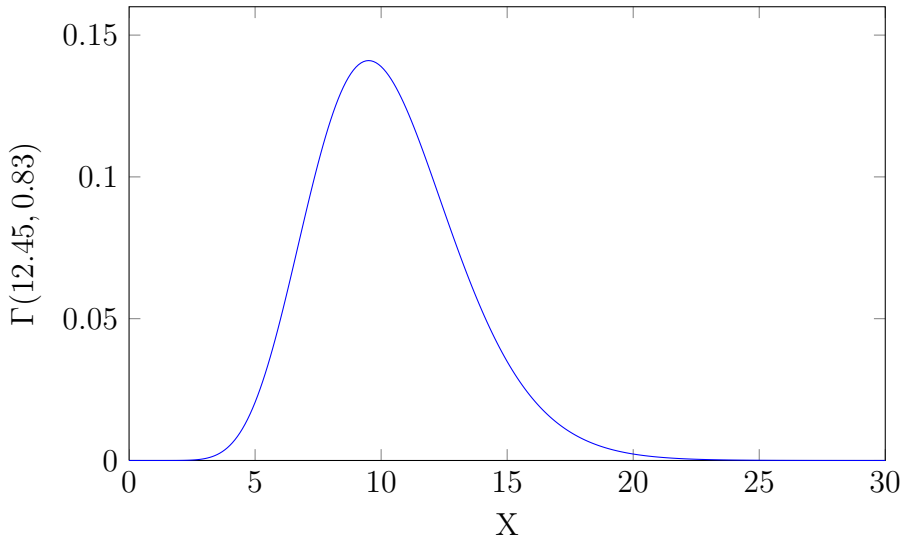

Figure 3.2.: $\Gamma_{phys} = (12.45, 0.83)$

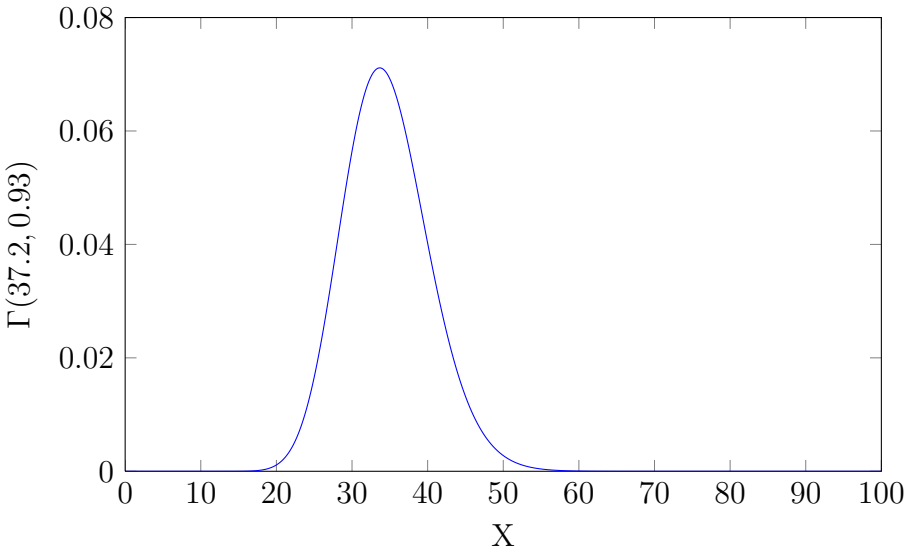

Figure 3.3.: $\Gamma_{path} = (37.20, 0.93)$

3.2 Experimental Conditions

To conduct the analysis, an objective function was established. In our objective function, our goal was to pair revenue with overtime costs while maintaining a perspective on the amount of waiting time each patient experiences. To accomplish this, we applied a wait time disutility, a cost incurred by the clinic for each hour a patient is made to wait. This function is of the form

$$F(S) = (r * V_{phys}) - (c_w * W_{total}) - (c_o * O_t) \tag{3.2}$$

where r is the reward for one layer of MMS or wound repair, c_w is the disutility of patient wait time, and c_o is the cost of clinic overtime. V_{phys} is defined as the total physician visits, W_{total} is the total patient wait time, and O_t is the amount of overtime. Here we consider only the physician visits because these are the sessions that generate revenue, as opposed to the pathology analysis. The following conditions were set for the analysis.

- $r = \$300, c_w = \$100/hr$, and $c_o = \$800/hr$
- $\beta = [1.0, 0.9, 0.8, .., 0.1]$
- Each simulation run corresponds to one day of clinic operation
- Clinic simulation for 100 days
- Patients scheduled to earliest available slots
- Slots are 30 minutes in length
- One patient per slot
- Number of patients scheduled, k, incremented from 1 until objective function value is negative

3.3 Simulation Results

By using the determined levels of r, c_w, c_o, simulation runs were conducted using the range of values of both k and β. β values were varied from 0.1 to 1.0 in increments of 0.1. For each set value of β, we begin by simulating only a single patient being scheduled, $k = 1$. When then increase the value of k for each set of runs until our objective function value is negative. Upon analysis, it was found that the objective functions were unimodal when varying k across a constant β for the values of r, c_w and c_o we selected. This allowed us to determine optimal levels of patient scheduling for each beta. These results are presented in Table 3.1 and Figure 3.4.

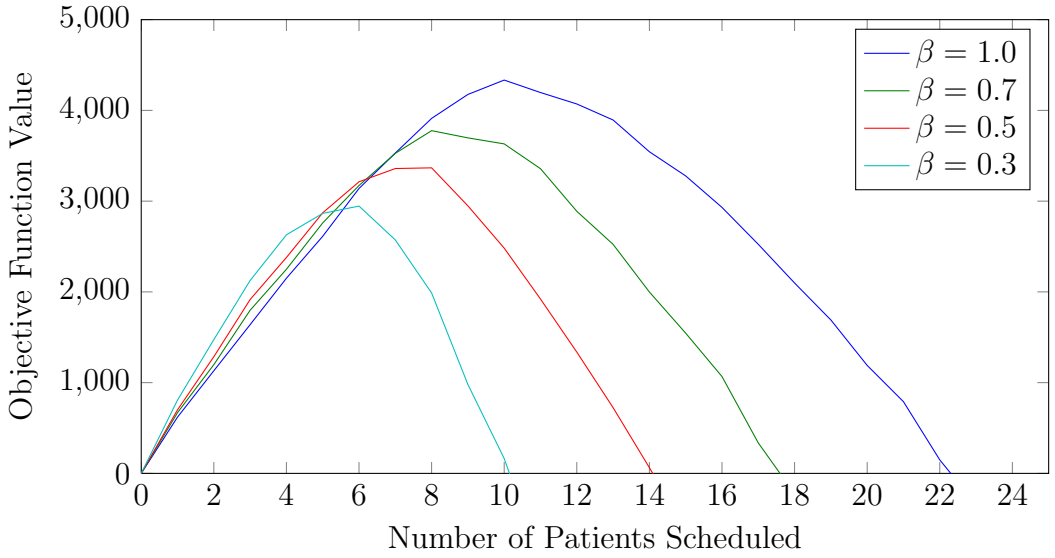

Figure 3.4.: Objective Functions for Various β

This optimal level, k_{opt}, was found to be a function of β with

$$k_{opt} = 3.163 \ln(\beta) + 9.806. \tag{3.3}$$

We should note some of the characteristics of this result. As our β coefficient increases, a patient's expected number of physician visits decreases. This characteristic is validated in these results, as demonstrated in Table 3.1, both the number of patients treated and the maximum value of the objective function increase. When

Table 3.1: Optimal Objective Function Values

β	Max $F(S)$	k_{opt}
0.1	2554.07	3
0.2	2740.57	4
0.3	2954.59	6
0.4	3150.67	7
0.5	3366.92	8
0.6	3635.44	8
0.7	3776.69	8
0.8	3991.83	9
0.9	4245.31	10
1.0	4334.14	10

considering this in the framework of our probability function, this is a reasonable result. Having patients repeat less in our system means less penalty for the disutility of waiting in the clinic, not only enabling more patients to be scheduled, but also adding additional reward without the waiting cost. What we have determined here is that with a brief study of a clinic's operations to determine its parameters including cost-reward structure, service times, and average patient repetitions, we can use this model to determine an optimal number of patients to schedule per day.

