APPLICATIONS OF
MODELLING AND SIMULATION
http://arqiipubl.com/ams eISSN 2600-8084 VOL 5, 2021, 145-155

An Overview of the Conceptual Simulation Modelling for Optimizing Funds Allocation in Health Care: Hip and Knee Replacement Case

Alexei Botchkarev

GS Research and Consulting, Toronto, Canada

Corresponding author: botchkarev.alexei.m@gmail.com

Submitted 05 May 2021, Revised 21 June 2021, Accepted 26 June 2021.

Abstract: Allocation of funds is one of the most important functions of the health care authorities. Performance of this function has implications not only for the macro-level health care system but also for individual patients whose wait times for medical procedures may be affected. In this paper, we consider conceptual simulation modelling of a multi-tier system of funds allocation. Allocation of funds is examined on a case of funding common surgical procedures: hip and knee replacements (HKRs). This study was motivated by the question: how to build a simulation model that will contribute to more effective allocation of HKRs funds? Effectiveness of funds allocation (optimization criteria) is viewed from the perspective of wait times reduction. The purpose of conceptual modelling was to provide a high-level description of the objects under consideration and understanding of the situations, systems, processes, their interactions, etc. Conceptual modelling revealed several key findings which need to be considered during implementation of simulation models that will allow better allocation of funds. From the methodology perspective, the study can be defined as a simulation-based healthcare resource allocation combinatorial optimization in a queueing-system environment.

Keywords: Fund allocation; Health economics; Hip and knee replacement; Optimization; Simulation and modelling.

1. INTRODUCTION

Osteoarthritis (OA) is the most common type of arthritis. It is also known as degenerative joint disease, degenerative arthritis, and wear-and-tear arthritis. OA presents a serious burden for the individuals, health care systems and societies as a whole [1]. Globally, out of 291 conditions, the hip and knee OA was estimated the 11th highest contributor to disability [2]. Hip and knee replacements (HKRs) have become common surgical procedures used in the treatment of severe osteoarthritis since 1970s. They are used when conservative interventions are not sufficient. There is a growth trend for the incidence of hip and knee replacement. The spread of the cases is due to the growth of the older segments of the population and increasing percentage of the patients younger than 65 years old attributed to obesity. A study presented at the 2018 Annual Meeting of the American Academy of Orthopaedic Surgeons (AAOS), predicted that by 2030 in the US, primary total hip replacement will grow by 171 percent and primary total knee replacement will grow by up to 189 percent, for a projected 635,000 and 1.28 million procedures, respectively [3].

Increased patients' demand may lead to longer waiting lists and extended wait times (WTs) for HKR surgeries. Many factors may contribute to WTs reductions, e.g.:

- Implementation of highly effective new HKR technologies and procedures.
- Better management of resources by enhanced scheduling of workforce, operating rooms, etc.
- Increasing capacity and resources (adding more hospitals, operating rooms, surgeons, nurses, anesthesiologists, etc.).
- Overcoming geographic variances of orthopedic surgeons' availability.
- Increasing total funding. Although this approach is known to offer only temporary improvements [4, 5].

Another factor, which did not receive due consideration in the academic literature is effective distribution of HKR funds. Exploring this factor motivated our study. We focused on building a simulation model that will contribute to more effective allocation of HKRs funds. For the purpose of this project, effectiveness of funds allocation (optimization criteria) is viewed from the perspective of WTs reduction. In this paper, we consider a multi-tier system of funds allocation in which health care authorities (HCAs) plans and distributes annual allotments to area/regional health care management units (AHCMUs) which plan, integrate and fund local health care providers, improving access and patient experience. Due to the complexity of the health care system and uncertainties of the investigated episode of care processes, the overall problem is analytically intractable, including queueing theory [5].

The purpose of this study was in developing a conceptual simulation modelling for optimizing annual HCA process of allocating HKR funds to AHCMUs, with a goal to enable wait time reductions. Conceptual modelling precedes development of an actual computer model and provides a high-level description of the objects under consideration and understanding of the situations, systems, processes, their interactions, etc. (e.g., [6, 7]). Conceptual simulation modelling guided research to the framework which facilitates imitating the health care processes based on the parameters of the health system (number of orthopedic surgeons, nurses, hospitals, operating rooms, etc.), input flows of patients, and a set of optimized allocated funds.

The paper is organized as follows. Next section provides the methodology of the study. Literature review outlines factors influencing HKR WTs and approaches to WTs minimization; modern simulation modelling methods are also reviewed. After that we focus on developing a conceptual model with description of the HKR health care processes, requirements to the simulation model and explain approaches to implementation of simulation optimization. The final sections present discussion of the results and concluding remarks.

