

# Chapter 22

## Healthcare Analytics



Maqbool (Mac) Dada and Chester Chambers

### 1 Introduction to Healthcare Analytics: Simulation Models of Clinics in Academic Medical Centers

Ancient understanding of biology, physiology, and medicine was built upon observations of how the body reacted to external stimuli. This indirect approach of documenting and studying the body's reactions was available long before the body's internal mechanisms were understood. While medical advances since that time have been truly astounding, nothing has changed the central fact that the study of medicine and the related study of healthcare must begin with careful observation, followed by the collection, consideration, and analysis of the data drawn from those observations. This age-old approach remains the key to current scientific method and practice.

#### 1.1 Overview of Healthcare Analytics

The development of technologies related to information capture and analysis over the past 20 years has begun to revolutionize the use of data in all branches of medicine. Along with better and easier methods for the collection, storage,

---

**Electronic supplementary material** The online version of this chapter ([https://doi.org/10.1007/978-3-319-68837-4\\_22](https://doi.org/10.1007/978-3-319-68837-4_22)) contains supplementary material, which is available to authorized users.

---

M. (Mac) Dada (✉) · C. Chambers  
Carey Business School, Johns Hopkins University, Baltimore, MD, USA  
e-mail: [mdadal@jhu.edu](mailto:mdadal@jhu.edu)

and interpretation of data, the new technologies have spawned a number of new applications.<sup>1</sup> For example, data analysis allows earlier detection of epidemics,<sup>2</sup> identification of molecules (which will play an unprecedented role in the fight against cancer<sup>3</sup>), and new methods to evaluate the efficacy of vaccination programs.<sup>4</sup>

While the capacity of these tools to increase efficiency and effectiveness seems limitless, their applications must account for their limitations as well as their power. Using modern tools of analytics to improve medicine and care delivery requires a sound, comprehensive understanding of the tools' strengths and their constraints. To highlight the power and issues related to the use of these tools, the authors of this book describe several applications, including telemedicine, modeling the physiology of the human body, healthcare operations, epidemiology, and analyzing patterns to help insurance providers.

One problem area that big data techniques are expected to revolutionize in the near future involves the geographical separation between the patient and the caregiver. Historically, diagnosing illness has required medical professionals to assess the condition of their patients face-to-face. Understanding various aspects about the body that help doctors diagnose and prescribe a treatment often requires the transmission of information that is subtle and variable. Hearing the rhythm of a heart, assessing the degradation in a patient's sense of balance, or seeing nuances in a change in the appearance of a wound are thought to require direct human contact. Whether enough of the pertinent data can be transmitted in other ways is a key question that many researchers are working to answer.

The situation is rapidly changing due to the practice of telemedicine. Market research firm Mordor Intelligence expects telemedicine, already a burgeoning market, to grow to 66.6 billion USD by 2021, growing at a compound annual growth rate of 18.8% between 2017 and 2022.<sup>5</sup> New wearable technologies can assist caregivers by collecting data over spans of time much greater than an office visit or hospital stay in a wide variety of settings. Algorithms can use this data to suggest alternate courses of action while ensuring that new or unreported symptoms are not missed. Wearable technologies such as a Fitbit or Apple Watch are able to continuously track various health-related factors like heart rate, body temperature, and blood pressure with ease. This information can be transmitted to medical

---

<sup>1</sup>The article in *Forbes* of October 2016 provided many of the data in this introduction—<https://www.forbes.com/sites/mikemontgomery/2016/10/26/the-future-of-health-care-is-in-data-analytics/#61208ab33ee2> (accessed on Aug 19, 2017).

<sup>2</sup><https://malariajournal.biomedcentral.com/articles/10.1186/s12936-017-1728-9> (accessed on Aug 20, 2017).

<sup>3</sup>[http://cancerres.aacrjournals.org/content/75/15\\_Supplement/3688.short](http://cancerres.aacrjournals.org/content/75/15_Supplement/3688.short) (accessed on Aug 20, 2017).

<sup>4</sup><https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4287086> (accessed on Aug 20, 2017).

<sup>5</sup><https://www.mordorintelligence.com/industry-reports/global-telemedicine-market-industry> (accessed on Aug 23, 2017).

personnel in real time. For example, military institutions use chest-mounted sensors to determine the points at which soldiers reach fatigue and can suggest tactical options based on this information.

While some wearable technologies are on the cutting edge and are thus often expensive, telemedicine can use cheap, sturdy hardware to make diagnoses easier. Electronic kits such as the Swasthya Slate,<sup>6</sup> which is used in community clinics in New Delhi, can be used by doctors to conduct blood sugar tests and electrocardiograms and monitor a patient's temperature and blood pressure.<sup>7</sup> In Kigali, Rwanda, digital health company Babylon is testing a service that will allow patients to video-call doctors rather than wait in lines at hospitals. In economies which suffer from large burdens on healthcare and a scarcity of trained professionals, interventions such as these can help save time, money, and lives.

The proper application of such technologies can prevent the wearer's lack of expertise from clouding data collection or transmission. Normally, doctors learn about physical symptoms along with the patient's experiences via face-to-face interactions. This adds a gap between the experience and its discussion, as well as the subjective interpretation of the patient. Its effectiveness also depends on the patient's ability to relay the information accurately. Direct recording of data can bridge these gaps ensuring that the doctor receives objective information while also understanding the patient's specific circumstances. This information transmission can be combined with additional elements including web-cameras and online voice calling software. This allows doctors and patients to remain in contact regarding diagnoses without the need for the patient to physically travel to a hospital or office. Thus, new metrics become possible, accuracy is increased, and time is saved, while costs are reduced. Additionally, such solutions may help provide proactive care in case of medical emergency.

The benefits of data analytics are not just limited to diagnosis. Data analytics also facilitates the leverage of technology to ensure that patients receive diagnoses in a timely fashion and schedule treatment and follow-up interactions as needed. Analytics already plays a key role in scheduling appointments, acquiring medicines, and ensuring that patients do not forget to take their medications.

The advantage of using big data techniques is not limited to the *transmission* of data used for diagnosis. Data analysis is key to understanding the fundamental mechanisms of the body's functions. Even in cases where the physical function of the body is well understood, big data can help researchers analyze the myriad ways in which each individual reacts to stimuli and treatment. This can lead to more customized treatment and a decrease in side effects. By analyzing specific interactions between drugs and the body, data analytics can help fine-tune dosages, reduce side effects, and adjust prescriptions on a case-by-case basis. Geno Germano,

---

<sup>6</sup><https://www.thebetterindia.com/49931/swasthya-slate-kanav-kahol-delhi-diagnostic-tests/> (accessed on Jul 16, 2018).

<sup>7</sup><https://www.economist.com/international/2017/08/24/in-poor-countries-it-is-easier-than-ever-to-see-a-medic> (accessed on Sep 15, 2018).

former group president of Pfizer's Global Innovative Pharma Business, said in 2015 that doctors might (in the near future) use data about patients' DNA in order to come up with personalized, specific treatment and health advice that could save time and ensure better outcomes.<sup>8</sup>

The ability to analyze multiple streams of related data in real time can be applied to create simulations of the human body. Using such simulations allows researchers to conduct experiments and gather information virtually, painlessly, and at low cost. In constructing artificial models of all or some parts of the body, big data techniques can harness computational power to analyze different treatments. Initiatives such as the *Virtual Physiological Human Institute*<sup>9</sup> aim to bring together diverse modeling techniques and approaches in order to gain a better, more holistic understanding of the body and in turn drive analysis and innovation.

Analytics is being used in the study and improvement of *healthcare operations* to enhance patient welfare, increase access to care, and eliminate wastes. For example, by analyzing patient wait times and behavior, data scientists can suggest policies that reduce the load on doctors, free up valuable resources, and ensure more patients get the care they need when and where they need it. Simulation techniques that predict patient traffic can help emergency rooms prepare for increased number of visitations,<sup>10</sup> while systems that track when and where patients are admitted make it easier for nurses and administrators to allocate beds to new patients.<sup>11</sup> Modern technologies can also help in the provision of follow-up care, that is, after the patient has left the hospital. Software and hardware that track important physical symptoms can notice deviation patterns and alert patients and caregivers. By matching such patterns to patient histories, they can suggest solutions and identify complications. By reminding patients regarding follow-up appointments, they can reduce rehospitalization.

Analytics is also needed to guide the use of information technologies related to updating patient records and coordinating care among providers across time or locations. Technologies like Dictaphones and digital diaries are aimed at collecting and preserving patient data in convenient ways. Careful analysis of this data is key when working to use these technologies to reduce redundant efforts and eliminate misunderstandings when care is handed from one provider to another.

