Matching Supply and Demand for Hospital Services

Diwakar Gupta
University of Minnesota
Industrial & Systems Engineering
guptad@umn.edu

Sandra J. Potthoff
University of Minnesota
School of Public Health
potth001@umn.edu
Editorial Scope

Topics

Foundations and Trends® in Technology, Information and Operations Management publishes survey and tutorial articles in the following topics:

- Applications of technology, information and operations
- B2B commerce
- Business process engineering and design
- Business process outsourcing
- Capacity planning
- Competitive operations
- Contracting in supply chains
- E-Commerce and e-business models
- Electronic markets, auctions, and exchanges
- Enterprise management systems
- Facility location
- Information chain structure and competition
- International operations
- Marketing/Manufacturing interfaces
- Multi-location inventory theory
- New product & service design
- Queuing networks
- Reverse logistics
- Service logistics and product support
- Supply chain management
- Technology management and strategy

Information for Librarians

Foundations and Trends® in Technology, Information and Operations Management, 2014, Volume 8, 4 issues. ISSN paper version 1571-9545. ISSN online version 1571-9553. Also available as a combined paper and online subscription.
Matching Supply and Demand for Hospital Services

Diwakar Gupta
University of Minnesota
Industrial & Systems Engineering
guptad@umn.edu

Sandra J. Potthoff
University of Minnesota
School of Public Health
potth001@umn.edu
Contents

1 Introduction .............................................. 3
  1.1 The US Context ............................................. 7
  1.2 Hospital Service Types ................................. 9
  1.3 Demand Variability ...................................... 10
  1.4 Supply Challenges ...................................... 13
  1.5 Themes and Organization ............................... 15

2 Operating Room Capacity Management .................... 19
  2.1 Introduction .............................................. 19
  2.2 Institutional Background ............................... 21
  2.3 Evidence from Data ...................................... 23
  2.4 Operational Challenges ................................. 29
  2.5 State of Practice ......................................... 34
  2.6 Current OM Approaches ................................. 38
  2.7 OM Opportunities ....................................... 55

3 ED Capacity Management .................................... 58
  3.1 Introduction .............................................. 58
  3.2 Institutional Background ............................... 59
  3.3 Evidence from Data ...................................... 65
  3.4 Operational Challenges ................................. 65

Full text available at: http://dx.doi.org/10.1561/0200000036
3.5 State of Practice ............................................... 69
3.6 Current OM Approaches ................................. 75
3.7 OM Opportunities .......................................... 88

4 Staffed-Bed Capacity Management .................. 90
  4.1 Introduction ................................................. 90
  4.2 Institutional Background .............................. 91
  4.3 Evidence from Data ...................................... 93
  4.4 Operational Challenges ............................... 97
  4.5 State of Practice ......................................... 100
  4.6 Current OM Approaches .............................. 110
  4.7 OM Opportunities ...................................... 118

5 Concluding Remarks ......................................... 120

References ......................................................... 124
Abstract

Hospitals are nearly synonymous with the idea of a healthcare delivery system. They provide three main types of services: surgical services, emergency services, and inpatient services. In addition, some hospitals have outpatient clinics and facilities in which specialist consults and surgical services are provided on an outpatient basis. What challenges do hospital managers face in matching supply and demand for hospital services while maintaining service quality and keeping costs low? To what extent and how has the Operations Management (OM) literature contributed to addressing these challenges? What are emerging trends in practice? What opportunities and additional challenges do they pose for the OM researchers?

This monograph is the authors’ attempt to address the questions posed above. Although service capacity could very well be defined in terms of the numerical count of physical and human resources, we have chosen to focus on the three main types of services that hospitals provide. In doing so, we expose the interconnectedness of these services and the challenges that arise due to the cascading effects of mismatches in any one area on all other hospital operations. Our goal is to expose key issues from practitioner perspectives, use representative data to highlight problems that are amenable to modeling using operations management tools, summarize state of the art in modeling such problems, and identify opportunities for future research.