2. METHOD

The purpose of conceptual modelling was to provide a high-level description of the objects under consideration and understanding of the situations, processes, their interactions, etc. that are essential for the systems perspective of the HCA allocation of funds for HKRs and effect WTs. Several methodologies were used to achieve the research objective of the conceptual modelling: identification of related peer-reviewed papers, critical literature review, critical thinking, and inductive reasoning.

The following issues were considered out of scope. We are not examining:

- allocation of funding and other resources outside of the HKR processes e.g., allocation of funds to health care as compared to other programs such as education, transportation or infrastructure.
- allocation of resources within health care, e.g., between hospitals and community care.
- distribution of funds between HKR and other interventions.
- Also, we are not attempting to minimize the total amount allocated for the HKR surgeries.

The purpose of the study can be presented in a more formal way mathematically:

$$\mathit{min}\ \mathrm{WT}(\mathrm{SAF}_n)$$

where:
$\mathrm{SAF}_n = \{u1_n, \ldots, uN_n\}$ is set of allocated funds - denotes feasible solution.
WT - wait time representing objective function.
N – number of AHCMU.
u1, …, uN – amounts of funds allocated to a correspondent AHCMU.

We have to acknowledge that wait time function depends on a number of arguments:

$$\mathrm{WT} = f\,(\mathrm{P}, \mathrm{OS}, \mathrm{OR}, \ldots, \mathrm{SAF})$$

where:
P – number of patients scheduled for HKR
OS – number of orthopedic surgeons
OR – number of operating rooms

However, due to the focus of our research, arguments (except SAF) of the WT function are set as constant parameters based on the current geographic realities, except the number of patients which is random.

3. LITERATURE REVIEW

Based on the research purpose and objectives, the review focused on the following knowledge areas:

- Approaches to defining WTs.
- Factors influencing HKR WTs and approaches to WTs minimization.
- Types of simulation modelling.
- Simulation optimization.

The search of relevant literature was conducted using Google Scholar. The following key words (and their combinations) were used: wait time knee/hip replacement; wait time factor health care; waiting time definition hip and knee replacement; wait time effect of funding; wait time hip replacement simulation; optimization of funds allocation; funds allocation simulation; queueing with resource limitations; optimization constraints, etc. Further search of related papers has been conducted through analysis of references in the papers found by Google. Several hundred papers were retrieved and reviewed with different extent of depth.

3.1 Approaches to Defining WTs

The total WT of a HKR patient starts when a primary care physician issues a referral to see a specialist (orthopedic surgeon) and ends when the HKR operations is done. The total WT period, which is important to the patient and contributes to the patient's satisfaction, is usually subdivided into several segments.

a) WT1 is the time between the physician's referral and the first specialist consult.
b) WT1A is the time between the first consult and decision to treat (make an operation).
c) WT2 is the time between decision to treat and surgery date or a time on the waiting list.

Although most common approach is to record distribution of the WTs of patients treated (after the surgery has been delivered and patient is taken off the list), it has been noted that WTs of the patients on the list at the time of reporting may be preferable [8] because the former approach reflects past efforts and the latter has a forward-looking potential. Commonly reported WTs measures include mean and median WT, wait times at different percentiles of the distribution, the number or proportion of patients waiting longer than the benchmark (e.g., [8, 9]).

It is worth noting that averages, although quite convenient for public reporting, may not be the best WT measures. Individual WTs are stochastic numbers with skewed probability distributions. Some authors suggest using measures that are more meaningful, e.g., the whole distribution of waiting times (e.g., [10]).

3.2 Factors Influencing HKR WTs and Approaches to WTs Minimization

Influencing factors can be broadly categorized in the following groups:

- Health system capacity-resources and their management.
- Demand for HKRs and its management.
- Health care authority policies and strategies.

3.2.1 Health System Capacity-Resources and Their Management

A critical limiting factor is the availability of orthopedic surgeons. The 2018 density of orthopaedic surgeons in the US is 9.25 per 100,000 population, with some states (e.g., Montana, Vermont) having over 14 surgeons, and some states (e.g., Mississippi, Texas) having only 7 surgeons per 100,000 population [11]. The second important factor is the availability of operating rooms (OR), beds for the recovering period and special equipment. Operating rooms are often referred to as a bottleneck resource of the hospital surgery operations [12]. Utilization of the operating rooms is also an important parameter. The third important factor is the availability of orthopedic surgeon's team members, e.g., nurses, anesthesiologists, etc.

