



A Simulation Modelling Study of Referral Distribution Policies in a Centralized Intake System for Surgical Consultation

Deborah A. Marshall¹ · Toni Tagimacruz¹ · Monica Cepoiu-Martin² · Jill Robert³ · Bernice Ring⁴ · Michael Burston⁵ · Suzanne Higgins⁴ · Monica Hess⁵ · Jonathan White⁴

Received: 1 March 2022 / Accepted: 1 December 2022 / Published online: 30 December 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Delays beyond recommended wait times, especially for specialist services, are associated with adverse health outcomes. The Alberta Surgical Initiative aims to improve the referral wait time—the time between a referral is received at the central intake to the time a specialist sees the patient. Using the discrete event simulation modelling approach, we evaluated and compared the impact of four referral distribution policies in a central intake system on three system performance measures (number of consultations, referral wait time and surgeon utilization). The model was co-designed with clinicians and clinic staff to represent the flow of patients through the system. We used data from the Facilitated Access to Surgical Treatment (FAST) centralized intake referral program for General Surgery to parameterize the model. Four distribution policies were evaluated – next-available-surgeon, sequential, "blackjack," and "kanban." A sequential distribution of referrals for surgical consultation among the surgeons resulted in the worst performance in terms of the number of consultations, referral wait time and surgeon utilization. The three other distribution policies are comparable in performance. The "next available surgeon" model provided the most efficient and robust model, with approximately 1,000 more consultations, 100 days shorter referral time and a 14% increase in surgeon utilization. Discrete event simulation (DES) modelling can be an effective tool to illustrate and communicate the impact of the referral distribution policy on system performance in terms of the number of consultations, referral wait time and surgeon utilization.

Keywords Surgical consultation · Wait time management · Discrete-event simulation model · Central intake · Referral and consultation · Access

Introduction

Delays beyond recommended wait times, especially for specialist services, are associated with adverse health outcomes and increased personal costs and costs to the healthcare

system [1–4]. Thus, access to surgical services, including consultation and surgery, is a priority area of improvement identified by Alberta Health Services (AHS). The Alberta Surgical Initiative (ASI) commits to ensuring that all scheduled surgeries in the province occur within clinically recommended timeframes by 2023. The ASI aims to improve the referral wait time, defined as the time between a referral sent by the referring physician to the time the patient consults with the specialist.

The Facilitated Access to Surgical Treatment (FAST) is a centralized intake and triage referral program that processes surgical referrals in the Edmonton Zone and intends to improve access to care [5]. The program has handled more than 52,000 referrals since its inception in 2015. Central intake systems have been shown to have a range of operational and quality-of-care benefits [6–8]. Distribution policies at the central intake are defined as the rules that govern the assignment of a patient to one of the surgeons providing

✉ Deborah A. Marshall
damarsha@ucalgary.ca

¹ Cumming School of Medicine, McCaig Bone and Joint Health Institute, University of Calgary, 3280 Hospital Drive NW, Calgary, AB T2N 4Z, Canada

² Department of Critical Care Medicine, Cumming School of Medicine, University of Calgary, Calgary, Canada

³ Surgery, Alberta Health Services, Bone & Joint Strategic Clinical NetworkTM, Alberta, Canada

⁴ Surgery Strategic Clinical NetworkTM, Alberta Health Services, Alberta, Canada

⁵ Alberta Health Services, Alberta, Canada

consultation for the condition specified in the referral. Currently, the FAST program operates using a "next-available-surgeon" policy. Referring providers are asked if they wish to refer to the surgeon with the shortest wait time or a particular surgeon of their choice. The vast majority of referring providers choose to refer to the "next available surgeon"; these patients are assigned to the surgeon who has both the expertise needed and the shortest waiting time based on recent data on their practice. Patients for whom a specific surgeon is requested are assigned to that surgeon regardless of waiting time. Various distribution policies have been used in the past by the central intake; these policies will be discussed in the next section.

In the healthcare literature, Gunal and Pidd [9], Katsaliaki and Mustafee [10], and Zang [11] published the most recent reviews of the literature on healthcare applications of DES. In these studies, operational decisions are "tested" in a simulation model to identify bottlenecks and match resources and demand to improve system performance [12–16]. The most common system performance measure is the time a patient waits between referral and consultation. In Canada, examples include a DES simulation model to aid capacity planning decisions and wait time reduction for General Surgery [17] and the analysis of the radiotherapy planning process [18].

This study evaluates and compares the impact of several referral distribution policies in a central intake system on

three system performance measures – the number of consultations, referral wait time and surgeon utilization.

Methods

We use discrete event simulation (DES) to represent the process of referral arriving at the central intake, being distributed to surgeons, and surgical consultations being completed, and compare four referral distribution policies for patients with gallbladder disease.

The modelling process

Our approach followed the key steps for simulation model development described in the literature [19–23]. The modelling process (Fig. 1) was a collaborative effort of a team of clinical staff, physicians and researchers from the AHS Surgery Strategic Clinical Network™ (SCN™) and researchers from the University of Calgary. We defined the system, its boundaries, the current distribution policy and the system performance measures and constructed the conceptual model of the current system through discussions and analyses of available data. The conceptual model was translated into a DES model implemented in the ARENA [24] software.

