

Capacity Allocation for Emergency Surgical
Scheduling with Multiple Priority Levels

by

Anisa Aubin

B. Sc. (Hons) Open University 2010

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTERS OF HEALTH SYSTEMS

TELFER SCHOOL OF MANAGEMENT

UNIVERSITY OF OTTAWA

SEPTEMBER 2012

©Anisa Aubin, Ottawa, Canada, 2012

Abstract

Research objective Emergency surgeries are serviced by three main forms of capacity: dedicated operating room time reserved for emergency surgeries, alternative (on call) capacity, and lastly, canceling of elective surgeries. The objective of this research is to model capacity implications of meeting wait time targets for multiple priority levels in the context of emergency surgeries.

Methodology Initial attempts to solve the capacity evaluation problem were made using a non-linear optimisation model, however, this model was intractable. A simulation model was then used to examine the trade-off between additional dedicated operating room capacity (and consequent idle capacity) versus increased re-scheduling of elective surgeries while keeping reserved time for emergency surgeries low. Considered performance measures include utilization of operating room time, elective re-scheduling, and wait times by priority class. Finally, the instantaneous utilization of different types of downstream beds is determined to aid in capacity planning.

Results The greatest number of patients seen within their respective wait time targets is achieved by a combination of additional on call capacity and a variation of the rule allowing low priority patients to utilize on call capacity. This also maintains lower cancelations of elective surgeries than the current situation.

Conclusions Although simulation does not provide an optimum solution it enables a comparison of different scenarios. This simulation model can determine appropriate capacity levels for servicing emergency patients of different priorities with different wait time targets.

Contents

Abstract	ii
Contents	vi
List of Tables	vi
List of Figures	vii
Acknowledgements	ix
1. Introduction	1
1.1. Research questions and objectives	8
2. Related literature	10
2.1. Surgical scheduling	10
2.2. Priority scheduling	17
2.3. Capacity planning	18
2.4. Gaps in literature	20
3. Methodology	22
3.1. Motivation for research - Non-Linear optimisation model	23
3.2. Simulation model	31
3.3. Advantages and limitations	44

4. Case study	46
4.1. The Ottawa Hospital	46
4.2. The Ottawa Hospital problem formulated	50
4.3. Current practice in emergency surgery scheduling at The Ottawa Hospital	53
4.4. Solving The Ottawa Hospital problem	55
4.5. Available data	57
5. Experimentation, analysis and results	65
5.1. Performance metrics	65
5.2. Scenario results	71
5.3. Verification	76
5.4. Validation	77
6. Discussion and Conclusion	85
6.1. Main findings	85
6.2. Conclusions and future studies	89
6.3. Glossary	93
6.4. Appendix	95
Bibliography	99

List of Tables

4.1 Attributes of data	62
4.2 OR capacity schedules	64
4.3 Factorial design of scenarios	64
5.1 Averaged daily utilisation of downstream capacities for each scenario	66
5.2 Dedicated capacity performance metrics for each scenario . .	75
5.3 On call performance metrics for each scenario	75
5.4 Canceled electives and P3s who exceeded their targets for each scenario	76
6.1 Summary table of reviewed literature	95

List of Figures

3.1 Patient flow map for the simulation model.	35
3.2 Patient flow map for the capacity of the OR	36
3.3 Animation layer of the model.	41
4.1 Arrival rates by day of the week	52
4.2 Flow map of the data from TOH	57
5.1 Instant utilisation of bed capacities.	67
5.2 Utilisation with combined regular and overtime.	67
5.3 Capacity utilisation by day of the week, in the base model	68
5.4 Capacity utilisation by hour of the day, in the base model.	69
5.5 P3 patients with relation to their wait time targets, in the base model.	70
5.6 Arrivals by priority and service type for 2 months of simulation data and 2 months historic data.	81
5.7 Waiting times for 2 months of the historic data.	82
5.8 Waiting times for P3 historic data and simulation.	82

5.9 Surgical durations compared from the model and TOH data	83
5.10 ICU and ward LOS stays by priority	84

Acknowledgements

I would like to especially thank Prof. J Patrick for his ongoing supervision and valuable feedback, and without whom this thesis would not have been complete. The staff at the TOH especially P. Doering, K. Woodhall, and J. Schubert for their continued support. Additionally, my family and friends for their consistent inspiration and encouragement, and putting up with me throughout, a special acknowledgement to my sister whose 'faith, hope and love' made it easier. All those who helped me with editing, you know who you are. I am deeply grateful. Thank you.

1. Introduction

A quality healthcare system is one which provides the appropriate care at the appropriate time to all individuals. It is often the experience of those attempting to access the healthcare system that although they receive the right care, it is not at what they would consider the “right time”. The chronic shortage of resources and/or inefficient management makes the efficient use of resources more important in order to address the ubiquitous nature of extended wait lists throughout the healthcare system. It is not too surprising therefore that there has been a significant rise in the use of operations research in healthcare management. With the rising costs of health on the minds of the public at large, individual taxpayers, patients and the government, it is essential that we use the available resources in the most efficient manner possible, whilst maintaining the high quality of healthcare Canadians have grown to expect.

The Parliament of Canada recognizes as part of the Canada Health Act that continued access to quality healthcare without financial or other barriers is critical to maintaining and improving the health and well-being of Canadians ((2)). Therefore, conforming to the Act would mean providing Canadians with access to healthcare in a timely manner. However, waiting times remain excessive. In 2005, the Supreme

Court of Canada ruled that the laws preventing private healthcare when long waiting times existed violated the Quebec Charter of Human Rights and Freedoms((1)). A recent survey conducted in 2011, stated that Canadians waited more than 3 or 4 weeks beyond what they would consider to be reasonable. According to this study, on average 2.8% of the Canadian population are waiting for a service at any given time. (This percentage was calculated by the total number of procedures and assuming that each person only waits for one procedure ((8))). In order to improve the quality of the Canadian healthcare system it is imperative that waiting times decrease, and that the use of private funds to purchase accelerated treatment does not become too tempting an option.

In November 2004 in Ontario, the Wait Time Target Strategy was announced by the Minister of Health and Long Term Care. Wait time targets represent time periods within which each category of patient ought to receive treatment based on accepted clinical practice. Wait time targets were set in five priority areas in which waiting times were seen to be especially challenging. These areas were cancer surgery, selected cardiac procedures, cataract surgery, hip and knee total joint replacements, and Magnetic Resonance Imaging (MRI)/ Computed Tomography (CT) scans. The purpose of these targets was to control the

maximum length of the waiting period for these areas. Incentives were then offered to those institutions which reported the targets, and as a result, reduced waiting times were reported ((30)). The Ontario government has now spent over \$614 million on this wait time strategy ((3)).

The practice of setting wait time targets continues to gain in popularity and it is now the mandate of each hospital in Ontario to ensure they meet these targets. Such targets exist for most surgical types within Canada as instructed by the Ontario Ministry of Health and Long-Term Care. Yet, it is not always clear how much capacity is necessary in order for this mandate to be achievable especially in instances such as emergency surgery, where demand is highly stochastic. In determining the necessary capacity required to meet specific waiting time targets, the nature of the demand and service times are important factors but the problem is further complicated when there exist multiple priority levels competing for a single shared resource.

This thesis focuses on patients requiring emergency surgery. These patients arrive without prior appointment and need to be serviced in a timely fashion. To ensure this timely service, wait time targets are set. Different targets are set for different priorities, and patients are assigned a priority depending on need.

In October 2007, the Government of Ontario expanded Ontario's wait times strategy to include time spent in the emergency department (ED). Incentives were provided to improve the performance of selected hospitals with high ED volumes and long ED wait times. An innovative pay-for-results program was created where hospitals had to reduce wait times for higher-acuity patients (Canadian Triage and Acuity Scale [CTAS] levels 1-3), reduce the number of patients with extremely long waits (greater than 24 hours) and ensure lower-acuity patients do not experience increased wait times. Patients categorized as CTAS level 1 require resuscitation, level 2 are emergent, level 3 are urgent, level 4 are less urgent and level 5 are non-urgent. Hospitals were also required to regularly track patient satisfaction and monitor quality of care ((28)).

Based on three demonstration pilots geared to improve ED flow, an ED performance improvement program was developed ((5)). New roles were developed for nurses working in the ED, some of which were related to the timely off-loading of ambulances as well as diagnosis and treatment off site in long-term care homes. Finally an Alternate Level of Care (ALC) plan with initiatives was developed to reduce the number of patients in hospitals who could be ALC patients ((5)). Projects have since been put in place, like the Chronic Disease Management Project, which aims to achieve fewer hospitalizations, better health outcomes,

increased efficiency and reduced use of the emergency department and duplicated services. These projects were put in place in an effort to help control the “demand” for ED services ((3)).

The Government of Ontario has set targets for the time spent in the ED until final treatment. Although there are currently only two wait time targets with respect to emergencies, one for complex cases and one for non-complex conditions, having targets displays significant development. These targets are not specifically linked to emergency surgery wait times, they exclusively measure the time patients wait in the emergency waiting room.

In the case of surgeries, elective surgeries have individual wait time targets published publicly. Elective surgery wait times are measured from the time that both surgeon and patient decide to proceed with surgery to the time that the procedure starts. It is noted that this should be adapted to reflect the initial recommendation for surgeon’s consultation as the start time, therefore capturing the wait time for the consultation as well.

In contrast, ED targets are strictly for the time spent in the emergency room (ER) being diagnosed, treated or waiting for admission to a hospital bed. These targets are 8 hours for complex conditions requiring more time for diagnosis, treatment or hospital bed admission,

and 4 hours for patients with minor or uncomplicated conditions requiring less time for diagnosis, treatment or observation. These times refer to the maximum time that 9 out of 10 patients should wait from the time they register to the time they leave the ED. The information collected on these targets compared with the baseline information collected in 2008 shows that although the targets are not yet reached, on all counts the values had improved ((4)). As of February 2012, 84% of the ED patients seen across Ontario, are within their targets for either complex or minor cases. The actual average time for all complex cases is 11.3 hours (target 8 hours) and for uncomplicated conditions is 4.5 hours (target 4 hours) ((4)). For the complex cases, data are divided by those patients who are admitted, and those who are not. Those admitted wait an average of 33.5 hours whereas patients who are not admitted wait 7.5 hours. This information is collected from 128 different hospital sites across Ontario who submit ED data to the National Ambulatory Care Reporting System (NACRS). The emergency department reporting system in Ontario also collects the information from over 20,000 ED visits at these hospitals. Information on their ED length of stay, ambulance off-load time, time to physician assessment in the ED, time to disposition decision and time to hospital admission is collected ((4)).

Few systems face more variability than healthcare. Even for planned procedures such as elective surgeries, it is not the demand that is fixed but the number of scheduled surgeries for a given day. The daily arrivals of new requests for surgery is a stochastic process. Similarly, very few medical procedures are deterministic in their actual service times. A CT scan scheduled for 15 minutes may take anywhere from 10 minutes to half an hour. This is the nature of healthcare where unexpected complications and delays are commonplace. Added complications occur when scheduling multiple priorities for a single service. This case is not just exclusive to surgeries but certainly applies there. There are a number of fixed resources involved in scheduling a surgery (such as operating room time) and a number of limited resources (such as staff, beds, equipment, etc.). The problem includes the allocation of patient surgeries to resources in a manner that is efficient for the hospital, convenient for the surgeons, and timely for the patients.

Patients are categorized by their requirements for surgery, level of urgency and order in which they arrive for service. Surgery schedules are generally composed of two main elements; a weekly schedule for elective surgeries, and blocks within that reserved for emergency surgeries. The focus of this thesis is on emergency surgeries and how

these can be serviced in their dedicated capacity within their respective wait time targets. Failing the availability of dedicated capacity emergency patients can be serviced alternatively through an available on call team. If neither of these options are feasible then there is the option of canceling an elective surgery or block. The challenge is to determine how much of each type of capacity to schedule for multiple priority emergency patients who need to be seen within different wait time periods.

1.1. Research questions and objectives.

The research questions that we seek to answer are: How can we mathematically determine the optimal capacity levels for treating emergency surgical patients with multiple priority classes within fixed wait time targets? Additionally, can we demonstrate the usefulness of our model by applying it to The Ottawa Hospital?

This thesis consists of six chapters which proceed as follows. A literature review of what has been researched with regards to surgical and priority scheduling and capacity planning can be found in chapter 2. Chapter 3 consists of methodology. The case study at The Ottawa Hospital is described in chapter 4. Experimentation, analysis and results are found in chapter 5. Discussion, conclusions and future studies

make up chapter 6. A glossary, appendix and bibliography follow these sections.

2. Related literature

The focus of this literature review is surgical and priority scheduling as well as capacity planning. Surgical scheduling involves setting surgical appointments both in advance and by appointment times. Priority scheduling looks at cases where different priorities of patients need to be scheduled simultaneously for a shared resource. Capacity planning applied in surgical cases differs in that it does not involve only a single resource but instead a number of resources for which the available capacity needs to be determined. Some of these resources needed for surgery include beds, surgical teams and surgeons, and operating rooms and other equipment. A decision has to be made as to which priority of patient gets the available capacity. The following literature is organized by the definition of the problem: surgical scheduling, priority scheduling, capacity planning and the methodologies used in each. A summary of those studied can be found in the Appendix (Page 95).

2.1. Surgical scheduling.

Surgical scheduling models dealing with advanced schedules have largely focused on resource utilisation rather than directly on capacity planning ((12)). Cardoen et al. ((12)) review methods of surgical scheduling and operation room planning. Literature they review are evaluated on numerous measures, some related either to the setting of

the problem or the technical features (such as methodology used) of the papers. They define planning as the process of reconciling supply and demand, thus dealing with capacity decisions. Their definition of scheduling is described as the defining sequence and time allocated to the activities of an operation. It is the construction of a detailed timetable that shows at what time or date each job should start and when each job should end ((37)). They state that less than half of the reviewed papers take into account downstream and upstream resources and facilities. Cardoen et al. question whether research should focus on realistic and complex problems in order to improve current practise or on examining simplistic versions of the problem that can be solved optimally.

The literature on surgical scheduling will be divided by the focus of the work; advanced scheduling, appointment scheduling and those applied to emergency surgical scheduling.

2.1.1. *Advanced scheduling.*

There are a number of papers that deal with the advanced scheduling of surgeries; answering the question of how far in advance to schedule each patient, and determining the allocation of capacity for surgeries by the end of their respective wait time targets. Wait times in the advanced schedule problem are measured as the number of days

until service, whereas in the appointment scheduling problem they refer to the time from patient arrival to scheduled service. The issue of advanced scheduling of patients to a service day after day received the most attention in the fields of outpatient scheduling and surgical scheduling. The majority of the work seeks to provide a block schedule, determining what specialty or surgeon operates on each day, that maximizes the availability downstream by reducing fluctuations in utilisation from day to day.

Blake ((11)) derives a quick and equitable master surgical schedule using an integer programming model which minimizes the undersupply of operating room time to each department. In this case, surgical departments were allocated blocks of time depending on the budgeted available room hours and nurse schedule of arrangements, assuring a consistent number of rooms per day.