All factors noted above pertain to the supply of resources. VanBerkel and Blake hold that WTs in general surgery are more dependent on availability of beds than on OR time [5]. Management of the resources is another factor influencing WTs. Poor management of resources can diminish existing capacity and increase WTs. Significant attention has been given to the resource planning and scheduling in hospitals, e.g., scheduling operating rooms, recovery beds, surgeons, nurses, appointments, improving appointment systems and processes (e.g., [12, 13]).

3.2.2 Demand for HKRs and its Management

Demand for HKRs is determined by the volume of population, its age-sex composition, health status, etc. All of these factors are underpinned by natural slowly changing over time processes and cannot be directly affected by the health system. Several measures have been suggested to manage, i.e., reduce demand. These include [9, 14]:

- Controlling the number of referrals to specialists. Primary care physicians would be expected to reduce the number of referrals to specialists by selecting the patients who would benefit the most from the operation.
- Diverting patients from public to private hospitals through private insurance [9].
- Allowing patients changing providers or having surgery abroad after certain WT.

Waiting list management is a common approach to manage demand. Surgeons prioritize waiting lists based on the patients' level of need [15, 16].

3.2.3 Health Care Authority (HCA) Policies and Strategies

HCAs are developing and implementing policies and strategies to influence supply and/or demand for HKRs and hence they can directly affect wait times. The following factors can be attributed to HCA involvement (especially, in the publicly funded health care systems):

- Long-term sustained investments in health system capacity, e.g., increasing number of surgeons' positions, hospitals, etc.
- Funding policies, e.g., paying directly for specific patient-based treatment activities – HKRs.
- Setting performance targets (e.g., WTs durations), monitoring and publicly reporting results.

All of the above approaches can provide positive effect. Although it should be noted that using any specific method does not guarantee successful outcome. There are always caveats (mostly based on ad hoc contexts). For example, investments should be long-term and sustainable, because a single episode of funds transfer is known to provided only temporary WT reduction [9]; likewise, setting performance targets with lack of target enforcement (e.g., penalties for non-compliance) may reduce potential effect of this approach [9]. Similarly, it was shown that the increase of bed capacity will reduce waiting lists, but this is true only in case that demand will not grow [17]. As a result, despite multiple HCA initiatives to tackling the WTs in many countries around the world, the evidence of success varies [8].

In the real world, factors mentioned above are not static, they exhibit dynamic interactive behaviour with temporary variations of capacity and demand, e.g., changing productivity of surgeons, staff not arriving on time, operating room is not ready, service time variability, variability of patient demand, late arrivals of patients or no-shows, etc. (e.g., [4, 18]. WTs may be negatively affected, if health system fails to rapidly adjust to these fluctuations (e.g., [19]).

3.3 Types of Simulation Modelling

Due to the complexity of the systems and processes involved, it was clear from the outset that neither assessments of the WTs nor optimization of funding can be analytically tractable. Simulation was a logical venue. Comparing different types of simulation techniques and defining determinants of their use for specific problems in the health care domain – is an ongoing area of research (e.g., [20-27]). Most often discrete-event simulation (DES), system dynamics (SD), agent-based simulation (ABS) and Monte Carlo simulation are considered.

DES modelling has been widely and successfully used to tackle operations research problems in health care. The objectives of the DES models focus on a variety of topics including patient flow analyses; scheduling, setting appointments, managing admissions and waiting list prioritization strategies; planning of healthcare services; capacity planning and asset allocation, e.g., number of surgeons, nurses, anesthesiologists, operating rooms, etc.; cost-effectiveness evaluations in health economics, etc. (e.g., [24, 25, 28-31]). An important challenge in DES (also applies to other simulation techniques) is in selecting an appropriate level of the model simplification: "appropriate simplification can be a surprisingly complex process…" [24]. Increased level of detail may allow for better representation of the real-life system or process but leads to higher complexity. The consequence of the stated notion is that the health care models rarely tackle hospital as a whole. Researchers commonly limit their models to specific processes (e.g., scheduling, forecasting patient arrival, etc.) or departments (e.g., orthopedic, emergency, etc.) and assume any outside connections as given [24].

Although some diseases, osteoarthritis among them, take patients through several health care sectors (primary care, hospitals, community care, home care), models that involve complete continuum of care (cross-sectorial) are hard to find (e.g., [32]. Another rare example is a DES simulation of WTs in a system of outpatient clinics [33]. DES has been often used with an aim to reduce waiting times mostly in an outpatient environment (e.g., [33, 34]). DES is noted to be useful for models with an individual patient focus and resource constraints (e.g., [24, 25]). DES is recommended specifically for the problems with queueing structure and examining waiting time-related performance [23]. Katsaliaki and Mustafee [25] note that Monte Carlo simulation is the most popular technique in healthcare simulation research based on the number of published papers. Commonly this research is focused on health risk assessment and cost-benefit studies.