There are many applications of analytics related to the *detection of health hazards and the spread of disease*: Big data methods help insurers isolate trends in illness and behavior, enabling them to better match risk premiums to an individual buyer's risk

---

<sup>8</sup><https://www.forbes.com/sites/matthewherper/2015/02/17/how-pfizer-is-using-big-data-to-power-patient-care/#7881a444ceb4> (accessed on Aug 21, 2017).

<sup>9</sup><http://www.vph-institute.org> (accessed on Sep 1, 2017).

<sup>10</sup><http://pubsonline.informs.org/doi/10.1287/msom.2015.0573> (accessed on Aug 21, 2017).

<sup>11</sup><http://www.bbc.com/news/business-25059166> (accessed on Aug 21, 2017).

profile. For instance, a US-based health insurance provider<sup>12</sup> offers Nest Protect,<sup>13</sup> a smoke alarm and carbon monoxide monitor, to its customers and also provides a discount on insurance premiums if they install these devices. Insurers use the data generated from these devices to determine the premium and also in predicting claims.

Information provider *LexisNexis* tracks socioeconomic variables<sup>14</sup> in order to predict how and when populations will fall sick. Ogi Asparouhov, the chief data scientist at LexisNexis, suggests that socioeconomic lifestyle, consumer employment, and social media data can add much value to the healthcare industry.

The use of *Google Trends*<sup>c</sup> data, that is, Internet search history, in healthcare research increased sevenfold between 2009 and 2013. This research involves a wide variety of study designs including causal analysis, new descriptive statistics, and methods of surveillance (Nutti et al. 2014). *Google Brain*,<sup>15</sup> a research project by Google, is using machine learning techniques to predict health outcomes from a patient's medical data.<sup>16</sup> The tracking of weather patterns and their connection to epidemics of flu and cold is well documented. The World Health Organization's program *Atlas of Health and Climate*<sup>17</sup> is such an example regarding the collaboration between metrological and public health communities.

By gathering diverse kinds of data and using powerful analytical tools, insurers can better predict fraud, determine appropriate courses of action, and regulate payment procedures. A comprehensive case study (Ideal Insurance) is included in Chap. 25 that describes how analytics can be used to create rules for classifying claims into those that can be settled immediately, those that need further discussion, and those that need to be investigated by an external agency.

These techniques, however, are not without their challenges. The heterogeneity of data in healthcare and privacy concerns have historically been significant stumbling blocks in the industry. Different doctors and nurses may record identical data in different ways, making analysis more difficult. Extracting data from sensors such as X-ray and ultrasound scans and MRI machines remains a continuing technical challenge, because the quality of these sensors can vary wildly.<sup>18</sup>

Big data techniques in healthcare also often rely on real-time data, which places pressure on information technology systems to deliver data quickly and reliably.

---

<sup>12</sup><https://www.ft.com/content/3273a7d4-00d2-11e6-99cb-83242733f755> (accessed on Sep 2, 2017).

<sup>13</sup><https://nest.com/smoke-co-alarm/overview> (accessed on Sep 2, 2017).

<sup>14</sup>[http://cdn2.hubspot.net/hubfs/593973/0116\\_Predictive\\_Modeling\\_News.pdf?t=1453831169463](http://cdn2.hubspot.net/hubfs/593973/0116_Predictive_Modeling_News.pdf?t=1453831169463) (accessed on Sep 3, 2017).

<sup>15</sup><https://research.google.com/teams/brain> (accessed on Sep 3, 2017).

<sup>16</sup><https://www.cnbc.com/2017/05/17/google-brain-medical-records-prediction-illness.html> (accessed on Sep 3, 2017).

<sup>17</sup><http://www.who.int/globalchange/publications/atlas/en> (accessed on Sep 2, 2017).

<sup>18</sup><https://pdfs.semanticscholar.org/61c8/fe7effa85345ae2f526039a68db7550db468.pdf> (accessed on Aug 21, 2017).

Despite these challenges, big data techniques are expected to be a key driver of technological change and innovation in this sector in the decades to come. The rest of this chapter will discuss in detail the use of data and simulation techniques in academic medical centers (AMCs) to improve patient flow.

## **2 Methods of Healthcare Analytics: Using Analytics to Improve Patient Flow in Outpatient Clinics**

Demands for increased capacity and reduced costs in outpatient settings create the need for a coherent strategy on how to collect, analyze, and use data to facilitate process improvements. Specifically, this note focuses on system performance related to patient flows in outpatient clinics in academic medical centers that schedule patients by appointments. We describe ways to map these visits as we map processes, collect data to formally describe the systems, create discrete event simulations (DESS) of these systems, use the simulations as a virtual lab to explore possible system improvements, and identify proposals as candidates for implementation. We close with a discussion of several projects in which we have used our approach to understand and improve these complex systems.

### **2.1 Introduction**

As of 2016, the Affordable Care Act (ACA) extended access to health insurance coverage to roughly 30 million previously uninsured Americans, and that coverage expansion is linked to between 15 and 26 million additional primary care visits annually (Glied and Ma 2015; Beronio et al. 2014). In addition, the number of people 65 and older in the USA is expected to grow from 43.1 million in 2012 to 83.7 million by 2050 (Ortman et al. 2014). This jump in the number of insured Americans coupled with the anticipated growth in the size of the population above the age of 65 will correlate with rising demand for healthcare services.

At the same time, Medicare and other payers are moving away from the older “fee-for-service” model toward “bundled payment” schemes (Cutler and Ghosh 2012). Under these arrangements, providers are paid a lump sum to treat a patient or population of patients. This fixes patient-related revenue and means that these payments can only be applied to fixed costs if variable costs are less than the payment. We expect the continued emergence of bundled payment schemes to accelerate the gradual move away from inpatient treatment to the delivery of care through outpatient settings that has been taking place for over 20 years. Consequently, a disproportionate share of the growth in demand will be processed through outpatient clinics, as opposed to hospital beds. This evolution is also seen as one of the key strategies needed to help get healthcare cost in the USA closer to the costs experienced in other developed countries (Lorenzoni et al. 2014).

An additional complicating factor is that healthcare delivery in the USA is often interspersed with teaching and training of the next generation of care providers. In 2007, roughly 40 million outpatient visits were made to teaching hospitals known as academic medical centers (AMCs) (Hing et al. 2010). Inclusion of the teaching component within the care process dramatically increases the complexity of each patient visit. The classic model of an outpatient visit where a nurse leads the patient to an examination room and the patient is seen by the physician and then leaves the clinic is not a sufficient description of the process in the AMC. Adding a medical resident or fellow (trainee) into the process introduces steps for interactions between the trainee and the patient as well as interactions between the trainee and the attending physician (attending). These added steps increase flow times, the number and levels of resources deployed, and system congestion (Boex et al. 2000; Franzini and Berry 1999; Hosek and Palmer 1983; Hwang et al. 2010). The delays added are easy to understand when one considers the fact that the trainee typically takes longer than the attending to complete the same task and many teaching settings demand that both the trainee and the attending spend time with each patient on the clinic schedule (Williams et al. 2007; Taylor et al. 1999; Sloan et al. 1983).

The addition of the teaching mission is not simply adding steps to a well-managed process. The added complexity is akin to changing from a single-server queueing system to a hybrid system (Williams et al. 2012, 2015). The trainee may function as a parallel (but slower) server, or the trainee and attending may function as serial servers such that a one-step activity becomes a two-step process, or decisions on how the trainee is intertwined in the process may be made dynamically, meaning that the trainee's role may change depending on system status.

In short, we are asking our current healthcare system to improve access to care to a rapidly growing and aging population as demand is shifted from inpatient to outpatient services in teaching hospitals using delivery models that are not well understood. While the extent to which this is even possible is debatable (Moses et al. 2005), it is quite clear that efforts to make this workable require thoughtful data analysis and extremely high-quality operations management (Sainfort et al. 2005).

The primary objective of this chapter is to lay out a strategy toward gaining an understanding of these complex systems, identifying means to improve their performance, and predicting how proposed changes will affect system behavior. We present this in the form of a six-step process and provide some details regarding each step. We close with a discussion of several projects in which our process has been applied.

