OPERATIONS MANAGEMENT TOOLS APPLIED TO THE OPERATING ROOM: A REVIEW OF CURRENT CONCEPTS AND A SINGLE CENTRE EXPERIENCE

by

Charles Yuan-Hui Lo
Bachelor of Science (Honours), University of British Columbia 2001
Doctor of Medicine (Honours), University of Toronto 2005

PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF BUSINESS ADMINISTRATION

In the
Faculty
of
Business Administration

SIMON FRASER UNIVERSITY

Summer 2009

Approval

Name: **Charles Yuan-Hui Lo**

Degree: **Master of Business Administration**

Title of Project: **Operations Management Tools Applied to the Operating Room: A Review of Current Concepts and a Single Centre Experience**

Supervisory Committee:

____

Dr. Andrew von Nordenflycht
Senior Supervisor
Assistant Professor
Faculty of Business Administration

____

Dr. Sudheer Gupta
Second Reader
Assistant Professor
Faculty of Business Administration

Date Approved: ______________________________________

Abstract

Operations management tools can be applied to the operating room setting in order to improve throughput of the system. This is important because of the limitation of resources and funds available to hospitals in the public healthcare system. Hospitals must deal with variability in demand and uncertainty surrounding scheduling; these considerations can be placed in a queuing theory framework to better design processing capacity to minimize wait times and maximize utilization. Lean techniques can be used to identify and reduce waste in processes. A single centre experience is presented to demonstrate real-world application of these tools and to suggest foci for potential improvement in other centres.

Keywords: operations research; Lean management; queuing theory; healthcare operations; operating room utilization; surgery wait times.

Dedication

This work is dedicated to Annica Tong for helping me to understand who I have been and who I might one day be.

Thank you to my parents for your perpetual support and patience.

Acknowledgements

Thank you to my essay supervisors, Drs. Andrew von Nordenflycht and Sudheer Gupta for their insight, support and guidance throughout this process. I am indebted to you and your professionalism and expertise.

I would like to thank the anaesthesiology residency program at the University of British Columbia and the program director, Dr. Clinton Wong, for allowing me the latitude in my training to explore my interests ancillary to medicine. I look forward to returning this investment through contributions to the enrichment of the profession and helping to fulfil our responsibility to continually strive for improvement in our practices.

Finally, I would like to thank my classmates and colleagues this year in the MBA program at Simon Fraser University. I have learned as much from you as I have from any text or lecture; to you, I wish ever increasing success in all your individual professional and personal pursuits.

Table of Contents

List of Figures

List of Tables

Glossary and Abbreviations

OR	Operating room
RN	Registered Nurse
TAT	Turnaround Time
FCST	First-Case Start Time
Poly-trauma	Multiple injuries resulting from trauma
Lean	Production practice that considers the expenditure of resources for any goal other than the creation of value for the end customer to be wasteful, and thus a target for elimination
Queuing Theory	The mathematical study of waiting in lines
Operating Suite	Collection of operating rooms and ancillary facilities including recovery room and preoperative holding area
Triage	Prioritization of medical care based on perceived level of urgency
Emergency Department	One gateway to hospital care and admission. High acuity care provided by emergency physicians and RNs
VGH	Vancouver General Hospital (also called the Vancouver Acute)
VA	Vancouver Acute site of the Vancouver Coastal Healthy Authority
Recovery Room	Facility for the recovery from anaesthesia and post surgical stabilization prior to transfer to hospital ward or home
Preoperative Holding Area	Facility where patients are marshalled for surgery and where RNs perform a detailed pre-surgical screening and patients are seen by anaesthesiologists and surgeons
Preadmission Clinic	Clinic where patients are seen in advance of surgery (days to weeks) in order perform further pre-surgical workup and optimization
Disposition	Planned or intended trajectory of patient care
Discharge	Release of a stable patient home
Inpatient	Patient that requires in hospital care for longer than one day
Outpatient	Patient that visits the hospital for care but returns home on the same day
Stable	Denotes a blood pressure, heart rate, breathing, pain control and level of consciousness that is within normal limits
Unstable	The converse of Stable. Requiring immediate or ongoing medical attention

1: Introduction

Resource limitations are real in the Canadian public healthcare system. In British Columbia, the media regularly reminds the public that health services are being cut in already under-serviced areas and that surgeries are being cancelled or operating rooms being closed because of budget limitations (Hunter, B.C. health authorities told to trim $360 million, 2009). Unlike the private healthcare model that exists in other parts of the world, the health services industry that we depend upon for medically essential care does not generate revenues to support its operation. Public hospitals and provincial health authorities are constrained by mandated budgets; they face increasing costs, but cannot pass those costs on to patients like a private enterprise would. Recent acceleration in rising healthcare costs, due partially to an aging population and concomitant increasing demand, increasing prices of pharmaceuticals, and capital expenditure on equipment and health information technology, has outpaced government spending in British Columbia. Stakeholders in the system, including physicians and nurses, are increasingly alert to the possibility that our current standard of quality may not be sustainable.

As health professionals charged with the responsibility of administering public health spending on behalf of the government but in the interest of the people, we physicians must consider novel approaches to reducing costs while constantly seeking improvement in quality and outcomes for our patients. The professional culture of medicine in Canada guides physicians not to make medical decisions on the basis of costs; one would not prescribe an inferior course of treatment solely because it is less expensive. This is borne from the single-payer system that has been deeply entrenched in the public system and that improves access and equality. At the same time, rationality informs most decision-making; unnecessary or marginally beneficial investigations are discouraged as a means to keep costs low. Another approach to minimizing costs is to address inefficiencies in the system; the removal of non-value added steps or increasing utilization may result in greater sustainability in the long-term.

Operations management techniques have only recently found their way to healthcare operations (Hunter, A clear vision for better health care, 2009). The reluctance to translate industrial efficiency frameworks to health and human systems is likely the result of cultural resistance and the belief that health processes are somehow unique. There exists a belief that humans, both the patients and the practitioners, are autonomous and should not be treated as widgets or cogs in a machine. Healthcare processes, however, share considerable similarity to industrial processes: there are inputs and outputs linked by processes and these processes are subject to variability and the development of queues. Clearly, healthcare processes are services rather than a manufacturing assembly line and can be appropriately described in service terms. The intention of this essay is to increase understanding of these health processes within an operations management framework. An example of the application of these concepts in a complex academic institution is used to illustrate potential opportunities in process redesign and efficiency gains. It is intended that by framing these strategies in operations management, that benefits may be better communicated and translated to other institutions.

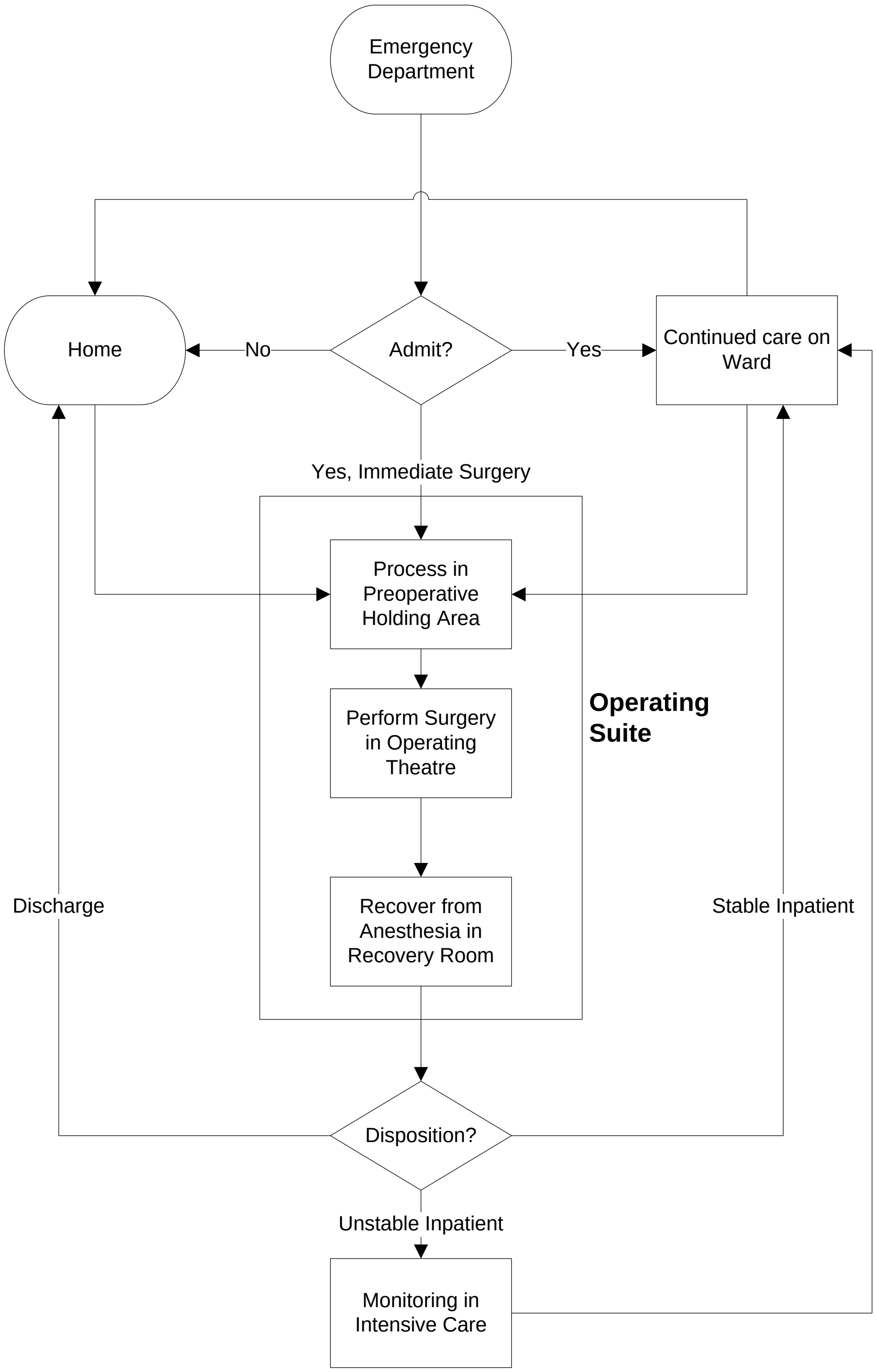
1.1 The operating room as a process

Operating rooms (ORs) are known to be one of the largest cost centres of hospitals accounting for over 40% of expenses (Denton, Viapiano, & Vogl, 2007). The costs associated with running the ORs include high fixed costs arising from the facilities, equipment, and other capital expenditure and high variable costs including nursing and allied health labour wages. Physicians and surgeons, in Canada, are for the most part paid directly from the province on the basis of a fee-for-service schedule and in fact are not considered employees of the hospital at which they work; rather physicians are granted “privileges” to use the hospital facilities for their medical practice.

Unlike in the United States, where most of the healthcare operations research is performed and where operating rooms function as profit centres, in Canada, the vast majority of hospitals are publicly owned and are provided with a fixed annual budget by each respective provincial government. Each hospital’s board determines the allocation of its funds and resources to the various departments and hospital areas. In British Columbia, an intermediate level of governance exists in that the province is currently divided into 5 regional and one province-wide health authorities. Canadian hospitals, not uniquely, are facing increasing constraints in annual budgets as the cost of healthcare delivery steadily increases (Hunter, B.C. health authorities told to trim $360 million, 2009) while governments pursue innovative yet costly initiatives such as reductions in CO_2 emissions and carbon neutrality (Livesmart BC, 2009). As operating rooms constitute a major component of hospital spending, it is logical to continually seek improvements to operating efficiency in order to limit costs while maintaining quality and continuing innovation.

The operating suite occupies a central position in the process of a hospital admission and surgical services are delivered in a number of different environments and contexts. Figure 1.1 describes the flow of patients through a hospital admission.

Figure 1.1: The operating suite is central to hospital operations



Entry to the operating room can occur as an inpatient or and an outpatient. Inpatients are admitted to the hospital prior to surgery through the emergency department or directly to the ward. They have their surgery performed and then following recovery in the post-operative care unit or the intensive care unit, they return to the hospital ward before returning home. Outpatients arrive to the operating room from home the morning of surgery and following surgery and postoperative recovery are returned home the same day. Outpatient procedures are typically simple or short surgeries with low postoperative pain control difficulties, low risks of complications, and are performed on generally healthy people. The need for surgery can be either elective or emergent.

Elective surgery describes a procedure for a condition that is not immediately life-threatening where a delay will not significantly increase risks of morbidity or mortality; elective surgery can be scheduled in advance and is almost exclusively performed during daytime hours. Elective surgery is by no means always simple and routine. Some elective procedures are highly complex and require the marshalling of extensive hospital resources and the cooperation of many services and can be associated with the potential of serious complications. The term “elective” is effective in communicating a connotation of degree of urgency amongst healthcare workers or in the administration of policies. The term, however, may belie the seriousness of the underlying condition or the degree of impairment or loss of function associated with injury; for example, many cancer surgery procedures (mastectomy for breast cancer, bowel resection for colon cancer, and others) or spine surgeries, or even open-heart coronary bypass surgery are considered elective as they can be scheduled in advance.

In contrast, emergency surgery describes procedures that must be performed within a certain time frame or risk the loss of "life, limb or organ." These surgeries enter the operating suite through the emergency room, the hospital ward, or through the intensive care units and often necessitate a long postoperative course in hospital. Examples of emergency surgery include the control of bleeding in penetrating chest trauma, craniotomy to control bleeding from a ruptured aneurysm, ruptured appendicitis, fractured femur from a fall, or organ transplantation amongst many others. These surgeries may arrive at any time throughout the day and on any day of the week and multiple emergency surgeries may arrive simultaneously at a tertiary referral hospital. Each situation or surgery is associated with a time frame within which surgery should optimally be performed in order to minimize risk of morbidity or mortality. This time frame will depend upon the specific type of underlying condition, the severity of the condition, and the co-morbid status of the patient, in other words, the other issues and extenuating status of the patient. The surgeon who assesses the patient and requests the operating room determines the time frame within which surgery must be performed; commonly, a surgical procedure required within 24 hours is considered an emergency. Requirements amongst emergency surgeries, however, can range from less than 24 hours to immediate. Appendix A lists a number of surgical procedures and their associated urgency.

1.2 Scheduling elective and emergency surgery

Efficient scheduling of surgery in a large hospital with several operating rooms is fraught with significant complexity. ORs typically have a planned utilization time beyond which labour costs accrue significantly; therefore a constant pressure exists to optimize the number of cases that are being performed within that time. However, even the elective surgical environment suffers from a high degree of variability. This variability stems from individual differences in patients, differences in their pathology, and differences in skill and technique of the practitioners including surgeons, anaesthesiologists, nurses, and trainees. Two systems for scheduling surgery are block scheduling and open scheduling (Denton, Viapiano, & Vogl, 2007).

1.2.1 Block scheduling

In block scheduling, individual surgeons or surgical departments are assigned blocks of time in individual ORs on a set schedule (for example, general surgeon Dr. Smith is assigned 8 hours of time in OR#2 each and every Monday). The surgeon then books her own elective cases into her allotted time such that the mean duration of her cases fits within the block of time. The surgeon may be responsible for filling the entire available time or his or her colleagues from the same surgical service may be allowed to fill in any remaining time. Commonly OR#2 would also be used for general surgery on other days but with another surgeon from that service occupying that time.

Block scheduling allows for specialization of operating rooms and its associated staff and equipment. For example, as neurosurgery differs significantly from orthopaedic joint reconstruction surgery, the highly specialized microscope used for micro-dissection can be housed inside the neurosurgery operating room, while the image intensifier or Xray machines can be located near the orthopaedic and trauma rooms. In the same arrangement, surgical scrub nurses can specialize in certain surgical subspecialties so that they perform with greater consistency and speed. The specialization of operating room nurses into teams of specialization has been shown to increase employee satisfaction amongst nurses (Sprengel, Snell, & Boissoneau, 1993).

1.2.2 Open scheduling

In the open scheduling system, all surgeons submit cases up until the day of surgery and these cases are scheduled into various ORs such that all are accommodated. The same operating room may be scheduled to accommodate different surgeons and different surgical specialties on any given day. Emergency surgeries are often scheduled into an emergency operating room in an open system.

One advantage of the open scheduling system is that the available operating room time can be fully utilized as long as there is a queue of cases available to be scheduled, unlike in the block system where surgeons are allocated time whether or not they have the need for it on a given day. On the other hand, emergency surgeries are subject to more randomness in arrivals and the time required to perform surgery, resulting in adverse consequences on operating room utilization, as discussed in Chapter 2.

1.2.3 Emergency scheduling

Emergency surgery complicates scheduling in operating rooms by introducing increased uncertainty. Emergency procedures can be booked at any time of the day and can have a variable window of time within which they must be done. Moreover, emergency surgeries, because of their unpredictable nature, can have different resource demands. For example, an emergency multiple digit (finger) replant surgery may require one or two surgeons, two nurses and an anaesthesiologist but may require several hours of operating time. In contrast, an emergency Caesarean section may require two surgeons, two or three surgical nurses, an anaesthesiologist, a paediatrician and a team of neonatal nurses but only 45 minutes of operating time. In a small community hospital with only a few operating rooms, the difficulty of accommodating extra or add-on surgery is easily apparent; if emergency surgery is to be done in a timely way, elective surgery will need to be delayed until the room or another room and its staff becomes available. This delay may extend several hours or the delay may force the postponement of the surgery to another day.