SD is mostly deterministic method and is appropriate for tackling healthcare problems at a macro-level for facilitating health policy development [35, 36]. It is less suitable for a detailed modelling because handling stochastic variations is not a strong feature of this method [23, 25]. Specifically, SD is used to model epidemics and disease prevention [26]. On the other hand, ABS is a relatively new approach and is not as popular as previously mentioned methods. It has individual focus (agents) and mostly uses deterministic approach [37, 38]. Queues in ABS exist implicitly [23].

Some real-world situations can benefit from combined use of simulation methods [39]. For example, DES can be enhanced with ABS for improved modelling of human behaviour. Integration of DES and ABS was used to optimize scheduling for outpatient orthopedic consultations [40]. A hybrid system dynamics and DES simulation was used to investigate boarding patients in acute hospitals [41].

3.4 Simulation Optimization

Simulation optimization includes several groups of techniques which can be used to find input parameters of a stochastic simulation that lead to optimization (most commonly, minimization) of the objective function not necessarily having explicit algebraic form (e.g., [42-44]). Simulation optimization also referred to as simulation-based optimization, parametric optimization, black-box optimization, optimization via simulation, hybrid simulation-optimization, etc. (e.g., [44]). Current state of understanding of the simulation optimization problems and types of solutions, approaches and algorithms can be found in reviews (e.g., [42, 44, 45]). Most relevant for our research are methods that deal with discrete decisions and single criteria.

Simulation optimization is known to be successfully applied to discrete event simulations, particularly, that model queues (e.g., [44]). Banditori et al used simulation optimization to address the master surgical scheduling problem [46]. Malik et al applied multi-objective combinatorial optimization based on a genetic algorithm to balance waiting time and health care costs [19]. Literature review informed decisions made during development and validation of the simulation model.

4. CONCEPTUAL MODELLING

The purpose of conceptual modelling is to provide a high-level description of the objects under consideration and understanding of the situations, systems, processes, their interactions, etc. (e.g., [6, 7]). At this stage, examination is independent of the simulation method (i.e., DES, SD, ABS, etc.). The results of the conceptual modelling will guide the selection of the simulation method. In this study, two aspects of the real-world situation need to be examined: processes of the HKR episode of care and financial processes of funds allocation. Although interconnected, both types of the processes exist within their own paradigms. The two aspects defined in the conceptual model will be integrated in the simulation model.

4.1 Health Care Processes

Figure 1 shows the scope of the HKR episode of care (e.g., [32]). The chart shows the complexity and cross-sectorial nature of the episode of care: i.e., primary care, acute care, community care, home care are involved [47]. Most detailed model should reflect these complexities.

