Chris here. Remember the TV show *Emergency!*? Before there was a *St. Elsewhere*, *ER*, or *Grey's Anatomy*, there was *Emergency!* That show was the best! Every week for 52 minutes on NBC, paramedics John Gage and Roy DeSoto got down and dirty in the streets of Los Angeles, saving lives and dishing out excitement as LA County Fire Department Squad 51. The paramedics were the stars of the show, but I couldn’t wait until they pulled into Rampart General Hospital to deliver their patients. That’s when the magic happened. In the emergency department (ED), Drs. Kelly Brackett and Joe Early stood ready in their long white coats, the picture of competence and authority, to take on whatever came their way. Rampart General! For me, that was as good as it got. That’s where I wanted to be.

Cut to 1999. I began my residency training in emergency medicine at Harbor-UCLA Medical Center, a busy county hospital with an amazing staff and some of the sickest patients you could imagine. And guess what—there’s a scene in the opening credits of *Emergency!* when Squad 51 pulls into the ambulance bay of Rampart General, lights flashing and sirens blaring, that was filmed at Harbor. My hospital! I was an emergency physician at Rampart General, my childhood fantasy having literally come true!

While the outside of the hospital looked like Rampart, the inside was more like County General from the show *ER*. Our ED was chaotic; people had gunshot wounds and heart attacks, and we raced from patient to patient, trying to keep up. In my three years there, I saw the floors in the waiting room being cleaned only once—there were always too many people to get in there with a mop. During my last year of residency, some of the writers from *ER* came to hang out with us and get story ideas (one of mine even ended up on season four).

We were so busy, and the hospital was so overcrowded, that people might have to wait eight or more hours to get seen. We usually got to the sick patients sooner, but not always. I remember one occasion when a frantic triage nurse called me into the crowded waiting room. People were on their feet, yelling, and the throng...
of patients had parted to create a path to a young-looking guy, maybe 25 years old, lying motionless on the ground. Another nurse was standing over the man as I reached him, a step ahead of my two co-residents. I felt for a pulse—there wasn’t one—and we lifted his body up onto a gurney that had appeared. I hopped onto the gurney and started CPR on the patient as the other residents rushed us back through the waiting room to the ED. We worked on him for another hour, trying to get his heart to start back up. He didn’t make it.

I’ve held a crystal-clear memory of that day ever since. At the time, it seemed like an amazing and heroic effort, riding on top of this dying patient through a crowd of onlookers, pumping on his chest to resuscitate him. It was like a television scene out of ER, with me—a younger, better-looking George Clooney—trying to save this guy’s life. Now, after decades of experience and perspective, I look back at that day with sadness and regret. This wasn’t a patient who died an unpreventable death in his home. He died in our waiting room—our house—with doctors and nurses all around. Our house was too crowded, and we failed him. At the time, we didn’t question if we could have done better; it wasn’t on our radar then. I had the best medical training available, but I never considered patient flow a necessary part of that training.

Cut to a few years ago. At another hospital several miles from where I now work, a patient walked into the ED waiting room with a head injury. The waiting room was busy. This patient didn’t seem to be too sick or in much pain, so he sat down to wait his turn. And then he died. He had an epidural hemorrhage; the blood inside his skull expanded and put too much pressure on his brain. This is a rare occurrence, and it is possible to be fooled by the reassuring appearance of a patient who is, in fact, really sick. However, there should never be a bad outcome because the hospital was too crowded to get to the patient in time. It’s unthinkable that someone would come to the hospital with a heart attack and die because we didn’t know what to do. Yet we know for a fact that hospital overcrowding can harm patients (as we will soon discuss), so why is it okay that we don’t know what to do about poor patient flow? No one should ever be hurt because of overcrowding if the tools exist to manage it. This is why patient flow is important to me, and it’s why we wrote this book.