Santibanez et al. ((36)) provide a mixed integer programming model that seeks to include demand related constraints both on bed availability and operating room capacity. They study the impact of simultaneously changing the master surgical schedule of multiple hospitals on throughput or the peak use of post-surgical resources. They restrict the amount of operating room blocks assigned to each surgical

specialty at each hospital between an upper and lower bound. Similarly, they state upper and lower throughput bounds for procedure types. What this research has shown is that an intelligent schedule can have a significant impact on reducing the variability in downstream resource consumption.

Together with an advanced schedule a more detailed appointment schedule is required in order to obtain a complete schedule for within each day of service. Within the literature this is classed as solving appointment scheduling problems.

2.1.2. Appointment scheduling.

Besides determining how far in advance to schedule patients, as summarized previously, the other problem is determining the order and the start times for patients previously scheduled for service today. This is referred to as the appointment scheduling problem. The majority of papers on surgical scheduling have primarily focused on determining the optimal start times and durations of appointments for a fixed number of patients. In this setting, the objective is to minimize some combination of idle time, patient wait time and overtime. Applications include outpatient and surgical scheduling.

Recently, Denton and Gupta ((17)) developed a two stage stochastic linear programming approach that also solves the appointment scheduling problem. Based on a large set of sample paths of potential service times, Denton and Gupta develop this linear program to set the start times for a fixed number of services based on a weighted average of idle time, wait time and overtime costs and propose a set of easy to implement heuristics. One result demonstrated by their model and that has been confirmed by others, is that provided services times are i.i.d. (independently and identically distributed) and wait time and idle time costs are the same for all appointments, the start times demonstrate a “dome” shape. That is, start times are bunched closer together at the beginning and the end of the day and are more spread out in the middle of the day. They also demonstrate numerically instances where scheduling based on mean length of service time is reasonable and highlight the importance of the variability of service time in setting an optimal schedule ((17)). Later they examine the effect that case sequencing has on patient waiting time, operating room idling time (surgeon waiting time) and operating room overtime ((18)).

Begen and Queyranne ((9)) provide a stochastic model that can determine the optimal appointment schedule (starting time of each appointment for a fixed number of appointments) as well as a methodology for determining the optimal schedule based on limited data. They develop an integer programming approach to the appointment scheduling problem. Initially they assume that the distribution of processing times are integer and follow a known discrete probability distribution. They demonstrate that an optimal solution can be found in polynomial time. They then relax the assumption of a known discrete probability distribution and determine the necessary sample size of surgical times to obtain provably near-optimal solutions with high probability.

Gul et al. ((24)) look at the impact of variable surgical times, using discrete event simulation to evaluate 12 different appointment schedules in an effort to minimize a weighted combination of wait times and overtime. Appointment times were determined by a recursion formula, known in the literature as job hedging ((39)). They then seek to improve these appointment schedules through a bi-criteria genetic algorithm. This paper expands into the area of advanced scheduling by allowing surgeries to be moved to another day.

2.1.3. *Emergency Surgical applications.*

From the reviewed literature limited research has been conducted in dealing with non-elective surgeries, those classified as urgent or emergent, those who can wait for service and those who need immediate attention. Emergency surgical schedules cannot be booked in the same way as described above due to the inability to anticipate emergency demand or to delay service by booking in advance.. Emergency surgery applications include those in which capacity is set aside for emergency surgeries to fit within the scheduled elective procedures given the available resources, multiple priorities, and the variable arrivals.

Gerchak et al. ((21)) look at surgical scheduling but rather than focusing on the block schedule they sought to determine the optimal number of elective surgeries to accept in a given day in order to ensure that sufficient capacity remains for emergency surgeries, assuming that elective patients arrive before emergencies. The trade-off is between revenue generated, overtime incurred and wait time costs. They demonstrate that a strict cut-off policy is not necessarily optimal as the number of elective surgeries to accept may increase with the number of elective surgeries yet to be served. What they do not do is determine the day-to-day impact of providing sufficient capacity for emergency surgeries or complications of multiple priority classes within the set of emergency surgeries.

Dexter ((19)) looks at making management decisions on the day of surgery with the intent of improving operating room efficiency and patient waiting times. This is achieved by using a list of ordered priorities, along with case durations determined by upper and lower prediction bounds.

2.2. Priority scheduling. The problem of scheduling is further complicated by the addition of multiple priorities of patients. A few papers look at this as part of their problem.

Dobson et al. ((20)) develop a stochastic model to determine the impact of reserving capacity for urgent outpatients on two service quality measures. These measures are the number of urgent patients (who need same day service) that are not handled during normal hours and the average length of the queue of non-urgent or routine patients (those who can wait for service). These service measures along with revenue are accounted for as numerical experiments optimize the performance of the system. They do not differentiate between the different types of appointments, the variation in service times and queuing dynamics.

Patrick et al. ((34)) use approximate dynamic programming to determine a policy for scheduling multiple priority classes to a single resource where each class is characterized by a specific wait time target. Patients are scheduled into a booking window of available treatment

days. They provide a solution that is easily translated into a readily implemented policy and derive bounds on the necessary capacity for this policy to function well. A potential application to surgery would be within block scheduling though stochastic service times would need to be taken into account.

2.3. Capacity planning.

There has been extensive research conducted in the field of capacity planning, not only in healthcare but also in other domains. Capacity planning in the advance scheduling problem answers questions such as, how much capacity is needed to meet the wait time targets for stochastic demand? In the appointment scheduling problem it is how much capacity is required to service the n patients scheduled for today? These questions become increasingly important as the type of surgery we are dealing with is emergency. Looking at the literature there is a gap when it comes to this problem. This problem has the characteristics of both advanced and appointment scheduling problems and yet it is neither one, but rather a capacity planning challenge where these emergency surgery patients are held in a queue until a decision is made to release them. The decision to release is based on different priority patients waiting different periods of time to be seen and the immediate available capacity.

Ayvaz and Huh ((7)) look at the allocation of a limited capacity of resources amongst several patient types. The different patient classes display differing reactions to delays in service. They use a dynamic programming approach, common in the inventory management literature. There are penalties incurred for patients lost in the system due to a high priority patient not been seen immediately (i.e. emergency patients) and a second penalty for the time a patient (of a lower priority, i.e. elective patient) waits in the system. A simple threshold heuristic is proposed protecting some amount of capacity for emergency patients in each period, and more if the elective backlog is cleared.

Green et al. ((23)) determine a policy through a Markov decision process (MDP) model for diagnostic imaging where there is a known number of booked appointments (outpatients) and the decision is whether to serve a scheduled outpatient or whether to give the current service slot to a waiting emergency patient or an inpatient. Decisions are made based on the number of waiting outpatients, inpatients and emergency patients and the optimal policy is found through formulating this as a dynamic program. They demonstrate that the optimal policy lies in the set of “monotone switching curve policies” where outpatients are served as long as the number of waiting outpatients exceeds a certain threshold and where that threshold is a function of the

number of waiting inpatients. As the number of waiting inpatients increases, the threshold naturally increases as well.

Problems where multiple resources are consumed in sequence are understandably quite complex and therefore difficult to model. Often the only methodology possible is simulation or a very stylized queuing model. Some attempts made in this area include the work by Oddeye et al. ((33)) who formulate a goal programming model with simulation in order to achieve optimal clinical work flow by eliminating bottlenecks. The model was successful in determining the impact of disruptions.

Creemers ((16)) uses a bulk serving queueing method to feed inputs into an optimisation model determining the expected waiting time of patients of different classes and then assigning server time slots to the different classes of patients by minimizing patient waiting times, varying by patient class. For large problems, a step-wise heuristic is developed. This formulation can be applied to the allocation of operating room time slots over different medical services in the hospital.

2.4. Gaps in literature.

The research on scheduling both in general and in emergency surgeries as well as priority scheduling and capacity planning cover many problems including setting appointment start times, durations of appointments, downstream capacities and optimum panel sizes. There is

an absence of literature that deals with multiple priority classes and multiple methods of services. There is a gap when looking at the meeting of wait time targets for different priority classes with multiple methods of service, combined with the effect on the downstream capacities within these cases. This research contributes to surgical scheduling and looks at the capacity requirements in order to meet pre-specified wait time targets. More specifically it adds to the few reviewed papers that look at the scheduling of emergency surgeries. A mathematical model can be used for emergency surgery scheduling to analyze the different required capacities to meet the demands of different priority classes within their respective set wait time targets.

3. Methodology

Capacity requirements need to be determined in order to meet wait time targets for emergency surgery scheduling. Initially an optimisation model was developed that proved intractable leading to the choice of simulation as a methodology.

Simulation models take on characteristics depending on how a system under investigation is conceptualized. Simulations are either dynamic or static depending on how they evolve over time, and either stochastic or deterministic depending on the random nature of the important behavior within the system. Additionally a simulation model can be continuous or discrete depending on how the system changes over time at discrete intervals/events or continuously over the studied period. Often a model may be a combination of both discrete and continuous.

The simulation model chosen for evaluating the emergency capacity planning problem can be classified as dynamic, stochastic and discrete. It is dynamic as the problem is changing over time, stochastic as there are random distributions used and discrete as it changes over time at discrete events, i.e. the simulation clock moves forward to each subsequent event occurring at a fixed point in time.

The initial attempt is described below, its purpose, assumption, the model, and the application to emergency surgery. Following that the simulation is described in general detail, the purpose, components, assumptions, and experimental design. The chapter ends with the advantages and limitations of the simulation methodology.

3.1. Motivation for research - Non-Linear optimisation model.

Initial attempts to solve this problem included an optimisation model to minimize costs associated with regular and overtime.

3.1.1. Purpose.

This model was chosen to obtain a single optimal solution. A desired solution determines capacity requirements with minimized costs while also meeting wait time targets. It is generally assumed that wait time targets are set for clinical purposes. This is certainly the situation in the studied case. While this is true, there are also managerial benefits to having a wait time target. Fundamentally to this research, the ability to smooth out demand over a number of time periods reduces the amount of capacity required in order to meet a given level of demand. In theory, given a base capacity, smoothing out demand over more days should reduce the amount of overtime required. The initial research question was set in order to determine one solution. The first

question to address is, what is the minimum wait time target that allows the user to maintain average daily overtime below a threshold? This question becomes more complicated as multiple priority classes compete for a single resource. The single objective is to choose wait time targets for each priority in such a way that the average daily overtime remains below a specified capacity level while also meeting clinical targets.

Early attempts to answer this question were to consider the challenge of simultaneously determining the best wait time targets and the regular capacity required to meet those targets. This problem was formulated by means of a non-linear optimisation model outlined below.

The problem was defined as multiple (i) priority classes competing for a single resource, where each priority class has its own wait time target. The challenge is to determine both the best wait time targets (T_1, \dots, T_I) for each priority and also the appropriate regular hour capacity to meet the demand. Determining the best choice of wait time targets is linked to the choice of capacity as they combine to determine the expected overtime cost.

3.1.2. *Assumptions.*

We will assume a simple booking policy that books the highest priority clients as soon as possible and books lower priority patients

into the day with the smallest amount of booked time that is within their wait time target. If no capacity is available, then the congestion in the system is relieved by servicing patients of the highest priority who have waited the duration of their wait time target through overtime.

We also assume that there exists a clinical maximum on the wait time target (C_i for priority $i \in I$) that cannot be exceeded. The point at which there is no longer any financial advantage to extending the wait time targets is the point we desire to choose. The new wait time target must be no more than the previous clinically set targets, determined by management. This is ensured by constraint (3.4).

We assume that overtime occurs independently to regular time. Constraint (3.3) determines sufficient capacity of overtime in order to meet the demand for each priority level with their respective wait time target. This is calculated by ensuring that demand that has to be seen that exceeds the regular capacity is seen during overtime. The daily regular capacity is compared to an option where one day is added to one wait time target, finding a solution in which this is less than the value of pushing a patient back by one day. The constraints of the model will determine if going into overtime, or extending one of the wait time targets by one unit is a more optimum decision (3.5 and 3.6).

Finally, it is assumed that there is no related cost incurred for idle time. Overall costs are determined by combining those associated with running regular and overtime capacity - overtime occurring when there is not enough regular capacity set and the allocated time is exceeded. Idle time occurs when too much regular capacity is allocated and is not used due to a lack of demand.

3.1.3. *Non-linear optimisation model.*

The objective is to minimize the cost of using different capacity. A schedule without some overtime is not running efficiently as there will be excessive idle time. A small amount of overtime is beneficial when optimising the scheduled time. u represents the capacity of regular time and C^R is the associated cost with one unit of regular time. Likewise v represents the capacity of overtime, and C^{OT} the associated unit cost. The following is the original model formulation created in order to determine the emergency surgical wait time targets and capacity requirements to meet them. The objective is a minimization of the total cost of regular and overtime.

Let X_{it} be the demand from priority class i on day t . If service times can be assumed to be deterministic then this can simply represent the number of requests for service by priority i patients on day t . If service

times are stochastic then this represents the total time of all requests from priority class i on day t . In addition, let

$$(3.1) \quad y_{iT} = \sum_{i=1}^T X_{it}$$

represent the total demand from priority class i over a period of T days.

Let δ be defined as the cost of increasing the wait time target one additional day. If $\delta_k = 0$ then the optimal solution is to let $T_i = C_i$ as each additional day added to the wait time within the clinical target reduces the overtime cost. However, the benefit of increasing the wait time target diminishes the further the wait time target is delayed and thus at some point the reduction in cost is insufficient to warrant the delay in service.

The model follows

$$(3.2) \quad \min_{u,v,\vec{T}} C^R u + C^{OT} v$$

subject to

$$(3.3) \quad E\left[\sum_{i=1}^I \frac{y_i T_i}{T_i} - u\right]^+ < v$$

$$(3.4) \quad T_i \leq C_i \quad \forall i \in \{1, \dots, I\}$$

$$(3.5) \quad E\left[\sum_{i=1}^I \frac{y_i T_i}{T_i} - u\right]^+ - E\left[\sum_{i=1}^I \frac{y_i T_i^*}{T_i^*} - u\right]^+ > \delta_k \quad \forall k \in \{1, \dots, I\}$$

where

$$(3.6) \quad T_i^* = \begin{cases} T_i, & \text{if } i \neq k; \\ T_i + 1, & \text{if } i = k. \end{cases}$$

3.1.4. *Application to emergency surgery.*

The complexity of this problem means that at this point it will not be possible to solve this model for emergency surgical cases. This formulated problem objective is to minimize the total costs associated with the scheduling of multiple priority classes using a balance of regular and overtime, which is an over simplified version of the problem. In the studied case this division of time is not used. Capacity is divided into dedicated, regular capacity that has the ability to run into

overtime, on call capacity, that is overtime which can be used at any time, and elective capacity, which is overtime that can be used only during the ‘regular’ time capacity. The definition of overtime here is problematic when it comes to emergency surgeries as it violates the assumption that overtime occurs independently of regular time.

