Research Statement

Sandeep Rath
sandeep_rath@kenan-flagler.unc.edu

1 Introduction

In this document, I present an overview of my research approach, highlight key contributions from my past research, and lay out the road map for my future research endeavors.

My primary research focus is within the domain of **Healthcare Operations**, where I delve into two distinct areas: **enhancing the efficiency of operating rooms** and **improving the delivery of primary and chronic care**.

My approach to research involves close collaboration with hospitals, providers, and technology firms to pinpoint critical issues related to resource allocation that arise in healthcare systems. My research encompasses a variety of methodologies. I combine predictive and prescriptive methods, including data-driven optimization, large-scale linear and integer programming, robust optimization, dynamic programming, and causal inference. My model-based approach culminates in actionable solutions and valuable insights for healthcare administrators.

As a quick reference [1], [2], and [3] are published in *Operations Research*, [4] is published in *Production and Operations Management*, [5] is under 2^{nd} round review at *M&SOM*, and [6] is under review at *Operations Research*. Furthermore, [7] is under preparation for submission. In addition, towards discussing my future research projects, I also include a brief discussion on current ongoing work in [8], [9], and [10].

[3] was in partnership with the Veterans Health Administration (VHA), Los Angeles. [2] and [4] were in collaboration with the University of California Los Angeles (UCLA) Ronald Reagan Medical Center. [1] was in partnership with University of North Carolina (UNC) Health and Opflow a healthcare technology firm, [5], [7], [8], [9], and [10] are in partnership with UNC Health Department of Psychiatry and Department of Family Medicine. [6] was in collaboration with University of Pittsburgh Medical Center (UPMC).

Several of these models were implemented at the partner health organizations. [3] was implemented at VHA, [2] was implemented at UCLA, [1] was implemented at UNC Rex Hospital and subsequently several other healthcare systems. Finally, [6] was implemented across 11 hospitals under UPMC. Other papers have significantly guided management and policy at partner health organizations.

In addition to the above work, [11], [12], and [13] have been published in clinical and practitioner-focussed publications.

2 Research Approach

Broadly, in my research, I perform the following steps with my research collaborators:

1. **Partnering with Healthcare Practitioners**

   First, I identify potential partners in healthcare systems who deal with resource allocation problems within healthcare systems. My partners have been researchers in population health, healthcare administrators, and healthcare providers such as primary care physicians, surgeons, and anesthesiologists.

2. **Identifying High-Impact Operational Decisions**

   Next, together with healthcare practitioners, I identify high-impact operational decision at their healthcare facility, which is currently being performed in an ad-hoc fashion or through naive heuristics. This status quo operational decision-making and potential improvement forms the starting point of my research problem formulation.

3. **Data Collection, Model Development, and Analysis**

   Next, I collect relevant retrospective data from EHR systems or other transactional data collected by healthcare systems towards their daily operations. I utilize methodologies from data-driven optimization, predictive statistical methods, and causal inference to develop analytical models that would provide decision support and management insight towards improving operations in the healthcare system.

4. **Practical and Theoretical Contribution**

   Implementing my work at these healthcare organizations has helped close the last-mile gap between theory and practice. Several of these implementations have led to significant cost savings, stable provider schedules, and improved patient outcomes for the healthcare systems. In addition to practical impact, the contribution of my existing work to the broader literature in Operations Management can be categorized as follows:

   (a) **Combining predictive and prescriptive methods** to enhance optimization algorithms towards high performance in complex high dimensional systems.

   (b) **Scaling optimization algorithms** to solve resource allocation problems at large healthcare systems.

   (c) **Identifying novel managerial insights** from situations that appear naturally in healthcare systems and contribute to enhancing our understanding of operations more broadly.

In the next section, I describe briefly the above aspects of my research papers.

3 Research Output

I organize this section in the two primary areas of healthcare operations I have contributed to: (1) improving operating room efficiency and (2) primary and chronic care delivery. Within each area, I describe the problem, my contribution, and the impact of my work. I also describe my ongoing work in each area.