The monograph underscores several important observations. First, hospital administrators need to consider forces affecting demand and supply for services both inside and outside the hospital walls. Specifically, the option to shape and smooth demand for services, which is often underutilized, may provide significant new opportunities to lower the cost of matching demand and supply. Second, hospitals need both careful advance planning, based on patterns observed in historical data, as well as dynamic response strategies to unfolding reality that forces inevitable deviations from plans. Effective dynamic response may require administrators to invest in building alliances with other hospitals and flexibility (e.g., contingent staff). Third, the role of hospitals is changing. Innovations in payment mechanisms that bundle payments...
to hospitals and doctors, and offer incentives for lowering costs, are creating the need to design and implement effective gainsharing plans. These same forces have also increased the importance for hospital administrators of choosing the right number and specialization of salaried physicians, and building alliances with both upstream (e.g. primary care clinics) and downstream (e.g. skilled nursing facilities) service providers. OM researchers have addressed some of these topics, but significant new opportunities abound.

---

Healthcare expenditures constitute an ever increasing percentage of gross domestic products (GDPs) of the 34 OECD\textsuperscript{1} countries about which such data are publicly available (OECD 2015). Ten countries spent between 10-11.8\% of their respective GDPs on healthcare in 2012. The United States (US) was an outlier at 16.9\%, or nearly 2.8 trillion dollars (The Henry J Kaiser Family Foundation 2014). Furthermore, inpatient care, the focus of this monograph, comprises nearly 30\% of healthcare expenditures (OECD 2013).

The supply of hospital beds in OECD countries has decreased steadily over the past two decades. Reasons for this decline vary across countries (see, e.g. McKee 2004 and Roberge et al. 2010), but a primary reason has been to control hospital expenditures (Kroneman and Siegers 2004). For example, the average number of acute-care hospital beds per 100,000 population has dropped from 410 beds in 2000, to 330 beds in 2012, and the average hospital length of stay has decreased from 8.4 days to 7.7 days. Over approximately the same time period, hospital discharges have increased from 15,160 per 100,000 population in 2000 to 15,800 in 2010 (OECD 2015). The average inpatient bed

\textsuperscript{1}The Organization for Economic Co-operation and Development.
occupancy rate is approximately 80% (OECD 2013). These patterns, i.e. higher rates of bed utilization coupled with shorter lengths of stay, have served to increase the importance of matching demand and supply for inpatient services.

Figures 1.1 and 1.2 summarize data on acute hospital bed supply, total hospital discharges, lengths of stay, and acute care bed utilization for five selected OECD countries (OECD 2013, 2015). The top panel in Figure 1.1 shows decreasing bed supply. The bottom panel in the same figure shows either stable or increasing discharges (demand) except in Canada. The top panel in Figure 1.2 shows stable or mixed patterns of acute care bed utilization, and the bottom panel shows declining average lengths of stay.

The statistics reported in Figures 1.1 and 1.2 are important because they show that decreased bed supply, shorter lengths of stay, and increased demand have affected bed occupancy rates differently in different countries. For example, although the United Kingdom (UK) and the US have similar number of acute care beds and discharges per capita, the average lengths of stay are approximately two days longer in the UK. These differences are consistent with the calculated bed occupancy rate of 85% in the UK versus 65% in the US. The statistics may leave one with the impression that there may be an oversupply of hospital beds in the US. However, these numbers must be interpreted carefully because the OECD calculation is based on the number of beds that a US hospital is licensed to staff. On any given day, hospitals adjust staffing levels to match demand for beds. Because US hospitals typically staff fewer beds than licensed, the staffed-bed utilization is usually much higher. In fact, it may be similar in the US to that of other OECD countries. These observations underscore the importance of understanding operational features even for interpreting aggregate data.

What is the appropriate number of beds per 100,000 population? This question has been the topic of debate in many countries. Some have expressed concern that the supply of beds has fallen too much, resulting in shortages (Cunningham and Sammut 2012). Others argue

\footnote{Separate acute-care hospital discharge data were not available to the authors.}
Figure 1.1: Hospital Bed Supply and Discharges for Selected OECD Countries
Figure 1.2: Hospital Lengths of Stay and Bed Utilization for Selected OECD Countries
1.1. The US Context

that the number of beds is adequate and that improved bed management is needed (Allder et al. 2010). Regardless of the perspective one takes, it is important to understand, shape, and respond to the demand for acute care services. A key challenge for hospital managers comes from the variability in the demand and supply for such services.