The optimum cost value is obtained by balancing the capacity between regular and overtime which in the emergency context are not linearly independent. It is the assumption about the independence of the regular and overtime set out in this model that is violated in the emergency surgery case. Overtime occurs in the case that regular capacity is not available and a patient’s wait time target has been exceeded. In the case of elective surgeries, average overtime would be calculated as the time the OR rooms are in use beyond the scheduled closing time of the rooms. This, however, is not the case for emergency surgeries. A high priority patient who exceeds their wait time may be seen within regular capacity and yet it would be considered overtime in the case of that patient if additional capacity is used. The stochastic nature of the problem and the added complexity of multiple priorities being serviced by a single resource require additional modifications which cannot be made at this time. Removing the assumption would change the objective of the model, and adding an additional constraint

to determine when this was the case would mean adding a non-linear function, making the model non-linear and much harder to solve, if possible.

An example of why the non-linear model cannot be solved for emergency surgery follows. If a high priority patient needs surgery when all capacity is occupied, and no current case finishes within the wait time target of that priority patient, then the newly arrived patient will be serviced using overtime capacity, i.e., on call capacity as currently used in hospitals. In this case of assigning emergency cases within the day, regular time capacity would be used but the associated costs would need to be that of using overtime. In addition to this overlap in the definition of time the units of capacity could, in many cases, be shorter than those of service times, that is, the time that it takes to operate on an emergency patient would not always fit within the capacity allocated by the optimisation model.

In conclusion, this model will not be solved at this time. The primary focus of the research is to answer the question of what capacity is required for the planning of emergency surgeries and therefore other methods can be used. The initial optimisation model cannot be validated for use in the emergency surgery scheduling and an alternative methodology that does not explicitly give an exact optimum solution

will be used in line with the complexity of the problem. A final complication is that capacity is not simply either regular or overtime, but within that there are different categories according to the specific resources used and the time of day. For these reasons the optimisation methodology can not be extended to the scheduling of emergency cases. Thus the chosen methodology will be outlined below.

For evaluating this problem a simulation model is used to model the reality of emergency patients through the surgical department. The simulation can provide utilisation rates of different capacity allocations as well as downstream capacities for each, which this model was not able to consider. Confidence intervals with the results from the simulation can be used in order to determine a favourable amount of capacity needed to service most patients within their wait time targets.

3.2. Simulation model. Simulation is perhaps one of the most commonly used analytical tools today. In principle it is not a complicated model process, and is often used when it is impossible or inconvenient to tackle a problem in a more analytical way. In simulation, a model of the system under investigation is designed, a computer program is written embodying this design and a computer is used to imitate the system's behavior under a variety of operating policies so that the most favourable one can be chosen. Simulated scenarios can

be carried out on the computer at a faster rate than in reality and much more cost effectively. Computer simulation is a method in which a structured design of experimentation can be conducted so that only when the best results match the desired outcome would the model be implemented in the real system. Realistic simulation models require long computer programs of increased complexity, resulting in quite a time-consuming creation process. Simulation can be used where it is not feasible for some reason to conduct direct experimentation or develop an analytical model.

Simulation is used for a number of reasons, one being that it is possible to simulate any situation, and the other that there is a wide selection of high-quality computer software to aid in the rapid development of a valid computer simulation. For this model, the chosen software is Arena v10. Arena is a strongly supported language and is used in both academic and industrial fields. Arena is competitive in the simulation marketplace, and is fundamentally a process description-based language, where aspects of the model are described as the entity, in this case patients, would experience them ((35)).

For this research a simulation model is created for evaluating emergency surgery capacity planning. The base time units of the model are minutes. A policy is evaluated by comparing the resulting wait

times and capacity utilisation across different scenarios. The goal of the simulation model is to detect the capacity levels for scheduling emergency surgeries in such a way that decreased alternative capacity is utilised, and patients' wait time targets are met. The problem is how to schedule emergency surgeries, and what capacities (how much and of what type) are needed in order to meet the wait time targets for different priority classes. Patients enter the system, and for high priority patients they wait in a queue for dedicated reserved capacity. If their wait time targets are reached, they are seen by another capacity either on call or by canceling elective capacity. Lower priority patients wait for the dedicated capacity to become available regardless of their wait time target. This rule is to reflect circumstances where on call capacity or canceling an elective patient would not be utilised for a low priority patient, and is a common approach amongst hospitals with the exception of a worsening condition which would also trigger a change in priority. By running the simulation repeatedly and using set performance measures the amount of appropriate capacity needed can be discussed as different scenarios are tested.

The policy to be used in determining the order of service from the waiting room is that of a priority queue, where each client is served closest to their target time. Patients arrive and if there is no available

capacity they wait in a queue. There are no reservations, that is, as soon as capacity is available the patients in the queue in order of highest priority can use it. If the wait time target is reached for one of the higher priority patients in the queue, that patient leaves the queue and is serviced by use of alternative capacity. This could be in the form of an on call team, canceling an elective patient, rescheduling elective surgeries or opening an additional room with additional staff. In case an additional on call surgeon is required, and the surgical type is one that is currently being used in the on call, a team and perhaps room would need to be made available and a scheduled elective surgeon would be used. This final option would still require an additional surgeon so this case is not regarded as an additional option but rather the same as canceling an elective. This scenario would mean at least delaying an elective surgery, if not completely rescheduling it.

3.2.1. *Purpose.*

The purpose of the model is to create a representation of the emergency surgery schedule and the Operating Rooms (OR) using collected historical data. This is achieved by creating patients with attributes and capturing their flow through the system in a realistic manner. The characteristics of this simulation can then be used to evaluate policy

changes. Below are the steps taken to ensure that the simulation represents reality as accurately as possible. The following section consists of patient flow and queues, the details of the resources used, and schedules.

3.2.2. Patient flow and queues.

Patient flow can be defined as the movement of a patient seeking treatment through the processes ((35)). Depending on the patients severity and the ability of the resource to service each patient, i.e. the duration of the surgery, the number of patients seen in the available capacity duration will fluctuate ((35)).

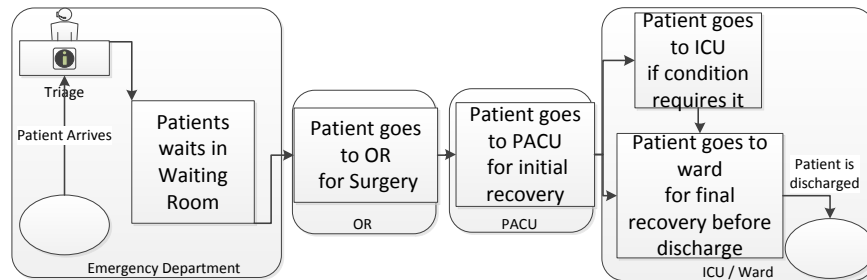


FIGURE 3.1. Patient flow map for the simulation model.

Upon arrival and after receiving a generic bed, patients are placed in a queue for surgery (Figure 3.2 on page 36). They leave this queue when either the dedicated emergency capacity of the OR is available, or if the patient is sufficiently urgent, when their wait time target is

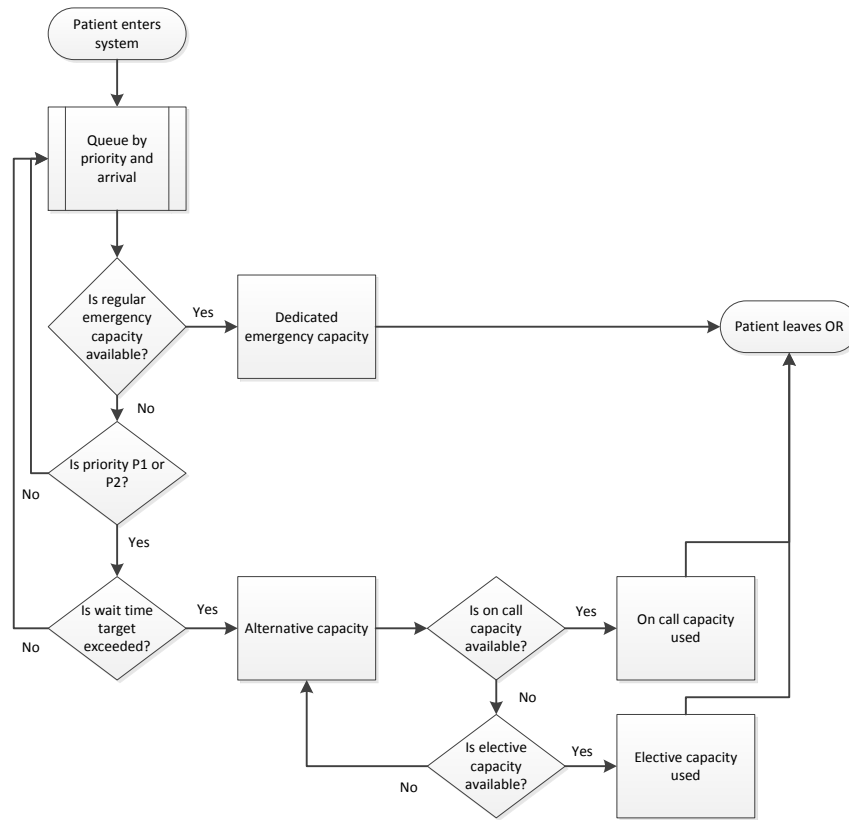


FIGURE 3.2. Patient flow map for the capacity of the OR

reached. The lowest priority patients remain waiting for dedicated capacity and do not leave the queue even when their wait time target is reached. This is to reflect the reality that an on call team would not be used to service a low priority patient even if the target wait time is reached. The only other case in which a high priority patient would remain in the wait room beyond their target is in the event that both the dedicated capacity and the canceling of an elective capacity

are unavailable, for instance during the night, and the on call capacity is currently in use. In this situation the patient waits for the on call capacity as there is no alternative available. The process described here is fairly typical of how hospitals service emergency surgical patients.

After surgery, all patients go to Post-Anesthesia Care Unit (PACU), after that either Intensive Care Unit (ICU) then a ward, or directly to the ward for final recovery before being discharged. Those patients who use the ICU between PACU and ward beds can be determined from data as a representative fraction for each priority.

3.2.3. *Model Resources.*

A resource is a service consumed by an entity that constrains the flow of entities in the system, i.e., with patients being entities in this case, resources are services provided by the hospital throughout the patients' stay, namely OR rooms, surgeons and beds ((35)).

For each of the three capacity types (dedicated, on call and elective) there are different quantities of resources available according to a schedule and the time of day. If there is more than one room available for dedicated OR (usually in the afternoon on a weekday), then multiple patients can simultaneously use the dedicated OR capacity.

Once attributes are assigned to patients, patients are given a bed. Beds are also resources, they are considered busy if occupied. The purpose of the model is to determine the required capacity to meet the wait time targets of emergency surgeries and also capture an accurate reading of the instantaneous number of beds utilised within each available capacity. The beds in the model are specifically designed to avoid a bottleneck in order to ascertain properly required capacity and so are uncapacitated, indicating to the hospital the number of beds that would be required in order to service the demand.

Upon arrival, patients occupy a general bed which is not released until they enter the operating room. Generic beds are released before the patient goes into the OR. At each change in designated area (PACU, ICU and ward) the type of bed changes with the patient. Records of the hourly usage of each type of bed on each specific day of the week are captured.

3.2.4. *Schedules.*

Schedules are easily created in Arena. Schedules are used for the creation of patients, or the schedule at which they arrive for service, and there are also schedules for the availability of resources such as the surgeons and rooms.

Patients enter the model according to a schedule based on the arrival data. That is, patients are divided according to their priority and also by the day of the week and the hour of the day when their first encounter occurred. Distributions are fitted separately for the duration of the surgical times for each priority and each service.

Schedules are made for each type of capacity and include available durations of those capacities. Once a schedule is created a resource can be assigned to that schedule, as such a room that is assigned to a dedicated OR schedule would open and close according to the schedule. Schedules created in this model were for the most part a week in length and the default is for the schedule to repeat itself upon completion unless indicated otherwise.

Below are the schedules for the base model. The schedule rule is to *ignore* when OR room is scheduled to close and a patient is in the process of being served. The rule states that the patient will finish with that resource before the room closes. The following schedule is not delayed by this rule so the next day will start as scheduled.

3.2.5. *Experimental design.*

The algorithm that was used for this model is as follows. First, set the available capacity and schedule, then simulate the policy for 90 days, collecting performance statistics after a warm up period of

30 days. Replicate this 10 times and from the results determine the average, standard deviation and 95% confidence intervals for the performance measures. Thus to measure the system the percentage of patients who are seen outside their wait time targets can be collected as a performance measure. Other performance measures are the usage of the alternative capacities, as well as the combination of regular and overtime capacity and utilisation of the downstream capacities. Overtime capacity is considered as the time after close that a dedicated room would remain in use for the completion of surgery. This would occur in the event that a patient was not able to have their surgery completed by the scheduled close time of the dedicated capacity.

3.2.6. *Animation and performance measures.*

The model consists of two layers or parts. There is the programmed part of the simulation model within which all the rules and processes are contained. The top layer or separate view of the same model consists of the animation of the model (Figure 3.3 on page 41) along with performance measures and details of the system. Animation is important for helping to verify that the model is working as intended and for validating the model's representation of the real system ((35)).

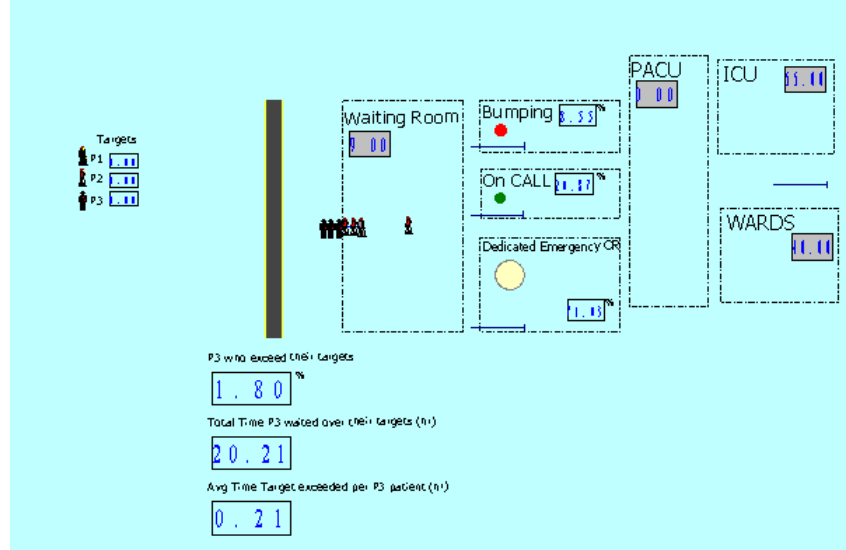


FIGURE 3.3. Animation layer of the model.

An animation was created for presenting the model, to demonstrate the purpose and provide an overview of the way the model works without showing all the extensive workings of the model. This animation includes only the process modules of the model (Figure 3.3 on page 41). Utilisation in progress is graphed for each capacity as seen in the lower part of (Figure 3.3 on page 41) with 0 being available and 1 being in use. Performance measures were also displayed as part of the animation. These were the percentage of patients that used each type of OR capacity, and the percentage of patients who exceeded their wait time targets. Calculations are performed by counting the total number of

patients as well as collecting time stamps associated with each priority when patients enter and leave the OR. Percentages can be measured both in the number of cases and in the total time taken. These measures can be compared to determine which policy changes should be implemented and the effect that each change in capacity has on the different performance measures. From the different scenario results a superior policy, indicating what capacity should be made available in order to meet the set wait time targets for the emergency surgeries at TOH is determined.