1. **Improving Operating Room Efficiency** Evidence suggests that a significant portion of healthcare expenditure is wasted because of operational inefficiencies at sites such as hospitals. In my work related to operating rooms, I address these costs of operational inefficiencies.

   ORs constitute 60% of cost and 40% of revenue at a hospital. Staffing and surgical instruments together constitute more than 40% of variable costs at ORs. In my current work, I have focussed on developing novel data-driven optimization models for improving staff planning and scheduling and reducing expenses related to managing surgical instruments. In the following paragraphs, I describe my work in these two areas.

   Data- driven Surgical Tray Optimization (Operations Research, 2023) [1]:

   Hospitals spend several million dollars annually on instrument sterilization, tray assembly, and repurchase costs. However, in most hospitals, less than 20-30% of reusable instruments supplied to surgery are used on average. Prior implementations of surgical tray rationalizations have typically been expert-driven. These tray rationalization methods are labor-intensive efforts typically focused on a few trays. On the other hand, past mathematical programming model-based studies have commonly tested models with simulated data because of the difficulty in obtaining actual instrument-level usage data. We obtained actual surgical instrument usage at a large multi-specialty hospital in partnership with OpFlow, a healthcare software company. We formulate a data-driven mathematical optimization model for surgical tray configuration and assignment to reduce costs of unused instruments, such as sterilization, instrument purchase, and tray assembly costs. We develop a solution methodology that scales to thousands of surgeries, thousands of instruments, and hundreds of surgical trays. This methodology decomposes the problem into a tray rationalization and an add-on tray creation step. At each step, we solve a restricted version of the complete problem. We perform extensive out-of-sample testing of our solution. Our model-based approach identifies improvements in tray configuration and assignment, which would lead to a 54% reduction in unused instruments per surgery compared to the current tray configuration used at this hospital. We also validated our model with an expert-recommended solution for a tray subset. Our model-based solution leads to 20% lower overage and 21% lower underage than the expert-recommended solution. We estimate projected annual cost savings of 35% in instrument sterilization, tray assembly costs, and instrument repurchase costs by using the recommendations of our model. We implemented our solution at the UNC Rex Hospital, and we report on the results of our implementation. This analysis has quantified the value of collecting point-of-usage data to be at least $1.39 million per year from using the model-recommended solution at the hospital. This model-driven solution was subsequently implemented across six other large healthcare systems across the U.S.

Staff Planning for Hospitals with Implicit Cost Estimation (Production & Operations Management, 2022) [4]:

Uncertainty in the demand for services and the specialized skill set of employees make staff planning for healthcare systems challenging. Hospitals make efficient use of labor through volume-flexible staffing strategies such as the use of floating resources and overtime. Such volume flexibility can help reduce hospital costs by reacting to changes as information about the future workload becomes available. Volume flexibility in hospitals has been used in staff planning for nurses and physicians—overtime of clinical staff with lower patient safety, higher employee burnout, and deteriorating employee health. Thus, staff planners often use additional employees who can be called on short notice to reduce reliance on overtime. The use of this contingency labor supply reduces the number of overtime hours. However, depending on the staffing policy, this may increase administrative costs. These consist of both explicit and implicit costs. Explicit costs represent the actual monetary payment made and recorded for an activity. Implicit costs could include the opportunity cost to the organization associated with staff idle time and the inconvenience to employees whose schedules change on short notice. Some staff planning systems have not been successful at large retail organizations. The principal challenge in implementing model-based staff planning systems is minimizing overall costs by incorporating the employees' explicit and implicit human costs. Not combining all the human costs would likely lead to failure in accepting and implementing these systems. In this study, we provide an approach to estimate the implicit costs in staff planning. Subsequently, we solve a two-stage stochastic staff planning problem that incorporates both explicit and implicit costs. This provides a framework for creating staff planning models that overcome the shortcomings of dynamic optimization models in situations where some cost parameters may be implicit, as is often the case in service organizations.