WHY WE NEED ANOTHER BOOK ABOUT PATIENT FLOW

Dozens of books and countless articles have been published on hospital overcrowding and patient flow. Healthcare leaders have been learning about Lean and Six Sigma for decades, and process improvement methodology is part of healthcare
master's programs across the country. Despite these efforts, overcrowding remains endemic in healthcare. Poor patient flow is a root cause of overcrowding, so it’s not clear how effective those dozens of books and countless articles have been. This book is different. It lays out a methodology that has existed for more than 30 years but has not been widely applied to healthcare—more on this later.

**WHY WE SHOULD CARE ABOUT PATIENT FLOW**

In 2006, the federal government asked the Institute of Medicine (IOM) to study overcrowding in the emergency medical system. In response, the IOM published three reports describing an overburdened emergency healthcare system that was approaching its limits (IOM 2007a, 2007b, 2007c). In the years since, the situation has only worsened. ED visits have skyrocketed, while the number of available hospital beds has diminished (US Department of Health, Education, and Welfare 1978; Rui and Kang 2014; National Center for Health Statistics 2016).

Hospital overcrowding is important because it has deleterious effects on almost every aspect of healthcare. Hospital overcrowding results in ED overcrowding, which delays treatment for emergency conditions. Patients may take longer to receive pain medication, antibiotics, and imaging studies, and they may have longer hospital stays once they are admitted (Mills et al. 2010; Pines et al. 2010; Singer et al. 2011; George and Evridiki 2015). ED overcrowding also results in more patients leaving the ED against medical advice or without being examined. Consequently, ED overcrowding results in increased morbidity and mortality (Bernstein et al. 2009; George and Evridiki 2015; Singer et al. 2011).

Overcrowding also comes at an enormous financial cost. When a hospital operates at or above capacity, its ED is often forced to close to new ambulance traffic. For every hour that an ED is closed to ambulance traffic, the hospital loses thousands of dollars in potential revenue from ED visits and admissions—ambulances will now take that revenue to another hospital. Hospital overcrowding is frequently caused by stays that exceed the norm for admitted patients as compared to national benchmarks. Most of the hospital’s revenue comes from the first few days of a patient’s stay. If a patient is benchmarked to stay three days, for example, but is not discharged until the fifth day, the hospital receives relatively little revenue for those last two days of the stay. Although reimbursement for the final two days decreases, the hospital’s costs do not. Moreover, the hospital is missing out on the opportunity to put a new patient in the bed and realize higher reimbursements.

The direct financial costs of overcrowding pale in comparison to the indirect costs. Overcrowding poses a significant risk of malpractice, given the higher rates
of morbidity and mortality as well as the risks associated with patients leaving against medical advice or before being evaluated. Patients equate long wait times with poor customer service, and bad word of mouth can drive business away.

Hospital overcrowding results in poor patient outcomes, morbidity, and mortality; is a major dissatisfier for both patients and staff; and costs millions of dollars. Overcrowding is bad for business and bad for care.

POOR PATIENT FLOW AND HOSPITAL OVERCROWDING: TWO SIDES OF THE SAME COIN

Patient flow is the most important variable affecting hospital overcrowding. Throughout this book, we use the terms “hospital overcrowding” and “poor patient flow” synonymously. Flow is so vital to patient care that in 2012, the Joint Commission established a patient flow standard (LD.04.03.11), which describes nine performance elements a hospital must meet to ensure compliance. Patient flow is also included in the Centers for Medicare & Medicaid Services quality performance measures. Although healthcare administrators, patient care providers, and hospital staff members understand the importance of patient flow, few healthcare systems have taken significant and enduring steps toward flow optimization.

Hospitals and healthcare systems continue to struggle with patient flow because it's difficult to fix. Improving patient flow requires a wholesale change in staff culture and a clear mandate from the highest levels of administration—not easy to achieve, especially when the mechanics of patient flow are poorly understood.