3.2.7. Recording from the model.

Modules were used at five different locations in the model to calculate information and results were written to excel files to allow the calculation of instantaneous utilisation. A module in the simulation model was created to capture data on an hourly basis. That is, on every hour the module recorded the hour of day, day of week, and number and type of each bed currently in use. This information was exported to an excel sheet, and from here distributions and proportions were calculated, and used to evaluate the different scenarios.

Waiting times were calculated from the model as the difference between when the surgery begins and the time the entity was created, i.e., started waiting. By counting the number of patients and the time

(hour during the day, as well as the day of the week), the percentage of cases performed outside of the dedicated emergency OR schedule can also be analyzed.

Patients of each priority were counted at different points in the model: as they leave the system, and as it is determined if their wait time is above or below their wait time target. The percentage of these patients is obtained and used as one of the performance measures. This metric is calculated by summing the total number of entities above the target, and counting the total number of patients to determine the percentage of those who are not serviced within their target. Patients were given an attribute after they leave their respective OR capacity in order to determine the denominator for this metric. Different attributes were given for each of the capacities and the percentages of patients who use each capacity were also calculated.

The priority of each patient was recorded to each excel sheet. In addition, recordings were made of what capacity serviced each patient in order to calculate the percentage of patients that utilised each type of OR capacity. This is given as a fraction of the patients that enter each capacity over of the number of patients that have left the waiting room.

Many simulations are conducted in order to compare policies or system configurations for the operating system. The controllable inputs which are thought to be the cause of the response from the system are called factors. These factors can be quantitative (i.e., number of available rooms or schedule) or qualitative (i.e., policy change). There may be a number of factors that are thought to interact and jointly affect the behavior of the system, requiring many runs of the model ((35)). Factors in this model include available capacity and the duration of this capacity, as well as the rule regarding the action that low priority patients take at the end of their wait time targets.

3.3. Advantages and limitations.

Simulation is a methodology that can provide a guide of what results could be expected from different scenarios without an exact optimum solution. The problem of determining emergency capacity required to meet wait time targets is a sufficiently complex problem that it cannot be computed exactly using a deterministic algorithm. The methodology chosen to aid in finding an educated approximation is a simulation. The method includes defining a domain of inputs and parameters, and with these a random distribution can be used to create patients to simulate the effects of changes made in the system under investigation. Multiple replications of the simulation are run in order to

get more accurate results. Initially, a base model representing current practise was created and once it was verified and validated, alternative scenarios were evaluated and compared to the base results using predetermined performance metrics in order to determine the best or most favourable scenario.

Simulation has the advantage of being flexible enough to be applied in any case study. For each case, different parameters need to be determined. As a limit to the methodology there is no optimum solution, so although numerous scenarios it is always possible that the optimal solution simply was not tested.

4. Case study

In order to apply the simulation model to a real scenario as well as to validate the model, a case study was conducted at The Ottawa Hospital.

4.1. The Ottawa Hospital.

TOH is a major teaching hospital with 1195 beds offering primary (general), secondary (specialized) and tertiary (most advanced) health-care services to adults and newborns of the National Capital and Eastern District area. The hospital consists of 3 campuses: The General, the Civic, and Riverside. The General includes the Ottawa Rehabilitation Center, the Cancer Center, and the Eye Institute. The Civic specializes in cardiac care and research having developed one of Canada's leading medical research programs. Riverside is a day facility specially for outpatients and specialty clinics. ((6))

Emergency surgeries only occur at the former two campuses. The Civic and General surgical suites are each comprised of 16 operating rooms and an emergency department. Currently these campuses run on three month block schedules comprised of a master schedule of four weeks. The block schedules come out in April, July, October and January. Usually on any given day not all the available rooms are scheduled. Elective surgery specialties are usually assigned a room for the whole

day. For emergency surgeries there are some rooms reserved on this master schedule. Previously The Ottawa Hospital was reserving eight hours on three days a week for emergencies. At this capacity level, there was not enough dedicated capacity to handle the emergency surgeries, thus the hospital management arbitrarily changed the capacity to five days a week of seven hours a day with the additional rooms on each Thursday and Friday compared to the previous schedule. Currently the hospital has 236 priority hours available per week at both the Civic and General. This includes an additional room primarily for non-ortho (Orthopedics) urgent but the first 3 hours of this room each weekday are dedicated to General Surgery. On occasion ortho may use the non-ortho, but the primary purpose for this time was to clear the backlog of waiting ortho patients. However, in some cases more urgent patients take priority. On the weekend the hospital has an emergency surgery room available from 8am to 3:30pm with the aim of closing the room at 4pm. An evening call back team is available everyday. Bringing one of these teams in for service means that overtime costs are incurred. Nurses work on a rotation and after their 8 hour shift they are on call for an additional 8 hours. Thus, there is always a team available on call 24 hours a day seven days a week.

4.1.1. *Wait time targets.*

There are three classes of priorities of emergency patients at The Ottawa Hospital (TOH); P1, those who need immediate service, P2, urgent cases that can wait for a short while, usually need to be seen the same day, and P3 the least urgent of the emergency cases. Depending on the condition of P3 patient, an undefined sub class of P3 are, in fact, not kept at the hospital but summoned back when capacity is available ideally within a week to 10 days. These are referred to as “walking wounded”. For each of the priorities the hospital has assigned wait time targets as follows: P1 - 2 hours, P2 - 8 hours, and P3 - 48 hours.

Emergency surgery wait time targets are set for each class or priority of patient and these values represent assigned time periods for each priority, within which a patient of that priority should be serviced. Once a patient is assigned a priority, and therefore a wait time target, they should not, in theory, wait beyond that target for service. However, in the event that the priority is changed due to a worsening medical condition the target and priority may be reduced. The other exception where targets change are in cases where the hospital has financial incentives associated with meeting certain priority patients, as mandated by the Government. On such occasions a certain type of patient (often a P3) whose wait times are collected for monitoring

(usually either a cancer patient or a hip and knee total joint replacement whose data are required by the Ontario Ministry of Health and Long-Term Care) will change to a higher priority class to ensure that their target is met. This changing of priority means that patients who are seen as P2's would have already waited almost 40 hours as a P3 before having their wait time target changed to 8 hours thus ensuring that the original P3 type patient has not exceeded the 48 hour wait time target. As there is no way to predict this behavior, and no data from the hospital to differentiate the P2 cases in which this has happened, we have assumed that this is not a contributing factor in the decision of the length of wait time target or capacity. It is assumed in the model that patients do not change priorities.

Emergency patients are considered those patients who use the operating rooms during allotted time for emergency surgery. Most of the time these patients would enter the hospital through the emergency department though in some instances an inpatient may also require emergency surgery. Some patients may have had an elective surgery scheduled for a future date, but due to unforeseen circumstances they arrive to the emergency department prior to their scheduled time, and use capacity reserved for emergency surgery. In such a case the patient would be considered an emergency patient and would have an

assigned emergency priority and associated wait time target. Patients who arrive in the emergency department are assigned a priority based on the severity of the case, as well as the type of injury and procedure required. Certain surgical procedures are always considered P1, other cases which may not ordinarily be of highest priority may be assigned a higher priority if there are additional complications or where large quantities of blood are lost or conditions are dire. Patients assigned a priority P3 have certain emergency needs but, by TOH standards, can wait longer for their service.

4.2. The Ottawa Hospital problem formulated.

The simulation is populated based on historic data from TOH. Information about the OR system and how it operates was collected from various sources, primarily TOH data and personnel working in their OR. The dynamics and rules of the model were defined and tested before the simulation results could be analyzed. The model used was composed of the data from the Civic site of TOH. The model for the General site would function as an individual model because patients are rarely transferred between sites and even less frequently emergency surgery patients.

Patients are created as emergency entities differentiated by priority. Patients are classified with attributes that make them similar to

the cases that would constitute the demand for emergency surgery. Over 500 primary procedures are grouped into the 15 services (Anesthesiology, Dental, Ear nose and throat, General surgery, Gynecology, Neurosurgery, Ophthalmology, Orthopedics, Other, Plastic, Radiology, Spinal, Thoracic, Urology, and Vascular). However, not all services occur at each site: due to resources, some are exclusively at one physical site. Anesthesiology only occurs at the General site, whereas Neurosurgery, Radiology and Spinal only occur at the Civic site of TOH. Patients are assigned a service type by a percentage that is calculated from the historical data depending on the proportion of patients that make up each service. The same percentage was used to assign services to the simulated patients (Appendix 6.4.1 on page 95). Within each service there are associated probability distributions of surgical times and a combination of setup/teardown times. These represent the duration it takes to get the OR ready for the patient and the patient in the OR ready for the surgeon, and teardown likewise refers to the time it takes to get the patient out after surgery and to prepare the room or equipment in readiness for the next patient. The setup and teardown times in the model are based on the percentage of patients in the data with procedures of corresponding setup and teardown combinations within each specific service type (Appendix 6.4.2 on page 96).

It is assumed that patient arrivals are only significantly different between weekdays and weekends. A total number of patient arrivals by time and day over the year was broken down into a cyclical weekly schedule of arrival rates. From analyzing the data the days of the week did not differ considerably from each other (See figure 4.1 where 1 represents Sunday), so only 2 schedules were created, one for the weekdays, and the other for weekends. To determine the number of hourly arrivals, the total weekday arrivals were summed over the year and then divided by 261 (the total number of weekdays for the captured data). For the weekend arrival rates the total weekend arrivals were divided by 104 (the total number of weekend days for the data period) to derive the hourly rate of weekend patients.

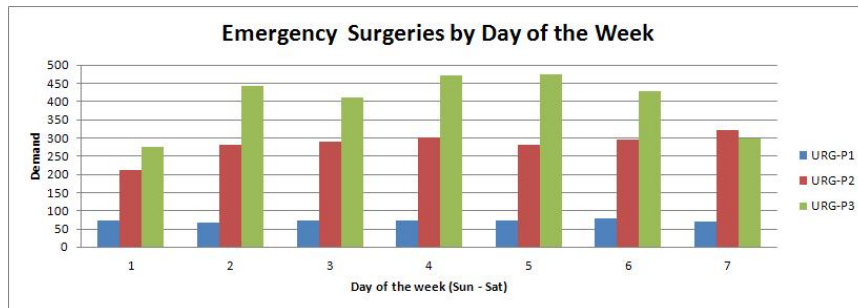


FIGURE 4.1. Arrival rates by day of the week

4.2.1. Patient flow.

Patient flow refers to how the patients move through the system. When patients arrive in the emergency room, they are triaged and

assigned a priority. They get a bed and wait in a queue until a surgeon and OR are available. There are different types of capacity for emergency surgery, these are dedicated OR, on call teams using unscheduled rooms, or canceling elective capacity. ‘Dedicated’ refers to the time set aside for emergency surgeries on the schedule. On call is a team available to come in to perform emergency surgeries if needed, and canceling elective capacity refers to the process of taking a room which was scheduled for an elective patient. This often occurs when that specific type of surgeon is needed. This is the order in which these options are exhausted. After surgery, patients use a PACU bed for the duration of their initial recovery. Depending on the specifics of their condition they next proceed to ICU and then to the ward (at each capacity using a specific bed type) or they proceed directly from PACU to the ward before they are discharged. The above description follows the generic description in the previous chapter quite closely.

4.3. Current practice in emergency surgery scheduling at The Ottawa Hospital.

The scheduling of emergency patients is a complex procedure. Often patients have prerequisites that need to be fulfilled before surgery can take place. For example, they may have to wait for their digestive system to be free from food etc, for anesthesiology. The board records

if and when they are ready to begin surgery. In the worst case scenario, the maximum number of simultaneous P1 arrivals is three from experience at TOH ((27)). In that situation one patient can be seen in the dedicated OR, the on call team would be called in to see the second P1 patient, and assuming that an appropriate surgeon type is currently scheduled or working in the elective capacity then when that desired surgeon and a room is available it will be used. The scheduled elective patients will be reshuffled to be completed at a later time or in a different room depending on the surgery type. If capacity allows, on occasion surgeons work between two rooms resulting in patients being serviced faster. This was not modeled in this thesis, as it is not predictable, nor scheduled.

A set of surgeons is created in the model, so that there exists one of each type for each service. There is always one of each type of surgeon on call. It is assumed that unless scheduled, no more than one surgeon of each type can be used by the on call capacity and that there is one of each surgeon available for the dedicated OR.

4.3.1. *Assumptions.*

It is assumed that once assigned priority, patients keep that priority for the duration of their encounter. Although in reality patients may

change priority due to condition, the data do not differentiate these patients and the assumption is kept.

4.4. Solving The Ottawa Hospital problem.

As a hypothetical situation in this scenario, if a patient arrives for surgery during daytime hours (i.e., Monday at 3pm), and there is no one in the waiting room then the patient goes straight for surgery using the dedicated capacity. If at the same time that this patient goes for surgery another 2 patients arrive, one of P1 and one of P3, they both receive beds and wait in a queue. P1 patients are ahead of the P3 patient as they have a higher priority. If the patient currently in the dedicated OR develops complications and is not finished at the time when the schedule determines that the room becomes unavailable (4pm), then the staff have to finish the surgery before they can finish their shift, and they still have to be back at the same scheduled time the next day. Meanwhile the P1 has almost waited for 2 hours, and cannot wait until the next day for the dedicated OR capacity to reopen, so the on call team is called in. This team is comprised of the nurses that have just finished their shift and a surgeon of the required surgical type. An unoccupied room is opened for this patient's surgical procedure. The P3 however would now be first in the queue for the dedicated OR when it would reopen on Tuesday 8am. Further, if during the

night there is a car accident and four P2 patients requiring ortho are brought into the emergency, they wait until 8am. At this time there is only one OR in dedicated capacity so one of the four can be seen straight away, however as they all need the same surgeon even when the second room opens at noon only one will be seen at a time. There is another inpatient who needs to be seen immediately and is therefore assigned the priority P1, although they did not enter the hospital from the ED. This patient uses the second room in dedicated capacity at noon. The car accident patients are seen quickly and the third patient of the four is in dedicated capacity when their target is reached (early afternoon as they were admitted early in the morning, 5am) the fourth patient is then seen using on call capacity. After these are all finished the only other patient waiting is the initial P3. Although this patient has not yet reached their wait time target, they will be seen in the dedicated OR as it is available. Each of the patients get a PACU bed for initial recovery, and then the inpatient along with the P1 patient go to ICU. The car accident patients go to the wards for a short time before they can be discharged, and the P3 remains in the ward for a few days before being discharged. This would be representative of the simulation in the warm up period; after a month of this behavior a

queue of P3 patients builds up and they are seen at every opportunity when dedicated capacity is vacant.

