8 USING SIMULATION IN AN ACUTE-CARE HOSPITAL: EASIER SAID THAN DONE

Michael W. Carter[1] and John T. Blake[2]

[1] Department of Mechanical and Industrial Engineering
University of Toronto
Toronto, Ontario, Canada M5S 3G8

[2] Department of Industrial Engineering
Dalhousie University
Halifax, Nova Scotia, Canada B3J 2X4

SUMMARY

Simulation, as it is typically taught, is a rather mechanical process. Students are taught to follow a recipe: analyze a system, design a model, convert the model to computer code, collect data, verify, validate, and analyze the output. In practice, many analysts find that simulation is an odd combination of art, science, and marketing. Using this technique appropriately, in any industry, involves more than simply following the text book. In our experience, health care provides some rather unique challenges for the modeler. This chapter describes four different practical examples of using simulation to analyze a problem in an acute care hospital. The specific examples are not described in detail, since the applications have appeared in other publications. The emphasis here is to present some of the obstacles that were encountered and the lessons learned.

KEY WORDS

Simulation, Nursing human resources, Surgical schedules, Emergency department modeling, Drug order entry

8.1 INTRODUCTION

Simulation has a vast range of application in health care. Anyone who has ever visited a hospital emergency room, undergone surgery, or even visited their family doctor will recognize that the provision of health care is a complex, stochastic process with an overall structure analogous to a network of queues. The heterogeneity of customers in this system, the vast range of potential paths through the network, and the time-sensitivity of service make health care a "textbook" application for simulation.

The application of simulation in a health care setting is not always as simple and straight forward as one might think from reading the standard texts. In this chapter we present four simulation studies and describe lessons learned during the projects. The objective of this chapter is not to describe how to conduct a simulation study, or to provide all details for the four projects, since this material appears elsewhere in the literature. Our goal is to give analysts an idea of the issues that arise when an operations research technique is applied to a health care setting.

The projects described in this chapter include: a study to evaluate the link between inpatient census and the surgical schedule; a study to evaluate the causes of, and solutions to, emergency room wait time in a pediatrics hospital; a pharmacy ordering model; and a generalized simulation model for an acute care emergency department. In each instance, the problem is described, and an overview of the solution methodology is presented along with a summary of results. Each section concludes with a summary of the lessons learned during the project.

8.2 EVALUATING THE IMPACT OF THE ELECTIVE SURGERY SCHEDULE ON RESOURCE ALLOCATION

8.2.1 Description of the application

Nursing, like many regulated health care professions, tends to go through human resource availability cycles. The length of time required to fully train a doctor or a nurse (four years or longer in many jurisdictions) means that decisions made today regarding training spaces in universities and colleges only have a noticeable impact five to ten years later. Of course, in the period of time between when the plans are made and come to fruition, the demand for health care professionals may have changed. This is a common problem in almost all medical human resource planning.

Nursing, as a profession, has a number of unique characteristics that make human resource planning more difficult still. The profession is disproportionately female, and thus child rearing and family responsibilities have an impact on participation in the market place. As with other professions, general economic conditions, quality of work-life issues, and random fluctuations in the labor market also affect participation.

In 1989, Toronto experienced a shortage of qualified nurses. A good economy, combined with a rapidly rising housing market in the metropolitan Toronto area, caused a net outflow of nursing personnel from downtown to suburban institutions. To deal with this problem, nursing leaders from five urban Toronto hospitals collaborated to discuss possible ways to attract and retain nurses in their institutions. The number one nursing complaint, the amount of money paid to nursing staff, was not open to change. The second most important issue was the work week; in short, nurses wanted less weekend work.

One of the members of the committee facetiously suggested, “If we did surgeries on Monday for people with length of stay of four nights, and did only day surgery on Friday, we could empty the wards on the weekend, and give nurses more weekends off.” The suggestion was clearly not practical, but the idea that we could change the surgical schedule to reduce the weekend ward census was thought to be interesting. A project was subsequently funded by a grant from the Ontario Ministry of Health and five Toronto teaching hospitals: Toronto Hospital for Sick Children, Toronto General Hospital, Sunnybrook Health Science Centre, Mount Sinai Hospital and Toronto Western Hospital. (Our co-investigator was Professor Linda O’Brien-Pallas, Faculty of Nursing, University of Toronto.)

The study lasted for two years in 1991-93 and involved developing a simulation model to use as a decision support tool [1, 2]. The model included the operating rooms, the recovery room, intensive care units and regular inpatient wards. We were primarily interested in surgical patients since 90% of all surgical patients were scheduled in advance and, therefore, were somewhat controllable. Conversely, it was felt that nothing could be done to control medical patients, since 90% of all medical patients were emergency admissions. In all of the hospitals in the study, operating room time was assigned on a “block booking” basis. Surgeons received blocks of operating room (OR) time (e.g., every Monday morning for three hours) and were free to schedule patients in any order within their assigned blocks. Typically, elective surgery took place Monday to Friday on the day shift with one or two rooms available nights and weekends for emergencies.

Given this arrangement, we concluded that by changing the weekly OR schedule, we could influence workload and census in the rest of the hospital. By extension, we argued, it should be possible to determine a schedule that would be optimal from a staffing perspective. Furthermore, because we did not anticipate making any changes to the number or length of assigned blocks, we assumed that our schedule would have no impact on patients or their clinical care. The only impact, as far as we could tell ahead of time, would be minor inconvenience to surgeons who might have their block time rearranged within the master surgical schedule.

With these assumptions in mind, we built a simulation model, a database and user interface for the simulation. The model included all scheduled surgical patients and allowed for emergency patients who could preempt elective surgery as well as medical patients competing for intensive care unit (ICU) beds. The database included an underlying nursing workload model that estimated the total hours in each ward given the patient mix and volume flowing through it. If there were no beds available or not enough nursing hours when a patient was to be admitted, elective surgery would be canceled. The model generally used a first-come-first-served logic for allocating scarce resources. A small percentage of patients were also canceled for other reasons.