This monograph examines challenges and innovations associated with managing hospital capacity related to emergency departments, operating rooms, and inpatient units. In the Operations Management (OM) literature, many such issues are examined using mathematical models. The monograph also provides a practice-relevant way of thinking about supply-demand matching, modeling approaches, and unsolved problems.

In this chapter we first briefly describe the macro US context of reimbursement strategies being used by payers to reduce demand for hospital services, and the concomitant provider responses. We then turn to the hospital level matching of demand and supply. We describe the typical services provided in acute-care hospitals, followed by the sources of demand variability, and hospitals’ challenges in shaping and responding to it. We conclude this chapter with the themes and organization of this monograph.

1.1 The US Context

The supply of US hospital beds varies significantly across geographical regions (Dartmouth College 2012). Discharge rates per 1,000 Medicare enrollees also vary more than two-fold. One of the goals of payment innovations is to reduce or eliminate unnecessary and duplicate services. Beginning in the mid-1980’s, Medicare reimbursement for hospitalization changed from a cost-plus method to a Diagnosis-Related-Group (DRG) episode-payment method. This decreased hospital bed supply, shortened the average length of hospital stay, and increased the use of post-acute transitional care (Fischer and Eustis 1989; Kosecoff et al. 1990; Qian et al. 2011). More recently, Medicare has implemented pay-

---

3Medicare is a US government healthcare insurance program for eligible residents who are at least 65 years old – see https://www.ssa.gov/medicare for details.
Introduction

ment penalties for hospitals whose 30-day re-admission rate for patients with certain conditions exceeds a threshold (Medicare Payment Advisory Committee 2014).

Medicare and private insurers are also piloting new reimbursement mechanisms to further decrease both inpatient and outpatient demand. These reimbursement methods pay for value, not production volume (Ryan and Press 2014). Pay-for-performance, risk-adjusted global payments, and bundled payments are examples of reimbursement innovations (Conrad et al. 2013) designed to promote integrated care, and align clinical and financial incentives across all providers involved in a patient’s episode of care.

Many of the strategies being developed by healthcare delivery systems in response to reimbursement changes lie outside of the hospital walls. These strategies focus on prevention and wellness in the outpatient setting to reduce the incidence rate of ED services and hospitalizations. One example is the patient-centered medical home, where interprofessional physician-led teams provide care to panels of patients (Nutting et al. 2011; Schoen et al. 2011). Another example is case management and care coordination, where a mid-level provider proactively works with patients and gets them to appropriate outpatient care at the right time (Lion et al. 2014; Mathews 2014; Sutherland and Hayter 2009). New models of physician compensation that focus on patient outcomes, such as certain biomarkers, and not just the number of patient encounters, are also being implemented (Bailit et al. 2015; Latham and Marshall 2015).

A challenge is that until all insurers switch to an outcome-based payment system, many health systems face conflicting reimbursement strategies. For many systems, only a small portion of their reimbursement stream is value based; most is still fee-for-service based. Therefore, when they implement strategies such as case management, they may penalize their bottom line (lower fees) and at the same time incur higher production costs.
1.2 Hospital Service Types

Most acute-care hospitals provide at least three major service types. These are emergency care (via emergency departments or EDs), surgical services (via operating rooms or ORs), and inpatient care (via inpatient beds) – see schematic in Figure 1.3. In this monograph, a separate chapter is devoted to each service type: Chapter 2 to ORs, Chapter 3 to EDs, and Chapter 4 to inpatient units.
Hospital beds are divided into inpatient units, each specializing in a different type of care. Examples include intensive care units (ICU), step-down units, general-care units, surgical units, obstetric units, and neonatal units. Some hospitals further specialize care by medical condition. For example, there may exist intensive care units or step-down units that focus on cardiac patients or patients with neurological conditions. Similarly, different operating rooms may be fitted with different specialized equipment for different types of surgeries. Finally, EDs differ from one hospital to another in terms of the type of specialized care they can provide (e.g., some hospitals are designated regional trauma centers) and the availability of diagnostic equipment.