4.5. Available data.

From the data we can generate distributions for demand and service rates, that is the rate that patients arrive, surgical durations by service type, and the durations of the times spent in downstream capacities, including PACU, ICU and ward. Using these distributions in the simulation model the effect of different scheduled capacity (dedicated OR and alternative capacity) with the currently set wait time targets can be evaluated. The attributes in the data that were used for creating these distributions are primarily priority, encounter start time, procedure start time, procedure length, procedure duration, service type, total post-operative ICU Length of Stay (LOS) and total post-operative non-ICU acute LOS.

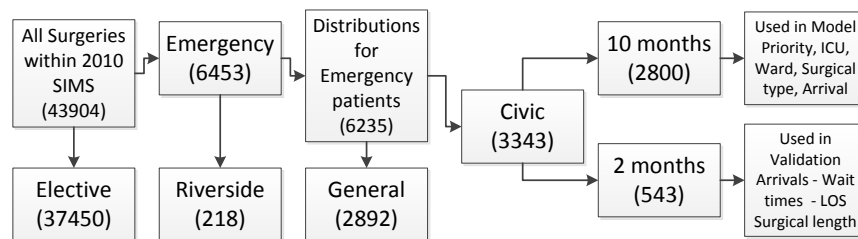


FIGURE 4.2. Flow map of the data from TOH

4.5.1. *Data collection.*

Data were provided for one year of surgical patients at TOH. From these data, the information for emergency surgeries was used and from that only the first 10 months of it was used to populate the model. The final two months were analyzed and used to validate the model. The data used to develop the simulation model are secondary data provided by the Ottawa Hospital. Data were collected from SIMS. The SIMS data set is linked with the Data Warehouse (DW) observations; these were comprised of both emergency and elective surgeries, categorized by the procedure type attribute. The SIMS data were used to collect the relevant information about emergency surgery procedures in order to assign rules to the simulation model to replicate this system as close to reality as possible. These data consisted of 25 variables, the ones which were used are elaborated below. Data were compiled on the 20th of July 2011.

4.5.2. *Data sources.*

There are data for five sites, CIVMOR (Civic OR), GENMOR (General OR), RIVCC (Riverside critical care unit), RIVMOR (Riverside OR) and GENEI (General eye institute). The first two are the only ones that were used for the model. Emergency Data were filtered by site (attribute variable), patients seen at the Riverside locations,

RIVCC and RIVMOR were removed, as those campuses lack the facilities to perform emergency operations. CIVMOR and GENMOR sites were kept and GENEI data were removed as surgeries performed there occur in a separate facility. The emergency patients from the Civic were then divided into 10 and 2 months.

Data for the Civic emergency patients from the Surgical Information Management System (SIMS) data set that were used for the model included 2800 patients (Figure 4.2 on page 57) of which 302 were P1, 963 were P2 and 1535 were P3. From these, the ones that had data for their stay in ICU were 153 or 50.66% of P1 patients, 92 or 9.55% of P2 patients and 75 or 4.89% of P3 patients. Distributions found from the data are used to determine the non-ICU, i.e., ward LOS in days, details of these can be found in appendix 6.4.3 on page 97. For the majority of the distributions, lognormal was the best fit. In the few cases where it was not the best fit, it was second or third best. For ease of importing the data into the model, the same distribution was chosen for each service and priority with different parameters for each. For some service and priority combinations there were very few values. For all of these cases the data points were combined and the resulting distribution that was used for each of them was LOGN (9.26, 113) (Appendix 6.4.1 on page 95).

4.5.3. *Characteristics of the data.*

Attributes of the data include Patient IDs specific to each patient, and Physician IDs indicating which surgeon was responsible for the operation. Case ID is unique for each case and each patient may have more than one of these case IDs. Encounter ID is given as a unique code for each time a patient receives a unique service. Again each patient has at least one encounter but is not limited to just one, this represents each time a patient is seen at the hospital. If Encounter ID are the same for different cases, that indicates that a patient was seen at one encounter for multiple issues, which was the case in about 819 encounters. Each patient has a unique identifier, and each time they come to the hospital they get an encounter ID and each ailment that is seen to is given a case ID. Patients can have multiple case IDs for the same encounter ID. They can have multiple encounters but not for the same case. Each encounter has a unique case ID. Multiple cases map to a single ICD-10 (International Classification of Disease) code. ICD codes are used to categorize procedures. A service type is associated with each procedure type.

The entry type of encounter is also important as it identifies the patients in terms of where they entered the system. Some surgeries that occur are only of the procedure type emergency but not entry

type emergency. As we are concerned with the required capacity that reduces wait times of emergency surgery patients all those patients for whom surgery is of type ‘emergency’ are considered. Priority for each emergency surgery is given in the attributes of the data.

The encounter start is one of the attributes used to calculate the patient’s wait time. That is, the time between when the patient is first seen (encounter start datetime), and when the patient’s treatment is started (procedure start time). Good distributions of the actual data were not achieved as many patients had wait times exceeding 6 days, these patients we assumed to be walking wounded. These patients theoretically utilize the dedicated emergency capacity at times when it would otherwise most likely be idle. The data do not differentiate between when patients return home or when patients have a long waiting time. The model was formulated in such a way that patients of the higher two priorities were seen at the latest by the time their targets were reached, and so this information was only used to validate and compare results.

4.5.4. *Discrepancies in the Data.*

The encounter type identifies the difference between inpatient or daycare. Many (19551 total) patients do not have a recorded encounter

TABLE 4.1. Attributes of data

Attribute	Details			
	Values taken	Min value	Max value	Total
Name				
Patient ID	#	1	35220	
Physician ID	#	1	211	
Case ID	#	1	44399	
Encounter ID	#	1	43085	
Encounter Type	INPATIENT			13797
	DAYCARE			30107
Encounter start	datetime	19Aug2009	31Dec2010	
Encounter end	datetime	01Jan2010	23Mar2011	
Entry Type	DIRECT			8638
	EMERGENCY			5092
	CLINIC			50
	DAY PROCEDURE			17
Site	CIVMOR			10867
	GENMOR			10679
	GENEI			2847
	RIVMOR			9399
	RIVCC			10112
	Procedure start	datetime	01Jan2010	
Procedure end	datetime		31Dec2010	
Primary Procedure	ICD-10 Code	A04	Z52	691
Service Type	#	1	15	
Procedure Type	ELECTIVE			37450
	EMERGENCY			6453
Procedure Length	# minutes	1	1440	
Priority	P#	1	3	
Total non-ICU acute LOS	# days	-19.27	161.94	
Total post-operative ICU	# days	-0.74	85.32	
Total post-operative LOS	# days	0	388.85	

end datetime and therefore LOS could not be accurately calculated from the encounter end time and procedure end time.

Information on the total non-ICU acute and post-operative ICU length of stay is given in days, with a total of 30107 patients missing

information for these attributes. The cleaned data were analyzed in order to populate and validate the proposed simulation model.

Data errors such as negative LOS(as seen in table 4.1 on page 62) and wait times were dealt with by assigning a value. In the case of negative wait times these values were reassigned the lowest non-negative value. There is no reasonable explanation for a negative wait time with the exception of data entry error.

The cleaned data set was restricted to those patients who have an allocated priority and were serviced for site specified non-elective surgery types that is, either the GENMOR or the CIVMOR for a total of 6235 patients. The model focused on the Civic data as both hospitals were very similar in their available capacity and scheduling. They have identical set wait time targets for their priority classes. A couple of surgical cases occurred only at one or the other site, from this point the model is formulated specific to the Civic site data, however, it is created in such a way that the General site can easily be implemented using the same model by recalculating the distributions from the data.

4.5.5. *Scenarios.*

The base model is the simulation model as described. It includes the wait time targets as set by the hospital. This simulation models the

TABLE 4.2. OR capacity schedules

OR Capacities and their Scheduled times	
Dedicated OR (2)	M-F, 12.30-16:00 and 7 days 8-16:00
On call (1)	7 days a week, 24 hours
Elective Capacity (1)	M-Sa, 8-16:00

TABLE 4.3. Factorial design of scenarios

Scenario	Additional capacity					Patients seen		
	On call			Dedicated		On call		
	12am-12pm	12pm-12am	24Hrs	8am-12pm	12pm-4pm	P1	P2	P3
Base						X	X	
A	X					X	X	
B		X				X	X	
C			X			X	X	
D				X		X	X	
E					X	X	X	
F						X	X	X
G	X					X	X	X

current capacity set at the hospital according to the master schedule.

It is defined as the available capacity in table 4.2 on page 64. For the other scenarios the available capacity and the way in which P3 are seen are varied. Below is a table of the other scenarios indicating which changes apply to each scenario (Table 4.3).

5. Experimentation, analysis and results

The simulation model built in Arena enabled data collection to evaluate the performance of the system under different conditions. The most favourable alternatives can be determined from the results obtained by running the simulation model. Performance measures are used to compare the different scenarios or variations from the base model. Specific measures are set to evaluate the percentage of patients using each alternative capacity and total time of use. Other measures look at the time that patients have to wait. Data were collected from the base model as a benchmark. In each scenario the same performance measures were evaluated. Below are descriptions of the model and the scenarios tested. First, the base case was defined, and from there the scenarios were broken into categories. The categories include those scenarios which vary available amounts of on call capacity, dedicated capacity and the rule for P3 patients upon reaching their wait time targets.

5.1. Performance metrics.

The performance metrics that were calculated mostly concern utilisation. The utilisations measure the balance between regular and overtime, the on call capacity, the amount of canceled elective capacity used, and downstream utilisation. In the model, P1 and P2 patients

TABLE 5.1. Averaged daily utilisation of downstream capacities for each scenario

	PACU	ICU	Wards
Base	1.28 +- 0.004	5.34 +- 0.004	50.55 +- 0.054
A	1.27 +- 0.004	5.25 +- 0.004	50.26 +- 0.049
B	1.26 +- 0.004	5.32 +- 0.003	50.92 +- 0.054
C	1.26 +- 0.004	5.23 +- 0.003	48.84 +- 0.048
D	1.28 +- 0.005	5.39 +- 0.005	49.61 +- 0.047
E	1.30 +- 0.004	5.34 +- 0.005	50.88 +- 0.051
F	1.28 +- 0.004	4.83 +- 0.004	49.91 +- 0.048
G	1.28 +- 0.003	5.21 +- 0.003	49.18 +- 0.047

are seen within their wait time targets, so the proportion of P3 patients who are not seen within their targets are compared between the different scenarios.

5.1.1. *Instantaneous utilisation of downstream capacities.*

Instantaneous utilisation was analyzed both daily and by hour of the day. As patients are only discharged during the day, the daily utilisation is the focus. The instant utilisation did not vary considerably between the daily and hourly calculations. Below is a table of the utilisation for each PACU, ICU and ward: calculation is given hourly for each day. For each day of the week, the average over the 10 replications is calculated, the average over the entire week is calculated and the 95 percent confidence interval is given below (Figure 5.1 on page 67 and Table 5.1 on page 66).

Downstream Capacity Planning by day of the week

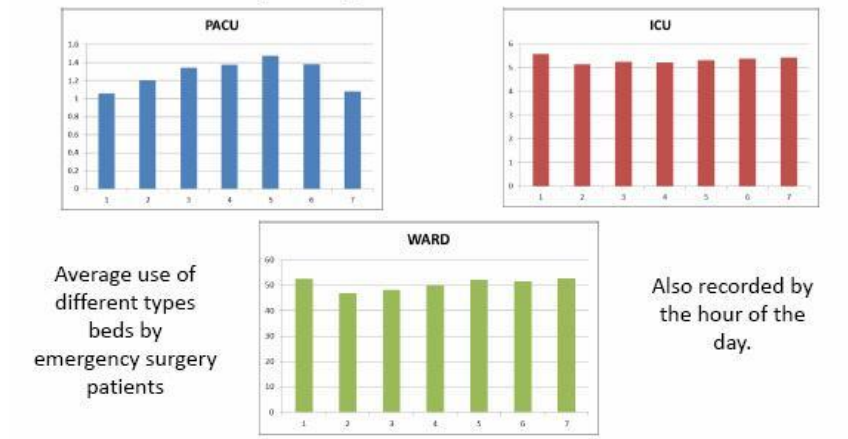


FIGURE 5.1. Instant utilisation of bed capacities.

5.1.2. Utilisation of capacities.

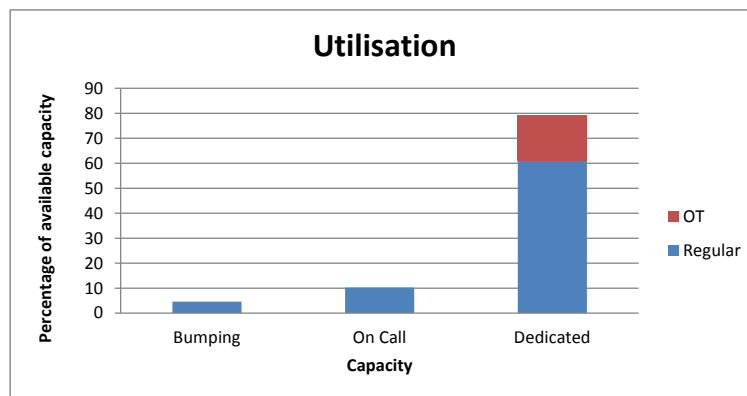


FIGURE 5.2. Utilisation with combined regular and overtime.

For this section the utilisation of OR capacities will be detailed. From the simulation runs recordings were made of the surgical times of each patient and the capacity by which they were serviced. These numbers were used to calculate the total time each capacity was used. This time is divided by the total available time for each of the three capacities (Figures 5.2, 5.3 and 5.4). The time in minutes out of the total available time of each capacity is given in Figure 5.2 on page 67. Figures 5.3 on page 68 and 5.4 on page 69 show the number of patients that use each capacity calculate both by day of the week, and by hour of the day.

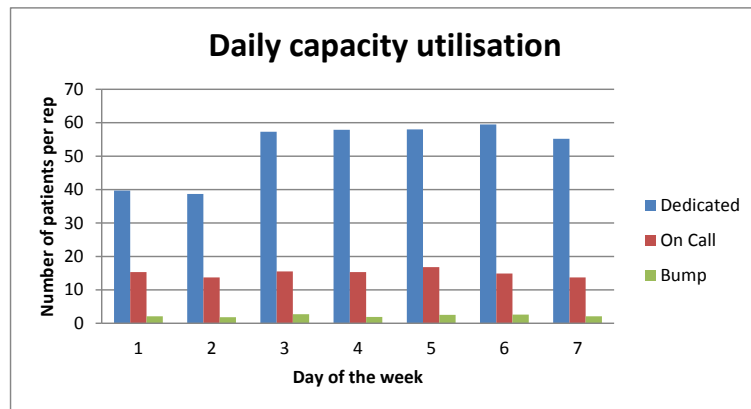


FIGURE 5.3. Capacity utilisation by day of the week, in the base model

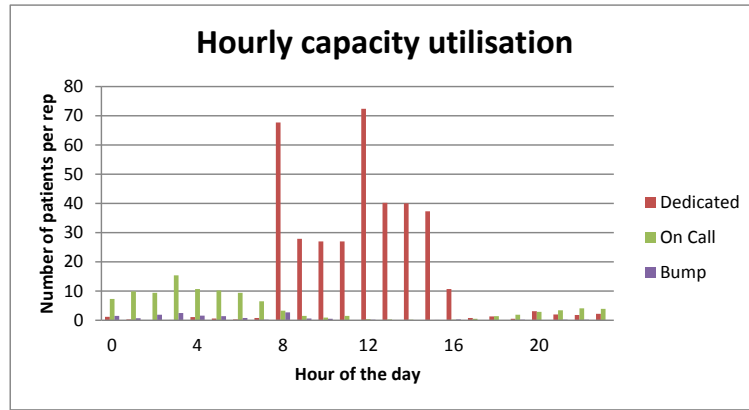


FIGURE 5.4. Capacity utilisation by hour of the day, in the base model.