We used a two-pronged approach to collecting data for the model. We spent several months in each hospital analyzing the process to understand how patients flowed through the facility, creating process flow charts and collecting unique site-specific data. We also took advantage of an existing database of discharge records, The Canadian Institute of Health Information (CIHI). CIHI is a third-party organization that stores a discharge summary of every patient admitted to a hospital in Canada. Institutions in Canada are required to contribute data to this source, which is used by hospitals for their own internal review as well as by federal and provincial authorities.

Through the interface, the user was able to set the surgical schedule, make adjustments to the surgeons' case mix, specify the number of beds and nurses on each ward and change a variety of control parameters. The simulation itself was driven on a data trace. Because of the often confounding factors relating to age, gender, disease, co-morbidity, treatment, and outcome, we reasoned that it would be more practical to dispense with the idea of developing and fitting distributions for key simulation variables such as length of stay, processing time, etc, since we could not assume independent and identically distributed observations. Instead, we decided to sample directly from a large list of patients available from hospital discharge records. Thus when we needed to "create" a patient in our model, we randomly selected a person from this existing list and simply associated all of that patient's demographic, treatment, and outcome data with the simulated patient. This mechanism, we felt, would make the simulation easy to port between sites and easy to validate.

After running random patients through the model for a two-week warm up period, we ran ten replications of two weeks. Upon the completion of the run, we produced summary statistics on estimated annual patient volumes, cancellations, emergencies, and patient census and nursing hours in each ward by day of week.

8.2.2 Challenges encountered

Timing/project cycle time Simulations typically look simple to build; or at least they look simple at the start of a project. Our project was originally designed for a two-year cycle. The pilot model took approximately 12 months to complete. Ports to other institutions, which were scheduled to take two months, took about four months apiece to complete. Thus, by the end of the project, more than four years had elapsed since its inception in 1989. The reality of 1993 was much different than the reality of 1989, particularly in the health care sector in Canada. While 1989 was the high point in an economic cycle, 1993 was a low point. Thus, by the time our program was ready, the government was cutting health care budgets, and hospitals were laying off nurses! Simply put, weekend workload and quality of work-life issues had dropped off the radar screen; people were much more interested in holding onto their jobs than getting the weekend off.

Fortunately, this unexpected turn of events did not detract from the value of the project. The model turned out to be very useful as a mechanism to balance the use of increasingly tight hospital resources. The simulation allowed users to experiment with various allocations of OR time and forecast the impact of ward census, nursing workload, ICU beds and recovery rooms. Several of the sites used our model to improve their operations.

For example, at the Toronto Hospital for Sick Children, in one ward, the census was double on Wednesday night compared to every other day of the week. By making a few minor adjustments, we were able to suggest an OR schedule that would balance the nursing workload over weekdays. As another example, we used the model to look at Christmas closing in 1994 after the Ontario government asked hospitals to close all elective surgery for two weeks as a cost reduction measure. Mount Sinai Hospital asked us to complete an analysis of residual demand for OR time and ward space due to emergency patients. We used the model to predict the staffing levels that would be needed to cover this demand for the two weeks. At Sunnybrook Health Sciences Centre the model was used in a number of planning scenarios, not to balance nursing workload, but to calculate production limits for their cardiovascular surgery program.

At one institution we ran into a problem validating our model. The simulation model suggested that the OR time currently allocated to the Ear, Nose and Throat (ENT) service could accommodate almost twice as many patients as they were actually serving. We searched for the cause of the discrepancy for several days in the simulation. Ultimately, we discovered that ENT had a habit of not always using all of their allocated time. Meanwhile, General Surgery was starving for OR time. Whenever ENT did not need the allocation, someone in General Surgery was happy to use it. The OR managers had not noticed the problem since all of the booked rooms were being fully utilized.

Data collection In any simulation, data collection, verification, and validation are major issues. In our experience in health care, no one ever had the right data in the form that we needed it. Health care information systems are typically designed to meet clinical requirements, not administrative needs. The CIHI data was a mixed blessing for our project. The CIHI data was universally available for all institutions, in a standard format, and from a single source, and was thus easy to access and import into the simulation. We did, however, find a number of weaknesses in the database which limited its applicability for a simulation study.

Since the discharge summary is only a summary of what happened to a patient, it was not always possible to entirely reconstruct a patient's process through the hospital from their discharge report. For example, a patient admitted as a medical patient for treatment of diabetes falls and breaks a hip during her hospitalization. If, at discharge the broken hip is considered to have contributed more to the patient's length of stay than the diabetes, the patient may then have been labeled as a surgical patient. Without complete access to a patient's record, reconstructing a patient's length of stay often involved some assumptions and some estimation.

Furthermore, we found that source data sent to CIHI was not always viewed by the institutions as reliable. (This is rather surprising given that the institutions themselves are responsible for abstracting and summarizing the data that is forwarded to CIHI.) Finally, the lag between when data was collected, abstracted, and made available to CIHI meant that we typically had to use patient abstracts that were at least a year old (and in one instance two years old) in the model. This led to a common complaint among potential users that the data was "too old" and "not representative of what we're doing now".

Every hospital was different Our model was designed to be flexible and to provide the ability to answer a wide variety of questions. We wanted to be able to test potential length-of-stay variations by age, disease, sex, etc. Furthermore, we wanted the model "portability process" to be as simple as possible. Our intention was to develop one generic model and simply move this model from place to place, plugging in new patient records and a small amount of site-specific data (i.e., the number of wards and the beds on each). In practice we found it very difficult to create a single, generic, general-purpose patient simulation. Each institution had a unique combination of services, programs, and unique "quirks" that made it difficult to directly move a model from one location to another. These quirks ranged from unique processing rules to arcane details of the physical plant.

