



Enhancing efficiency and workflow in oncology outpatient services through simulation-based optimization

Chenhao Zhou¹ · Xin Ding¹ · Weiwei Chen¹ · Lei Lei¹ · Jacqueline Norrell² · Alison Tray² · Andrew M. Evens²

Received: 7 May 2025 / Accepted: 16 March 2026 / Published online: 26 May 2026
© The Author(s) 2026

Abstract

Outpatient chemotherapy clinics face rising demand and growing patient dissatisfaction, creating pressures to improve efficiency while sustaining quality care. This study uses an agent-based simulation of a National Cancer Institute (NCI)-designated cancer center's infusion clinic to evaluate the effectiveness of three proposed operational changes: setting up an on-site laboratory, adapting nurse staffing based on anticipated demand increases, and balancing appointment scheduling via simulation optimization. The model, calibrated with historical data, tracks patient visit length and nurse utilization. Our study shows that combining an on-site lab with optimized scheduling reduces average visit length by about 75–90 minutes (approximately 30%) under various demand levels, smoothing workflow peaks. The study also shows that modest increases in nursing staff ensure timely patient throughput with 7–10% decreases in nurse utilization. Joint implementation of these operational changes outperforms individual changes, highlighting important synergies and the value of an integrated operations strategy. Overall, the study demonstrates a generalizable simulation-based framework to guide oncology process redesign and strategic decisions such as capacity expansion or facility relocation.

Keywords Healthcare process redesign · Oncology · Process improvement · Simulation optimization

1 Introduction

The increasing prevalence of cancer worldwide has sparked a surge in demand for oncology services, including the Outpatient Chemotherapy Process (OCP), a critical component of cancer treatment. In 2018, an estimated 9.8 million patients required first-course chemotherapy annually, which is expected to increase to 15 million by 2040 (Wilson et al., 2019). This indicates a 53% demand surge for chemotherapy services, corresponding to roughly a 2% annual increase in demand, with many of these patients residing in low-income or middle-income countries (Wilson et al., 2019). Chemotherapy requires intricate coordination and

✉ Chenhao Zhou
chenhao.zhou@rutgers.edu

¹ Department of Supply Chain Management, Rutgers University, 100 Rockefeller Rd, Piscataway 08854, NJ, USA

² Rutgers Cancer Institute, 195 Little Albany Street, New Brunswick 08901, NJ, USA

management due to its profound physical and psychological impacts on patients (Dollinger, 1996; Meaney et al., 2017; Todres and Wojtiuk, 1979), as well as operational challenges like the variability in treatment times, different cancer types, various chemotherapy protocols, and limited resources such as the availability of pharmacists and nurses (Heshmat and Eltawil, 2017; Mustapha et al., 2016). The challenge is further complicated by the need to coordinate across multiple departments, such as the administration of chemotherapy drugs in the pharmacy and blood tests in the laboratory, thus necessitating the improvement of operational efficiency for cancer care (Corsini et al., 2023; Heshmat et al., 2022; Lamé et al., 2016).

Existing research efforts have focused primarily on achieving operational objectives, such as minimizing waiting times, reducing idle time, and improving resource utilization (Zhan et al., 2017). Mathematical modeling and optimization methods have been employed for scheduling and resource allocation in healthcare settings, guided by well-defined objectives and constraints (typically linear), and underscore the value of optimization in improving oncology clinic operations (Liang et al., 2015). However, some empirical research highlights that delays in outpatient chemotherapy are propelled by three structural drivers: volatile appointment scheduling, misaligned resource allocation, and fragmented process design. These weaknesses manifest as slow registration and insurance checks, lengthy laboratory turnaround, drug-compounding queues, nurse shortages, and constrained chair capacity directly erode patient satisfaction (Hendershot et al., 2005). Since these three drivers interact in unpredictable, highly dynamic, and nonlinear ways, mathematical optimization models can become intractable for many practical problems (Chen et al., 2023; Jarrett et al., 2020). As a result, simulation has emerged as a useful tool for accurately representing these dynamics over time and providing detailed and exploratory insights (Goswami et al., 2024; Jalalimanesh et al., 2017).

Simulation models have demonstrated their effectiveness as analytical tools in healthcare operations, aiding in exploring operational dynamics and patient flow within oncology settings (Alvarado and Ntamo, 2018; Corsini et al., 2023). However, existing simulation studies typically evaluate policies under fixed and steady-state conditions, sidestepping the realities of demand growth and ongoing process drift. This reveals gaps in strategic planning for future demand and operational adjustments, suggesting that more adaptive and forward-looking simulation models can be developed to ensure that oncology clinics meet changing healthcare needs without sacrificing quality or efficiency (Woodall et al., 2013). Another limitation is to capture human behavior and interactions during simulation. Recent studies have combined agent-based simulation (ABS) with traditional discrete event simulation (DES) by modeling each patient and caregiver as an autonomous entity whose decisions evolve with system state (Corsini et al., 2023). A third limitation lies in the what-if analysis used in simulation studies to evaluate the impact of various changes on a system's performance. Such analyses often focus on a limited number of hypothetical scenarios, leaving many other potential situations unexplored. This can restrict the comprehensiveness of the analysis and potentially overlook interactions of different components of the system (Hadid et al., 2022).

To address these gaps, we develop an ABS model of an academic oncology infusion clinic that captures patient flow and resource interactions. The model is calibrated using empirical operational data and validated against it. We further embed simulation-based optimization to explore appointment patterns under prospective demand growth and utilization constraints. Unlike prior work that isolates scheduling, staffing, or process redesign, our approach evaluates these components jointly to uncover interaction effects and operational trade-offs. The methodology used to evaluate efficiency and process improvements, along with the recommendations drawn from the case study, offer valuable insights for other oncology outpatient clinics experiencing increased demand in the coming years.

The rest of this paper is organized as follows. Section 2 provides a review of literature on oncology process optimization, highlighting the gaps in current research and the need for improved simulation analysis. Section 3 details the development and specification of our simulation model, including the processes at the Rutgers Cancer Institute of New Jersey and model validation. Section 4 presents experimental settings, the findings of our simulation study, and their implications for improving the operations of oncology outpatient units. Finally, Section 5 concludes the paper, discussing the limitations of our study and suggesting directions for future research.

2 Literature review

There is an extensive body of literature on the use of simulation in healthcare operations. Given that this paper focuses on oncology clinics with unique characteristics, we limit our review to simulation studies on outpatient chemotherapy.

Table 1 illustrates that while simulation research on outpatient chemotherapy has grown steadily, the studies are dispersed across various topics. Researchers have employed what-if analysis to explore improved appointment schedules, optimized staffing plans, and more efficient day-to-day workflows. However, few studies have simultaneously examined all three components holistically and used simulation optimization techniques to search for optimized system designs, particularly in the context of rising patient demand. In the following sections, we review the literature for each component individually.

Appointment Scheduling. Scheduling of chemotherapy appointments is a complex puzzle, as clinics juggle patient preferences, varying treatment durations, and downstream tasks (labs, pharmacy) that precede each visit. The objective is to create appointment timetables that maximize the utilization of chairs and staff while minimizing patient waiting time. Most studies on this topic use mathematical models to optimize schedules, while simulation remains the principal means to evaluate same-day scheduling rules under real-world variability. Historical data have been used to test alternative templates and recommend a balanced mix of long and short infusions Ahmed et al. (2011). Staggering start times and capping simultaneous arrivals increased throughput and were implemented in practice Woodall et al. (2013). A hybrid study first optimized a timetable and then simulated infusion-time uncertainty, identifying schedules that remained effective under randomness Liang et al. (2015). Heuristics are assessed in the same way: “longest infusion first” rule performs well under random deferrals but is sensitive to pharmacy delays Garaix et al. (2020), whereas sequencing patients by decreasing infusion length raises efficiency only when arrivals and drug preparation are perfectly on time Suss et al. (2018). In all of these cases, simulation functions mainly as an evaluation stage, validating pre-specified rules rather than driving an iterative search for optimal schedules.