The system represented by the model is the referral to surgical consult system focused on the management of

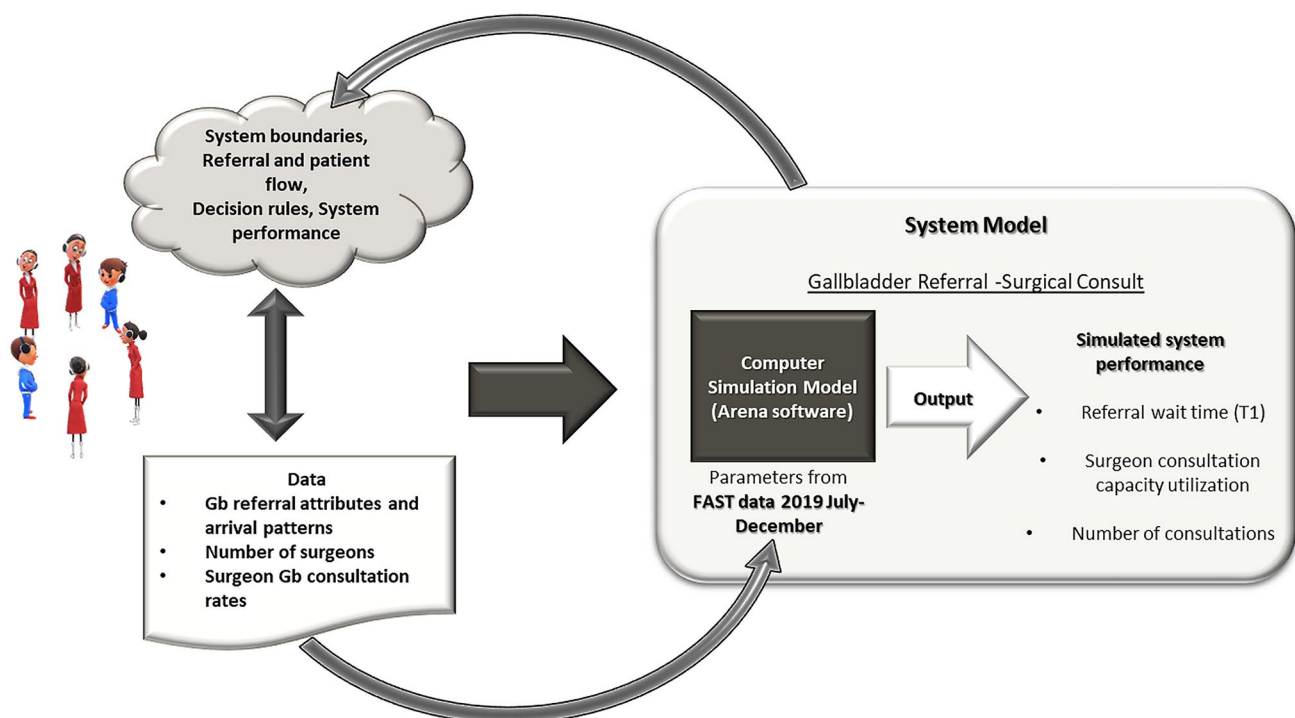


Fig. 1 Modelling Process

patients with gallbladder disease. This disease was chosen as an example of a common condition that often requires surgical consultation and on which recent data was available from the FAST program. The *entities* in the simulation model are the referrals arriving at the central intake and the patients linked to the referrals, each assigned to a surgeon for a surgical consult. The relevant *events* are the receipt of the referral at the central intake, the assignment of patients to surgeons for a surgical consult, the surgical consultation, and the patient's exit from the surgical consultation.

The relevant system *performance measures* collected in our simulation model were the mean referral wait time, the mean number of surgical consultations in the system and the surgeon consultation capacity utilization. The mean referral wait time was the number of days between the receipt of the referral at the central intake and the consultation with the surgeon, averaged across all the simulated referrals received after the model warm-up period and averaged across all the simulation replications. The mean number of surgical consultations was the number of consultations after the warm-up period averaged across all the simulation replications. Finally, the consultation capacity utilization is the average capacity utilization of each surgeon for all replications, averaged across all surgeons. Surgeon consultation capacity utilization is the ratio of the surgeon's time used for a consultation to the surgeon's available time for consultation.

The distribution policy used in our model is the next available surgeon. We run the ARENA model for 30 replications taking advantage of the software's feature to generate "random samples" for each replication from a defined. Figure 2 represents the flow of the entities in the simulation model.

Model parameters, assumptions, calibration and validation

The model was parameterized using the FAST program data collected between July and December 2019. The July to December period represents a "typical" or stable period of operation in terms of the volume of referrals and surgical consultations. A total of 659 gallbladder referrals were received in this period, for an average of 108 gallbladder referrals a month, distributed for a surgical consultation to 39 surgeons. The daily rate was assumed to fit a Poisson distribution, which satisfies the assumptions of independence between referral arrivals with a known average arrival rate and orderliness [25]. The number of arrivals, $x \sim \text{Poisson}(\frac{108}{30})$. From the data available, the probability that a referral requested a preferred surgeon was 14%. In building our simulation model, we assumed that a preferred surgeon would be requested with the same probability as observed in the data. Data on the existing waitlists for each surgeon were also provided to inform conditions at the beginning of the simulation for the period observed.

Because data on each surgeon's available gallbladder consultation capacity was unavailable, we calibrated our model by adjusting the observed consultation capacity. We used a fixed factor to multiply each surgeon's observed consultation capacity to achieve a reasonable level of agreement between the observed mean referral wait time (65 days, SD = 41, N = 659) and the simulated mean referral wait time (66 days, 95%CI[42,91], n = 30). The clinical staff, physicians and researchers involved in the AHS Surgery SCN validated the logical structure and results of the model.