For example, at The Toronto Hospital for Sick Children, the managers suspected that an old bank of elevators that frequently broke down significantly impacted transportation time! In this case, we included the elevator in our model.

When we initially developed the pilot simulation at The Toronto Hospital for Sick Children, we decided to restrict the model to patients who had only one surgical procedure. The number of cases of multiple surgeries there was quite small. However, Sunnybrook is a regional trauma center, and multiple procedures are relatively common. So, we needed to modify the Sunnybrook model to allow for multiple surgical procedures.

<u>Stakeholders</u> Getting the buy-in of all stakeholders is always a key component to any simulation project. However, when working in a health care setting, acceptance by all stakeholder groups is especially important. In this particular project, our assumption that the schedule rearrangement would be a minor issue turned out to be incorrect. Physicians, as a rule, control the creation of the master surgical schedule and guard it jealously. Schedule changes are almost never thought to be a matter of minor inconvenience.

In fairness to physicians, the schedule dictates both their income and their work schedule. That is why, in practice, the issue is so controversial that in most of the hospitals that we have worked in, the administration simply allocates total O.R. time to each service (cardiology, general surgery, orthopedics, etc.). The doctors in each service then decide among themselves how to allocate specific blocks of time. This solves some of their political issues but, as a consequence, the administration relinquishes any control over daily work flow balance.

8.3 CHILDREN'S HOSPITAL OF EASTERN ONTARIO (CHEO)

8.3.1 Description of the application

CHEO is a pediatrics teaching hospital affiliated with the University of Ottawa. In 1993, the hospital's Emergency Department expressed concern that up to 20% of patients were forced to wait at least two hours before being seen by a physician. The issue was one of quality of service rather than quality of care, since all patients are triaged promptly and urgent cases are seen right away. Long waits are generally associated with patients having "runny noses" and other minor complaints. However, with provincial budget cuts looming, managers at CHEO felt it important to maintain good public relations.

The Vice President of Ambulatory Care (VPAC) called us in May 1993. She had received eleven process improvement suggestions from staff members. Suggestions ranged from overhauling patient flow to making changes to the physical layout of treatment rooms. One suggestion called for installing video games in the waiting room so patients would not realize how long the wait was. While the VPAC thought that many of the suggestions were interesting, she needed a mechanism to provide quantitative analysis of the options.

To determine the impact of the various strategies on patient wait time a full-scale patient simulation model was developed [3]. The model included all of the major patient processes in the emergency department (ED): patient arrival, registration, triage, assessment, testing, treatment, and admission or discharge processes. Our main evaluation criteria were the average wait time and the distribution of these times for each of the four triage categories defined by the hospital: emergency, urgent, deferrable, and medical walk in.

In terms of modeling effort, the simulation itself was relatively simple. However, data collection, model validation, and output analysis required significant effort. One of the first things we discovered when we started collecting data at the hospital was the highly fractured nature of work in the ED. CHEO is a teaching hospital. The ED was staffed by one to three physicians, called Casualty Officers (COs), five to seven nurses, and a number of residents. Normally, each patient was seen by a nurse, a resident, and the CO who reviewed the resident's assessment. On any given shift there were ten patient treatment rooms available for use. Patients in these rooms were under the care of one CO who might also have had responsibility for providing medical education to one or two medical residents.

We noted immediately that it was extremely rare for any worker (physician, nurse, or resident) to complete a work cycle on one patient from start to end with no interruptions. More commonly, we observed that nursing and physician services were delivered in small discrete batches spaced over a fairly long time period. For example, a physician might assess a patient and order a test. A nurse might then collect a sample or transport the patient to another area. During the time the test was being performed, the physician would move on to treat other patients. When the results of the test became available, the physician would read the results, interpret them, and order a treatment or send the patient home.

Since a physician had five to ten patients "on the go" at any time, work cycles became quite fractured. Indeed, physician work cycles were nothing short of chaotic given the additional requirement of also providing medical education for students and residents. The physician was, for example, required to confirm the resident's diagnosis, provide him or her with background about a disease state or treatment option, and then confirm test results, treatment opinions, or patient instructions. Since the casualty officer was legally responsible for the patient's care, no part of the treatment process could occur without the permission of the CO. In fact, we later found that COs spent about as much time interacting with residents as with patients.

Once we had the model working and validated, we started a designed experiment. The factors that we varied included the number of COs on shift, the number of residents on shift, and the queuing discipline used to select patients from the waiting list. We did not find much of a factor effect for queue discipline, but we did note a strong negative effect for the number of COs on shift and a strong positive effect for the number of residents on shift.

As described earlier, resident education was a major component of work in the ED. Our experiment indicated that the work created by resident education was so great that eliminating all residents from the ED would substantially reduce patient waiting time! In fact, our model indicated that adding one additional CO, or eliminating all residents, would result in approximately the same improvement in waiting time.

Obviously, eliminating residents from a teaching hospital is not a practical alternative, but the results indicated that waiting time could be impacted by a number of different scenarios, including different numbers of physicians, different shift schedules, and/or the addition of a hospital "walk-in clinic" to treat patients with minor injuries.

These scenarios led us to one of our more interesting results. As part of our plan to rearrange physician schedules, we prepared a simple plot of patient arrival times for each day of the week. We compared this to the COs' shift schedule. We found that demand peak (patients) often occurred several hours before the staffing peak. For example, on Sundays, the peak patient arrival period was between 10 a.m. and 1 p.m., but the peak staffing levels were scheduled for 5 p.m. to 7 p.m. Needless to say, the wait times for patients arriving in the afternoon were extremely long because a queue had been building all day. We were able to make significant improvements simply by staggering the doctors' start times.