Resource Allocation. Simulation models are widely used to pinpoint bottlenecks in nurses, pharmacists, infusion chairs, and ancillary spaces and then test ways to relieve them. Early work showed that increasing chair–nurse dyads steadily reduces waiting, but with diminishing returns Matta and Patterson (2007). A comprehensive study of a cancer agency demonstrated that the availability of nurses, chairs, and physicians jointly govern delays, yet the greatest leverage came from rebalancing existing resources rather than adding new capacity Santibáñez et al. (2009). At the Duke Cancer Institute, simulation identified nurse unavailability as the dominant bottleneck; coupling a mixed-integer scheduling model with the simulation improved flow without adding additional staff Woodall et al. (2013). Similar

Table 1 Representative simulation-based studies in oncology clinic operations (2005–2025)

Study (Year)	Scheduling	Resource	Process	Demand	Method
Matta and Patterson (2007)	✓	✓			DES
Santibáñez et al. (2009)	✓	▲	✓		DES
Ahmed et al. (2011)	✓	▲			DES
Lu et al. (2012)			✓		DES
Masselink et al. (2012)			✓		DES
Yokouchi et al. (2012)	✓	▲			DES
Woodall et al. (2013)	▲	✓		✓	DES+OPT
Liang et al. (2015)	✓	✓	▲		DES+OPT
Baril et al. (2016)			✓		DES
Alvarado and Ntaimo (2018)	✓	✓			DES+OPT
Suss et al. (2018)			✓		DES
Liu et al. (2019)	▲	✓		✓	DES
Baril et al. (2020)		✓			DES+DOE
Lamé et al. (2020)			✓		DES
Heshmat and Eltawil (2021)	✓	✓			DES+OPT
Slocum et al. (2021)	✓				DES
Hadid et al. (2022)	✓	✓		▲	DES+OPT
Corsini et al. (2023)		✓	✓	▲	ABS+DOE

✓ indicates that a topic is explicitly addressed in the paper; ▲ indicates that a sensitivity or minor test is conducted in the paper

DES = discrete-event simulation; ABS = agent-based simulation; DOE = design of experiment;

OPT = optimization

scenario tests revealed that simply hiring more nurses was less effective than adjusting arrival patterns Ahmed et al. (2011), underscoring the interaction between staffing and scheduling. Hybrid optimization–simulation frameworks push the analysis further: one study balanced chair utilization across heterogeneous regimens by jointly choosing schedules and resource mixes Liang et al. (2015); another DES-based model re-assigned nurses dynamically to raise effective capacity Alvarado and Ntaimo (2018). Design-of-experiments map performance across nurse-to-chair ratios and offer robust staffing rules Baril et al. (2020), while a capacity-planning tool links nurse–pharmacist mixes to service targets as demand grows Liu et al. (2019). Recent multi-objective studies combined simulation with meta-heuristics or stochastic optimization to simultaneously minimize waiting time, overtime, and cost Heshmat and Eltawil (2021); Hadid et al. (2022). These studies confirm that smart rostering and judicious capacity additions can accommodate high volumes more effectively than simple staffing increases, yet most models still evaluate pre-set resource scenarios rather than searching the full design space iteratively.

Process Redesign. Beyond scheduling and staffing, process redesign interventions target the structure and sequence of tasks in the chemotherapy visit. Real-time pharmacy–clinic coordination and parallel tasking reduce the overall length of stay Lamé et al. (2020). Pharmacy batching and pre-mix policies further compress compounding time and smooth workload peaks Lu et al. (2012); Masselink et al. (2012). Lean-oriented studies show that co-locating services and removing redundant steps accelerate patient flow Baril et al. (2016); Santibáñez et al. (2009). An agent-based model demonstrates that synchronizing external

pharmacy deliveries with chair availability markedly lowers waiting time Corsini et al. (2023). Additional redesigns, installing an on-site pharmacy, dedicating pre-treatment nurses, and resequencing lab-consultation-infusion activities also expedite throughput Suss et al. (2018). These studies show that altering task sequencing, information flow, or physical layout can yield substantial, clinic-wide performance gains, yet they are typically evaluated in isolation from scheduling and staffing decisions, highlighting the need for integrated analysis.

Despite the progress summarized above, two main research gaps continue to limit the practical relevance of simulation studies. First, most simulation studies optimize only one operational component while treating others as fixed parameters. A few papers hint at the value of a broader view, such as evaluating schedules alongside nurse capacity Santibáñez et al. (2009), co-optimizing slotting and nurse assignment Alvarado and Ntaimo (2018), or jointly tuning slot patterns and staffing in a multi-objective scheme Hadid et al. (2022). In addition, global forecasts anticipate a sustained increase in chemotherapy volume over the coming decades Wilson et al. (2019); nevertheless, most models tune parameters to current attendance levels and report gains that may erode as caseloads grow. To our knowledge, no existing work integrates all three components while also assessing robustness under increasing demand.

Second, most studies use what-if analysis to evaluate limited scenarios, while some adopt an optimization-then-simulation sequence: a heuristic or mathematical program proposes a schedule, which is further validated through simulation Ahmed et al. (2011); Masselink et al. (2012); Slocum et al. (2021). Only a few papers embed optimization inside the simulation loop, and even then, the search space remains narrow: one calls a mixed-integer model while keeping resources fixed Liang et al. (2015), and another optimizes start dates before simulation refines daily slots Heshmat and Eltawil (2021). As a result, few studies have attempted to realize the full potential of simulation optimization, which can search a large-scale design space and explore multi-objective trade-offs among waiting time, overtime, and cost, particularly subject to prospective demand scenarios.

These omissions are not merely academic: hospital leaders must choose portfolios of interventions that will hold up under future growth and shifting workflows. To fill this gap, the present study develops an ABS model with detailed operational data from the Rutgers Cancer Institute and searches for optimized system designs using simulation optimization. By capturing agent-level interactions and testing policy bundles under demand growth scenarios, the study provides a comprehensive, demand-aware evidence base for outpatient chemotherapy management and offers actionable guidance on which combinations of components deliver resilient improvements.

3 The simulation model

In this section, we will present the general process for studying efficiency and process improvement in oncology clinic settings, illustrated through the case study at Rutgers Cancer Institute. We first introduce the background of this case study and then present the three simulation model development steps.

3.1 Background and general approach

The Rutgers Cancer Institute, affiliated with RWJBarnabas Health (RWJBHealth), is part of the Robert Wood Johnson Medical School. It is one of the nation's 57 National Cancer Institute-designated Comprehensive Cancer Centers and the only one in New Jersey. Through

its affiliation with RWJBHealth, the Rutgers Cancer Institute provides adult and pediatric patients access to the most advanced cancer treatment options, including clinical trials, precision medicine, immunotherapy, complex surgical procedures, and sophisticated radiation therapy techniques. This partnership ensures that patients across the RWJBHealth network can benefit from the Rutgers Cancer Institute's specialized oncology services. It manages over 100,000 patient visits annually, reflecting its significant role in cancer care in the region.

The Rutgers Cancer Institute has experienced significant growth over the past several years in terms of patient visits and treatments. Such growth trends are expected to continue in the coming years as the center transitions into New Jersey's first freestanding cancer hospital in a new 12-story, 510,000-square-foot pavilion. However, current service lines face challenges in operational inefficiencies, which lead to long waiting times, understaffing, poor care coordination, and a lack of access to timely appointments. The utilization of resources (e.g., clinic room utilization and infusion chair turnover) has also not reached its potential. To improve patient experience, the Rutgers Cancer Institute leadership, including leading physicians, chief nursing officers, and vice presidents, has recognized that existing workflows are incapable of addressing variations in visit length across providers and service units. They have also recognized the pressing need for standardized operations and compliance with procedures, such as signing chemo orders and lab requests. As a result, the Rutgers Cancer Institute team partnered with Rutgers Business School (RBS) to identify bottlenecks in clinical workflows, aiming to enhance resource utilization, reduce waiting time, and improve the overall patient experience.

We followed a structured approach in developing the simulation model, encompassing three stages: process study, model development, and model validation Abdou et al. (2012). The process study stage identifies specific issues and defines the key processes that the model will simulate. The model development stage involves the creation of the simulation model, which includes the design of the environment's layout and the definition of agents and their communication mechanisms to ensure realistic interactions and behaviors. Finally, the model validation phase involves rigorous validation, comparing the model's output against empirical data or theoretical standards to confirm its fidelity to the real-world phenomena it aims to model. Later sections will systematically dissect each phase, detailing our approach to model development.

3.2 Process study

The study focuses on the busiest location at the Rutgers Cancer Institute, the ambulatory unit in East Tower, where all hematologic malignancy and stem cell transplant patients are managed. Our goal is to improve patient treatment processes, evaluated using two key performance metrics: *daily patient visit length* (abbreviated as *patient visit length* for the remainder of this paper) and *daily nurse utilization* (abbreviated as *nurse utilization*).

- Patient visit length is defined as the time from a patient's check-in to check-out, and the daily average visit length is the mean of all patient visits within a day.
- Nurse utilization is calculated as the total time nurses spend serving patients divided by the total available nurse-hours in a day. Each nurse's available hours are determined by their working hours multiplied by the maximum number of patients they can serve simultaneously. For example, if nine nurses work eight hours each and can serve two patients at a time, the total available nurse-hours for the day is $9 \times 8 \times 2 = 144$.

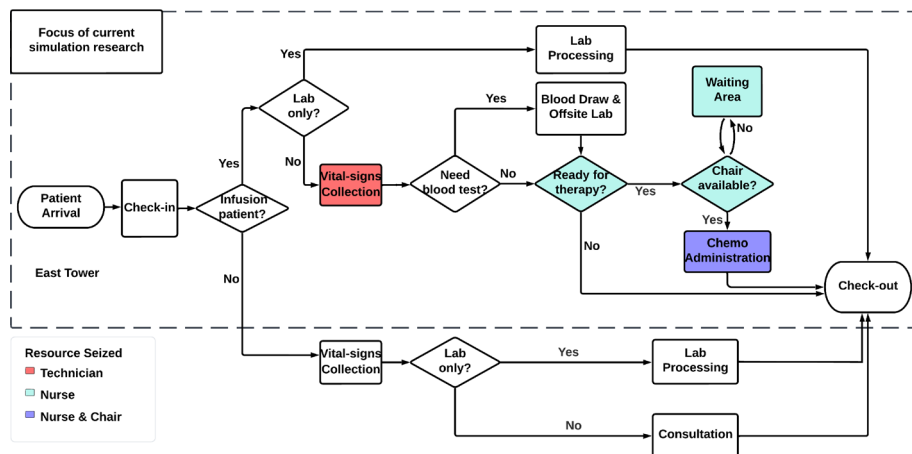


Fig. 1 Flowchart of the patient visiting processes

Note that these metrics specifically measure performance for infusion (chair) visits and standalone lab visits, as they directly impact the quality of patient care and operational efficiency of oncology clinics.