Overcrowding is typically most visible in the ED, because the ED is the only place in the hospital that cannot close its doors when it gets too full. When other units in the hospital are busy, admitted patients who are waiting for beds on inpatient units back up in the ED. These patients take up bed space in the ED for hours or even days, forcing new ED patients to spill over into hallway gurneys and chairs, or to spend long hours in the waiting room. A stroll through most hospital inpatient wards will show only as many patients as there are available beds, but a casual observer in a busy ED will see patients everywhere: hallways, chairs, and makeshift annexes. Because hospital overcrowding is most visible in the ED, it is often considered the ED’s problem to fix. However, as anyone who works in the ED will tell you, ED overcrowding is more often a symptom of a broken hospital system. The entire hospital must therefore be engaged in flow to address ED overcrowding in a meaningful way.

Why is overcrowding such a difficult problem to fix? Many books and conferences offer solutions, but their advice doesn't always translate into successful, actionable, lasting change. Bright, capable people have not been able to
definitively solve the overcrowding issue. It might be because we don’t know how; or we think we know how, but we don’t.

In the past, I worked in a hospital system that was committed to using Lean. Leadership created an entire innovation department, staffed with engineers who were Lean experts. They held educational seminars for hospital employees and scheduled kaizen (improvement) workshops for departments throughout the system. The people who received this training were smart and hardworking, but five years later, there wasn’t much to show for their efforts.

The hospital system missed out on opportunities for enduring improvement because it failed to implement effectively. The hospital did not have an overarching strategy, so while the Lean experts skillfully managed individual projects, they were underutilized and deployed randomly. Lean was used for one-off projects based on the biggest fires that had to be put out, which manager hollered the loudest, or who put in a request for resources first. Responding in this fashion rarely yielded long-term, sustainable results. Although the “Lean way” did create some positive gains, its methods were never institutionalized. When the innovation department moved on to other projects, new ways reverted back to old ones.

My medical director at the time experienced the implementation gap firsthand when he was invited to a kaizen on ED flow. My director, Michael Anthony (“Tony” to his friends, and not to be confused with the bassist for Van Halen), trained on the East Coast and brought a blunt, matter-of-fact sensibility with him to the Pacific Northwest. Tony had a finely tuned BS detector, and it started ping-
ing almost immediately on the first day of the kaizen.

Essentially, **kaizen** is a Japanese term for improvement. A kaizen workshop is a Lean tool for continuous improvement. The workshop is typically held over five days and seeks to identify and implement rapid, incremental changes in a specific area of a more complex system. Along with Tony, other participants in the kaizen included the ED nurse manager, the director of emergency and critical care services, the director of social work, the trauma nurse manager, and the Lean expert who was going to help the group “fix” patient flow in the ED.

Unfortunately, the problem with the ED flow had little to do with the ED. The real problem was on the inpatient side, where there were not enough beds to accommodate admitted patients, who then overflowed into the ED. The ED could stop receiving ambulances, but it couldn’t close its doors to walk-ins, who made up the majority of ED patients. Tony’s ED was overcrowded because the boarded patients left no room for new ED patients. Fixing the ED flow required reengineering inpatient flow, and that was beyond the scope of the kaizen.

My favorite part of this story is when the Lean engineer analyzed the ED using queuing theory and recommended we triple our physician coverage. Tony heard...
this and apparently went ballistic. He knew that physicians often sat idle because there was nowhere to see new patients. Adding physicians wouldn’t create new spaces to see patients; it would only increase provider costs and decrease provider satisfaction.

The group made a few recommendations and implemented them successfully, but the recommendations didn’t meaningfully affect overcrowding. As Tony said later, the group nibbled at the edges instead of taking a big bite out of the real problem.

It’s all right if a kaizen produces only incremental changes, but the changes need to be embedded in the culture or they will not stick. If positive changes are not institutionalized, people will revert back to their previous way of doing things. Kaizen is a continuous process of ongoing improvement (much like the Theory of Constraints, as we will discuss throughout this book). It isn’t effective to make a change and then stop, move on to something else, and assume any gains made will continue indefinitely.