Using data from the operating services department at the UCLA Ronald Reagan

Medical Center, our model shows that using our estimation and staff planning framework can reduce overall staffing costs by up to 16% or $2.17 million annually. We also provide managerial insights related to the relative scale of these costs, hiring decisions by service, sensitivity to cost parameters, and improvements in predicting the booked time durations.

Integrated Anesthesiologist and Room Scheduling for Surgeries (Operations Research, 2017) [2]:

Management at UCLA Ronald Reagan Medical Center's surgical services department felt that the daily resource allocation decision played a significant role in overall department cost and the service quality delivered to patients. They believed these aspects could be significantly improved by developing an analytical model-based approach that considered the fundamental complexities in this environment and applied historical surgical data to decide resource assignment and scheduling. This paper describes the development, implementation, and evaluation of a model-based decision support system that uses a data-driven robust optimization procedure to determine the daily scheduling of anesthesiologists and rooms for elective surgeries at the UCLA RRMC.

We developed a model that considered two key types of parallel resources required for surgeries, ORs, and anesthesiologists, and we simultaneously optimized their assignment and sequencing. We develop an efficient solution method using robust optimization to provide effective solutions to staffing and scheduling at large OR suites. This data-driven robust optimization approach successfully solved the full-scale problem for the entire surgery suite at the RRMC within 25 minutes, with a performance gap within 5% from the lower bound. It also significantly outperforms the best benchmark procedures in the literature.

Finally, we developed a model-based decision support system, which has been validated and implemented at the UCLA RRMC. To the best of our knowledge, this is the first implementation of a robust optimization-based decision support system in the healthcare industry. This decision support system has considerably improved upon current practice and has resulted in average daily cost savings of around 7% or estimated to be $2.2 million annually. Further, the insights from our work have had a notable impact on decision-making at the hospital.

Multilocation, Dynamic Staff Planning for a Healthcare System (Under Review at Operations Research) [6]:

The Department of Anesthesiology at UPMC centrally manages anesthesia services for 11 hospital locations in the greater Pittsburgh area, staffing up to 100 physicians per day. The department's management was interested in implementing a data-driven on-call system to reduce overtime and idle time costs, which led to our involvement. In this paper, we develop, validate, and implement a data-driven on-call system for anesthesiology staffing across the 11 UPMC hospitals.

We formulate the problem as a multi-stage robust mixed-integer program incorporating on-call flexibility to address demand uncertainty. The optimized dynamic staffing plan has been successfully implemented in the University of Pittsburgh Medical Center healthcare system, leading to estimated annual cost savings of 12% compared to current practice, or about $800,000 annually. We also provide managerial insights into the importance of improved forecasts, location flexibility, and the impact of fairness constraints. The proposed methodology is generalizable to other areas of healthcare staffing, such as nurse staffing.

Existing staffing models in healthcare staffing typically focus on aggregate staff planning, where the decision variable is the number of staff assigned to facilities. Such models face challenges during implementation as their formulation cannot enforce individual-level constraints such as location feasibility and fairness constraints. The problem's scale is also large, with binary assignment variables for over 100 anesthesiologists across 11 hospital locations and uncertainty at two different times. Therefore, existing models that solve the staffing problem by methods such as Sample Average Approximation (SAA) would not apply to our problem. We overcome these challenges by formulating the staff planning problem at the individual level and exploiting the problem's structure to reformulate the situation into a large-scale mixed-integer linear program. We then propose a nested column and constraint generation method to solve this problem computationally efficiently.

2. **Improving Primary and Chronic Care Delivery** In my work on primary and chronic care delivery, I address the challenges of EHR workload on primary care physicians and the challenges of integrating chronic care within primary care. In particular, I investigate the management of patients with co-morbid diabetes and depression within primary care, and the screening, testing, and care of HIV patients. In the following paragraphs, I describe my work in these areas.