## ***2.2 A Representative Clinic***

To make the remainder of our discussion more concrete, let us introduce a representative unit of analysis. Data associated with this unit will be taken from a composite of clinics that we have studied, but is not meant to be a complete representation of

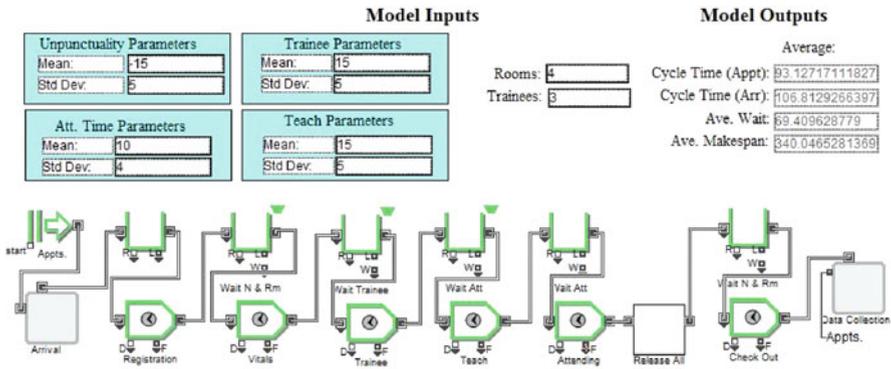


Fig. 22.1 Copied from DES of representative clinic

any particular unit. Consider a patient with an appointment to see the attending at a clinic within an AMC. We will work with a discrete event simulation (DES) of this process. DES is the approach of creating a mathematical model of the flows and activities present in a system and using this model to perform virtual experiments seeking to find ways to improve measurable performance (Benneyan 1997; Clymer 2009; Hamrock et al. 2013; Jun et al. 1999). A screenshot from such a DES is presented in Fig. 22.1 and will double as a simplified process map. By simplified, we mean that several of the blocks shown in the figure actually envelop multiple blocks that handle details of the model. Versions of this and similar models along with exercises focused on their analysis and use are **linked to this chapter**.

Note that the figure also contains a sample of model inputs and outputs from the simulation itself. We will discuss several of these metrics shortly.

In this depiction, a block creates work units (patients) according to an appointment schedule. The block labeled “Arrival” combines these appointment times with a random variable reflecting patient unpunctuality to get actual arrival times. Once created, the patients move to Step 1. Just above Step 1, we show a block serving as a queue just in case the resources at Step 1 are busy. In Step 1, the patient interacts with staff at the front desk. We will label this step “Registration” with the understanding that it may include data collection and perhaps some patient education. In Step 2, a nurse leads the patient into an examination room, collects data on vital signs, and asks a few questions about the patient’s condition. We will label this step “Vitals.” In Step 3, a trainee reviews the patient record and enters the examination room to interact with the patient. We label this step “Trainee.” In Step 4, the trainee leaves the exam room and interacts with the attending. We label this step “Teach.” During this time, the trainee may present case information to the attending, and the pair discusses next steps, possible issues, and the need for additional information. In Step 5, the trainee and attending both enter the exam room and interact with the patient. We label this step “Attending.” Following this step, the trainee, attending, and room are “released,” meaning that they are free to be

assigned to the next patient. Finally, the patient returns to the front desk for “Check Out.” This step may include collection of payment and making an appointment for a future visit.

In order to manage this system, we need an understanding of its behavior. This behavior will be reflected in quantifiable metrics such as cycle times wait times, and how long it will take to complete the appointment schedule (makespan). Note that cycle times may be calculated based on appointment times or patient arrival times. Both of these values are included among the model outputs shown here. While this model is fairly simple, some important questions may be addressed with its use. For example, we may make different assumptions regarding the attending’s processing time and note how this changes the selected output values. This is done by altering the parameters labeled “Att. Time Parameters” among the model inputs. For this illustration, we assume that these times are drawn from a log-normal distribution and the user is free to change the mean and standard deviation of that distribution. However, one benefit of simulation is that we may use a different distribution or sample directly from collected activity time data. We will discuss these issues later. This model can also be used as part of a more holistic approach to address more subtle questions, including how the added educational mission affects output metrics and what is the best appointment schedule for this system. In the next section, we lay out a more complete approach to handling more complex questions such as these.

### ***2.3 How to Fix Healthcare Processes***

Much of the early development of research-oriented universities in the USA was driven by the need for research related to healthcare (Chesney 1943). Consequently, when working with physicians and other healthcare professionals in the AMC, a convenient starting point for the discussion is already in place. Research in most parts of healthcare addresses questions using randomized trials or pilot implementations. These typically center on formal experiments which are carefully designed and conducted in clinical or laboratory settings. This experiment-based approach to research has proven to be highly effective and is assumed by many to be the best way to produce evidence-based results on medical questions relating to issues including the efficacy of a new drug or the efficiency of a new technique. One way to get buy-in from practitioners in the AMC is to take a very similar approach to issues related to patient flow.

At the same time, operations research (*OR*) has had a long history of using tools to improve service delivery processes. *OR* employs a predictive modeling investigative paradigm that uses mathematical equations, computer logic, and related tools to forecast the consequences of particular decision choices (Sainfort et al. 2005). Typically, this is done in abstraction without a formal experiment. This approach permits the consideration of alternative choices to quickly be evaluated and compared to see which are most likely to produce preferred outcomes. Many traditional areas of *OR* are prevalent in clinic management. These topics include

appointment scheduling (Cayirli et al. 2006), nurse rostering problems (Burke et al. 2004), resource allocation problems (Chao et al. 2003), capacity planning (Bowers and Mould 2005), and routing problems (Mandelbaum et al. 2012).

Given this confluence of approaches and needs, it seems natural for those working to improve healthcare processes to employ *OR* techniques such as DES to conduct controlled, virtual experiments as part of the improvement process. However, when one looks more closely, one finds that the history of implementations of results based on *OR* findings in AMCs is actually quite poor. For example, a review of over 200 papers that used DES in healthcare settings identified only four that even claimed that physician behavior was changed as a result (Wilson 1981). A more recent review found only one instance of a publication which included a documented change in clinic performance resulting from a simulation-motivated intervention (van Lent et al. 2012).

This raises a major question: Since there is clearly an active interest in using DES models to improve patient flow and there is ample talent working to make it happen, what can we do to make use of this technique in a way that results in real change in clinic performance? Virtually any operations management textbook will provide a list of factors needed to succeed in process improvement projects such as getting all stakeholders involved early, identifying a project champion, setting clear goals, dedicating necessary resources, etc. (Trusko et al. 2007). However, we want to focus this discussion on two additional elements that are a bit subtler and, in our experience, often spell the difference between success and failure when working in outpatient clinics in the AMC.

First, finding an important problem is not sufficient. It is critically important to think in terms of finding the right question which also addresses the underlying problem. As outside agents or consultants, we are not in a position to pay faculty and staff extra money to implement changes to improve the system. We need a different form of payment to motivate their participation. One great advantage in the *AMC* model is that we can leverage the fact that physicians are also dedicated researchers. Thus, we can use the promise of publications in lieu of a cash payment to induce participation.

Second, we need to find the right combination of techniques. Experiments and data collection resonate with medical researchers. However, the translation from “lab” to “clinic” is fraught with confounding factors outside of the physician’s control. On the other hand, *OR* techniques can isolate a single variable or factor, but modeling by itself does not improve a system, and mathematical presentations that feel completely abstract do not resonate with practitioners. The unique aspect of our approach is to combine *OR* tools with “clinical” experiments. This allows clinicians to project themselves into the model in a way that is more salient than the underlying equations could ever be. The key idea is that value exists in finding a way to merge the tools of *OR* with the methodologies of medical research to generate useful findings that will actually be implemented to improve clinic flow.

## 2.4 *The Process Improvement Process*

Given this background, we need a systematic approach to describing, analyzing, and predicting improvements in performance based on changes that can be made to these systems. In order to do this, we need to accomplish at least six things, which form the statement of our method:

1. Describe processes that deliver care and/or service to patients in a relevant way.
2. Collect data on activity times, work flows, and behavior of key agents.
3. Create a DES of the system under study.
4. Experiment with both real and virtual systems to identify and test possible changes.
5. Develop performance metrics of interest to both patients and care providers.
6. Predict changes in metrics which stem from changes in process.

We now turn to providing a bit more detail about each of these steps.

### **Step 1: Process Description**

Much has been written concerning process mapping in healthcare settings (Trusko et al. 2007; Trebble et al. 2010). In many instances, the activity of process mapping itself suggests multiple changes that may improve process flow. However, some insights related to the healthcare-specific complications of this activity warrant discussion.

Perhaps the most obvious way to develop a process map is to first ask the agents in the system to describe the work flow. We have found that this is absolutely necessary and serves as an excellent starting point but is never sufficient. Agents in the system often provide misleading descriptions of process flow. In many cases, physicians are not fully aware of what support staff do to make the process work, and trainees and staff are often quite careful to not appear to contradict more senior physicians. To get high-quality process descriptions, we must gather unbiased insights from multiple levels of the organization. Ideally, this will include support staff, nurses, trainees, and attendings. In some cases, other administrators are valuable as well, especially if there is a department manager or some other person who routinely collects and reports performance data. It is ideal to have all of these agents working on the development of a process map as a group. However, if this cannot be done, it is even more vital to carefully gather information about process flows from as many different angles as possible.