Other major recommendations that came from this project included adding an additional four hours of CO time daily to the main ED and implementing a fast-track clinic for low-acuity patients. We estimated that these improvements would reduce patient wait time by as much as 20%. Although the approval process took over a year, the hospital did eventually hire a new casualty officer due, in large part, to our analysis.

8.3.2 Challenges encountered

Data collection The fractured nature of work in the ED presented a data collection problem for us. While good theoretical and practical models of nursing workload are available, no corresponding workload standards exist for physicians. As a result, it was very difficult to determine, for example, the demand for physician time resulting from a patient presenting symptoms of asthma.

Furthermore, the highly fractured nature of work cycles made manual data collection a difficult task. For example, much of the work a physician performed on a patient's file was done when the physician was distant from the patient (e.g. reading x-rays, interpreting test results, discussing with nurses or residents). Thus, measuring physician contact time was not an entirely accurate method of determining workload.

"Job shadowing" also presented some difficulties. For example, the nature of patient confidentiality precluded an observer from direct access to many types of patient-physician encounters. All in all, identifying accurate physician workload was a difficult task. We were, however, able to satisfy our data requirements through a combination of statistical work sampling and job shadowing. One of the project team members undertook the work sampling procedure, which could be performed without the observer necessarily having to be in the vicinity of the patient and the physician. In addition, the hospital provided us with two nurse instructors, who performed a physician job shadow. As clinicians, both physicians and patients accepted the nurses. In the end, we were able to build a reasonable data sample using the two techniques.

In this application as well as most of the others, we discovered that the length of time required for any particular task is extremely variable. When things get busy in the ED, everyone tends to work a little faster. In particular, the casualty officers spend much less time teaching as the demand increases. This is not surprising, but it creates some serious modeling challenges. One way to avoid this issue, as we did, was to use process times based on data collected during the busy times. Our real objective in this study revolved around queue length during busy times. As a result, our simulated patients were treated faster than the real patients during relatively quiet times.

<u>Time frame</u> A key challenge we faced with this project was finding the time to collect data, build the model, and run a reasonable set of scenarios. While the project originally was envisioned to be a short term two-week project, in the end we spent almost a year working on the model and its various components. Building the actual simulation model, as it turned out, was not particularly difficult or time-consuming. In fact, it took us about two weeks to build. The time consuming aspect of the project was data collection. To complete data collection, it was necessary to: identify the data necessary to run the simulation, make appropriate simplifying assumptions, define the method by which this data should be collected, assign personnel to data collection, and then collect the data.

Once the model was up and running, we found it was not possible to simply complete a set of runs, write up the results, and put the project behind us. Management at the hospital viewed the model as a useful planning tool. As the planning process developed at the hospital, we were asked to run the model under different assumptions and scenarios. Coincidentally, as we developed and ran these scenarios, our understanding of the ED process increased and we were able to point out to management results we felt were interesting. This resulted in a collaborative arrangement between management and modelers which, while fruitful, extended the project completion date.

8.4 MODELING THE DRUG ORDER ENTRY PROCESS FOR INPATIENTS

8.4.1 Description of the application

Currently, in the vast majority of hospitals in North America, doctors still prescribe medications for hospital inpatients by scribbling notes on paper. In one study published in 1998, Ash, Gorman and Hersh [4] found that fewer than 2% of U.S. hospitals had Computerized Physician Order Entry (CPOE) completely or partially available and required its use by physicians. The initial cost of implementing CPOE is one major obstacle for hospitals. At Brigham and Women's Hospital, the cost of developing and implementing CPOE was approximately $1.9 million, with $500,000 in maintenance costs per year. Installation of even "off the shelf" CPOE packages requires a significant amount of customization for each hospital and can be very expensive [5]. Finally, there may be cultural obstacles to CPOE implementation. For example, many physicians resist the idea of ordering prescriptions via computer instead of by hand. Although summary results were not available, the Leapfrog Group hospital survey [6] indicated that most U.S. hospitals are in the process of implementing CPOE.

On the surface, the manual Medication Administration Process appears quite simple. The physician writes a prescription on paper at the bedside and puts the order in the patient's chart. The nurse retrieves the order, transcribes it onto the "Medication Administration Record" (MAR) and leaves a copy of the order in a tray in the ward to be picked up by pharmacy technicians at routine times throughout the day. A pharmacist reviews the order and transcribes it into a computer with access to electronic patient records and decision support capability. The order is prepared in the pharmacy and delivered to the ward. The nurse checks drugs against the MAR and administers to the patient. The nurse records the administration on the MAR.

What is wrong with this picture? The doctor relies on memory/knowledge to determine the dose of the medication, to think of patient allergies and to remember possible drug interactions. The nurse may not know that an order has been written or that the drug has arrived. The multiple transcriptions increase the possibility of error and are not value-added work. The physical transport of the order wastes time. If the nurse cannot read the order, s/he must check with the doctor. If the pharmacist has any questions about medication or dosage, s/he must page the nurse and/or the doctor and hold the prescription until the order is confirmed. We believe that the process could be greatly improved if the doctor entered the order directly into a computer, using a handheld device, at the bedside.

Dr. Glen Geiger is a physician in Internal Medicine at Sunnybrook & Women's College Health Science Centre in Toronto. In 1999, Glen initiated a study where he asked doctors and nurses in his service to record process times on the drug orders. He discovered that over 25% of the orders were not administered within the targeted time frame. Most failures were not even close. These were process errors; they do not include cases where patients received the wrong drugs. The results of this study were a surprise to hospital leaders and continue to surprise health care professionals from other areas. Many issues that we discovered at Sunnybrook are common to most manual drug order entry processes.

It is fairly obvious that physician order entry will dramatically reduce cycle time for the process and reduce the workload of all parties – with the possible exception of the physician. Thus, we needed to convince the doctors that the system would dramatically improve the process without significantly increasing their workload. We decided to use simulation to quantify the potential for process improvement. We believed that it would be an important tool for demonstrating the advantages to physicians.