In the remainder of the discussion in this chapter, we focus on hospital beds to explain institutional features that make demand-supply matching challenging. Similar issues also arise for other service types. As shown in Figure 1.3, a patient may be admitted to an inpatient bed either after visiting the ED, or after a surgery, or following a physician’s direct-admit request to the hospital. During their hospital sojourn, patients are usually moved several times — either to a lower care-intensity unit if their condition improves, or to a higher care-intensity unit if their condition deteriorates. The movement of patients from EDs, ORs, or direct admission through one or more inpatient units is often referred to as the patient flow and its study has attracted much attention in the OM literature.

1.3 Demand Variability

Patient flow creates random inter-related demand for ORs, ED care, and inpatient beds because many patients require combinations of these services. ED arrivals requiring either surgery, or inpatient care, or both, are generally unscheduled and stochastic. In addition to emergent cases, ORs face urgent demand from hospitalized patients, and deferrable demand from patients who schedule arrival at a future date for surgical services. The vast majority of ORs are staffed Monday through Friday for either 8, 10 or 12 hours, starting at 7 AM. Post-surgery patients begin to need inpatient beds sometime after mid-morning hours.
1.3. Demand Variability

missions through EDs tend to build at about the same time. In contrast, discharges typically happen in the afternoon. This causes a systemic misalignment in the demand and supply for hospital beds.

An example of this phenomenon can be seen in Figure 1.4, which shows the total annual number of admissions and discharges by hour for an example acute-care hospital. Admission activity rises in the morning and is fairly steady throughout the day. But discharge activity peaks sharply in the afternoon (about 1 pm). This mismatch can cause shortage of beds at certain times of the day even though the hospital may have enough staffed beds on average.

![Graph](image.png)

**Figure 1.4:** Total Annual Patient Admission and Discharge Activity by Hour of Day

Hospitals track bed occupancy, which is referred to as census. Figure 1.5 shows both the typical time-of-day and day-of-week census pattern of a 500-certified-bed hospital. The daily peak of census increases and occurs later in the day as the week unfolds because patients admitted on previous days are still in the hospital and new patients arrive. The variability around the mean census also increases later in the week. However, by Friday afternoon, the goal is to discharge eligible patients to minimize nurse staffing needs over the weekends, result-
ing in reduced weekend census. Many surgeons prefer to operate earlier in the week to minimize the need to make hospital rounds during weekends. Operating Rooms conduct only urgent or emergency surgical cases on weekends, which leads to lower inflow of patients to inpatient units.

![Figure 1.5: Hospital Census by Hour of Day and Day of Week](image)

Some of the variation observed in Figure 1.5 is natural, i.e. it is caused by the circadian rhythm. The demand for beds from the emergency department or direct admissions also exhibits natural variability. The natural variability described above has both predictable and unpredictable components (Allon et al. 2013). The EDs, for example, have predictable arrival patterns by the time of day, and tend to be busier on some days of the week. There is also unpredictable variability (randomness) in ED-arrival patterns, which can cause excessive wait at certain times of the day.
1.4 Supply Challenges

The vast majority of surgical cases are neither urgent or emergent, which means that patients could wait a short amount of time without harming their health. The bulk of the OR demand variability is driven by surgical scheduling protocols in effect at a hospital. This type of variability is often referred to as artificial variability, because it is the result of hospital policies, procedures and practices. Artificial demand variability can be predicted and steps can be taken to smooth it over time (Litvak 2005). From a management perspective, it is useful to understand different types of variability – natural or artificial; predictable or unpredictable (Allon et al. 2013; Van den Schrieck et al. 2014).

1.4 Supply Challenges

Hospital census data is used to develop long-term, medium-term, and short-term budgets and staffing plans. Historically, hospitals used the midnight census for these purposes, although the midnight census does not capture the workload variability that may occur during the day shifts. The use of midnight census was driven mostly by the ease of data collection in the era before computerized bed management and electronic-medical-record systems. This practice is changing rapidly because hourly census data is readily available in many hospitals.