Large centres face similar dilemmas, although a larger number of rooms are available to accommodate add-on surgical cases, these rooms are already maximally scheduled to optimize the number of cases being performed. If already maximally utilized, diverting an operating room to attend to an emergency case will still force a delay and possible postponement just as in the small hospital. Large centres, because of the nature of the tertiary referral process (wherein complex or high-acuity cases are funnelled from local and regional hospitals into a small number of specialized centres), face the situation of emergency add-on surgery far more frequently and often with more complex, time and resource consuming procedures.

The question that can then be asked is: does there exist a system of scheduling that accounts for the variability and uncertainty that emergency surgeries introduce to the already variable nature of operating room scheduling that is able to maximize the utilization of the resources of the operating room, while still achieving other important goals, including sustaining high quality of care (as measured in general by morbidity and mortality or more specifically by the rate of complication or the need for “redo” surgery), minimizing excessive costs (such as overtime pay for surgical teams), and maintaining or even reducing surgical wait times?

The following chapter will describe operations management in the context of the operating room setting. Lean management and queuing theory will be introduced as frameworks that help to shape strategic decisions about redesigning various operating room processes and scheduling in order to deal with this unique variability. Additionally, metrics will be discussed that guide improvements and form the basis of benchmarking within the industry. The third section of this paper will discuss the efforts of the Vancouver General Hospital to redesign several of its practices, framed within an operations approach, in order to better achieve targets for its performance, utilization, and quality in the operating room.

2: Operations Management in the Operating Room Setting

2.1 Queuing theory

The operating room and the progress of a patient through the operative experience can be thought of as a service process that has an input (the patient requiring surgery), a service (the operation and recovery period), and an output (the patient discharged home or to the hospital ward or intensive care unit, or patient expiry). As with any resource constrained process there is a limited capacity and should demand overwhelm capacity a queue will necessarily form. The mathematical tool devised to optimize service processes in which queues form was first described by A. K. Erlang (Queuing Theory, 2009) in the early 1900s and is termed "queuing theory". In attempting to determine the number of switches required to handle telephone calls arriving at a switchboard, Erlang concluded that the two key variables to determining the optimal, or fewest, number of switches required to produce a tolerably short wait time for callers were the length of each telephone call and the time between arrivals of calls. Applications of this theory have been extended to countless management and business scenarios including determining the number of bank tellers required to limit the length of time patrons stand in line, or optimizing the supermarket checkout and the development of express checkout lanes.

There have been many applications to healthcare as well, including of particular interest, application to emergency room wait times (Schlechter, 2009) and level of staffing (Fomundam & Herrmann, 2007) as well as the surgical treatment of trauma patients requiring emergency operations. The latter study (Tucker, Barone, Cecere, Blabey, & Rha, 1999) looked specifically at the need to call in a second operating room team at night in order to deal with the formation of a queue of trauma surgery cases. This solution deals with only a single attribute of the system as described by queuing theory, the number of servers or channels. The number of channels describes the capacity of the system to process inputs simultaneously (in this context, number of operating rooms). There are other parameters of service processes and queue formation that can be characterized and effectively modified to alter the performance of the system, as elaborated below (Anupindi, Chopra, Desmukh, Van Mieghem, & Zemel, 2006).

The emergency surgery process can be described in terms of queuing theory by the parameters (1) arrival of patients which can be measured as a rate or an interval time between arrivals, (2) service which is specified best by the length of time required for the surgical procedure, (3) the number of service channels which describes the number of available operating rooms, (4) the number of queues which are allowed to form and (5) queue discipline (e.g., the ability of one surgery to “bump” another based on personnel, equipment, logistical reasons and the prioritization of certain cases over others).

Arrivals to the operating room are subject to considerable variability and uncertainty. Variability in this context refers to the potential distribution of all possible events. Intuitively, the arrival of emergency patients does not follow a predictable pattern and there is considerable randomness to arrivals such that inputs cannot be considered deterministic. Uncertainty describes the somewhat different element of errors in prediction that arise from insufficient information.. For this essay variability and uncertainty will be treated as the same and will be used interchangeably. They both introduce the need for greater capacity and/or inventory into the queuing system.

Historical patient arrival data for individual institutions may suggest some patterned behaviour but overlaid by elements of randomness; inputs to this queuing model have been considered by other authors as "stochastic" (Lamiri, Xie, Dulgui, & Grimaud, 2008) and can be described in terms of a mean (the average number of patients arriving in a given time period) and a standard deviation. Similarly, the length of time required for a surgical procedure is not constant and again there is considerable variability and uncertainty. For example, a patient who has sustained a life threatening poly-trauma may occupy an operating room for several hours if his injuries are survivable or repairable or he may succumb to his injuries within minutes of starting surgery if the sustained trauma is severe enough. The extent of the injuries may be unknown until a surgeon has been able to directly assess them and so accurate predictions of the length of time required in the operating room cannot be achieved with any certainty, Some surgeries can be completed in less than one hour and some may require several hours. Herein is one substantial difference between elective surgical scheduling and emergency scheduling; there is less, although not zero, variability in service processing time for elective procedures and arrival times are controlled. This allows for greater utilization rates as will be discussed below.

The fifth attribute of prioritization of service and queue discipline warrants special consideration with respect to emergency surgery. The term emergency surgery encompasses a broad range of urgencies and surgeries are attached a prioritization based upon urgency and the length of delay associated with increases in morbidity or mortality.

Many examples of prioritization exist in medicine that are not necessarily seen in many other service or queuing examples. The emergency department triage system designed and utilized in Canada, the Canadian Triage and Acuity Scale is an example of prioritization wherein the most serious emergency cases such as chest pain, or loss of consciousness are seen first over less serious emergencies such as cuts or other injuries. In emergency surgery scheduling, cases are given a prioritization (in decreasing order) of (1) less than one hour or STAT case, (2) less than eight hours, (3) less than 24 hours, or prioritization (4) "urgent elective." Increased complexity to the queuing model is introduced by this set of prioritizations: cases cannot be processed through the queue strictly on a first-in, first-out basis, nor can lower priority (priority 3 or 4) cases be put on hold indefinitely in order to wait and accommodate more urgent cases. Within a given level of prioritization, a first-in, first-out system can be used but even within this simplification there may be numerous adjustments based on availability of specialty nursing staff, preparation or turnover of specialized tools and equipment, and agreements between surgeons to "bump" one another, or availability of a particular surgeon to do a particular procedure as the surgeon may be otherwise occupied in the emergency department or elsewhere in the hospital.

A copy of the prioritization protocol for the Vancouver Acute hospital is shown in Appendix B and a graphical depiction is seen in Figure 3.2 in Chapter 3.

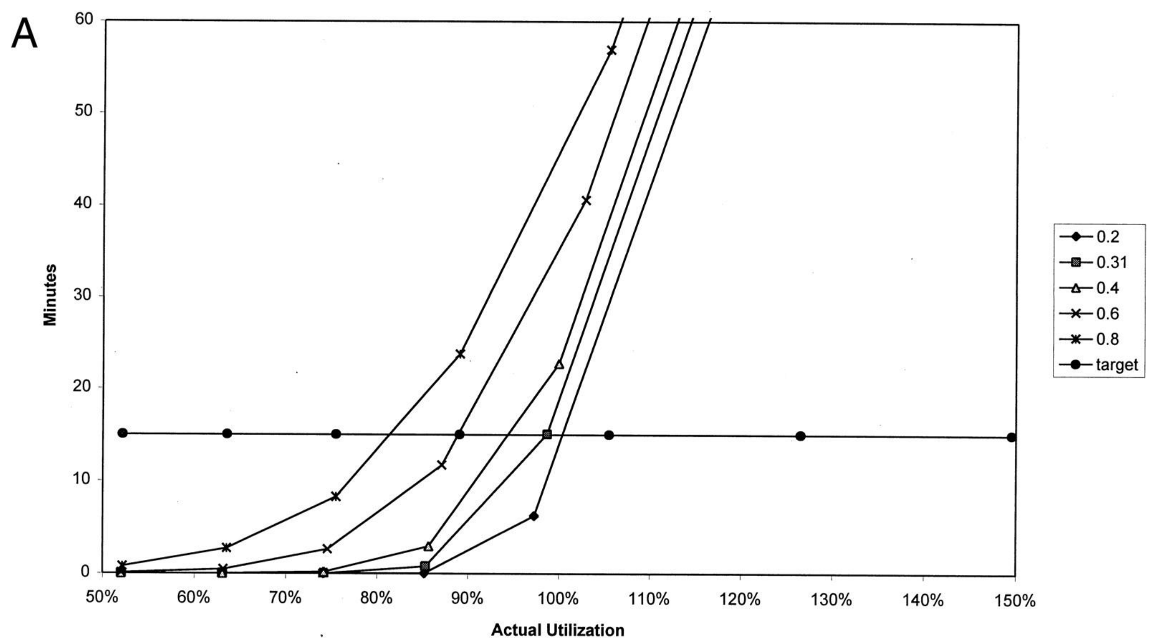
2.2 Utilization

Further, the process can be described in terms of performance characteristic and benchmarked in terms of metrics, of which, utilization is a key measure. The raw utilization of the operating room resources is equal to the time that patients are occupying the operating room (service length) divided by the available resource hours (Kindscher & Rockford, 2009). The adjusted utilization describes the same calculation but factoring in setup and turnover time for each procedure; this represents a truer reflection of utilization of the operating room.

Described in terms of queuing theory, utilization is the throughput divided by the resource capacity. Strum, *et al.* defined the terms overutilization and underutilization as applied to OR scheduling (Strum, Vargas, & May, 1999). Overutilization is the time used by scheduled cases past the end of scheduled time and underutilization is the converse. While utilization is commonly used to track surgery performance and to allocate blocks of surgical time, there is no clear optimal utilization, where the economic costs of staffing potential overutilization is balanced with the risk of underutilization, although some authors have attempted to determine it under specific conditions (Tyler, Pasquariello, & Chen, 2003). Most operating rooms in the United States seek a utilization of 70 to 85%. More aggressive utilization targets force the trade-off of availability of operating time for surgeons to book new patients and risk decreases in patient satisfaction (Kindscher & Rockford, 2009). Expressed in terms of efficiency, a low utilization in the face of high fixed costs is a reflection of wasted resources and inefficiency while utilization in excess of 100% represents additional costs associated with overtime nursing wages or surgeon and anaesthesiologist discomfort and resistance (*ibid*). Even high utilization (85% or higher) is considered inefficient because of the inability to deal with variable demand; in effect, a highly utilized system lacks adequate surge capacity to meet excessive inputs.

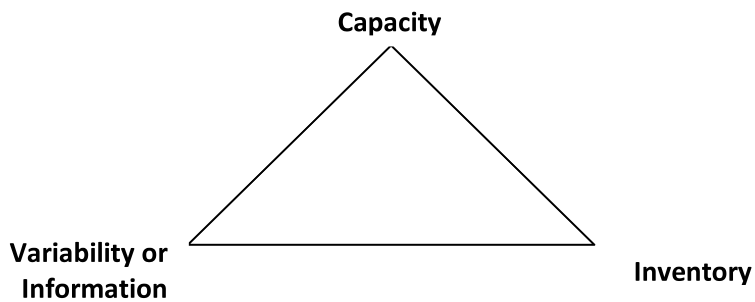
It has been shown that an increase in variability of case duration reduces the achievable utilization (Tyler, Pasquariello, & Chen, 2003). This trade-off between variability and utilization can be depicted by a throughput-delay curve such as the one shown in Figure 2.1; Tyler has modelled this trade-off as it relates to the operating room setting (*ibid*):

Figure 2.1: Tyler's model of a throughput delay curve for OR utilization



This model helps to illustrate one of the fundamental trade-offs in the operations management triangle described by Lovejoy (1998). This author described the relationship between capacity, variability and inventory (queue length) schematically as:

Figure 2.2: Lovejoy's OM Triangle



The principle behind the Operations Management triangle can be derived mathematically with Little's Law and the Pollaczek-Khinchin (PK) formula (Appendix C). The PK formula, while derived for a single-server model and based on an assumption of a first-in, first out process, can be generalized to the more complex scenario such as emergency operation scheduling.

Qualitatively, it suggests that there are necessary trade-offs between capacity, variability, and inventory or buffers. An ideal system would be one where capacity is fully utilized despite maximal variability and minimal inventory; applied to the operating room, the ideal system would be able to deal with any surgical case with 100% utilization of its operating room resources without the need to maintain a long queue of waiting surgical cases. This of course is not a feasible system; trade-offs necessitate that if variability is high, then capacity is underutilized and the buffer size (patient queue) is large. If the goal of the system is to minimize queue size, then either capacity must be increased or variability decreased, or both (Lovejoy, 1996).

We have already discussed how emergency surgery is associated with a greater amount of variation for operation duration and so utilization is expected to be low or waiting time for surgery is expected to be high. Posed as a queuing problem then, we may ask, can any of the parameters of the queue be manipulated in order to achieve a more optimal design of tolerable waiting times, cost-effective capacity utilization and reasonable accommodation of variability?

The added complexity to this system is that it is not a simple, single-server, first-in, first-out process. Additional operating rooms may be employed at night for special situations (transplant surgery, cardiac surgery, other subspecialty procedures), and there is a formal prioritization of certain cases over others. Emergency scheduling is especially complex in large tertiary referral hospital where emergency demand is high both day and night and elective surgeries are already scheduled to optimize utilization. The experience of one large teaching hospital, the Vancouver Acute site of the Vancouver Coastal Health Authority in British Columbia will be discussed in Chapter 3. Vancouver Acute, also known as the Vancouver General Hospital (VGH) has put in place initiatives that attempt to deal with the variability component of the OM triangle in order to achieve targets of tolerable wait times for urgent (priority 4) and emergent (priority 1 to 3) surgical cases.

2.3 Lean management in the operating room

In addition to queuing theory, another relevant operations management framework is the concept of “lean production”. Lean tools and techniques have recently and increasingly been applied to a wide variety of healthcare applications including office appointment scheduling, hospital bed allocation and emergency room visits (Fomundam & Herrmann, 2007), (Schlechter, 2009). Lean and other operations tools have also found application and benefit in the operating room primarily with regards to supply chain management and reducing inventory levels to reduce unnecessary costs and in minimizing excessive movement of patients and healthcare providers to improve efficiency through such strategies as the implementation of factory flow and fixed position processing (Slack, Chambers, & Johnston, 2007), and ensuring quality while reducing costs (Spear, 2005). There are numerous tools in the Lean framework in addition to queuing theory that can be carried over to the planning and sequencing of emergency surgery in a complicated tertiary referral centre.

The Toyota Production System (TPS) on which the Lean framework is based, describes seven categories of waste in a manufacturing or production system (McLaughlin & Hays, 2008). Of the seven, six have direct parallels in healthcare services and some can be applied in a more abstract fashion. Waste in the operating room can arise from overproduction, excessive waiting, unnecessary transportation, unnecessary motion of providers, over-processing and inventory.

Overproduction refers to initiating production prior to the demand; in the OR this refers to setting up instruments or specialized equipment for a surgical case that has not been confirmed or that is eventually bumped by another higher priority case. Waste is incurred in having to spend time and cost repackaging and sterilizing open equipment.

Waiting is perhaps the most obvious dimension of waste in the OR; patients may wait several hours in the emergency room after a surgical booking is made but prior to being transported to the operating room. This waiting time incurs costs upstream by preventing patient flow in the emergency room, and incurs costs to the patient because prolonged delays may result in increased morbidity. Waiting can also occur at or prior to several other steps in the process as illustrated in value stream map, Figure 2.3. Waiting occurs on the ward prior to being called to the operating room, in the preoperative holding area, in the operating room prior to induction of anaesthesia, and after surgery but before being cleared to enter the recovery room and more. Waiting parallels the concept of queue formation in queuing theory.

Transportation and motion of providers lead to waste by contributing to additional time not being utilized in direct care of the patient. For example, waste arises if the circulating nurse must make repeated trips into the operating room sterile core to retrieve specialized equipment, incurring procedure delays, opportunities for miscommunication, and opportunities for injury, when the same equipment may be stored directly in the operating room or immediately adjacent to the room.

Over-processing in the OR environment refers to redundant or unnecessary steps in surgical care ranging from excessive or duplicate documentation, to performing surgery and finding no pathology. The role of over-processing in medicine may differ from its role in manufacturing. In manufacturing processes, steps that are deemed to not add value may be eliminated without affecting the finished product. In medicine, over-processing is built into the system because it serves as contingency against human error. Where there are multiple layers of redundancy, it is less likely that an error will be carried forth (Gaba, Fish, & Howard, 1993).

Inventory, in the context of the Lean production system, refers to the holding or purchasing of raw materials including supplies, for example prosthetic joints of many different sizes, coronary stents, and pharmaceuticals. Inventory can also refer to work-in-process and as such can also refer to patients in a queue (which then relates to waiting time).