In the summer of 2001, four students, including three industrial engineers and one medical student, were hired to perform a detailed analysis of the prescription process. The students spent two months documenting the current process through interviews and direct observation. They then conducted a two week data collection during which all drug orders for a thirty-six bed Internal Medicine ward were tracked to facilitate the creation of a simulation model. The results of the detailed tracking confirmed Dr. Geiger's earlier results. In particular, many medication orders were not administered to patients in a reasonable amount of time [7].

One of the surprising discoveries was that this seemingly simple process was actually quite complex. For example, a different process was used when a doctor phoned an order in to the nurse as opposed to when the order was written. The day and night processes are different because the pharmacy is closed at night. At night, instead of placing an order with the pharmacy, nursing staff can access a night cupboard for commonly required medications.

There are also communication issues. For instance, pharmacists regularly visit the ward and review patient charts. Pharmacists sometimes write a "P" on the order. Some of the nurses knew that the "P" meant the pharmacist had reviewed the order. Others thought it meant they had "Pulled" the order. Also we found some confusion surrounding a physical flag attached to the chart. When the doctor writes an order s/he puts the flag up. Unfortunately, there is only one flag on each patient chart, and it is used for all orders. When multiple orders (e.g. drugs, lab tests, imaging, etc.) are in the file, the possibility exists that the nurse will find only the first one and put down the flag. Nurses check the complete chart every two hours, but errors sometimes occur. One order was in the chart for two days before the students pointed it out to the physician.

Timeliness of medication orders can be measured in several ways. For example, suppose a doctor prescribes medication for a patient at 11 a.m. to be administered three times a day at 6 a.m., 2 p.m. and 10 p.m. We could consider the delivery to be late if it was not back in time for the 2 p.m. dose administration. Pharmacokinetic practice says that a dose of medicine can be administered up to four hours late (for example, at 6 p.m.), half of the dosage interval, and still be on time. From a process perspective, we estimated that a prescription should not take longer than two hours to fill. All three measures were used in the study for determining whether an order was filled on time.

8.4.2 Challenges encountered

Lack of control From a quality perspective, we were quite surprised with the apparent lack of control of the prescription process. Since no written documentation was available, to determine how the process worked we simply asked everyone what he or she did. There was no formal training for nurses or doctors. New staff members learned by word of mouth. Virtually everyone we spoke with had a different view of the process. Moreover, there is no standardization across the hospital; each ward had apparently developed its own set of procedures. We attributed this to the perception that the process was "simple" and therefore did not require formal documentation and training.

Need for greater modeling detail In the validation of our simulation model, we could not get the turnaround times for medication orders in the model to match the times that we observed in practice. Initially, the average time in the model was 225 minutes, while the true average from the data was 262 or 16% higher. This seemed odd since the distributions in the model were based on statistically fitting the same data.

A major cause of this discrepancy originated in the pharmacy portion of the model. Initially, we had assumed that the pharmacy part of the process would be fairly reliable once the orders arrived there, so we chose to model the pharmacy as a black box. Since the pharmacy was computerized, we expected the process would provide consistent results and that when an order was picked up, it would be processed expeditiously and delivered back to the ward. We were surprised to discover dramatic variations in turnaround times.

Several months after the initial study, and long after the summer students had returned to school, we concluded that we needed to expand the scope of the analysis and perform a detailed process analysis of the pharmacy area. We discovered several anomalies. There was an 8 a.m. rush in the pharmacy to fill all of the orders that had accumulated overnight. The pharmacy processed each ward as a group of orders, and the sequence of the wards varied daily. Many long delays occurred when a particular ward was left toward the end of the sequence. We also learned that the pharmacy's workforce was highly variable. The pharmacy could not tell us how many pharmacists were working at any one time; the assignments varied hour-by-hour and day-by-day. Furthermore, it was found that a major complication was created by orders requiring clarification. We discovered that over 13% of orders required the pharmacist to call the doctor. These orders would be set aside temporarily while the pharmacist paged the doctor. We were unable to collect meaningful statistics on how long it took to get an answer to a page; however, it appears that about 5% of orders took more than three hours, and many of these were not resolved in the same day. Moreover, many of the pharmacists processed the simpler orders first, and saved clarifications until later in the day when they had some spare time.

Technology implementation As described above, the motivation for the simulation was to be able to demonstrate the potential process improvements that accrue from using automated physician order entry. Sunnybrook has already purchased the software to implement the automated prescription entry process. However, the system is still in development and the user interface must be customized. We cannot complete the simulation without first performing experiments with the interface to determine the distribution for access time. We do not expect it to take long, but this is likely to be the central measure of success for the physicians. We expect to have a pilot version ready by Spring 2004.

8.5 THE CROWDED STUDY: CAUSES AND RELATIONSHIPS OF OVERCROWDING AND WAITING IN DIFFERENT EMERGENCY DEPARTMENTS

8.5.1 Description of the application

Waiting times and overcrowding in the Emergency Department (ED) have become increasingly serious problems over the past several years. In the United States, surveys of hospital directors have reported ED overcrowding in almost every state [8, 9]; ED overcrowding has also been reported in Europe [10]. In most hospitals Emergency Department overcrowding is a symptom, rather than a cause, of the problem. For example, overcrowding in the province of Ontario in Canada is often attributed to patients who have been admitted to hospital, but who are waiting in the ED until a ward bed becomes available. Beds are often blocked in the wards because of discharge delays (e.g. waiting for test results, waiting for nursing home space, rehabilitation beds or home care). Thus, to really understand how the ED functions and why it backs up, it is necessary to develop a detailed process analysis specifically focusing on the impact of bed blockers.