To obtain an accurate picture of patient flows, the RBS team worked closely with the Rutgers Cancer Institute team and spent three months conducting on-site observations and data collection. Specifically, a group of graduate students worked with nurses and the associate chief nursing officer to observe and log patient flows from June to August 2022. They also collected first-hand information on major concerns from patients and feedback from nurses. In addition, the RBS team analyzed Rutgers Electronic Health Record (EHR) data to supplement the field study. Through cross-validation based on the two data sources, we first developed the process flow map. Figure 1 depicts the processes captured for each patient category, as well as resources seized in certain processes (color coded).

As seen from Figure 1, the patient journey begins at the front desk, where staff record patient information and validate insurance details. Two main categories of patient visits are identified: those involving chemotherapy infusions (termed “routine” visits) and those involving other clinical services without chemotherapy (termed “clinical” visits). Clinical visits refer to appointments for non-chemotherapy care (e.g., consultations or other treatments with physicians), whereas routine visits involve regular chemotherapy infusions or related treatments (such as blood transfusions) administered in chemotherapy chairs. Patients must pass blood tests to prepare for the upcoming chemotherapy. Based on these requirements, we can categorize patients into three groups, where their specific processes are described below with nodes marked by square brackets ([]) corresponding to those in Figure 1:

- Infusion patients (chair visits) who undergo blood tests and chemotherapy on chemo-chairs. As shown in Figure 1, these patients proceed from [Check-in] to [Vital-signs Collection] for assessment. If patients have not completed their blood test prior to the visit, they will proceed with the [Blood Draw & Offsite Lab] step. During this step, blood samples are collected and sent to the hospital laboratory affiliated with the Rutgers Cancer Institute for off-site processing. Nurses are notified of the lab results upon completion and verify if the patient is ready to proceed to therapy. To ensure quality care and patient satisfaction, each nurse is assigned a maximum of two patients. If blood test results

Table 2 Summary statistics of the EHR data from the Rutgers Cancer Institute

Patient Category	October	November	December
Chair Visits	851	893	887
Lab Visits	223	231	318
Total Visits	1,074	1,124	1,205
Unique Patients	545	565	620
Resource Pool	Numbers		
Technicians	2		
Nurses (on average)	9		
Chemotherapy Chairs	21		
Lab Rooms	4		

indicate that infusion is contraindicated, patients will not proceed to the therapy step. Instead, they exit the queue via [Check-out], and any previously allocated infusion slot, as well as the nursing and chair resources, are released. If the patient is ready for therapy and when a therapy chair becomes available, the patient then proceeds to the [Chemo Administration] step, where the assigned nurse will collect the prescribed chemotherapy treatments from the pharmacy and administer them to the patient on the assigned therapy chair.

- Infusion patients (lab visits) who only require lab procedures and do not occupy chemo-chairs, such as blood draws or injections, without additional treatments on the same day. This includes patients who come for blood tests for future chemotherapy sessions or routine activities like injections. These patients follow the path [Check-in] → [Lab Processing] → [Check-out]. They are free to leave immediately after their lab procedures.
- Consultations patients who receive clinical services from providers, excluding chemotherapy. They visit doctors for comprehensive assessments and consultations in specific labs or procedure rooms, following a path of [Check-in] → [Vital-signs Collections] → [Consultation] or [Lab Processing] → [Check-out], different from an infusion visit after the department check-in.

To supplement our observations and refine the learned process, we further extracted patient-level timestamps from the institute's electronic health record (EHR) database for all ambulatory encounters between October and December 2022. Table 2 summarizes the numbers of visits related to infusions (chair visits) and standalone lab visits. Note that visits consisting solely of provider consultations (with no infusion or lab procedure that day) are excluded, as the focus of this study is on the infusion clinic workflow, and consultation patients utilize a distinct set of resources separate from those used for infusion and lab visits. Table 2 also reports the availability of key resources, such as chemotherapy chairs and lab rooms. Specifically, the data reveals that the clinic handled a growing number of monthly patient visits, with 1,074 visits in October, 1,124 in November, and 1,205 in December. The unique patients also increased from 545 in October to 620 in December. Additionally, the clinic was supported by 4 lab rooms and 21 chemotherapy chairs throughout the period. On average, the clinic's daily operations were facilitated by 9 nurses and 2 vital clinicians.

Based on input from the clinical team of the Rutgers Cancer Institute, chair visits/chemotherapy treatments were the main bottleneck in clinic operations. Following their insights, we mapped the EHR data onto the process flow outlined in Figure 1. Specifically, we calculated the distribution of patients across three service categories and the correspond-

ing length of stay. The analysis confirmed that chair visits/chemotherapy treatments account for 45% of total patient encounters and have the longest average length of stay of about 250 minutes. In comparison, lab visits only account for 29% of total patient encounters, with a shorter average length of stay of less than 170 minutes. The distribution of service categories underscores the importance of simulating patient flow and clinic operations for chair visits/chemotherapy patients.

3.3 Model development

Based on the flowchart and data learned in the process study, we proceeded to develop a simulation model in AnyLogic. AnyLogic is a versatile simulation software that supports agent-based, discrete event, and system dynamics modeling. It facilitates the creation of complex models with a user-friendly interface and provides powerful tools for visualizing and analyzing simulation results.

Our AnyLogic model creates a nurse agent pool, where each nurse can handle up to 2 concurrent patients. If all nurses are fully occupied, incoming patients wait in the queue. Chairs are modeled as resources with the pool size of chairs limited to 21. When a patient completes labs or vital checks, the model checks nurse and chair availability before proceeding to the infusion stage. If resources are unavailable, the patient enters a queue until a nurse agent and a chair resource both become available. In practice, once a pharmacist verifies the order and prepares the chemotherapy, a nurse physically collects the medication and proceeds with the patient's infusion. We therefore incorporate the pharmacy compounding and pick-up time into one combined measure labeled "chemo administration." This composite duration reflects both drug preparation and nurse infusion activity. Preliminary observations and discussions with clinical staff indicated that pharmacy was not the primary bottleneck in this particular study: nurses typically collected medications without significant queuing. In addition, pharmacy delays are treated as an exogenous factor: they may influence our results, but they are not an operational level that the clinic can directly change or control. However, should future demand shifts or operational data reveal that pharmacy is a distinct constraint (e.g., limited compounding hoods or increased complex regimens), the model can be extended to treat pharmacy as an independent resource pool with capacity parameters. Such an extension might reveal additional optimization opportunities, particularly if pharmacy delays interact with nurse scheduling during peak hours.

Figure 2 shows a 3-dimensional (3D) view of the simulation developed, illustrating the patient flow through various stages of the healthcare process, including registration, waiting, consultation, and discharge. Note that the figure is a schematic representation; in the actual facility, the functional areas are distributed across separate rooms. The nurse station, which serves as the coordination hub for caring staff, is also depicted in the model. This 3D visualization helps stakeholders better understand the system dynamics and identify potential areas for improvement.

To ensure the reliability and validity of simulations, it is essential to select appropriate probability distributions with accurately estimated parameters to drive the simulation. This is particularly important in systems characterized by significant randomness or uncertainty. For this study, we employed the Python 3 *fitter* library to determine the time distributions for each random process described in Figure 1. The *fitter* package, which leverages SciPy's suite of statistical functions, automates the fitting process by comparing multiple candidate distributions using the following metrics:



Fig. 2 3D view of the Rutgers Cancer Institute simulation model. *Note:* Labels indicate the key stages of the patient flow: check-in, vital-signs collection, waiting area, lab rooms, infusion chairs, nurse station, and check-out

Table 3 Fitted distributions and test statistics for selected processes in simulation

Process	Distribution	SSE ($\times 10^{-5}$)	Kolmogorov–Smirnov test
Vital-signs Collection	Burr ¹	1.40	$D = 0.033, p = 0.179$
Blood Draw & Offsite Lab	Double lognormal ²	0.70	$D = 0.031, p = 0.184$
Chemo Administration	Weibull ³	2.10	$D = 0.044, p = 0.443$

¹ Burr ($c = 2.18, d = 1.06, \text{loc} = 0.00, \text{scale} = 13.32$)

² Double lognormal ($w = 0.65, s_1 = 0.68, \lambda_1 = 42.31, s_2 = 0.24, \lambda_2 = 69.42$)

³ Weibull ($c = 1.17, \text{loc} = 1.68, \text{scale} = 148.17$)

- The Akaike Information Criterion (AIC). AIC measures the relative quality of a statistical model for a given set of data, assessing the trade-off between the model's goodness of fit and complexity Akaike (1974).
- The Bayesian Information Criterion (BIC). BIC is similar to AIC but includes a stronger penalty for models with more parameters, which helps select simpler models with better predictive performance Schwarz (1978).
- Kolmogorov-Smirnov (KS) goodness-of-fit test. The KS test evaluates how well the observed data conforms to the expected distribution, ensuring they fit well and remain robust in their predictive capabilities.