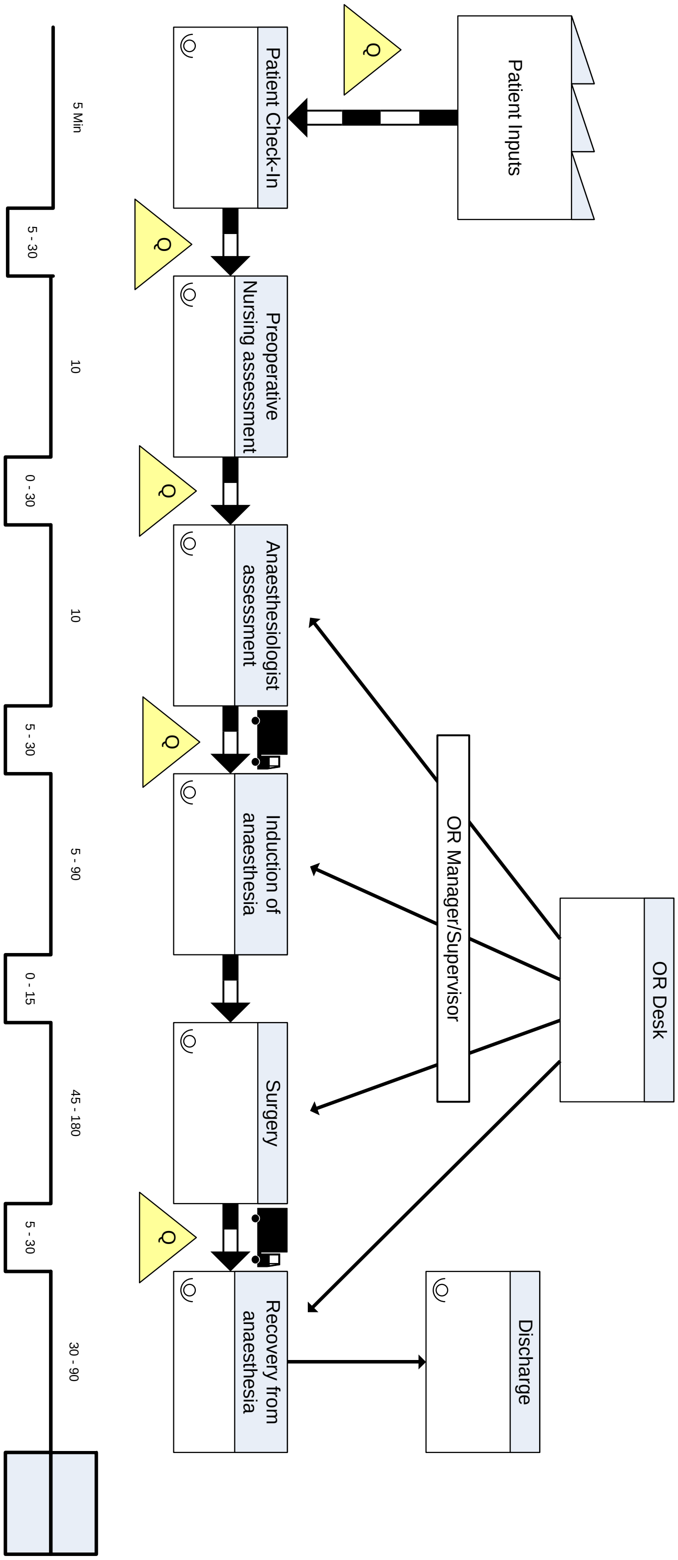
Further, because of the uncertainty inherent in diagnosis and imperfect diagnostic testing, few tests are truly 100% specific or sensitive. That is, in any diagnosis there is a chance that the result may be, in evidence based medicine parlance, a false-positive (type 1 error) or a false-negative (type 2 error). In order to minimize the number of false-negative results, our diagnostic tests may result in a greater number of false-positive results. These false-positives may rarely result in unnecessary surgery. A classical example is that of appendicitis. Prior to the invention of the CT scanner and the abdominal ultrasound, acute appendicitis was entirely a clinical diagnosis based upon a classical history and certain physical examination findings. Patients were often urgently operated on in the middle of the night to find that in fact the appendix was normal. This was, and often still is, considered appropriate practice because it minimizes the number of missed cases of true acute appendicitis. Historically, the quality of a general surgeon was measured as much by the number of unnecessary surgeries he performed as the number of actually necessary ones.

2.3.1 Value-added steps in the OR process

The operating room and the services provided therein can be approached from two different perspectives. It can be viewed in terms of a manufacturing system with inventories such as surgical instruments and devices and work in process, and transportation of equipment and the movement of workers, all of which can be targeted for efficiency improvements; this approach is well described in operations management research. By contrast, the system can be viewed more globally as a process through which an input, the patient, progresses through a series of steps resulting in a completed surgery, or the output.

The process can be illustrated through a value stream map. The value stream map includes all the steps in the process experienced by the patient including value-added and non-value-added steps. By identifying specific steps that are unnecessary and non-value-added, significant waste can be eliminated. The value stream map of a surgical patient entering the process of the operating room can be depicted as in Figure 2.3:

Figure 2.3: Value stream map of operating room process



Whether a process step is deemed of added value in the private healthcare setting or pay-for-service setting such as we see often in the United States, can be based upon whether or not the patient would be willing to pay for that particular service step. Clearly, a step such as an anaesthesiologist's preoperative examination, or administration of prophylactic antibiotics would be considered as value-added steps, but lying on a stretcher in the preoperative holding area, having fasted for 12 hours while the surgery immediately preceding yours runs overtime, would not be value-added. In a single-payer system where the payer is a third party such as we have in Canada or in the case of private insurance, this question of value-added from the patient perspective becomes less well defined. One option is simply to continue to consider the process from a payer perspective; would a third party or government payer pay for a particular process step. However, for another party to determine whether or not they are willing to pay for the service, some more objective measure may need to be introduced. Instead, another option may be to consider whether or not the process step directly or indirectly affects the outcome of the patient. A step that is considered value-added is one that reduces mortality and/or morbidity and a step that is non-value-added is one that reduces or has no effect on patient outcome. In both scenarios, the most common non-value-added activity is waiting (McLaughlin & Hays, 2008). Patients would not independently pay to wait longer and for most surgical procedures, waiting beyond a certain length of time has a negative impact on outcomes.

2.3.2 Standardization of processes

A core component of Lean is the standardization of work processes. In the hospital operating room setting, processes can be standardized in the form of written documentation, and step-by-step codification of best practices. Clinical pathways for specific types of surgeries are examples of standardized work. The Massachusetts General Hospital (MGH) developed a care path for coronary artery bypass surgery that detailed the optimal timing of specific interventions and clearly identified goals of care for the patient in the day prior to surgery, the day after surgery and in the post-surgical days following. This initiative resulted in a 1.5 day reduction in average length of stay and cost savings (Wheelwright & Weber, 2004). Many hospital intensive care units use standardized pre-printed order sets for such routine processes as admission to the ICU, dialysis, feeding, and electrolyte replacement. These standardized orders do not dictate medical care or replace decision making but instead ensure no requisite orders are forgotten and that significant errors in notation or interpretation by ancillary staff are not made. This in effect improves efficiency by saving time, reducing subsequent redundancy, and decreasing the rate of error.

2.3.3 The operating room as the bottleneck

A bottleneck dictates the throughput of the entire system. In British Columbia, waitlists extend months for some procedures (Ministry of Health Services, 2009) and preoperative patients occupy beds meant for postoperative recovery because there is insufficient operating room availability. In the process of both elective surgery and emergency surgery, the resources that comprise the operating suite act as the bottleneck. These resources include staff as well as equipment and physical space; all of these are restrained by a limited budget that does not adequately reflect demand ; (Hunter, B.C. health authorities told to trim $360 million, 2009). These resources are represented by the number of available operating rooms and the hours that they are available. It is expected that improving the performance of the bottleneck operation should result in overall increased throughput and improved patient flow through the hospital.

2.3.4 Benefits of Lean

Up to this point, we have described the Lean operations framework and the theory underlying queuing. These tools will be qualitatively applied to the operating room setting in particular to describe how these tools can improve patient flow even in emergency cases that cannot be prescheduled and are complicated by significant variability in types of surgery and uncertainty in length of surgery duration.

One potential means to improve patient flow therefore is through the reduction of variability of cases presenting for surgery by limiting the allowable types of surgery. If a further reduction of variability cannot be achieved, then an alternative may be to reduce uncertainty surrounding the variability by improving our tools of prediction. Unfortunately, there remains a component of demand that is clearly random. Trauma events such as motor vehicle accidents or gun violence, or accidents leading to mass casualty are not predictable and therefore operations that are able to maximize utilization and efficiencies despite variability must be incorporated into operating room practices.

2.4 Benchmarking

Management consultants in the industry have identified a number of key performance indicators for operating room performance (OR Benchmarks, 2009). The standard indicators used in benchmarking within the industry include (McKesson, 2009):

- Prime-time (0700H to 1700H) Utilization
- Start-Time Accuracy
  - First case of the day (+/- 5 minutes indicated on OR schedule)
  - Subsequent cases (+/- 15 minutes from time indicated on OR schedule)
- Case Time Effectiveness
  - Patient incision to close (skin-to-skin) time
  - Turnover time
- Estimated Case-Duration Accuracy (+/- 15 minutes from estimated time)
- Add-on Rate Day of Surgery (percent of cases added on day of surgery)
- Total Volume
- Quality Measures
  - Percent return to surgery within 24 hours
  - Compliance with surgical time out
  - Percent patients screened prior to surgery

These key performance indicators suggest processes that can be improved. In the following section strategies will be discussed that attempt to Lean the OR process by emphasizing value-added steps and minimizing non-value-added steps or that aim to reduce or work around the variability of demand and uncertainty placed upon operating rooms dealing with emergencies.

2.5 Strategies

Lean strategies have been applied to inventory and supply chain management successfully in the operating room (Noon, Schiffer, & Crane, 2009), (Patterson, 2009) (Epstein & Dexter, 2000). Lean techniques applied to scheduling may similarly improve efficiency and increase utilization. Consideration will be given to strategies that apply to elective, scheduled procedures and those that address the uncertain demand of emergency add-on procedures.

The strategies that will be described include (1) reducing wait times (non-value added steps) through aiming to start cases on time and reducing turnaround times, (2) precisely predicting the actual length of surgical procedures (reducing uncertainty), (3) adding emergency operating rooms dedicated to specific surgical services (increasing processing channels), and (4) incorporating hybrid scheduling (improving utilization by reducing unscheduled time).

2.5.1 On-Time Starts

Tardy operating room starts have long been recognized as a source of waste in the system and attempts to quantify the impact of these late starts have been made (Dexter & Epstein, 2009), (Overdyk, Harvey, Fishman, & Shippey, 1998). Dexter demonstrated that each minute of tardiness in the start of the first case translated into almost 1.2 minutes of extra operating room costs for a room scheduled to run for at least eight hours.

In one academic centre, the measured incidence of tardiness was approximately 70% in one quarter with 15% of first cases starting twenty or more minutes later than scheduled. The first case of the day, one would expect, would be associated with the least amount of uncertainty and would be most expected to start on time; however, the observed incidence of tardiness is strikingly high (Leibovitz, 2003). Reducing tardiness and ensuring on time starts is clearly an obvious target for improvement.

Numerous factors contribute to late starts. Organized simply, these fall into three broad categories: (1) personnel factors, (2) patient factors, and (3) equipment factors (Leibovitz, 2003), (Kindscher & Rockford, 2009), (Shelver & Winston, 2001).

2.5.1.1 Personnel

The operating room nursing staff is responsible for setting up and laying out the operating room and surgical instruments prior to the first case of the day. After the room is set up, one nurse is responsible for seeing the patient in the preoperative area and bringing the patient to the room. Prior to the patient being brought to the room however, the patient must be seen and assessed by the anaesthesiologist and often is seen by the surgeon who must mark the appropriate site of surgery if the procedure is unilateral. In many centres, the sequence cannot be altered; for example, the surgeon must see the patient and mark the site of surgery before that patient can proceed into the operating room or in other situations, the surgeon must be physically in the building before anaesthesia can be induced. Delay in any of these personnel (nursing staff, anaesthesiologist, surgeon) will delay the start of surgery.

Operating room culture has been identified as the most responsible contributor for staff-related tardiness of first case starts. It has been shown (Shelver & Winston, 2001) that nursing staff members require support and encouragement to reduce tardniness and absenteeism. A surgeon who is repeatedly unavailable at the expected start time will effectively discourage the nursing staff from being ready at the specified time which over time and in turn will lead the surgeon to the conclusion that the room is never ready on time anyways so there is no need to be on time either (Truong, Tessler, Kleiman, & Bensimon, 1996).

In order to address personnel contributions, buy-in by the key stakeholders is required. Some centres have been successful in reducing staff related tardiness by obtaining support from surgeons, anaesthesiologists, and nursing staff (Shelver & Winston, 2001). Buy-in can be encouraged by both intrinsic and extrinsic incentive and reward. An intrinsic incentive may be the promise of recognition amongst peers while an example of an intrinsic reward is the improved interpersonnel dynamic and team work that is encouraged by a smooth start each day. The most direct extrinsic reward is monetary; by starting the surgery on time, the day is more likely to finish early or on time. For the surgeon and anaesthesiologist who are remunerated per procedure, the early finish allows them to move on to the next task and for the nursing staff in many centres, an early finish allows them to leave early. The simplest incentive is to pay nurses their full wage yet allow them to leave early if their slated surgeris are complete. The Shelver group at St. Luke's Hospital (2001) initiated a star system where the names of surgeons, anesthesiologists and nurses who were responsible for on-time starts were posted publically and stars were pinned next to the names each time an on-time start was achieved.

2.5.1.2 Patient

Patient-specific factors that may cause delay of the first case include late arrival of patients to the admitting department, unprocessed paperwork or incomplete laboratory investigations, and difficult intravenous access. While it is out of the control of the operating room if a patient arrives to the hospital tardy for their surgery, effort can be made to streamline or increase the capacity of the admission process in order to accommodate late arrivers in order to reduce any impact on the schedule. In terms of value-added, the admissions step of the process is not considered a value-added step but it is necessary. The minimal amount of time possible should be spent in admitting. Commonly surgical admitting is physically separated from main hospital admitting and multiple patients may be processed by multiple channels simulatenously. Laboratory investigations that are initiated in the preoperative holding area should be prioritized understanding that some tests require 45 minutes or more to complete. Preadmission and preassessment by the anaesthesiologist in a clinic on a day prior to surgery has significantly improved the admission process and decreased the incidence of incomplete investigation and substantial delays (Conway, Golberg, & Chung, 1992). The admission paperwork can be completed prior to surgery such that on the morning of surgery, the patient need only check in at the desk and not complete any extra paperwork.

2.5.1.3 Equipment

The vast majority of operating facilities utilize the case cart system of organizing surgical supplies (Golen, 2001), (OR Benchmarks, 2000). Sterilized equipment is packed in advance of the surgery onto rolling carts that are delivered from the central processing department of a hospital to the operating room. Commonly, case carts for each individual surgery planned for the day are delivered during the night so that the minimal time is spent during turnaround periods retrieving supplies. The case carts are assembled for each particular procedure based on a “pick-list” that the surgeon and operating room nurses have worked together to formulate. If multiple similar procedures are planned for the same day, facilities required an inventory of duplicate supplies such that there is adequate time for used equipment to return to central processing, be sterilized, and repackaged onto a subsequent case cart.

Case carts for the first case of the day are typically more reliable in terms of availability of equipment because there is ample time to prepare and package equipment. Delays can occur however if equipment from the first case cart must be used for an emergency procedure immediately prior to the scheduled first case and insufficient time is available to “flash” or quickly sterilize the item. As well, errors in preparation of the case carts may lead to delays.

2.5.2 Reducing turnaround times

The time between surgical cases is time when the operating room is not being utilized for its primary purpose, the performance of surgery. Unanticipated long or slow turnover of the operating room contributes to delayed completion of the elective surgical slate, cancellations of procedures, and cost over-runs. Inefficiencies in the turnover process become compounded when an operating room is responsible for a high number of procedures per day.

Unfortunately, there are differences in measures of turnover time because of the lack of a standardized definition for the beginning and end of the turnover period. The definition differs from surgeon to surgeon and from surgeon to anaesthesiologist to operating room manager. One definition of the turnover time is the time between surgical closure of one case until the surgical incision of the next, this can also be termed the non-operative time (Harders, Malangoni, Weight, & Sidhu, 2006). This definition is the one commonly adopted by surgeons as it discretely divides the available operating room time into two divisions: surgical and non-surgical time. The surgical time for each procedure is simply the time between first incision to closure, commonly called the "skin-to-skin" time. Surgical skin-to-skin time is the basis for schedule planning and operating room allocation; however, difficulty in estimating surgical time and its surrounding variability is a key factor leading to inefficiency is scheduling. This will be discussed in detail in the subsequent section.

Decreasing the turnover time between surgical cases is equivalent to the Lean technique of Single Minute Exchange of Dies (SMED) that was developed by Shigeo Shingo for automobile manufacturing in Japan (Slack, Chambers, & Johnston, 2007). The goal of SMED is to compress time for set-up activities in order to decrease batch size, and inventory size. When applied to a service industry such as a hospital operating room, where batch size and inventory are less relevant measures, SMED aims to reduce set-up time which will in turn increase capacity. Non-surgical time, or turnover time in the former definition, is comprised of a number of discrete steps. These steps can be considered in terms of internal and external set-up. By converting internal set-up to external set-up and streamlining both, significant reductions in turnover time may be achieved.