A number of people have done ED simulations in the past, but have generally assumed that the processes outside the ED have little direct impact on its overall operation. Jun et al [11] present an extensive survey of simulation applications in health care. In fact, several simulation studies have been conducted to specifically analyze the issue of overcrowding in the ED. A priority queuing model was developed in one study to evaluate the potential impact of adding a fast-track facility to an emergency department [12]. Simulation modeling has also been employed to examine the relationship between hospital bed capacity and emergency admissions rates [13], with the finding that bed shortages can be expected when average bed occupancy rates exceed 85%. Simulations have been successfully applied to investigate the impact in the ED of nurse scheduling on utilization and patient length of stay (LOS) [14-16]. Based on these studies, recommendations were made for changing policies on staff scheduling, triage procedures and nursing responsibilities. Using the simulation model, the potential savings from the proposed changes were quantified.

The study described earlier in this chapter [3] also modeled the flow of patients through an ED. For all of the ED simulations mentioned, the patient LOS in the ED is assumed to be an exogenous variable, sampled once for each patient from a statistical distribution based on historical data. This assumption is reasonable given the complexity associated with most emergency departments. One can usually construct and validate these models quite adequately. However, this method does not allow decision makers to investigate the impact of changing non-ED components on the overall process flow. For example, if the time required to complete an external consult was reduced, or the process for MRIs was improved, what impact would that have on wait times in the ED or throughout the entire hospital?

In fact, our analysis suggested that the ED is a very complex entity, referred to by some of the doctors on our study team as “organized chaos”. In 2002, a team including operations researchers, ED physicians, a statistician and an epidemiologist received funding for a two-year study to analyze the detailed processes in ten Ontario hospital emergency departments. The Causes and Relationships of Overcrowding and Waiting in Different Emergency Departments (CROWDED) study was designed to include detailed data to promote better allocation decisions for scarce resources such as doctors, nurses, and examination rooms. The hospitals were selected to represent a cross-section of geography and clientele. Three large teaching hospitals, four community hospitals, and three rural emergency departments were selected for inclusion in the study. Two full time research assistants were hired for one year to collect data by directly observing patients, doctors and nurses. We conducted three trips to each site. There was a pre-visit of 2-3 days to study the layout, understand the policies, meet people, and put up posters to educate and inform people about the study. Data collection was conducted in two separate one week periods at different times of year to get a sense of pattern changes over time. The project was designed to construct a generic model of an ED that can provide detailed decision support for a wide range of process flow issues.

8.5.2 Challenges encountered

<u>Doctors are difficult to track</u> As mentioned earlier in the CHEO study, it is often difficult to tie physician workload to a specific patient. Doctors consult on the phone, read x-rays, view images on-line, chat with nurses and residents, as well as performing many other activities; all of these activities are done in the course of a patient's treatment, but rarely happen when the physician is proximate to the patient. However, since doctors are probably the scarcest ED resource, it is important to determine accurate workload information for them. In the CHEO study, we chose to implement a work sampling method supplemented with a job shadow provided by a small group of nurses. In the CROWDED project, we had significantly more resources at our disposal, and we were determined to get very accurate workload information.

Many physician and patient processes could not be observed directly. The observers needed to use indirect means of observation, such as consulting the patient chart or the "white board" that keeps track of patient progress in the ED. In some study sites, we had access to the hospital's electronic order entry/patient tracking systems. This also helped the observers to track the patients' pathways. However, in both paper and electronic documentation, it was found that recorded times did not usually reflect the actual time or duration of a process. For example, nurses or ward clerks might log an order for blood work into the computer at a certain time, but might not collect the blood until much later. The time recorded on the chart frequently corresponds to the time the order was entered; there is no information about the actual start and end time of the process.

<u>Missing data</u> A related issue we discovered during the course of the project was that it was quite difficult to collect complete, accurate flow data on all ED patients. The observers estimated that some data was missing for approximately 10-15% of patients in the study.

For example, critically ill patients may stay in the ED for a long time. The CROWDED study did not employ 24 hour observation, so process flow data tended to be incomplete. The observers tried, wherever possible, to fill in blanks using the patient's chart, but it was difficult to get good time estimates. When patients remain in the ED outside of the period of direct observation, the patient pathway through the ED will always have some missing data. However, even if some patient data was missing they were usually able to record a minimum data set including admission or discharge time along with any other charted information. Patients remaining in the ED for longer than a single observation shift tended to be admitted patients, or patients that required lengthy observation.

Trauma cases were also difficult to track. Because treatment for trauma cases needs to be started immediately, charting is usually performed after the fact. Moreover, trauma cases are generally handled behind closed doors. On the assumption that it is inappropriate for data collectors to be inside a trauma room or that observation may impede patient care, it was decided to forego direct observation of trauma cases. Instead the points of time of "trauma begins" and "trauma ends" were used as a way to track the many processes that could not be directly observed.

Similarly, acute patients may also receive treatment or undergo tests according to medical directives behind closed doors/curtains. In these cases, many different processes may be happening. The observers used the charts after the fact to determine which processes had occurred. This usually provided reasonable results in terms of what happened, but not always when it was done. It was sometimes possible to estimate start and/or end times if, for example, the observers saw nursing staff gathering up supplies or equipment prior to a process, but there was a lot of guesswork.

In addition to "closed door" treatments carried out by staff on trauma and acute patients, another challenge was that processes for many patients happen simultaneously. The research assistants were only able to observe the processes of one patient at any given time. Sometimes in the case of an acute patient such as, for example, a heart attack victim, a team of nurses and doctors might perform a series of treatments until the patient is stabilized. To capture all these processes required observation of that particular patient for an extended period of time. During that time, other things could be happening to other patients which were not observed or recorded.