5.1.3. *Canceling elective capacity.*

For the canceling elective capacity we evaluated the percent of the time this option was utilised, the number of patients who used it, and the total time that emergency surgeries used that capacity. Canceled elective capacity used is calculated as a combination of the surgical durations of those patients who were seen using this capacity. In this case the total time is calculated as a percentage of the equivalent of one scheduled OR of elective surgery per day.

5.1.4. *P3 waiting beyond their target times.*

The model configurations restrict P1 and P2 patients from waiting beyond their wait time targets. Thus, the focus will be on P3s, as they can wait for surgery beyond their wait time targets. Again, for these patients we examined the percentage of patients who were not seen within their targets. This measure was calculated by finding the average over the 10 replications of P3 patients. The proportion of those who waited in excess of 2880 minutes (the P3 wait time target) is then calculated. The number of these patients, and their average time waiting for service were evaluated.

Below is a graph (Figure 5.5 on page 70) of the current P3 patient wait times. Wait times are measured as the times from when patients arrive at the ED to when they receive surgery. The vertical line represents the P3 wait time target and the patients to the right are representative of patients who had to wait beyond their targets.

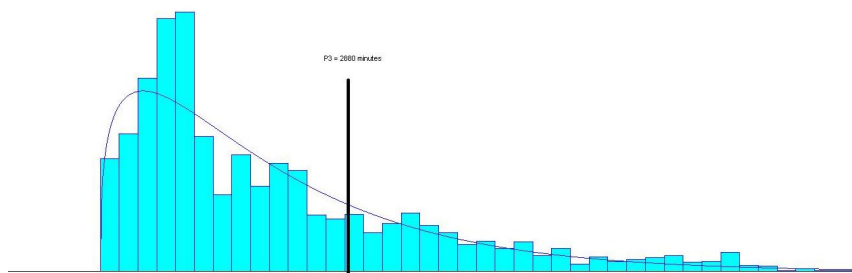


FIGURE 5.5. P3 patients with relation to their wait time targets, in the base model.

5.2. Scenario results. Refer back to Table 4.3 on page 64 for the details of the scenarios. The results from the different scenarios were compared to evaluate what variations improved upon the results from the base case. The first set of variations in which the on call capacity was varied are outlined below. Table 5.3 on page 75 shows the effect that each scenario has on the utilisation of on call capacity.

5.2.1. Varying on call capacity.

The first three scenarios are those in which the on call capacity is varied. These scenarios provide results with the most reduced canceled elective capacity from the base case, as well as decreased on call utilisation. However, these results were not beneficial for the number of P3 seen within their wait time targets.

Scenario A

This scenario varies from the base model by the addition of on call capacity for the first 12 hours of the day. From 12am to 12pm, a second room is made available to be used as on call capacity, thus doubling the initial capacity for this time period. It is assumed that the number of available surgeons is not altered. This scenario has the most favourable use of dedicated capacity, with a decrease in overtime usage from the base case. However, the P3 who exceed their targets in this scenario is worse than the base case.

Scenario B

In this scenario additional on call capacity is added for the last 12 hours of the day. From 12pm to 12am a second on call room is made available. The earlier 12 hours are the same capacity as the base model with one on call room available. Slightly more P3 patients are seen within their targets than scenario A. This scenario results in low elective capacity usage but is still not favourable for P3 patients.

Scenario C

In this scenario a second on call capacity is added for 24 hours a day 7 days a week. Thus doubling the entire on call capacity from the base model. Again assuming the same surgeon availability as the base case. Although this capacity has the lowest on call utilisation the results are not twice as favourable as scenarios A and B. Again as for this group of scenarios there is a reduced used of elective capacity but a great number of P3 patients are still not seen within their targets.

5.2.2. Varying dedicated capacity.

The results for each scenario are displayed below (Table 5.2 on page 75) showing the effect on the dedicated capacity. Below are also the scenarios for which the available dedicated capacity was varied from the base case. For these scenarios the number of P3 patients that are seen within their targets are greatly increased, but the elective capacity

usage levels are close to that of the base case. On call utilisation is not greatly improved upon from the base case.

Scenario D

In this scenario 4 hours of additional dedicated capacity are added to the schedule in the am during the weekdays. Resulting in 2 full rooms being available 5 days a week. This scenario provide the lowest results for overtime utilisation, and the fewest P3 patients seen beyond their targets. However, elective capacity utilisation is higher than in the base case, and dedicated utilisation is low.

Scenario E

This scenario adds 4 hours to the dedicated capacity. However, the added time is in the afternoon resulting in a total of 1 room 8am to 12pm for dedicated capacity and then 3 rooms from 12pm to 4pm for dedicated capacity during the weekdays. This scenario results in the lowest utilisation of dedicated capacity, implying high idle time. Overtime utilisation is also high. This scenario is favourable for P3 patients.

5.2.3. Varying the servicing rule for P3 patients.

In the base case the P3s were only serviced by dedicated capacity, allowing them to be serviced by on call capacity would result in more P3 patients being serviced within their wait time targets. Below are

two scenarios in which this is the case. In these scenarios there are low numbers of P3 seen outside their wait time targets, and low utilisation of dedicated capacity overtime.

Scenario F

In this scenario the rule is changed so that P3 patient who reach their targets can utilize the on call capacity during the daytime hours, thus not restricting higher priorities from being serviced as at this time canceling elective capacity is also available. Although this scenario provides low numbers of P3 patients exceeding their targets, there are high levels of on call and elective capacity utilisation. To balance this a final scenario is created as a combination with this scenario.

Best case: Scenario G

In this final scenario the most favourable two scenarios A and F are combined. Scenario A uses minimal amounts of canceled elective capacity. When compared to scenario B, dedicated capacity is used for a greater proportion of time, and slightly fewer patients are seen using overtime capacity in scenario A. Scenario F ensures a higher number of P3 patients are seen within their wait time targets. In scenario G there are an additional 12 hours of on call capacity available from 12am to 12pm, as well P3 patients have the ability to be serviced by on call capacity during the daytime hours. In this scenario the number of

P3 patients seen beyond their targets and the use of both on call and elective capacity are much reduced from the base case. In summary, this can be seen in Table 5.4 on page 76 which displays the results of each scenario related to P3 patients serviced outside their wait time targets, and additionally the performance metrics for canceling elective capacity, in both time and number of patients.

TABLE 5.2. Dedicated capacity performance metrics for each scenario

	% Reg time	% OT time	% Reg patients	% OT patients
Base	78.28 +- 0.17	23.59 +- 0.07	56.67 +- 0.08	17.67 +- 0.04
A	60.41 +- 0.15	18.88 +- 0.08	56.25 +- 0.04	18.02 +- 0.05
B	59.91 +- 0.09	18.53 +- 0.04	56.81 +- 0.05	18.03 +- 0.03
C	58.85 +- 0.14	18.36 +- 0.08	56.26 +- 0.05	17.67 +- 0.05
D	52.76 +- 0.05	11.24 +- 0.02	61.01 +- 0.06	14.34 +- 0.02
E	45.66 +- 0.07	18.93 +- 0.06	57.24 +- 0.06	18.25 +- 0.04
F	60.47 +- 0.08	16.86 +- 0.04	54.66 +- 0.05	17.06 +- 0.03
G	59.16 +- 0.09	16.64 +- 0.05	55.14 +- 0.07	17.07 +- 0.03

TABLE 5.3. On call performance metrics for each scenario

	% On call time	% On call patients
Base	10.46 +- 0.04	22.23 +- 0.04
A	7.14 +- 0.02	23.96 +- 0.05
B	6.41 +- 0.01	21.97 +- 0.04
C	5.37 +- 0.01	24.40 +- 0.06
D	9.77 +- 0.02	21.58 +- 0.04
E	10.39 +- 0.02	21.87 +- 0.04
F	11.28 +- 0.08	24.81 +- 0.06
G	7.84 +- 0.02	25.59 +- 0.07

TABLE 5.4. Canceled electives and P3s who exceeded their targets for each scenario

	Canceling elective capacity		P3 patients over target
	% Elective time	% Patients seen in elective	% Patients seen beyond target
Base	4.67 +- 0.03	3.43 +- 0.02	25.94 +- 0.33
A	2.59 +- 0.03	1.76 +- 0.01	27.44 +- 0.34
B	3.97 +- 0.03	3.20 +- 0.02	22.57 +- 0.39
C	2.59 +- 0.03	1.67 +- 0.02	20.65 +- 0.36
D	4.98 +- 0.04	3.07 +- 0.02	3.09 +- 0.05
E	4.21 +- 0.03	2.63 +- 0.02	3.91 +- 0.07
F	5.52 +- 0.04	3.47 +- 0.02	6.15 +- 0.09
G	3.68 +- 0.04	2.19 +- 0.02	2.78 +- 0.04

5.3. Verification.

Verification was done to ensure that the model is performing as intended. This verification consists of model debugging to ensure that if there are errors in the simulation code they can be located. Errors of importance include improper flow control or entity creation, failure to release resources and incorrectly implemented statistics. These errors could be either logical or arithmetic. A part of model debugging included scenario repetition with varying factors to ensure that the model simulated with anticipated behavior. Individual modules in the simulation code were also tested ((35)).

This model was continually verified throughout the building stage. As sections of the model were built patients were measured and counted

both in and out of modules. This step ensured that the flow was reliable according to the rules incorporated into the simulation and that patients were not building up in queues where there should not be any blockage. In the outputted results it was also verified that P3 patients were only serviced by the dedicated OR capacity. Finally, the model was verified by changing the wait time targets to reflect the actual times within which 80% of patients were actually seen. By making this change and by inspecting how the results compared with the original data, the model was verified. It was clear that the model was a reliable representation of the reality in terms of the resource utilisation of the various beds type and ORs.

The bed utilisation is what seemed average for the given situation according to the surgical management staff at TOH who supplied data sources ((27)). The model is therefore performing as intended with real scenario inputs, and expected results of the current practice were were obtained using model parameters of actual wait times.

5.4. Validation.

To increase the scientific nature of this research the notion follows that models should be wholly tested or validated before use, ensuring that each model is appropriate for the task which it is intended. Validation is important from a practical perspective as well as theoretical

one. A model is developed with the view of encompassing a system. This system may previously be in existence and the model can shed light into alternative operations. Alternatively, a target system could exist and the model could demonstrate preferable modes of operation, or new considered designs of a system which does not yet exist could be modeled. The latter two options are more challenging to validate as there does not exist a reality with which results can be compared. Theoretically the validation is performed to determine if the model accurately depicts the real system. Ignoring the internal workings of the model, there can be two parts to this validation: Is the output an accurate reflection of the real system? And, do components of the model represent existing known behavior or a valid theory? The latter is more important in cases where a real system is not already in place and other validation is not possible. These questions can be answered by showing the model to different personnel associated with the system who can offer advice concerning the realism of the simulation model. Further observations of the system are performed to ensure model validity with respect to actual system performance. A simple technique is to statistically compare the output of the simulation model to the output from the real system and analyze whether there is a significant and practical difference between them. This validation application is

to a system that is already in place and can provide alternative modes of operations, where the components of the model represent existing behavior.

A meeting was set to discuss these results with experts as expert opinion was the initial chosen method for validation. Findings were discussed and the workings of model were demonstrated. A comparison of the way in which the model used real data to show the utilisation of beds and numbers of patients waiting was then compared to the same model with fixed wait time targets. The most relevant validation provided at this meeting was that the initial SIMS data did not represent the real waiting times of emergency surgeries. Waiting times from the data were much longer due to the changing priorities and walking wounded patients that were not captured in the data.

Even for P1 and P2 patients the model was well-received and thought to be representative of the current and potential future situation.

Following this meeting, additional means were employed to further validate the model. This testing was conducted using data that had been previously divided into 10 months and 2 months. The 10 month data was used to populate the model and the 2 month data was used to validate the model. This validation was based on four criteria: patient arrivals, wait times, surgical durations and the downstream LOS.

5.4.1. *Patient arrivals.*

Arrivals calculated by service type and priority class for an average period of 2 months of simulated data were compared with 2 months of historic data. Figure 5.6 on page 81 shows the arrivals by service type for each priority. It can be concluded that the simulation model is an accurate representation of the patient arrivals from the records kept in 2010.

5.4.2. *Patient wait times.*

As previously validated by the expert opinion of TOH executive, the wait times were not an accurate representation of the reality at TOH. Considering the distribution of wait times by priority class, it was not expected that P1 or P2 patients would align with TOH data (Figure 5.7 on page 82). The servicing rule implemented in the model meant that these patients were seen at the latest by the time their targets were reached with very few patients exceeding this. P3 patients however, had wait times that were closely in line with those from TOH 2 months data (Figure 5.8 on page 82).

