# Health System Innovation: Analytics in Action

**Martin S. Copenhaver,**[a,b] **Michael Hu,**[c] **Retsef Levi,**[b] **Kyan Safavi,**[a]
**Ana Cecilia Zenteno Langle**[a]

[a]Massachusetts General Hospital, Boston, Massachusetts 02114; [b]MIT Sloan School of Management,
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; [c]Operations Research
Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

**Contact:** mcopen@mit.edu, https://orcid.org/0000-0002-9988-260X (MSC); hum@mit.edu,
https://orcid.org/0000-0002-6802-168X (MH); retsef@mit.edu,
https://orcid.org/0000-0002-1994-4875 (RL); ksafavi@partners.org,
https://orcid.org/0000-0002-0883-7329 (KS); azentenolangle@partners.org (ACZL)

**Abstract**     This tutorial focuses on the implementation of data- and analytics-driven innovation in
             health systems. In particular, we focus on innovation around operations, system design,
             and optimization of large delivery systems. Additionally, the tutorial discusses a formal
             project management framework, general principles, and key success drivers that enable
             field implementations of high-impact, analytics-driven projects. These discussions are
             specifically centered on facilitating collaboration between academics with analytics ex-
             pertise, clinicians, and administrative leaders in healthcare systems, as well as policy
             makers. To illustrate the usage of these ideas in practice, we describe three projects done in
             collaboration between Massachusetts General Hospital and the Massachusetts Institute of
             Technology Sloan School of Management.

**Keywords**   healthcare • health system • healthcare operations management • service operations •
             hospital operations • analytics • optimization

## 1. Introduction

Over the last decade, *analytics* disciplines and methodologies have been increasingly leveraged
to drive system-level innovation in many industries and areas, such as airline, retail, mar-
keting, and finance. The healthcare industry, albeit behind, is making big strides. Given the
increasing and uncontrolled government and private spending, as well as the quality and safety
challenges that health systems around the world face, the need for disruptive innovation is
more pressing than ever. Furthermore, the widespread dissemination of electronic medical
record systems (EMRs); advances in the understanding and mapping of the human genome;
and various emerging digital, mobile, and wearable technologies are enabling an un-
precedented accumulation of rich health-related data on individuals and populations.

   Stimulated by these emerging trends, academics from various communities and disciplines
have been increasingly researching the potential that data and analytics have to inform and
drive innovation in the healthcare industry. Current research efforts are focused on various
application domains, such as the development and optimization of clinical care and thera-
peutics, improved operations management and design of health systems, changing public
policy and regulations, as well as the design of financial schemes and incentives. Yet, thus far,
much of this emerging body of work has stayed at the academic level of models, simulations,
use cases, and small-scale pilot experiments, whereas widespread adoption and dissemination
have been relatively slow and limited. The healthcare industry and health systems are
complex, are highly regulated, and often involve high stakes, all of which can serve as inhibitors

and barriers to system innovation. At the same time, much of the academic work in these areas has not yet fully and consistently incorporated various organizational, cultural, and regulatory considerations that are critical to the practical relevance of the work and its ultimate potential impact on practice.

This tutorial is focused on the *implementation* of data- and analytics-driven innovation in health systems, specifically innovation around operations, system design, and optimization of large delivery systems. The tutorial describes vignettes from projects conducted as part of a multiyear collaboration between Massachusetts General Hospital (MGH) and the Massachusetts Institute of Technology (MIT). The goal is not merely to highlight the technical details and the respective analytics approaches and solutions but to also describe a robust process of collaborative work between academics with analytics expertise, clinicians, and administrators in large medical centers. Moreover, the tutorial attempts to identify and discuss key considerations and success drivers as to how analytics related academic research can lead to field implementations of innovation in complex health systems. This will hopefully provide some principles indicating how academics who master analytics methods can effectively collaborate with clinician and administrative leaders, as well as front-line clinical teams, to disseminate related research into innovation in the operations management and system design of large medical centers and potentially other large health systems.

To appropriately position the work described in this tutorial, it is worthwhile to discuss the major areas of study within the broader and evolving work of healthcare analytics. The discussion will be brief and focused on the major research questions, methodologies, and typical results in each stream of work. The examples mentioned are merely a small sample and do not represent the entirety of relevant work.

## 1.1. Diagnostics, Clinical Care Protocols, and Therapeutics

There is an increasing body of work that aims to leverage data to inform how clinical care protocols could be optimized and personalized to individual patients and specific patient segments. This includes a wide range of medical conditions and care elements, such as diagnostics, care interventions, drug selection, and optimal treatment parameters. Typical questions include the following: What is the dynamic policy that optimally uses diagnostic resources? What is the optimal combination of drugs for each patient segment, and what are the appropriate patient segments? What is the best timing to deploy certain clinical interventions? Can certain medical conditions be diagnosed earlier based on available data? This body of work relies heavily on data-driven work, as well as on underlying models of disease progression. To tackle these questions, there is a wide range of optimization, statistical, and machine learning frameworks and methodologies that are being applied, such as *partially observed Markov decision processes* (e.g., Ayer et al. [12], Bertsekas and Tsitsiklis [20], Schaefer et al. [77]), *robust optimization* (e.g., Chan and Mišić [27], Nohadani and Roy [69]), *multiarm bandit models* (e.g., Bastani and Bayati [17]), as well as existing and newly developed machine learning models and algorithms, particularly sparse models (e.g., Bertsimas et al. [22], Mahmoudi et al. [61]).

The potential impact of this work is high. In particular, the hope is that physicians will adopt the specific models or at least the insights that emerge from the models to improve the design and delivery of diagnostic and clinical care protocols. Indeed, the medical community is increasingly interested in this type of work, and many physicians actively seek out academic partners with analytics skills and expertise. However, there remain multiple significant barriers that often inhibit and slow the rate and scale of implementation of this type of work. First, the efficacy analysis of the work often relies on pure modeling, simulations, or (retrospective) observational data analyses. These have objective limitations but, more importantly, fall below the stringent standard commonly accepted in the medical world to rely on *randomized control experiments* to assess efficacy. Moreover, the relevant data sets are often

private, and it is challenging to integrate multi-institutional data to obtain more powerful and reproducible analyses. Whereas models can inform potential randomized trials, there are substantial logistical and regulatory challenges to conduct randomized field experiments with newly proposed (model-driven) protocols. Furthermore, there are often significant challenges to incorporating newly developed tools and models into the existing workflows and operational information technology (IT) systems (e.g., EMR) currently used by physicians and clinical teams. This is another major barrier for adoption. Finally, although physician culture is changing rapidly, there are still many physicians who are resistant to the adoption of such tools, because of a lack of familiarity with analytics methods and, potentially, for fear that their traditional roles as sole decision makers will be diminished.

## 1.2. Public Policy and Health-Related Public Resource Allocation

There is a large and evolving body of work that aims to inform public health policy and regulations. One stream of work in this category aims to develop models that more accurately assess healthcare costs (e.g., Bertsimas et al. [23]) or to inform government (and private) payment schemes (e.g., Fuloria and Zenios [30]). Another important stream of work in this space is focused on the optimization of the policies and systems responsible for *organ transplantations* and other scarce resources (e.g., Agarwal et al. [2], Ashlagi and Roth [9], Ata et al. [10, 11], Bertsimas et al. [21], Kaplan and O'Keefe [48], Kim et al. [51]). This work covers multiple aspects, including issues such as the design of improved criteria and algorithms to assign organs to patients, incentive schemes to improve overall system performance, and logistics optimization. Data used for this type of work include publicly available data, such as claims, as well as private data on patients who belong to specific hospital systems or one provided by a government agency as part of sponsored research (or a formal research contract). From a methodological point of view, this stream of work relies on various modeling approaches in stochastic systems (e.g., random graphs and queuing models), game-theoretic settings to capture the various stakeholder incentives, and classical optimization techniques and frameworks. In addition, this body of work often employs a range of empirical methods to establish causal inference and efficacy analysis, often supported by computerized (data-driven) simulations.

It is worth noting that at least some of the work on organ transplantations has, in fact, resulted in field implementations and tangible impact on practice (e.g., Anderson et al. [7], Ata et al. [11]). One potential driver for these impressive success stories is the fact that in this area, there are available curated data sets (some are anonymized) and, even more important, commonly accepted simulation settings to assess newly proposed policies. In other cases, obvious interfaces for dissemination of the work into practice do not currently exist, with the typical barriers including a lack of structured organizational interfaces, as well as a lack of culture and expertise in using data analytics.

## 1.3. Health System Operations and Design

This large body of work is concerned with questions related to improving clinical and operational processes and tackling system-level issues in healthcare organizations and networks. This work includes the design and optimization of a broad range of activities, such as scheduling and staffing practices and algorithms, network design, strategic and real-time resource allocation, patient placement and flow management, and other clinical and management processes. There are several themes on the methodology front. One is concerned with descriptive empirical analysis that employs various econometric methods to identify causal mechanisms (Ang et al. [8], Kim et al. [49, 50], Song et al. [82], Terwiesch et al. [83]). Although typically not providing specific prescriptive algorithms, these studies often provide high-level insights on the effectiveness of existing practices and how they could potentially be improved.

A second theme is concerned with developing predictive algorithms, for example, to assess the risk of specific patients to be readmitted or not show up to an appointment. Another example would be predictive models to anticipate how long patients will stay in the hospital (e.g., Carter and Potts [24], Liu et al. [60], McEvoy et al. [63], Mohammadi et al. [65], Robinson et al. [73], Shi et al. [80]). Other predictive algorithms are developed in the context of specific decisions or interventions and aim to inform which patients should be prioritized. The papers under this theme employ a range of statistical and machine learning models, and some also assume certain structural models to capture the dynamics of the underlying systems. A third theme is focused on the development of prescriptive optimization algorithms that support specific strategic or operational processes with the goal of providing decision support tools to clinical and administrative teams (e.g., Ayer et al. [13], , Hu et al. [42], Lee et al. [53, 54], Li et al. [58], Negoescu et al. [68], Saghafian et al. [75], Shi et al. [80], Sir et al. [81], Zenteno Langle et al. [92]). These papers employ a range of optimization models and algorithms, as well as simulation techniques, often incorporating inputs from predictive models. Most of the work in this area is done based on private data sets that belong to specific systems (e.g., scheduling data, EMR records and time stamps, radio-frequency identification [RFID] systems and claims). This is perhaps the largest body of work among the three categories, and the related work goes back several decades. There is no doubt that this area of work has tremendous potential impact on the quality and cost of health systems. In fact, an increasing number of health systems recognize the importance of developing internal analytics capabilities and collaborating with analytics experts. This trend is partially affected and incentivized by ongoing changes in market payment schemes and the transition, in some states, from fee-for-service to risk contracts. However, excluding relatively few exceptions (e.g., Ayer et al. [14], Cohn et al. [28], Lee and Zaider [52], McEvoy et al. [63], Shi et al. [80], Sir et al. [81], Thomas et al. [84], Thompson et al. [85], Woodall et al. [90], Zenteno Langle et al. [92]), most of the existing work does not ultimately get implemented in the field, and even in cases when it does, these are typically small-scale pilots without clearly documented sustained, system-level impact. There are multiple hypothesized barriers that lead to the current lack of implementations, which this tutorial discusses in detail in the subsequent section. Furthermore, cross-system dissemination is even more challenging.