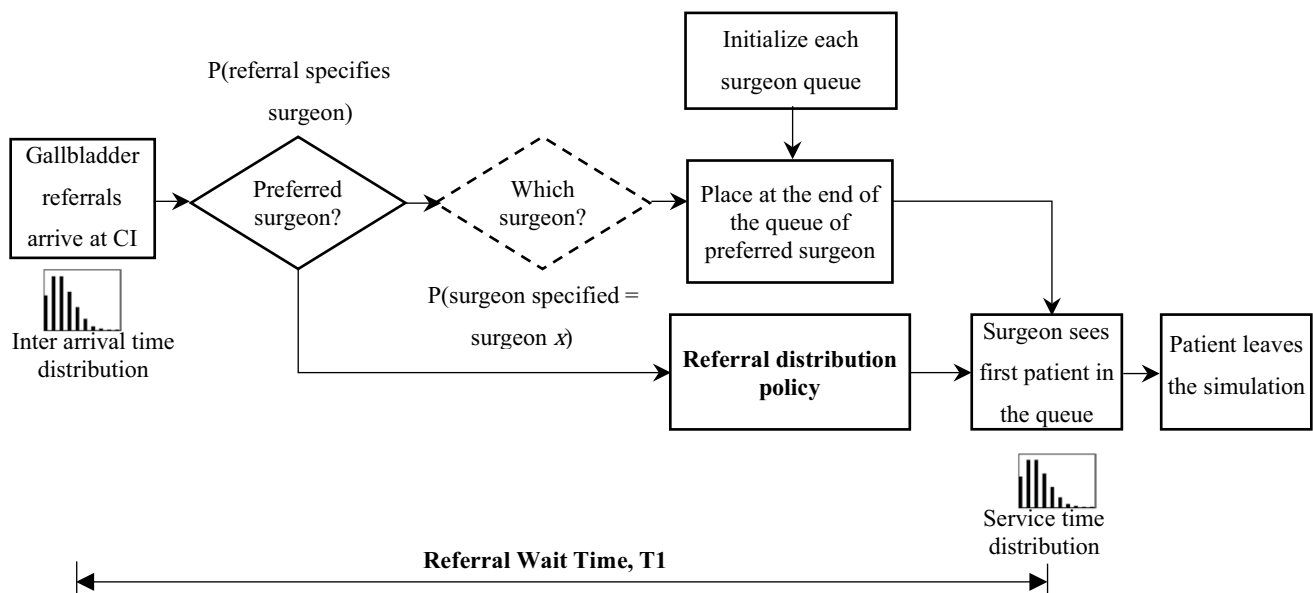
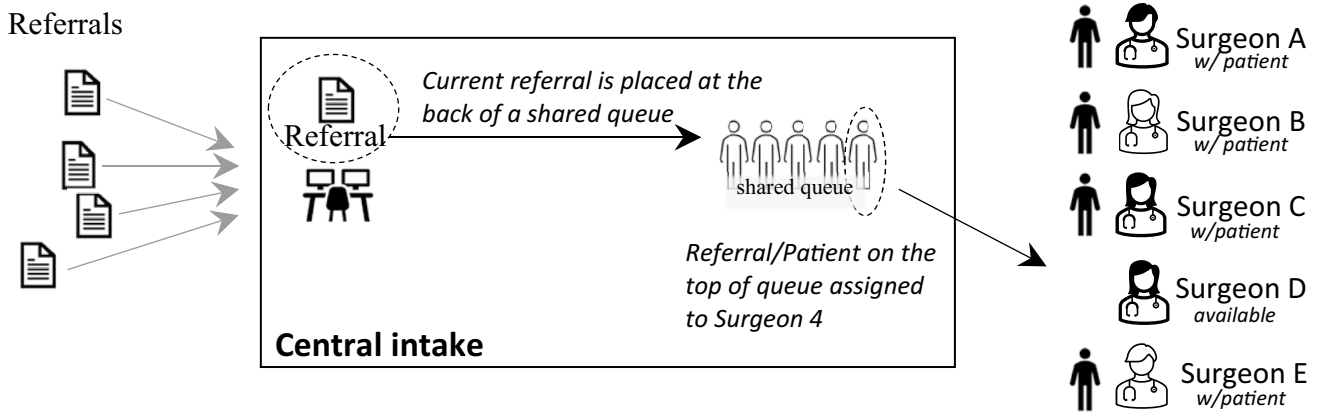


Fig. 2 Flow of arrival of referrals for gallbladder surgical consultation to surgical consultation in the simulation model



The arriving referral is placed in the back of the shared queue because there are other referrals waiting for a surgical consult. When surgeon D becomes available, the patient on the top of the queue is assigned to surgeon D.

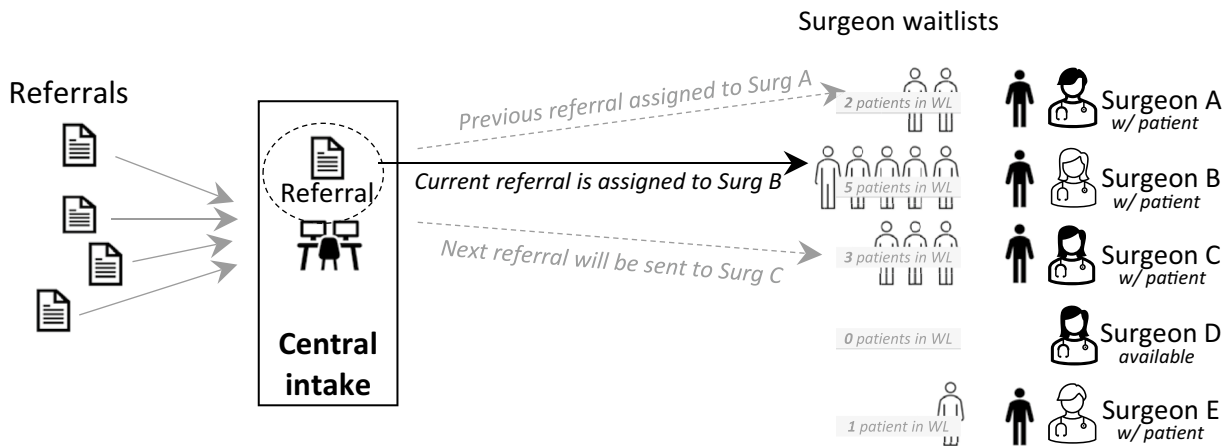
Fig. 3 Next-available-surgeon scenario example with five surgeons