First of all, hospitals may find themselves unable to cope with extreme variability in ED arrivals, making it necessary to declare \textit{ambulance divert}. During an ambulance-divert period, ambulances are advised to take patients to other nearby facilities. Second, the daily challenges in demand-supply matching of beds become especially difficult when post-surgical patients and ED patients requiring hospitalization cause downstream shortage of staffed inpatient beds, called \textit{bed block}. When bed blocks occur, post-surgery patients may be boarded in the Post-Anesthesia Care Unit (PACU), and ED patients may continue to occupy ED beds until inpatient beds become available. PACU boarding can delay the surgical schedule, while ED boarding can lead to ambulance diversions. Because boarding delays the start of inpatient care, it can in turn result in longer length of hospital stay – Singer et al.
Introduction

(2011) have identified a positive association between patients’ delay in securing beds and their lengths of stay.

Capacity, defined as the availability of staff, is also random, as scheduled staff may end up being sick, or additional staff may not be available to come in on short notice in response to unplanned excess demand. Hospitals typically staff a float pool of nurses to cover unplanned absences, and move nurses to various hospital units throughout the day to handle census fluctuations. Therefore, the inpatient-bed demand-supply matching problem is often described as a multi-faceted and difficult problem.

Much attention has been devoted in the OM literature to modeling and optimizing capacity planning problems for each of the three major hospital service types of ORs, EDs, and inpatient beds. For example, many OM papers have focused on using hourly census data, rather than midnight census, to develop staffing and scheduling models that better match daily and hourly seasonality. In addition, there are several papers that study the problem of efficient patient flow across the hospital. Flow optimization can be achieved with the help of a combination of strategies including capacity flexing and demand management. The former typically means matching peaks and valleys in demand for services with the appropriate provision of time-varying staffing levels, whereas the latter typically involves input control, e.g. by exercising control over either the number of each type of surgical procedures performed, or direct admits to inpatient units, or controlling ED demand via ambulance diversions. The strategy of smoothing peaks and valleys in demand for hospital services has been a relatively less-studied problem, but the literature on this topic is growing. Similarly, recovering from unanticipated deviations in plans is also an important but less-studied topic within the OM literature.

The commonly used economic lever of dynamic pricing for demand management is generally considered to be an unethical strategy for hospitals. However, US hospitals do negotiate unit prices with insurers and payers, which are typically much less than the hospital’s full-retail prices (the latter are commonly known as charge masters\(^4\)). Typ-

\(^4\)Uninsured patients are often billed according to the charge master.
ically, this means that the insurer designates a particular hospital as a preferred or in-network service provider. Patients pay a lower co-pay and/or deductible when using an in-network service provider and the service provider offers a discount to the insurer for each unit of service. Thus, demand management for US hospitals using pricing schemes entails deciding which insurers to align with and what discounts to offer on which services. However, hospitals usually do not offer different prices to different patients depending on operational factors such as the level of congestion, which is often how demand is shaped in other service industries.

Healthcare delivery organizations are adopting innovative strategies to better understand and meet demand. For example, some hospitals are using radio-frequency-identification (RFID) location technology to create data that identifies the location of each patient, staff member, and piece of equipment. These data are then analyzed to parameterize and validate detailed computer simulation models of the entire hospital’s operations. By using such models to analyze proposed changes in service protocols and capacity, hospitals are able to identify the best options before investing resources. Such efforts can lead to improved OR utilization, accurate nurse staffing, and provably-good patient placement rules (Hendrich et al. 2008; Rosen et al. 2014; Wamba 2012 and Yao et al. 2012).

1.5 Themes and Organization

A hospital’s capacity may be described by its physical and human resources. The former include the numerical counts of operating rooms, beds, diagnostic and monitoring equipment, and laboratories, whereas the latter include the counts of nurses, doctors, health technicians, and administrative and support staff. In this monograph, we take a different approach. We characterize capacity in terms of the three major types of services that hospitals provide: (1) Operating Room (OR) Capacity; (2) Emergency Department Capacity; and (3) Inpatient Staffed-Bed Capacity. This approach allows us to emphasize the interactions between the physical and human resources, and their organization, to satisfy
Introduction

different types of patients’ needs. This viewpoint also helps to shine light on the interconnectedness of provisioning capacity for these three service types. Commensurately, this monograph consists of 4 additional chapters. Chapter 2 focuses on operating rooms’ demand and capacity management issues, Chapter 3 on emergency departments, and Chapter 4 on inpatient units. We conclude the monograph in Chapter 5.