As already mentioned, the goal of this tutorial is to discuss several general principles and success drivers that enable field implementations of high-impact, analytics-driven innovation in large medical centers and delivery systems, specifically through collaboration between academics with analytics expertise, clinicians and administrative leaders in healthcare systems, and policy makers. The rest of the tutorial is structured as follows. Section 2 provides a discussion of the structure and methodology of the MGH-MIT collaboration and particular key success drivers to enable system-level, analytics-driven innovation. This section is followed by three sections, each discussing a particular vignette of a specific project. In particular, the three examples that will be discussed are as follows:

i. In Section 3, the implementation of a new scheduling system at the *MGH Cancer Infusion Center*

ii. In Section 4, the development of a machine learning–based algorithm to *predict hospital discharges in the next 24 hours* as part of *a redesign of the discharge process* in the hospital

iii. In Section 5, the design of a system-level strategy to improve the management of *heart failure patients*

The discussion of each sample project includes a description of the underlying problem that the project aims to address the development of the solution approach and the implementation strategy. In particular, the discussion of specific projects aims to illustrate how various principles and success drivers discussed in Section 2 play a role in the specific context.

## 2. Collaboration Model and Success Drivers

### 2.1. Background

MGH was founded in 1811 in Boston, Massachusetts, as the first teaching hospital of Harvard Medical School and is, together with Brigham and Women's Hospital, a founding member of Partners Healthcare. It is consistently ranked among the top five hospitals in the United States by *U.S. News and World Report*, and it is widely regarded as one of the leading healthcare institutions in the country. With over 1,000 licensed beds, MGH admits approximately 50,000 inpatients, performs more than 42,000 procedures and records close to 108,000 emergency room visits per year. Additionally, MGH serves over 1.5 million outpatients annually in clinics located across several specialized treatment facilities. Finally, as an *academic medical center* (AMC), MGH administers numerous residency training and educational programs and maintains an extensive research portfolio (Massachusetts General Hospital [62]).

The *MGH-MIT Collaboration* started over a decade ago between MGH and the MIT Sloan School of Management. Over the years, it has evolved to have a formal structure within MGH that is managed by a recently constructed vice president- (VP-) level leadership role, the *Vice President for Healthcare Systems Engineering*, with direct guidance, funding, and oversight from the president of the hospital as well as other executive leadership. At its inception, the collaboration focused on the *perioperative* environment (i.e., the system responsible for surgical activities), but after several successful data-driven implementations, it expanded its reach to pursue projects in the hospital more broadly, including multiple outpatient projects. Over the last four years, the collaboration as a strategic priority has centered its attention on projects related to improving inpatient capacity management by addressing major patient crowding and flow management challenges that MGH has been facing over the years.

The first project implementation occurred at the end of 2011; it was enabled by work to optimize surgeons' day of surgery with the goal of smoothing out the weekly patient census on the surgical floors. On the basis of the recommendations of the developed optimization framework, about 35% of the surgeons at MGH changed their day of surgery (Zenteno Langle et al. [92, 93]). Since then, the collaboration has implemented nine multiyear, system-level projects, almost all based on extensive data analytics and decision support tools that were operationalized in the field. These projects have resulted in a major positive impact on the AMC's operational efficiencies and quality of care. In addition, there are currently five projects that are in the implementation phase and several more that are in the planning phase. The collaboration's projects have involved many of the services and units of the hospital, in both the inpatient and outpatient arenas, and engaged many clinical and administrative leaders and teams. Other examples of projects that were implemented but are not discussed in detail in this tutorial include

i. the development of improved algorithms to assign hospital beds to admitted patients (Hiltrop [40], McNichols [64], Ugarph [88]),

ii. the design of a new outpatient nononcology infusion center (Ghobadi et al. [32]),

iii. the optimization of surgical packs and supplies inventory levels (Ben-Zvi [18], Schlanser [78]), and

iv. the optimization of primary care physicians' session schedule to smooth staff workload (Patel [71]).

### 2.2. Project Management Methodology

The MGH-MIT Collaboration has developed a proven methodology to identify, manage, and implement its projects.

**2.2.1. Project Identification.** Each project starts with the identification of a relatively broad problem domain based on multiple criteria, including

i. hypothesized significant system-level impact in high-priority areas according to the priorities of the hospital senior leadership,

ii. assessment that the solution is likely to be nontrivial and require major system-level data analyses and changes,

iii. the existence of relevant operational and clinical data that allow rigorous analyses and hypothesis testing, and

iv. organizational maturity to make the necessary changes, and the presence of appropriate "local" (unit-level) leadership that is ready to embrace and lead the anticipated change process.

The verification of these "conditions" is done through a series of discussions and meetings, as well as appropriate data analyses. At the successful conclusion of the identification phase, the initial project team is constructed. The project team typically consists of

i. clinical and administrative leaders from the relevant affected units and other stakeholders likely to be involved in the project;

ii. representatives of the MGH leadership of the collaboration; and

iii. MIT students, postdoctoral fellows, and faculty.

There is also a set cadence of weekly meetings to plan, conduct, and discuss the data analyses and project management; they also serve to monitor progress and provide leadership guidance. Although the goal of each project is to ultimately develop and implement practical solutions in the hospital system, the projects also have explicit and formally defined academic research goals. In particular, many aspects of the projects are managed as a research collaboration with the aim to mentor and enable the research output of MIT students and postdoctoral fellows, as well as the hospital's residents and fellows.

**2.2.2. Project's Problem Definition.** As mentioned above, the initial identification phase is focused on a broad problem domain. The subsequent phase aims to study this problem domain and develop a more concrete and specific problem definition. This process can take several months and is informed by extensive process and system mapping, experts' inputs, and extensive data analyses and analytics. The decision is informed by multiple factors, among which are

i. data-driven identification and validation of hypotheses regarding system-level root causes of observed "organizational symptoms,"

ii. refined understanding and measurement of desired performance metrics to be affected and what their current state is,

iii. assessment of the organizational maturity to accept different levels of change, and

iv. longer-term vision regarding the most effective sequence of efforts that maximizes the overall likelihood to implement system-level change and obtain significant impact.

This phase is conducted by the project team with oversight of the senior leadership of the hospital and the involved units (typically the hospital's president and senior VPs). At the conclusion of this phase, the team's composition could be modified to adapt to the refined project's problem definition and the hypotheses regarding anticipated activities.

**2.2.3. Development of Prototype Solution Approaches and Tools.** This phase is conducted in an iterative manner over multiple months, following research hypothesis testing methodology and extensive data analyses. This step is also informed by literature review of relevant past work. In addition, there is extensive development of appropriate analytics tools. The typical outcome of this phase is not a singular prescribed solution but rather a robust modeling framework and a set of decision support tools that include

   i. end-to-end modeling tools that

     a. capture the relevant potential changes to the system and workflow processes,
     b. enable reliable assessments of the anticipated impact of these changes, and
     c. delineate and quantify the magnitude of the main trade-offs decision makers should consider; and

  ii. validated prototypes of appropriate decision support tools that enable the newly proposed system and process design.

At the conclusion of this phase, there is a sequence of organizational presentations and discussions to obtain feedback and buy-in from all relevant stakeholders. Finally, we work to obtain the hospital's leadership approval to move on to the implementation phase.

**2.2.4. Development and Execution of Implementation Plan.** Once a project transitions to the implementation phase, the project team is expanded to include additional relevant stakeholders and other units that are critical for this process. The team continues to iteratively refine the solution approaches and decision support tools, as more considerations and constraints are identified. Depending on the nature of the project and the specific developed solution approaches and tools, the implementation plan could include a sequence of pilots with the goal of building organizational confidence and buy-in, test various key elements of the newly proposed system changes, and smooth the transition from the current state to the desired new state. One of the major focus areas during this phase is the detailed planning and testing of new workflow processes. Specifically, there is focused effort to identify and develop practical plans to integrate the new decision support tools into the existing hospital IT infrastructure that supports the workflow processes. Furthermore, there is extensive work to transition the analytics and decision support tools and models from the development phase, in which MIT team members are heavily involved, to an operational state in which the expectation is that the tools will be operated and maintained primarily (and hopefully solely) by hospital staff.

**2.2.5. Monitoring Results and Impact.** This phase starts during the implementation phase and continues afterward. The goal is to monitor the outcomes and impact of the project on the corresponding predetermined project's performance metrics, as well as others. It is critical to establish measurement methodology and robust processes and infrastructure, not only for research purposes (this is typically the phase when academic research articles are written) but also to sustain the project's impact and allow further improvements and refinements over time. Moreover, this is also critical to inject a culture of data-driven decision making and to build organizational confidence, which are key for the ability to pursue future, potentially more ambitious, projects.

Availability of data is key to all of the project phases outlined above. The MGH-MIT Collaboration has devoted significant efforts to develop a robust data infrastructure that supports the work. These efforts include processes to allow team members timely access to the relevant data sources and data sets. For example, MIT students and postdoctoral fellows are enrolled (including appropriate training) as nonemployees at MGH to allow them easy access to the hospital data systems. Rigorous and robust Institutional Review Board and data management processes are implemented to enable the work. Most important, there is extensive effort to constantly identify, understand, and integrate new relevant data sources. The collaboration team has already integrated very rich data from multiple modalities and sources. This includes, among other things, clinical orders and records, unstructured clinical notes, scheduling data, claims, operational time stamps recorded in the EMR, RFID data, and even computer click-level data. The integration of the data enables first-of-a-kind analytics and analyses that are often key to the development and implementation of the innovative system-level solutions and processes.

## 2.3. Key Success Drivers

The goal of this section is to briefly highlight several key success drivers that enable the implementation of system-level, analytics-driven innovation, particularly in the context of collaborative research work between academics with analytics expertise and large healthcare delivery systems. Some of the subsequent key success drivers are likely to be relevant in broader settings.

### 2.3.1. Organizational Interface and Senior Leadership Involvement.
One of the common failure modes of collaborative efforts, particularly related to the ultimate ability to reach field implementation and more generally obtain overall impact, is the lack of an appropriate organizational interface to the health system and leadership involvement. Specifically, because of the organizational complexity of health systems and the entrepreneurial research culture, particularly at AMCs, it is not uncommon that interactions with individual clinicians could give the wrong impression regarding the importance of the underlying problems that are being addressed and their corresponding system-level root causes. Another challenge is the potential cultural and language gaps between analytics experts and clinicians. Although there is no unified recipe to build the appropriate organizational interface and leadership involvement, and the process of creating those could take time and evolve gradually, there are still several key issues and questions that should be addressed. One key input is an assessment regarding where the specific interfacing unit and individuals are placed within the respective health system, and what their organizational perspective and decision power are likely to be. The latter requires very good understanding of the structure, dynamics, and decision-making processes of the relevant health systems. In general, there should be a reasonable match between the proposed scope of the collaboration and what is realistic to expect from the existing organizational interface. Finally, it is important to develop a common language and mutual understanding of the different cultures.

### 2.3.2. Focus on Decision and Workflow Processes.
Another common failure mode of collaborative research efforts is the lack of clarity as to which underlying decision processes are studied and addressed. In fact, as much as analytics could be a key enabler of system-level innovation, it is important to realize that it is only a means to an end. The ultimate goal should always be defined as the development of improved system-level decision and workflow processes. Too often efforts are focused on the development of advanced analytics tools, such as sophisticated predictive models, without sufficient (or any) understanding of the underlying decision and workflow processes. As an example, consider the relatively extensive work conducted on predicting the length of stay (LOS) of patients in hospitals (Azari et al. [15], Barnes et al. [16], Carter and Potts [24], Gustafson [37], Liu et al. [59], Morton et al. [66], Robinson et al. [73], Tu and Guerriere [86], Walczak et al. [89]). Most of these algorithms are not very useful, as it is not at all obvious how a prediction of LOS can be leveraged to improve hospital operational and clinical decisions. In fact, in most hospitals, current decision and workflow processes cannot leverage predictions of LOS that are conducted many days before the patient is actually supposed to be discharged, unless they are very accurate (see a more detailed related discussion in Section 4). Another typical example is work that does not incorporate first-order hard constraints that stem from either clinical or cultural factors. For example, theoretically speaking, the "optimal" surgical schedule in a hospital will likely include surgeries on each one of the days of the week, including Sunday. However, this is not likely to be conceivably feasible because of a blend of cultural and logistics reasons. Thus, it is critical to integrate data and analytics within the system's context and have deep understanding of the relevant clinical, organizational, and cultural factors. More examples of such constraints are provided throughout the discussions of the different vignettes.