<u>Layout issues</u> The layout of the ED sometimes created problems for data collection. Some EDs were physically spread out which made it difficult to see what was happening to a patient or to observe the doctor/nurse treating

the patient. In one ED in our study, the physical layout was divided into a "major" and a "minor" side. During peak times, doctors would be assigned exclusively to one side or the other. However, in off-peak times, one doctor would float between both. It became impossible to follow the doctor, the nurses and the patients simultaneously in this environment. At another site the ED had a number of separate areas. The segregation of the ED made tracking difficult for the observers.

<u>Fast-track clinic</u> Some study sites had an off-site "fast-track" clinic (FTC) or an "urgent care" center, separate from the ED. Again, this physical separation made it difficult to track patients.

At one site, the hospital had a fast-track clinic operating from 2pm - 10pm on weekdays. The FTC was in a separate area from the ED, but had four beds and was staffed by a Nurse Practitioner[1]. During its hours of operation, less acute patients came to the ED, saw the triage nurse, were registered by ED staff, but then headed to the FTC for treatment. The fact that the FTC is external makes it harder to observe the patient flow process. It was tempting to simply ignore patients who were sent to a FTC; however, we believe it is important to model it as an internal process, using ED resources. In particular, one of our model decision variables may be to consider adding two FTC physicians, or having some shared resources work in the FTC and the regular ED.

<u>Wait time before triage</u> When we consider the question of ED wait time, part of that measure involves patients waiting before triage or registration. Predictably, none of the study hospitals tracked or had data on "time before triage". While we believe serious patients are seen immediately and all patients are triaged expeditiously, we asked our observers to sit in the waiting room and conduct a separate study of "time-to-triage" to determine the magnitude of this issue. Observing time-to-triage; however, meant that observers could not track patients inside the ED due to layout and sight-line issues. The results of our preliminary studies indicated patients frequently line up to be triaged, but critically-ill patients were not overly delayed.

<u>Unplanned critical events</u> In any study, blind luck (good or bad) sometimes comes into play. In the CROWDED study the data collection process was facilitated by a custom designed PDA application. After the first few site visits were completed, the PDA programmer made some minor adjustments to the application. Subsequently, after three days of collection following the

---

[1] A Nurse Practitioner is a Registered Nurse who has taken a graduate level program and who can perform many of the functions that are commonly associated with doctors.

adjustment, the observers discovered that a bug in the revised program blocked the transfer of all patient demographic information (name, age, gender, ID number, etc.) to the production database. This was a serious issue in terms of validation and completion of missing data elements.

Additionally, a number of unforeseeable public health issues arose during the collection process. The ED at one site was closed for several weeks because of an outbreak of the Norwalk virus, which interrupted our data collection. To make matters worse, after three days of data collection at a different site the next week, one of the observers became ill with Norwalk like symptoms. She went into voluntary quarantine, and the other observer attempted to collect what she could for the remainder of the visit.

However, the worst setback occurred in March 2003 when Toronto was hit with the SARS virus (Severe Acute Respiratory Syndrome). We had to pull the observers out of all study hospitals for almost two months. Even hospitals distant from Toronto were closed to non-essential personnel. Moreover, patient volumes in EDs throughout the province decreased in response to patient fear of SARS. Things started to return to normal after a short period, but we needed to extend the data collection for two months, hire an extra observer, and adopt an aggressive visit schedule to make up for lost time.

Preemption When the ED gets busy, some processes can be preempted by more critical needs. When a doctor or nurse comes back to the interrupted process, they may have to start the entire process again. For example, while a nurse would not interrupt an IV start, he/she might interrupt an assessment. When the nurse later returns to complete the assessment, it is usually necessary to repeat some elements. One of our team members, an ED physician, believes the process can almost grind to a halt when things get busy. Physician assessments and nursing assessments are frequently interrupted. In our study the observers attempted to track starts and ends for all processes, even those that were incomplete, but this was an imperfect solution.

Administrative issues Despite our best efforts, staff members at the hospitals were often suspicious about the intentions of our study. It was perceived as a study created by the provincial government to streamline the costs of health care and reduce employment. Many staff members at hospitals believed the study would never be used to benefit health care or that the study was misguided. Our research assistants were conscientious in assuring the participating hospital staff that we were performing an independent study funded by CIHR, and not by their hospital administration or the provincial government, and that we were doing our best to accurately represent the processes in their departments so that patient flow could be improved without compromising patient care.

We also found that turnover in key management positions in the ED was a factor in our project. During our one year study we had the primary contact change at over half of the hospitals. Despite the fact that all of the study hospitals had agreed to be part of the project, we often discovered when we went to visit a site the current managers had no knowledge of our project, and we needed to begin the sales pitch again. In one case, we needed to reshuffle our data collection schedule because new managers did not know we were coming.

<u>Security</u> Data security was a very critical component of our study. During data collection, our team needed personal information to allow matches between paper and electronic hospital records. PDAs used for data collection were downloaded daily into the laptop computers and backed-up daily on a password protected CD-ROM, which was kept in a safe, secure location. Upon return to the lab, the data was copied from the laptop onto a master CD-ROM, which was kept in a locked drawer. The data on the laptop was then stripped of all personal identifiers (name, address, ID numbers, etc.) to ensure patient anonymity.

8.6 DISCUSSION/CONCLUSION

Health care is an enormous business offering a wealth of potential applications for simulation and other operations research techniques [17]. However, health care is a business unlike any other business. In our experience the context in which a decision making situation arises has a significant impact on the way in which it is solved. Nowhere is this truer than in health care. We believe that, because analysts and clinicians speak different languages, operations research has made fewer inroads into this field than in more traditional industries. However, our experience also suggests that OR techniques can be successfully applied in the health care setting. The secret is to understand the unique nature of the health care business and its impact on models, decision makers, and the development of implementable policies.

In this chapter we have used four simulation projects to highlight the practical lessons of applying operations research in health care. Analysts should remember that decision making in hospitals is characterized by multiple players; seeking the council and incorporating the objectives of all decision makers is vital in this environment. In this industry data collection systems may not be designed to provide administrative data; collecting data on patient flow and operational performance metrics may require some patience and may extend the project life cycle. Finally, while many processes and procedures are fundamentally similar regardless of the institution, there are usually enough local quirks to render multi-site “cookie-cutter” models infeasible.