Alternative distribution policies

Next-available-surgeon When a referral arrives at central intake, it is assigned to a surgeon, who can immediately provide a consultation. If no surgeon is available to provide the consultation or if other referrals are waiting, the referral is placed at the back of a shared queue. As soon as a surgeon becomes available, the referral at the top of the shared queue is assigned to the available surgeon (Fig. 3).

For referrals that request a preferred surgeon, the referral is placed in a queue for that surgeon. When the requested surgeon becomes available, the patient with the longest waiting time in the shared queue and the surgeon's queue will be seen by the surgeon.

Sequential distribution When a referral arrives at central intake, it is assigned to the surgeon next to the surgeon on the list who was assigned the previous referral (Fig. 4). Referrals



If the previous referral was assigned to surgeon A, the current referral is assigned to surgeon B. The following referral will be assigned to surgeon C, followed by surgeon D, then surgeon E/last surgeon and then back to the first surgeon on the list.

Fig. 4 Sequential scenario example with five surgeons

that request a preferred surgeon are sent to that surgeon. If the next referral does not request a preferred surgeon, the assignment resumes from the last surgeon in the sequence.

Blackjack distribution In this scenario, all surgeons have the same maximum waitlist size. Individual surgeons are assigned additional referrals so that all surgeons have a similar waitlist (Fig. 5). When all surgeon's waitlists have the maximum number, an arriving referral is held in a shared queue. According to their arrival order, referrals on the shared queue are assigned to surgeons' waitlists.

For referrals that request a preferred surgeon, the referral is placed in a queue for that surgeon. A surgeon may exceed the maximum waitlist size if a significant number of referrals request that surgeon.

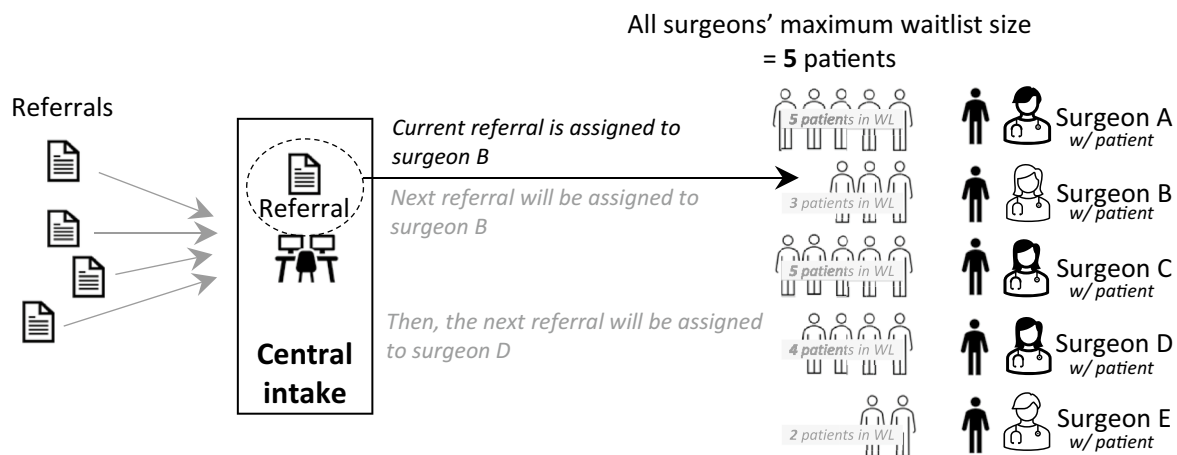
Kanban distribution Each surgeon has a waitlist size equal to their monthly consult rate in this scenario. When a referral arrives at the central intake, it is assigned to the next surgeon on the list whose waitlist size is less than his monthly consult rate (Fig. 6). When all surgeons' waitlists are full, an arriving referral is held in a shared queue and assigned to surgeons according to their arrival order.

For referrals that request a preferred surgeon, the referral is placed in a queue for that surgeon. A surgeon may exceed her waitlist size if a significant number of referrals request that surgeon.