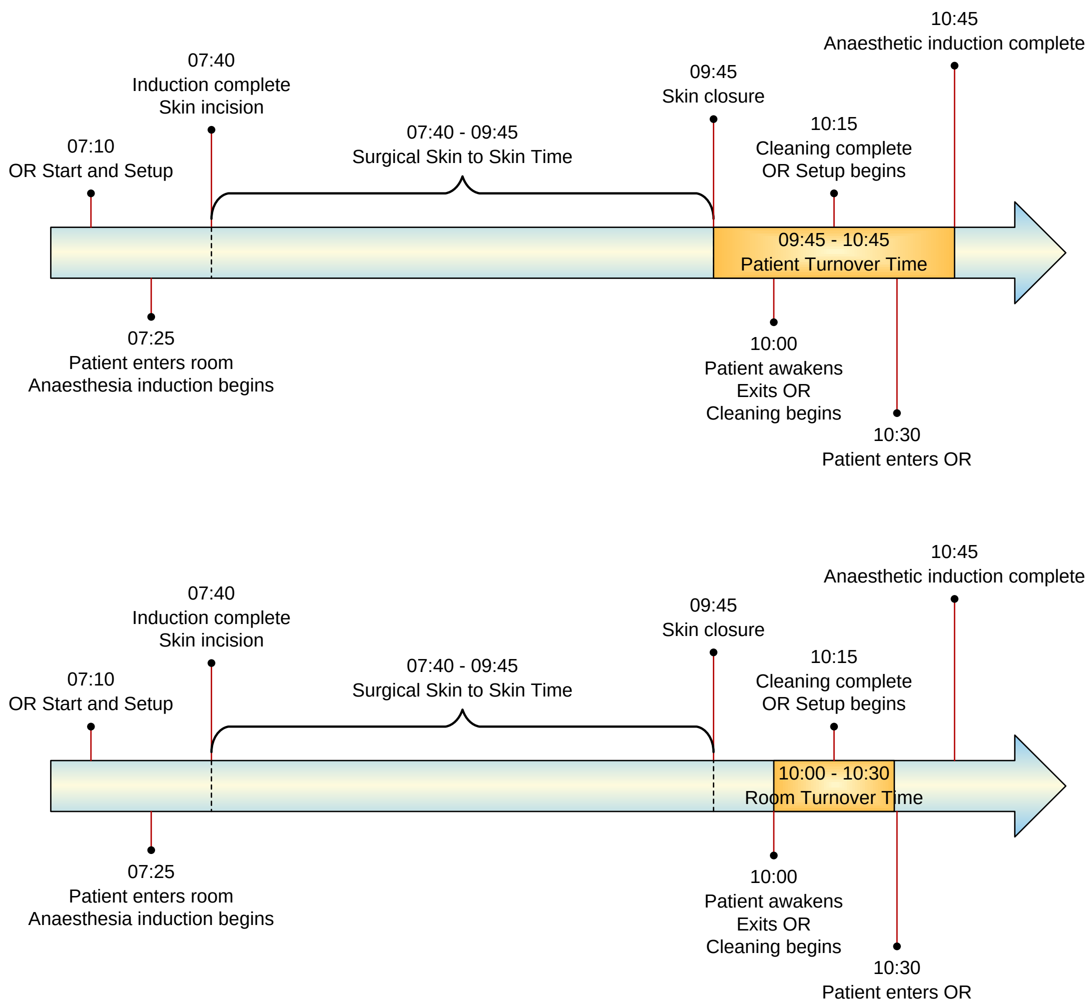
Following surgical closure, the surgical field is cleaned and an appropriate dressing is applied to the surgical wound and drains are stitched into place as necessary; if the patient has undergone a general anaesthetic for surgery, the patient will be awoken from anaesthesia and breathing tubes and the ventilator will be removed. The patient then needs to be transferred from the surgical table to a stretcher for transfer to the post-anaesthetic recovery room. This process, depending on the complexity of the life support medications and systems associated can be time consuming and dangerous and has in itself been described from an operations management perspective (Naik, 2006). Following the transfer of the patient from the operating room by the anaesthesiologist, the equipment used in the last surgery are collected and sent for cleaning in the sterilization and processing department of the hospital. Dedicated cleaning staff is then responsible for a complete sterilization of reusable equipment (monitors, tables, lights, etc) and replacement and restocking of supplies and medication inside the operating room. While the room is being processed, the anaesthesiologist will have ensured the safe transfer of care to the recovery room staff and will then proceed to assessment and preparation of the next patient. This may involve a history, a physical examination, and potentially premedication or placement of a regional anaesthetic technique. Simultaneously, the nursing staff is responsible for unpacking and preparing surgical equipment for the next case and then bringing the patient from the preoperative area to the operating room. A preoperative checklist is increasingly becoming a necessary component of this process and has recently become a mandate of the World Health Organization (World Health Organization, 2009).

Upon entering the operating room, the anaesthesiologist will decide if any specialized monitoring or pain control techniques will be employed during the case. These procedures, including the placement of invasive blood pressure monitors, epidural catheters, or electroencephalogram electrodes may be time consuming and may vary from case to case. These procedures are followed by the induction of anaesthesia and positioning of the patient; positioning is another time-consuming and potentially dangerous step especially when positioning includes flipping the asleep patient into a prone position or other positions necessitated by the surgical procedure. Finally, the patient is prepped and draped and the final surgical time-out is often performed at this time.

Internal set-up includes activities that can only be performed while the process is halted. These activities include the sterilization and cleaning of the operating table and floor, the unpacking of supplies for the second case and the transfer of the patient to the recovery room; clearly none of these activities can be performed in the midst of the surgical procedure. However, some activities can be and should be external, that is, they can be performed while the process is active. Shifting internal activities into external activities reduces the downtime of the system (Leslie, Hagood, Royer, Reece Jr, & Maloney, 2006). Such activities include restocking of anaesthetic drugs, or surgical supplies such as sutures, having the next patient assessed in advance by a nurse or another anaesthesiologist so that these activities do not need to occur in the turnover period.

In an early study, Mazzei determined that at his university hospital, skin incision was occurring between 21 and 49 minutes after the scheduled start time for the first surgery of the day (Mazzei, 1994). In this study, the author attempted to differentiate room turnover and patient turnover. Room turnover is the time from patient out to next patient in, and patient turnover is the time from end of surgery to the induction of the next patient; this is illustrated in Figure 2.4:

Figure 2.4: Surgical and turnover times



The room turnover time was almost uniformly 36 minutes while the patient turnover time was highly variable but generally one hour. It can be expected that cases with greater complexity or critically-ill patients may incur longer patient turnover time because this measure includes the anaesthesia component of surgical care and the inherent clinical variability that exists in medicine.

Attempts to decrease turnover time have been well documented in the anaesthesiology, surgical, and operations management literature (Mazzei, 1994), (Leibovitz, 2003). The Vancouver General Hospital initiative to standardize the turnaround time will be discussed further in the following chapter.

2.5.3 Accurate OR booking times

By far the greatest degree of variability in attempting to adhere to the scheduled elective slate of surgery arises from actual surgical times that differ from the expected surgical time. In both the block and open scheduling system, surgeons book procedures and attach a predicted surgical time to each procedure. The allocated block of time can be filled with whatever combination of surgeries that the individual surgeon, or a group of surgeons, chooses to combine provided that the surgical times do not exceed the available number of hours. A number of different means can be used to project surgical times.

One system is to have surgeons predict operative times individually for each individual patient. This will take into account both surgeon factors and patient factors which may shorten or lengthen the expected duration. However, it has been demonstrated that surgeons are generally optimistic about their projections and that schedule overruns and cancellations are frequent when surgeons attempt to book the list themselves (Pandit & Carey, 2006). In one study, Pandit demonstrated that when asked to book a surgical list, surgeons over ran the scheduled time 50% of the time, and 34% of lists suffered a cancellation. This author suggests that while surgeon's predictions should be used in scheduling, they should not be responsible for assembling the schedule.

One factor may be confusion about the terminology. It has already been shown that no standard definition of turnaround time is widely accepted. The same lack of standardization plagues the definition of surgical time as well. When predicting a surgical time, some surgeon may include the anaesthesia time (setup and induction) as well as the time for waking the patient up or applying dressings and bandages. Other surgeons will submit their prediction as simply skin-to-skin time and not include any of the ancillary activities.

An alternative system is to track each surgeon's historical surgical times for each procedure and book future schedules based on individual averages. Of course, this system cannot account for patient variability; often the complexity and therefore duration of surgery cannot be judged until after the opening incision is made. If operative times are simply based on historical averages, the surgeon is unable to tailor scheduling requests to reflect a predicted difficult case. By taking historical averages, provided the surgeon has performed a sufficient volume of cases, the expectation is that some cases will run long and some will run short but that on the average, extremes will be offset.

The broadest means by which to schedule duration is by using historical averages for each type of surgery across a pool of surgeons. For example, a laparoscopic gallbladder removal on average is a 45 minute procedure, a clipping of a single cerebral aneurysm averages 3.5 hours, or a multi-level spinal fixation with rods averages 6 hours. Again, individual patient factor variations are not accounted for and now surgeon differences are unaccounted. This system of predicting future surgical times has been found to be ineffective in scheduling cases to finish on time (Zhou, Dexter, Macario, & Lubarsky, 1999). Difficulty arises in predicting surgical duration in teaching institutions where resident surgeons are involved in performing surgical procedures. Junior surgical residents are granted gradually increasing responsibility in surgery based on proven performance but a learning curve is expected and throughput is slower when junior trainees are involved.

Clearly, no single system can account for all degrees of variation; a combination of historical times and surgeon customization may be the solution that provides the most accurate picture. This system would allow the surgeon, when booking a case, to select the procedure from a list of his commonly performed procedures; attached to each procedure would be a mean length of time calculated from historical data on cases that that particular surgeon had already performed. Prior to confirming the procedure and the associated length of time, the surgeon would have the option of increasing or decreasing the booking time based on their impression of factors specific to that patient. For example, a surgeon may average 3.0 hours for clipping a single cerebral aneurysm but believes, based on the location of a particular aneurysm as seen on CT scanning, that the dissection will be more challenging than usual and will therefore add 30 minutes to the booking time. By having the system generate a mean time first and then allowing customization afterwards, there is an anchoring effect that prevents overly optimistic or aggressive booking of the schedule and perhaps avoiding overruns.

2.5.4 Dedicated emergency rooms

In a tertiary referral hospital or large academic institution, the emergency operating room is available 24 hours a day to handle urgent procedures. The demand for emergency surgery is weighted heavily towards particular surgical services with general surgery and orthopaedic trauma surgery as the highest users of the emergency operating room. It has been recognized that while still emergent, many orthopaedic injuries that require repair such as hip fractures, or removal of infected orthopaedic hardware, can be medically stabilized for a period of time prior to requiring surgery. As a result, many centres have developed dedicated resources for orthopaedic trauma and other services (plastic surgery, ophthalmology, gynaecology) to deal with emergency cases that have arisen over the previous 24 hours but that can be handled during daytime hours. These resources are deemed "protected time" for each individual service and can be booked by surgeons on that particular service on an open booking system.

Evidence suggests that for those orthopaedic injuries that can wait until the next day, there may be numerous benefits in waiting (Lemos, et al., 2009). From a patient safety perspective, a group from MGH noted that for hip fracture surgery and femoral nailing, surgery performed at night had a significantly higher incidence of surgical complication than when performed in the daytime, a result that can be attributable to surgeon fatigue and fewer ancillary support services (Bhattacharyya, et al., 2006). In terms of utilization, the shift from overnight surgery to next day surgery resulted in a 6% reduction in over-utilization. The additional benefit is in recruitment and retaining of surgeons; a facility that offers the opportunity to perform emergency surgery during the day rather than at night is more attractive for surgeons.

2.5.5 Hybrid block and open booking

Hybrid scheduling is a technique that incorporates the benefits of both the block scheduling method and the open scheduling method. Effectively, a portion of the operating rooms are dedicated to specific surgical services and continue to be booked as block time while another portion of the operating rooms function with an open booking system. Surgical services that have regular demand for operating time would benefit from securing block time while others services that do not have a regular input of surgical demand would only schedule into the open room. Surgical services in the latter category include the burn surgery and trauma surgery services.

While the hybrid model appears to provide substantial advantage over a pure block or open system, it is not widely used (Kindscher & Rockford, 2009). One likely explanation for this is that each surgical service is self-interested in preserving their operating room time and a loss of block allocation results in competition for scarce room availability.

Operations management provides a framework for describing OR processes and suggests steps in the processes that require improvement. Two tools in particular that are highly relevant to the operating room are queuing theory and Lean management. Strategies that incorporate these tools may offer meaningful benefits in terms of efficiency and improved utilization of resources Moreover, operations management provides performance indicators on which hospitals can be reasonably compared to one another.

The ease of application of these techniques will likely vary from institution to institution and depend heavily on local factors including culture and complexity. An example of how some of these strategies were implemented at a large tertiary hospital may provide insight into potential challenges as well as the extent of potential improvements.

3: The Vancouver General Hospital Experience

Many of the operations management techniques discussed can be illustrated by the real world experience of the Vancouver General Hospital. Over the preceding 18 months, this large institution has begun to address several flaws in its operations. Some of the challenges faced by managers at this institution may serve to guide other centres in addressing similar process re-design.

In particular, the Vancouver Hospital (1) introduced a flex operating room, distinct from the existing emergency operating room that helped to reduce the queue of patients awaiting emergency surgery, (2) formalized a policy for starting the first case of the day on time, (3) standardized turnaround time, and (4) is attempting to implement a more precise scheduling system.

3.1 Vancouver Hospital Background

The Vancouver Acute Hospital is the major tertiary referral centre of the Coastal Health Authority of British Columbia. It is the largest of the academic teaching hospitals fully affiliated with the medical school of the University of British Columbia. The Vancouver Acute (VA) site, formally known as the Vancouver General Hospital (VGH) has 955 fully accredited inpatient beds and is the second largest hospital in Canada. The annual volume of patients seen in the emergency department is approximately 66,000 per year. The health authority is responsible for administering services to over one million people on the South coast of British Columbia and operates on a budget of approximately $2.8 billion.

Surgery is performed at two sites; there are 19 operating rooms in the main operating suite in the Jimmy Pattison (JP) pavilion of the Vancouver Acute (VA) site and another 8 operating rooms at the UBC Hospital site. The UBCH is a lower acuity hospital where elective day-care[1] surgery is performed on outpatients and elective surgery is performed on generally healthier inpatients. The UBCH does not perform emergency surgery and patients admitted through the emergency department requiring emergency surgery are transferred to the main VA site. In addition to the 19 main operating rooms, surgical staffs (anaesthesiologists, nurses, and surgeons) are distributed to other remote sites including the interventional radiology suite, interventional cardiology, and the lithotripsy suite. Review of operations for 2007 revealed that more than 16,000 scheduled surgical procedures were being performed during surgical primetime (0700h to 1930h) each year. Approximately 8,000 emergency procedures are also performed yearly at the VA site.

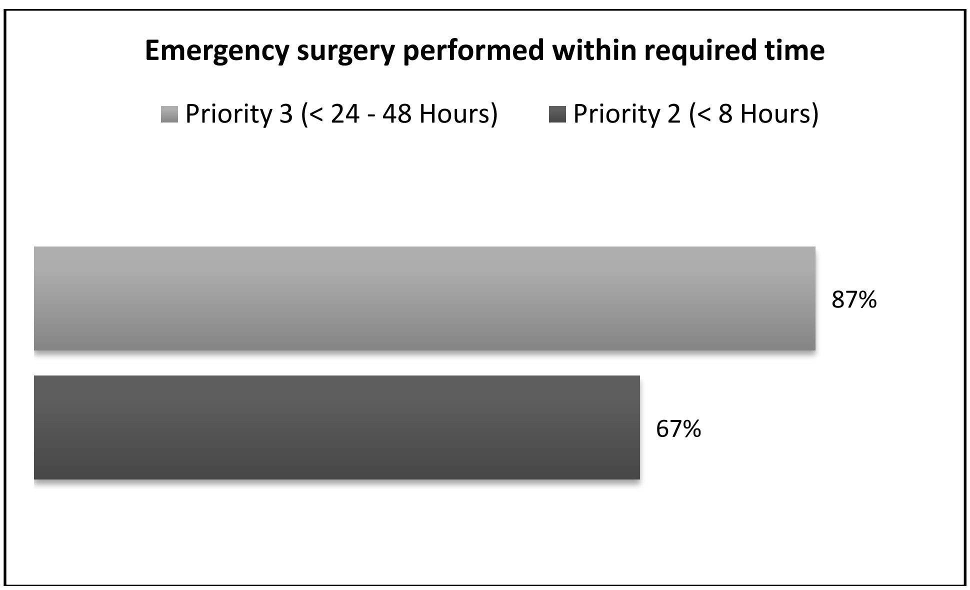
The scope of surgeries performed at the Vancouver General Hospital includes comprehensive general and trauma surgery, cardiac, thoracic, gynaecological, orthopaedic (trauma and reconstructive), ophthalmologic, urologic, ENT, plastic, vascular and neurosurgical procedures. No paediatric or obstetrical care is provided at the VGH; dedicated regional centres provide these services in the vicinity. The site is also the provincial centre for lung, liver and kidney transplantation.