Second, we have found that no matter how much information about the process has been gathered, direct observation by outside agents working on the process improvement process is *always* required. We have yet to find a process description created by internal agents that completely agrees with our observations. Healthcare professionals (understandably) put patient care above all other considerations. Consequently, they make exceptions to normal process flows routinely without giving it a second thought. As a result, their daily behavior will almost always include several subtleties that they do not recall when asked about process flow.

### ***Step 2: Data Collection***

In our experience, this is the most time-consuming step in the improvement process. Given a process map, it will be populated with some number of activities undertaken by various agents. The main question that must be asked at this stage is how long each agent spends to complete each step. This approach makes sense for several reasons: First, the dominant patient complaint in outpatient settings is wait times. Thus, time is a crucial metric from the patient's perspective. Second, many systems have been developed which accumulate costs based on hourly or minute-by-minute charges for various resources (Kaplan and Anderson 2003; Kaplan and Porter 2011; King et al. 1994). Consequently, time is a crucial metric from the process manager's perspective as well. Therefore, how long each step takes becomes the central question of interest.

We have utilized four ways to uncover this information. First, agents within the system can be asked how long a process step takes. This is useful as a starting point and can be sufficient in some rare instances. On the other hand, quizzing agents about activity times is problematic because most people think in terms of averages and find it difficult to measure variances. This can only be done after a sufficient number of observations are in hand.

We have also used a second approach in which the caregivers record times during patient visits. For example, in one clinic, we attached a form to each patient record retrieved during each clinic session. In Step 1, staff at the front desk record the patient arrival time and appointment time. The nurse then records the start and end times of Step 2 and so on. This approach can be automated through the use of aids such as phones or iPad apps, where applicable. However, this approach introduces several issues. Recording data interrupts normal flow, and it is not possible to convince the participants that data recording compares in importance to patient care. As a consequence, we repeatedly see instances where the providers forget to record the data and then try to "fill it in" later in the day when things are less hectic. This produces data sets where mean times may be reasonable estimates, but the estimates of variances are simply not reliable.

A third approach to data collection often used in AMCs is to use paid observers to record time stamps. This approach can generate highly reliable information as long as the process is not overly complex and the observer can be physically positioned to have lines of sight that make this method practical. This approach is common in AMCs because they are almost always connected to a larger university and relatively high-quality, low-cost labor is available in the form of students or volunteers. While we have used this technique successfully on multiple occasions, it is not without its problems. First, the observers need to be unobtrusive. This is best done by having them assigned to specific spaces. If personnel travel widely, this becomes problematic. For example, a radiation oncology clinic that we studied had rooms and equipment on multiple floors, so tracking became quite complex. Second, the parties serving patients know they are being observed. Many researchers have reported significant improvements to process flow using this approach, only to find that after the observers left, the system drifted back to its previous way of functioning and the documented improvement was lost.

We have also used a fourth approach to data collection. Many hospitals and clinics are equipped with real-time location systems (RTLS). Large AMC's are often designed to include this capability because tracking devices and equipment across hundreds of thousands of square feet of floor space are simply not practical without some technological assistance. Installations of these systems typically involve placing sensors in the ceilings or floors of the relevant spaces. These sensors pick up signals from transmitters that can be embedded within "tags" or "badges" worn by items or people being tracked. Each sensor records when a tag comes within range and again when it leaves that area. When unique tag numbers are given to each caregiver, detailed reports can be generated at the end of each day showing when a person or piece of equipment moved from one location to another. This approach offers several dramatic advantages. It does not interfere with the care delivery process, the marginal cost of using it is virtually 0, and since these systems are always running, the observation periods can begin and end as needed.

In closing, we should highlight three key factors in the data collection process: (1) data collection needs to be done in a way that does not interfere with care delivery; (2) audits of the data collection system are needed to ensure accuracy; and (3) sufficient time span must be covered to eliminate any effects of the "novelty" of the data collection and its subsequent impact on agent behaviors.

### **Step 3: Create a DES of the System**

We have often found it useful to create DES models of the systems under study as early in the process as possible. This can be a costly process in that a great deal of data collection is required and model construction can be a nontrivial expense. Other tools such as process mapping and queueing theory can be applied with much less effort (Kolker 2010). However, we have repeatedly found that these tools are insufficient for the analysis that is needed. Because the variances involved in activity times can be extremely high in healthcare, distributions of the metrics of interest are important findings. Consequently, basic process analysis is rarely sufficient and often misleading.

Queueing models do a much better job of conveying the significance of variability. However, many common assumptions of these models are routinely violated in clinic settings, including that some processing times are not exponentially distributed, that processing times are often not from the same distribution, and that if arrivals are based on appointments, inter-arrival times are not exponentially distributed.

However, none of these issues pose the largest challenge to applying simple process analysis or queueing models in outpatient clinics. Consider two additional issues. First, the basic results of process analysis or queueing models are only averages which appear in steady state. A clinic does not start the day in steady state—it begins in an empty state. It takes some time to reach steady state. However, if one plots average wait times for a clinic over time, one quickly sees that it may take dozens or even hundreds of cases for the system to reach steady state. Clearly, a clinic with one physician is not going to schedule hundreds of patients for that resource in a single session. Thus, steady-state results are often not informative.

Second, if activity times and/or the logic defining work flow changes in response to job type or system status, then the results of simple process analysis or queueing models become invalid. We have documented such practices in multiple clinics that we have studied (Chambers et al. 2016; Conley et al. 2016). Consequently, what is needed is a tool that can account for all of these factors simultaneously, make predictions about what happens when some element of the system changes, and give us information about the broader distribution of outcomes—not just a means for systems in steady state. DES is a tool with the needed capabilities.

A brief comment on the inclusion of activity times in DES models is warranted here. We have used two distinct approaches. We can select an activity time at random from a collection of observations. Alternatively, we can fit a distribution to collected activity time data. We have found both approaches to work satisfactorily. However, if the data set is sufficiently large, we recommend sampling directly from that set. This generates results that are both easier to defend to statisticians and more credible to practitioners.

#### ***Step 4: Field and Virtual Experiments***

It is at this point that the use of experiments comes into play, and we merge the *OR* methodology of DES with the experimental methods of medical research. The underlying logic is that we propose an experiment involving some process change that we believe will alter one or more parameters defining system behavior. We can use the DES to predict outcomes if our proposal works. In other cases, if we have evidence that the proposed change works in some settings, we can use the DES to describe how that change will affect system metrics in other settings. The construction of these experiments is the “art” of our approach. It is this creation that leads to publishable results and creates novel insights.

We will provide examples of specific experiments in the next section. However, at this juncture we wish to raise two critical issues: confounding variables and unintended consequences. Confounding variables refer to system or behavioral attributes that are not completely controlled when conducting an experiment but can alter study results. For example, consider looking at a system before an intervention, collecting data on its performance, changing something about the system, and then collecting data on the performance of the modified system. This is the ideal approach, but it implicitly assumes that nothing changed in the system over the span of the study other than what you intended to change. If data collection takes place over a period of months, it is quite possible that the appointment schedule changed over that span of time due to rising or falling demand. In this example, the change in demand would be a confounding variable. It is critically important to eliminate as many confounding variables as you can before concluding that your process change fully explains system improvement. DES offers many advantages in this regard because it allows you to fix some parameter levels in a model even if they may have changed in the field.

It is also critical to account for unintended consequences. For example, adding examination rooms is often touted as a way to cut wait times. However, this also makes the relevant space larger, increasing travel times as well as the complexity of

resource flows. This must be accounted for before declaring that the added rooms actually improved performance. It may improve performance along one dimension while degrading it in another.

DES modeling has repeatedly proven invaluable at this stage. Once a DES model is created, it is easy to simulate a large number of clinic sessions and collect data on a broad range of performance metrics. With a little more effort, it can also be set up to collect data on the use of overtimes or wait times within examination rooms. In addition, DES models can be set up to have patients take different paths or have activity times drawn from different distributions depending on system status. Finally, we have found it useful to have DES models collect data on subgroups of patients based on system status because many changes to system parameters affect different groups differently.