We evaluated a wide range of distributions for each random process. We randomly selected about 60% of unique dates from historical data for parameter fitting (training), constituting 1,549 patients for chair visits. The remaining data are used as the testing group in the model validation in Section 3.4. The best-fit distribution was chosen based on combined AIC/BIC rankings and KS goodness-of-fit tests. When these metrics have different rankings, priority is given to the KS test results, as they provide the most direct measure of the model's accuracy.

Table 3 presents the best-fit distributions for three key processes in the simulation, along with their goodness-of-fit statistics. Specifically, for each fitting, the SSE (Sum of Squared Errors) and the D -statistic are small, with p -value exceeding 0.05, indicating a strong fit to the data. Figure 3 illustrates the fitted probability density functions along with their corresponding empirical distributions, visually confirming the high quality of the fit. Note that we specifically included bimodal distributions for the Blood Draw & Offsite Lab process because its empirical data showed two distinct modes, reflecting delays in both blood draws and offsite lab processing. A double lognormal distribution was selected as the best fit.

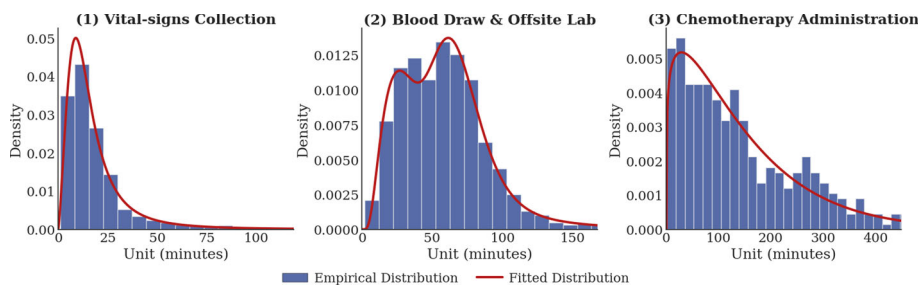


Fig. 3 Empirical distributions and fitted probability density functions for three key processes

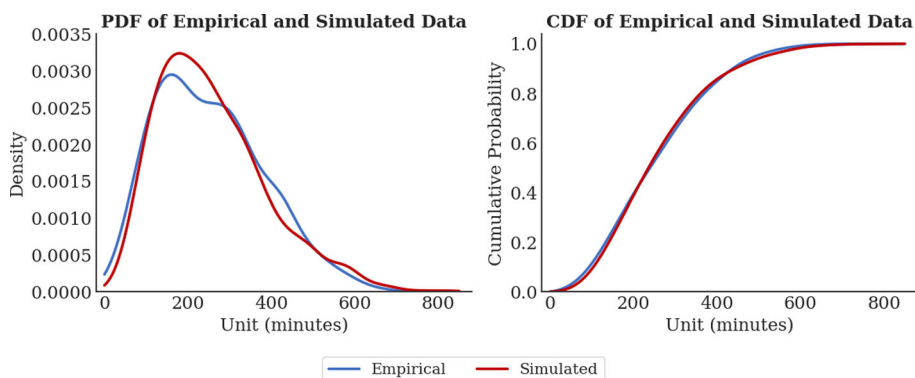


Fig. 4 Distributions of visit lengths: empirical data vs. simulated data

3.4 Model validation

In a simulation study, model validation is critical to assess how accurately the model represents the real-world system. We conducted simulations over the testing group (40% of the data) to validate the model. Patient visit length is selected as the primary metric in this study, defined as the total time elapsed from a patient's arrival at the front desk to its departure from the system.

To examine the similarity between the simulation results and empirical data in the testing group on patient visit length, we first performed a visual check of the distributions. Figure 4 provides a side-by-side comparison of the probability density function (PDF) and cumulative density function (CDF) for patient visit lengths from the test set and the simulated data. This visual representation illustrates the close agreement between the two distributions, offering a straightforward demonstration of the model's ability to replicate real-world behavior.

Next, we statistically compared the patient visit lengths produced by the simulation with the empirical data. Table 4 summarizes the descriptive statistics (mean and standard deviation (SD)) of both samples, as well as the results of a Welch t -test (with test statistic t) and a Kolmogorov–Smirnov (KS) test (with test statistic D), both conducted at a significance level of 0.05.

More specifically, it is observed from Table 4 that the mean and SD computed using the simulated data are very close to the counterparts computed using the empirical data. A Welch t -test shows that $t = -0.674$ with $p = 0.501$, indicating that we fail to reject the null hypothesis that two samples have equal means. Further, the 95% confidence interval

Table 4 Comparison of empirical and simulated patient visit lengths

Statistic	Empirical	Simulated
Mean (min)	253.2	258.3
Standard deviation (min)	127.1	128.2
Goodness-of-fit		
Welch <i>t</i> -test	$t = -0.674, p = 0.501$	
95% CI for mean difference (min)	[-19.74, 9.74]	
Kolmogorov–Smirnov test	$D = 0.052, p = 0.426$	

for the difference of means is [-19.74, 9.74], indicating that any practically meaningful deviation is unlikely. Additionally, we employed a two-sample KS test to compare whether two samples (empirical and simulated) come from the same distribution. The results show that the maximum absolute vertical distance between two empirical CDFs, captured by D , is relatively small, with $D = 0.052$. Further, the p -value of 0.426 exceeds the chosen significance level of 0.05, indicating that we fail to reject the null hypothesis that the two samples originate from the same underlying distribution. Collectively, these findings show that there are no statistically significant discrepancies between the simulated and empirical data.

In summary, the distribution visualization and statistical analysis results confirm that the simulation model developed models well patient flow at the Rutgers Cancer Institute. The close agreement between the simulated data and the real-world observations demonstrates the model's fidelity in capturing the essential characteristics of patient visit lengths, thereby providing a robust foundation for further analysis and optimization.

4 Simulation analysis and managerial insights

This section presents our experimental settings, the primary simulation results, and the associated managerial insights.

4.1 Methodologies

We employed both what-if analysis and simulation optimization for the simulation analysis. What-if analysis is used to simulate and compare a limited number of scenarios, which is a typical approach in simulation studies. However, what-if analysis can become overwhelming for complex situations involving multiple independent variables (e.g., hourly visitor arrival rates), where the design space to be tested and compared is large. To this end, simulation optimization techniques can search for the optimal design from a large design space, significantly enhancing the depth and breadth of the analysis. This method helps to understand potential outcomes better and validates a larger space of independent variables, resulting in a more comprehensive and robust analysis. It is necessary to use it in this study since we aim to understand and explore the interactions among several components in the oncology settings, as introduced earlier. By simulating different combinations of variables, simulation optimization provides insights into the best actions to take under uncertainty.

To solve the simulation optimization problem, we used OptQuest, developed by OptTek Systems Inc. OptQuest is a commercial implementation of scatter and tabu search, providing starting points for any gradient-based local solver for nonlinear programming (NLP) problems Ugray et al. (2007). These techniques have been identified as outperforming traditional methods such as genetic algorithms and simulated annealing Egea et al. (2009); Zúñiga et al. (2020). By integrating OptQuest within our scheduling model, we can explore and optimize appointment schedules more effectively than the current arrival schedule for the schedule balancing dimension.

Our study applied analysis across three dimensions: process redesign, resource level, and schedule balancing. The assumed improvement strategies included: 1) process redesign: establishing an on-site lab to enhance the blood work process and improve visit length performance; 2) resource level: increasing the number of nurses to improve resource availability; and 3) schedule balancing: adjusting the appointment schedule to balance the arrival rate between morning and afternoon hours.

Four factors are controlled and varied in the simulation studies:

1. Time Distribution of Blood Test Processing (α): This factor assesses whether each phase of the blood test process should maintain the current off-site time distribution (Level A) or transition to a new on-site time distribution (Level B). The objective is to evaluate the impact of changing the current process on patient time allocation.
2. Number of Nurses per Day (β): This pertains to the daily count of healthcare workers, presently set at 9 (Level A). The aim is to analyze performance changes when considering a variation in the number of nurses. The hospital seeks to understand the effects of adding one nurse (Level B) or two nurses (Level C) at the required time and assess the benefits of altering the healthcare team composition.
3. Schedule Balance (γ): This dimension addresses whether to maintain the current appointment capacity per hour (Level A) or adjust to an optimized number of individuals per hour at the optimized visit time through simulation optimization (Level B). Hospitals aim to understand the impact of schedule balancing and assess the necessity of enhancing this mechanism.
4. Daily Outpatient Volume (δ): This factor encompasses the daily appointment capacity, currently set at the 100% demand level (Level A), as well as 110% (Level B) and 120% (Level C) for demand increases. Considering anticipated demand growth, the hospital will evaluate the effects of process re-engineering and resource realignment in response to higher appointment volumes.

We conducted 500 simulation replications for each scenario to ensure robust statistical analysis using different random seeds to simulate distinct days. The above factors used in the design of experiments are shown in Table 5.

4.2 Simulation results and analysis

We now present the detailed setup of our experiments and analyze the simulation results.