**2.3.3. Data Availability and Interpretation.** Another key aspect is the availability of data, which can be a major barrier for implementing change or even conduct meaningful, analytics-based work. Once data are accessible, it is key to develop a deep understanding of the underlying system's context and related workflow processes to have appropriate *data interpretation*. In particular, most of the data leveraged to conduct analytics work in health systems is generated by clinical and operational processes that are system specific. This is critical to be able to use data appropriately. For example, consider analytics work that is concerned with patients' delays and wait times in the emergency department (ED). A typical reason for patient delays in the ED is the lack of available hospital beds. As such, it would be natural to assess the respective wait times based on the typically recorded time stamp of the first time a bed was requested for the respective patient by the ED team. However, hospital systems are often highly utilized, and clinical teams, anticipating long delays, could be tempted to aggressively place bed requests well before the respective patient is clinically ready to leave the ED. Not accommodating for these endogenous dynamics can easily lead to false analyses and ineffective analytics tools.

**2.3.4. The Integration of Expert Inputs and Analytics.** Current clinical and operational decision and workflow processes center around the dynamics of humans providing service to humans. This human-to-human interaction is a key characteristic of health systems with deep cultural roots. Moreover, the human service providers in health systems are typically highly qualified experts that are often trained around the premise of full individual autonomy and ownership of decisions related to the care of their patients. Moreover, there is often individual ethical and legal liability imposed on providers. All of this makes the appropriate integration of analytics and decision support tools with experts' input critical to the ability to ultimately implement them. The integration of expert inputs is central to multiple design aspects in the development of analytics and decision support tools. First, it should affect modeling and methodology choices with clear preference to interpretable models and tools. Second, it should affect the tools' outputs that need to be adapted to match the decision and workflow processes used by (or at least natural to) clinicians and staff members. Finally, the integration of data and model inputs with experts' input is a nontrivial task and, in fact, often presents some important open research questions.

**2.3.5. Focus on Trade-offs and Improvement Opportunities.** Whereas traditionally academic disciplines related to analytics are often focused on the notion of *optimal solutions*, this notion often does not exist in complex health systems that operate under multiple and likely contradicting objectives and metrics. Thus, most attempts to develop tools and models that prescribe specific optimal solutions are likely to fail. Instead, decision support tools that are designed to enable and support organizational efforts to improving systems and processes are far more attractive. In particular, such models should provide and support at least three functions:

i. reliable predictions of the anticipated outcomes of different decisions ("what-if analyses"),
ii. highlighting of decision trade-offs, and
iii. identification of system-level improvement opportunities.

To implement these functions, there is need to consider not only the "mathematics" of the underlying tools but also, and equally important, the "communication" aspects of the tool. The latter includes consideration related to data visualization and the creation of appropriate dashboards.

# 3. Vignette 1: Outpatient Capacity Management in an Infusion Center

We begin our first of three vignettes by focusing on the example of an outpatient infusion unit at Massachusetts General Hospital (Lennes et al. [55], Reib [72], Zenteno Langle et al. [94]). The effective management of outpatient healthcare resources is an area ripe for the application of data-driven, *practically implementable* operations research methodologies. This is particularly true in light of growing interest in and demand for systems-based care wherein areas such as centralized resource planning and population health management are critical components. In this section, we will illustrate several key principles related to outpatient capacity management in the context of infusion centers.

## 3.1. Overview

One of the cornerstones of modern cancer treatment is chemotherapy administration, which is most often administered through intravenous infusion or through pills in an outpatient facility or infusion unit. In addition to chemotherapy, patients also receive related or supportive therapies in the infusion unit such as blood transfusions, hydration, antibiotic therapy, and pain management treatments. With these combined, infusion appointments may last anywhere between 30 minutes and 12 hours. Treatments are generally scheduled in series up to 6 weeks in advance, with checkups with the clinical team before some of the infusion visits to ensure that patients are strong enough to receive treatment or that they are responding well to it.

The MGH Cancer Center (CC) has over 100 medical oncologists and nurse specialists spread across multiple divisions, and it operates 60 chairs in its main campus outpatient infusion unit. In 2013, this infusion unit was experiencing a 2.5% annual growth rate in the number of infusion appointments and, at the same time, severe overcrowding every weekday, especially between 10 a.m. and 2 p.m.; the unit was underutilized otherwise (see Figure 1). During congested hours, patients would experience long wait times, and staff members would be severely strained, creating frustration, job dissatisfaction, and concern for safety issues. Moreover, even though resources were clearly underutilized at the beginning and the end of the day, the unrelenting midday congestion created a false perception of insufficient capacity, effectively preventing volume growth and reducing access to care. This disproportionate midday appointment load is not an uncommon problem for large oncology infusion units. Indeed, some cancer centers are actively asking patients to try to book appointments at the beginning or at the end of the day to avoid crowded hours (Dorland et al. [29]).

**Figure 1.** (Color online) Average scheduled daily utilization (initial state).
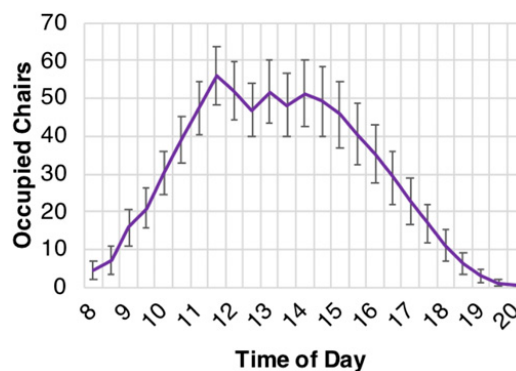


*Notes.* Average number of scheduled occupied chairs by hour into the day (±1 standard deviation) for preimplementation phase. Physical capacity is 60 chairs.

To target this problem, we designed a data-driven online scheduling algorithm that aims at generating a more balanced, and thus more predictable, intraday resource utilization. The study is based on a detailed patient flow analysis of the infusion unit's patient flow and its relation to the CC clinics that refer patients to it.

## 3.2. Initial Analysis

The Cancer Center patient flow is highly complex, as patients need to coordinate appointments within many different areas, depending on their clinical needs. In general, patients typically need to get their blood drawn and wait for the results to inform whether they are in adequate clinical conditions to receive chemotherapy. Next, they may see their oncologist or a nurse practitioner to review the treatment plan (practice-to-chemotherapy, or PTC, visit) or go directly to infusion for treatment (infusion-only).

At the time of this study, the practice and infusion unit operated separate, in-house scheduling services. When a patient's oncologist ordered an infusion regimen, the patient would receive a list of dates in which her PTC and infusion-only visits should occur. The patient then would proceed to meet with a scheduler in the practice who, in turn, would call a scheduler in the infusion unit to coordinate with the patient. Thus, the practice scheduler had no visibility of the infusion unit utilization when requesting appointments. Additionally, schedulers, who often develop relationships with patients, would regularly offer what is perceived as convenient times for patients, or try to accommodate their preferences as much as possible.

Myopic scheduling is not unique to MGH; previous studies have identified that failing to incorporate the utilization of the appropriate resources into appointment scheduling practices is a main inefficiency driver in outpatient infusion settings (Aboumater et al. [1], Chabot and Fox [26], Gruber et al. [35], Kallen et al. [47], Turkcan et al. [87]).

In summary, scheduling infusion treatments presents the following challenges:

1. *Multiresource scheduling.* All visits require blood draws prior to treatment. Additionally, three of four infusion visits are coupled with an appointment in the oncology practice, where the patient is assessed before receiving chemotherapy.

2. *"Online" scheduling.* The coupled-appointment system requires the Cancer Center to make scheduling decisions as the appointment needs are revealed. That is, each patient is assigned both the dates and the times of her appointments at the same time.

3. *Operational and clinical constraints.* Not only does the schedule need to respect the infusion unit constraints (e.g., nurse staffing levels) but also the clinic times in the oncology practice, as well as pharmacy capacity where treatments are personally tailored to each patient once he or she has arrived and been approved for treatment.

4. *Patient-centered culture.* There is a very strong practice of accommodating patient requests and needs, sometimes in disadvantage of the system's overall performance.

5. *Schedule variability.* The infusion unit accommodates previously unknown treatments that are added to the schedule on the same day, same-day cancellations, or modifications to treatment as a result of the patients' clinical conditions. Additionally, it provides treatment for patients that are part of clinical trials, for which the treatment is highly nonstandardized.

## 3.3. Solution Approach and Related Work

The problem of scheduling a new infusion (of known duration) on a specific day (1) given an existing collection of scheduled infusions and (2) with the objective of minimizing the peak utilization during the day can be naturally modeled as an online optimization problem. Such a framework is useful for several reasons:

1. It gives a correspondence between the online (true) version of the problem and its offline (optimistic) counterpart. Consequently, the offline mixed integer optimization problem of minimizing peak utilization given the entire collection of infusion appointments scheduled

during the day provides a lower bound on the best possible online policy. This can subsequently be used as a benchmark in simulated comparisons. Furthermore, the solution to the offline problem provided insights into what a "good" schedule looks like and ultimately guided the design of the online algorithm.

2. It accommodates the introduction of constraints. In light of the observations made earlier, some of the constraints include respecting clinicians' schedules (for the infusions that are "linked" to clinician appointments) and only scheduling "sensitive" chemotherapies[1] toward the middle of the day when the most resources can be made available.

**3.3.1. Related Work.** There is a rich body of literature on online algorithms for scheduling (see references in Jaillet and Wagner [44]) as well as scheduling applications in the healthcare setting (e.g., Cayirli and Veral [25], Gupta and Denton [36]). Scheduling of chemotherapy infusions has also received a variety of attention, including for treatment planning (e.g., Agur et al. [3], Shi et al. [79]) as well as other inter- and intraday scheduling considerations (e.g., Gocgun and Puterman [33], Santibáñez et al. [76]).

Although some of the techniques and principles of our work are common to the studies mentioned above, the main difference is that we focus on building an intraday schedule *in an online fashion*; that is, patients are told the date and time of their appointment at the same time. This is important for the MGH Cancer Center Administration because, following patient preferences, the practice and the infusion appointments occur on the same day, which implies that both of them need to be booked simultaneously. This is contrary to the approach of other centers, where these appointments are decoupled to reduce delays and wait time. Also, in our study we did not attempt to determine the dates in which patients receive treatment, as this interday timing is typically guided by clinical guidelines. We take these as given and determine only the times of their appointments. Finally, although we do not account for pharmacy and nursing resources explicitly in our model, minimal adjustments to their current setup are shown to be enough to accommodate the proposed solution.

**3.3.2. MinGap Policy.** A simple heuristic that we refer to as "MinGap" was devised to (approximately) solve the online optimization problem; namely, given an existing collection of scheduled infusions for a day and a new appointment needing to be scheduled,

1. select the appointment time that minimizes the gap between the (max) peak utilization and (min) trough utilization *during that appointment* while not increasing overall peak utilization during the day; and

2. in the case that all feasible appointments increase overall peak utilization, choose the one that minimizes the gap between the peak utilization and trough utilization during the appointment.