[1] Day-care surgery refers to surgery performed on outpatients that do not require admission overnight to hospital. Typically patients are healthy and procedures are routine.

3.2 Vancouver Hospital description of the problem

At the start of 2008, it was recognized that the VA site was failing to meet critical benchmarks for the performance of emergency surgeries. In particular, two shortcomings were recognized. First, surgeries booked by surgeons as priority 3 (ideally done within 24 hours but must be done within 48 hours) were not being done in less than 48 hours and on occasion were not being completed for three or four days after the initial booking; similarly alarming, a large number of surgeries scheduled as less than 8 hours (priority 2) were not being performed within that time frame either. Figure 3.1 demonstrates the failure to complete emergency surgeries within the delineated time requirements.

Figure 3.1: Failure to meet emergency surgery time requirements (Third Quarter 2007)



At VGH, a designation of less than 8 hours (Priority 2) means that surgery is required within that time window in order to save life, limb or organ. Failure to achieve these targets suggests that patient safety was being jeopardized in the delay. Recalling the process map discussed earlier, the result of prolonged waits of three to four days for priority 3 and 4 cases was that patients were often occupying inpatient medical and surgical beds which in turn hindered the flow of patients from the emergency department (ED) to hospital wards, backing up the ED, and producing long waits for paramedics to drop off patients or forcing the diversion of ambulances to other hospitals.

The second major failure recognized was the high number of cancellations of elective surgeries on the day of scheduled surgery because of time overruns (diversions from the scheduled slate) and the need to accommodate emergency procedures. Cancellation of surgery has ramifications on all stakeholders (patients, physicians, the hospital facility and administration, and the government payer). Patients clearly suffer from cancellation; patients are instructed to fast from midnight preceding the morning of surgery and because the surgery that is most likely cancelled due to overrun of the schedule is the last surgery of the day, cancelled patients may have fasted up to 15 hours before being informed of the cancellation. Furthermore, VA being a tertiary referral hospital where complex surgeries are performed, many surgical patients travel from distant regions of British Columbia and when cancelled, are inconvenienced in having to reschedule their travel or place an extra cost onto the health authority if they are admitted to hospital until the surgery can take place. These physical inconveniences are minor in comparison to the emotional stress of anticipation resulting from the unexpectedly prolonged wait.

Surgeons are inconvenienced in having to clear up surgical time in their elective slates. In most cases, surgeons at VA are granted only one day of elective operating room time a week; clearly, reorganizing their elective slates poses a challenge. This challenge is reflected in lengthy wait-times for particular surgeons and surgeries. Finally, schedule overruns short of case cancellation force overtime on the surgical staff, most notably the operating room nurses. While overtime is paid at a premium in this health system, perpetual demands of overtime on nursing staff had led to diminished morale in the department and alarming attrition and the inability to attract new staff. The operating room management had already acknowledged that the state of overtime was not sustainable for the staff; further, overtime wages were a significant strain on the operating room departmental budget (personal communication, June 22, 2009) which threatened the closure of entire operating rooms.

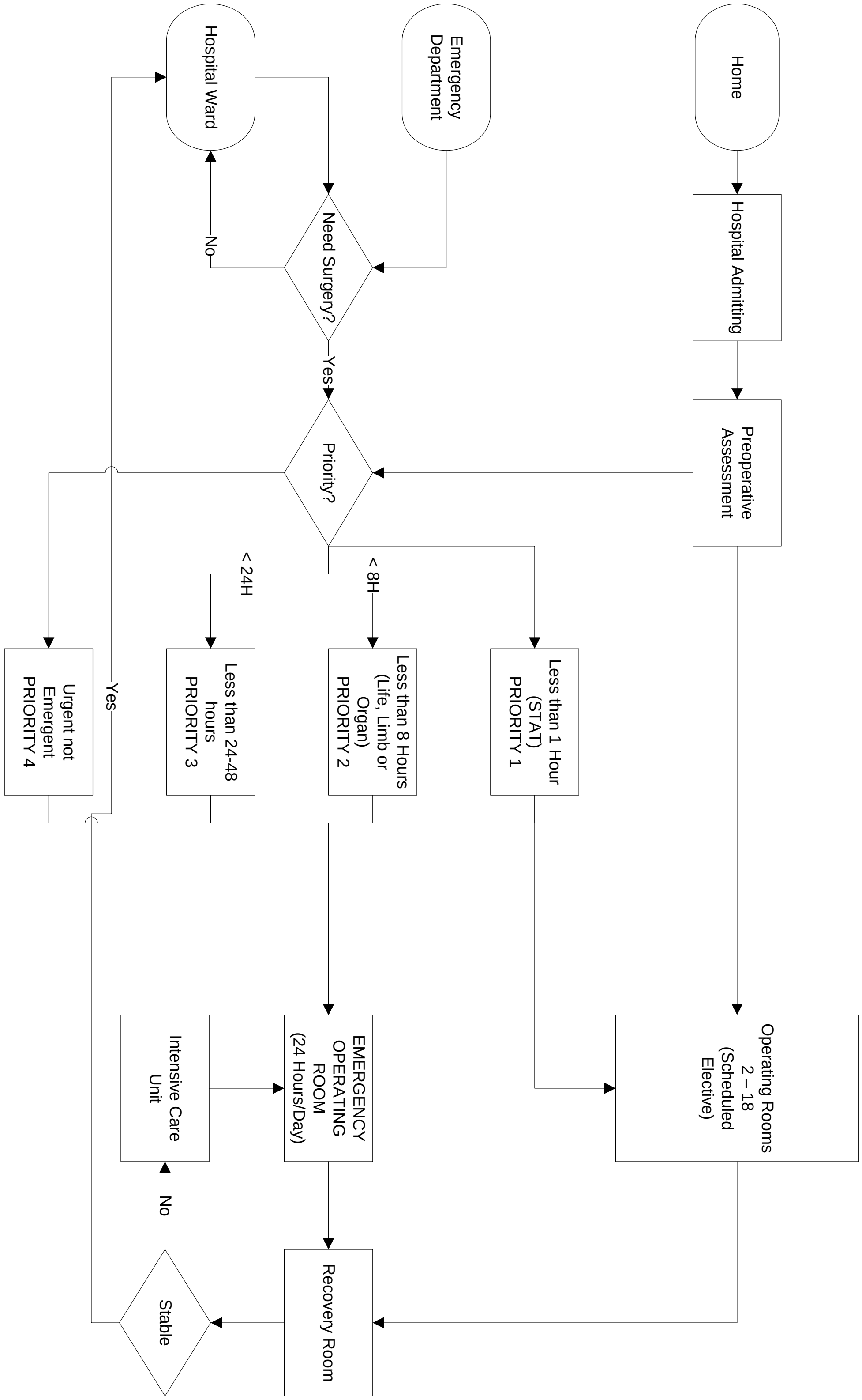
Operating room managers and medical directors recognized that in order to achieve appropriate benchmarks on timely completion, maximize utilization within a given budget, and maintain employee morale, various steps within the operating room process required revision. The areas specifically identified and targeted were those previously discussed: standardizing turnaround time, emphasizing and enforcing on-time starts, and initiating accurate surgical booking times. All of these were overseen by a newly created non-medical, managerial position of slate expediter. While not explicitly intentioned to conform to any framework, the steps taken in addressing system failures and attempts to improve flow can be well described in terms of Lean management and queuing theory. The resulting improvements in reduction of schedule overruns and increased achievement of emergency time targets (described in a later section) can be emulated in other centres through the implementation of queuing strategies that reduce variability and increase processing channels, and Lean strategies that minimize non-value added steps in the process.

3.2.1 The Sweeper Room reduces the queue, improves bottleneck

The operating suite[2] of the VGH, when represented as a service process has two distinct components. There is the elective scheduled component and the emergency component; while they can be considered separately and running in parallel, there is some overlap. Figure 3.2 depicts the “as-is” parallel elective and emergency processes.

[2] The operating suite refers to the collection of operating rooms including the ancillary activities (preoperative assessment, post anaesthetic recovery, etc)

Figure 3.2: VGH JPOR elective and emergency surgery as-is process



Emergency surgeries that cannot be performed overnight are carried forward to the next morning wherein more emergency cases are continually added. Emergency cases may infringe on the elective schedule during the day if a stat priority 1 case presents and the emergency room is already occupied for an extended period of time. The emergency case would bump into the elective slate and force delay or cancellation.

The queuing problem arises at night because of the limited resources available overnight and the unpredictable demand placed on those resources by emergency admissions. While up to three operating rooms can run simultaneously overnight, the demand placed on nursing and anaesthesiology resources is not sustainable from a workload or morale perspective; on most nights, one operating room is in operation after 2300h

Most surgical specialties are represented at the hospital with the exception of paediatric and obstetrical surgery. Obstetrical surgery (emergency Caesarean-section, and post-partum bleeding) was eliminated from the Vancouver Acute site twenty years ago because of the additional demand it placed on surgical services overnight; obstetrical services for much of the city of Vancouver were consolidated at the British Columbia Women`s Hospital following the closure and amalgamation of a number of smaller facilities. Even without emergency C-sections, the variability of demand is high; the most frequent emergency surgeries performed are repairs of orthopaedic injury including hip and long bone fractures and spine instability, laparotomy for general surgery, stabilization of multi-trauma, neurosurgery, and vascular surgery for limb salvage. While seasonal variation exists (summer months are associated with a greater incidence of gun violence, and winter months are associated with more fractures resulting from falls), there is too much uncertainty to alter resource levels overnight (having a second operating room open and regularly available) based on predictions of demand.

In discussing the necessary trade-offs that exist in a queuing system, we have seen that there are several ways to address concerns with the formation of long queues. These strategies include increasing the capacity of the system by increasing the number of processing channels, reducing variability or uncertainty (in length of processing, and in time between inputs), or by simply reducing the amount of input (the demand) by directing some patients to alternate facilities.

In not meeting targets for completing emergency surgery, it was apparent that the queue forming each night at VGH had grown too long. Traditionally, an open scheduling system has been used for emergency surgery with bookings from surgeons starting immediately following the elective slate in the afternoon. The table below (Table 3.1) illustrates a typical slate of emergency surgeries accumulated by the afternoon and awaiting completion.

Table 3.1: Overnight emergency bookings for the JPOR at the VGH (listed in order of booking)

Surgery	Urgency (Time)	Priority	Estimated Duration
IM Nailing of a hip fracture	< 24 – 48	3	1.5
Appendectomy for appendicitis	< 8	2	1.5
Craniotomy for intracranial haemorrhage	<1	1	2.0
Abdominal washout for sepsis	< 8	2	2.0
Laparotomy and bowel resection for ischemic bowel	< 8	2	2.5
Tracheostomy	< 24	3	1.5
Open reduction and internal fixation humeral fracture	< 24	3	2.0
Open reduction and internal fixation femoral mid-shaft fracture	< 24	3	2.0
Laparotomy for massive bleeding due to gunshot wound	< 1	1	2.5
		Total:	**17.5 Hrs**

Cases would be done on a first-in, first-out system based first on prioritization level. Patients requiring immediate surgery (priority 1) were operated on first; surgeons would on occasion ask to bump other surgeons in the queue if they felt that their patient required more urgent care. Urgent cases could be performed in multiple ORs in the afternoon because of the availability of anaesthesiologists and nurses until 1900h; after 1900h however, only one anaesthesiologist and one team of operating room nurses are typically on site.

After 1900h and throughout the night, inputs to the queue continue but the service is reduced to one channel. A second operating room can be opened up (by calling in an anaesthesiologist and a team of nurses) for stat surgery if the team present is already occupied and unavailable in a timely manner; but this comes at significant increased cost (Tucker, Barone, Cecere, Blabey, & Rha, 1999). Increasing capacity, by calling in the second team, is only available for true stat surgery such as trauma, craniotomy, or transplant and cannot be made available every night for less urgent emergency cases (less than 8, or less than 24-48). The result is that while stat surgery requirements are met, less urgent emergency cases may need to be delayed until the operating room becomes available; and prioritization of other more urgent surgeries in front of waiting procedures can produce an extensive backlog of cases designated as priority 2 or 3. The as-is process is mapped in Figure 3.2.

In order to deal with emergency cases that were accumulating overnight or over several nights, the VGH initiated the Sweeper (Urgent/Flex) Room. Conceptualized as a resource rather than a physical room, the Sweeper Room is essentially an anaesthesiologist and a team of operating room nurses available from 0700h to 1700h that is tasked with sweeping up cases that remained from the night before but that were not booked into the original emergency operating room that runs 24 hours. These were primarily priority 3 or 4 cases that in the block scheduling system were difficult to schedule and often were delayed several days. In relating the effect of the Sweeper Room to the framework of a queuing system, the Sweeper Room produces three results:

- Decreases variability
- Increases the number of queues
- Increases the processing channels

In order to fund the Sweeper Room in the trial period (third and fourth quarters of 2008), elective surgical time was clawed back from all surgical services. This resulted in an estimated reduction of 350 elective scheduled cases per year, which represents a 2% reduction in elective procedures. Additional funding has subsequently been obtained to restore the entire cancer surgery portion of the previously reduced services. The majority of surgeons were supportive of the new initiative, despite a small reduction in elective time, because of the opportunity to improve emergency surgery performance. Furthermore, the reduction of elective cases scheduled did not exceed the number of forced cancellations that were previously occurring because of schedule overruns.

3.2.1.1 Decreased variability

In making available additional dedicated operating room time in the daytime, lower priority emergency cases that were previously booked into the open emergency room could be electively postponed until the next day. A large number of surgeries and in fact, entire types of surgery could be shifted to the daytime without impeding the availability of the high priority emergency OR; this includes the majority of orthopaedic trauma surgeries, plastic surgery procedures and stable gynaecological procedures. The reduction in the variety of surgeries-in both emergency channels-is beneficial in minimizing the queue because estimates of duration could be more precise and a smaller range of equipment or resources would be needed and could be close by, reducing turnaround time; moreover, nursing staff would face fewer types of surgeries, avoiding some procedures that required specialized experience or expertise. This allows for more rapid set-ups and smoother performance.

3.2.1.2 Increased number of queues

In effect, diverting priority 3, 4 and some priority 2 cases until 0700h produced two distinct queues where previously there was only one. With the Sweeper Room (also referred to as the Flex Room), one queue would form for the low priority urgent cases, and a second queue would be formed for priority 1 and 2 cases.

Prior to the Sweeper Room, the VGH had already implemented a number of protected operating rooms. Protected ORs or protected time refers to dedicated operating room time and facilities granted to a surgical service to book openly. The combination of open scheduling and the more common block scheduling is the hybrid system previously discussed. Many surgeries previously performed at night, including hip fractures, wrist fractures and other orthopaedic procedures, had already been shifted to the orthopaedic trauma protected time in the daytime. This model of care has been demonstrated to improve outcomes because of a reduction in surgical complications arising from surgeon fatigue and other factors (Lemos, et al., 2009).

The Sweeper Room works in a similar way to divert cases from the emergency slate to an openly booked room but rather than being protected for a single surgical service, all services would have access to these resources.