Table 5 Factor levels applied in the design of experiments

Factor	Level A	Level B	Level C
Blood Draw & Offsite Lab Time Distribution (α)	Double lognormal ¹	Triangular ²	–
Number of Nurses (β)	9	10	11
Schedule Balance (γ)	Current	Optimized	–
Daily Outpatient Volume (δ)	100%	110%	120%

¹ Double lognormal ($w = 0.65$, $s_1 = 0.68$, $\lambda_1 = 42.31$, $s_2 = 0.24$, $\lambda_2 = 69.42$)

² Triangular (min = 5, mean = 10, max = 20)

4.2.1 Impact of on-site laboratory

Our data analysis and the Rutgers Cancer Institute practices have identified variability and inefficiencies in blood draw and processing times, primarily caused by the current practice of sending blood samples for off-site processing. Currently, the process involves collecting patient samples at the Rutgers Cancer Institute, sending them to an external laboratory for analysis, and then receiving the results back. Specifically, the historical data show that the service time for the “Blood Draw & Offsite Lab” step at the patient flowchart in Figure 1 has an empirical mean of 59.5 minutes and a standard deviation of 36.7. In other words, patients wait close to 60 minutes on average for the blood test results before proceeding to the chemotherapy in a chair. Such a long and inconsistent waiting time can delay subsequent treatment steps of this patient, as well as remaining patients assigned to the same chair. Thus, the Rutgers Cancer Institute has been seeking a more streamlined process to reduce the blood test turnaround times to resolve this bottleneck. Under consideration, one potential solution is to explore the feasibility of establishing an on-site blood lab to increase processing speed and reduce overall patient flow time rather than outsourcing blood lab work. Therefore, we use our simulation model to test the effectiveness of an on-site lab compared to the current practice of an outsourced lab.

More specifically, we tested two scenarios in the simulation: one based on the current off-site lab setup and the other on a proposed on-site lab setup. The off-site setup is characterized by the service time distribution shown in Table 5. For the on-site lab, the Rutgers Cancer Institute team anticipates that the proposed on-site process will offer faster turnaround times and more streamlined workflows than the current off-site process. The time duration of the proposed on-site process follows a triangular distribution (5, 10, 20), where 5 is the minimum time, 10 is the most likely time, and 20 is the maximum time. This distribution reflects the anticipated variability, with most process times expected to be around 10 minutes.

We used the aforementioned service time distributions to drive the simulation experiments for two scenarios. Since the Rutgers Cancer Institute is interested in evaluating performance based not only on current patient demand but also on anticipated future demand surges, we conducted the simulation under three different levels of daily outpatient volume (δ): 100% (current patient arrival levels), 110%, and 120%. The latter two are achieved by increasing the current hourly average visit rate by 10% and 20%, respectively. Evaluating average patient visit length and nurse utilization under different demand levels is practically significant for hospital operations, as it helps ensure that our strategies remain effective under future high-demand conditions.

The simulation results are presented in Table 6. It is seen from the table that the on-site blood test processing strategy significantly reduces average patient visit length compared to

Table 6 Simulation results: off-site vs. on-site lab

α	δ					
	Level A (100%)		Level B (110%)		Level C (120%)	
	Avg. Visit Len. (min)	Nurse Util. (%)	Avg. Visit Len. (min)	Nurse Util. (%)	Avg. Visit Len. (min)	Nurse Util. (%)
Off-site lab	256.35	54.13%	266.13	59.30%	280.87	65.53%
On-site lab	214.32	52.90%	224.61	57.92%	238.35	64.01%
Sensitivity Analyses for On-site Lab Triangular Distribution Parameters (min, mode, max)						
(4.5, 9, 18)	212.69	52.73%	224.71	58.15%	237.70	63.51%
(5.5, 11, 22)	214.08	53.10%	226.56	58.59%	237.25	63.69%
(5.5, 10, 18)	211.39	53.05%	221.71	58.31%	234.17	63.25%
(4.5, 10, 22)	212.52	52.83%	222.29	58.21%	234.00	63.18%

the off-site strategy. Concerning nurse utilization, it marginally decreases from off-site to on-site labs, as the nurse resources are only seized by the patient after the lab results become available (see Figure 1). Notably, both off-site and on-site labs show increases in nurse utilization as outpatient volume increases from 100% (Level A) to 120% (Level C). Nurse utilization rises from 52.90% to 64.01% with on-site lab, an increase of 11.11%, whereas it jumps from 54.13% to 65.53% with off-site lab, an 11.40% increase.

While off-site lab time distribution was derived from empirical data, the on-site lab processing time uses a triangular distribution with parameters (5, 10, 20), as provided by the Rutgers Cancer Institute team. To test the robustness of our findings, we performed a sensitivity analysis on these parameters using four alternative settings: (1) shifting the entire distribution left by 10%: (4.5, 9, 18); (2) shifting the entire distribution right by 10%: (5.5, 11, 22); (3) widening the distribution by 10% with a fixed mode: (4.5, 10, 22); and (4) narrowing the distribution by 10% with a fixed mode: (5.5, 10, 18). The results of these sensitivity analyses, provided in Table 6, confirm that the outcomes are comparable to those with the default parameters, demonstrating the robustness of the model.

Regarding average patient visit length, the on-site strategy consistently outperforms the off-site approach across all levels of outpatient volume. The data reveal that, on average, patient visit length decreases by more than 15% with on-site blood testing compared to off-site testing. This advantage remains stable regardless of changes in outpatient volume, indicating that the demand changes did not significantly influence the advantage of transitioning to on-site blood testing. The predicted average visit length reduction between the on-site and off-site strategies is consistent across all three demand levels.

The data strongly suggest that replacing off-site blood testing with an on-site strategy at the Rutgers Cancer Institute would enhance operational efficiency and offer patients a better overall visit experience. The consistent reductions in patient average visit length across varying demand levels indicate that the on-site strategy is well-suited to efficiently manage increased patient volumes. Given the upcoming move of the Rutgers Cancer Institute to a new free-standing cancer pavilion location, the transition to an on-site blood testing model presents an opportune moment to update workflow processes with relatively low initial implementation difficulty. This proactive change is especially important as the Rutgers Cancer Institute anticipates operational challenges associated with increasing demand. By integrating on-site blood testing into the new facility's design and processes, the Rutgers

Table 7 Simulation results: varying nurse levels

β	δ					
	Level A (100%)		Level B (110%)		Level C (120%)	
	Avg. Visit Len. (min)	Nurse Util. (%)	Avg. Visit Len. (min)	Nurse Util. (%)	Avg. Visit Len. (min)	Nurse Util. (%)
9	256.35	54.13%	266.13	59.30%	280.87	65.53%
10	255.97	48.72%	265.02	53.37%	279.44	58.98%
11	255.95	44.29%	264.84	48.52%	279.17	53.62%

Cancer Institute can better accommodate a growing patient load, streamline patient care, and reinforce its commitment to delivering high-quality, patient-centered services.

4.2.2 Impact of increasing nurse level

Under the current operational configuration, each nurse can attend to a maximum of two patients simultaneously. However, the average number of on-duty nurses daily is nine, with twenty-one treatment chairs available. This discrepancy indicates that the existing chair needs to have more nursing staff during peak patient volume periods. Consequently, this can lead to significant waiting times, where patients receive their blood test results but may experience delays in subsequent care due to a lack of available nursing staff. This observation is supported by feedback from the nurses, who have reported experiencing excessive workloads during peak periods.

Excessive workloads prolong the overall service delivery process and may also detrimentally impact the quality of care provided. In anticipation of the Rutgers Cancer Institute's relocation to a new facility and the expected increase in patient demand, the hospital has proposed testing the feasibility of increasing the nursing staff size. This investigation assesses whether an expanded nursing team can alleviate the anticipated process delays and enhance nurse utilization. To this end, we used the simulation model to vary the number of nurses available each day, allowing us to test and compare the effects of different nursing team sizes on patient visit lengths and nurse utilization. Specifically, we developed three primary scenarios based on varying nurse staffing levels: 9, 10, and 11 nurses. We examined three patient demand levels within each staffing level scenario: 100%, 110%, and 120% of the current demand, similar to the setting in Section 4.2.1.

Table 7 highlights the effects of increasing the number of nurses on visit length and utilization under varying outpatient demand levels.

The simulation results show that at the current demand level (100%), increasing the number of nurses from 9 to 11 results in insignificant average visit length improvements and significant utilization drops. We tested the system's performance when patient demand increased by 10% and 20%. The results indicated that as patient demand increased, adding more nurses yielded consistent outcomes across different demand levels. That is, increasing the number of nurses generally decreases utilization while barely mitigating the adverse effects on average visit length as patient demand increases. In addition, the analysis underscores that higher demand amplifies nurses' need to maintain efficiency and utilization. It is important to note that the current nurse levels, ranging from 9 to 11, can adequately meet the current demand, even if it slightly increases. However, if the nurse levels drop to 8 or 7, they will become a bottleneck

resource, when we anticipate a noticeable increase in both nurse utilization and the average length of visits (to be further discussed in Section 4.3).

4.2.3 Impact of balancing the schedule

Currently, the Rutgers Cancer Institute schedules patients in the morning to complete daily tasks more quickly and avoid overtime. However, our observations and data indicate that this scheduling approach leads to a high demand for nurses and treatment chairs around noon. This results in significantly longer waiting times for patients arriving around that time. To address the issue of overly concentrated scheduling, the hospital is seeking a potential solution to distribute the number of visits more evenly throughout the day.