   <u>*Physician's Choice of When to Perform EHR Tasks*</u> *(Under 2nd round review at M&SOM) [5]*:

   Physicians spend more than 5 hours a day working on Electronic Health Record (EHR) systems and more than an hour doing EHR tasks after the end of the workday. Numerous studies have identified the detrimental effects of excessive EHR use and after-hours work, including physician burnout, physician attrition, and appointment delays. However, EHR time is not purely an exogenous factor as it depends on physician usage behavior that could have important operational consequences.

Interestingly, prior literature has not considered this topic rigorously.

We combine EHR audit log data and appointment time stamp data from UNC Family Medicine to obtain a novel, detailed process-level data set of EHR use with respect to appointment progression for over 150,000 appointments from 74 physicians. We model the pre-appointment EHR work (prework), in-room EHR work with the patient in the room (multitasking), post-appointment EHR work (postwork), and end-of-day EHR work (eod) as a simultaneous system of identifiable equations. We analyze the causal effect of performing prework and postwork on total and after-hours EHR workload. We consider these two outcomes as they significantly influence physician burnout.

We find that the effect of working on EHR systems depends on whether the work is done before or after an appointment. Prework reduces total EHR workload and after-work hours spent on EHR. Postwork reduces after-work hours on EHR but increases total EHR time. We find that increasing idle time between appointments can encourage both prework and postwork, but physicians use idle time between appointments more for postwork than for prework.

Our results not only help us understand the timing and structure of work on secondary tasks, more generally, but also will help healthcare administrators create EHR workflows and appointment schedules to reduce physician burnout associated with excessive EHR use

Managing Collaborative Care for Diabetes and Depression.(Working paper) [7]:

About 27% of patients with diabetes also suffer from depression, and the presence of co-morbid depression could increase the cost of care for diabetes by up to 100%. Several randomized clinical trials have demonstrated that physical and mental health are more likely to improve for diabetes patients suffering from depression when regular treatment for depression is provided in a primary care setting (called Collaborative Care). An important operational lever in managing Collaborative Care is allocating the care manager's time to enroll patients based on their requirements, which influences the revenue, costs, and patient health outcomes. We present a mathematical modeling approach that determines the optimal allocation of a care manager's time and quantifies the costs and benefits of Collaborative Care. In particular, we model Collaborative Care management at the clinical level as an infinite horizon Markov Dynamic Program. The objective is a weighted sum of total patient QALYs and the clinic profits. The model incorporates insurance payment, resource utilization costs, and disease progression of comorbid diabetes and depression. We derive structural properties of the optimal allocation of the care manager's time. Using these structural properties, we develop a practical and easy-to-implement approach that performs close to the optimal solution. We calibrate the model with data obtained from a large academic medical center and show that our solutions could potentially improve total QALYs and clinic profits compared to current practices. We also perform sensitivity analysis to different payment models to derive insights relevant to healthcare policy.

Planning for HIV screening, testing, and care *(Operations Research, 2016) [3]*:

We analyzed the planning problem for HIV screening, testing, and care. This problem consists of determining the optimal fraction of patients to be screened in every period as well as the optimum staffing level at each part of the healthcare system to maximize the total health benefits to the patients measured by Quality-Adjusted Life-Years (QALYs) gained. We modeled this problem as a nonlinear mixed integer programming program comprising of disease progression (the transition of the patients across health states), system dynamics (the flow of patients in different health states across various parts of the health care delivery system), and budgetary and capacity constraints. We applied the model to the Greater Los Angeles (GLA) station in the Veterans Health Administration (VHA) system. We found that a Center for Disease Control recommended routine screening policy in which all patients visiting the system are screened for HIV, irrespective of risk factors, may not be feasible due to budgetary constraints.