3.2.1.3 Increase number of processing channels

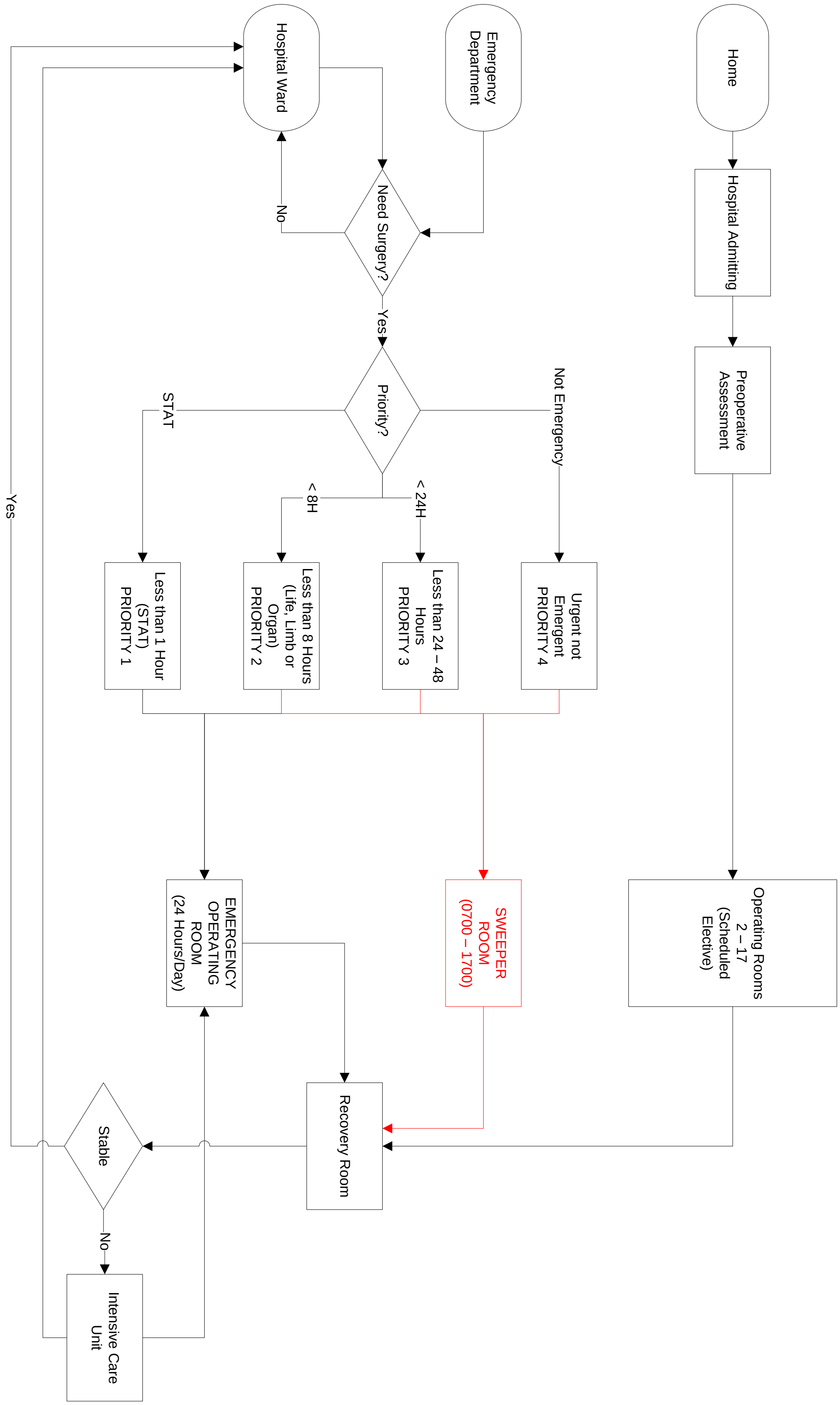
The flexibility of the Sweeper Room allows for increased processing of all emergency surgeries because it increases the capacity of the system. It acts as increased surge capacity for the unpredictable demand of emergencies. During the daytime hours (0700h to 1900h) low urgency surgeries can be performed in the flex space preventing long delays for these procedures while the emergency room can be utilized for higher priority or stat cases. When needed though, it also provides increased capacity for high priority cases should multiple priority 1 or 2 cases present and need to be performed simultaneously.

This system is different than simply increasing emergency room capacity by doubling the number of emergency operating rooms. While the mandate of the Sweeper

Room is to ensure that all true emergencies (priority 1 and 2) are performed in a timely way, it is not commonly utilized to perform these cases itself, although it can. Instead, it acts to relieve the congestion on the bottleneck room by diverting cases that would normally be done in that room (priority 3 and 4) and processes them through a secondary channel. Conversely, if two operating rooms were dedicated to performing urgent priority 1 and 2 procedures, low priority procedures would continue to be heavily delayed.

Depicted graphically, the process map for the new process appears in Figure 3.3 below. Perhaps the most beneficial secondary effect of the Sweeper Room was the increased ability to address priority 4 urgent elective cases. These cases were often complicated surgical problems that required the coordination of multiple surgical services such as the combined efforts of vascular, general, orthopaedic and plastic at the same time. These cases were traditionally difficult to schedule in the block or protected time system because no individual service was willing to allocate its dedicated resources to a combined procedure. The result was often lengthy delays and prolonged hospital admissions occupying resources while awaiting an opening in the emergency room (personal communication, June 30, 2009). The Sweeper Room was clearly ideal to address this demand because of its open scheduling system and availability to all surgical services.

Figure 3.3: VGH Operating Room process with the addition of the Sweeper Room.



3.2.2 Implementing on-time starts and standard turnarounds

Several factors leading to the need to cancel elective surgeries were identified. Two were determined to be easily targeted for correction, on-time starts and reducing the time required to turn-around operating rooms between cases. These two steps in the service process are clearly non-value added and contribute to waiting times and decrease throughput through cancellation and reduction of utilization.

3.2.2.1 First Case Start Time

The Vancouver Acute hospital implemented the First Case Start Time (FCST) policy in 2008. By observation, several of the 19 regular operating rooms were routinely starting the morning case five or more minutes after the scheduled start time of 0725h. According to the slate expediter charged with monitoring late starts, the most regularly cited surgical subspecialties failing to start on-time were the general surgery and gynaecology operating rooms (personal communication, June 30, 2009).

The development of a culture amongst the personnel was identified as the key contributor to the recurrent late starts. Set-up of the surgical instruments and the layout of the room is intended to begin at 0710h when the team of operating nurses arrive. The necessary equipment for the elective schedule of surgeries would have already been delivered the night before from the sterile processing department and are packaged inside the operating suite core, ready to be individually unpacked for each scheduled surgery that day. The nursing staff is responsible for unpacking the equipment, setting up the room appropriately for the given procedure and then retrieving the patient from the preoperative holding area such that the patient enters the room at 0725h for induction of anaesthesia.

Observations by the nursing staff and physicians suggested where delays were arising. These included, short staffing as a result of “sick-calls”, prolonged in-services or morning huddles that extended into the 0710h set-up time, being pre-occupied setting up the OR and neglecting to check-in and pick up the patient on time, and the patient not being ready because the surgeon had not marked the surgical site; most clearly however, surgical teams that were guilty of starting late on a regular basis, simply worked slower with a sense of lower urgency. A culture of tardiness and a cycle of complacency had been established.

In order to address these deep cultural behaviours, the FCST policy was implemented. Motivation and buy-in can be difficult when implementing significant change, especially when demanding an increase in productivity. In the case of the VA experience, motivation would need to come from both intrinsic reward as well as extrinsic penalty. Nursing staff were involved in the process of developing the FCST as it has been recognized that input from front-line workers is key to buy-in and success of the project (Noon, Schiffer, & Crane, 2009). Nurses were able to voice their complaints about delays stemming from the preoperative holding area, and physicians complained about nurses waiting too long before retrieving the patient. The final motivation for improvement was the strict implementation of a case cancellation policy. The policy states that if the final scheduled case of the day in an individual operating room is expected to start one hour or more after the scheduled time, then it will be cancelled. This became a strong motivator for surgeons to start on-time and to encourage the nursing staff to adhere to start times and rapid turnover.

Changes were made in order to better align practices in the preoperative area with the goal of on-time starts. These changes included allowing anaesthesiologists to mark the surgical site in advance or in lieu of the surgeon if a regional anaesthesia technique was to be used, having more patients seen by an anaesthesiologist in the preadmission clinic in order to cut down time for assessment on the morning of surgery, and having operating room nurses complete a portion of the pre-surgical checklist in the preoperative holding area.

In order to motivate individual stakeholders further, a record of the arrival time of nurses, surgeons and anaesthesiologists was kept by the slate expediter. Regular tardiness would need to be explained in person to the slate expediter and formal letters were sent to surgeons whose operating rooms regularly started late.

3.2.2.2 Turnaround Time

The time required to clean and re-prepare the operating room between the exit of a patient until the entrance of the next patient has been termed the turnover time or the turnaround time (TAT) at the Vancouver Acute hospital. Several activities must occur in order for a room to turnaround between cases and these activities involve numerous groups of individuals. At the VGH, these activities include:

1) the anaesthesiologist must take the postoperative patient to the recovery room and ensure stability then transfer care to the RN; the anaesthesiologist must then proceed to the see the next patient in the preoperative area, and then return to the OR where he must prepare the anaesthetic for the upcoming case;

2) the OR nurses must count every item of surgical equipment used in the surgery, safely handle and repack surgical supplies from the first case so that they can be processed in sterile processing, and complete documentation and then following cleaning of the OR, nurses must unpack equipment and lay-out the room for the next procedure, and then check-in the next patient;

3) cleaning staff enter the room immediately after the patient has exited the room and perform a thorough cleaning and sanitization of the floors, machinery and equipment, and surgical table; perioperative assistants (PAs) replace supplies consumed in the earlier cases and assemble and position equipment with the guidance of the nurses and surgeons;

4) the surgeon and surgical team must dictate the events of the previous surgery, see the next patient. While not directly affecting turnaround, surgeons will often leave the operating suite to address other pressing issues in the hospital including consultations and assessments on the ward or in the emergency room. As well, surgeons and their teams often take the opportunity during turnaround periods to eat. The immediate availability of the surgical team does impact the ability of the OR to proceed to the next case if the surgeon has not seen the patient and marked the surgical site.

Prior to the TAT policy, no formal guidelines for these tasks were delineated and this turnaround period varied in length. The target of the TAT policy (appendix) was to have all turnarounds completed within 30 minutes. A reduction in TAT, equivalent to the concept of minimizing set-up time in the Lean/JIT framework reduces overall throughput time and allows for more cases to be performed or prevents cancellations. When applied to the 24 hour a day emergency operating room, a mean reduction of turnaround time (or set-up time, in the Lean philosophy) of 10 minutes when applied over the course of 60 cases (an average weekly emergency case load) translates into almost ten hours of additional operating room availability.

In order to achieve this goal, the policy outlined several steps that should be taken in order to convert many of the set-up activities from internal to external set-up (Slack, Chambers, & Johnston, 2007) or to improve efficiency through parallel processing (Harders, Malangoni, Weight, & Sidhu, 2006), (Friedman, Sokal, Chang, & Berger, 2006); a number of them are listed here:

- At VGH, resident anaesthesiologists or surgeons can be dispatched to the preoperative holding area to assess upcoming patients prior to the completion of the first case in order to save time during turnaround periods.
- A policy was implemented that allowed the anaesthesiologist to mark the site of surgery thus relieving the surgeon of this responsibility in case they were still occupied in the previous case.
- Communication between the operating room and the preoperative care area made it possible to have patients ready 30 minutes in advance of the next case even if the slate was running ahead of schedule.
- The anaesthesiologist would communicate to the recovery room prior to the completion of surgery that special equipment such as a ventilator would be required and set-up so as to prevent delays at the end of the case.

- Anaesthesiologists and surgeons were made to understand that the turnaround period was not a “break-period” and that all staff should be immediately available within the turnaround time.
- Cleaning staff and PAs were organized into teams with each team responsible for only 6 operating rooms, all located in the same area of the operating suite in order to minimize unnecessary movement of staff
- Visual displays outside of rooms would alert cleaning staff of surgeries that were nearing completion so that teams of cleaners would be immediately ready to commence cleaning.

3.2.3 Predictive Scheduling: the X-Time initiative

Of the steps in the value chain, the step with the greatest degree of variation in length is the actual surgical time. It is well understood that different surgeries will require different lengths of time and so allotment is made in the elective schedule for this. For example a complicated resection for pancreatic cancer requires upwards of 4 hours while a laparoscopic cholecystectomy for gallstones requires one hour. There are several ways in which block time can be allocated in order to achieve different levels of high utilization including longest to shortest and others (Dexter & Epstein, 2003), (Dexter, 2006) (Dexter & Traub, 2002). At the VA hospital, surgical specialties schedule their own block time such that the aggregate lengths of the surgeries planned and the necessary turnaround times in between procedures fits within the allocation of the block. The problem that became clear in discussion with surgeons was that there was no standardization of the estimates of surgical times and moreover, and perhaps more fundamentally, there was no standard definition of what activities surgeons were factoring into their estimates of surgical time. Some surgeons were estimating their skin-to-skin time while just as many were estimating the patient-in to patient-out time. Clearly, as we have already discussed, differences in definition will significantly alter scheduling.

Several attempts had previously been made to control this variation including using the surgical times generated from the Operating Room Management Information System (ORMIS). This system fails to recognize significant surgeon to surgeon differences as well as patient complexities and variation.

The VGH operating room has recently attempted to initiate a new program aimed at solving this scheduling problem; the program which is called Predictive Scheduling has three key steps. The first step of this ongoing initiative has been to establish clear definitions. The second step has been to measure and account for the non-surgical variation and attempt to standardize it; and the final challenge has been to obtain buy-in from the surgeons and anaesthesiologists.

3.2.3.1 Defining X-Time

Historically, operating room booking has been done by patient-in to patient-out times; this time includes surgical time as well as preparatory time by the anaesthesiologist, and occasionally other allied services like neurology, radiology, or perfusion services. In some complicated surgeries or when in high-risk patients where epidural catheters or invasive monitors need to be placed in the patient, or the patient needs to be positioned in the prone position, this preparatory time can be substantial (30 to 90 minutes or rarely more). In many instances surgeons are unaware of the extent or magnitude of this preparatory work and therefore cannot accurately predict its duration. It is much more practical to have the surgeon project only his component of the booking time, that is, the skin-to-skin time. The remainder of the time has been uniquely termed at the VA as the X-Time.

- Patient In to Patient Out time = A
- Skin-to-Skin time = B
- A – B = **X Time**

The X-Time is comprised mostly of anaesthesia related activities prior to skin incision including establishing intravenous access, placement of monitors and invasive access monitors, epidural catheters or regional anesthetic blocks, and positioning the patient; and, activities after skin closure including application of dressings and casts, awakening the patient from anaesthesia, and transfer of patients from the surgical table to the stretcher or bed.

3.2.3.2 Variations in X-Time

Times of patient entry to the OR, skin incision time, closure time and patient exit were recorded for over ten thousand elective procedures performed in a six month period in 2008. Procedures were classified by procedure code such that approximately 800 unique procedures were identified. For each unique procedure code, a mean X-Time and standard deviation were calculated; and groups of similar X-Times were placed together in groups called "buckets". Targets for each bucket were also assigned; these were termed "Olympic bucket times." Appendix D shows a list of representative calculations.

An observation was made that for each unique type of procedure, X-Times were similar with only a small standard deviation in comparison to standard deviations in actual historic skin-to-skin time. This suggests that X-Times have limited variability for a given procedure and that an expected X-Time can be assigned to each surgery which would be predictive of the actual X-Time. For example, the mean X-Time for five vessel coronary artery bypass grafts (CABGX5) was 79.3 minutes with a standard deviation of 14.1 minutes. In scheduling future procedures, the X-Time can be factored into the booking to more accurately predict the duration of the surgery. For this CABGX5, an X-Time of 75 minutes would be assigned and the surgical time would be predicted separately. Grouping similar X-Times together into buckets of 30 minutes, 45 minutes, 60 minutes, or 75 minutes allows for ease of booking and as a tool for benchmarking performance.

In effect, defining the X-Time and assigning each procedure to a bucket separates the more consistent portion of the booking time from the more highly variable surgical time.

3.2.3.3 Achieving buy-in by surgeons and anaesthesiologists

In separating the X-Time from the surgical skin-to-skin time, the predictive scheduling initiative removes one source of uncertainty from the equation. Surgeons have in the past underestimated the time required by their anaesthesiology colleagues to set-up and induce anaesthesia in a complicated patient. With this scheduling system, surgeons will no longer be asked to factor the anaesthesia time or the X-Time into the surgical booking, only the skin-to-skin surgical time. Cases will instead be booked as:

- Skin-to-Skin Time + X-Time + Turnaround Time = Booking Time

The X-Times have been shown to have only limited variability regardless of the anaesthesiologist involved in the case and so can therefore be applied strictly by historical data. The turnaround time (TAT) is the 30 minutes required to change the room over as previously described.

Surgeons would be asked to book their cases as previously discussed using a combination of historic performance times with alteration based on individual patient complexities. The X-Time for that particular surgery would then be applied with the option for the anaesthesiologist seeing the patient in the preadmission clinic (one week in advance of the surgery) to augment the booking in 15 minute increments based on specific individualized concerns.

Despite initial support, the predictive scheduling system has been difficult to implement because of lack of commitment from the surgeons and their offices. While having been informed of the need to book only surgery skin-to-skin time, many offices have failed to do so and have continued to attempt to book total OR time. In the absence of surgeon input, the booking system has reverted back to using ORMIS historical times for surgical booking. Moreover, the leader of the predictive scheduling program, the Medical Director of Perioperative Services has completed his tenure and the program has been delayed while a replacement is sought.

In mediating this scale of change, a champion is required. This may be best achieved through a trial program that includes a limited number of surgeons and surgical specialties. Success of a trial would be measured in a reduction of the number and length of overruns of the schedule and an absolute reduction in the number of case cancellations in comparison to the same period using the historic system. Alternatively, the schedule could be shadow booked with the predictive scheduling approach and a comparison could be made to see which system predicted more precisely and accurately the duration of the cases. Convincing a small number of surgeons, especially departmental heads, to conform to the new system may help in further buy-in from the remainder of the surgeons.

3.3 Measuring outcomes

In the last 18 months, the Vancouver Acute hospital has attempted to make significant changes in order to lean its operations and to improve performance. Interpreting the outcome of these initiatives is uniquely challenging for the hospital because of the lack of appropriate benchmarks.

The most commonly used metric for operating room efficiency is utilization. It has been suggested (Accenture, 2007) that for a large tertiary or trauma hospital, raw utilization is targeted at between 70 and 75% in order to compromise between efficient use of resources and maintaining the capacity to deal with surges in demand; these benchmarks, however, are derived from data collected in the private hospital system in the United States. In comparison, the VA utilization data suggests that raw operating room utilization was already 81% prior to implementation of any changes. Factoring in turnaround times, utilization rises to 95%, leaving an idle time of only 5% during daytime hours. This is significantly higher than utilizations seen in comparable institutions in the United States (Tyler, Pasquariello, & Chen, 2003).