Our approach focuses on balancing the schedule to optimize patient flow and nurse workload, aiming to minimize patient visit length while maintaining nurse utilization within a reasonable range. The Rutgers Cancer Institute team recommends that anticipated nurse utilization should be maintained below 60% for current and 110% of increased outpatient demand levels, and should not exceed 65% as demand increases to 120%. This ensures that the quality of treatment remains high even with increased demand. To achieve this, we propose distributing patient appointments more evenly throughout the day, instead of concentrating them during specific periods. This approach aims to minimize peaks and valleys in patient arrivals, resulting in a smoother workflow and better resource management. By distributing appointments evenly, we ensure that nurses are consistently utilized throughout their shifts, which reduces idle time and prevents overburdening during peak hours.

Given the large size of the design space for this problem, we formulated a simulation optimization problem to find the optimal arrival rates across the operating hours, such that the patient's visit length metric is minimized. Specifically, we consider the operating periods in the set $T = \{8, 9, \dots, 15\}$, where each member of T is the start time of a 1-hour period. Our goal is to determine the optimal patient arrival rates x_t for each hour $t \in T$ to maximize efficiency while meeting service constraints. The visit length of patient i , observed in replication j of the simulation with arrival rate $\mathbf{x}$, is denoted by $L_{i,j}(\mathbf{x})$, where each replication simulates a day of operations. The nurse service duration of patient i , observed in replication j of the simulation with arrival rate $\mathbf{x}$, is denoted by $D_{i,j}(\mathbf{x})$. We must ensure that the expected daily nurse utilization does not exceed the threshold U , and that the number of patients served is the daily outpatient arrival-rate load C . Our optimization model incorporates these factors to find the best $\mathbf{x}$ values that satisfy all constraints while accounting for the variability in service times. The notation used in the model is summarized in Table 8.

The respective simulation optimization problem can be expressed mathematically as follows:

$$\min_{\mathbf{x}} \quad \mathbb{E} \left[\frac{\sum_{i=1}^{n_j} L_{i,j}(\mathbf{x})}{n_j} \right] \quad (1)$$

$$\text{s.t.} \quad \mathbb{E} \left[\frac{\sum_{i=1}^{n_j} D_{i,j}(\mathbf{x})}{W_{\text{total}}} \right] \leq U \quad (2)$$

$$\sum_{t \in T} x_t = C \quad (3)$$

$$0 \leq x_t \leq 8 \quad \forall t \in T \quad (4)$$

Constraint (2) ensures that the expected nurse utilization does not exceed the given threshold. Constraint (3) ensures that the sum of total patient arrival rates equals daily outpatient

Table 8 Notation for the simulation optimization model

Notation	Description
$T = \{8, 9, \dots, 15\}$	Set of time periods, representing one-hour operating periods from 8 AM to 3 PM
$\mathbf{x} = (x_8, x_9, \dots, x_{15})$	Decision variable; a vector of patient arrival rates during each operating hour $t \in T$
n_j	Number of patients observed in the j -th replication of the simulation
$L_{i,j}(\mathbf{x})$	Visit length of the i -th patient observed in the j -th replication of the simulation, given decision variable $\mathbf{x}$
$D_{i,j}(\mathbf{x})$	Nurse service duration for the i -th patient observed in the j -th replication of the simulation, given decision variable $\mathbf{x}$
W_{total}	Total daily available nurse-hours, representing the total capacity of onsite nurses, where each nurse's available hours are their working hours times the maximum number of patients they can serve simultaneously.
C	Daily outpatient arrival-rate load; $C = C_0$ representing the current daily patient arrival (i.e., the summation of current hourly arrival rates)
U	Utilization threshold

arrival-rate load C . Note that C equals the current daily outpatient arrival rate C_0 when $\delta = 100\%$, and $C = \delta C_0$ as we increase δ . Constraint (4) limits the hourly patient arrival rate to the maximum hourly slot in the hospital appointment system. These constraints collectively ensure a balanced workload, controlled arrival rates, and adherence to the system's daily capacity.

To develop balanced schedules, we solved the simulation optimization problem using OptQuest in AnyLogic. Note that the solution space for $\mathbf{x}$ is inherently large because x_t can take on any continuous value in the range of $[0, 8]$, making the simulation-optimization problem computationally challenging. To reduce the computational complexity, we discretized the solution space of each x_t in increments of 0.1 within $[0, 8]$. A finer granularity could be used to further refine the results, although this would increase computational complexity.

Table 9 compares the results of the optimized schedule with the current schedule. The optimized patient scheduling scenario resulted in more consistent and optimized average visit length and nurse utilization across scenarios without adding nurses while using an off-site lab. Specifically, in the optimized schedule, the average visit length was significantly reduced from 256.35 minutes to 225.78 minutes, and nurse utilization dropped from 54.13% to 51.87%. This strategy optimizes resource allocation and efficiently handles varying levels of demand, making it an effective solution for improving patient flow and nurse workload management. Table 10 presents the optimized outpatient appointment schedule driven by the hourly arrival rate, which shows a more balanced patient arrival schedule compared to the previous schedule.

In addition to the results presented for $\delta = 100\%$, we further optimized the arrival rates for $\delta = 110\%$ and $\delta = 120\%$. Specifically, we first simulated the system by increasing the current hourly arrival rate according to each value of δ , and then ran the simulation-based optimization model (1)–(4) by setting $C = \delta C_0$. The last two columns in Table 9 present the simulation results for 110% and 120% demand levels. As seen, similar trends have been observed at 110% and 120% demand levels compared to that at the current demand level.

Table 9 Simulation results: current vs. optimized schedule

γ	δ					
	Level A (100%)		Level B (110%)		Level C (120%)	
	Avg. Visit Len. (min)	Nurse Util. (%)	Avg. Visit Len. (min)	Nurse Util. (%)	Avg. Visit Len. (min)	Nurse Util. (%)
Current	256.35	54.13%	266.13	59.30%	280.87	65.53%
Optimized	225.78	51.87%	228.54	57.33%	232.93	62.37%

Table 10 Hourly arrival rates from simulation optimization

Time	Current (patients/hr)	Optimized (patients/hr)
8am	7.6	5.5
9am	8.6	3.3
10am	7.0	4.6
11am	5.3	5.2
12pm	3.3	2.3
1pm	1.8	2.9
2pm	1.7	6.5
3pm	0.9	5.9
Total	36.2	36.2

This approach reduces expected visit length and ensures more consistent use of nursing resources throughout the day. As a result, the nurses' workload becomes more manageable, reducing burnout and enhancing job satisfaction. In addition, this strategy is highly cost-effective because it utilizes existing resources more efficiently without incurring additional costs from hiring more staff or restructuring an on-site lab. By optimizing existing resources, the Rutgers Cancer Institute can improve patient care and operational efficiency, making it a sustainable and cost-effective solution for managing patient flow and resource allocation. The results suggest that the Rutgers Cancer Institute should consider implementing a balanced strategy to inform scheduling and resource allocation.

4.3 Managerial insights

In this study, our research team applied simulation models to evaluate the effectiveness of setting up an on-site lab, increasing nurse staffing, and balancing patient scheduling in addressing ongoing operational challenges and anticipated demand increases at the Rutgers Cancer Institute. The analysis provides insights for the Rutgers Cancer Institute on how to leverage existing resources to improve its operational efficiency and patient experience.

Currently, as part of the RWJBHealth system, the Rutgers Cancer Institute is using an external lab located in a neighboring RWJBHealth hospital to process all patients' blood samples. Sharing the lab across healthcare facilities might sound economically viable to the healthcare system; combining blood samples from the ambulatory unit of the Rutgers Cancer Institute with those from the inpatient-based hospital, however, creates a bottleneck as patients have to wait for the lab results before proceeding with chemotherapies. Our simulation results show that setting up an on-site lab significantly reduces average patient

visit length, suggesting that shifting from a shared lab to a designated lab helps address the bottleneck at the Rutgers Cancer Institute. This approach also demonstrates a more stable response to increasing outpatient demand, indicating a robust system capable of handling fluctuations in patient volume without overburdening resources. Our analyses suggest that healthcare systems should reevaluate their resource-pooling strategies. While such strategies may help maximize resource utilization at the system level, they could also lead to undesirable outcomes for local clinical units.

Traditionally, healthcare providers plan to staff more nurses to address increasing patient demand with the hope of reducing visit length and increasing patient turnover. While it is logical to assume that adding labor and other resources helps speed up patient flow, such a relationship is not guaranteed. Our study compares multiple scenarios and shows that the strategy helps reduce utilization, yet is less effective in reducing patient visit length. Specifically, increasing the number of nurses appears to be least effective when switching from 10 to 11 nurses in our simulation. Adding one nurse only decreases waiting time by less than one minute across all three scenarios, suggesting that staffing more nurses can be ineffective in reducing waiting time or improving patient experience. Hospital administrators must evaluate the cost/benefit tradeoff of alternative staffing strategies and choose the optimal strategy per their performance objective (e.g., reducing utilization and shortening visit length are likely to require different strategies).