3.4 Numerical Analysis of Simulation Results

Our next area of analysis was to examine the behavior of simulated clinic under various combinations of our costs, c_w and c_O and our system parameters β and k. We can begin by exploring β and its implications. We already know that a decrease in β corresponds to an increase in re-entrant probability for a set value of n. However,

what we do not know is what is the expected number of treatments for a specific value of β

In order to calculate the expected number of treatments, we needed to evaluate our probability of re-entrance. When looking at the nature of our system, we can view re-entering the system as a "failure" and exiting the system as a "success". Thinking like this, we can draw a comparison between our probability distribution and the geometric distribution. The geometric distribution as we are referencing it is defined as the probability distribution of the number of Bernoulli trials, X, needed to get one success. In our instance however, we do not have Bernoulli trials with a constant probability of success. Our probability changes with each increment of n. Because of this, we cannot just raise our probability of failure to the $(k-1)$. What we need to do is compute a product from 1 to $k-1$ of our failure probabilities while evaluating our probability of success at k. We begin by looking at these probabilities at a few early points, defining p as the probability of re-entering the queue and $q = (1-p)$ as the probability of exiting the queue. These basic calculations are demonstrated in Table 3.2. Knowing that expected values take the form of $E[X] = xp(x)$ we know

Table 3.2: Probability Calculations

n	$P(N = n)$
2	$q(2) * p(1)$
3	$q(3) * p(2) * p(1)$
4	$q(4) * p(3) * p(2) * p(1)$
$\vdots$	$\vdots$
n	$q(n) * p(n-1) * \cdots * p(1)$

that our expected value will be of the following form.

$$E[N] = \sum_{n=2}^{\infty} nq(n)p(n-1)\cdots p(1) = \sum_{n=2}^{\infty} nq(n) \prod_{m=1}^{n-1} p(m). \tag{3.4}$$

Now, we know that our probability of failure is our probability of re-entering the queue, so $p = p_r$ and $q = (1 - p_r)$. Substituting these into our expected value equation yields the following final result.

$$E[X] = \sum_{n=2}^{\infty} n(1 - e^{-\beta(n-1)}) \prod_{m=1}^{n-1} e^{-\beta(m-1)} \tag{3.5}$$

The relationship between β and $E[n]$ is demonstrated in Table 3.3. For convenience, Figure 3.1 has been reproduced in Figure 3.5.

Table 3.3: $E[n]$ Values for β

β	$E[n]$
1	2.17818
0.9	2.23056
0.8	2.29374
0.7	2.3715
0.6	2.4696
0.5	2.59812
0.4	2.77497
0.3	3.03813
0.2	3.48606
0.1	4.51188

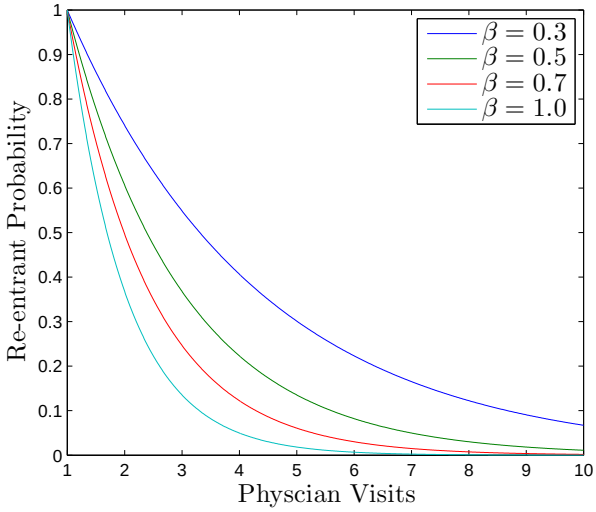

Figure 3.5.: Re-entrant Probability for Various β

For our analyses, we use the following conditions.

- $r = \$300, c_w = \$100/hr$, and $c_o = \$800/hr$
- $\beta = [1.0, 0.9, 0.8, .., 0.1]$
- Each simulation run corresponds to one day of clinic operation
- Clinic simulation for 100 days

We begin by exploring R, which we will define as the total revenue. If we label each patient k, each patient receives n_k services. We have a total of K patients. As such, we can write the following equation.

$$R = \sum_{k=1}^{K} n_k r \tag{3.6}$$

This equation makes it clear that R is a function of our reward, r, and the number of patients serviced, K. Similarly, we can construct such an equation for our waiting time.

$$C_w = \sum_{k=1}^{K} w_k c_w = c_w W_{total} \tag{3.7}$$

Here w_k corresponds to the amount of waiting time for patient i and C_w is the total waiting cost. We will define overtime more simply and only in relation to the clinic.

$$C_O = c_o O_t \tag{3.8}$$

Our first goal was to examine the effect on the current practice model if we assumed a constant R. As mentioned, R is a function of the number of patients serviced and how many services each one has. As such, to fix R, we can instead fix $E[N]$, where N is the sum of all services. Therefore, instead of fixing R, we fix the total number of services provided. This is done for each β. The difficulty in this arises in that we are dealing with patients, so we must operate with only integer values of k. To do this, we assembled a set of k values, one for each β that leads to the closes levels of E[N] possible. These are outlined in Table 3.4. These values lead to $\overline{E[N]} = 23.20$, $\sigma^2_{E[N]} = 0.87$ and a spread of 2.20. The simulation results of this analysis can be viewed in Table 3.5 and Figure 3.6.

As can been seen in Figure 3.6, with a constant $E[N]$, as β increases, we see a steadily decreasing objective function value, $F(S)$. This can be attributed to the increasing (or more negative, since the weights for C_w and C_O are negative) total costs. The small variations in the changes can be attributed to the inconsistencies of our R value. What this is telling us is that as our re-entrant probability increases, even if we attempt to maintain a constant $E[N]$, our wait time and overtime costs

Table 3.4: Constant R Values

β	k	$E[N]$
1	11	23.9600
0.9	10	22.3056
0.8	10	22.9374
0.7	10	23.7150
0.6	9	22.2264
0.5	9	23.3831
0.4	8	22.1998
0.3	8	24.3050
0.2	7	24.4024
0.1	5	22.5594

increase. So under these set experimental conditions, we have observed that for a constant R, as β decreases, C_w and C_O increase, and therefore $F(S)$ decreases.

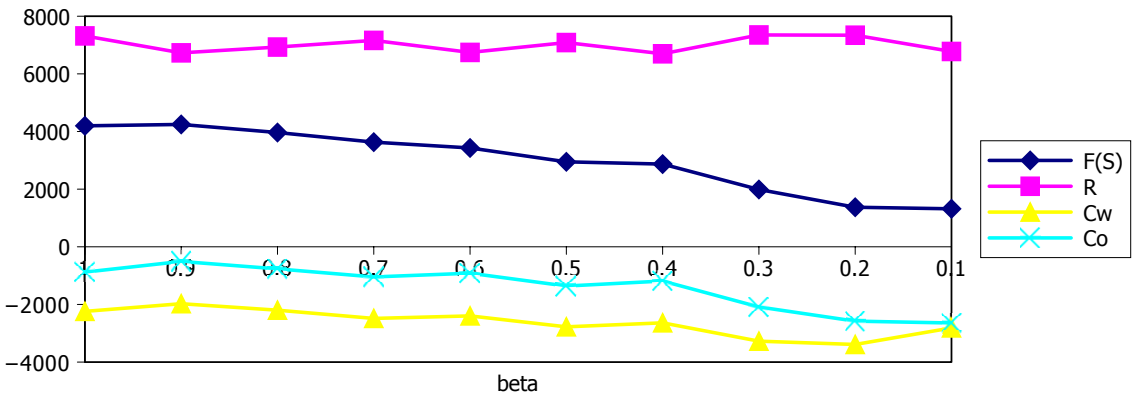

Figure 3.6.: Simulation Results with Constant R

Table 3.5: Simulation Results for Constant R

β	k	$F(S)$	R	C_w	C_O
1	11	4196.82	7317	-2240.9	-879.2
0.9	10	4245.31	6729	-1974.5	-508.8
0.8	10	3964.69	6930	-2199.2	-766.4
0.7	10	3630.73	7161	-2485.0	-1045.6
0.6	9	3429.19	6744	-2397.9	-916.8
0.5	9	2948.06	7086	-2776.7	-1361.6
0.4	8	2869.30	6696	-2638.4	-1188.0
0.3	8	1988.02	7350	-3274.4	-2087.2
0.2	7	1372.35	7341	-3389.1	-2579.2
0.1	5	1318.17	6777	-2814.8	-2644.0

A second perspective was obtained by maintaining a constant k for each β value. For this, we selected an average value of $k = 8$. The results for this are displayed in Table 3.6 and Figure 3.7.

As can be seen in Table 3.6 and Figure 3.7, again as β decreases, we see an overall decrease in $F(S)$. This is due to the fact that while our R value does increase exponentially, both C_w and C_O also increase exponentially. As such, under these experimental conditions we have observed that for a constant k, as β decreases, R, C_w and C_O increase, and therefore $F(S)$ decreases . These two combine into a slightly counter-intuitive view of our system. One would expect that having more difficult procedures leads to more treatments which would increase the overall profit. However, what is being observed is that with the inclusion of waiting time, having patients that spend less time in the system individually makes more economic sense. This clearly point to the importance of the relationship between our revenue, R, and our waiting and overtime costs, C_w and C_O. These results must be treated carefully. If such a view were to be adopted by clinics on a large scale, patients who are expected to require

Table 3.6: Simulation Results for Constant k

β	k	$F(S)$	R	C_w	C_O
1	8	3911.972	5184	-1242.8	-29.6
0.9	8	3901.874	5298	-1336.7	-59.2
0.8	8	3850.082	5415	-1457.9	-107.2
0.7	8	3776.693	5604	-1640.7	-186.4
0.6	8	3635.437	5880	-1891.6	-352.8
0.5	8	3366.923	6228	-2192.4	-668.8
0.4	8	2869.305	6696	-2638.4	-1188.0
0.3	8	1988.021	7350	-3274.4	-2087.2
0.2	8	439.576	8298	-4344.9	-3513.6
0.1	8	-3549.223	10959	-6850.9	-7657.6