The reason why utilization is substantially higher at the VA and likely similar centres in Canada in comparison to American counterparts is likely the result of the restraints of the public system in which the Vancouver General Hospital operates. Resources are more highly constrained in the public system such that the decision of how many operating rooms to open and staff is determined chiefly by the size of the budget; in contrast, in a private institution the driver for capacity is primarily demand. In the private system, where many surgical procedures are a source of revenue generation, utilization can be kept artificially low in order to maintain availability for add-on cases.

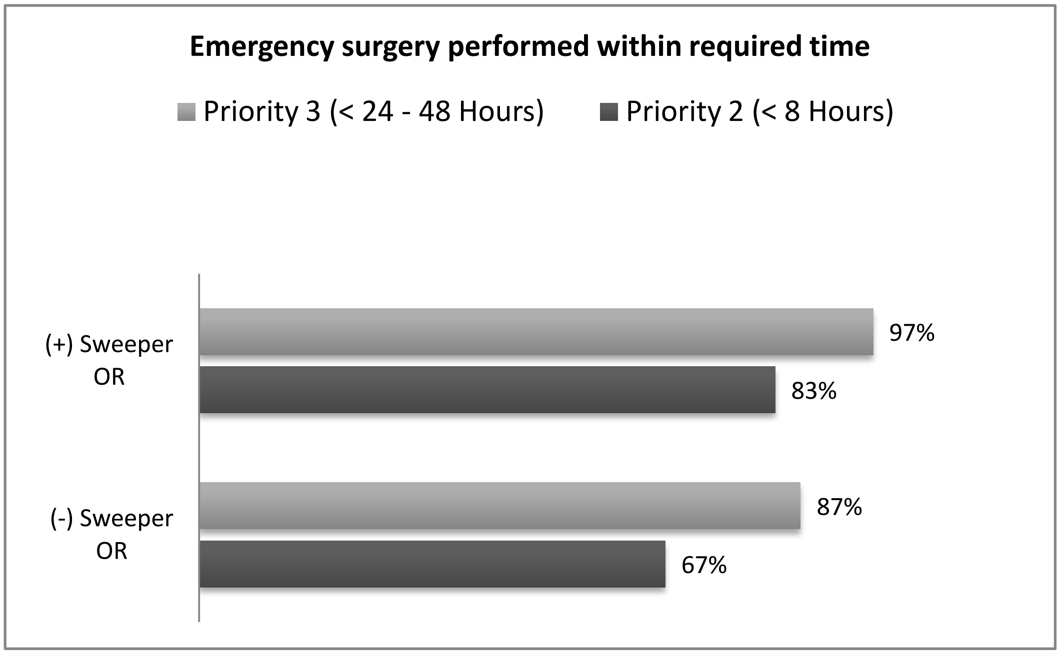
That point being taken, such a high utilization may be the reason that a large number of cancellations were occurring and that urgent and lower priority emergency surgeries were not being performed on-time. By stretching hospital resources to optimize utilization, there is a trade-off with surge capacity because variables such as demand and duration of emergency surgeries are difficult to predict. Comparison with similar Canadian institutions may suggest whether VGH is over or underutilizing its operating facilities. If for example, a similar centre is able to sustain high utilization without a similar frequency of cancellation then causes for this difference need to be sought. Utilization benchmarking with American counterparts is more difficult because of the different treatment of surgical demand. Utilization tends to be lower in American hospitals (in the 70% range) because when utilization approaches 80%, additional operating room facilities are opened (OR Benchmarks, 2000). Interestingly, frequency of case cancellation is not measured as a key performance indicator in most benchmarking studies in the United States likely because of the relative infrequency of cancellation.

3.3.1 Morbidity and Mortality

In order to determine if improvements have actually resulted from performing more emergency surgeries on-time, the most important measure would be a decrease in morbidity or mortality. These are the standard measurements utilized in medicine to determine efficacy and quality of treatment. Unfortunately, data to determine mortality would require a comprehensive review of all charts for emergency procedures and would require follow-up even post hospital discharge; this review has not been performed. Morbidity is even more difficult to measure; morbidity post-operatively is not uncommon due to the substantial co-morbidity that exists in the emergency surgery population and numerous other confounders will influence outcomes, including heart disease, respiratory illness, traumatic injury and many more. Moreover, a patient that is admitted to hospital post-operatively is monitored and medically managed for potential complications and many of the potential morbidities that would have arisen from long delays are avoided by expert management. However, this additional medical management may be costly for the medical system and minimizing the need for it is still essential.

In the absence of reliable morbidity and mortality data from first-hand experience at the Vancouver Acute hospital, the appropriate surrogate performance goals are the timelines established for each priority level of emergency surgery. Ideally one hundred percent of priority 1 cases would be performed within one hour, and one hundred percent of priority 3 cases would be performed within the 48 hour window. While stat surgery (within one hour) was always being performed within the time required, we have already seen that the Vancouver Acute hospital was failing to meet the required guidelines for emergency surgery in over 30% of cases booked as priority 2 and approximately 15% of the time in priority 3. Following implementation of the Sweeper/Flex OR, there has been a substantial increase in the proportion of emergency cases started within the specified guidelines.

Figure 3.4: Improvement in timely surgery (Third Quarter 2008 over 2007)



An additional benefit was that priority 4 cases which were previously difficult to schedule because of the high utilization of the elective slate and a back-log of emergency cases found a place in the Sweeper Room; this resulted in a 400% increase in the number of Priority 4 cases that were performed in the six month trial period ending in January of 2009.

3.3.2 Staff satisfaction

One of the primary goals in reducing schedule overruns was to address growing dissatisfaction amongst operating room nurses and poor morale leading to attrition. Secondarily, surgeons are continually seeking more operating room time in order to address ever-increasing waitlist pressure. Satisfaction with changes must be considered an essential component to long-lasting improvement.

Satisfaction amongst the stakeholders has yet to be measured following the initiatives described above. Moving forward, feedback should be sought from nurses, doctors, and managers as to whether their goals are being adequately addressed. Certain absolute measures may be useful as surrogates for satisfaction. As reducing the heavy burden of overtime work on the nursing staff is a goal, some easily measured performance indicators include the number of nurses joining or leaving the operating room (attrition and retention rates) and the number or frequency of "sick-calls."

A formal measurement of satisfaction with working conditions is part of the annual review of performance for individual nurses and channels for voicing concerns are widely available.

4: Conclusion

Faced with constrained resources, administrators and medical practitioners need to find ways to simply do more with less. In the Canadian public healthcare system, this has long been a recognized challenge and evidence suggests that utilization of resources such as operating rooms (as well as hospital beds) is high when benchmarked with comparable American institutions. Nevertheless, there may be shortcomings and even failures to meet clinical and evidence-based guidelines for satisfactory service in terms of best practices.

In order to narrow the gap in service quality including meeting guidelines for emergency surgery and limiting queues for elective surgery by preventing cancellations, there may be opportunities found within the processes and operations that already exist in the system. Lean management techniques and theories originating from the Toyota Production System suggest that throughput and efficiency gains can be made through the reduction in waste. In the context of healthcare processes, steps in the processes that do not add value to the patient experience represent waste. Waiting time is the most common illustration of a non-value added step that occurs frequently even in the course of a single surgical visit. Other forms of waste include over-processing and excessive movement of the practitioners performing the tasks. All of these areas of waste offer an opportunity for improvement.

Waiting times, in addition to contributing to waste and not adding value, may result in poorer outcomes. The example of the Vancouver General Hospital has been given to demonstrate that process re-design can be implemented through a qualitative queuing theory framework in order to enhance throughput of the system. Through diverting resources from the elective operating rooms and implementing a flexible operating room resource that decreased demand on the processing ability of the 24 hour emergency operating room, significant improvements were made to achieving on-time emergency surgery. This is expected to result in improved surgical outcomes and may therefore reduce subsequent costs incurred during hospitalization and reduce long-term sequelae.

Some general lessons can be taken from the application of the operations management tools discussed herein (Lean management and queuing theory frameworks).

- There is a necessary trade-off between capacity, inventory, and variability
- Identify and reduce areas of waste including excessive queuing
- Identify how internalized setup activities can be externalized in order to minimize non-value added steps
- Appropriate key performance indicators are required for benchmarking
- Involve stakeholders in making changes and foster a culture of improvement and recognition of performance improvement

Healthcare operations have been resistant to outside influence because of the belief that systems involving professional practitioners and a high degree of variability in patient factors could not be adequately described in traditional production or service terms. It is becoming increasingly clear, however, that healthcare services must be optimized to meet similar goals as industrial processes including quality, described in terms of timely delivery of services resulting in best practice and optimal outcomes, maximal throughput despite variability in processing times, and maximal utilization which can be achieved through reductions in non-value added activities and accurate scheduling of tasks. As costs associated with the delivery of care increase because of demand, budgets will continue to limit our ability to deliver this care. It is the responsibility of those actually delivering this care to find ways to optimize our day-to-day operations. Management tools such as Lean and queuing theory may be the way to start to address and communicate these opportunities.

Appendices

Appendix A

Excerpts taken from the Vancouver Acute Operating Room Policy:

I. General Surgery

Priority 1.
Massive bleeding
Bowel obstruction with mesenteric ischemia
Trauma (Open / Closed) - bleeding and / or unstable
Peritonitis with septic shock
Necrotizing fasciitis

Priority 2.
Perforated viscous
Bowel obstruction
Trauma (Open / Closed) – less urgent bleeding/+ DPL
Acute appendicitis
Incarcerated hernia
Ischio rectal abscess
Peri-diverticular abscess
Cholecystitis with septicemia
Severe systemic sepsis

Priority 3.
Bowel obstruction of less acute nature
Acute cholecystitis
Uncomplicated abscess

III Neurosurgery

Priority 1.
Acutely deteriorating LOC due to:
- intracranial mass e.g. Hematoma, tumour, abscess
- acute hydrocephalus

Insertion of ICP monitor
Repair of compound brain wounds
Post op complications: bleeding, cranial decompression
Acute cord compression
Acute vascular obstruction
Intracranial aneurysm repair after sub-arachnoid bleed

Priority 2.
Acute compound skull fracture
Vascular procedures to prevent incipient stroke

Priority 3.
Closed depressed skull fracture
Repair CSF leak
Brain biopsy for infection
Reoperation for post-op infection
Stabilization of spine
Severe tic doloureaux
Acute disc syndrome

Priority 4. Removal of percutaneous hardware

IV. Orthopedics

Priority 1. Orthopedic injuries with ischemia, massive bleeding or neurologic compromise
Necrotizing fasciitis/gas gangrene
Acute compartment syndromes
Dislocated hip (traumatic)
Massive post-op hemorrhage
Dislocated hip or knee arthroplasty with neurologic or vascular compromise
Periprosthetic fracture with neurologic or vascular compromise

Priority 2. Open fractures/fracture –dislocations/joint injuries
Multiple long bone fractures
Multiple fractures in polytrauma patients
Acute septic joints
Dislocated hip arthroplasty
Dislocated knee arthroplasty
Acutely infected hip and knee arthroplasty
Wound lesions, hematomas, infections after arthroplasty and/or oncology
Femoral shaft fracture
Displaced hip/talus fracture in young patients
Irreducible joint dislocations/fracture –dislocations

Priority 3. Closed fractures
Amputations for tumours with severe pain, severe sepsis, or gangrene
Acute ligamentous and tendinous injury repairs and reconstruction
Redebridements and delayed wound closure
Delayed closures and debridement

Priority 4. Locked knee
Impending fractures
Periprosthetic fractures after hip and knee arthroplasty

Appendix B

Prioritization of emergency surgery protocol for VGH:

Scheduling of Surgical Cases

1. Purpose:

To define standard scheduling practices and ensure accurate scheduling of surgical cases across all Operative Sites

A. Elective Case: A case that can be scheduled at least two weeks prior to the preferred surgical date.
B. Urgent Case: A case that must be scheduled within ***48hrs***.
C. Emergency case: A case that must be scheduled within ***24*** hrs.
D. OR Locations: ECC – Eye Care Center
JPOR – Jim Pattison Pavilion OR's (VGH site)
UBCH – UBCH OR's (UBC site)
E. Block Scheduling – OR block time made available to each service and scheduled for a fixed time period. Each division is responsible for determining individual surgeon time and informing the OR booking clerk. ***Services must use block time before booking in to any other available time.***
F. If block time cannot be utilized by a service they must inform OR Booking a minimum of two weeks in advance.
G. Protected OR time – Block time provided to a service to accommodate urgent cases.
H. Operating room Slate – A list of scheduled surgical cases arranged according to site, date and OR room number
I. Surgical case – The information provided on the slate will include:
    i. Patient name
    ii. MRN number
    iii. Account #
    iv. Age, Birth date
    v. Sex
    vi. Post-op location
    vii. Surgeon
    viii. Anesthesiologist, anesthetic type
    ix. Procedure code, procedure description
    x. Diagnosis
    xi. Case #
    xii. Admitting process
    xiii. Date of Pre-Admission clinic visit
    xiv. Special considerations re isolation precautions and blood products
    xv. Comment field with special instructions regarding patient care or equipment required

J. The time allocated will be defined as follows:
    i. Setup/ start: time allotted for setting up the case in the room
    ii. Patient in – time the patient enters the room
    iii. Patient out – time the patient is expected to leave the room

iv. Total room time: Includes the time between when the patient enters and leaves the room.(*This includes patient positioning, anesthesia induction and extubation time, skin to skin time as well as preparing the patient for leaving the room e.g. dressings and drape removal etc.)

K. The difference between patient out time and the next case patient set up time is the time allotted for cleaning the room

L. OR Booking Package – The package that must be submitted to the OR Booking clerk to indicate the need for a patient to be scheduled for elective surgery. The package contains the OR booking form, procedure consent, a history and signed patient questionnaire.

2. Process

A. Scheduled cases: The OR booking package must be submitted to the booking clerk within 14 calendar days of the anticipated surgical date. If the block time is not fully booked within the 14 days then the service that has that time allocated will be notified. The office will have 48 hrs to indicate if a different surgeon within that department will be using the remaining time and the time must be booked and a complete package received within 48 hrs. Time not scheduled within 10 days of the surgical date will be made available for urgent bookings.

B. Urgent cases: Patients that require surgery within ***48 hrs*** and the surgeon does not have block time available may request access to urgent time. The request must include the patient name, birth date, location, surgeon, procedure code and description, skin to skin time and anticipated LOS. Urgent case time will be offered based on a first come/ first served basis by the OR booking clerk supervisor. (***These cases will then become scheduled and accommodated through the Pre-Admission assessment process if appropriate, considering the patient's status, location and capacity within the Pre-Admission clinic.)***

C. Slate " cut off " time – The time after which the slate is considered final and no further changes may be made except to have patients cancelled who are not able to proceed to the OR as expected. The slate cut off for all sites will be 1200, 48 hrs ahead of the calendar date.

D. Any ***unscheduled*** OR time made available after slate cut off will be allocated for ***urgent and emergency cases***.

Emergency cases will be booked through the OR control desk and scheduled for surgery using the designated "Emerg" rooms

Priorities for Emergency Surgery – Guidelines

All emergency surgeries are to be classified by the attending surgeon according to the following priorities when the surgery is booked on the emergency list.

Priority 1	Required within 1 hour to save life, limb, or organ
Priority 1b	Required within 4 hours, Trauma Surgery only
Priority 2	Required within 8 hours to save life, limb, or organ
Priority 3	Ideally done within 24 hours, but not more than 48 hours, then on the slate
Priority 4	Urgent elective to facilitate medical treatment, patient discharge or bed utilization, bumped (not cancelled) surgery.

Appendix C

Little's Law applies to any queuing system in steady state. Steady state in this context means that over time the number of inputs is equal to the number of outputs. The Law simply states that the expected number of items in the system is equal to the rate of arrival of inputs multiplied by the average processing time of an input.

$$L = \lambda W$$

L is the number of items in the system
λ is the arrival rate
W is the average time each item spends in the system (including the queue)

For a single server, first-in, first-out system, the PK formula provides the expected length of time spent in the queue before processing (inventory buffer):

$$Wq = \frac{1}{\mu}\left(\frac{\rho}{1-\rho}\right)\left(\frac{c_a^2 + c_s^2}{2}\right)$$

Wq is the waiting time in queue
1/μ is the average service time (actual process)
ρ is the expected number in service or utilization $\left(\rho = \lambda/\mu\right)$, in other words, the number in service is equal to the rate of arrival multiplied by the processing time
$\left(\frac{c_a^2+c_s^2}{2}\right)$ represents the variability seen in rate of input and variability in processing times. This term can be defined as V (the variability factor)

The total expected inventory (items in queue and in work in process) is:

$$L = Lq + \rho$$

Then, separating the queue from work in process and substituting Little's Law, gives:

$$L = \lambda W_q + \rho$$

Substituting the PK formula then gives:

$$L = \lambda\left(\frac{1}{\mu}\left(\frac{\rho}{1-\rho}\right)\left(\frac{c_a^2 + c_s^2}{2}\right)\right) + \rho$$

$$L = \rho\left(1 + \left(\frac{\rho}{1-\rho}\right)\right)\left(\frac{c_a^2 + c_s^2}{2}\right)$$

$$L = \rho\left(1 + \left(\frac{\rho}{1-\rho}\right)\right)V$$

This demonstrates the trade-off of inventory (L), utilization (ρ), and variability (V). When variability rises, utilization must fall in order to maintain the same inventory (queue length); similarly, in order to increase utilization, variability must decrease otherwise queue length will increase. This is the basis of the OM triangle.