**Figure 2.** (Color online) Smoothed average daily utilization.
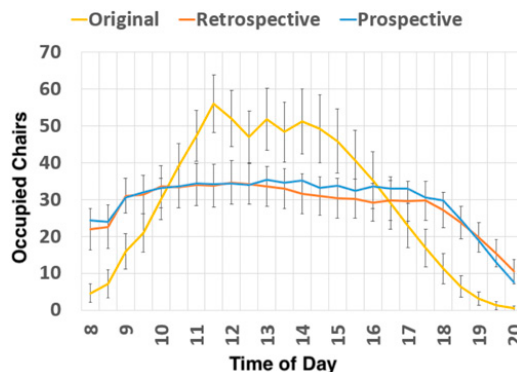


*Note.* "Original" as in Figure 1, "Retrospective" is the optimal offline policy, and "Prospective" is the MinGap policy with a single option offered.

Despite its simplicity, the MinGap policy seems to perform exceptionally well relative to an offline optimal solution (for comparative performance in a simulation, see Figure 2). Moreover, the resulting projected distribution of appointments worked well with other key components of CC's working model: primary nursing, pharmacy, and ancillary staff. The CC's Infusion Unit follows a *primary nursing model*, under which patients are matched with a primary nurse with whom their subsequent infusion appointments are preferentially, though not exclusively, scheduled. This model aims to increase continuity of care and improve patients' overall experience. Although this was not an explicit constraint in our model, given that most nurses have 10- to 12-hour shifts, distributing appointment start times throughout the day mostly helped patients to get appointments with their primary nurse. The additional load of appointments at the beginning and tail-end of the day could also be worked out by redistributing existing resources in staggered shifts.

Finally, another key aspect of the MinGap policy is that it can be adapted to offer patients more choice. In particular, the algorithm can rank possible appointment start times according to how much they minimize the gap between the peak and the utilization throughout the treatment duration. In this way, schedulers give patients a more active choice in their treatment planning while offering the best possible appointments from a resource-utilization perspective.

## 3.4. Implementation and Outcomes

One of the critical barriers to implementation (and innovation) of operations research projects related to scheduling in healthcare applications is the connection between the EMR and the scheduling algorithm/software. This is particularly difficult because the EMR software system is often owned by a third-party vendor, thus raising issues surrounding intellectual property and ownership if any direct modifications were to be made to the EMR. Furthermore, most EMRs have built-in scheduling capabilities, which complicates the decision of whether to implement a stand-alone system. Such a system would necessarily require both reading data contained in the EMR and feeding any scheduling decision back to the EMR, as this acts as a reference point for other departments that may want to schedule appointments for the same patients.

Eventually, leadership opted to partner with the MGH Laboratory for Computer Science to create a stand-alone web-based application that connects to the hospital's EMR. The resulting tool, called "OptIn" (for "Optimized Infusion"), retrieves real-time data from the EMR, including basic patient, infusion, and clinician information; checks for available appointment times; runs the heuristic; and offers up to four appointment options in sets of two for the

**Figure 3.** (Color online) Average pre- and postimplementation scheduled utilization.
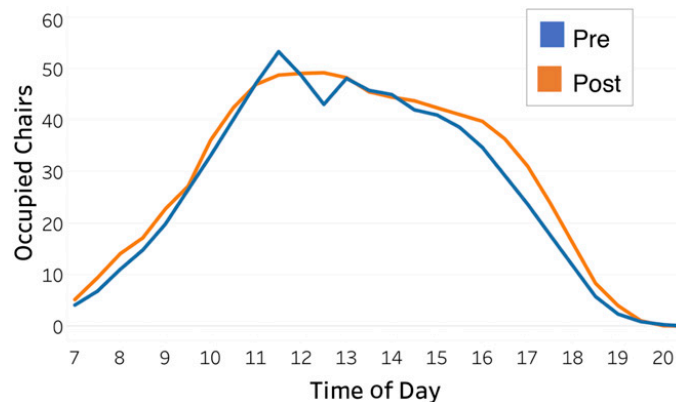


*Note.* Average number of scheduled occupied chairs by hour into the day. "Pre" is before implementation of the OptIn tool (October 2016–April 2017); "Post" is after implementation of the tool (January 2018–December 2018).

**Table 1.** Average pre- and postimplementation scheduled utilization, by the numbers.

| Period | Median patients/day | Median chair hours/day | Median peak |
|---|---|---|---|
| Pre | 134 | 375.5 | 54 |
| Post | 146 | 402 | 53 |

*Note.* Differences between "Pre" and "Post" distributions of these three measures (number of patients per day, number of chair-hours per day, and peak number of chairs) are statistically significant at the 0.05 level via a Kolmogorov–Smirnov test for equality of distributions.

selected appointment date (whenever possible), ranked as described above. The scheduler, who speaks with the patient, is instructed to offer these appointments. If none of these work for the patient, the scheduler may choose to override OptIn and select a different set of appointments. The final choice is then placed in the EMR.

OptIn was rolled out in phases in the spring of 2017. Since then, its implementation has had a marked effect on the scheduled utilization (see Figure 3 and Table 1). There are several observations of (statistical[2]) significance:

1. The average number of scheduled patients per day has increased by 9.5%, from 129.7 to 142, and the average number of scheduled chair hours per day increased by 8%, from 363.5 to 392.8.

2. Previously, the peak average chair occupancy occurred at 11:30 a.m. Following OptIn, this has decreased by an average of 4.6 chairs, a noticeable reduction at the previously most congested time of the day.

3. Previously, the infusion center had a standard lunch break time, leading to a decrease in scheduled chair occupancy around noon that rebounded at 1:00 p.m. On the basis of the findings of this project, staff lunch breaks were changed to be staggered; as a result, after implementation, the average occupancy at 12:30 p.m. has increased by 6.6 chairs.

4. Average scheduled occupancy has increased in the early morning and afternoon hours. For example, at 5:00 p.m., the average occupancy has increased by 6.9 chairs. This suggests that the OptIn approach has made more effective use of the underutilized chair time.

5. Even with the increased visit and chair hour volume, the average chair hours that each patient actually spends in the infusion unit decreased from 3.6 hours to 3.1 hours, signaling the reduction of operational delays as resources are used more evenly throughout the day.

These results reflect *all* scheduled appointments during calendar year 2018, regardless of whether OptIn system's recommendations were followed. The average adoption rate of the system is approximately 74%.

We close this section by commenting that this set of results is only a small part of the ongoing discussion surrounding the use of this work in the infusion center. Indeed, related analysis of staffing levels, actual versus scheduled utilization, and compliance with the OptIn system, among many other interrelated topics, are all critical to the continued success of the effort.

## 3.5. Discussion

The case of the infusion center scheduling redesign involves several features that are relevant to management of outpatient resources more broadly. One such aspect relates to the fundamental task at hand—namely, that it is necessary to build a *process* and not just a *model*. In particular, it is critical to develop an intimate knowledge of the underlying processes in order to understand precisely what the impact of underlying model assumptions are and, consequently, whether the model's implications are valid. For example, a model that fails to incorporate information about the "linked" infusion appointments would not realistically capture a central, complicating aspect of the system.

Furthermore, it is necessary to understand proactively how a proposed solution might affect the system as a whole. For example, what burden does this place on existing staff, and how

does it incorporate into their workflows? As discussed in Section 2, these are absolutely critical components to address in order to ensure that a solution is not only implemented but also can actually be *sustained* over time. Along the same lines, designing tools that can interact with a health system's EMR is crucial. Although the current version of OptIn can only read information from the EMR, its easy-to-use interface is a good incentive for schedulers to use the tool; this is especially true for new staff, who are not yet familiar with all the intricacies of infusion scheduling. For members of the operations research community, such work should not be seen as a roadblock or obstruction to their creative work; instead, we see this as an opportunity to design effective, practical, and innovative solutions that can also influence the direction and scope of the operations research field.

Another aspect of the infusion center example relates to the management of patients' preferences and overall capacity. In particular, given the scarce nature of capacity (i.e., resources), the behavior of granting patients' preferences for midday appointments has led to a system that is congested midday and otherwise underutilized. In other words, the decision made by individual patients can have, in aggregate, an adverse effect on the system as a whole. The approach taken herein suggests one way of mitigating some of these effects while increasing access to these scarce resources to a larger number of patients than was previously possible. Finally, it is important to note that this approach actually *increases* the number of appointments that can be scheduled under the primary nursing model; consequently, this approach has potential net benefits for *all* patients, not just those who might otherwise not have been able to be scheduled.

Finally, we conclude this vignette by noting the findings from this have been deployed in several other settings at MGH. On one front, this approach has been used to change the scheduling practices at outpatient infusion centers at two affiliated hospitals. On another front, it has also helped drive the creation of an entirely new infusion center at MGH (Ghobadi et al. [32]) for other infusions not necessarily related to oncology.

## 4. Vignette 2: Discharge Prediction and Inpatient Capacity Management

In the second vignette, we turn our attention to the inpatient setting and the problem of managing inpatient capacity. In particular, we focus on a redesign of the *discharge process* for surgical patients at MGH by predicting hospital discharges in the next 24 hours. For complete details of the relevant work, see Safavi et al. [74] and Zanger [91].

### 4.1. Genesis of Work

As hospitals continue to grapple with increased demand for their services together with intense pressure to control costs, operational efficiency has become a key element to provide timely access to existing inpatient resources. This is particularly true for hospitals such as MGH that frequently operate at or near operational capacity. The inpatient floor beds of the hospital are the common destination for the major sources of patient inflow to the hospital. Without beds available, this flow is disrupted, and the hospital's ability to serve patients is threatened.

In the United States, patients in the ED may spend many hours or even days waiting for a bed, a phenomenon that has been associated with delays in care and increased mortality. In parallel, patients in the intensive care unit (ICU) who have recovered from their critical illness cannot be transferred to the inpatient floor, causing the ICU to fill up with those who do not need its services, thus denying precious beds to the sickest patients in the hospital. A similar chain reaction occurs for patients who are in the postanesthesia care unit (PACU) recovering after surgery. After their recovery, patients in the PACU and requiring admission to the hospital are transferred to the inpatient floors, but if beds are unavailable, they remain in place, denying those beds to upstream patients completing their surgeries in the operating room. As a result, the operating room becomes stagnant, and new patients cannot receive

surgery. Finally, large academic centers provide a regional service by receiving patients from smaller outlying community hospitals when the patient's needs exceed their available capabilities. Large medical centers may be the only location where the expertise and technology exists that they need. Without beds available at the large medical center, this critical route is closed off, and patient care may be significantly delayed.

With these pressures being experienced on a near daily basis, bed capacity management has become critical for large medical centers to fulfill their mission to serve patients. To continue to fulfill the demand for their services, hospitals must increase the efficiency with which they transition patients through their stay. One critical aspect of hospital outflow is the patient discharge process. This is the process whereby patients who are nearly ready to leave the hospital receive the remaining care they require just prior to being safe for transition home or to a posthospital facility. This critical release valve for the congested inpatient floors is a major focus of large medical centers. However, one of the central complicating factors in addressing discharges is that the discharge process is often managed on a floor-by-floor or team-by-team basis; this is in stark contrast to elective surgical admissions, which form a large part of the hospital inflow, that are often scheduled in advance and centrally managed.

To be performed efficiently, the discharge process requires transparency, prioritization, and coordination. There must be transparency into which patients should be considered for discharge each day and which barriers remain to be addressed. Furthermore, there must be a scheme to prioritize the remaining care that should take place so that the resources of the care team can be allocated in a timely manner that enables discharge the same day. Finally, to perform this work, there must be coordination across a multidisciplinary team that is unified in its understanding of which patients to focus on and what to focus on doing.