#### ***Step 5: Metrics of Interest***

A famous adage asserts, “If you can’t measure it, you can’t manage it.” Hence, focusing on measurements removes ambiguity and limits misunderstandings. If all parties agree on a metric, then it is easier for them to share ideas on how to improve it. However, this begs an important question—what metrics do we want to focus on? In dealing with this question, Steps 4 and 5 of our method become intertwined and cannot be thought of in a purely sequential fashion. In some settings, we need novel metrics to fit an experiment, while in other settings unanticipated outcomes from experiments suggest metrics that we had not considered earlier.

Both patients and providers are concerned with system performance, but their differing perspectives create complex trade-offs. For example, researchers have often found that increase in face time with providers serves to enhance patient experience (Thomas et al. 1997; Seals et al. 2005; Lin et al. 2001), but an increase in wait time degrades that experience (Meza 1998; McCarthy et al. 2000; Lee et al. 2005). The patient may not fully understand what the care provider is doing, but they can always understand that more attention is preferable and waiting for it is not productive. Given a fixed level of resources, increases in face time result in higher provider utilization, which in turn increases patient wait times. Consequently, the patient’s desire for increased face time and reduced wait time creates a natural tension and suggests that the metrics of interest will almost always include both face time and wait time.

Consider one patient that we observed recently. This patient arrived 30 min early for an appointment and waited 20 min before being lead to the exam room. After being led to the room, the patient waited for 5 min before being seen by a nurse for 5 min. The patient then waited 15 min before being seen by the resident. The trainee then spoke with the patient for 20 min before leaving the room to discuss the case with the attending. The patient then waited 15 min before being seen by the resident and the attending together. The attending spoke with the patient for 5 min before being called away to deal with an issue for a different patient. This took 10 min. The attending then returned to the exam room and spoke with the patient for another 5 min. After that, the patient left. By summing these durations, we see that the patient was in the clinic for roughly 100 min. The patient waited for 20 min in the waiting room. However, the patient also spent 45 min in the exam

room waiting for service. Time in the examination room was roughly 80 min of which 35 min was spent in the presence of a service provider. Thus, we can say that the overall face time was only 35 min. However, of this time only 10 min was with the attending physician. Consideration of this more complete description suggests a plethora of little-used metrics that may be of interest, such as:

1. Patient punctuality
2. Time spent in the waiting room before the appointment time
3. Time spent in the waiting room after the appointment time
4. Wait time in the examination room
5. Proportion of cycle time spent with a care provider
6. Proportion of cycle time spent with the attending

The key message here is that the metrics of interest may be specific to the problem that one seeks to address and must reflect the nuances of the process in place to deliver the services involved.

#### ***Step 6: Predict Impact of Process Changes***

Even after conducting an experiment in one setting, we have found that it is extremely difficult to predict how changes will affect a different system simply by looking at the process map. This is another area where DES proves quite valuable. For example, say that our experiment in Clinic A shows that by changing the process in some way, the time for the Attending step is cut by 10%. We can then model this change in a different clinic setting by using a DES of that setting to predict how implementing our suggested change will be reflected in performance metrics of that clinic in the future. This approach has proven vital to get the buy-in needed to facilitate a more formal experiment in the new setting or to motivate implementation in a unit where no formal experiment takes place.

## ***2.5 Experiments, Simulations, and Results***

Our work has included a collection of experiments that have led to system improvements for settings such as that depicted in Fig. 22.1. We now turn to a discussion of a few of these efforts to provide context and illustrations of our approach. Figure 22.1 includes an arrival process under an appointment system. This is quickly followed by activities involving the trainee and/or nurse and/or attending. Finally, the system hopes to account for all of these things when searching for an optimized schedule. We discuss a few of these issues in turn.

### ***Arrival Process***

We are focusing on clinics which set a definite appointment schedule. One obvious complication is that some patients are no-shows, meaning that they do not show up for the appointment. No-show rates of as much as 40% have been cited in prior works (McCarthy et al. 2000; Huang 1994). However, there is also a subtler issue of patients arriving very early or very late, and this is much harder to account for. Early work in this space referred to this as patient “unpunctuality” (Bandura

1969; White and Pike 1964; Alexopoulos et al. 2008; Fetter and Thompson 1966; Tai and Williams 2012; Perros and Frier 1996). Our approach has been used to address two interrelated questions: Does patient unpunctuality affect clinic performance, and can we affect patient unpunctuality? To address these questions, we conducted a simple experiment. Data on patient unpunctuality was collected over a six-month period. We found that most patients arrived early, but patient unpunctuality ranged from  $-80$  to  $+20$ . In other words, some patients arrived as much as 80 min early, while others arrived 20 min late. An intervention was performed that consisted of three elements. In reminders mailed to each patient before their visit, it was stated that late patients would be asked to reschedule. All patients were called in the days before the visit, and the same reminder was repeated over the phone. Finally, a sign explaining the new policy was posted near the registration desk. Unpunctuality was then tracked 1, 6, and 12 months later. Additional metrics of interest were wait times, use of overtime, and the proportion of patients that were forced to wait to be seen (Williams et al. 2014).

This lengthy follow-up was deemed necessary because some patients only visited the clinic once per quarter, and thus the full effect of the intervention could not be measured until after several quarters of implementation. To ensure that changes in clinic performance were related only to changes in unpunctuality, we needed a way to control for changes in the appointment schedule that happened over that time span. Our response to this problem was to create a DES of the clinic, use actual activity times in the DES, and consider old versus new distributions of patient unpunctuality, assuming a fixed schedule. This allowed us to isolate the impact of our intervention.

Before the intervention, 7.7% of patients were tardy and average tardiness of those patients was 16.75 min. After 12 months, these figures dropped to 1.5% and 2 min, respectively. The percentage of patients who arrived before their appointment time rose from 90.4% to 95.4%. The proportion who arrived at least 1 min tardy dropped from 7.69% to 1.5%. The range of unpunctuality decreased from 100 to 58 min. The average time to complete the session dropped from 250.61 to 244.49 min. Thus, about 6 min of overtime operations was eliminated from each session. The likelihood of completing the session on time rose from 21.8% to 31.8%.

Our use of DES allowed us to create metrics of performance that had not yet been explored. For example, we noticed that the benefits from the change were not the same for all patients. Patients that arrived late saw their average wait time drop from 10.7 to 0.9 min. Those that arrived slightly early saw their average wait time increase by about 0.9 min. Finally, for those that arrived very early, their wait time was unaffected. In short, we found that patient unpunctuality can be affected, and it does alter clinic performance, but this has both intended and unintended consequences. The clinic session is more likely to finish on time and overtime costs are reduced. However, much of the benefit in terms of wait times is actually realized by patients that still insist on arriving late.

### ***Physician Processing Times***

Historically, almost all research on outpatient clinics assumed that processing times were not related to the schedule or whether the clinic was running on time. Is

this indeed the case? To address this question, we analyzed data from three clinic settings. One was a low-volume clinic that housed a single physician, another was a medium-volume clinic in an AMC that had one attending working on each shift along with two or three trainees, and the last one was a high-volume service that had multiple attendings working simultaneously (Chambers et al. 2016).

We categorized patients into three groups: Group A patients were those who arrived early and were placed in the examination room before their scheduled appointment time. Group B patients were those who also arrived early, but were placed in the examination room after their appointment time, indicating that the clinic was congested. Group C patients were those who arrived after their appointment time. The primary question was whether the average processing time for patients in Group A was the same as that for patients in Group B. We also had questions about how this affected clinic performance in terms of wait times and session completion times.

In the low-volume clinic with a single physician, average processing times and standard errors (in parentheses) were 38.31 (3.21) for Group A and 26.23 (2.23) for Group B. In other words, the physician moved faster when the clinic was behind schedule. Similar results have been found in other industries, but this was the first time (to the best of our knowledge) that this had been demonstrated for outpatient clinics.

In the medium-volume clinic, the relevant values were 65.59 (2.24) and 53.53 (1.97). Again, the system worked faster for Group B than it did for Group A. Note the drop in average times is about 12 min in both settings. This suggests that the finding is robust, meaning that it occurs to a similar extent in similar (but not identical) settings. Additionally, remember that the medium-volume clinic included trainees in the process flow. This suggests that the way that the system got this increase in speed might be different. In fact, our data show that the average amount of time the attending spent with the patient was no more than 12 min to begin with. Thus, we know that it was not just the behavior of the attending that made this happen. The AMC must be using the trainees differently when things fall behind schedule.

In the high-volume clinic, the parallel values were 47.15 (0.81) and 17.59 (0.16). Here, we see that the drop in processing times is much more dramatic than we saw before. Again, the message is that processing times change when the system is under stress and the magnitude of the change implies that multiple parties are involved in making this happen. In hindsight, this seems totally reasonable, but the extent of the difference is still quite startling.