Reference

Schmidt, Glen M. Teaching the OM triangle. *Operations Management Education Review*. 2005 (1): 87-104.

Appendix D

Selection from data collected on X-Times of various procedures (representative buckets):

Procedure Code	Procedure Description	# of Cases	Min X Time	Max X Time	Std Dev X Time	Avg X Time	Olympic Average	Time Bucket (Avg)	Time Bucket (Olympic
37271	MASTECTOMY PARTIAL W/ AXILLARY NODE DISSECTION	44	24	60	8.8	37.5	35.5	45	30
62089	ORBITAL BIOPSY	7	22	66	16.2	37.6	25.0	45	30
39337	URETEROSCOPY COHERENT HOLMIUM YAG LASER	15	20	66	15.3	37.7	31.9	45	30
500249	VAGINAL RECTOCELE REPAIR	4	24	61	16.2	37.8	16.5	45	30
620882	ORBITOTOMY ANTERIOR COMPLICATED	5	24	62	14.4	37.8	20.6	45	30
34903	THIGH SOFT TISSUE TUMOR EXCISION	33	5	92	14.6	37.9	34.9	45	30
500192	MIS PELVIC LYMPHADENECTOMY	42	25	59	7.6	37.9	35.9	45	30
30051	INGUINAL NODE DISSECTION	10	25	70	14.2	38.4	28.9	45	30
310396	AICD - PACEMAKER	7	12	48	12.2	38.4	29.9	45	30
30047	HERNIORRHAPHY INCISIONAL	24	24	88	18.3	52.1	47.4	60	45
54709	FEMUR # INSERTION NAIL SYNTHES OSTEOTOMY	9	32	63	10.5	52.3	41.8	60	45
39065	PROSTATECTOMY RETROPUBIC	8	23	97	21.1	52.5	37.5	60	45
320161	DISCECTOMY ANTERIOR CERVICAL W/ILIAC CREST BONE GRAFT	9	39	61	7.5	52.7	41.6	60	45
320183	DEEP BRAIN STIMULATION STAGE I & II	19	18	104	20.8	52.8	46.4	60	45
34919	HIP TOTAL PRIMARY VERSYS/TRILOGY (MINI/STANDARD INCISION)	10	39	86	16.2	52.8	40.3	60	45
34116	KNEE TOTAL REVISION	63	31	88	12.1	53.5	51.6	60	45
24053	MAGNETIC RESONANCE IMAGING	13	8	195	48.1	53.6	38.0	60	45
30014	COLOSTOMY	13	24	93	24.3	54.5	45.5	60	45
32046	VENTRICULO-PERITONEAL SHUNT	32	38	99	13.3	54.6	50.3	60	45
50209	MIS OOPHORECTOMY	81	22	73	8.7	39.2	38.0	45	45
24051	GDC COILING	64	15	210	26.4	67.3	63.7	75	60
30981	RADIOFREQUENCY LIVER ABLATION	12	45	117	20.2	67.8	54.3	75	60
360024	THORACOTOMY	12	23	115	26.3	67.8	56.3	75	60
36032	MIS THORACOSCOPY (PLEUROSCOPY) POSSIBLE THORACOTOMY	72	35	148	21.1	67.9	65.4	75	60
34959	HIP TOTAL REVISION GMRS/TRILOGY	12	50	90	14.0	68.7	57.0	75	60
30507	MIS SEGMENTAL LIVER RESECTION	19	38	112	21.1	69.8	61.9	75	60
39040	RENAL TRANSPLANT RECIPIENT ST I	51	20	169	28.3	70.4	66.7	75	60
34953	HIP TOTAL REVISION ZMR XL/TMR	12	56	98	11.9	70.8	57.9	75	60
391013	CYSTECTOMY RADICAL W/ILEAL LOOP MALE	29	41	137	19.4	70.8	64.7	75	60
310574	CABG X5 & IMA OR BIMA	22	20	93	15.8	71.3	66.2	75	60
39345	MIS NEPHRECTOMY LIVING RELATED KIDNEY DONOR	43	38	83	10.2	57.3	54.5	60	60
99010	CAROTID THROMBOENDARTERECTOMY	135	27	152	14.5	57.4	56.1	60	60
34932	FEMUR DISTAL RESECTION & RECONSTRUCTION W/ GMRS	2	48	67	13.4	57.5	57.5	60	60
44041	SPINE LUMBAR LAMINECTOMY	37	31	110	15.5	57.7	53.9	60	60
99001	ABDOMINAL AORTIC ANEURYSMECTOMY	47	62	183	25.4	96.0	90.8	75	75
30021	HEPATIC ARTERY ANEURYSM REPAIR	1	97	97		97.0	97.0	75	75
30251	HEPATOPLASTY OF HEPATIC CYST MARSUPIALIZATION +/OR DRAINAGE	2	81	118	26.2	99.5	99.5	75	75
310392	VALVE AORTIC & MITRAL & TRICUSPID REP	3	81	118	18.5	99.7	99.7	75	75
320281	CRANIOTOMY ARTERIOVENOUS MALFORMATION RESECTION	14	54	223	41.3	100.2	80.4	75	75

Appendix E

Turnaround Time (TAT) Policy

Turn Around Time Policy – JP OR

Goal: Turn Around Time [TAT] will be 30 min or less (TAT = patient out time until next patient in time)

To achieve this goal:

1. The RNs in the Perioperative Care Centre [PCC] should have the patient ready to go to the OR 30 min prior to scheduled start time or sooner if OR slate is running ahead of schedule.
2. The OR RNs should call PCC ASAP if the room is running ahead of schedule by > 15 min.
3. The OR RN transporting the patient from the PCC to the OR should complete their check and have the patient in the OR within 30 min (or less if possible) of end of the last case.
4. Anesthesiologists* should assess their next patient as soon as they can after handing over care of their previous patient to the PAR staff. Notwithstanding that VA is a teaching institution, Fellows, Residents and Students must not delay cases from starting at the appropriate time.
5. Whenever possible the slating anesthesiologist will organize breaks for the anesthesiologists. When this is not possible it is expected that the anesthesiologists will take appropriate breaks in cooperation with the Nursing and Surgical staff, and timed to minimize delays.
6. The patient must not be moved from PCC until seen by the Anesthesiologist* unless the Anesthesiologist* calls into the OR and tells the RN that the patient can be moved to the OR, or the Anesthesiologist is already in the room getting prepared and tells the RNs that they may bring the patient to the OR.
7. Surgeons** should be immediately available to help position the patient and start the surgery once the patient has been anesthetized and the appropriate monitoring lines have been inserted.
8. Surgeons** should complete their patient assessment and mark the surgical site, (if required), in the PCC as soon as they have completed their previous case. The patient must not be moved from PCC until marking has been completed. Notwithstanding that VA is a teaching institution, Fellows, Residents and Students must not delay cases from starting at the appropriate time.
9. Between cases Surgeons** must either be in the OR suite or inform the OR RNs as to their whereabouts and how they wish to be contacted. Once contacted the Surgeon** should return immediately to their OR.
10. Perioperative assistants should be available promptly when required to help in their ORs.
11. Cleaning staff should commence cleaning the OR as soon as the patient leaves the OR.
12. OR RNs will only call the PAR when the patient is ready to be transported to the PAR. (Calling before that wastes time in the PAR that could be used by other patients).
13. If the patient requires special equipment in the PAR equipment such as a ventilator, the Anesthesiologist* will inform the PAR at least 15 min ahead of time (they may ask one of the OR RNs to do this on their behalf).
14. If no PAR RN is available to accept the patient, but there is space available, the patient should be transported to PAR and the Anesthesiologist*, and if necessary, the OR RN (at the discretion of the Anesthesiologist) will remain with the patient until a PAR RN is free to take over care, thus allowing housekeeping staff to clean the OR and the RNs to start the set up for the next case. In this event, arrangements must be made by the PAR staff as soon as reasonably possible to relieve the Anesthesiologist* so that they can proceed to assess their next patient.
15. When the TAT in an OR extends beyond the target time, the reasons will be documented by the Slate Expediter. This will include the names of all individuals who are non-compliant with the above policy. This information will be reviewed on a regular basis by the Perioperative management team and followed up with the appropriate manager and / or department head for action.

* "Anesthesiologists" include fellows or residents designated by the staff anesthesiologist to assume this responsibility

** "Surgeons" include fellows or residents designated by the staff surgeon to assume this responsibility.

Draft for Discussion - Confidential

Bibliography

Works Cited

Anupindi, R., Chopra, S., Desmukh, S., Van Mieghem, J., & Zemel, E. (2006). *Managing Business Process Flows: Principles of Operations Management.* Upper Saddle River: Prentice Hall.

Bhattacharyya, T., Vrahas, M. S., Morrison, S. M., Kim, E., Wiklund, R. A., Smith, R. M., et al. (2006). The value of the dedicated orthopaedic trauma operating room. *The Journal of Trauma: Injury, Infection and Critical Care* , 1336-1340.

Conway, J. B., Golberg, J., & Chung, F. (1992). Preadmission anaesthesia consultation clinic. *Canadian Journal of Anesthesia* , 1051-1057.

Denton, B., Viapiano, J., & Vogl, A. (2007). Optimization of surgery sequencing and scheduling decisions under uncertainty. *Health Care Management Science* , 13-24.

Dexter, F. (2006). A researcher responds to block FAQs. *OR Manager* , 16-17.

Dexter, F., & Epstein, R. H. (2003). Scheduling of cases in an ambulatory center. *Anesthesiology Clinics of North America* , 387-402.

Dexter, F., & Epstein, R. H. (2009). Typical savings from each minute reduction in tardy first case of the day starts. *Anesthesia and Analgesia* , 1262-1267.

Dexter, F., & Traub, R. D. (2002). How to schedule elective surgical cases into specific operating rooms to maximize the efficiency of use of operating room time. *Anesthesia and Analgesia* , 933-942.

Epstein, R. H., & Dexter, F. (2000). Economic analysis of linking operating room scheduling and hospital materials management information systems with just-in-time inventory control. *Anesthesia and Analgesia* , 337-343.

Fomundam, S. F., & Herrmann, J. W. (2007). *A survey of queuing theory applications in healthcare.* College Park: Institute for Systems Research.

Friedman, D. M., Sokal, S. M., Chang, Y., & Berger, D. L. (2006). Improving efficiency through parallel processing. *Annals of Surgery* , 243: 10-14.

Gaba, D. M., Fish, K. J., & Howard, S. K. (1993). *Crisis Management in Anesthesiology.* Los Angeles: Churchill Livingstone.

Golen, J. (2001). Making a case for a case cart system. *Herman Miller for Healthcare* .

Harders, M., Malangoni, M., Weight, S., & Sidhu, T. (2006). Improving operating room efficiency through process redesign. *Surgery* , 509-516.

Hunter, J. (2009, July 23). *A clear vision for better health care.* Retrieved July 25, 2009, from The Globe and mail: http://www.theglobeandmail.com/news/national/british-columbia/a-clear-vision-for-better-health-care/article1223207/

Hunter, J. (2009, July 15). *B.C. health authorities told to trim $360 million.* Retrieved July 25, 2009, from The Globe and Mail: http://www.theglobeandmail.com/news/national/british-columbia/bc-health-authorities-told-to-trim-360-million/article1219190/

Kindscher, J., & Rockford, M. (2009). Operating Room Management. In R. Miller, *Miller's Anesthesia.* Los Angeles: Elsevier.

Lamiri, M., Xie, X., Dulgui, A., & Grimaud, F. (2008). A stochastic model for operating room planning with elective and emergency demand for surgery. *European Journal of Operational Research* , 1026-1037.

Leibovitz, P. (2003). Why can't my cases start on time? *AORN Journal* , 594-597.

Lemos, D., Nilssen, E., Khatiwada, B., Elder, G. M., Reindl, R., Berry, G. K., et al. (2009). Dedicated orthopedic trauma theatres: effect on morbidity and mortality in a single trauma centre. *Canadian Journal of Surgery* , 87-91.

Leslie, M., Hagood, C., Royer, A., Reece Jr, C., & Maloney, S. (2006). Using lean methods to reduce OR turnover times. *AORN Journal* , 849-854.

Livesmart BC. (2009, June 30). *Carbon Neutral Action Reports: Health Authorities.* Retrieved July 25, 2009, from Livesmart BC: http://www.livesmartbc.ca/government/health_authorities.html

Lovejoy, W. S. (1998). Integrated operations: a proposal for operations management teaching and research. *Production and Operations Management* , 106-124.

Lovejoy, W. S. (1996). Variability, Buffers, and Inventory. University of Michigan, Business School.

Mazzei, W. (1994). Operating room start times and turnover times in a university hospital. *Journal of Clinical Anesthesia* , 405-408.

McKesson. (2009). *OR Benchmarks Collaborative*. Retrieved July 26, 2009, from McKesson: http://sites.mckesson.com/orbc/

McLaughlin, D. B., & Hays, J. M. (2008). *Healthcare Operations Management.* Chicago: Health Administration Press.

Ministry of Health Services. (2009). Access to Surgery in British Columbia. The cutting edge. *January 15-16, 2009 Conference Report* (pp. 1-43). Vancouver: Ministry of Health Services.

Naik, G. (2006, November 14). *A hospital races to learn lessons of Ferrari pit stop.* Retrieved November 14, 2006, from The Wall Street Journal Online: http://online.wsj.com/article_print/SB116346916169622261.html

Noon, C. E., Schiffer, H., & Crane, J. (2009). How to have a Leaner OR department. *OR Manager* , *25*, 11-12.

OR Benchmarks. (2009). *OR Benchmarks Collaborative*. Retrieved July 26, 2009, from OR Benchmarks: http://www.orbenchmarks.com/ORB_Collaborative/index.htm

OR Benchmarks. (2000). ORs benchmark activities for the day of surgery. *OR Manager* , 1-3.

Overdyk, F. J., Harvey, S. C., Fishman, R. L., & Shippey, F. (1998). Successful strategies for improving operating room efficiency at academic institutions. *Anesthesia and Analgesia* , 896-906.

Pandit, J., & Carey, A. (2006). Estimating the duration of common elective procedures: implications for operating list management. *Anaesthesia* , 768-776.

Patterson, P. (2009, May). Getting the staff's buy-in for Lean. *OR Manager* , 13-14.

Queuing Theory. (2009, July 25). In *Wikipedia, the free encyclopedia.* Retrieved July 30, 2009, from http://en.wikipedia.org/wiki/Queueing_theory

Schlechter, K. L. (2009, March 2). *Hershey Medical Centre to open redesigned emergency department.* Retrieved July 30, 2009, from Pennlive.com: http://www.pennlive.com/midstate/index.ssf/2009/03/hershey_med_to_open_redesigned.html

Shelver, S. R., & Winston, L. (2001). Improving surgical on-time starts through common goals. *AORN Journal* , 506-511.

Slack, N., Chambers, S., & Johnston, R. (2007). *Operations Management.* Madrid: Prentice Hall.

Spear, S. J. (2005). Fixing health care from the inside, today. *Harvard Business Review* , (9) 78-93.

Sprengel, A., Snell, W. J., & Boissoneau, R. (1993). Specialty Surgical Teams. Results of a study. *AORN Journal , 58* (6), 1170-1180.

Strum, D., Vargas, L., & May, J. (1999). Surgical subspecialty block utilization and capacity planning. *Anesthesiology* , 1176-85.

Truong, A., Tessler, M., Kleiman, S., & Bensimon, M. (1996). Late operating room starts: experience with an education trial. *Canadian Journal of Anesthesia* , 1233-1236.

Tucker, J. B., Barone, J. E., Cecere, J., Blabey, R., & Rha, C.-K. (1999). Using Queueing Theory to Determine Operating Room Staffing Needs. *The Journal of Trauma: Injury, Infection, and Critical Care* , 71-79.

Tyler, D. C., Pasquariello, C. A., & Chen, C.-H. (2003). Determining optimum operating room utilization. *Anesthesia and Analgesia* , 1114-1121.

Wheelwright, S., & Weber, J. (2004). *Massachusetts General Hospital: CABG Surgery (A).* Boston: Harvard Business School Publishing.

World Health Organization. (2009). *Patient Safety*. Retrieved July 28, 2009, from World Health Organization: http://www.who.int/patientsafety/safesurgery/en/

Zhou, J., Dexter, F., Macario, A., & Lubarsky, D. (1999). Relying solely on historical surgical times to estimate accurately future surgical times is unlikely to reduce the average length of time cases finish late. *Journal of Clinical Anesthesia* , 601-605.

Interviews

Ms. Tamara Holtham, RN, Patient Services Manager JP Operating Rooms. June 22, 2009.

Dr. Anthony Boulton, MD, FRCPC, Anaesthesiologist, Former Medical Director of Perioperative Services, JP Operating Rooms. June 30, 2009.

Mr. Albert Csapo, Former Slate Expediter, JP Operating Rooms. June 30, 2009.

Works Consulted

Patrick, Jonathan and Puterman, Martin. Reducing wait times through operations research: Optimizing the use of surge capacity. 2007.