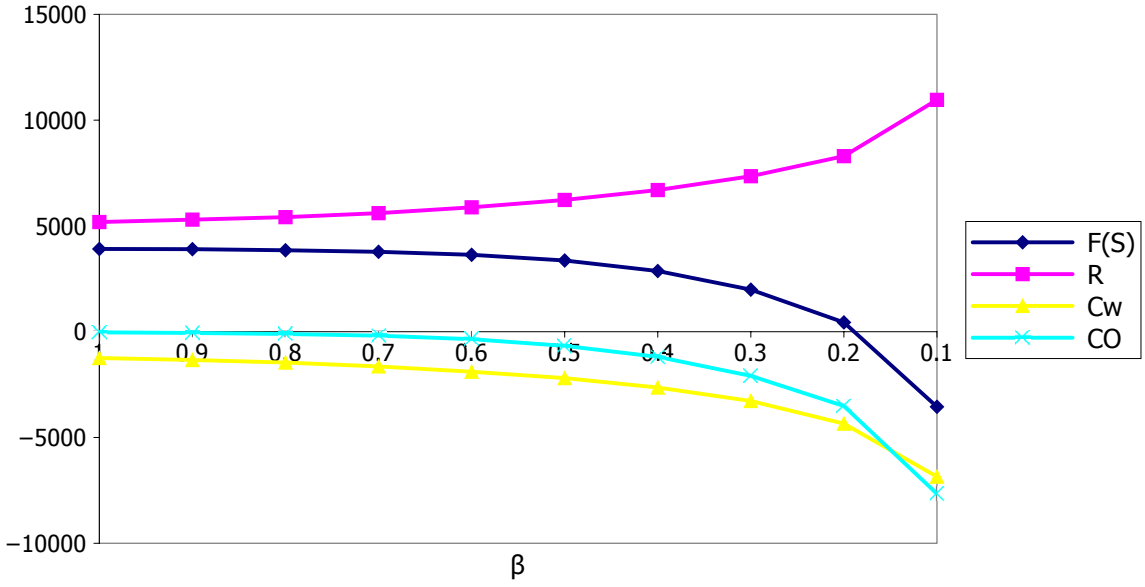

Figure 3.7.: Simulation Results with Constant k

more complex procedures could be shunned by clinics attempting to maximize the customer experience. Further work is planned to evaluate the impact of a changing ratio of these values.

4. PROPOSED QUALITATIVE SCHEDULING CONSTRAINTS

The next area of study was to attempt to create heuristics based on Webb and Rivera's *WAR Score* [28]. These heuristics would attempt to analyze patient characteristics and score patients based on their level of risk for extensive subclinical spread. As mentioned previously, it is extremely difficult to determine the amount of subclinical spread simply from a carcinoma's physical presentation. In response to this, we have identified patient characteristics that place patients at increased risk for subclinical spread. To begin these types of analyses, we began by developing methods for generating random patients. This was accomplished using data sets in Batra's two papers, *Predictors* and *Risk Scale* [30, 31]

4.1 Method of Patient Generation and Scoring

The two data sets presented in the previously mentioned papers were combined in order to obtain the percent of MMS patients who had each of the characteristics identified as relevant by Batra [30]. The probabilities for each of the patient characteristic groups are in Tables 4.1, 4.2, and 4.3.

To use these tables, we sum the values, assigning each characteristic a segment of [0,1]. We then generate a random uniform [0,1] number and assign to that patient the characteristic whose area the generated number falls into. A new random number is generated for each test. Our next characteristic to consider is the size of the carcinoma. We found that carcinomas follow a truncated normal distribution with the parameters $\mu = 11.9mm$ and $\sigma = 8.9mm^2$ [28]. As such, the size is generated according to this distribution. Scoring the patients is then completed as described in

Table 4.1: Gender Probability [30, 31]

Gender	**Probability**
Male	0.552
Female	0.448

Table 4.2: Type Probability [30, 31]

Type	**Probability**
Nodular BCC	0.533
Morpheaform BCC	0.990
Basosquamous BCC	0.027
Recurrent BCC	0.066
SCC	0.150
SCC in situ	0.041
Recurrent SCC	0.011
Other	0.017

Table 4.3: Location Probability [30, 31]

Location	**Probability**
Nose	0.313
Ear, Helix	0.043
Ear, Non-helix	0.041
Eyelid	0.051
Lip	0.048
Forehead	0.105
Cheek	0.183
Chin	0.013
Eyebrow	0.014
Temple	0.070
Neck	0.021
Trunk	0.023
Extremity	0.032
Scalp	0.042

the two papers *WAR Score* and *Risk Scale* [28,31]. These rules for Batra's Risk Scale are displayed in Table 4.4 and WAR Score rules are displayed in Table 4.5.

For the Risk Scale Predictor Scores [30, 31], the highest possible score is selected. For example, if a patients has a $10 - 20mm$ (score of 5) basosquamous BCC on the nose (score), the patient will receive a score of 19. In contrast, the scores provided by the WAR score [28] are additive. Here, if a patient has a $2.5cm(+2)$ recurrent $(+1)$ tumor on the ear $(+1)$ of an aggressive subtype $(+1)$, the patient would receive a score of 5.

For the Risk Scale analyses we the divide patients into "Risk Groups" based on these values. These risk scores are then used to do two additional calculations, assigning values for high/low risk lesions for extensive subclinical spread and the

Table 4.4: Batra Risk Scale Predictor Scores [31]

Predictive Characteristic(s)	**Point**
Basosquamous BCC on nose	19
Morpheaform BCC on nose	19
Morpheaform BCC on cheek	18
Recurrent BCC on nose	17
Lesion on ear helix	16
Size $> 20mm$	14
Recurrent BCC in men	13
Neck tumors in men	12
Eyelid	9
Nodular BCC on nose	9
Temple	8
Size $10 - 20mm$	5

Table 4.5: Webb and Rivera (WAR) Scoring [28]

Size (cm)	**Point**	**Occurence**	**Point**	**Location**	**Point**	**Subtype**	**Point**
$0 - 0.9$	+0	Primary	+0	Other	+0	Non-aggressive	+0
$1 - 1.9$	+1	Recurrence	+1	Nose, ear,	+1	Aggressive	+1
$2 - 2.9$	+2			eyelid, lip			
≥ 3.0	+3						

expected number of Mohs layers required. These are done in the same manner as the generation of patient characteristics, with high/low risk determined by data pulled directly from *Risk Scale* [31] and the number of Mohs layers expected derived from data in both *Risk Scale* [31] and *Predictors* [30]. For high risk/low risk, if a patient meets the probability for extensive subclinical spread, they are deemed a high risk

patient, otherwise they are considered low risk. Low risk patients are defined as expecting only 1-2 Mohs layers to fully excise the lesion. High risk patients are expected to require 3-6 layers. Based on the point values, we assign patients to the Risk Groups outlined in Table 4.6 and give them probabilities of being at high risk for extensive subclinical spread.

Table 4.6: Risk Group Extensive Subclinical Spread Probabilities [30] [31]

Risk Score	**Risk Group**	**High Risk Probability**
<5	1	0.10
5-8	2	0.15
9-12	3	0.23
13-16	4	0.33
17-20	5	0.44

From the data in *Risk Scale* [31] and *Predictors* [30] we were able to derive the probabilities, $p(n)$, on how many layers patients in both the high and low risk groups could expect to have. The results are featured in Table 4.7.

Table 4.7: Risk Level Expected Layers [30] [31]

Low Risk		**High Risk**	
n	$p(n)$	n	$p(n)$
1	0.35	3	0.65
2	0.65	4	0.22
		5	0.07
		6	0.06

4.2 Set of Proposed Scheduling Constraints

Currently, twelve sets of constraints are being used in the simulation. The primary sets are based on the WAR Score [28] and two alternate versions of Batra's Risk Scale [31]. Each one of these then has three additional variations. We chose the following variations for each set: Alternating Difficulty, Longest Expected Processing Time First (LEPTF) and Shortest Expected Processing Time First (SEPTF). Here when we speak of expected processing time, what we are using as an indicator of this is either the WAR Score [28] or the Risk Group. Longer expected processing time means the patient is in a group more prone to needing additional layers. As such, the higher the WAR Score or the Risk Group, the longer the expected processing time. The opposite is true for shortest expected processing time. The logic behind alternating the difficulty comes in attempting to isolate the patients we expect to require more re-entrances into the system. This way they are spread out in the system and should not cause as significant of a bottleneck. LEPTF was chosen to place all difficult patients earliest with the objective of reducing the risk of overtime in the system. SEPTF was chosen because in many scheduling applications, this is often used as a method to reduce overall system time. The basis of the WAR Score is to rank procedures by how difficult they may be [28]. The only problem with this definition is that "difficulty" is a subjective term. This is mostly determined by how long the patient will have to remain at the clinic, but other factors do interact. It functions by reviewing the characteristics of a patient and assigning point values to certain characteristics and summing the patients score. All sets of rules are for a given number of patients per day (e.g. 8 in this study). In his paper, Webb proposes a basic application of the WAR score which utilizes the following rules for a total of 8 patients.

1. A maximum of two patients per day with a WAR score of 3+
2. A maximum of two patients per day with a WAR score of 2
3. All remaining patients must have a WAR score of 0 or 1.

In our execution, we will be scheduling a total of 8 patients. Table 4.8 outlines the possible patient sets. Sets are grouped horizontally. For example, if we have a call-in

Table 4.8: Possible WAR Score Combinations from Webb [28]

WAR Scores		
0 or 1	2	3+
4	2	2
5	2	1
6	2	0
5	1	2
6	1	1
7	1	0
6	0	2
7	0	1
8	0	0

system and three patients with score of 2 and three patients with scores of 3 or larger call in, one of each of these patients will be scheduled the next day. This will lead to a maximum of two 3+ patients, two 2 patients, and a between four and eight 0-1 patients. This appears to be a reasonable on-the-fly application, but we have proposed two alternate forms of this scoring system due to the fact that the patients are still scheduled on a FCFS basis, thus leaving the difficulty randomly distributed throughout the day. To counter this, we proposed scheduling patients using the Webb heuristic (WAR Score [28]) but following a few extra step relating to the ordering of the schedule. These versions are seen in Table 4.9.