A one-size-fits-all approach to patient flow typically doesn’t work. The causes of poor patient flow differ from hospital to hospital, and solutions also need to be specific to each hospital. People publish books, write articles, and hold conferences touting turnkey solutions to your problems—without knowing what your problems are. What healthcare needs is a methodology for identifying specific problems and solutions for individual hospitals. Tackling flow is like the adage “If you give a man a fish, you feed him for a day, but if you teach a man to fish, you feed him for a lifetime.” There are many resources that give you a fish, but they don’t teach you to fish. Lean, for example, is a popular, thoroughly studied, and widely implemented methodology that has existed for decades. Yet in Japan, the birthplace of Lean and the Toyota Production System, only 20 percent of manufacturers have implemented Lean because it doesn’t make sense in every setting (Goldratt 2009). One size does not fit all.

To improve patient flow, you must first understand patient flow. Understanding patient flow requires applying a methodology to your system to identify where it is broken. It’s well and good to learn how to set up an observation unit or talk about moving boarders to inpatient hallways (both popular topics in the literature and in flow courses), but these strategies aren’t helpful unless observation units and hallway boarders are solutions relevant to your particular problems. To see improvements, you must invest time to understand the problem. Instead of jumping into projects, you must exercise patience.

A good starting point would be a methodology that provides real, lasting change that can be implemented on a manageable scale and expanded as people gain experience with flow.
OUR SOLUTION: THE THEORY OF CONSTRAINTS!

In 1984, Israeli physicist Eliyahu Goldratt wrote a book called *The Goal* (Goldratt and Cox 2016) that introduced the concept of the Theory of Constraints (TOC). Simply put, his theory posits that the productivity of any system is limited at a point in time by a bottleneck resource. Identifying, optimizing, and breaking the bottleneck improves the performance of the entire system.

At this point I want to introduce a concept, then ask you to forget it until much later in the book. This idea is important enough to bring up now but too complicated to discuss thoroughly until we have covered the basics:

You can maximize a system's performance by using a particular resource, called the *constraining resource*, to synchronize the activity of the entire system—much like a conductor synchronizes the activity of all the musicians in an orchestra.

The constraining resource may or may not be the bottleneck resource. This is a nuanced but important point, and we will return to the constraining resource later in the book. For now, though, we'll concentrate on bottlenecks.

TOC is a process of ongoing improvement, because there will always be a bottleneck in place in any system. Once the bottleneck is broken, another will take its place, which must in turn be identified and broken. *The Goal* was written specifically about a manufacturing plant, but the concepts introduced in the book are widely applicable to all business domains—including healthcare.

I read *The Goal* over a decade ago, and it is no exaggeration to say that the book changed my life! Serendipitously, around the time I completed *The Goal*, I also became the director of our hospital's flow committee. Our hospital was in bad shape—bad enough to make me the director of flow. Our inpatient lengths of stay were too long (several patients had been in the hospital for over a year), and our ED was closed to ambulance traffic an average of 60 hours every month. Up to that point, my experience in process improvement had consisted solely of reading *The Goal*, but it was one more book than anyone else had read on flow, so (in my own mind, at least) I was the most qualified to lead our hospital's flow team.

Within months of applying TOC to our hospital's system, we had virtually eliminated ambulance diversion. Our rates of ED patients who left without being seen fell sharply, and our inpatient length of stay dropped to levels below the national benchmark. We accomplished these achievements during a devastating influenza season and a period of record ED patient volumes, ambulance traffic, and patient admissions. We no longer had to close to ambulance traffic, which put less pressure on the EDs at other hospitals. Our efforts helped reduce the diversion rates at almost every other hospital in the region. Our department became so
productive that we were able to add full-time ED nursing positions for the first time in years. Even though we were busier, the work felt much easier. Staff members were happier, employee retention was at an all-time high, and we improved the hospital’s bottom line by several million dollars in the first year alone. Most important, we sustained the improvements for years. Since that time, we have introduced TOC to many of our other hospitals.