**4.1.1. Discharge Processes at MGH.** In its current state, MGH relies on diverse, multidisciplinary clinical teams to perform the discharge process without a standard, data-driven process. Clinical teams manually perform the task of identifying which patients are candidates for discharge each day. This process is time consuming and based on the subjective input of team members. Despite discharge requiring numerous team members to perform tasks and coordinate their efforts in a timely manner, the current process does not allow for transparency in which patients have been identified or their remaining barriers to discharge. Furthermore, this manual process is time consuming and takes teams away from numerous other care priorities such as attending to acutely ill patients on the inpatient floors. All of this is further complicated by the fact that surgical patients at MGH are on over a dozen inpatient floors spread across five buildings, and any one attending physician can have patients in multiple physical locations, each being cared for by a different nursing team.

As a consequence, the discharge process is highly inefficient. Without a standardized process to identify patients and their barriers, the likelihood for intraday and interday delays in patient discharge is substantial. Moreover, without systematically identifying expected discharges each day, it is impossible for care teams to study their discharge processes and performance relative to an expected standard. This means that teams and hospital administrators are blind to their performance and the factors that lead to a lack of discharge, making improvement extremely challenging.

## 4.2. Model Development

To address these challenges, we developed a *discharge prediction app* ("DPA"). Each day, the DPA predicts which patients will leave the hospital within 24 hours, thereby providing transparency into which patients are candidates for early discharge and enabling clinical teams to leverage its information to increase transparency, prioritization, and coordination in the discharge process. In this section, we detail the development of the DPA predictive model, highlighting underlying principles in its development and execution.

**4.2.1. Principles of Model Design and Development.** To be successful, the DPA's model had to receive inputs that were aligned with those that clinical decision makers use to identify patients who are candidates for discharge. Moreover, such a model design would need to create outputs that are aligned with the clinical intuition that decision makers use and are thus aligned with downstream actions by clinicians. Without such an alignment, the model not only risks lacking the inputs to make accurate predictions but also may create an output that is not interpretable by clinical team members or connected to important downstream actions. In other words, consistent with the principles of Section 2, the DPA needed to be developed with a focus on existing workflows and data sources while ensuring that the data were interpreted in its appropriate context.

The model was developed in two stages along these lines. First, clinical leaders on the project team helped serve as proxies to the clinical decision makers responsible for the discharge process. These team members developed a framework for the selection of data inputs such that each data element selected would be relevant to understanding whether the patient was a candidate for discharge. As a result, data inputs superfluous to the discharge process needed to be eliminated. The framework classified data as either representing *milestones* marking the patient's progression toward discharge or *barriers* that must be surmounted prior to the patient being safe to transition out of the hospital. For example, the ability of the patient to eat a regular diet after abdominal surgery represents a significant milestone in the patient's recovery from surgery and a behavior that marks a sustainable posthospital state in which the patient can be independent of hospital care. By contrast, when a patient has a fever, these data are interpreted with great caution by clinicians, as they may represent the early signs of a new infection and require further diagnostic testing and evaluation for which the patient most often needs to remain in the hospital.

**4.2.2. Technical Considerations and Model Output.** Using the framework of "milestones" and "barriers" led to hundreds of input feature variables sourced from the EMR, including demographic, administrative, clinical, and environmental data. These data were used to frame a prediction problem of predicting the likelihood of any individual patient to be discharged within 24 hours. We considered a variety of possible classes of binary-classification-based predictive models for such a task, such as (regularized) logistic regression and feed-forward neural networks (multilayer perceptron) with slightly more complex architectures. We ultimately selected a neural network model using standard software packages and cross-validation criteria of the area under the receiver-operating characteristic curve (AUC), a measure of a binary classification model's discriminative ability.

The resulting model, which we refer to as the DPA, assigns to every patient in the hospital a prediction score that is correlated with his or her likelihood of discharge within 24 hours. Each morning, the DPA provides clinical teams with the following:

1. a ranked list of patients in order of their likelihood of discharge that day that can be sorted by hospital floor, clinical service, and responsible physician;
2. a ranked list of patients in order of their likelihood of discharge the next day;
3. a daily prediction of the total number of patients that will be discharged that day; and
4. a comprehensive list of barriers to discharge, shown for each patient, such as abnormal laboratory results, imaging studies that are pending completion, physical therapy needs, or insurance approval issues.

The initial list of barriers shown was generated as follows. Among those barriers present, which ones (if resolved) result in a *higher* likelihood of being discharged? By looking one at a time over all possible barriers, a list of barriers can be generated for every patient. This scheme, although simple, provides a first approximation to interpreting the results of the DPA.

**4.2.3. Further Model Development.** In the second stage of the model development, we sought to include the multidisciplinary clinical teams who would ultimately be the end users of the DPA. The aim was to elicit differences existing between the model's outputs and the intuition of clinical team members, including physicians, nurses, physical therapists, case managers, social workers, and others. Each of these team members holds a unique expertise in perceiving the many different aspects of what may prevent a patient from being discharged from the hospital.

During a 100-day period, the output of a preliminary version of the DPA was provided to members of this multidisciplinary team during their daily morning conference. Team members were instructed to review the list and determine whether they would agree that the patient was a candidate for discharge that day. Moreover, for patients who were not identified as likely to be discharged, the team was instructed to identify candidates who they saw as likely ready for discharge that day. Finally, project and clinical team members discussed for each of these patients which barriers to discharge were relevant and compared these with the barriers displayed by the DPA. Over the course of this period, over 40 additional data items were identified that were added to the predictive model, each of which was aligned with the clinical intuition of the members of the clinical team.

Another focus of the prepilot stage was in understanding whether the barriers to discharge displayed by the DPA were represented in a manner that was interpretable, relevant, and actionable for the clinical team members. Given the limited time of clinical team members to devote to discharge assessment and execution, it was critical to display only highly interpretable and actionable barriers. For example, physical therapy–related barriers are highly relevant to discharge decisions. In addition to making for a stronger and more easily interpretable model, the project team needed to understand the broader process and workflow within which the DPA would be incorporated. This knowledge was important to making critical implementation decisions.

For example, the time at which the DPA would need to provide its output was based on when key decision points were made throughout the day related to discharge. There was a balance that needed to be struck between having the output available for team members to use during these critical decision points versus including more information about the patient as it became available later in the day. The later the prediction, the more information the model would have to work with and, potentially, the stronger the prediction. It is important to note that, however, such a model would be accurate though useless to decision makers on the clinical team. Ultimately, it was determined that the DPA would need to be available for when the care team rounded at 6:00 a.m. This time represented the one critical time point where the nurse and physicians would meet for each patient to discuss the plan for the day, including which candidates should be discharged and what care would be required to get them to that transition point.

**4.2.4. Final Model Selection.** Based on the two stages of model development, the final version of the DPA used more than 900 feature variables. Consistent with earlier model development, the model architecture was selected using cross-validation, resulting in a relatively simple neural network architecture with a single hidden layer. All together, the final model selected had an out-of-sample AUC of 0.84.

Furthermore, it is worth noting that there are a variety of competing metrics of interest that need to be used to inform the final choice of model. These should be considered in the context of the process described above. Specifically, the question is how to use the model to generate the list of predicted discharges to the clinical teams and what the relevant performance metrics are. The DPA generates a likelihood prediction $\in [0, 1]$ for each patient; we must either choose to list a fixed number of patients who we believe are likely to be discharged or present all patients with a likelihood above some threshold. Each of these algorithmic choices has practical implications for metrics such as false-positive rate, false-negative rate, and so on, and these need to each be considered carefully in turn to guide the final results presented to clinicians.

With the selected neural network model, we elected to present all patients with a score above a threshold $\tau$, where $\tau$ was chosen so that the corresponding true-positive rate was approximately equal to the positive predictive value; equivalently, $\tau$ was chosen so that the number of predicted positives (true positives and false positives) was equal to the number of actual positives (true positives and false negatives). Based on this criterion, the final threshold chosen was $\tau = 0.5$. With such a threshold, the out-of-sample true-positive rate was 0.512, the true-negative rate was 0.855, the positive predictive value was 0.565, and the negative predictive value was 0.827. Each of these metrics has its own interpretation; for example, the true-positive rate of 0.512 means that on a patient's day of discharge, the DPA correctly identifies that patient as being likely to be discharged 51.2% of the time.

From the use of the DPA, there are a variety of specific hospital metrics that can be impacted and are measurable. Examples of these include

1. the rate of early morning discharges,
2. the overall average hospital length of stay,
3. the rate of critical ED congestion events,
4. the total number of outside hospital transfer diversions,
5. the frequency of ED boarding (patients waiting in the ED for a floor bed more than two hours after the decision to admit),
6. the frequency of ICU transfer delays, and
7. the frequency of PACU transfer delays.

We conclude the discussion of model selection by pausing to reflect on the fundamental prediction task at hand. In particular, per the setup above, the DPA predicts whether a patient will be discharged; that is, it predicts the decision to discharge a patient. It is worth noting that, this is *not the same* as predicting whether a patient is *clinically ready* to be discharged from the hospital. This suggests that "false-positive" predictions from the DPA might actually reveal opportunities for system-level improvements in discharge processes. Indeed, as discussed later regarding outcomes, we found this to be the case.

**4.2.5. Related Work.** There has been significant historical interest in the problem of predicting how long patients will stay in the hospital (Azari et al. [15], Carter and Potts [24], Gustafson [37], Harutyunyan et al. [38], Huang et al. [43], Jiang et al. [46], Liu et al. [59, 60], Mohammadi et al. [65], Morton et al. [66], Robinson et al. [73], Tu and Guerriere [86], Walczak et al. [89]). Although these models can provide insight into the drivers of long length of stay, they can be difficult to translate into actionable, operationally focused tools that can guide the day-to-day management of bed capacity. In light of those challenges, there has also been work to predict likelihood of discharge within some time window (Barnes et al. [16], Levin et al. [56]).

The work described herein differs from the related work in several key ways. First and foremost, our approach has focused on the design of an operational tool that can guide day-to-day decision making. Furthermore, the framework of milestones and barriers to discharge allows us to present results that directly reflect clinically meaningful indicators for the discharge process as opposed to raw, unprocessed clinical information. Finally, this work utilizes a significant and diverse array of data sources for the prediction task using the same EMR data that drive clinical decision making.

## 4.3. Implementation and Ongoing Work

The DPA was implemented in several phases with differing objectives in each instance. As discussed above, the prepilot phase focused on using a preliminary version of the model to provide output to a multidisciplinary team of key stakeholders representing different aspects of the discharge process. This process was instrumental in creating a dialogue between various

stakeholders and in gathering feedback on which data elements are (and are not) reliable indicators of clinical measures.

**4.3.1. Pilot Implementation.** Following the model development phases, the improved DPA was deployed without providing predictions to end users. That is, it did not require anything from frontline clinical teams, including using the DPA or altering their workflow. The aim of this pilot was to measure the potential impact of the DPA by quantifying the gap between predicted discharges and actual discharges and subsequently identifying the root cause of these mismatches. The mismatches were subcategorized (using a manual case review by clinicians) into those that represented acute changes in the clinical status of the patient appropriately requiring additional days in the hospital versus scenarios in which the care team could have discharged the patient with an improved process.

One of the most important, though not initially anticipated, outcomes of the DPA was that it created an entirely new "database" that could enable identification and classification of problems plaguing the discharge process. The initial conception of the value of the tool was around its real-time use to impact discharge efficiency. However, the DPA also identified dozens of cases each day in which a patient was not discharged despite being predicted as such by the tool (i.e., "false positives"). Thus, the DPA automatically created a cohort of patients for further study into why patients did not get discharged on the day they were expected to be. Such an analysis of this cohort is particularly useful for understanding and quantifying variation in clinical pathways for patients (see Safavi et al. [74]).