The important thing to note here is that patients are still taken on at a FCFS basis and will only be rejected if they violate the stated rules. For example with initial heuristic from Webb, if there are already two 3+ patients scheduled and a third comes up, the patient will be rejected. However, this is not saying that we will

Table 4.9: WAR Score Alternate Rules [28]

WAR Score Variant 1: Alternating
Maximum of 4 patients per day with WAR score 2+
Fill remaining patient slots with WAR score 0 or 1
Alternate patients of 2+ and 0,1 starting with 2+
WAR Score Variant 2: LEPTF
Maximum of 4 patients per day with WAR score 2+
Fill remaining patient slots with WAR score 0 or 1
Longest Expected Processing Time First
WAR Score Variant 3: SEPTF
Maximum of 4 patients per day with WAR score 2+
Fill remaining patient slots with WAR score 0 or 1
Shortest Expected Processing Time First

always have four 2+ patients and four 0,1 patients. This applies to all of the versions here.

The second and third sets of heuristics are derivations from Batra's *Risk Scale* [31]. As stated previously, in this paper she presents a point system to help determine the chances that a patient experiences extensive subclinical spread, defined as requiring three or more Mohs layers to fully excise. The risk groups are segmented into groups in two different manners in the two sets of heuristics based on the *Risk Scale*. These policies are outlined in Tables 4.10 - 4.12.

The reasoning behind regrouping the risk groups lies in the fact that we want these heuristics to be able to be used on-the-fly by physicians. Having different assigning rules for each of the five risk groups seemed overly complicated for such an approach. However, the actual regrouping decisions were more difficult, hence the result of two separate policies. The rules for each of these polices are very similar to those of the WAR score and its alternative versions. The primary difference between

Table 4.10: Policy Risk Group Organization

Risk Group	**Policy A Rank**	**Policy B Rank**
1	1	1
2	1	1
3	1	2
4	2	2
5	2	3

Table 4.11: Risk Scale Policy A Versions

Risk Score Policy A	**Policy A Variant 1: Alternating**
Maximum of 4 patients of rank 2 All remaining patients rank 1	Maximum of 4 patients of rank 2 All remaining patients rank Alternate patients of ranks 1 and 2
Policy A Variant 2: LEPTF	**Policy A Variant 3: SEPTF**
Maximum of 4 patients of rank 2 All remaining patients rank 1 Longest expected processing time first	Maximum of 4 patients of rank 2 All remaining patients rank 1 Shortest expected processing time first

ranking the patients by their WAR score and their Risk Score is that the WAR score emphasizes the amount of time that the patient will take to treat, from first cut to final healing decision, while the Risk Score is completely dependent on the number of Mohs layers that are expected to be needed to fully excise the lesion. Again, the initial versions of the Risk Score policies focus only on the numbers of each specific rank that we are allowed to schedule each day. The alternate versions of it focus on ordering the patients in specific ways, attempting to improve on the constrained FCFS method that was originally used. The alternating of difficult/long procedures with easier/shorter procedure is designed to provide the physician with extra time to complete the more difficult case while still completing another case. It is possible

Table 4.12: Risk Scale Policy B Versions

Risk Score Policy B	**Policy B Variant 1: Alternating**
Maximum of 2 rank 3 patients	Maximum of 2 rank 3 patients
Maximum of 2 rank 2 patients	Maximum of 2 rank 2 patients
All remaining patients rank 1	All remaining patients rank 1
	Alternate rank 2+ and rank 1
Policy B Variant 2: LEPTF	**Policy B Variant 3: SEPTF**
Maximum of 2 rank 3 patients	Maximum of 2 rank 3 patients
Maximum of 2 rank 2 patients	Maximum of 2 rank 2 patients
All remaining patients rank 1	All remaining patients rank 1
Longest expected processing time first2	Shortest expected processing time first

though that the shorter/easier patients will end up being bottlenecked by the more difficult cases surrounding them. In an attempt to alleviate this, the second variation was proposed. By scheduling all difficult cases early, they may delay and interact with each other, but the effect may be less severe than the effect these delays would have on the shorter cases. The third variation is to schedule the easiest cases earlier in the morning in an attempt to drive down over all waiting times by minimizing the initial bottleneck.

4.3 Simulation and Numerical Analyses of Proposed Constraints

Analysis of the effectiveness of the heuristics was accomplished through a simulation model. The experimental conditions are as follows.

- $r = \$300, c_w = \$100/hr$, and $c_o = \$800/hr$
- Each simulation run corresponds to one day of clinic operation

Table 4.13: Time in System

Heuristic	**Min**	**Avg**	**Max**
FCFS	156.29	400.737	604.269
RiskA	156.181	399.489	602.219
RiskA: Alt	158.426	401.954	602.342
RiskA: LEPTF	156.136	392.196	585.450
RiskA: SEPTF	145.515	382.105	585.729
RiskB	156.415	394.636	592.919
RiskB: Alt	157.983	396.202	592.782
RiskB: LEPTF	158.346	393.085	585.393
RiskB: SEPTF	145.211	380.663	585.702
Webb	155.973	397.822	600.318
Webb: Alt	156.451	398.672	600.280
Webb: LEPTF	159.094	397.759	595.018
Webb: SEPTF	146.449	388.629	594.943

- 8 patients scheduled per day
- Patients scheduled to first 8 slots
- Slots are 30 minutes in length
- Clinic simulation for 1000 days

Groups of random patients were generated, sorted (if applicable), and fed into the simulation model. Several key performance measures were analyzed for each heuristic, including average wait times, average time in system, and resource utilization. These measures can be viewed in Tables 4.13, 4.14, 4.14 and 4.16. All results are in minutes.

As can be seen from the results, the impact of the proposed heuristics is minimal at best. In relation to the time each patient spends in the system, the proposed

Table 4.14: Time in Phys Queue

Heuristic	**Min**	**Avg**	**Max**
FCFS	0	13.446	68.677
RiskA	0	13.473	68.677
RiskA: Alt	0	13.473	68.677
RiskA: LEPTF	0	13.734	68.677
RiskA: SEPTF	0	13.734	68.677
RiskB	0	13.624	68.677
RiskB: Alt	0	13.624	68.677
RiskB: LEPTF	0	13.734	68.677
RiskB: SEPTF	0	13.734	68.677
Webb	0	13.505	68.677
Webb: Alt	0	13.505	68.677
Webb: LEPTF	0	13.575	68.677
Webb: SEPTF	0	13.575	68.677

Table 4.15: Time in Path Queue

Heuristic	**Min**	**Avg**	**Max**
FCFS	0	120.857	216.662
RiskA	0	120.851	216.649
RiskA: Alt	0	122.008	216.844
RiskA: LEPTF	0	122.248	217.858
RiskA: SEPTF	0	117.515	212.99
RiskB	0	121.211	216.669
RiskB: Alt	0	121.969	216.559
RiskB: LEPTF	0	122.644	218.568
RiskB: SEPTF	0	116.839	212.744
Webb	0	120.716	216.754
Webb: Alt	0	121.102	216.405
Webb: LEPTF	0	122.196	218.458
Webb: SEPTF	0	117.964	213.081

Table 4.16: Average Utilization Levels

Heuristic	**Physician**	**Pathology**
FCFS	0.426	0.956
RiskA	0.426	0.956
RiskA: Alt	0.426	0.957
RiskA: LEPTF	0.430	0.956
RiskA: SEPTF	0.429	0.954
RiskB	0.428	0.956
RiskB: Alt	0.429	0.956
RiskB: LEPTF	0.430	0.956
RiskB: SEPTF	0.429	0.954
Webb	0.426	0.956
Webb: Alt	0.426	0.956
Webb: LEPTF	0.428	0.956
Webb: SEPTF	0.427	0.955

scheduling sets due generally perform better, but only marginally so. This is a general trend throughout all of the presented results. Overall, the SEPTF methods do outperform the others in overall time in system, but this does not necessarily translate to both wait time in the pathology and physician queues. We believe these results may be due to a weakness in our patient generation data.

One of the potential difficulties associated with using the Batra data is that both sets of data are collected from the same hospital, namely the Beth Israel Deaconess Medical Center in Boston, MA [30] [31]. One of the leading causes of skin cancer is overexposure to sunlight [2], which varies from location to location. For example, one would expect to see higher incidence rates of overexposure in areas known for being sunny, such as Florida, Texas, or Australia, and less in areas that are not sunny, such as Maine, Massachusetts, or Canada. This difference may lead to the data being

generated having a specific skew associated with the area in which it was collected. To further analyze the Batra data, frequency data was calculated on three specific characteristics: the expected number of Mohs layers, the calculated WAR score, and the Batra Risk Score. These are represented in Figures 4.1, 4.2, and 4.3.