Currently, the Rutgers Cancer Institute schedules most patient visits in the morning to ensure sufficient buffer time to complete all visits within the same day. While this approach appears accommodating and patient-friendly, it creates multiple issues. First, it creates unnecessary crowding as patients have to wait extra time due to the above-mentioned resource constraints in the morning. Based on the team's observation, the clinic is busiest between 9 am and 12:30 pm. The clinic appears to have minimal patient visits after 3 pm on weekdays. The imbalance in patient scheduling also creates challenges in staffing. Since nurses work on a shift basis, most are scheduled for the morning shifts, leaving the afternoon time slots underutilized. Our simulation suggests that the Rutgers Cancer Institute should adopt a balanced patient schedule, which involves distributing appointments more evenly throughout the day rather than concentrating them in the morning. This approach offers significant benefits, including reduced flow time and improved nurse utilization. By maintaining consistent and optimal visit length and nurse utilization across different scenarios, the balanced schedule strategy proves to be more economical and easier to implement, requiring no additional investment or extra nurses, provided there is cooperation among staff.

The optimal strategy for maximizing performance involves combining the on-site strategy with the updated, balanced schedule, which consistently outperforms merely applying the on-site strategy with the existing schedule at all demand levels—100%, 110%, and 120% (see Table 11). This strategy saves 75.32 minutes of average visit length and slightly reduces nurse utilization at the current demand (100%). When the updated patient schedule is accepted at the current demand level (100%), it further enhances performance by reducing average visit length by an additional 33.29 minutes, representing a 15.5% improvement compared to relying solely on the on-site strategy. As demand increases, the time savings become even more substantial: 39.30 minutes at 110% demand and 48.35 minutes at 120% demand. This indicates that the proposed approach is not only the best choice at current demand levels but also becomes increasingly valuable as demand escalates. The more significant average visit length reductions under higher demand scenarios highlight that this strategy is better in managing the rising number of patient visits, offering a more robust solution for handling patient surges effectively.

Table 11 Average visit length (minutes) and nurse utilization under various configurations

Configuration			Daily Outpatient Volume (δ)							
On-site Lab (α)	Nurse Number (β)	Schedule Balance (γ)	Level A (100%)			Level B (110%)			Level C (120%)	
			Avg. Visit Len. (min)	Nurse Util. (%)	Nurse Util. (%)	Avg. Visit Len. (min)	Nurse Util. (%)	Nurse Util. (%)	Avg. Visit Len. (min)	Nurse Util. (%)
✓			256.35	54.13%	54.13%	266.13	59.30%	59.30%	280.87	65.53%
			214.32		52.90%	224.61	57.92%	57.92%	238.35	64.01%
		✓	181.03		52.12%	185.31	57.23%	57.23%	190.00	62.41%
✓	✓	✓	180.96	46.59%	46.59%	185.07	51.55%	51.55%	189.51	56.06%

Interestingly, Table 11 also suggests that merely increasing the number of nurses, while decreasing nurse utilization by 5–7%, does not significantly enhance visit length metrics when integrated with these two strategies at various demand levels. Specifically, at 100% demand, adding one nurse only reduces the average visit length by 0.07 minutes, representing a negligible improvement. Similarly, at 110% and 120% demand levels, the reductions remain negligible at 0.24 and 0.49 minutes, respectively. This challenges the traditional perception that adding nurses necessarily yields the greatest marginal gains in reducing waiting times, suggesting instead that under certain constraints, other strategies (e.g., on-site lab, balanced scheduling) may be more effective in achieving the desired operational outcomes. However, while increasing nurse levels may not substantially improve efficiency under current demand conditions, it could become a critical factor in optimizing performance as patient demand increases beyond current capacity or nurse level drops below current levels of nine. To support this argument, we conducted additional experiments by reducing the number of nurses from nine to seven. As a result, the average visit length increased by 5.71, 8.50, and 10.45 minutes across the three demand levels. This indicates the presence of a critical threshold below which nurses become the primary system bottleneck. At this point, jointly optimizing all three factors becomes essential.

This unintuitive finding underscores the complexity of oncology operations, where high levels of interaction between different departments result in intricate personnel dynamics. Conventional assumptions may not fully address these complexities, highlighting that judgments based merely on experience are insufficient to improve the actual workflow meaningfully. In such settings, staff from various specialties must coordinate closely to manage patient care, often leading to situations where changes in one area, such as increasing nurse staffing, do not yield straightforward improvements in efficiency. Instead, these interactions can produce unexpected outcomes, as the workflows of nurses, doctors, and other healthcare professionals are deeply intertwined. This further emphasizes the necessity of simulation-based studies in tackling the challenges of OCP management.

The Rutgers Cancer Institute was highly receptive to our recommendations, and in line with the gleaned insights, the clinic has already begun trialing the first phase of improvements by focusing on laboratory efficiency. In particular, the simulation model runs and the associated data outputs revealed that laboratory services were a prominent pain point. Due to this, a new “fast track” system was developed to expedite laboratory services for patients during their ambulatory visits, which often include infusion of anti-cancer agents in addition to oncology provider visits. The associated intervention led to a significant reduction in patient waiting time, with more than 30% improved productivity for this complex care component at the cancer center. The Rutgers Cancer Institute is evaluating other recommendations for implementation in the future, particularly as the Rutgers Cancer Institute will be transitioning to the new location. This move will provide an opportunity to introduce a balanced scheduling strategy and make further adjustments as operations stabilize, ensuring that the clinic can continue to meet increasing patient demands efficiently and effectively.

5 Conclusion and future research

This study addressed critical issues in process optimization and resource allocation efficiency at the Rutgers Cancer Institute, providing specific improvement strategies through a systematic design of experiments approach. By analyzing factors such as size of nursing staff, appointment schedule balancing, and patient arrival rates, we identified the effective-

ness of individual factors and the optimal combination of these factors, considering their interrelationships and synergistic effects.

Unlike previous studies, our approach did not focus solely on the impact of individual factors on process improvement. Instead, we explored potential optimal configurations of factor combinations. Additionally, we tested these recommendations under multiple demand scenarios to ensure the robustness and long-term efficacy of the strategies, verifying that the proposed improvements are effective across different operational settings. Specifically, by implementing on-site lab work, adjusting nursing staff levels, and balancing the appointment schedule, we proposed operational improvements for the Rutgers Cancer Institute. These measures are designed to enhance process efficiency, reduce patient visit length, and maintain nurse utilization within an optimal range, ultimately improving patient satisfaction and reducing operational costs without compromising quality.

Although we incorporated relevant operational factors and process flows into the simulation and subsequent optimization, we acknowledge that other elements, such as physician preferences, patient preferences, and chemotherapy treatment plans, could potentially influence our results. Our model assumes relatively homogeneous treatment protocols and does not explicitly capture physician-specific or patient-specific preferences that may affect scheduling and service times. In practice, treatment plan complexity varies across patients (e.g., multi-drug regimens versus single-agent therapies), and some patients may have preferences for specific nurses or time slots. Incorporating such heterogeneity could reveal additional scheduling constraints or opportunities. For instance, clustering complex cases in off-peak hours might reduce average visit length for routine patients while ensuring adequate attention for high-acuity cases. Future extensions could integrate patient acuity scores or regimen complexity indices as model parameters. However, this information was not readily available during our study period or was restricted due to Health Insurance Portability and Accountability Act (HIPAA) considerations.

While our optimization framework focuses on minimizing average visit length subject to nurse utilization constraints, healthcare operations involve broader trade-offs. Reducing visit length generally improves patient satisfaction by decreasing time spent in clinical settings; however, aggressive scheduling optimization could increase staff workload stress, potentially affecting care quality and employee retention. Cost considerations also interact with these metrics: on-site laboratory integration reduces visit length but requires capital investment and ongoing operational expenses. Similarly, adding nursing staff decreases nurse utilization but increases labor costs. A comprehensive implementation would benefit from multi-objective optimization that explicitly models these trade-offs, allowing decision-makers to explore the Pareto frontier of solutions balancing efficiency, cost, staff well-being, and patient experience.

Additionally, while pharmacy mixing processes can significantly impact patient flow in many oncology clinics, the leadership at the Rutgers Cancer Institute reported adequate pharmacy resources, with multiple compounding hoods and pharmacists on shift. In the current state, the pharmacy was not identified as a critical bottleneck. Preliminary observations and discussions with clinical staff indicated that nurses typically collected medications without significant queuing. However, as volume increases or if operational changes shift workflow timings, pharmacy delays could become more prominent. Future expansions of our simulation will explicitly model pharmacy compounding steps and resource constraints to capture their potential effect on overall patient cycle times. Such an extension might reveal additional optimization opportunities, particularly if pharmacy delays interact with nurse scheduling during peak hours.

Furthermore, the findings from this study are not only relevant to the Rutgers Cancer Institute but also have broader implications for oncology departments and healthcare institutions

worldwide. Many of the challenges identified—such as managing possible patient demand surges, improving resource utilization, and reducing waiting times—are shared across oncology clinics, irrespective of geographic location. The ABS model used in this research can be adapted to various healthcare environments, making it a valuable tool for different oncology clinics, regardless of size or resource level. In settings with limited infrastructure, strategies like balanced scheduling and optimized resource allocation can provide significant improvements without requiring substantial financial investments.

Beyond oncology departments, the principles outlined in this study can also be applied to other areas of healthcare, such as outpatient clinics, emergency departments, and surgical units, where similar operational inefficiencies and resource management challenges exist. By strategically managing resources and patient flows, healthcare administrators can improve service delivery and patient outcomes in various medical environments.