In each of Chapters 2 – 4, we motivate the OM-relevant problems by exploring data. In subsequent sections, we present a practitioner-oriented description of key issues, discuss the state of the practice, highlight operations problems that can be gleaned from typical datasets, and discuss contributions to the literature from these perspectives. Specifically, each of Chapters 2 – 4 has the following main sections: (1) Introduction, (2) Institutional Background, (3) Evidence from Data, (4) Operational Challenges, (5) State of Practice, (6) Current OM Approaches, and (7) OM Opportunities. Three themes are interwoven throughout the discussion in these chapters. These are

1. supply-demand matching across Operating Rooms (ORs), Emergency Departments (EDs) and Inpatient Staffed-Beds are interconnected,

2. practitioner responses have focused both on improving efficiency and effectiveness within hospital walls, and on service-chain re-design to reduce demand for hospital services, and

3. it is fruitful to think of different types of variabilities – natural or artificial, predictable or unpredictable, and different types of mitigation strategies, internally focused or service-chain focused – when addressing operational challenges (see, e.g. Section 1.3).

Next, we briefly describe the contents of each chapter.

In Chapter 2, we first present lists of questions that fall under the purview of different stakeholders concerned with operating rooms’ capacity management. This underscores the fact that different stakeholders have different objectives, which may not align, resulting in poor overall performance. Next, we use data from three hospitals to highlight the inherent variability in the mix of surgical procedure types and
1.5. Themes and Organization

their durations. We also demonstrate the typical consequences of block assignments, which is a common approach used in the US hospitals to allocate OR capacity to surgeons. Based on these ideas, we identify a series of OM-relevant decision problems, formulate models, and discuss possible solution methods and related literature. One of the key issues we identify is that although surgical cases are often booked one at a time (in an online fashion), the vast majority of papers in the OM literature consider the situation in which all cases that need to be booked for a particular day are known in advance (i.e. an offline version of the problem).

Long wait times in Emergency Departments (EDs) have attracted a great deal of attention from policy makers and the public. In Chapter 3, we begin by explaining the origins of EDs and an institutional perspective on factors responsible for severe ED capacity shortages. Next, we explain how patients typically flow through an ED and operational performance metrics. We use evidence from hospital data to identify three key operational issues. What mechanisms can best shape the demand for ED services? How can the ED process itself be improved? How can an adequate supply of downstream beds be assured for ED patients who require hospitalization? Chapter 3 is focused on state of practice, OM models, and open challenges as they relate to these three problems.

In Chapter 4, we turn to the problem of providing adequate staffed bed capacity to meet demand in a cost effective manner. Whereas the number of physical beds (referred to as the number of licensed or certified beds) changes infrequently and may be subject to government-agency approval, the number of staffed beds varies by day and by shift in an effort to avoid shortages and overages. Shortages may result in poor patient care as well as increased cost. The latter occurs because contingent workforce wage rates are higher than those of regular employees with the same level of expertise. Excess capacity leads to higher staffing costs. In this chapter, we identify four key decisions that have a direct bearing on staffing performance. These are (1) unit size and scope decisions, (2) the choice of performance metrics, and (3) policies concerning planned time off and shift schedules, and (4) policies con-
cerning patient placement and movement. For each scenario, we present the state of the practice and the state of the OM research, and identify open research problems.

In concluding the monograph in Chapter 5, we summarize our findings and highlight the central theme from the three preceding chapters, which is that by simultaneously considering OR scheduling policies, ED demand prediction, and medium and short-term staffing plans, OM researchers have an opportunity to make a significant impact on the future delivery of healthcare services. We hope that this monograph will make it easier for new researchers to get acquainted with key issues, and with the state of practice, research, and open problems.


References


References


References


References


Kane, N. and Siegrist, R. Understanding Rising Hospital Inpatient Costs: Key Components Of Cost And The Impact Of Poor Quality. *Blue Cross Blue Shield*, 2002.


Full text available at: http://dx.doi.org/10.1561/0200000036


References


Full text available at: http://dx.doi.org/10.1561/0200000036
References


References


**References**