Results

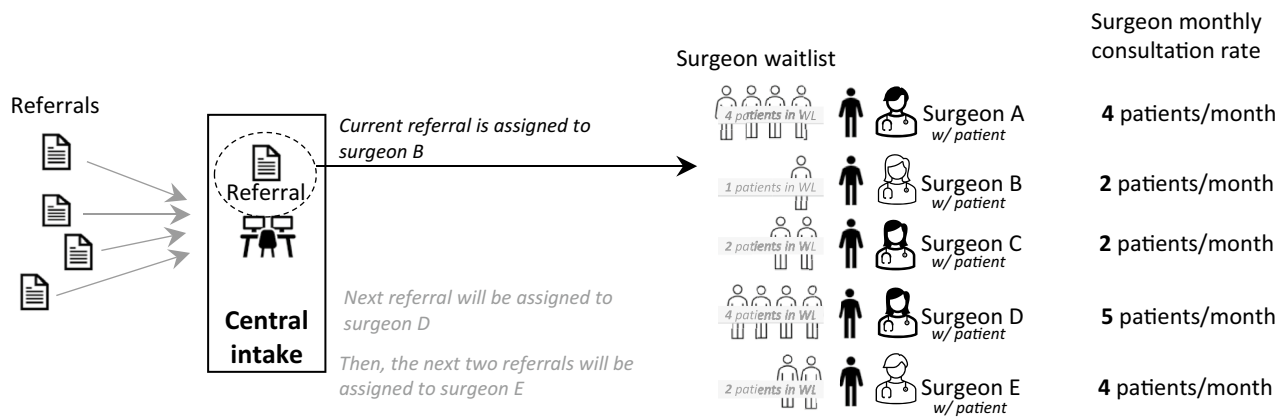
We ran the DES model for each of the four scenarios for 30 replications. Each replication had a simulation run of 3 years, including a 1-year warm-up. Table 1 summarizes the performance measures for the *Next-available-surgeon*, *Sequential*, *Blackjack*, and *Kanban* scenarios using the list of surgeons arranged randomly.

The number of gallbladder referrals assigned by central intake to surgical consult and number of gallbladder consultations The mean number of gallbladder referrals received at the central intake for all scenarios is comparable, as can be inferred from the overlapping 95% confidence interval of their means. The mean number of gallbladder consultations was significantly lower (inferred from the non-overlapping 95% CI) in the *Sequential*, 1,686 consultations 95% CI [1,686, 1,711] than in the *Next-available-surgeon*, 2,652 consultations



For example, the maximum waitlist size is five patients. When a referral arrives, it is assigned to the next surgeon in the list whose waitlist size is less than five patients. The current referral is assigned to surgeon B because surgeon A has a maximum of 5 patients on his waitlist and surgeon B only has three patients. Assuming that the next referral comes while the number of patients in each of the surgeon's waitlists remains the same, the next referral is still assigned to surgeon B. The number of patients on surgeon B's waitlist equals five. Suppose another referral arrives with no change in the number of patients in each of the surgeons' waitlists. This referral will be assigned to surgeon D and not surgeon C because surgeon D has only four patients on his waitlist while surgeon C already has five patients on his.

Fig. 5 Blackjack scenario example with five surgeons



The arriving referral is assigned to surgeon B because his waitlist size is lesser than his monthly consultation rate. Assuming that the next referral arrives while the number of patients in each of the surgeon's waitlists remains the same, the next referral will be assigned to surgeon D. The next two referrals will be assigned to surgeon E.

Fig. 6 Kanban scenario example with five surgeons

95% CI [2,633, 2,672], *Blackjack*, 2,567 consults 95% CI [2,545, 2,588], and *Kanban*, 2,595 consults 95% CI [2,574, 2,615] scenarios.

Mean referral wait time The mean referral wait time for the *Sequential* scenario, 172 days 95% CI [169, 175], is significantly higher (inferred from the non-overlapping 95% CI) than the other three scenarios – 66 days 95% CI [42, 91] for the *Next-available-surgeon* and 72 95% CI [66, 77] days for *Blackjack* and *Kanban*).

Mean consult capacity utilization The results show that the surgeons' consult capacity is fully utilized in the *Next-available-surgeon* scenario. The same can be said about the *Blackjack*, and *Kanban* scenarios, but with some variation across the 39 surgeons (SD = 3% for *Blackjack* and SD = 4% for *Kanban*). For the *Sequential* policy, the capacity utilization is much lower (86%) when compared to the other three scenarios, with a higher variation across the 39 surgeons (SD = 24%).

Table 1 Comparison of the performance measures of the four scenarios

	Next-available-surgeon	Sequential	Blackjack	Kanban
Mean number of gallbladder referrals received at Central Intake (2 years), 95% CI 95% CI [LL, UL] n = 30	2,628 referrals [2,614, 2,642]	2,624 referrals [2,608, 2,640]	2,617 referrals [2,601, 2,634]	2,620 referrals [2,602, 2,638]
Performance measures				
Mean number of gallbladder consults (2 years), 95% CI [LL, UL], n = 30	2,652 consults [2,633, 2,672]	1,686 consults [1,686, 1,711]	2,567 consults [2,545, 2,588]	2,595 consults [2,574, 2,615]
Mean Referral Wait Time, T2 to T5, n = 30				
Mean, 95% CI [LL, UL]	66 days [42, 91]	172 days [169, 175]	72 days [66, 77]	72 days [66, 77]
Minimum	35 days	0	1 days	6 days
25 th percentile	48 days	0	24 days	39 days
50% percentile (median)	58 days	60 days	47 days	54 days
75 th percentile	65 days	305 days	88 days	76 days
90 th percentile	72 days	513 days	167 days	115 days
95 th percentile	84 days	629 days	279 days	164 days
Maximum	953 days	1,016 days	994 days	950 days
Mean Consult Capacity Utilization				
Mean (SD), n = 39	100% (SD = 0)	86% (SD = 24%)	99% (SD = 3%)	99% (SD = 2%)

Table 2 Mean referral wait times of the four scenarios for patients who specified a surgeon in their referral and those who did not