This retrospective analysis of the tool's output and its mismatch within the reality of clinical practice proved invaluable to identifying process improvement themes that were previously invisible to the hospital. As such, it provided the project team the ability to study individual groups of patients within the hospital being taken care of by different specialties of teams, each of which had different discharge workflows and processes and each with their own strengths and weaknesses. Without requiring anything from the clinicians, the DPA allowed the project team and, ultimately, senior leadership to study the discharge inefficiencies in each service prior to the implementation of the tool. Although some of these challenges can be addressed using the DPA, others require the application of different sorts of process improvement tools.

**4.3.2. Ongoing and Future Work.** Following the pilot implementation, there has been work on a variety of fronts to improve the DPA and expand its scope. Here, we consider a few aspects: transition to a real-time predictive tool, use by clinical teams to guide workflow, and expansion of the patient population for the DPA.

As noted earlier, the DPA was initially designed to provide predictions early in the morning (about 6:00 a.m.) for use in morning rounds by clinical teams. However, there are other potential uses for the DPA at different times in the day. For example, the DPA could provide another output at 8:30 a.m., when multidisciplinary rounds would take place involving the case managers, social workers, physical therapists, and physicians. As another example, at 3:00 p.m., the DPA could provide clinical team members with an opportunity to touch base about patients who were expected to leave and their status as well as what challenges arose during the day and whether the patient would be a candidate for early morning discharge the following day. Transitioning to such an approach requires using *live* data streams that are continually updated, as opposed to the nightly updates that have been used in our previous DPA work. Such a transition requires a variety of work and technical support in validating data and understanding the database structures and design necessary to facilitate such a real-time tool.

On another front, we are actively working to incorporate the DPA into the workflow of clinical team members. As part of this, the DPA has been piloted with inpatient floors' nursing directors and case managers with the aim of having these team members use its output to actually impact the discharge process. This required these team members to incorporate a new tool into their daily workflow, pay close attention to its results, and use the output to spark

discussion among the multidisciplinary team. Furthermore, in combination with the transition to a real-time DPA, we are also working with physicians to further validate the DPA and integrate it into their workflows.

Finally, we have also been working to expand the DPA to a larger patient population beyond surgical patients, such as general medicine patients or nonprocedural patients cared for by surgical teams. This raises a variety of interesting clinical and technical challenges, as there can be significantly more clinical variability in patients' care pathways during hospitalization for the patient population at large.

## 5. Vignette 3: Managing the Inpatient-Outpatient Relationship for Heart Failure Patients

In this final vignette, we turn our attention to the connection between resources in the outpatient and inpatient settings, specifically in the context of managing chronically ill heart failure patients. For complete details of the relevant work, see Al-Meer [5] and Furtado [31].
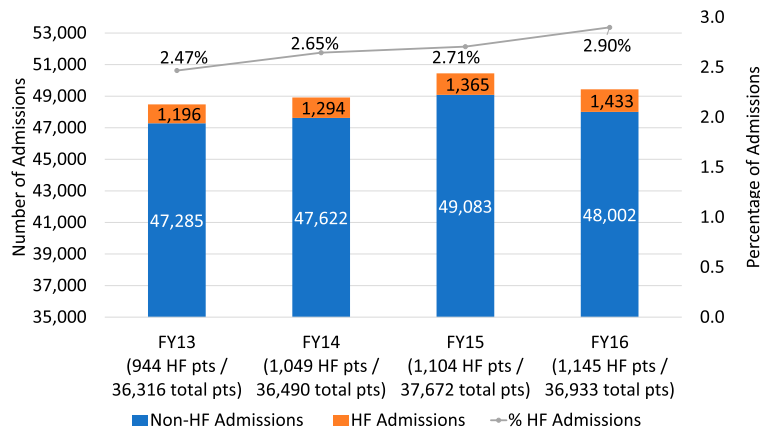
### 5.1. Heart Failure

Heart failure (HF) is a chronic medical condition that affects 6.5 million adults in the United States (Mozaffarian et al. [67]). This progressive condition affects the heart's ability to either fill with or pump blood. HF commonly manifests with symptoms including shortness of breath, fatigue, and swelling. The implications of HF can be devastating as the condition can affect not only the entire circulatory system but also vital organs. Because of this, patients' health generally deteriorates over time after experiencing HF. Unfortunately, there is no cure for HF, so it is important for healthcare practitioners to properly manage the condition to mitigate its negative effects.

**5.1.1. Heart Failure in the United States.** The HF population in the United States has steadily increased over time and is projected to increase by 46% from 2012 to 2030 to over 8 million people (Benjamin et al. [19], Heidenreich et al. [39]). HF patients generated more than 57 million HF-related hospital admissions from 2001 to 2014 (Akintoye et al. [4]). Hospital length of stays for HF admissions are also longer than length of stays for other medical condition admissions (Houchens et al. [41], Ziaeian et al. [95]). This has contributed to the ever-increasing costs of HF treatment in the United States, which are projected to inflate to $70 billion dollars annually by 2030 (Heidenreich et al. [39]). In an effort to control healthcare costs, the Affordable Care Act introduced the Hospital Readmission Reduction Program (HRRP) in 2012. This program provides financial incentives for hospitals to reduce readmissions. As a result, many health systems have dedicated significant resources to reduce readmissions, especially for HF, which has the highest rate of 30-day readmissions according to Medicare studies (Jencks et al. [45]).

**5.1.2. Heart Failure at Massachusetts General Hospital.** MGH has approximately 50,000 hospital admissions every year, with 2%–3% of these admissions having HF as a primary diagnosis. This percentage has steadily increased over recent years (see Figure 4). Furthermore, HF hospitalizations account for a relatively large percentage (3%–4%) of the hospital's bed-days (see Figure 5). This is in part because HF admissions have longer average length of stays compared with non-HF admissions (8.7 versus 5.9 days, respectively, in 2016). Beginning in 2012, Section 3025 of the Affordable Care Act and the HRRP began reducing payments to hospitals with excess readmissions. For this reason, many improvement efforts around the nation are focused on reducing readmissions. However, MGH observes that only 20% of their HF admissions are readmissions. It is therefore important for MGH to find ways to reduce all HF admissions. Furthermore, HF is classified as an *ambulatory care–sensitive condition*, which means that effective ambulatory care can reduce utilization of inpatient

**Figure 4.** (Color online) HF admissions at MGH, FY2013–FY2016.



resources. This claim is supported by many specialized HF clinics and HF management programs reporting reduced hospitalization rates and shorter length of stays.

## 5.2. Outpatient Access for Heart Failure at MGH

Analysis of the care paths that MGH HF patients follow indicates that 90% of HF admissions originate from ED visits, with 92% of these HF ED visits resulting in hospital admissions. Examining the events preceding each HF admission from the ED, we see that in 51% of these cases, the patient had no completed outpatient appointment (of any kind) 2 weeks prior to his or her hospitalization. For patients who had a scheduled appointment during the 2 weeks prior to their hospitalization, 21% of them failed to complete the appointment. This absence of outpatient activity immediately preceding the HF hospitalization indicates an opportunity to capture these patients through ambulatory interventions.

We also used scheduling and hospital admission data to examine scenarios in which patients call in the week prior to their subsequent HF admission in order to schedule an outpatient appointment. These individual cases can be categorized into two groups depending on the appointment date: the appointment date is (1) before or (2) after the patient's subsequent HF admission. We first identify that the time between the initial call and the HF admission date between these two groups is not statistically different. This suggests that there is no significant difference in the urgency of the appointment requests in the two groups. The scheduling lead time (the time between the call and the appointment date), on the other hand, is statistically different between the two groups. In particular, cases in which the appointment date falls after the HF

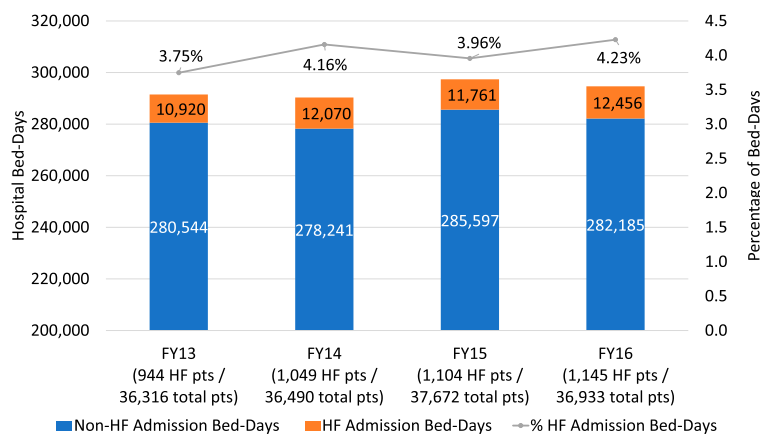**Figure 5.** (Color online) HF admission bed-days at MGH, FY2013–FY2016.

**Table 2.** Summary of heart failure model variables.

| Category name | Variable name |
| --- | --- |
| Patient characteristics | *Age Squared* |
| | *Gender* |
| | *English Speaker* |
| | *Marital Status—Single* |
| Medications | *Total Number of Medications* |
| | *HF-related* |
| | *Antipsychotic* |
| Clinical indicators | *Blood Pressure* |
| | *Pulse* |
| | *Ejection Fraction* |
| | *≥1 EF Ever Recorded* |
| Hospital utilization | *Cardiology Outpatient Visits* |
| | *Overall HF Outpatient Visits* |
| | *Time Since HF Diagnosis* |
| | *Number of HF Hospital Admissions* |
| | *Number of non-HF Hospital Admissions* |
| | *Time Since Last EF Measure* |
| Socioeconomic factors | *History of Substance Use Disorder* |
| | *Estimated Income* |
| | *Distance to PCP Clinic* |
| Patient engagement | *Resident Owned* |
| | *iCMP Enrollment* |
| | *Number of Canceled/No-Show Appts* |
| | *Patient Gateway Activity* |
| Comorbidities | *History of Diabetes* |

*Note.* iCMP, or the Integrated Care Management Program, provides care coordination services to patients with complex health issues at MGH.

admission generally had longer lead times (average of 14 days versus 1 day). These results suggest that long outpatient appointment lead times are a potential contributing factor to HF admissions.

## 5.3. Predicting Heart Failure Hospital Admissions

To implement ambulatory interventions for reducing HF admissions, it is important to identify high-risk HF patients in the primary care population. To this end, we developed a logistic regression model with 32 covariates (see Table 2) to predict six-month hospital admission risk. The model performance was evaluated using the standard receiver operating characteristic AUC. The AUC of our model on out-of-sample data is 0.78, which outperforms other comparable hospitalization risk models (Amarasingham et al. [6], Greenwald et al. [34] Levy et al. [57], O'Connor et al. [70]).

## 5.4. Implementation and Next Steps

Upon review of the prediction model results, we believe using our prediction scores to flag high-risk HF patients will allow primary care staff to more effectively triage and manage the patients. These flags in the electronic health record will not only alert practice staff about a patient's HF condition when there is an encounter but also allow the practice to proactively reach out to high-risk patients who may otherwise be out of touch with the practice. Stratifying patients by their risk score will also help identify candidate HF patients for additional healthcare services such as telemonitoring. We hypothesize that these cumulative efforts will help HF patients be more engaged in an ambulatory setting and reduce their likelihood of admission.

Although we have a clear objective of flagging high-risk HF patients, implementing these flags in the EHR is surprisingly challenging. There is no simple feature in the EHR that enables new visible flags in the banner of patients' charts, and creating such a flag requires both institutional approval and technical development that can span over a year from start to finish. In the meantime, we have developed a workaround that enables us to display a manually written free-text note in a patient's chart that identifies him or her as a high-risk HF patient. This is far from ideal, however, as it requires someone to insert the free-text note into each patient's chart individually. These types of technical challenges and obstacles further highlight the fact that analytics are only a means to an end, as described in Section 2. Without establishing the necessary interfaces for the analytics to be deployed and used, it is highly unlikely that any meaningful impact will be achieved.