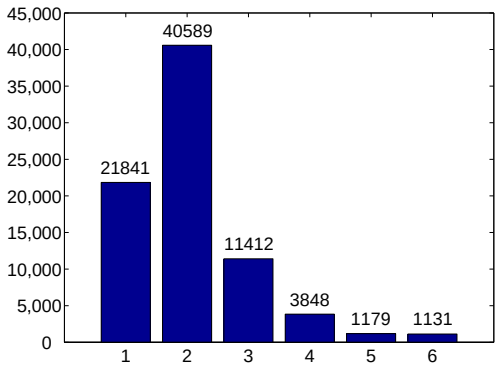

Figure 4.1.: Expected Layer Frequencies

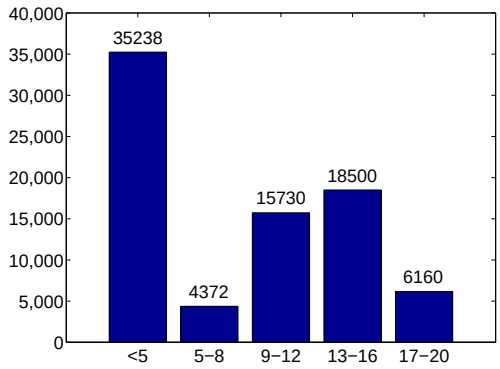

Figure 4.2.: Risk Score Frequencies

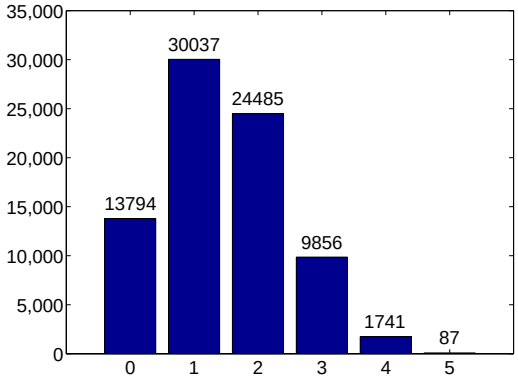

Figure 4.3.: WAR Score Frequencies

The important characteristic to note here is the major skew to lower scores and less layers. When reviewing patients who experience extensive subclinical spread, defined previously as requiring three or more Mohs layers, this characteristic applies to only 21.96% of all patients generated. Since the heuristics proposed based on Batra's *Risk Score* [31] are focused around controlling the number of patients with a high risk of extensive subclinical spread the physician would see in a day, it is reasonable that

Table 4.17: Number of Changes Made to Schedule

Model	**Min**	**Avg**	**Max**
Webb	13,232	13,673	14074
Risk A	879	992	1093
Risk B	21934	22,397	22975

the proposed heuristics would have little to no effect on the average amount of time spent at the clinic by patients. With so few at risk patients existing in the patient pool, few patients will not be scheduled due to violating the constraints on at risk patients. A similar statement can be made about the WAR Score data [28]. The rules for scheduling patients according to the WAR score heuristics control patients of rank two and three or higher. A maximum of two patients with WAR scores of both 2 and 3+ can be scheduled a day. With a daily schedule consisting of eight patients, that means a maximum of 25% of the scheduling capacity can be assigned to each of these categories, with 50% remaining for patients of scores 0 and 1. With patients of rank 2 populating only 30% of the patient pool and patients with scores 3+ consisting of a mere 14% of the population, it becomes readily apparent how the scheduling rules will have little effect on the actual patients being seen by the physician.

To further explore this area, a comparison was done between the initial FCFS patient lists and the unsorted (first version) of each of the heuristic sets to see how man patients were being dropped or moved in the lists. To do this, 100 sets of the total patient lists were generated and compared. The total patients in each set is 80,000. The results can be seen in Table 4.17. The results demonstrate that few patients out of the pool are actually being moved or declined based on the added rules with only 17.09%, 1.24%, and 28.00% being changed on average in each model, respectively. This has been identified as a weakness in the data, but various solutions to the problem exist. Proposed solutions to this difficulty include either obtaining similar data from literature from other parts of the country and world or artificially generating data with

a wider variety of frequency distributions. With the minimal literature that exists in the area, obtaining additional data sets without conducting additional surveys around the globe seems to have a minimal chance of occurrence. However, it should be very simple to produce alternative data sets and see their effect on the model. This can be accomplished by directly altering the WAR score, Risk Score, and Expected Layers, or by simply selecting characteristics that push these scores up and increasing their probability of occurring. Finally, the proposed heuristics can simply be made much stricter in order to account for the data. Additionally, the constraints could be reformulated for various types of data sets to increase their generalizability.

5. PROPOSED SEQUENTIAL SCHEDULING POLICY FOR MMS CLINICS

Our proposed scheduling policy consists of a call-in system in which patients are generated and then scheduled. At the time of each patient's scheduling, an objective function is analyzed in order to choose the best slot for the patient to be scheduled for based on the current schedule and the patient type.

5.1 Parameters and Definitions of Scheduling Policy

Our system in this approach has the following characteristics.

- Patients are scheduled sequentially and cannot affect a preceding patient's appointment time.
- We use a myopic policy, only considering the upcoming patient and not any future call-in schedule.
- A schedule corresponds to a single day's operation
- Our algorithm attempts to maximize a combination of the revenue, wait time disutility, and overtime costs.
- We assume homogeneous patient behavior in both no-show, p_{ns}, and re-entrant, p_r, probabilities.
- We have a 2 server, 2 stage system consisting of physician and pathology, but we reduce to 1-1 for tractability.
- We assume patient punctuality.
- We do not fix the number of patients that can be scheduled each day.

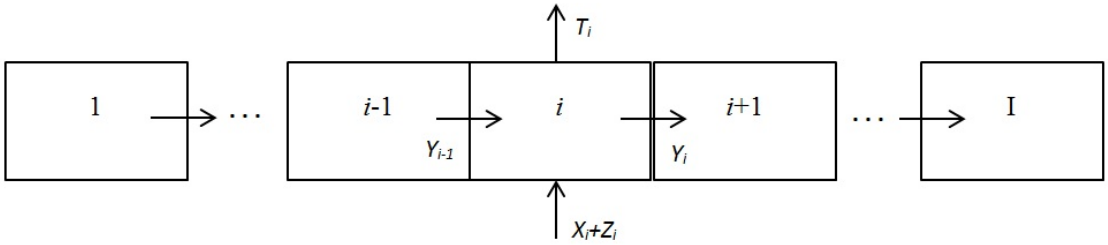

Figure 5.1.: The system.

In most clinic modeling approaches, a scheduling matrix, S of size $I \times J$ is used, where S_{ij} represents the number of patients of type j scheduled in slot i. However, since we assume completely homogeneous patients, $J = 1$, so we instead have a scheduling vector S_i. We assume a standard workday of 8 hours with 30 minute slots, leading to $I = 16$ slots per day. We also define λ as the expected number of patients that can be serviced during each slot's time period, assuming exponential service time. This characteristic can also be utilized to control the length of the slots, changing it in relation to the length of the slot. For example, if we originally have 30 minute slots and expect to be able to serve 2 patients during each slot (approximate service time of 15 minutes), changing the length of our slot to 15 minutes would correspond to a change in λ from 2 to 1. λ is also used to account for changes in service time. Our system is demonstrated in Figure 5.1.

The system variables are defined as follows.

- I: number of time slots
- p_{ns}: the probability of a patient not showing up for a scheduled slot
- p_r: the probability that a new treatment is needed after the treatment of a patient in slot i, the new treatment will assigned to slot $i + 1$.
- D: the length of each time slot

- λ: the expected number of treatments completed provided there are infinitely many treatments assign to this slot and assuming exponentially distributed service time for each treatment (this means the expected service time is $\frac{D}{\lambda}$)

It should be noted here that with the variables defined as above, all patients have identical no-show and re-entrant probabilities and that each time slot i is assumed to be equal in length.

- S_i: our schedule vector, S_i is the number of patients assigned to slot i during the call-in period
- S^n: our schedule after the n^{th} optimal assignment
- X_i: number of scheduled patients who show up for slot i. X_i depends on the schedule S and is a random variable
- L_i: number of patients treated in slot i, provided the queue is long enough
  - L_i is a Poisson random variable because we assume exponential service time

Due to the fact that a treatment completed in slot i has a probability of generating an additional treatment in slot $i+1$ (our re-entrants), we define Z_i in the following manner.