As we saw in the previous section, there is an unintended consequence of this system behavior as it relates to patient groups. Patients that show up early should help the clinic stay on schedule. This may not be so because these patients receive longer processing times. Thus, their cycle times are longer. Patients that arrive late have shorter wait times and shorter processing times. Thus, their cycle times are shorter. If shorter cycle times are perceived as a benefit, this seems like an unfair reward for patient tardiness and may explain why it will never completely disappear.

### ***Impact of the Teaching Mission***

The result from the previous section suggests that the way that the trainee is used and managed within the clinic makes a difference when considering system performance. To explore this point further, we wanted to compare a clinic without trainees with a similar clinic that included trainees. This is difficult to do as an experiment, but we were lucky when looking at this question. An attending from a clinic with no trainees was hired as the director of a clinic in the AMC that included trainees. Thus, we could consider the same attending seeing the same patients in both settings. One confounding variable was that the two clinics used different appointment schedules (Williams et al. 2012).

We collected data on activity times in both settings. Given these times, we could seed DES models of both clinics and compare results. Within the DES, we could look at both settings as though they had the same appointment schedule. If we consider the two settings using the schedule in place for the AMC, we see that the average cycle time in the AMC was 76.2 min and this included an average wait time of 30.0 min. The average time needed to complete a full schedule was 291.9 min. If the same schedule had been used in the private practice model, the average cycle time would be 129.1 min and the average wait time would be 83.9 min.

The capacity of the AMC was clearly greater than it was in the private practice model. This is interesting because the flow times in the private practice setting using the schedule that was optimized for that setting were much lower. It turns out that the total processing time for each patient was greater in the AMC, but the capacity was higher. This is explained using parallel processing. In the AMC setting, the attending spent time with one patient, while trainees simultaneously worked with other patients. We were able to conduct a virtual experiment by changing the number of trainees in the DES model. We found that having one trainee created a system with cycle times that were much greater than the private practice model. Using two trainees produced cycle times that were about the same. Using three trainees created the reduced cycle times that we noticed in practice. Using more than three trainees produced no additional benefit because both clinics had only three available exam rooms. This enabled us to comment on the optimal number of trainees for a given clinic.

The use of DES also highlighted a less obvious result. It turns out that the wait time in this system was particularly sensitive to the time taken in the step we labeled “Teach.” This is the time that the trainee spends interacting with the attending after interacting with the patient. In fact, we found that reducing this time by 1 min served to reduce average wait time by 3 min. To understand this phenomenon, recall that when the trainee and the attending are discussing the case while the patient waits in the examination room for 1 min, the three busiest resources in the system (the trainee, the attending, and the examination room) are simultaneously occupied for that length of time. Thus, it is not surprising that wait times are sensitive to the duration of this activity, although the degree of this sensitivity is still eye-opening.

### ***Preprocessing***

Given that wait times are extremely sensitive to teaching times, we created an experiment designed to alter the distribution of these times. Instead of having the

trainee review the case after the patient is placed in the examination room and then having the first conversation about the case with the attending after the trainee interacts with the patient, we can notify both the trainee and attending in advance which patient each trainee will see. That way, the trainee can review the file before the session starts and have a conversation with the attending about what should happen upon patient arrival. We also created a template to guide the flow and content of this conversation. We refer to this approach as “preprocessing” (Williams et al. 2015).

We recorded activity times using the original system for 90 days. We then introduced the new approach and ran it for 30 days. During this time, we continued collecting data on activity times.

Before the intervention was made, the average teach time was 12.9 min for new patients and 8.8 min for return patients. The new approach reduced these times by 3.9 min for new patients and 2.9 min for return patients. Holding the schedule as a constant, we find that average wait times drop from 36.1 to 21.4 min and the session completion time drops from 275.6 to 247.4 min.

However, in this instance, it was the unintended consequences that proved to be more important. When the trainees had a more clearly defined plan about how to handle each case, their interactions with the patients became more efficient. The trainees also reported that they felt more confident when treating the patients than they had before. While it is difficult to measure this effect in terms of time, both the trainees and the attending felt that the patients received better care under the new protocol.

### *Cyclic Scheduling*

Considering the works mentioned above, one finding that occurred repeatedly was that the way the trainee was involved in the process had a large impact on system performance and how that was done was often state dependent. Recall that we found that the system finds ways to move faster when the clinic is behind schedule. When a physician is working alone, this can be done simply by providing less face time to patients. When the system includes a trainee, an additional response is available in that either the attending or the trainee can be dropped from the process for one or more patients. Our experience is that doctors strongly believe that the first approach produces huge savings and they strongly oppose the second.

Our direct observation of multiple clinics produced some insights related to these issues. Omitting the attending does not save as much time as most attendings think because the trainee is slower than the attending. In addition, the attending gets involved in more of these cases than they seem to realize. Many attendings feel compelled to “at least say hi” to the patients even when the patients are not really on their schedule, and these visits often turn out to be longer than expected. Regarding the second approach, we have noticed a huge variance in terms of how willing the attending is to omit the trainee from a case. Some almost never do it, while others do it quite often. In one clinic we studied, we found that the trainee was omitted from roughly 30% of the cases on the clinic schedule. If this is done, it might explain why a medium-volume or high-volume clinic within the AMC could reduce cycle times after falling behind schedule to a greater extent than the low-volume clinic

can achieve. This can be done by instructing the trainee to handle one case while the attending handles another and having the attending exclude the trainee from one or more cases in an effort to catch up to the clinic schedule.

Accounting for these issues when creating an appointment schedule led us to the notion of cyclic scheduling. The idea is that the appointment schedule can be split into multiple subsets which repeat. We label these subsets “cycles.” In each cycle, we include one new patient and one return patient scheduled to arrive at the same time. A third patient is scheduled to arrive about the middle of the cycle. If both patients arrive at the start of the cycle, we let the trainee start work on the new patient, and the attending handles the return patient without the trainee being involved. This was deemed acceptable because it was argued that most of the learning comes from visits with new patients. If only one of the two patients arrives, the standard process is used.

Process analysis tools produce some results about average cycle times in this setting, but since wait times are serially correlated, we want a much clearer depiction of how each patient’s wait time is related to that of the following patients. Considering the problem using a queuing model is extremely difficult because the relevant distribution of activity times is state dependent and the number of cycles is small. Consequently, steady-state results are misleading. Studying this approach within a DES revealed that average makespan, wait times, and cycle times are significantly reduced using our cyclic approach and the trainee is involved in a greater proportion of the cases scheduled.

### 3 Conclusion

While a great deal of time, effort, and money has been spent to improve healthcare processes, the problems involved have proven to be very difficult to solve. In this work, we focused on a small but important sector of the problem space—that of appointment-based clinics in academic medical centers. One source of difficulty is that the medical field favors an experimental design-based approach, while many *OR* tools are more mathematical and abstract. Consequently, one of our core messages is that those working to improve these systems need to find ways to bridge this gap by combining techniques. When this is done, progress can be made and the insights generated can be spread more broadly. Our use of DES builds on tools of process mapping that most managers are familiar with and facilitates virtual experiments that are easier to control and use to generate quantitative metrics amenable to the kinds of statistical tests that research physicians routinely apply.

However, we would be remiss if we failed to emphasize the fact that data-driven approaches are rarely sufficient to bring about the desired change. Hospitals in AMCs are often highly politicized environments with a hierarchical culture. This fact can generate multiple roadblocks that no amount of “number crunching” will ever overcome. One not so subtle aspect of our method is that it typically involves embedding ourselves in the process over some periods of time and interacting

repeatedly with the parties involved. We have initiated many projects not mentioned above because they did not result in real action. Every project that has been successful involved many hours of working with faculty, physicians, staff, and technicians of various types to collect information and get new perspectives. We have seen dozens of researchers perform much more impressive data analysis on huge data sets using tools that were more powerful than those employed in these examples, only to end up with wonderful analysis not linked to any implementation. When dealing with healthcare professionals, we are often reminded of the old adage, “No one cares how much you know. They want to know how much you care.” While we believe that the methodology outlined in this chapter is useful, our experience strongly suggests that the secret ingredient to making these projects work is the attention paid to the physicians, faculty, and especially staff involved who ultimately make the system work.

## Electronic Supplementary Material

All the datasets, code, and other material referred in this section are available in [www.allaboutanalytics.net](http://www.allaboutanalytics.net).