A QUICK WORD ABOUT LEAN AND SIX SIGMA

Lean and Six Sigma are additional methodologies used to improve operations throughout a system. Similar to TOC, they both have their origins in manufacturing but have been applied widely to the healthcare sector. Each has its own unique strengths and challenges, and both share some similarities with TOC. Lean and Six Sigma are both better-known entities in the healthcare domain, perhaps owing to earlier adoption or first-mover advantage—or maybe they had better publicists.

Each methodology can be characterized in the simplest terms as follows.

<table>
<thead>
<tr>
<th>Methodology</th>
<th>Focus</th>
<th>Waste</th>
<th>Variability</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lean</td>
<td>Focus</td>
<td>Waste</td>
<td></td>
</tr>
<tr>
<td>Six Sigma</td>
<td>Waste</td>
<td></td>
<td>Variability</td>
</tr>
</tbody>
</table>

Each methodology can offer improvements if implemented correctly. More recently, however, practitioners have found tremendous operational advantages from combining all three methodologies. This makes sense, as one method can act synergistically with the others.

For example, Lean emphasizes waste elimination and empowerment of frontline workers, making it a powerful tool for gaining efficiencies, leveraging insights, and increasing employee participation. However, the emphasis on waste elimination may make a system anemic, with too much focus on cost reduction and not enough on increased output. Moreover, Lean does not offer an obvious starting point for system-wide flow improvement; there is no prioritization in the search for waste. Lean does not emphasize synchronization of resources across a system, nor does it necessarily break down silos.

Six Sigma seeks to reduce variability in a system, thereby reducing defects and bad outcomes. Although Six Sigma’s emphasis on quality can result in better products or improved outcomes, it may result in only modest improvements to patient flow. Six Sigma also lacks a clear prioritization mechanism.

One of the greatest advantages of TOC is that it looks at the entire system and answers two key questions: “Where do we start?” and “Which resource should we
concentrate on?” With its emphasis on focus, TOC makes it easier to determine how and where to get started and, hence, gets faster results!

Synthesizing the three methodologies offers the greatest advantage to hospital systems. We recommend first using TOC to observe the system and determine where to begin and on which critical resource to focus. Once TOC has answered these questions, you can incorporate properties of Lean (eliminating waste, but only as it affects the bottleneck) and Six Sigma (primarily to reduce variability in usage of the bottleneck resource) to realize maximum gains in efficiency.

USING THIS BOOK

The purpose of this book is to teach you how to use TOC to improve flow in your hospital, medical office, urgent care center, or clinic, but anyone committed to improving patient flow and decreasing hospital overcrowding should also read The Goal: It has been part of the core curriculum in business schools and corporate offices for 30 years. The book you are currently reading outlines the principles described in The Goal and applies those principles to healthcare. We discuss

• how TOC differs from Lean;
• how to identify bottlenecks in your own hospitals and clinics; and
• how to exploit and break bottlenecks, improve patient flow, enhance care, and strengthen the bottom line.

Throughout the book, we use common bottlenecks to illustrate certain concepts. Although these examples may not be the bottlenecks in your organization, the methodology still applies to your system. Finally, we explore constraint management (remember that thing I told you to forget?) through use cases—specifically involving ED flow, inpatient flow, and outpatient settings—to illustrate fully synchronized systems. As you read through each chapter, we encourage you to think about how the principles discussed relate to your hospital or clinic setting.

SUMMARY

• Hospital overcrowding has detrimental effects on both the cost and quality of patient care.
• Overcrowding is endemic in our healthcare delivery system.
• Methodologies exist to improve flow, but they are often applied incorrectly or incompletely and have not always been particularly effective.
• Three of the most well-established methodologies are TOC, Lean, and Six Sigma.
• The emphasis of TOC is focus, the emphasis of Lean is waste, and the emphasis of Six Sigma is variability.
• Aspects of each methodology can be combined to achieve synergistic improvements in patient flow.

REFERENCES