Health care is a fascinating industry to work in. The authors have, over the past decade, devoted themselves to applying operations research to health care and have enjoyed the experience immensely. It is our desire that the lessons we learned will prove useful for others following in this field.

Acknowledgments

The authors would like to thank the Ontario hospitals that participated in the studies described in this chapter: The Toronto Hospital for Sick Children, the Toronto General Hospital, Mount Sinai Hospital, Sunnybrook and Women’s College Health Sciences Centre, the Children’s Hospital of Eastern Ontario (Ottawa), Markham-Stouffville Hospital (Markham), London Health Science Centre, Sudbury Regional Hospital, Stevenson Memorial Hospital (Alliston), South Muskoka Memorial Hospital (Bracebridge), Windsor Regional Hospital, Quinte Health Care Corp (Belleville), The Royal Victoria Hospital (Barrie) and Kingston General Hospital.

Thanks are also due to co-investigators and research assistants: Prof. Linda O’Brien-Pallas, Prof. Linda McGillis-Hall, Dr. Michael Schull, Dr. Glen Geiger, Prof. Greg Zaric, Dominic Fernandes, Yaron Derman, Cathy Wong, Carolyn Busby, Jessica Law, Tracy Ayow, Angelo Gentile Gosia Szymanski, Patricia Ricci, Aviv Gladman, Leila Peyravan and Rhea Plosker.

Finally, the authors would like to thank the Ontario Ministry of Health and Long Term Care, the Natural Sciences and Engineering Research Council of Canada, the Canadian Institute for Health Research and the Social Science and Humanities Research Council of Canada for their assistance funding these works.

References

[1] Blake, J.T., M.W. Carter, L.L. O'Brien-Pallas, and L. McGillis-Hall (1995). A surgical process management tool. *Proceedings of the 8th World Congress on Medical Informatics MEDINFO 95.*

[2] Carter, M.W., L.L. O'Brien-Pallas, J.T. Blake, L. McGillis, and S. Zhu (1992). Simulation, scheduling and operating rooms. *Proceedings of the 1992 Simulation in Health Care and Social Services Conference,* J.G. Anderson, Ed., Simulation Council Inc., San Diego, 28-30.

[3] Blake, J.T., M.W. Carter, and S. Richardson (1996). An evaluation of emergency room wait time issues via computer simulation. *INFOR,* 34, 263-273.

[4] Ash, J.S., P.N. Gorman, and W.R. Hersh (1998). Physician order entry in U.S. hospitals. *Proceedings of the AMIA Annual Symposium,* 235-239.

[5] Bates, D.W., L.L. Leape, D.J. Cullen, N. Laird, et al. (1998). Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. *Journal of the American Medical Association,* 280, 1311-1316.

[6] http://www.leapfroggroup.org, [online document] Accessed on June 29, 2003.

[7] Wong, C., G. Geiger, Y.D. Derman, C.R. Busby, and M.W. Carter, (2003). Redesigning the medication ordering, dispensing, and administration process in an acute care academic health science centre. *Proceedings of the 2003 Winter Simulation Conference,* S. Chick, P.J. Sánchez, D. Ferrin, and D.J. Morrice, Eds., New Orleans, LA, 1894-1902.

[8] Derlet, R.W. and J.R. Richards (2000). Overcrowding in the nation's emergency departments: Complex causes and disturbing effects. *Annals of Emergency Medicine,* 35, 63-68.

[9] Andrulis, D.P., A. Kellermann, E.A. Hintz, B.B. Hackman, and V.B. Weslowski (1991). Emergency departments and crowding in United States teaching hospitals. *Annals of Emergency Medicine,* 20, 980-986.

[10] Miro, O., M.T. Antonio, S. Jimenez, A. De Dios, M. Sanchez, and A. Borras (1999). Decreased health care quality associated with emergency department overcrowding. *European Journal of Emergency Medicine,* 6, 105-107.

[11] Jun, J., S. Jacobson, and J. Swisher (1999). Applications of discrete event simulation in health care clinics. *Journal of the Operational Research Society,* 50, 109-123.

[12] Siddharthan, K., W.J. Jones, and J.A. Johnson (1996). A priority queueing model to reduce waiting times in emergency care. *International Journal of Health Care Quality Assurance,* 9, 10-16.

[13] Bagust, A., M. Place, and J.W. Posnett (1999). Dynamics of bed use in accommodating emergency admissions: Stochastic simulation model. *British Medical Journal,* 319, 155-158.

[14] Kumar, A.P. and R. Kapur (1989). Discrete event application – Scheduling staff for the emergency room. *Proceedings of the 1989 Winter Simulation Conference,* MacNair, E.A., K.J. Musselman, and P. Heidelberger, Eds., IEEE, Washington, DC, 1112-1120.

[15] Rossetti, M.D., G.F. Trzcinski, and S.A. Syverud (1999). Emergency department simulation and the determination of optimal attending physician staffing schedules. *Proceedings of the 1999 Winter Simulation Conference.* Farrington, P.A., H.B. Nembhard, D.T. Sturrock, and G.W. Evans, Eds., Phoenix, AZ, 1532-1540.

[16] Kirtland, A., J. Lockwood, K. Poisler, L. Stamp, and P. Wolfe (1995). Simulating an emergency department is as much fun as .... *Proceedings of the 1995 Winter Simulation Conference,* Alexopoulus, C., K. Kang, W.R. Lilegdon, and D. Goldman, Eds., Arlington, VA.

[17] Carter, M.W. (2002). Health care management – Diagnosis: mismanagement of resources. *OR/MS Today,* April, 26-32.