- Model 22.1: Model1.mox
- Model 22.2: Model1A.mox
- Model 22.3: Model2.mox
- Model 22.4: Model3.mox

## Exercises

In “Using Analytics to Improve Patient Flow in Outpatient Clinics,” we laid out a six-step approach to improving appointment-based systems in outpatient academic medical centers. These exercises involve simplified versions of discrete event simulations (DESs) of such settings. Their purpose is to illustrate and conceptualize the process. Completion of these exercises should highlight many issues and subtleties of these systems and help the reader develop ideas that best fit with their setting of interest.

### *Introduction*

Simplified versions of several models referenced in Sect. 22.2.5 of the reading can be considered to explore the issues discussed there. These DES models have been developed in ExtendSim version 9.0.<sup>19</sup> Complete versions of the underlying software are available from the vendor, and a variety of academic pricing models are

---

<sup>19</sup>Download trial version from <https://www.extendsim.com/demo> (accessed on Jul 16, 2018).

available. A wide variety of texts and tools are also available to assist the potential user with details of software capabilities including Strickland (2010) and Laguna and Marklund (2005). However, the models utilized in this reading are fairly simple to construct and can be easily adapted to other packages as the reader (or instructor) sees fit. For ease of exposition and fit with the main body of the reading, we present exercises corresponding to settings described earlier. Hints are provided in the Hints for Solution word file (refer to book's website) that should help in going through the exercises. The exercises allow the reader to explore the many ideas given in the chapter in a step-by-step manner.

### ***A Basic Model with Patient Unpunctuality***

Service providers in many settings utilize an appointment system to manage the arrival of customers/jobs. However, the assumption that the appointment schedule will be strictly followed is rarely justified. The first model (Model 1; refer to book's website) presents a simplified process flow for a hypothetical clinic and embeds an appointment schedule. The model facilitates changes to the random variable that defines patient punctuality. In short, patients arrive at some time offset from their appointment time. By adjusting the parameters which define the distribution of this variable, we can represent arrival behavior. You may alter this model to address the following questions:

**Ex. 22.1** Describe clinic performance if all patients arrive on time.

**Ex. 22.2** Explain how this performance changes if unpunctuality is included. For this example, this means modeling actual arrival time as the appointment time plus a log-normally distributed variable with a mean of  $\mu$  and a standard deviation of  $\sigma$  minutes. A reasonable base case may include  $\mu = -15$  min, and  $\sigma = 10$  min. (Negative values of unpunctuality mean that the patient arrives prior to the appointment time, which is the norm.) Note how changes to  $\mu$  and  $\sigma$  affect performance differently.

**Ex. 22.3** Explain how you would create an experiment (in an actual clinic) to uncover how this behavior changes and how it affects clinic performance.

**Ex. 22.4** Explain how you would alter Model 1 to report results for groups of patients such as those with negative unpunctuality (early arrivers), those with positive unpunctuality (late arrivers), and those with appointment times near the end of the clinic session.

**Ex. 22.5** The DES assumes that the patient with the earlier appointment time is always seen first, even if they arrived late. How would you modify this model if the system "waits" for late patients up to some limit, " $w$ " minutes rather than seeing the next patient as soon as the server is free?

### ***An Academic Model with Distributions of Teaching Time***

The process flow within the academic medical center (AMC) differs from Model 1 in that it includes additional steps and resources made necessary by the hospital's teaching mission. Simple process analysis is useful in these settings to help identify the bottleneck resource and to use management of that resource to improve system performance. However, such simple models are unable to fully account for the impact of system congestion given this more complex flow. For example, idle time is often added because one resource is forced to wait for the availability of another.

Using a DES of such systems may be particularly valuable in that they facilitate various forms of sensitivity analysis which can produce novel insights about these issues. Use Model 2 (refer to book's website) of the AMC to address the following questions:

**Ex. 22.6** How do the average values of cycle time, wait time, and makespan respond to changes in teach time?

**Ex. 22.7** Describe the linkage between utilization of the trainees in this system and the amount of time they spend with patients. How much of their busy time is not explained by value-adding tasks?

**Ex. 22.8** Describe the linkage between the number of trainees and the utilization of other key resources in the system.

**Ex. 22.9** Explain how you would create an experiment (in an actual clinic) to uncover how changing the educational process is linked to resident productivity.

**Ex. 22.10** How would you alter Model 2 to reflect a new approach to trainee education aimed at increasing the share of their time that adds value to the patient?

#### ***State-Dependent Processing Times***

Experience with many service systems lends support to the notion that the service provider may be motivated to “speed up” when the system is busy. However, common sense also suggests that this is not sustainable forever. With these facts in mind, it is important to think through how we might measure this behavior and how we may monitor any unintended consequences from such an approach. With this in mind, Model 3 (refer to book's website) includes a reduction to processing times for the attending when the system is busy. Consider this model to address the following questions:

**Ex. 22.11** How do average values of cycle time, wait time, and makespan change when the attending gets faster in a busy system?

**Ex. 22.12** Instead of reducing face time, consider adding examination rooms to the system instead. Is there any evidence produced by the DES to suggest that one approach is better than the other?

**Ex. 22.13** Describe the comparison between decreasing processing times when the system is busy to changing processing times for all settings.

**Ex. 22.14** Explain how you would create an experiment (in an actual clinic) to explore how this behavior affects patient flow and service quality. What extra factors do you need to control for?

**Ex. 22.15** How would you alter Model 3 to separate the effects of patient behavior (including unpunctuality) from the effects of physician behavior (including changing processing times)?

#### ***Cyclic Scheduling***

Personnel creating an appointment schedule are likely to favor having a simple template to refer to when patients request appointment times. Consequently, there is administrative value in having a logic that is easy to explain and implement. Again, this is more difficult to do in the AMC since the process flow is more complex. Return to the use of Model 2 and modify it as needed to address the following questions:

**Ex. 22.16** Study the existing appointment schedule. Develop the “best” schedule if there is no variability to consider. (You may assume that average activity times are always realized.)

**Ex. 22.17** How does your schedule perform when patient unpunctuality is added, and how will you adjust your schedule to account for this?

**Ex. 22.18** Assuming that patients are always perfectly punctual and only attending time is variable, look for a schedule that works better than the one developed in Exercise 22.16.

**Ex. 22.19** Explain how you would create an experiment (in an actual clinic) to explore ways to reduce this variability. What extra factors do you need to control for?

**Ex. 22.20** How would you alter Model 2 to include additional issues such as patient no-shows, emergencies, work interruptions, and open-access scheduling?

### Conclusion

It is important to note that DES models are only one tool that can be applied to develop a deeper understanding of the behavior of complex systems. However, adding this approach to the “toolbox” of the clinic manager or consultant should provide ample benefits and support for ideas on how to make these systems better meet the needs of all stakeholders.

## References

- Alexopoulos, C., Goldman, D., Fontanesi, J., Kopald, D., & Wilson, J. R. (2008). Modeling patient arrivals in community clinics. *Omega*, 36, 33–43.
- Bandura, A. (1969). *Principles of behavior modification*. New York, NY: Holt, Rinehart, & Winston.
- Benneyan, J. C. (1997). An introduction to using computer simulation in healthcare: Patient wait case study. *Journal of the Society for Health Systems*, 5(3), 1–15.
- Beronio, K., Glied, S. & Frank, R. (2014) J Behav Health Serv Res. 41, 410. <https://doi.org/10.1007/s11414-014-9412-0>
- Boex, J. R., Boll, A. A., Franzini, L., Hogan, A., Irby, D., Meservey, P. M., Rubin, R. M., Seifer, S. D., & Veloski, J. J. (2000). Measuring the costs of primary care education in the ambulatory setting. *Academic Medicine*, 75(5), 419–425.
- Bowers, J., & Mould, G. (2005). Ambulatory care and orthopaedic capacity planning. *Health Care Management Science*, 8(1), 41–47.
- Burke, E. K., De Causmaecker, P., Berghe, G. V., & Van Landeghem, H. (2004). The state of the art of nurse rostering. *Journal of Scheduling*, 7(6), 441–499.
- Cayirli, T., Veral, E., & Rosen, H. (2006). Designing appointment scheduling systems for ambulatory care services. *Health Care Management Science*, 9(1), 47–58.
- Chambers, C. G., Dada, M., Elnahal, S. M., Terezakis, S. A., DeWeese, T. L., Herman, J. M., & Williams, K. A. (2016). Changes to physician processing times in response to clinic congestion and patient punctuality: A retrospective study. *BMJ Open*, 6(10), e011730.
- Chao, X., Liu, L., & Zheng, S. (2003). Resource allocation in multisite service systems with intersite customer flows. *Management Science*, 49(12), 1739–1752.
- Chesney, A. M. (1943). *The Johns Hopkins Hospital and John Hopkins University School of Medicine: A chronicle*. Baltimore, MD: Johns Hopkins University Press.