Consequently, we used the model to develop and evaluate managerially relevant policies within the existing capacity and budgetary constraints to improve the current risk-based screening policy of screening only high-risk patients. Our computational analysis showed that the GLA station can substantially increase (20% to 300%) the QALYs gained by using these policies over risk-based screening. The GLA station has already adopted two policies that could yield better patient health outcomes over the next few years. In addition, our model insights have influenced this station's decision-making process.

4 Future Work

In my current research projects, I continue my work in operating room management and primary and chronic care delivery.

In addition to staff planning and inventory management at ORs, an important decision is planning for Operating Room capacity. OR management typically provides surgeons with reserved time in the operating room to schedule their surgeries. These reserved times are called 'OR block time.' In order to make the best use of OR capacity, surgeons are often required to release unused OR block time several days in advance. In [9], I am extending my work towards improving operating room efficiency by collaborating with the UNC Health System Surgical Services department to develop a data-driven approach to OR block release.

In [8], we are extending our work on the operational impact of physician EHR use by analyzing the impact of EHR task standardization tools on clinic operational outcomes such as appointment delays. In [10], we are developing appointment schedules for primary care that incorporate both secondary tasks like EHR work and primary tasks like patient face-to-face time.

My Works

[1] Deshpande, V., Mundru, N., Rath, S., Knowles, M., Rowe, D., and Wood, B.C. Data-driven surgical tray optimization to improve operating room efficiency. *Operations Research*, 2023.

[2] Rath, S., Rajaram, K., and Mahajan, A. Integrated anesthesiologist and room scheduling for surgeries: Methodology and application. *Operations Research*, 65(6):1460–1478, 2017.

[3] Deo, S., Rajaram, K., Rath, S., Karmarkar, U.S., and Goetz, M.B. Planning for hiv screening, testing, and care at the veterans health administration. *Operations research*, 63(2):287–304, 2015.

[4] Rath, S. and Rajaram, K. Staff planning for hospitals with implicit cost estimation and stochastic optimization. *Production and Operations Management*, 31(3):1271–1289, 2022.

[5] Celik, U., Rath, S., Kesavan, S., and Staats, B.R. Does physician's choice of when to perform EHR tasks influence total EHR workload? *Under 2nd round review at M&SOM, available at SSRN: https://ssrn.com/abstract=4323956*, 2023.

[6] Rath, S., Rajaram, K., Hudson, M., and Mahajan, A. Multilocation, dynamic staff planning for a healthcare system: Methodology and application. *Under review at Operations Research, available at SSRN: https://ssrn.com/abstract=4538869*, 2023.

[7] Rath, S., Swaminathan, J.M., and Coleman, C. Managing collaborative care for diabetes and depression. *Under preparation for submission, Available at SSRN 3944940*, 2021.

[8] Celik, U., Rath, S., Kesavan, S., and Staats, B.R. Reducing appointment delays: The impact of standardized ehr usage on physician timeliness. *Work in Progress*, 2023.

[9] Rath, S. and Lim, M.R. Data-driven surgical block release planning. *Work in Progress*, 2023.

[10] Rath, S., Kesavan, S., and Staats, B.R. Appointment scheduling with offline discretionary work. *Work in Progress*, 2023.

[11] Wood, B.C., Konchan, S., Gay, S., Rath, S., Deshpande, V., and Knowles, M. Data analysis of plastic surgery instrument trays yields significant cost savings and efficiency gains. *Annals of Plastic Surgery*, 86(6S):S635–S639, 2021.

[12] Knowles, M., Gay, S.S., Konchan, S.K., Mendes, R., Rath, S., Deshpande, V., Farber, M.A., and Wood, B.C. Data analysis of vascular surgery instrument trays yielded large cost and efficiency savings. *Journal of Vascular Surgery*, 73(6):2144–2153, 2021.

[13] Rajaram, K. and Rath, S. Managing the operating room suite at ucla ronald reagan medical center. *Impact*, 3(2):31–35, 2017.

Last updated on October 26, 2023