- Z_i: number of patients who re-enter the system following treatment in slot $i-1$
- Y_i: number of patients who overflow from slot i
  - $Y_i = max\{Y_{i-1} + X_i + Z_i - L_i, 0\}$
- T_i: number of patients treated in slot i
  - $T_i = min\{Y_{i-1} + X_i + Z_i, L_i\}$

5.2 Execution of Sequential Scheduling Algorithm

We will indicate a day's schedule after n patient call-ins as $S^n \in \mathbb{R}^I$, so S_i^n represents the number of patients scheduled in slot i after n call-ins. Additionally, we will define an assignment vector Δ^i of length I which consists of a one at the ith position and zeros at all others. We will also set U as the set of all slots i. The model then uses a cost-reward structure to create a value function that is dependent on our schedule, S. We set r as the reward for each treatment of a patient and c_i as the cost of a patient overflowing from slot i. It should be noted that $c_i = c_w$ for all slots except that $c_i < c_I$ due to the cost of overflow from the last slot being more expensive due to overtime costs. In this instance, $c_I = c_O$. As such, we can structure our value function with the following characteristics: reward for each patient treated, costs for each patient overflow, net reward for each expected treatment during the overtime period. This characterization leads us to the following function.

$$f(S) = E[r\sum_{i=1}^{I}(X_i + Z_i) - \sum_{i=1}^{I} c_i Y_i + (r - c_I)Z_{I+1} + (r - c_I)(Y_i + Z_{i+1})\frac{p_r}{1 - p_r}] \quad (5.1)$$

5.2.1 Probability Derivations

To compute this function, we need the expectations of X_i, Y_i, and Z_i. The expectation of X_i is simple enough, since it is a binomial random variable following (S_i, p_a). To do this calculation, we can set Q_{ik}, the probability of k patients arriving in slot i. We will set $p_a = 1 - p_{ns}$.

$$Q_{ik} = Pr\{X_i = k\} = \frac{S_i!}{k!(S_i - k)!} p_a^k (1 - p_a)^{(S_i - k)} \quad (5.2)$$

Using this, we can calculate the expected value of X_i

$$E[X_i] = \sum_k k Q_{ik} = \sum_{k=0}^{k \leq S_i} k \frac{S_i!}{k!(S_i - k)!} p_a^k (1 - p_a)^{(S_i - k)} \quad (5.3)$$

If we attempt to define Z_i in the following manner, we quickly realize that this is not possible without conditioning the probability on T_i. However, we can still state that

Z_{i+1} is a binomial random variable with parameters T_i and p_r. As such, let us define b_n as a random variable following Binomial(n, p_r) and set $Z_{i+1} = b_{T_i}$. With this in mind we can now look at the joint distribution of (Y_i, Z_{i+1}) which can be calculated recursively.

$$\begin{aligned} P(Y_i = y, Z_{i+1} = z) &= \sum_t P(Y_i = y, b_{T_i} = z, T_i = t \\ &= \sum_t P(Y_i = y, b_t = z, T_i = t) \\ &= \sum_t P(Y_i = y, T_i = t) P(b_t = z) \end{aligned} \tag{5.4}$$

Knowing the parameters and distribution of $P(b_t = z)$, we can now separate $P(Y_i = y, T_i = t)$ and solve for it.

$$\begin{aligned} &P(Y_i = y, T_i = t) = \\ &= \sum_{y_{i-1}, z_i, x_i, l_i} P(Y_i = y, T_i = t | Y_{i-1} = y_{i-1}, Z_i = z_i, X_i = x_i, L_i = l_i) \cdot \\ &P(Y_{i-1} = y_{i-1} Z_i = z_i, X_i = x_i, L_i = l_i) \\ &= \sum_{y_{i-1}, z_i, x_i, l_i} P(Y_i = y, T_i = t | Y_{i-1} = y_{i-1}, Z_i = z_i, X_i = x_i, L_i = l_i) \cdot \\ &P(Y_{i-1} = y_{i-1} Z_i = z_i) P(X_i = x_i) P(L_i = l_i) \end{aligned} \tag{5.5}$$

Using the definitions of Y and T we can then write

$$P(Y_i = y, T_i = t) = \begin{cases} 1 \text{ if } & y = max\{y_{i-1} + x_i + z_i - l_i, 0\} \text{ and} \\ & t = min\{y_{i-1} + x + z_i, l_i\} \\ 0 \text{ otherwise} & \end{cases} \tag{5.6}$$

Using this equation, we can state the following.

$$P(Y_0 = y, Z_1 = z) = \begin{cases} 1 & if \text{ y=0 } and \text{ z=0} \\ 0 & otherwise \end{cases} \tag{5.7}$$

As such, by utilizing (5.4), (5.5), (5.6) and (5.7), we can compute $P(Yi = y, Z_{i+1} = z)$ recursively. Thus, to find the expectation of both Y_i and Z_i, we need to use these conditioned probabilities. We progress as follows.

$$E[Y_i] = \sum_y y \sum_z P(Y_i = y, Z_{i+1} = z) \tag{5.8}$$

$$E[Y_i] = \sum_y y \sum_z \sum_t P(Y_i = y, T_i = t) P(b_t = z) \tag{5.9}$$

In this instance, we have the expression for $P(Y_i = y, T_i = t)$ and we know that b_t follows Binomial(t, p_r). So similarly for Z,

$$E[Z_i] = \sum_z z \sum_y \sum_t P(Y_i = y, T_i = t) P(b_t = z) \tag{5.10}$$

We can similarly define an expression for T_i.

$$E[T_i] = \sum_t t \sum_y P(Y_i = y, T_i = t) \tag{5.11}$$

5.2.2 The Scheduling Policy

The primary outputs of this model then are a vector of objective function values for each time a patient is added and an overall schedule for the total number of patients scheduled. The scheduling algorithm executes in the following manner.

1. Set $S_i = 0 \; \forall \; i \in U$
2. Wait for nth call
3. nth call occurs
4. For each $i \in U$
   (a) Set $S_i^n = S^{n-1} + \Delta^i$
   (b) Compute $f(S_i^n)$
5. If $f_i^n \geq f^{n-1}$

(a) $i^* = \text{argmax} f_i^n, \;\; S^n = S^{n-1} + \Delta^{i^*}$

(b) Set $n = n + 1$, Go to Step 2

(c) Else Stop.

The objective function is unimodal as we add patients, so that point at which it obtains its maximum value is used to select our schedule. The schedule is then accepted up to the number of patients at which our optimal objective function value occurs. It should be noted that while our scheduling policy penalizes based on expected overflow, our simulation model objective function is based upon total waiting time. As such, it would be useful to have expressions for the expected total waiting time for a given schedule to enable a more accurate comparison of the expected results of the schedule and the simulated results.

5.2.3 Expected Total Waiting Time Derivations

We begin by looking at the expected waiting time for all patients in a specific slot i. If we state that there are N patients in this slot, then we can state that $X_i + Y_{i-1} + Z_i = N$. We can also define μ_n as the actual service time for the n^{th} patient in this slot assuming that the slot were of infinite length. Looking at this, we know that the first patient, $n = 1$ has no waiting time in this slot as they are the first to be serviced. Every other patient though must wait for all others before them to complete service. As such, the nth patient will need to wait for $n - 1$ services to be completed before his or her service begins. However, as defined previously, each slot is of length D, thus making D the maximum amount of waiting time per patient in each slot. This leads to the following statements: the total waiting time for the nth patient in slot i is defined as $min(D, \sum_{k=0}^{n-1} \mu_k,)$, and $\mu_0 \equiv 0$. We know that the service time for each patient is independent and identically distributed following an assumed exponential distribution. As such, the waiting time for the nth patient,

given that $n \geq 2$ is a truncated $Gamma(\frac{\lambda}{D}, n-1)$ random variable. The expected value of this function is as follows.

$$E[W_i(n)] = \int_0^D \frac{\frac{\lambda}{D} e^{-\frac{\lambda}{D}x} (\frac{\lambda}{D}x)^{n-2}}{(n-2)!} dx \tag{5.12}$$

Using this equation, we can state the expected total waiting time in any slot, which we will denote as TW_i, as follows.

$$\begin{aligned} E(TW_i) &= E[E(TW_i|N)] \\ &= \sum_{N \geq 2} P(X_i + Y_{i-1} + Z_i = N) \cdot \left[\sum_{n=2}^{N} \int_0^D \frac{\frac{\lambda}{D} e^{-\frac{\lambda}{D}x} (\frac{\lambda}{D}x)^{n-2}}{(n-2)!} dx \right] \end{aligned} \tag{5.13}$$

This equation though is not generalizable to the waiting in the overtime period. In the overtime period, we can state the expected number of treatments required by a patient as $\frac{1}{1-p_r}$. Following this logic, the expected length of the overtime period is $\sum_{N \geq 1} P(Y_I + Z_{I+1} = N) \cdot \frac{N}{1-p_r}$. We will denote the total waiting time for all patients in the overtime period as $WO(N)$ given that there are N patients in the system at the beginning of the overtime period. By definition $WO(0) = WO(1) = 0$, as there will be no waiting time in the overtime period if there are not at least two patients remaining in the system when it begins. Using this definition, we can calculate this waiting time recursively.

$$E[WO(N)] = \frac{D}{\lambda} \cdot (N-1) + (1-p_r) E[WO(N-1)] + p_r \cdot E[WO(N)], \; N \geq 1. \tag{5.14}$$

This equation has multiple parts. This first term indicates the amount of waiting time for the $N-1$ patients who are waiting during the first treatment in the overtime period. After this treatment is completed though, that patient will leave the system with a probability of $(1-p_r)$. The leads us to the second term, the expected waiting time of the remaining $N-1$ patients. However, the patient who completes treatment also has a probability, p_r, of re-entering the system. If this were to occur, the expected

waiting time in the overtime period would remain the same, leading to the third term. Re-arranging the terms of both sides of (5.14), we have the following.

$$\begin{aligned} E[WO(N)] &= \frac{D}{\lambda \cdot (1-p_r)} \cdot (N-1) + E[WO(N-1)],\ N \geq 1. \\ &= \frac{D}{\lambda \cdot (1-p_r)} \cdot [0+1+2+\cdots+N-1] + E[WO(0)] \\ &= \frac{D}{\lambda \cdot (1-p_r)} \cdot \frac{N(N-1)}{2}. \end{aligned} \tag{5.15}$$

Thus, the total expected waiting time is given by the following.

$$\sum_{N \geq 1} \frac{D}{\lambda \cdot (1-p_r)} \cdot \frac{N(N-1)}{2} \cdot P(Y_I + Z_{I+1} = N) \tag{5.16}$$

5.2.4 Model Weaknesses

The problems that we see with this model can be described in terms of its levels of abstraction. The primary area in which the model deviates is the inclusion of only one service type, the physician. As such, the model lacks the double service characteristic of the actual clinic, limiting the applicability of its results. Also, the model assumes a constant re-entrant probability for all patients. This is also not completely accurate, as the more times a patient has been treated by the physician, the lower his/her probability of re-entering the queue should be, according to our re-entrant probability model. Also, the calculations in the scheduling policy are based around overflow, rather than expected wait time. While we do have expressions for the expected wait time, the policy still conducts its calculation using overflow. However, we do still believe that such a model resembles the system enough to justify a comparison of objective function values obtained by the current practice simulations and our generated schedules.