**5.4.1. Next Steps.** Sharing our findings with stakeholders in both the primary care and cardiology departments has supported further collaborative efforts to improve comanagement of HF patients between primary care physicians (PCPs) and cardiologists, redesign triaging protocols to more effectively escalate patients' HF-related complaints, and enable knowledge sharing between the departments to establish HF management best practices.

## 6. Final Remarks

A tremendous number of problems that health systems face can be addressed using data analytics and operations research techniques. However, most of the relevant literature in this space typically ends at theoretical modeling, simulations, and analyses, but it does not reach large-scale field implementations with tangible impact on practice.

This tutorial attempts to offer principles and key success drivers underlying a formal project management framework established in partnership between MGH and MIT. The partnership has used the aforementioned framework to successfully complete multiple projects over the span of many years that have yielded system-level implementations driven by academic analytics research.

Finally, the tutorial attempts to stimulate community-wide discussions about how collaborative models can be scaled to increase the academic footprint and impact of analytics-related work on the healthcare industry.

### Acknowledgments

### Endnotes

[1]Examples include chemotherapies being administered in high-risk clinical trials.

[2]All findings referenced herein are statistically significant at the 0.05 level as measured via the Kolmogorov–Smirnov test for equality of distributions.

### References

[1] H. J. Aboumater, L. E. Winner, R. O. Davis, P. B. Trovitch, M. M. Berg, K. M. Violette, W. A. Messersmith, K. K. Maylor, and C. U. Lehmann. No time to waste: Decreasing patient wait times for chemotherapy administration using automated prioritization in an oncology pharmacy system. *American Journal of Managed Care* 14(5):309–316, 2008.

[2] N. Agarwal, I. Ashlagi, E. Azevedo, C. Featherstone, and O. Karaduman. What matters for the productivity of kidney exchange? *AEA Papers and Proceedings* 108(May):334–340, 2018.

[3] Z. Agur, R. Hassin, and S. Levy. Optimizing chemotherapy scheduling using local search heuristics. *Operations Research* 54(5):829–846, 2006.

[4] E. Akintoye, A. Briasoulis, A. Egbe, S. M. Dunlay, S. Kushwaha, D. Levine, L. Afonso, D. Mozaffarian, and J. Weinberger. National trends in admission and in-hospital mortality of patients

with heart failure in the United States (2001–2014). *Journal of the American Heart Association* 6(12):e006955, 2017.

[5] M. A. Al-Meer. Reducing heart failure admissions through improved care systems and processes. Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge, 2017.

[6] R. Amarasingham, B. J. Moore, Y. P. Tabak, M. H. Drazner, C. A. Clark, S. Zhang, W. G. Reed, T. S. Swanson, Y. Ma, and E. A. Halm. An automated model to identify heart failure patients at risk for 30-day readmission or death using electronic medical record data. *Medical Care* 48(11):981–988, 2010.

[7] R. Anderson, I. Ashlagi, D. Gamarnik, M. Rees, A. E. Roth, T. Sönmez, and M. U. Ünver. Kidney exchange and the alliance for paired donation: Operations research changes the way kidneys are transplanted. *Interfaces* 45(1):26–42, 2015.

[8] E. Ang, S. Kwasnick, M. Bayati, E. L. Plambeck, and M. Aratow. Accurate emergency department wait time prediction. *Manufacturing & Service Operations Management* 18(1):141–156, 2016.

[9] I. Ashlagi and A. E. Roth. New challenges in multihospital kidney exchange. *American Economic Review: Papers & Proceedings* 102(3):354–359, 2012.

[10] B. Ata, J. Friedewald, and A. C. Randa. Organ transplantation. T. Dai and S. Tayur, eds. *Handbook of Healthcare Analytics: Theoretical Minimum for Conducting 21st Century Research on Healthcare Operations.* John Wiley & Sons, Hoboken, NJ, 187–215, 2018.

[11] B. Ata, A. Skaro, and S. Tayur. OrganJet: Overcoming geographical disparities in access to deceased donor kidneys in the United States. *Management Science* 63(9):2776–2794, 2017.

[12] T. Ayer, O. Alagoz, and N. K. Stout. A POMDP approach to personalize mammography screening decisions. *Operations Research* 60(5):1019–1034, 2012.

[13] T. Ayer, C. Zhang, A. Bonifonte, A. C. Spaulding, and J. Chhatwal. Prioritizing Hepatitis C treatment in U.S. prisons. *Operations Research* 67(3):853–873, 2019.

[14] T. Ayer, C. Zhang, C. Zeng, C. C. White III, R. J. Vengazhiyil, M. Deck, K. Lee, D. Moroney, and Z. Ozkaynak. American Red Cross uses analytics-based methods to improve blood-collection operations. *Interfaces* 48(1):24–34, 2018.

[15] A. Azari, V. P. Janeja, and A. Mohseni. Predicting hospital length of stay (PHLOS): A multi-tiered data mining approach. Vreeken J, ed. *2012 IEEE 12th International Conference on Data Mining Workshops. IEEE*, Piscataway, NJ, 17–24, 2012.

[16] S. Barnes, E. Hamrock, M. Toerper, S. Siddiqui, and S. Levin. Real-time prediction of inpatient length of stay for discharge prioritization. *Journal of the American Medical Informatics Association* 23(e1):e2–e10, 2015.

[17] H. Bastani and M. Bayati. Online decision-making with high-dimensional covariates. *Operations Research*, forthcoming, 2019.

[18] N. Ben-Zvi. OR$^2$: Operations research applied to operating room supply chain. Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge, 2014.

[19] E. J. Benjamin, S. S. Virani, C. W. Callaway, A. M. Chamberlain, A. R. Chang, S. Cheng, S. E. Chiuve, et al. Heart disease and stroke statistics - 2018 update: A report from the American Heart Association. *Circulation* 137(12):e67–e492, 2018.

[20] D. P. Bertsekas and J. N. Tsitsiklis. *Neuro-Dynamic Programming.* Athena Scientific, Belmont, MA, 1996.

[21] D. Bertsimas, V. F. Farias, and N. Trichakis. Fairness, efficiency, and flexibility in organ allocation for kidney transplantation. *Operations Research* 61(1):73–87, 2013.

[22] D. Bertsimas, A. O'Hair, S. Relyea, and J. Silberholz. An analytics approach to designing combination chemotherapy regimens for cancer. *Management Science* 62(5):1511–1531, 2016.

[23] D. Bertsimas, M. V. Bjarnadóttir, M. A. Kane, J. C. Kryder, R. Pandey, S. Vempala, and G. Wang. Algorithmic prediction of health-care costs. *Operations Research* 56(6):1382–1392, 2008.

[24] E. M. Carter and H. W. W. Potts. Predicting length of stay from an electronic patient record system: A primary total knee replacement example. *BMC Medical Informatics and Decision Making* 14(1):26–38, 2014.

[25] T. Cayirli and E. Veral. Outpatient scheduling in healthcare: A review of literature. *Production and Operations Management* 12(4):519–549, 2003.

[26] G. Chabot and M. Fox. The creation of a patient-classification system in an outpatient infusion center setting. *Oncology Nursing Forum* 32(3):535–538, 2005.

[27] T. C. Y. Chan and V. V. Mišić. Adaptive and robust radiation therapy optimization for lung cancer. *European Journal of Operational Research* 231(3):745–756, 2013.

[28] A. Cohn, S. Root, C. Kymissis, J. Esses, and N. Westmoreland. Scheduling medical residents at Boston University School of Medicine. *Interfaces* 39(3):186–195, 2009.

[29] L. Dorland, C. Hardecki, B. Hess, C. McMahon, and L. A. Schneider. *Cancer Infusion Therapy Patient Handbook*. Rogel Cancer Center, University of Michigan, Ann Arbor, 2018.

[30] P. C. Fuloria and S. A. Zenios. Outcomes-adjusted reimbursement in a health-care delivery system. *Management Science* 47(6):735–751, 2001.

[31] J. D. Furtado. Applications of healthcare analytics in reducing hospitalization days. Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge, 2018.

[32] K. Ghobadi, A. C. Zenteno, A. R. Marshall, P. F. Dunn, R. Levi, and J. H. Stone. Translating a biologic revolution into an organizational overhaul. *NEJM Catalyst* (February 5), https://catalyst.nejm.org/biologic-therapy-revolution-organizational-overhaul/, 2017.

[33] Y. Gocgun and M. L. Puterman. Dynamic scheduling with due dates and time windows: an application to chemotherapy patient appointment booking. *Healthcare Management Science* 17(1): 60–76, 2014.

[34] J. L. Greenwald, P. R. Cronin, V. Carballo, G. Danaei, and G. Choy. A novel model for predicting rehospitalization risk incorporating physical function, cognitive status, and psychosocial support using natural language processing. *Medical Care* 55(3):261–266, 2017.

[35] M. Gruber, K. Kane, L. Flack, J. Abbotoy, J. Recchio, K. Williamson, K. Horan, and P. McCarthy. A "perfect day" work redesign in a chemotherapy and infusion center. *Oncology Nursing Forum* 30(4):567–568, 2003.

[36] D. Gupta and B. Denton. Appointment scheduling in healthcare: Challenges and opportunities. *IIE Transactions* 40(9):800–819, 2008.

[37] D. H. Gustafson. Length of stay: Prediction and explanation. *Health Services Research* 3(1):12–34, 1968.

[38] H. Harutyunyan, H. Khachatrian, D. C. Kale, and A. Galstyan. Multitask learning and benchmarking with clinical time series data. *Scientific Data* 6(1):Article 96, 2019.

[39] P. A. Heidenreich, N. M. Albert, L. A. Allen, D. A. Bluemke, J. Butler, G. C. Fonarow, J. S. Ikonomidis, et al. Forecasting the impact of heart failure in the United States: A policy statement from the American Heart Association. *Circulation: Heart Failure* 6(3):606–619, 2013.

[40] J. Hiltrop. Modeling neuroscience patient flow and inpatient bed management. Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge, 2014.

[41] R. L. Houchens, D. Ross, and A. Elixhauser. Using the HCUP national inpatient sample to estimate trends. HCUP Technical Report 2005-05, Healthcare Cost and Utilization Project (HCUP), Agency for Healthcare Research and Quality, Rockville, MD, 2015.

[42] W. Hu, C. W. Chan, J. R. Zubizarreta, and G. J. Escobar. An examination of early transfers to the ICU based on physiologic risk score. *Manufacturing & Service Operations Management* 20(3): 531–549, 2018.

[43] J. Q. Huang, P. M. Hooper, and T. J. Marrie. Factors associated with length of stay in hospital for suspected community-acquired pneumonia. *Canadian Respiratory Journal* 13(6):317–324, 2006.

[44] P. Jaillet and M. R. Wagner. Online optimization—An introduction. J. J. Hasenbein, ed. *Risk and Optimization in an Uncertain World*, INFORMS TutORials in Operations Research. INFORMS, Hanover, MD, 142–152, 2010.

[45] S. Jencks, M. V. Williams, and E. A. Coleman. Rehospitalizations among patients in the Medicare fee-for-service program. *New England Journal of Medicine* 360(14):1418–1428, 2009.

[46] X. Jiang, X. Qu, and L. Davis. Using data mining to analyze patient discharge data for an urban hospital. R. Stahlbock and S. F. Crone, eds. *Proceedings of the 2010 International Conference on Data Mining*. CSREA Press, Athens, GA, 139–144, 2010

[47] M. A. Kallen, J. A. Terrell, P. Lewis-Patterson, and J. P. Hwang. Improving wait time for chemotherapy in an outpatient clinic at a comprehensive cancer center. *Journal of Oncology Practice* 8(1):e1–e7, 2012.