- Clymer, J. R. (2009). *Simulation-based engineering of complex systems* (Vol. 65). New York, NY: John Wiley & Sons.
- Conley, K., Chambers, C., Elnahal, S., Choflet, A., Williams, K., DeWeese, T., Herman, J., & Dada, M. (2018). *Using a real-time location system to measure patient flow in a radiation oncology outpatient clinic*. *Practical radiation oncology*.
- Cutler, D. M., & Ghosh, K. (2012). The potential for cost savings through bundled episode payments. *New England Journal of Medicine*, *366*(12), 1075–1077.
- Fetter, R. B., & Thompson, J. D. (1966). Patients' wait time and doctors' idle time in the outpatient setting. *Health Services Research*, *1*(1), 66.
- Franzini, L., & Berry, J. M. (1999). A cost-construction model to assess the total cost of an anesthesiology residency program. *The Journal of the American Society of Anesthesiologists*, *90*(1), 257–268.
- Glied, S., & Ma, S. (2015). *How will the Affordable Care Act affect the use of health care services?* New York, NY: Commonwealth Fund.
- Hamrock, E., Parks, J., Scheulen, J., & Bradbury, F. J. (2013). Discrete event simulation for healthcare organizations: A tool for decision making. *Journal of Healthcare Management*, *58*(2), 110.
- Hing, E., Hall, M. J., Ashman, J. J., & Xu, J. (2010). National hospital ambulatory medical care survey: 2007 Outpatient department summary. *National Health Statistics Reports*, *28*, 1–32.
- Hosek, J. R., & Palmer, A. R. (1983). Teaching and hospital costs: The case of radiology. *Journal of Health Economics*, *2*(1), 29–46.
- Huang, X. M. (1994). Patient attitude towards waiting in an outpatient clinic and its applications. *Health Services Management Research*, *7*(1), 2–8.
- Hwang, C. S., Wichterman, K. A., & Alfrey, E. J. (2010). The cost of resident education. *Journal of Surgical Research*, *163*(1), 18–23.
- Jun, J. B., Jacobson, S. H., & Swisher, J. R. (1999). Application of discrete-event simulation in health care clinics: A survey. *Journal of the Operational Research Society*, *50*(2), 109–123.
- Kaplan, R. S., & Anderson, S. R. (2003). *Time-driven activity-based costing*. SSRN 485443.
- Kaplan, R. S., & Porter, M. E. (2011). How to solve the cost crisis in health care. *Harvard Business Review*, *89*(9), 46–52.
- King, M., Lapsley, I., Mitchell, F., & Moyes, J. (1994). Costing needs and practices in a changing environment: The potential for ABC in the NHS. *Financial Accountability & Management*, *10*(2), 143–160.
- Kolker, A. (2010). Queuing theory and discrete event simulation for healthcare: From basic processes to complex systems with interdependencies. In Abu-Taieh, E., & El Sheik, A. (Eds.), *Handbook of research on discrete event simulation technologies and applications* (pp. 443–483). Hershey, PA: IGI Global.
- Laguna, M., & Marklund, J. (2005). *Business process modeling, simulation and design*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Lee, V. J., Earnest, A., Chen, M. I., & Krishnan, B. (2005). Predictors of failed attendances in a multi-specialty outpatient centre using electronic databases. *BMC Health Services Research*, *5*(1), 1.
- van Lent, W. A. M., VanBerkel, P., & van Harten, W. H. (2012). A review on the relation between simulation and improvement in hospitals. *BMC Medical Informatics and Decision Making*, *12*(1), 1.
- Lin, C. T., Albertson, G. A., Schilling, L. M., Cyran, E. M., Anderson, S. N., Ware, L., & Anderson, R. J. (2001). Is patients' perception of time spent with the physician a determinant of ambulatory patient satisfaction? *Archives of Internal Medicine*, *161*(11), 1437–1442.
- Lorenzoni, L., Belloni, A., & Sassi, F. (2014). Health-care expenditure and health policy in the USA versus other high-spending OECD countries. *The Lancet*, *384*(9937), 83–92.
- Mandelbaum, A., Momcilovic, P., & Tseytlin, Y. (2012). On fair routing from emergency departments to hospital wards: QED queues with heterogeneous servers. *Management Science*, *58*(7), 1273–1291.

- McCarthy, K., McGee, H. M., & O'Boyle, C. A. (2000). Outpatient clinic wait times and non-attendance as indicators of quality. *Psychology, Health & Medicine*, 5(3), 287–293.
- Meza, J. P. (1998). Patient wait times in a physician's office. *The American Journal of Managed Care*, 4(5), 703–712.
- Moses, H., Thier, S. O., & Matheson, D. H. M. (2005). Why have academic medical centers survived. *Journal of the American Medical Association*, 293(12), 1495–1500.
- Nuti, S. V., Wayda, B., Ranasinghe, I., Wang, S., Dreyer, R. P., Chen, S. I., & Murugiah, K. (2014). The use of Google trends in health care research: A systematic review. *PLoS One*, 9(10), e109583.
- Ortman, J. M., Velkoff, V. A., & Hogan, H. (2014). *An aging nation: The older population in the United States* (pp. 25–1140). Washington, DC: US Census Bureau.
- Perros, P., & Frier, B. M. (1996). An audit of wait times in the diabetic outpatient clinic: Role of patients' punctuality and level of medical staffing. *Diabetic Medicine*, 13(7), 669–673.
- Sainfort, F., Blake, J., Gupta, D., & Rardin, R. L. (2005). *Operations research for health care delivery systems. WTEC panel report*. Baltimore, MD: World Technology Evaluation Center, Inc..
- Seals, B., Feddock, C. A., Griffith, C. H., Wilson, J. F., Jessup, M. L., & Kesavalu, S. R. (2005). Does more time spent with the physician lessen parent clinic dissatisfaction due to long wait times. *Journal of Investigative Medicine*, 53(1), S324–S324.
- Sloan, F. A., Feldman, R. D., & Steinwald, A. B. (1983). Effects of teaching on hospital costs. *Journal of Health Economics*, 2(1), 1–28.
- Strickland, J. S. (2010). *Discrete event simulation using ExtendSim 8*. Colorado Springs, CO: Simulation Educators.
- Tai, G., & Williams, P. (2012). Optimization of scheduling patient appointments in clinics using a novel modelling technique of patient arrival. *Computer Methods and Programs in Biomedicine*, 108(2), 467–476.
- Taylor, D. H., Whellan, D. J., & Sloan, F. A. (1999). Effects of admission to a teaching hospital on the cost and quality of care for Medicare beneficiaries. *New England Journal of Medicine*, 340(4), 293–299.
- Thomas, S., Glynn-Jones, R., & Chait, I. (1997). Is it worth the wait? a survey of patients' satisfaction with an oncology outpatient clinic. *European Journal of Cancer Care*, 6(1), 50–58.
- Trebble, T. M., Hansi, J., Hides, T., Smith, M. A., & Baker, M. (2010). Process mapping the patient journey through health care: An introduction. *British Medical Journal*, 341(7769), 394–397.
- Trusko, B. E., Pexton, C., Harrington, H. J., & Gupta, P. (2007). *Improving healthcare quality and cost with six sigma*. Upper Saddle River, NJ: Financial Times Press.
- White, M. J. B., & Pike, M. C. (1964). Appointment systems in out-patients' clinics and the effect of patients' unpunctuality. *Medical Care*, 133–145.
- Williams, J. R., Matthews, M. C., & Hassan, M. (2007). Cost differences between academic and nonacademic hospitals: A case study of surgical procedures. *Hospital Topics*, 85(1), 3–10.
- Williams, K. A., Chambers, C. G., Dada, M., Hough, D., Aron, R., & Ulatowski, J. A. (2012). Using process analysis to assess the impact of medical education on the delivery of pain services: A natural experiment. *The Journal of the American Society of Anesthesiologists*, 116(4), 931–939.
- Williams, K. A., Chambers, C. G., Dada, M., McLeod, J. C., & Ulatowski, J. A. (2014). Patient punctuality and clinic performance: Observations from an academic-based private practice pain centre: A prospective quality improvement study. *BMJ Open*, 4(5), e004679.
- Williams, K. A., Chambers, C. G., Dada, M., Christo, P. J., Hough, D., Aron, R., & Ulatowski, J. A. (2015). Applying JIT principles to resident education to reduce patient delays: A pilot study in an academic medical center pain clinic. *Pain Medicine*, 16(2), 312–318.
- Wilson, J. C. T. (1981). Implementation of computer simulation projects in health care. *Journal of the Operational Research Society*, 32(9), 825–832.