5.3 Simulation Analysis of Generated Schedules

Analysis of the schedules produced through our call-in model was conducted through the same simulation model used for the current practice analyses. Experimental conditions for the simulation are as follows.

- $r = \$300, c_w = \$100/hr$, and $c_o = \$800/hr$
- Each simulation run corresponds to one day of clinic operation
- Slots are 30 minutes in length
- Clinic simulation for 100 days

For tractability of our scheduling policy, it was necessary to set p_r as a constant, rather than $p_r(n, \beta)$. In order to determine the appropriate value of p_r, we determined the expected number of physician visits per patient for our $p_r(n, \beta)$ for the same set of β values as in our current practice analysis. The expected values in Table 3.3 were used to determine an equivalent constant re-entrant probability value.

Using the constant probabilities of re-entrance, schedules were then developed for each level of β. These schedules were then executed in our simulation. In every execution of the model, we have defined our no-show and re-entrant probabilities as in our current practice model, specifically, $p_n s = 0.1$ and $p_r = p_r(n, \beta)$ as defined in Equation 3.1. For our primary performance measures we selected the objective function value and the average patient wait time in system. For analysis, we selected our schedule with $\lambda = 1.5$, $I = 16$ with 30 minute slots, corresponding to a standard 8 hour work day. The results are presented in Table 5.1.

As can be seen in Table 5.1, in situations with larger β (and therefore lower re-entrant probability), our mathematical model produces favorable results in most cases, allowing us to schedule similar or larger numbers of patients without increasing the patient time in system. However, as β approaches its smallest values, our objective function and time in system values become worse than that of the current practice simulation. This could easily be a result of the level of abstraction mentioned about

Table 5.1: Mathematical Model Simulation Results

β	k	$F(S)$	Time in System
1	9	4426.783	1.973
0.9	9	4476.702	1.970
0.8	8	4197.070	1.930
0.7	8	4153.828	2.793
0.6	8	4131.296	2.372
0.5	8	3883.962	2.682
0.4	7	3574.197	3.035
0.3	7	3030.522	3.772
0.2	6	2641.956	4.613
0.1	5	1497.722	6.313

the model. Since we are forced to utilize a constant probability in our calculations rather than one that accounts for number of times a patient has already re-entered the system, it is possible that this reduces are generalizability to small β values. A comparison of these results to our optimal current practice simulations is presented in Table 5.2. Of course, our most important comparison is between the $F(S)$ values of the two methods. Generally, the algorithm produces superior results to the current practice method. The same can be said for the amount of time each patient can expect to spend in the system. These do not hold however for the smallest values of β, the instances were our re-entrant probability is the largest. This could though be related to the abstraction of the model, especially considering the constant re-entrant probability used. We have though however demonstrated a baseline that in most instances, our scheduling algorithm can produce superior results.

Next, we wished to examine some of the relationships between the various cost and reward weights and the function of our scheduling algorithm. We began this analysis using two unlikely relationships, $c_O = r$ and $c_O < r$. In our test of $c_O = r$, we found

Table 5.2: Comparison of Current Practice and Mathematical Model Results

	Current Practice			**Scheduling Policy**		
β	k	$F(S)$	**Time**	k	$F(S)$	**Time**
1	10	4334.14	2.459	9	4426.783	1.973
0.9	10	4245.31	2.587	9	4476.702	1.970
0.8	9	3991.83	2.718	8	4197.070	1.930
0.7	8	3776.69	2.693	8	4153.828	2.793
0.6	8	3635.44	3.044	8	4131.296	2.372
0.5	8	3366.92	3.490	8	3883.962	2.682
0.4	7	3150.67	3.754	7	3574.197	3.035
0.3	6	2954.59	4.191	7	3030.522	3.772
0.2	4	2740.57	3.991	6	2641.956	4.613
0.1	3	2554.07	4.794	5	1497.722	6.313

when these two measures are equivalent, the system reaches a sort of equilibrium. After scheduling the k patients in the optimal manner, the system will schedule an infinite number of patients in the final slot, assuming since the reward is the same as the cost, a net value of 0 is added each time. This represents a weakness in the model. It fails to incorporate the wait time of patients scheduled in the overtime slot. For $c_O < r$, the results are slightly different. In this case, since $c_O < r$, our objective function is no longer unimodal with respect to k in our scheduling algorithm. Instead, it is unbounded and will eventually increase linearly with each patient added. The linear amount at which it increases has been found to be independent of c_W and β (or the equivalent constant p_r) while being dependent c_O only. The p_r only changes the amount of patients needed to reach the linear increase. However, we have been unable to determine the relationship between the linear increase and c_O.

In the future, our objective is to continue this strain of analysis into the entire parameter space. We are currently working on the derivation of the expected profit

for a given schedule, and therefore the expected revenue, waiting time, and overtime. These derivations should enable us to further experiment with and analyze the behavior of the scheduling algorithm to be more able to accurately control its function, thus enabling it to perform even better against the current practice model.

6. CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

We have demonstrated three tools in this paper: a method of modeling and simulating current MMS clinic practice in order to identify optimal number of patients to schedule, a set of proposed heuristics to be used to increase the patient experience and an algorithmic scheduling policy. Our current practice model consisted of a simulation of an MMS clinic. By estimating service times for the two aspects of MMS, physician visits and pathology, and developing a probability distribution for patient re-entrant probability with a controlling shape factor β, we were able to produce a working simulation of the MMS clinic. We conducted analysis through an objective function that accounted for revenue based on the number of procedures completed, the disutility of patient wait time, and overtime operating costs. Our objective function was unimodal for all values of β analyzed, allowing us to identify an optimal number of patients to schedule for each β value. Such a tool could be used today by examining the types of patients seen by a physician at an MMS clinic and identifying an average number of physician services that the patients require. Then by limiting the maximum number of patients seen to this optimal value, the clinic should be able to improve the overall patient experience while also improving overall profits. However, the model did have unexpected impacts, for example indicating that under the conditions of our objective function, a clinic should attempt to schedule large numbers of simple procedures rather than a smaller number of more complex procedures. Such findings, if put into practice could limit treatment availability to specific patient groups, so such results must be treated carefully.

Our second tool developed was a set of scheduling constraints designed for on-the-fly use in an MMS clinic. A series of twelve heuristics were developed and compared

to a FCFS scheduling policy. Using a simulation model similar to the current practice model but based on patient characteristics rather than a re-entrant probability, we were able to conduct an analysis of the heuristics. The proposed heuristics proved to have little effect on the patient experience, with wait times and utilization levels remaining nearly constant, even in comparison with the FCFS method. However, after further analysis it was revealed that this was an effect of the data used in generating random patients. The data obtained from literature lead to a patient distribution that was skewed towards less expected layers, leading to our algorithms having a minimal effect on the schedules themselves. This leads us to the conclusion that while these heuristics may have some level of applicability depending on the types of patients being seen at a clinic, they suffer from a lack of generalizability. If further data sets can be obtained, additional analysis of these heuristics could be conducted to prove their viability. Additionally, reformulating these constraints for specific data sets may be possible.

Third, we have demonstrated a myopic sequential scheduling policy for MMS clinics that maximizes a profit function based on revenue, wait time disutility, and overtime costs. Our objective in the development of this model was to demonstrate superior results to those of our current practice model in a parameter space including β, number of patients, and our cost and reward weights. While the scheduling policy was used to compare against the current practice model, there were several differences between the way the two modeled the system. These included using a constant probability of re-entrance rather than a constant one and using exponential service times rather than the developed Gamma distributed times. The schedules generated in this model proved to be unimodal under specific cost/weight structures in which the system was stable. Under these stability conditions, schedules were produced and those with the maximum profit function value were selected as the optimal schedules. These schedules were then executed in the MMS clinic simulation. Results were found to indicate that under most values of β, the scheduling policy outperformed the current practice model. In the future, through the derivation of the expected profit

from the algorithm, we hope to be able to develop a series of theorems describing the function of our system so that we can better understand how it functions in all areas of the previously mentioned parameter space.

With the development of these three tools, we believe that we have provided a substantial framework for work in the future. This paper provides multiple areas of research, including simulation, heuristic development, and algorithmic scheduling policies. We believe that we have laid a basis that will enable these tools to be applied rigorously in the future to Mohs Micrographic Surgery. MMS is a valuable tool in the treatment of NMSCs, and we hope that with the development and application of tools similar to those presented here, we can not only benefit the clinic's operations, but improve the overall patient experience.

6.2 Future Work

Beyond the imminent areas of future work mentioned previously, there are several areas in which research of this nature can be continued. One of the major objectives we would like to implement in the future is a sensitivity analysis of the performance of our scheduling policy based on all of the parameters in the system, namely k, β,r,c_w, and c_o Such a sensitivity analysis will enable us to determine what parameters have the most dramatic effect on the operation of our scheduling policy. This would then allow a reformulation of the policy so as to combat any weakness.