5.4.3. *Surgical durations of the patients.*

Sample distributions of the length of surgeries by service type were calculated. Many graphs were made to compare simulation results and historical data for each surgical type. Examples of two of the service

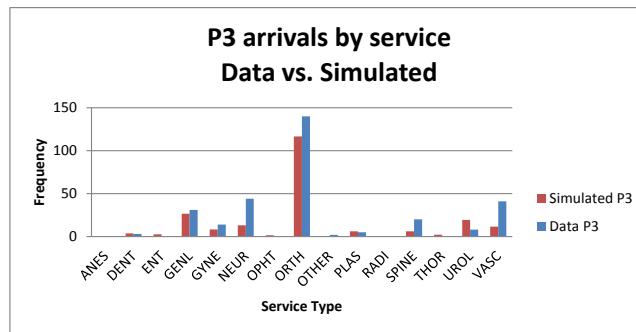
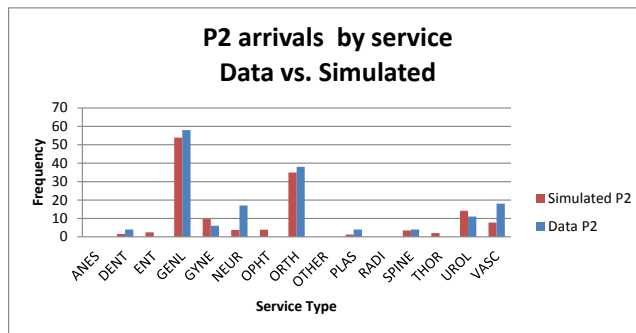
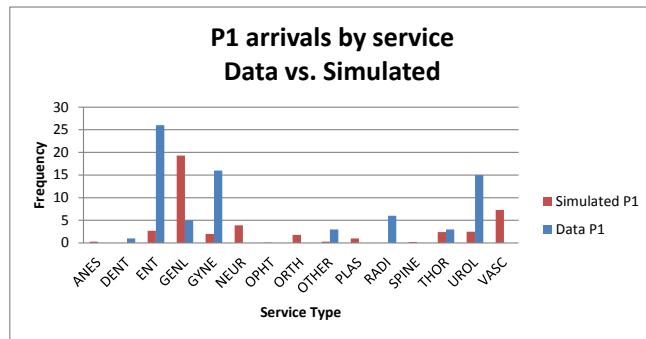


FIGURE 5.6. Arrivals by priority and service type for 2 months of simulation data and 2 months historic data.

types, general surgery and urology, which were nicely distributed are

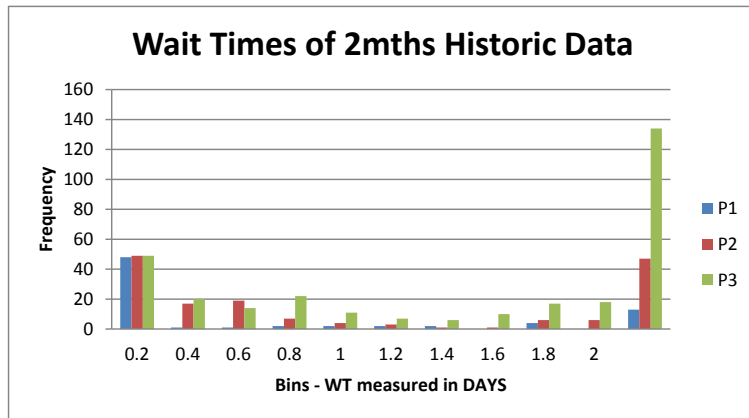


FIGURE 5.7. Waiting times for 2 months of the historic data.

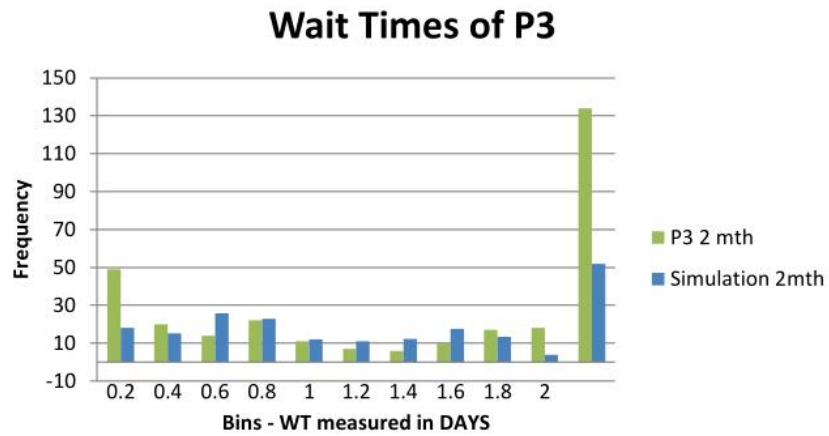


FIGURE 5.8. Waiting times for P3 historic data and simulation.

shown (Figure 5.9 on page 83). Differences within the graphs most likely arise due to the validation period only being two months.

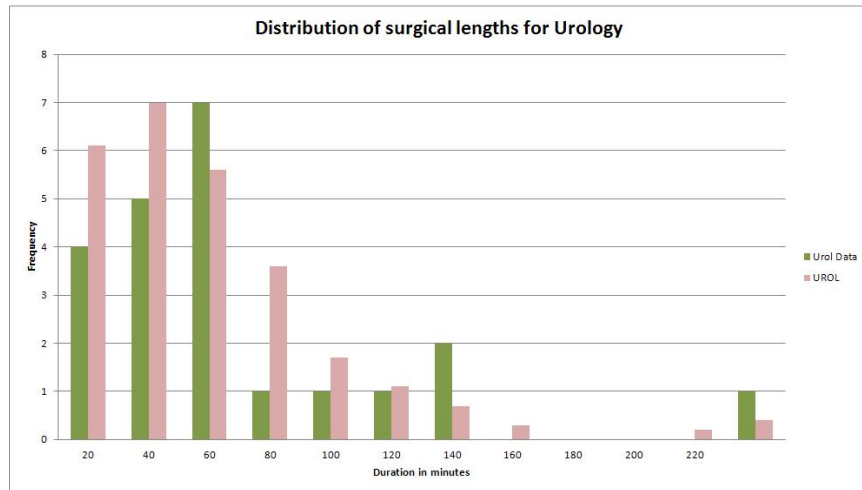
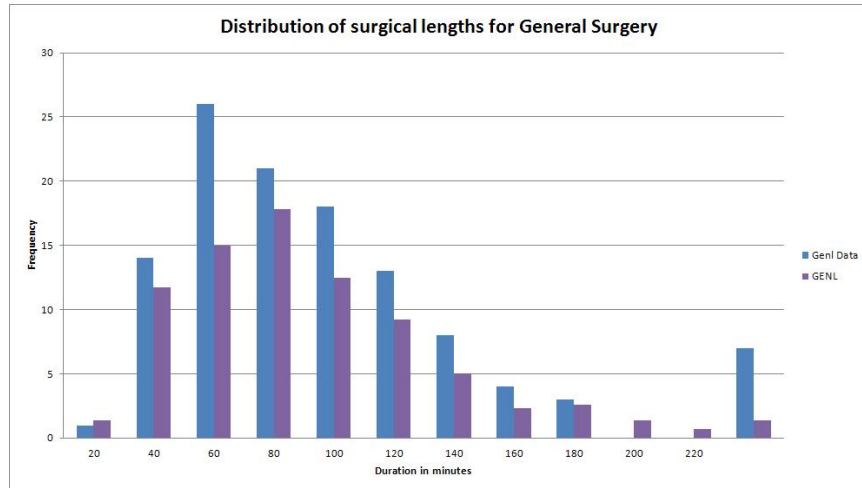


FIGURE 5.9. Surgical durations compared from the model and TOH data

5.4.4. *LOS in downstream capacities.*

The length of stay post-surgery was divided into PACU, ward and/or ICU. As there was no information on the PACU times given in the data, triangular estimates were used as provided by hospital personnel

for PACU durations. In this case, only the ICU and ward LOS are compared with the 2 months of data.

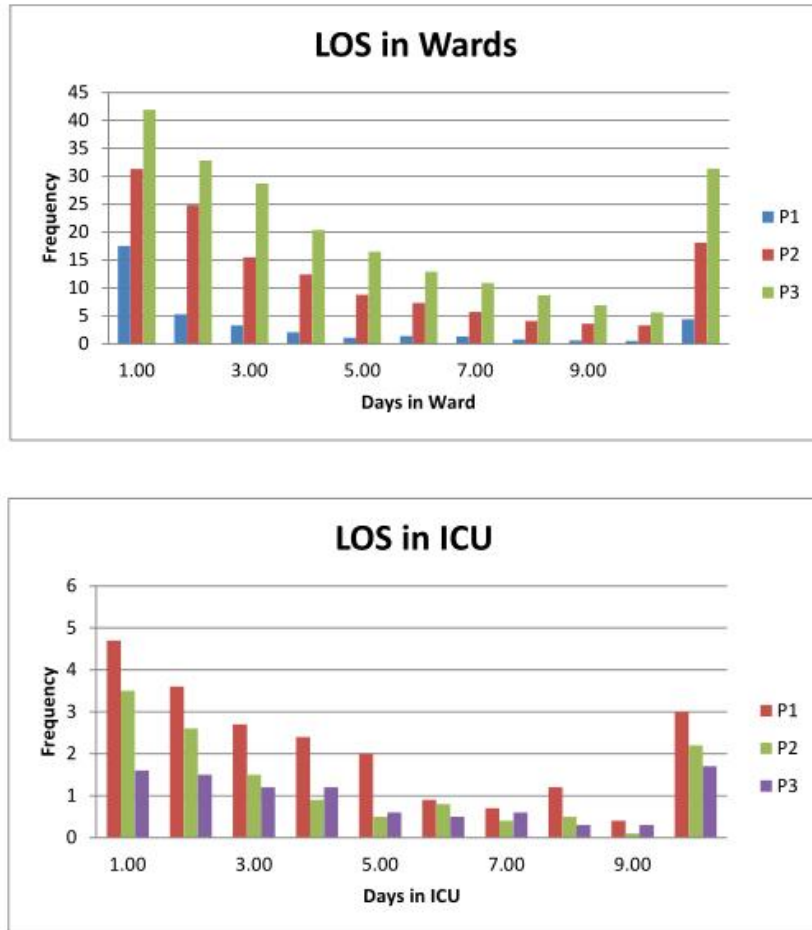


FIGURE 5.10. ICU and ward LOS stays by priority

In conclusion, when compared to the data from TOH, the distributions from the simulation model were well matched to the 2 month validation data and could validate the model on multiple counts.

6. Discussion and Conclusion

Below is the discussion of the results, the implications that it has for both TOH and for emergency surgeries in general, followed by a section on conclusions, limitations and future studies.

6.1. Main findings.

The emergency capacity at TOH needs to be increased in order to see patients within their wait time targets, as suggested by the simulation model. In the model P1 and P2 patients were not seen outside of their wait time targets however, these patients sometimes used overtime to have their surgeries completed. In the base case P3 patients often exceeded the wait time target. Although varying the dedicated capacity ensured a greater number of patients were seen, this scenario is associated with the cost of canceling elective capacity, whereas additional on call capacity reduced the number of patients who canceled elective capacity because it is available over a longer period of time. A combination of the P3 rule and additional on call in the early hours of the day is the best combination in order to reduce the number canceled elective surgeries while servicing more patients within their wait time targets.

6.1.1. *The Ottawa Hospital policy implications.*

This model may be the right model and respond to the objectives of the study but what is its usefulness for TOH and their day to day scheduling and servicing? This research provides TOH with a guideline of the most effective projected capacity changes, as well as the expected impact of different capacity changes. If TOH wants to meet more of the wait time targets for emergency surgical patients then an increased capacity is needed and the best results are achieved by the addition of an extra 12 hours of on call capacity assigned to the first 12 hours of the day and allowing P3 patients to use on call capacity during the daytime hours.

6.1.2. *Implications for emergency surgery scheduling.*

Emergency surgeries can be complex, and determining the capacity required to service them is highly important. Emergencies occur around the clock and across the country. They occur whether there is scheduled capacity or not. The way in which the capacity may be scheduled, and the method of determining the amount of available capacity to be set aside, is independent of the occurrence of the emergency surgery. Therefore, it is an important problem in every hospital regardless of their overall size or available capacity.

The objective is to have sufficient capacity in order to meet the demand within their targets, while avoiding an excess amount of scheduled capacity that results in under utilized resources. The problem is complicated because not every emergency surgery is of the same urgency. Thus, different classes of patients have different wait time targets. During scenarios in which the available on call capacity was increased, performance measures obtained in terms of patients seen within their targets were better than those scenarios in which the available dedicated capacity was increased. However, as more patients are seen with additional capacity the demand for downstream resources would also be increased.

For emergency surgeries here and as seen in the literature, allowing some managerial leeway through a booking window (as opposed to servicing demand directly) can have the effect of leveling the demand and reducing overtime, or alternative capacity in this research. Assigning patients a wait time target immediately allows for time to make a scheduling decision. Wait time targets however, need to be reasonable in relation to the amount of available capacity. Capacity needs to be sufficient to ensure that these targets can, for the most part, be met.

6.1.3. *Limitations.*

A common problem that the addition of capacity may have is that an unrealized demand may become apparent as more capacity is made available. Improvements made at one site or campus may not ensure more patients are seen within their wait time targets due to an increase in demand. For example, perhaps more patients who would have otherwise been in a different jurisdiction would come to this hospital because of its improved services, or patients who perhaps were on the border of being classified as emergency patients may be more likely to use the services, and no greater proportion of patients are seen within their targets than previously observed.

In order to determine new results of the capacity requirements to meet different wait time targets and different priorities the model would have to be altered and re-evaluated. Initially the new priority categories of patients would need to be defined in the model. Information would need to be collected containing the data with the changed priorities including encounter start and end time, and procedure start and end time. The patient's priority and LOS would also need to be collected. From these data the distributions of patients who are serviced by each priority could be calculated and imputed into the model. The model would then need to be re-run and performance metrics collected with

these different parameters. New results could then be evaluated as a part of a future study.

6.2. Conclusions and future studies.

The study conducted was in regards to the problem of capacity planning for emergency surgeries with multiple priority levels and a need to meet wait time targets. A simulation model was developed and the collected data were applied to it after it was cleaned. Data were used to populate the model from the emergency surgeries that occurred at the Civic site of TOH. Patients in this model were created with one of the three priorities and respective wait time targets as well as other attributes; in order to individualize them. P1 and P2 patients were not able to exceed their targets as the use of alternative capacity at this time was used in order to prevent their targets being exceeded. Dedicated capacity, capacity carved out of the elective surgical schedule used for the surgeries at TOH, are ORs reserved solely for emergency surgeries. Those patients waiting are seen in order of highest priority and by first arrival. If patients require immediate service (i.e., their wait time target has been reached) and dedicated capacity is in use, then alternative capacity is used. First, an on call team and failing that elective capacity would be utilised by canceling or rescheduling an elective surgery. In this base case P1 and P2 patients were seen

within their targets and P3 patients were seen by first available capacity generally soon after their target time. Performance metrics such as the percent and total of time that alternative capacities were used, and the percentage of time that P3 patients exceeded their wait times were evaluated. Different scenarios were run by adding additional capacity and allowing P3 patients to be seen within the on call capacity. Several scenarios were run and the results of each were compared using performance metrics which included waiting time for P3 patients beyond that of their targets. Utilisation of each of the capacities, including overtime usage of the dedicated capacity, as well as downstream utilisation of resources were measured. The two most favourable scenarios in terms of reduced canceled elective capacity and increased numbers of patients seen within their targets were those that included the addition of on call capacity in the am, and the permitting of P3 patients to use on call capacity during the daytime hours. These were both applied, creating the final most favourable solution in which both these variations were combined. The effect of this scenario is that the use of canceling elective capacity is reduced, and a lower percentage of P3 patients are seen outside of their wait time targets.

This simulation model demonstrates the relationship between the desired wait time targets and necessary capacities in order to meet

these targets. The results show that all P1 and P2 patients would be seen within their targets, the percentage of canceled elective capacity would be reduced and the proportion of P3 patients not seen within their targets would also be decreased. For the wait time targets that were used at TOH there would need to be an available downstream capacity of at least 2 PACU beds, 6 ICU beds and 51 ward beds in order to avoid congestion due to emergency surgery patients. This research filled a gap in the reviewed literature on waiting time management and capacity planning. Few studies of this type look at multiple priority cases and although wait time targets are increasingly important few studies incorporate them in planning problems. This research provides a model that incorporates both multiple demand classes and multiple supply classes allowing the user to accurately determine the necessary capacity in order to meet pre-specified wait time targets.