[48] E. H. Kaplan and E. O'Keefe. Let the needles do the talking! Evaluating the New Haven needle exchange. *Interfaces* 23(1):7–26, 1993.

[49] S. H. Kim, C. W. Chan, M. Olivares, and G. Escobar. ICU admission control: An empirical study of capacity allocation and its implication for patient outcomes. *Management Science* 61(1):19–38, 2015.

[50] S. H. Kim, C. W. Chan, M. Olivares, and G. Escobar. Association among ICU congestion, ICU admission decision, and patient outcomes. *Critical Care Medicine* 44(10):1814–1821, 2016.

[51] S. P. Kim, D. Gupta, A. K. Israni, and B. L. Kasiske. Accept/decline decision module for the liver simulated allocation model. *Healthcare Management Science* 18(1):35–57, 2015.

[52] E. K. Lee and M. Zaider. Operations research advances cancer therapeutics. *Interfaces* 38(1):5–25, 2008.

[53] E. K. Lee, F. H. Pietz, C.-H. Chen, and Y. Liu. An interactive web-based decision support system for mass dispensing, emergency preparedness, and biosurveillance. *Proceedings of the 2017 International Conference on Digital Health.* ACM, New York, 137–146, 2017.

[54] E. K. Lee, T. L. Wu, T. Senior, and J. Jose. Medical alert management: A real-time adaptive decision support tool to reduce alert fatigue. *American Medical Informatics Association Annual Symposium*, AMIA, Bethesda, MD, 845–854, 2014.

[55] I. Lennes, A. C. Zenteno Langle, M. Duk, A. Levy Carlis, M. Bloom, E. Souza, R. Levi, and D. P. Ryan. Modeling and improving chemotherapy patient flow. *Journal of Clinical Oncology* 32(30, Supplement):175, 2014.

[56] S. R. Levin, E. T. Harley, J. C. Fackler, C. U. Lehmann, J. W. Custer, D. France, and S. L. Zeger. Real-time forecasting of pediatric intensive care unit length of stay using computerized provider orders. *Critical Care Medicine* 40(11):3058–3064, 2012.

[57] W. C. Levy, D. Mozaffarian, D. T. Linker, S. C. Sutradhar, S. D. Anker, A. B. Cropp, I. Anand, et al. The Seattle heart failure model. *Circulation* 113(11):1424–1433, 2006.

[58] F. Li, D. Gupta, and S. Potthoff. Improving operating room schedules. *Healthcare Management Science* 19(3):261–278, 2016.

[59] V. Liu, P. Kipnis, M. K. Gould, and G. J. Escobar. Length of stay predictions: Improvements through the use of automated laboratory and comorbidity variables. *Medical Care* 48(8):739–744, 2010.

[60] X. Liu, M. Hu, J. E. Helm, M. S. Lavieri, and T. A. Skolarus. Missed opportunities in preventing hospital readmissions: Redesigning post-discharge checkup policies. *Production and Operations Management* 27(12):2226–2250, 2018.

[61] M. Mahmoudi, G. Caracciolo, A. Li, H. Poustchi, R. Safavi-Sohi, M. Vasighi, R. Chiozzi, et al. Multi-nanoparticle array and machine learning enable accurate identification of early stage cancer. Working paper, Harvard Medical School, Boston.

[62] Massachusetts General Hospital. Hospital overview. Accessed July 10, 2019, https://www.massgeneral.org/about/overview.aspx, 2019.

[63] M. D. McEvoy, J. P. Wanderer, A. B. King, T. M. Geiger, V. Tiwari, M. Terekhov, J. M. Ehrenfeld, W. R. Furman, L. A. Lee, and W. S. Sandberg. A perioperative consult service results in reduction in cost and length of stay for colorectal surgical patients: Evidence from a healthcare redesign project. *Perioperative Medicine* 5(3):3–12, 2016.

[64] S. T. McNichols. Reducing intraday patient wait times through just-in-time bed assignment. Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge, 2015.

[65] I. Mohammadi, H. Wu, A. Turkcan, Y. Toscos, and B. N. Doebbeling. Data analytics and modeling for appointment no-show in community health centers. *Journal of Primary Care & Community Health* 9(January–February):1–11, 2018.

[66] A. Morton, E. Marzban, G. Giannoulis, A. Patel, R. Aparasu, and I. A. Kakadiaris. A comparison of supervised machine learning techniques for predicting short-term in-hospital length of stay among diabetic patients. X. Chen, G. Qu, P. Angelov, C. Ferri, J. Lai, and M. A. Wani, eds. *Proceedings of the* 13th International Conference on Machine Learning and Applications (ICMLA). IEEE, Piscataway, NJ, 428–431, 2014.

[67] D. Mozaffarian, E. J. Benjamin, A. S. Go, D. K. Arnett, M. J. Blaha, M. Cushman, S. R. Das, et al. Heart disease and stroke statistics—2016 update: A report from the American Heart Association. *Circulation* 133(4):e38–e360, 2015.

[68] D. M. Negoescu, K. Bimpikis, M. L. Brandeau, and D. A. Iancu. Dynamic learning of patient response types: An application to treating chronic diseases. *Management Science* 64(8):3469–3488, 2018.

[69] O. Nohadani and A. Roy. Robust optimization with time-dependent uncertainty in radiation therapy. *IISE Transactions on Healthcare Systems Engineering* 7(2):81–92, 2017.

[70] C. M. O'Connor, V. Hasselblad, R. H. Mehta, G. Tasissa, R. M. Califf, M. Fiuzat, J. G. Rogers, C. V. Leier, and L. W. Stevenson. Triage after hospitalization with advanced heart failure: The escape (evaluation study of congestive heart failure and pulmonary artery catheterization effectiveness) risk model and discharge score. *Journal of the American College of Cardiology* 55(9):872–878, 2010.

[71] V. J. Patel. MGH Internal Medicine Associates: Primary care redesign. Unpublished master's thesis, Massachusetts Institute of Technology, 2015.

[72] W. L. Reib. Increasing patient throughput in the MGH Cancer Center Infusion Unit. Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge, 2015.

[73] G. H. Robinson, L. E. Davis, and R. P. Leifer. Prediction of hospital length of stay. *Health Services Research* 1(3):287–300, 1966.

[74] K. Safavi, T. Khaniyev, J. Zanger, A. C. Zenteno Langle, M. S. Copenhaver, B. Daily, R. Levi, and P. Dunn. Machine learning for hospital capacity challenges: Predicting surgical discharges and identifying opportunities for improvement. Working paper, Massachusetts General Hospital, Boston, MA, 2019.

[75] S. Saghafian, W. J. Hopp, M. P. Van Oyen, J. S. Desmond, and S. L. Kronick. Patient streaming as a mechanism for improving responsiveness in emergency departments. *Operations Research* 60(5): 1080–1097, 2012.

[76] P. Santibáñez, R. Aristizabal, M. L. Puterman, V. S. Chow, W. Huang, C. Kollmannsberger, T. Nordin, N. Runzer, and S. Tyldesley. Operations research methods improve chemotherapy patient appointment scheduling. *Joint Commission Journal on Quality and Patient Safety* 38(12):541–553, 2012.

[77] A. J. Schaefer, M. D. Bailey, S. M. Shechter, and M. S. Roberts. *Modeling Medical Treatment Using Markov Decision Processes.* M. L. Brandeau, F. Sainfort, W. P. Pierskalla, M. L. Brandeau, F. Sainfort, and W. P. Pierskalla. *Operations Research and Health Care*, International Series in Operations Research & Management Science, Vol. 70. Springer, Boston, 593–612, 2004.

[78] M. R. Schlanser. Optimization of surgical supply inventory and kitting. Master's thesis, Massachusetts Institute of Technology, Cambridge, 2013.

[79] J. Shi, O. Alagoz, F. S. Erenay, and Q. Su. A survey of optimization models on cancer chemotherapy treatment planning. *Annals of Operations Research* 221(1):331–356, 2014.

[80] P. Shi, J. Helm, J. Deglise-Hawkinson, and J. Pan. Timing it right: Balancing inpatient congestion vs. readmission risk at discharge. Working paper, Purdue University, West Lafayette, IN. http://dx.doi.org/10.2139/ssrn.3202975, 2019.

[81] M. Y. Sir, D. Nestler, T. Hellmich, D. Das, M. J. Laughlin Jr, M. C. Dohlman, and K. Pasupathy. Optimization of multidisciplinary staffing improves patient experiences at the Mayo Clinic. *Interfaces* 47(5):425–441, 2017.

[82] H. Song, A. Tucker, R. Graue, S. Moravick, and J. Yang. Capacity pooling in hospitals: The hidden consequences of off-service placement. *Management Science*, ePub ahead of print August 15, https://doi.org/10.1287/mnsc.2019.3395, 2019.

[83] C. Terwiesch, K. C. Diwas, and J. M. Kahn. Working with capacity limitations: Operations management in critical care. *Critical Care* 15(4):308–313, 2011.

[84] B. G. Thomas, S. Bollapragada, K. Akbay, D. Toledano, P. Katlic, O. Dulgeroglu, and D. Yang. Automated bed assignments in a complex and dynamic hospital environment. *Interfaces* 43(5): 435–448, 2013.

[85] K. M. Thompson, T. J. D. Tebbens, M. A. Pallansch, S. G. F. Wassilak, and S. L. Cochi. Polio eradicators use integrated analytical models to make better decisions. *Interfaces* 45(1):5–25, 2015.

[86] J. Tu and M. Guerriere. Use of a neural network as a predictive instrument for length of stay in the intensive care unit following cardiac surgery. *Computers and Biomedical Research* 26(3):220–229, 1993.

[87] A. Turkcan, B. Zeng, and M. Lawley. Chemotherapy operations planning and scheduling. *IIE Transactions on Healthcare Systems Engineering* 2(1):31–49, 2012.

[88] E. Ugarph. Patient flow optimization in the Department of Medicine at MGH. Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge, 2017.

[89] S. Walczak, W. E. Pofahl, and R. J. Scorpio. Predicting hospital length of stay with neural networks. D. J. Cook, ed. *Proceedings of the 11th International Florida Artificial Intelligence Research Symposium (FLAIRS) Conference.* AAAI Press, Menlo Park, CA, 333–337, 1998.

[90] J. C. Woodall, T. Gosselin, A. Boswell, M. Murr, and B. T. Denton. Improving patient access to chemotherapy treatment at Duke Cancer Institute. *Interfaces* 43(5):449–461, 2013.

[91] J. Zanger. Predicting surgical inpatients' discharges at Massachusetts General Hospital. Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge, 2018.

[92] A. C. Zenteno Langle, T. Carnes, R. Levi, B. J. Daily, and P. F. Dunn. Systematic OR block allocation at a large academic medical center. *Annals of Surgery* 264(6):973–981, 2016.

[93] A. C. Zenteno Langle, T. Carnes, R. Levi, B. Daily, D. Price, S. Moss, and P. Dunn. Pooled open blocks shorten wait times for nonelective surgical cases. *Annals of Surgery* 262(1):60–67, 2015.

[94] A. C. Zenteno Langle, R. Levi, W. Reib, E. A. Souza, B. Daily, M. Bloom, G. A. Estey, M. C. Jernigan, E. P. Hochberg, and I. Lennes. Smoothing resource utilization in a large cancer center infusion unit. Working paper, Massachusetts General Hospital, Boston, MA, 2019.

[95] B. Ziaeian, G. F. Kominski, M. K. Ong, V. M. Mays, R. H. Brook, and G. C. Fonarow. National differences in trends for heart failure hospitalizations by sex and race/ethnicity. *Circulation: Cardiovascular Quality and Outcomes* 10(7):e003552, 2017.