Performance measure (30 replications)	Next-available-surgeon		Sequential		Blackjack		Kanban	
	w/o specific surgeon	w/ specific surgeon	w/o specific surgeon	w/ specific surgeon	w/o specific surgeon	w/ specific surgeon	w/o specific surgeon	w/ specific surgeon
Mean number of gallbladder consults (2 years), 95% CI [LL, UL]	2,338 cons [1,501, 3,175]	314 cons [202, 427]	1,483 cons [952, 2,013]	216 cons [138, 293]	2,271 cons [1,458, 3,083]	296 cons [190, 402]	2,292 cons [1,472, 3,112]	303 cons [194, 411]
Referral Wait Time, T2 to T5								
Mean, 95% CI [LL, UL]	55 days [49, 61]	148 days [139, 158]	169 days [90, 247]	194 days [111, 277]	61 days [45, 76]	156 days [92, 220]	61 days [45, 76]	156 days [92, 220]
0 th percentile (min)	35 days	38 days	0 days	0 days	1 days	4 days	8 days	10 days
25 th percentile	47 days	61 days	0 days	7 days	23 days	52 days	39 days	50 days
50% percentile (median)	56 days	76 days	57 days	95 days	43 days	104 days	53 days	81 days
75 th percentile	63 days	143 days	301 days	335 days	79 days	252 days	72 days	185 days
90 th percentile	69 days	376 days	509 days	536 days	134 days	443 days	99 days	411 days
95 th percentile	71 days	520 days	623 days	662 days	194 days	595 days	123 days	545 days
Maximum	75 days	953 days	998 days	977 days	908 days	960 days	607 days	949 days

Table 2 presents the expected referral wait times for the four scenarios by type of referral: referrals that requested a preferred surgeon and those that did not.

In all scenarios, the patients who requested a preferred surgeon waited longer than those who did not. In the Next-available surgeon policy, patients who requested a

surgeon waited 93 days longer than those who did not. In the Sequential scenario, those who requested a surgeon waited, on average, 21 days more than those who did not. For both the Blackjack and Kanban scenarios, patients who requested a surgeon waited, on average, 95 more days than those who did not.

Table 3 Sensitivity analysis varying Waitlist size Blackjack scenario performance measures' sensitivity to fixed waitlist size

	Waitlist size = 5	Waitlist size = 1	Waitlist size = 10	Waitlist size = 15
Mean number of gallbladder referrals received at Central Intake (2 years), 95% CI 95% CI [LL, UL] n = 30	2,617 referrals [2,601, 2,634]	2,622 referrals [2,604, 2,640]	2,614 referrals [2,599, 2,630]	2,623 referrals [2,602, 2,644]
Performance measures				
Mean number of gallbladder consults (2 years), 95% CI [LL, UL], n = 30	2,567 consults [2,545, 2,588]	2,638 consults [2,619, 2,658]	2,533 consults [2,518, 2,549]	2,480 consults [2,460, 2,500]
Mean Referral Wait Time, T2 to T5, n = 30				
Mean, 95% CI [LL, UL]	72 days [66, 77]	68 days [62, 73]	117 days [115, 118]	149 days [148, 150]
Minimum	1 days	2 days	0 days	0 days
25 th percentile	24 days	34 days	32 days	33 days
50% percentile (median)	47 days	48 days	75 days	108 days
75 th percentile	88 days	71 days	144 days	199 days
90 th percentile	167 days	111 days	262 days	341 days
95 th percentile	279 days	168 days	402 days	477 days
Maximum	994 days	941 days	1,028 days	1,019 days
Mean Consult Capacity Utilization				
Mean (SD), n = 39	99% (SD = 3%)	100% (SD = 1%)	99% (SD = 5%)	98% (SD = 6%)

Table 4 Kanban scenario performance measures' sensitivity to the number of months of "waitlisted patients"

	Waitlist size for Scenario D, Kanban, (1 X monthly Capacity)	Waitlist size = Roundup (2 X monthly Capacity)	Waitlist size = Roundup (3 X monthly Capacity)
Mean number of gallbladder referrals received at Central Intake (2 years), 95% CI 95% CI [LL, UL] n = 30	2,620 referrals [2,602, 2,638]	2,614 referrals [2,594, 2,634]	2,610 referrals [2,595, 2,624]
Performance measures			
Mean number of gallbladder consults (2 years), 95% CI [LL, UL], n = 30	2,595 consults [2,574, 2,615]	2,572 consults [2,556, 2,588]	2,530 consults [2,514, 2,547]
Mean Referral Wait Time, T2 to T5, n = 30			
Mean, 95% CI [LL, UL]	72 days [66, 77]	85 days [83, 88]	103 days [101, 104]
Minimum	6 days	2 days	0 days
25 th percentile	39 days	47 days	57 days
50% percentile (median)	54 days	70 days	94 days
75 th percentile	76 days	97 days	126 days
90 th percentile	115 days	141 days	171 days
95 th percentile	164 days	200 days	223 days
Maximum	950 days	954 days	957 days
Mean Consult Capacity Utilization			
Mean (SD), n = 39	99% (SD = 2%)	99% (SD = 4%)	99% (SD = 5%)

Impact of Blackjack and Kanban policy parameters on performance measures

In both scenarios, the parameter for the size of the waitlist for the Blackjack and the period for the calculation of waitlist size for the Kanban scenario was chosen based on discussions with the FAST program team. The following results show the impact on the performance measures of changes in these parameter values.

Blackjack distribution The size of the waitlist for all surgeons was pegged to 5 patients. Table 3 shows the result of the simulations for waitlists of 1, 10, and 15 patients. All the performance measures are sensitive to the size of the waitlist.

Kanban policy For each of the surgeons represented in the simulation, the number of gallbladder consultations per month is the size of the waitlist. Table 4 shows the result for a waitlist size equivalent to two months and three months' worth of consultations for each surgeon.