Acknowledgements This research program was financially supported by Dr. Andrew M. Evens. The authors gratefully acknowledge his generous support and contributions to the project. Special thanks are also extended to the interns from the Master of Science in Healthcare Analytics and Intelligence program – Dhvani Mashru, Kunal Mehta, Nayib Dumar, Gurleen Kour, and Iram Contractor – for their efforts in collecting observation data. Additionally, we are grateful for the generous support provided by the Dean's Office at Rutgers Business School, which was instrumental to the successful completion of this project.

Declarations

Conflicts of Interest The authors report there are no competing interests to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdou, M., Hamill, L., & Gilbert, N. (2012). In: Heppenstall, A. J., Crooks, A. T., See, L. M., Batty, M. (eds.) *Designing and Building an Agent-Based Model*, pp. 141–165
- Ahmed, Z., ElMekkawy, T., & Bates, S. (2011). Developing an efficient scheduling template of a chemotherapy treatment unit: A case study. *Australasian Medical Journal*, 4(10), 575–582.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Alvarado, M., & Ntaimo, L. (2018). Chemotherapy appointment scheduling under uncertainty using mean-risk stochastic integer programming. *Health Care Manag Sci*, 21(1), 87–104.
- Baril, C., & Gascon, V. J. M. (2020). Design of experiments and discrete-event simulation to study oncology nurse workload. *IIESE Transactions on Healthcare Systems Engineering* 10(1), 74–86
- Baril, C., Gascon, V., Miller, J., & Côté, N. (2016). Use of a discrete-event simulation in a Kaizen event: A case study in healthcare. *European Journal of Operational Research*, 249(1), 327–339.
- Chen, W., Gao, S., Chen, W., & Du, J. (2023). Optimizing resource allocation in service systems via simulation: A bayesian formulation. *Production and Operations Management*, 32(1), 65–81.
- Corsini, R. R., Costa, A., Fichera, S., & Pluchino, A. (2023). A configurable computer simulation model for reducing patient waiting time in oncology departments. *Health Systems*, 12(2), 208–222.
- Dollinger, M. (1996). Guidelines for hospitalization for chemotherapy. *The Oncologist*, 1–2, 107–111.
- Egea, J. A., Vázquez, E., Banga, J., & Marti, R. (2009). Improved scatter search for the global optimization of computationally expensive dynamic models. *Journal of Global Optimization*, 43, 175–190.

- Garaix, T., Rostami, S., & Xie, X. (2020). Daily outpatient chemotherapy appointment scheduling with random deferrals. *Flexible Services and Manufacturing Journal*, 32(1), 129–153. <https://doi.org/10.1007/s10696-018-9326-x>
- Goswami, A., Baveja, A., Ding, X., Melamed, B., & Roberts, F. (2024). An integrated framework for modeling pharmaceutical supply chains with disruptions and risk mitigation. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-024-06381-y>
- Hadid, M., Elomri, A., Jouini, O., Kerbache, L., Saleh, A., & Hamad, A. (2022). Multi-objective simulation-based optimization for effective management of the outpatient chemotherapy process. *IFAC-PapersOnLine*, 55, 1639–1644.
- Hadid, M., Elomri, A., Mekaw, T. E., Jouini, O., Kerbache, L., & Hamad, A. (2022). Operations management of outpatient chemotherapy process: An optimization-oriented comprehensive review. *Operations Research Perspectives*, 9, Article 100214.
- Hendershot, E., Murphy, C., Doyle, S., Van-Cleef, J., Lowry, J., & Honeyford, L. (2005). Outpatient chemotherapy administration: Decreasing wait times for patients and families. *Journal of Pediatric Oncology Nursing*, 22, 31–37.
- Heshmat, M., & Eltawil, A. (2017). A new approach to solve operations planning problems of the outpatient chemotherapy process. In: 4th International Conference on Control, Decision and Information Technologies (CoDIT), pp. 544–549
- Heshmat, M., & Eltawil, A. (2021). Solving operational problems in outpatient chemotherapy clinics using mathematical programming and simulation. *Annals of Operations Research*, 298(1–2), 289–306.
- Heshmat, M., Eltawil, A. B., & Abdelghany, M. (2022). Operations planning in outpatient chemotherapy with hybrid simulation modelling. *International Journal of Simulation and Process Modelling*, 18(4), 296–303.
- Jalalimanesh, A., Haghighi, H. S., Ahmadi, A., & Soltani, M. (2017). Simulation-based optimization of radiotherapy: Agent-based modeling and reinforcement learning. *Mathematics and Computers in Simulation*, 133, 235–248.
- Jarrett, A. M., Faghihi, D., Hormuth, D., Lima, E., Virostko, J., Biros, G., Patt, D., & Yankeelov, T. (2020). Optimal control theory for personalized therapeutic regimens in oncology: Background, history, challenges, and opportunities. *Journal of Clinical Medicine* 9
- Lamé, G., Jouini, O., & Cardinal, J.S.-L. (2016). Outpatient chemotherapy planning: A literature review with insights from a case study. *IIE Transactions on Healthcare Systems Engineering*, 6, 127–139.
- Lamé, G., Jouini, O., & Cardinal, J. S. (2020). Combining soft systems methodology, ethnographic observation, and discrete-event simulation: A case study in cancer care. *Journal of the Operational Research Society*, 71(10), 1545–1562.
- Liang, B., Turkcan, A., Ceyhan, M. E., & Stuart, K. (2015). Improvement of chemotherapy patient flow and scheduling in an outpatient oncology clinic. *International Journal of Production Research*, 53(24), 7177–7190.
- Liu, E., Ma, X., Saure, A., Weber, L., Puterman, M. L., & Tyldesley, S. (2019). Improving access to chemotherapy through enhanced capacity planning and patient scheduling. *IIE Transactions on Healthcare Systems Engineering*, 9(1), 1–13.
- Lu, T., Wang, S., Li, J., Lucas, P., Anderson, M., & Ross, K. (2012). A simulation study to improve performance in the preparation and delivery of antineoplastic medications at a community hospital. *Journal of Medical Systems*, 36(5), 3083–3089.
- Masselink, I. H., Mijden, T. L., Litvak, N., & vanBerkel, P. T. (2012). Preparation of chemotherapy drugs: Planning policy for reduced waiting times. *Omega*, 40(2), 181–187.
- Matta, M. E., & Patterson, S. S. (2007). Evaluating multiple performance measures across several dimensions at a multi-facility outpatient center. *Health Care Manag Sci*, 10(2), 173–194.
- Meaney, C., Frein, A. M. D., & Heckmann, P. (2017). Supporting integrated care through national standardisation of chemotherapy protocols. *International Journal of Integrated Care*, 17, 515.
- Mustapha, K., Gilli, Q., Frayret, J.-M., Lahrichi, N., & Karimi, E. (2016). Agent-based simulation patient model for colon and colorectal cancer care trajectory. *Procedia Computer Science*, 100, 188–197.
- Santibáñez, P., Chow, V. S., French, J., Puterman, M., & Tyldesley, S. (2009). Reducing patient wait times and improving resource utilization at british columbia cancer agency's ambulatory care unit through simulation. *Health Care Manag Sci*, 12(4), 392–407.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464.
- Slocum, R. F., Jones, H. L., Fletcher, M. T., McConnell, B. M., Hodgson, T. J., Taheri, J., & Wilson, J. R. (2021). Improving chemotherapy infusion operations through the simulation of scheduling heuristics: a case study. *Health Systems*, 10(3), 163–178.
- Suss, S., Bhuiyan, N., Demirli, K., & Batist, G. (2018). Achieving level patient flow in an outpatient oncology clinic. *IIE Transactions on Healthcare Systems Engineering*, 8(1), 47–58.

- Todres, R., & Wojtiuk, R. (1979). The cancer patient's view of chemotherapy. *Cancer Nursing*, 2(4), 283–286.
- Ugray, Z., Lasdon, L., Plummer, J., Glover, F., Kelly, J., & Martí, R. (2007). Scatter search and local nlp solvers: A multistart framework for global optimization. *INFORMS Journal on Computing*, 19(3), 328–340.
- Wilson, B. E., Jacob, S., Yap, M. L., Ferlay, J., Bray, F., & Barton, M. (2019). Estimates of global chemotherapy demands and corresponding physician workforce requirements for 2018 and 2040: a population-based study. *The Lancet Oncology*, 20(6), 769–780.
- Woodall, J. C., Gosselin, T., Boswell, A., Murr, M., & Denton, B. T. (2013). Improving patient access to chemotherapy treatment at duke cancer institute. *Interfaces*, 43(5), 449–461.
- Yokouchi, M., Aoki, S., Sang, H., Zhao, R., & Takakuwa, S. (2012). Operations analysis and appointment scheduling for an outpatient chemotherapy department. In: Proceedings of the 2012 Winter Simulation Conference (WSC), pp. 1–12
- Zhan, Q., Fu, J., Fu, F., Zhang, J., & Wang, C. (2017). Survival and time to initiation of adjuvant chemotherapy among breast cancer patients: a systematic review and meta-analysis. *Oncotarget*, 9, 2739–2751.
- Zúñiga, E. R., Moris, M. U., Syberfeldt, A., Fathi, M., & Rubio-Romero, J. C. (2020). A simulation-based optimization methodology for facility layout design in manufacturing. *IEEE Access*, 8, 163818–163828.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.