6.2.1. *Future studies.*

Targets need to provide value from both the point of view of the hospital (i.e. financial) and from the point of view of the patients, as a health concern. The focus can be on finding the optimum value of this point, benefiting both parties as much as possible. Potentially, the emergency cases in which the model applied could be isolated in order to solve the optimisation model.

In future, with this more complete understanding of the scheduling process and different effects of small changes to the available capacity a second attempt at the initial non-linear model can be made. As a result of this simulation model, the operating system can be better understood, perhaps allowing for the initial optimisation model to be modified. It would need to be simplified to a point at which it could be solved and still maintain enough complexity for the results to be implementable. The objective function will need to contain more than just the capacity of regular and overtime as much more detail is needed in order to find a solution that can be useful in capacity planning of emergency surgeries and yet still solvable.

6.3. Glossary.

ATC Access to Care

ALC Alternative Level of Care

Capacity, patients required available capacity in order to be serviced, there are three types of capacity in this model; dedicated, alternative and canceling elective capacity.

CIVMOR Civic site OR, one of the attributes in the data

CT Computed Tomography

ED Emergency Department

ER Emergency Room

GENEI General site Eye Institute, one of the attributes in the data

GENMOR General Operating Room, one of the attributes in the data

ICD-10 International Classification of Disease - 10, a field of the data.

ICU Intensive Care Unit, the place where patients recover after they have come from PACU and before being transferred to the wards.

IP Integer program, a type of methodology.

LOGN Lognormal distribution

LOS Length of Stay, period of time a patient stays in a unit before being transferred or discharged.

LP Linear Program, a type of methodology.

MDP Markov Decision Program, a type of methodology.

MIP Mixed Integer Program, a type of methodology.

MRI Magnetic Resonance Imaging

OR Operating room

ORTHO Orthopedics, a service type from the data attributes.

OTHER A service type attribute for patients who do not fit within any of the service groups.

P1,P2,P3 Priority 1, 2 and 3 respectively.

PACU Post Acute Care Unit, where the patient goes immediately after service for a short recovery period before being transferred to either ICU or the ward.

RIVCC Riverside site Critical Care Unit, one of the attributes in the data

RIVMOR Riverside Operating Room, one of the attributes in the data

SIMS Surgical Information Management System, from where the Data was collected.

TOH The Ottawa hospital, the facility in which this model was applied.

Teardown The time after a patients operation in which the patient has left the room and it is cleaned and restored to be ready for next use.

WTIS Wait time information system, data collected by the provincial government to measure wait times for ALC, sugical and diagnostic imagine procedures

6.4. Appendix.

TABLE 6.1. Summary table of reviewed literature

Author	Problem			Application			Methodology			
	Cap	Adv	Apt	Gen	Surg	LP	Que	Sim	MDP	ADP
Ayvaz and Huh ((7))	X				X					X
Green et al. ((23))	X		X		X				X	
Oddeye et al. ((33))	X				X	X				
Creemers ((16))	X				X		X			
Cardoen et al. ((12))	X	X	X	X	X	X	X	X	X	X
Dobson et al. ((20))		X		X		X				
Conforti et al. ((15))		X		X		X				
Patrick et al. ((34))		X		X						X
Saure et al. ((38))		X		X						X
Belien et al. ((10))		X			X	X				
Blake ((11))		X			X	X				
Santibanez et al. ((36))		X			X	X				
Chow et al. ((14))		X			X	X		X		
Gerchak et al. ((21))		X			X	X				
Ho and Lau ((26))			X	X		X				
Green and Savin ((22))			X	X			X			
Muthuruman and Lawley ((32))			X	X		X				
Zeng et al ((40))			X	X		X				
Kim and Giachetti ((29))			X	X		X				
Gupta and Wang ((25))			X	X					X	
LaGanga and Lawrence ((31))			X	X				X		
Denton and Gupta ((17))			X	X		X				
Begen and Queyranne ((9))			X		X	X				
Charnetski ((13))			X		X			X		
Gul et al. ((24))			X		X			X		
Dexter ((19))			X		X					

6.4.1. Data for surgical types by priority.

P ercentages of patients of different services by priority			
Service type	Priority		
	P1	P2	P3
ANES	0.16286645	0	0.088941595
DENT	0.16286645	1.15658363	1.363771124
ENT	5.537459283	1.512455516	1.482359917
GENL	42.18241042	39.81316726	11.74029054
GYNE	6.51465798	6.806049822	3.290839016

NEUR	9.446254072	2.846975089	5.692262081
OPHT	0.325732899	2.846975089	0.622591165
ORTH	4.397394137	24.42170819	53.78001779
OTHER	0.325732899	0.222419929	0.118588793
PLAS	2.931596091	1.112099644	2.460717462
SPINE	0	0	0.029647198
THOR	1.628664495	2.669039146	3.261191817
UROL	5.374592834	1.245551601	0.978357545
VASC	6.026058632	9.919928826	9.190631485

6.4.2. Data for surgical setup and teardown times.

Using expert opinion procedure types were matched with standard setup and teardown times. Below are the percentages of setup and teardown times by service types. These were found by matching the given procedure types with their respective setup and teardown times for the emergency patients. These were then categorized and the occurrence of each individual pair was counted and divided by the number for that service type.

Service	Setup (mins)	Teardown (mins)	Percentage
ANES	10	5	100%
DENT	10	5	100%
ENT	10	5	81.4%
	5	5	11%
	25	15	6.7%
	20	10	0.9%
GENL	10	5	70.6%
	20	10	22.9%
	15	10	4.1%
	20	20	0.8%
	45	45	0.8%
	25	15	0.5%
	25	20	0.2%
	40	20	0.1%
GYNE	0	0	38.2%
	15	15	34.9%
	5	5	10.9%
	10	10	6.9%
	15	10	6.6%
	20	15	1.3%

	10	5	0.6%
	20	10	0.6%
NEUR	10	5	47.1%
	20	10	46.2%
	10	10	6.1%
	20	90	0.6%
OPHT	10	5	71.8%
	20	10	26.8%
	20	5	1.4%
ORTH	20	10	72.8%
	10	5	22.3%
	25	15	3.8%
	15	15	0.7%
	5	5	0.4%
OTHER	5	5	100%
PLAS	10	5	83.3%
	20	10	16.7%
RADI	5	5	100%
SPINE	20	10	65.6%
	25	15	16.1%
	10	10	13.9%
	25	10	1.6%
	10	15	1.1%
	20	15	1.1%
	10	5	0.6%
THOR	20	10	54.3%
	5	5	37.2%
	10	5	8.5%
UROL	5	5	43.3%
	10	5	37.4%
	10	10	13%
	20	10	5.1%
	15	10	1.2%
VASC	10	5	75.1%
	20	10	23.7%
	10	10	0.5%
	25	15	0.5%
	15	15	0.2%

6.4.3. *Distributions for ward LOS.*

Service type	Priority		
	P1	P2	P3
DENT	LOGN(9.26, 113)	LOGN(2.46, 5.42)	LOGN(1.15, 0.965)
ENT	LOGN(6.2, 128)	LOGN(2.28, 6.9)	LOGN(9.26, 113)
GENL	LOGN(39.7, 523)	LOGN(4.42, 8.93)	LOGN(7.37, 15.6)
GYNE	LOGN(1.7, 2.71)	LOGN(1.26, 1.75)	LOGN(0.72, 0.889)
NEUR	LOGN(37.1, 1e+003)	LOGN(11.2, 23.6)	LOGN(7.7, 9.71)
OPHT	LOGN(9.26, 113)	LOGN(9.26, 113)	LOGN(9.26, 113)
ORTH	LOGN(5.07, 13.6)	LOGN(8.19, 10.8)	LOGN(9.46, 15.2)
OTHER	LOGN(9.26, 113)	LOGN(9.26, 113)	LOGN(9.26, 113)
PLAS	LOGN(2.96, 2.51)	LOGN(9.98, 67.3)	LOGN(6.18, 29.5)
SPINE	LOGN(10.3, 21.3)	LOGN(14.5, 24.4)	LOGN(8.78, 11.4)
THOR	LOGN(9.26, 113)	LOGN(9.26, 113)	LOGN(9.26, 113)
UROL	LOGN(3.31, 4.07)	LOGN(4.78, 7.41)	LOGN(3.46, 5.66)
VASC	LOGN(21.3, 130)	LOGN(11.7, 28.3)	LOGN(9.65, 15.1)

Bibliography

- [1] Chaoulli v quebec, attorney general (2012). URL [*en.wikipedia.org/wiki/Chaoulliv.Quebec\(AttorneyGeneral\)*](http://en.wikipedia.org/wiki/Chaoulliv.Quebec(AttorneyGeneral))
- [2] Department of justice - canada health act (2012). URL <http://laws-lois.justice.gc.ca/eng/acts/C-6/page-1.html>
- [3] Ministry of community and social services (2012). URL [*http : //www.ontario.ca/en/yourgovernment/009831.html*](http://www.ontario.ca/en/yourgovernment/009831.html)
- [4] Ontario wait times - provincial summary (2012). URL <http://edrs.waittimes.net/En/ProvincialSummary.aspx?view=0>
- [5] Ontarios chronic disease prevention and management framework (2012). URL [*http : //www.toronto.ca/health/resources/tcpc/pdf/conferencelee.pdf*](http://www.toronto.ca/health/resources/tcpc/pdf/conferencelee.pdf)
- [6] The ottawa hospital (2012). URL [*http : //en.wikipedia.org/wiki/TheOttawaHospital*](http://en.wikipedia.org/wiki/TheOttawaHospital)
- [7] Ayvaz, N., Huh, W.: Allocation of hospital capacity to multiple types of patients. *Journal of Revenue and Pricing Management* **9**, 386–398 (2010)
- [8] Barau, B., Rovere, M., Skinner, B.: Waiting your turn: Wait times for health care in canada,2011 report. *Research Studies* **01** (2011)
- [9] Begen, M., Queyranne, M.: Appointment scheduling with discrete random durations. Under review
- [10] Belien, J., Demeulemeester, E.: Building cyclic master surgical schedules with leveled resulting bed occupancy. *European Journal of Operational Research* **176**, 1185–1204 (2007)
- [11] Blake, J., Donald, J.: Mount sinai hospital uses integer programming to allocate operating room time. *Interfaces* **32**, 63 (2002)
- [12] Cardoen, B., Demeulemeester, E., Belin., J.: Operating room planning and scheduling: A literature review. *European Journal of Operational Research* **201**, 921–932 (2010)
- [13] Charnetski, J.: Scheduling operating room times with early and late completing penalty cost. *Journal of Operations Management* **5**, 91–102 (1984)
- [14] Chow, V., Puterman, M., Salehirad, N., Huang, N., Atkins, D.: Reducing surgical ward congestion through improved surgical

- scheduling and uncapacitated simulation. *Production and Operations Management* **20**, 418–430 (2011)
- [15] Conforti, D., Guerriero, F., Guido, R.: Non-block scheduling with priority for radiotherapy treatments. *European Journal of Operational Research* **201**, 289–296 (2010)
- [16] Creemers, S., Belien, J., Lambrecht, M.: The optimal allocation of server time slots over different classes of patients. *European Journal of Operational Research* **219**, 508–521 (2012)
- [17] Denton, B., Gupta, D.: A sequential bounding approach for optimal appointment scheduling. *IIE Transactions* **35**, 1003–1016 (2003)
- [18] Denton, B., Viapiano, J., Vogl, A.: Optimization of surgery sequencing and scheduling decisions under uncertainty. *Health Care Management Science* **10**, 13–24 (2007)
- [19] Dexter, F., Epstein, R., Traub, R., Xiao, Y.: Making management decisions on the day of surgery based on operating room efficiency and patient waiting times. *Anesthesiology* **101**, 1444–1453 (2004)
- [20] Dobson, G., Hasija, S., Pinker, E.: Reserving capacity for urgent patients in primary care (2011). To be Published in *Production and Operations Research*
- [21] Gerchak, Y., Gupta, D., Henig, M.: Reservation planning for elective surgery under uncertain demand for emergency surgery. *Management Science* **42**, 321–334 (1996)
- [22] Green, L., Savin, S.: Reducing delays for medical appointments: A queuing approach. *Operations Research* **56**, 1526–1538 (2008)
- [23] Green, L., Savin, S., Wang, B.: Managing patient service in a diagnostic medical facility. *Operations Research* **54**, 11–25 (2006)
- [24] Gul, S., Denton, B., Fowler, J., Huschka, T.: Bi-criteria scheduling of surgical services for an outpatient procedure center (2010)
- [25] Gupta, D., Wang, L.: Revenue management for a primary-care clinic in the presence of patient choice. *Operations Research* **56**, 576–592 (2008)
- [26] Ho, C., Lau, H.: Minimizing total cost in scheduling outpatient appointments. *Management Science* **38**, 1750–1764 (1992)
- [27] Hospital, T.O.: Program analyst, clinical programs (2012). Personal correspondence
- [28] Hudson, A.: Commentary: Ontario’s efforts to reduce time spent in hospital emergency departments (2012). URL <http://www.longwoods.com/content/20885>
- [29] Kim, S., Giachetti, R.: A stochastic mathematical appointment overbooking model for healthcare providers to improve profits. *IEEE Transactions on Systems, Man, and Cybernetics-Part A*:

- Systems and Humans **36**, 1211–1219 (2006)
- [30] Kirby, M.: Review of ontarios wait time information system. Ministry Report (2007)
 - [31] LaGanga, L., Lawrence, S.: Clinic overbooking to improve patient access and increase provider productivity. *Decision Sciences* **38**, 251–276 (2007)
 - [32] Muthuraman, K., Lawley, M.: A stochastic overbooking model for outpatient clinical scheduling with no-shows. *IIE Transactions* **40**, 820–837 (2008)
 - [33] Oddoye, J., Jones, D., Tamiz, M., Schmidt, P.: Combining simulation and goal programming for healthcare planning in a medical assessment unit. *European Journal of Operational Research* **193**, 250–261 (2009)
 - [34] Patrick, J., Puterman, M., Queyranne, M.: Dynamic multi-priority patient scheduling for a diagnostic resource. *Operations Research* **56**, 1507–1525 (2008)
 - [35] Rossetti, M.: *Simulation Modeling and Arena*. John Wiley and Sons (2010)
 - [36] Santibanez, P., Begeen, M., Atkins, D.: Surgical block scheduling in a system of hospitals: An application to resource and wait list management in a british columbia health authority. *Health Care Management Science* **10**, 269–282 (2007)
 - [37] Slack, N.: *The Blackwell Encyclopedic Dictionary of Operations Management*. Wiley Blackwell (1999)
 - [38] Werker, G., Saure, A., French, J., Shechter, S.: The use of discrete-event simulation modelling to improve radiation therapy planning processes. *Radiotherapy and Oncology* **92**, 76–82 (2009)
 - [39] Yellig, E., Mackulak, G.: Robust deterministic scheduling in stochastic environments: the method of capacity hedge points. *International Journal of Production Research* **35**, 369–379 (1997)
 - [40] Zeng, B., Turkcan, A., Lin, J., Lawley, M.: Clinic scheduling models with overbooking for patients with heterogeneous no-show probabilities. *Annals of Operations Research* **178**, 121–144 (2009)