Discussion and conclusion

The simulation model described here determined that the *Next-available-surgeon* distribution policy performed best in terms of reducing waiting times for consultation and maximizing the utilization of resources. These results are consistent with the literature on distribution strategies,

and other authors have demonstrated the effectiveness of the *Next-available-surgeon* strategy in reducing wait times [26, 27]. The *Sequential* referral distribution policy resulted in the worst performance of the four distribution policies simulated. It was associated with the lowest number of surgical consultations delivered (1, 686 consults compared to the next lowest, Blackjack, 2,567 consults), the longest referral wait time (172 days, compared to the next longest, Blackjack and Kanban, 72 days), the lowest mean surgical consult capacity utilization (86% compared to almost 100% for the others) and the largest surgical consult utilization variation among surgeons (SD = 24% compared to the next largest variation, Blackjack = 3%). The other two distribution policies studied performed similarly to the *Next-available-surgeon* but were more sensitive to variation in waitlist size. Based on the simulation model results described here, we cannot recommend the *Sequential* distribution of referrals. The *Sequential* distribution performed the worst of all those examined and would be expected to reduce the efficient utilization of resources, increase variation between surgeons and reduce the timeliness of access for patients.

The *Sequential* strategy would be expected to be the least effective approach as it takes no account of a surgeon's existing waitlist and tends to increase variation between surgeons, adding patients to waitlists that are already long and not allocating more patients to surgeons with short waitlists. This is a significant finding, as many surgical groups use this strategy to distribute referrals between surgeons, assuming that a "one-for-me, one-for-you" policy is the "fairest" way to allocate the work. The results of our study suggest

that, in fact, this method of consult distribution results in more patients with referrals sent at an earlier date waiting longer than patients with a later referral date. Thus, there is a longer average wait for patients and a significant under-utilization of surgeon consultation capacity. This problem may be further compounded if all surgeons in the group do not share information about their waitlist for consultation or receive referrals from other sources, leading to significant inequity and variation between surgical practices. Surgical groups using a sequential model to distribute referrals might consider adopting an alternative model that is more efficient for reducing wait times.

Another interesting finding is that requesting a preferred surgeon for consultation increases the average time the patient waits to be seen for all the scenarios examined. This should be considered by both referring providers and patients when making choices about how they wish their referral to be handled. There may be a good reason that a patient elects to wait longer to see a preferred surgeon suggested by their doctor, but other patients may simply wish to be seen as quickly as possible.

The *Next-available-surgeon* distribution policy had the lowest mean referral wait time and the highest mean consultation capacity utilization across all surgeons. Still, these measures were similar in the *Blackjack* and *Kanban* models, suggesting that all three distribution strategies may have benefits to consider.

Implementation considerations also contribute to the choice of distribution policy. For the *Next-available-surgeon* scenario, referrals that do not request a preferred surgeon were held in a shared queue; regular monitoring of surgeon availability is required for this strategy, which may not be feasible for all programs. For the *Blackjack* scenario, the wait time and capacity utilization are sensitive to the consult waitlist size for each surgeon. For the *Kanban* scenario, the wait time and utilization are sensitive to the time horizon considered in calculating the waitlist size for each surgeon. Clinical network managers' experience with the "Blackjack" model suggests that the "minimum waitlist" number requires adjustment over time to meet demand and may result in difficulty in maintaining performance in the long run. The *Kanban* strategy is founded on the idea that previous practice patterns will predict future performance. This may be difficult to maintain when surgeons' practice patterns change or are highly variable.

Further work is required to identify a flexible, practical, robust model which maximizes the benefits of the *Next-available-surgeon* while maintaining the feasibility of the data collection, decision-making and monitoring efforts. Regardless of the distribution strategy chosen, we encourage all surgeons to engage in conversation about how they handle their consultation practice as a group, share information about their waiting times, and take steps to reduce

variation and increase efficiency to reduce waiting times and increase equitable access for their patients. We also encourage referring physicians to consider these results based on real-world data when submitting a surgical referral and to discuss the implications with their patients in choosing the next available surgeon versus requesting a preferred surgeon.

Considering the overall patient journey, a limitation of the study is that it models only a segment. The impact of front-end processes (e.g. referring physician decisions and activities, CI triaging procedures) and back-end processes (e.g. surgical capacity –operating room and surgeon availability) are significant factors affecting the overall patient wait time.

The main limitation of the study is its limited generalizability. The model was constructed using data and other information from actual referrals and consultations for patients with suspected or confirmed gallbladder disease. It is also somewhat artificial to study a single type of referral in isolation, as most surgeons provide care for patients with a range of conditions; we did not develop a model to account for referrals being received for a mixture of different conditions at the same time. The model does not include all the complexities of real-world practice. Many referrals may arrive simultaneously instead of discretely, coming from various sources to various destinations and in which surgeons' practices may change over time. However, studying a single type of referral allows this study to be a concise illustration of how distribution policies in a central intake program can directly impact system performance. It serves as a foundation to build upon and develop more complex, robust models in the future.

Funding This research was funded by Alberta Health Services through the Strategic Clinical Networks™. Deborah A. Marshall received salary and research support as a Canada Research Chair and as the Arthur J.E Child Chair Professor.

Data Availability The data used to parameterize the simulation study are available on request from the corresponding author, DM. The data are not publicly available due to restrictions.