Another area of development we would like to address is the current use of exponential service times. It is a common assumption due to its ease of calculation, but often does not translate into real world application well. As such, we would like to re-develop the model to account for general service times. This would allow this study to be generalized to a wide range of applications.

An additional area of future work is into the utilization of service constraints, as in Turkcan 2011 [34]. The application of service constraints to our system would again improve the generalizability of study. As with the rest of this problem though,

the difficulty arises in the re-entrant problem. Similarly, the inclusion of flexibility in scheduling for personal preference would be an excellent addition to this study. Patients do not typically simply accept whatever slot is assigned to them. They usually have specific times they are available. This inclusion would definitely increase the applicability and useful of this model in a real world setting.

Additionally we would like to change the implementation of the re-entrant probability in the scheduling policy. Currently, all patients in the system are assumed to have the same β. This would not be the case in a real clinic. If we were able to account for varying β values, such as Muthuraman accounts for varying no-show probabilities [36], we could develop and test schedules under varying patient load types. Additionally, allowing for varying β would enable a study of changing physician practices as well.

Finally, with the addition of some of the work mentioned above, implementation would be the next logical step. In studies like this one, the eventual goal is always implementation of your system in an actual clinic setting. We believe that with the above work, it would be feasible to use a system such as the one proposed here in a real-world application.

LIST OF REFERENCES

LIST OF REFERENCES

[1] H.W. Rogers, M.A. Weinstock, A.R. Harris, M.R. Hinckley, S.R. Feldman, A.B. Fleischer, and B.M. Coldiron. Incidence estimate of nonmelanoma skin cancer in the united states, 2006. *Archives of dermatology*, 146(3):283–287, 2010.

[2] Skin Cancer Foundation. Skin cancer facts. `http://www.skincancer.org/skin-cancer-information/skin-cancer-facts\#melanoma`. Accessed August 8,2012.

[3] L. Ravitskiy, D.G. Brodland, and J.A. Zitelli. Cost analysis: Mohs micrographic surgery. *Dermatologic Surgery*, 38:585–594, 2012.

[4] J. Paoli, S. Daryoni, A.M. Wennberg, L. Molne, M. Gillstedt, M. Miocic, and B. Stenquist. 5-year recurrence rates of mohs micrographic surgery for aggressive and recurrent facial basal cell carcinoma. *Acta dermato-venereologica*, 91(6):689–693, 2011.

[5] J. Alcalay. The value of mohs surgery for the treatment of nonmelanoma skin cancers. *Journal of Cutaneous and Aesthetic Surgery*, 5(1):1–3, 2012.

[6] J.M. Abide, F. Nahai, R.G. Bennett, et al. The meaning of surgical margins. *Plastic and reconstructive surgery*, 73(3):492–497, 1984.

[7] B.E. Pennington and D.J. Leffell. Mohs micrographic surgery: established uses and emerging trends. *Oncology*, 19(9):1165–1171, 2005.

[8] D.L. Shriner, D.K. McCoy, D.J. Goldberg, and R.F. Wagner. Mohs micrographic surgery. *Journal of the American Academy of Dermatology*, 39(1):79–97, 1998.

[9] UW Health. Mohs surgery sequence. `http://www.uwhealth.org/mohs/mohs-advantage/10633`. Accessed April 13,2012.

[10] J. Oliveira. A genetic algorithm with a quasi-local search for the job shop problem with recirculation. *Applied Soft Computing Technologies: The Challenge of Complexity*, pages 221–234, 2006.

[11] S. Bertel and J.C. Billaut. A genetic algorithm for an industrial multiprocessor flow shop scheduling problem with recirculation. *European Journal of Operational Research*, 159(3):651–662, 2004.

[12] T. Cayirli and E. Veral. Outpatient scheduling in health care: a review of literature. *Production and Operations Management*, 12(4):519–549, 2003.

[13] D. Gupta and B. Denton. Appointment scheduling in health care: Challenges and opportunities. *IIE transactions*, 40(9):800–819, 2008.

[14] R.E. Neale, M. Davis, N. Pandeya, D.C. Whiteman, A.C. Green, et al. Basal cell carcinoma on the trunk is associated with excessive sun exposure. *Journal of the American Academy of Dermatology*, 56(3):380–386, 2007.

[15] H.W. Lim and K. Cooper. The health impact of solar radiation and prevention strategies. *Journal of the American Academy of Dermatology*, 41(1):81–99, 1999.

[16] V. Madan, J.T. Lear, and R.M. Szeimies. Non-melanoma skin cancer. *The Lancet*, 375(9715):673–685, 2010.

[17] M.A. Mooney and D.M. Elston. Mohs micrographic surgery. `http://emedicine.medscape.com/article/1125510-overview`, May 2011. Accessed April 13,2012.

[18] J.D. Richmond and R.M. Davie. The significance of incomplete excision in patients with basal cell carcinoma. *British journal of plastic surgery*, 40(1):63–67, 1987.

[19] N.W.J. Smeets, G.A.M. Krekels, J.U. Ostertag, B.A.B. Essers, C.D. Dirksen, F.H.M. Nieman, and H.A.M. Neumann. Surgical excision vs mohs' micrographic surgery for basal-cell carcinoma of the face: randomised controlled trial. *The Lancet*, 364(9447):1766–1772, 2004.

[20] T.L. Bialy, J. Whalen, E. Veledar, D. Lafreniere, J. Spiro, T. Chartier, and S.C. Chen. Mohs micrographic surgery vs traditional surgical excision: a cost comparison analysis. *Archives of dermatology*, 140(6):736–742, 2004.

[21] J. Cook, J.A. Zitelli, et al. Mohs micrographic surgery: a cost analysis. *Journal of the American Academy of Dermatology*, 39(5 Pt 1):698–703, 1998.

[22] B.A.B. Essers, C.D. Dirksen, F.H.M. Nieman, N.W.J. Smeets, G.A.M. Krekels, M.H. Prins, and HA Neumann. Cost-effectiveness of mohs micrographic surgery vs surgical excision for basal cell carcinoma of the face. *Archives of dermatology*, 142(2):187–194, 2006.

[23] E.P. Tierney, C.W. Hanke, et al. Cost effectiveness of mohs micrographic surgery: review of the literature. *Journal of drugs in dermatology: JDD*, 8(10):914–922, 2009.

[24] R.P. Rapini. Pitfalls of mohs micrographic surgery. *Journal of the American Academy of Dermatology*, 22(4):681–686, 1990.

[25] P.G. Lang Jr, J.D. Osguthorpe, et al. Indications and limitations of mohs micrographic surgery. *Dermatologic clinics*, 7(4):627–644, 1989.

[26] C. Garcia, J. Holman, and E. Poletti. Mohs surgery: Commentaries and controversies. *International journal of dermatology*, 44(11):893–905, 2005.

[27] D.S. Gareau, Y.G. Patel, Y. Li, K.S. Nehal, B. Huang, and M. Rajadhyaksha. Multimodal confocal mosaicing of basal cell carcinomas in mohs surgical skin excisions. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 6431, pages 64310U–7, 2007.

[28] A.E. Rivera, J.M. Webb, and L.J. Cleaver. The webb and rivera (war) score: A preoperative mohs surgery assessment tool. *Archives of dermatology*, 148(2):206–210, 2012.

[29] S.J. Miller. Biology of basal cell carcinoma (part i). *Journal of the American Academy of Dermatology*, 24(1):1–13, 1991.

[30] R.S. Batra and L.C. Kelley. Predictors of extensive subclinical spread in nonmelanoma skin cancer treated with mohs micrographic surgery. *Archives of dermatology*, 138(8):1043–1051, 2002.

[31] R.S. Batra and L.C. Kelley. A risk scale for predicting extensive subclinical spread of nonmelanoma skin cancer. *Dermatologic Surgery*, 28(2):107–112, 2002.

[32] T. Cayirli, E. Veral, and H. Rosen. Assessment of patient classification in appointment system design. *Production and Operations Management*, 17(3):338–353, 2009.

[33] W.R. Lawrence and R. Rachel. Scheduling doctors' appointments: Optimal and empirically-based heuristic policies. *IIE Transactions*, 35(3):295–307, 2003.

[34] A. Turkcan, B. Zeng, K. Muthuraman, and M. Lawley. Sequential clinical scheduling with service criteria. *European Journal of Operational Research*, 214(3):780–795, 2011.

[35] S. Chakraborty, K. Muthuraman, and M. Lawley. Sequential clinical scheduling with patient no-shows and general service time distributions. *IIE Transactions*, 42(5):354–366, 2010.

[36] K. Muthuraman and M. Lawley. A stochastic overbooking model for outpatient clinical scheduling with no-shows. *IIE Transactions*, 40(9):820–837, 2008.

[37] K. Beer. Mohs surgery. `http://www.palmbeachcosmetic.com/mohs_surgery.html`. Accessed February 5, 2013.

[38] J. Rios and P. Lapinski. Mohs surgery. `http://www.rios-lapinski.com/mohs.cfm`. Accessed February 5, 2013.

[39] Washington University School of Medicine. Mohs micrographic surgery for skin cancer. `http://wuphysicians.wustl.edu/dept.aspx?pageID=4\&ID=21`. Accessed February 5, 2013.