References

1. Bichel A, Erfle S, Wiebe V, Axelrod D, Conly J. Improving patient access to medical services: preventing the patient from being lost in translation. *Healthc Q*. 2009;13:61–8.
2. Harrington DW, Wilson K, Rosenberg MW. Waiting for a Specialist Consultation for a New Condition in Ontario: Impacts on Patients' Lives. *Healthc Policy*. 2014;9:90–103.
3. McIntyre D, Chow CK. Waiting Time as an Indicator for Health Services Under Strain: A Narrative Review. *INQUIRY-J Health Car* 2020;57:0046958020910305.
4. Moir M, Barua B. The Private Cost of Public Queues for Medically Necessary Care 2020. Fraser Institute. 2020;10.
5. Alberta Health Services. Edmonton and Area - Facilitated Access to Surgical Treatment - General Surgery [Internet]. Alberta Health Services. 2021 [cited 22 Dec 2021]. Available from: <https://www.>

- albertahealthservices.ca/findhealth/Service.aspx?id=1080550&serviceAtFacilityID=1129243
6. Damani Z, Conner-Spady B, Nash T, Tom Stelfox H, Noseworthy TW, Marshall DA. What is the influence of single-entry models on access to elective surgical procedures? A systematic review. *BMJ Open*. 2017;7:e012225.
 7. Damani Z, Bohm E, Quan H, Noseworthy T, MacKean G, Loucks L, et al. Improving The Quality Of Care With A Single-Entry Model Of Referral For Total Joint Replacement: A Pre-implementation/Postimplementation Evaluation. *BMJ Open*. 2019;9:e028373.
 8. Greenwood-Lee J, Wild G, Marshall D. Improving accessibility through referral management: setting targets for specialist care. *Health Systems*. 2017;6:161–70.
 9. Günel MM, Pidd M. Discrete event simulation for performance modelling in health care: a review of the literature. *J Simul*. 2010;4:42–51.
 10. Katsaliaki K, Mustafee N. Applications of simulation within the healthcare context. *Journal of the Operational Research Society*. 2011;62:1431–51.
 11. Zhang X. Application of discrete event simulation in health care: a systematic review. *BMC Health Serv Res*. 2018;18:687.
 12. Rohleder TR, Lewkonja P, Bischak DP, Duffy P, Hendijani R. Using Simulation Modeling to Improve Patient Flow at an Outpatient Orthopedic Clinic. *Springer. Health Care Manag Sci*. 2011;14(2):135–45. ISSN: 1386-9620, 1572-9389. <https://doi.org/10.1007/s10729-010-9145-4>. Accessed 1 Oct 2021.
 13. Parks JK, Engblom P, Hamrock E, Satjapot S, Levin S. Designed to Fail: How Computer Simulation Can Detect Fundamental Flaws in Clinic Flow. *Journal of Healthcare Management*. 2011;56:135–46.
 14. Lu T, Wang S, Li J, Lucas P, Anderson M, Ross K. A Simulation Study to Improve Performance in the Preparation and Delivery of Antineoplastic Medications at a Community Hospital. *J Med Syst*. 2012;36:3083–9.
 15. Bair AE, Song WT, Chen Y-C, Morris BA. The Impact of Inpatient Boarding on ED Efficiency: A Discrete-Event Simulation Study. *J Med Syst*. 2010;34:919–29.
 16. Debacker M, Van Utterbeeck F, Ullrich C, Dhondt E, Hubloue I. SIMEDIS: a Discrete-Event Simulation Model for Testing Responses to Mass Casualty Incidents. *J Med Syst*. 2016;40:273.
 17. VanBerkel PT, Blake JT. A comprehensive simulation for wait time reduction and capacity planning applied in general surgery. *Health care manag sc*. Springer; 2007;10:373–85.
 18. Babashov V, Aivas I, Begen MA, Cao JQ, Rodrigues G, D'Souza D, et al. Reducing Patient Waiting Times for Radiation Therapy and Improving the Treatment Planning Process: a Discrete-event Simulation Model (Radiation Treatment Planning). *Clin Oncol*. 2017;29:385–91.
 19. Law AM, Kelton WD. *Simulation Modelling and Analysis*. 3rd Edition. McGraw-Hill, Boston ; Montreal : McGraw-Hill; 4th ed; 2000.
 20. Banks J, Carson JS, Nelson BL, Nicol DM. *Discrete-event system simulation*. Upper Saddle River, N.J.: Prentice Hall, Upper Saddle River, NJ: Prentice Hall; 3rd ed; 2001.
 21. Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Möller J. Modeling Using Discrete Event Simulation: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force–4. *Med Decis Making*. 2012;32:701–11.
 22. Jacobson SH, Hall SN, Swisher JR. *Discrete-Event Simulation of Health Care Systems*. In: Hall RW, editor. *Patient Flow: Reducing Delay in Healthcare Delivery* [Internet]. Boston, MA: Springer US; 2006 [cited 2021 Sep 27]. p. 211–52. Available from: https://doi.org/10.1007/978-0-387-33636-7_8
 23. Law AM. *Simulation Modeling and analysis*. Boston ; Montreal: McGraw-Hill, Boston ; Montreal : McGraw-Hill; 4th ed; 2007.
 24. *Systems Modelling Corporation*. *Arena*. Rockwell Automation; 2000. Rockwell Automation; 2000.
 25. Wolff RW. *Stochastic Modeling and the Theory of Queues*. 1st edition. Englewood Cliffs, N.J: Pearson; 1989.
 26. Kim JH, Ahn H-S, Righter R. Managing Queues With Heterogeneous Servers. *J Appl Probab*. 2011;48:435–52.
 27. Rosberg Z, Makowski AM. Optimal routing to parallel heterogeneous servers-small arrival rates. *IEEE T Automat Contr*. 1990;35:789–96.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.