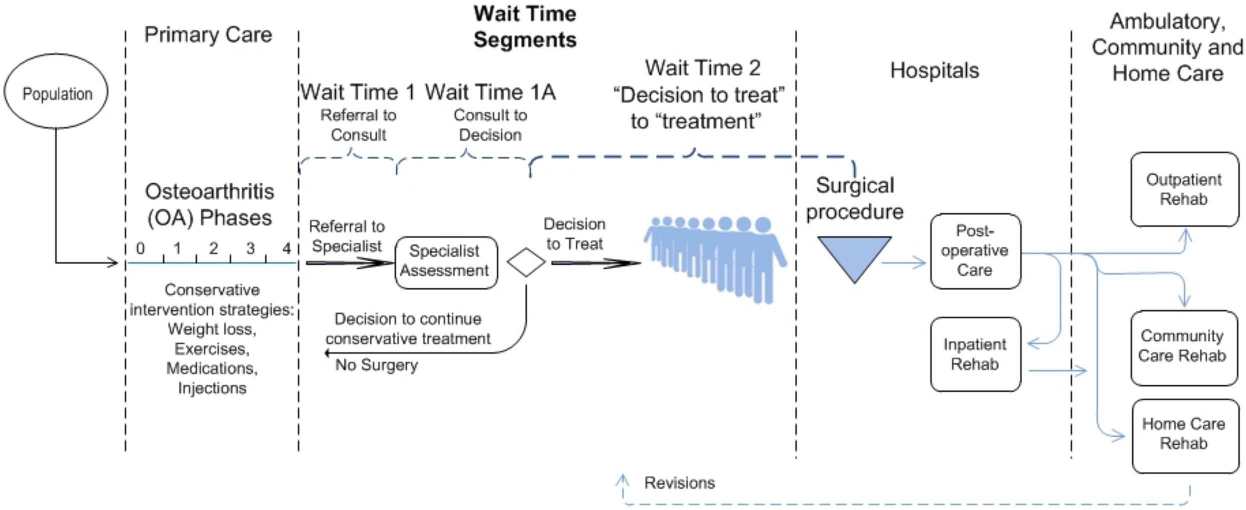

Figure 1. Scope of the hip and knee replacement episode of care

Patients enter the healthcare system through the primary care. Conservative intervention strategies are delivered by family physicians. At a certain stage of the disease, the patient is referred to a specialist (orthopedic surgeon) who assesses the patient and makes a decision to treat (conduct HKR surgery) or continue conservative treatment.

Based on the purpose of this project we will be interested in the WT the patient experiences. Figure 1 shows the overall WT has several segments (referral to consult, consult to decision, decision to treat to treatment – operation). Our focus is on the third segment: from the day the surgeon makes a decision to make an HKR operation and the patient agrees to undergo the procedure to the day the surgical procedure is delivered. This segment is shown as WT2 on the chart and signifies the period a patient stays in a queue or on a waiting list.

On the day of the operation, a patient is admitted to the hospital. Actual HKR surgery lasts 1.5-2.5 hours. Total time a patient stays in the operating room (including preparation and recovery) is 2 – 4 hours. After a surgery, a patient stays in acute care for 4-5 days. We assume that availability of hospital beds has been considered within the hospital resource planning and scheduling phase – so there are no delays for patients. Also, the time period after operation is not included in WTs and do not need to be included in the model. Exclusion of the inpatient time from the model was also prompted by the current trend of the length of stay reduction. For example, Cox et al [48] reported results of the program with the goal to reduce the overall length of stay (LOS) after primary hip and knee arthroplasty to less than 3 days with the overall patient satisfaction. After discharge from acute care, rehabilitation takes 30 to 90 days in ambulatory, community or home care. Certain percentage of the patients are readmitted for revision surgeries due to procedure-related complications. Mahajan et al show that the average 30-day and 90-day readmission rates were 5.33% and 7.12%, respectively [49]. HKR revisions present additional workload for the health system and negatively affect WTs.

The HKR episode of care described above illustrated the sequence of the processes from the patient's perspective. From the system's perspective, it is important to note that there is no centralized intake system or a single list of patients waiting for HKRs. Most HKR waiting lists (queues) are formed, managed and served by individual orthopedic surgeons [4, 34]. This observation leads to major implication for the structure of the overall modelling framework. It should include two levels: lower level of individual queue components and higher (HCA) level integrated queueing system which is comprised of the individual queue components.

Figure 2 shows queueing system component - waiting list of an individual orthopedic surgeon. WTs are realized in the patient-level queueing system components with hospital-level resource-sharing constraints. The study of WTs should focus on this level. All other WTs calculations - at hospital, AHCMUs or HCA levels - are straightforward aggregations of the individual patient WT. Suitability of the patient-level simulation models was confirmed in academic literature (e.g. [50, 51]).

It should be noted that system components follow priority queue behaviour and not first-in, first-out (FIFO) discipline. When a surgeon adds new patients to the list, they are placed in the queue based on the severity of symptoms (not necessarily to the end of the queue). Also, if patients' condition changes, while they are waiting, their position on the list may be changed by the surgeon, i.e., the waiting list can be undergoing a continuous modification. These types of lines are commonly referred to as managed queues. Certain features of the queueing system components' (customers/patients and server/surgeon) behaviour should be noted. Firstly, each patient is listed on only one waiting list. Secondly, patients do not "change" the lines (surgeons) – not to lose the priority. Thirdly, quitting the line by a patient is rather rare event. Although, (as large percentage of patients represent older groups and waiting period may extend for more than six months) leaving the line is possible due to changing patient's condition (different priorities arise) or death.

Figure 3 shows integrated queueing system. Individual queueing system components of the same hospital (or facility, if a hospital has more than one facility) are interdependent as they share resources (operating rooms, operating teams, etc.). The number of the queueing system components in the integrated system corresponds to the number of practicing orthopedic surgeons. This number will be validated using administrative databases registering actual surgery activities. The number of surgeons does not directly correlate with the overall volume of the HKRs. Independency of surgeons translates into their productivity variations due to professional preferences, personal choices on the amount of work, etc. The volume of HKRs has large variability across practicing surgeons.

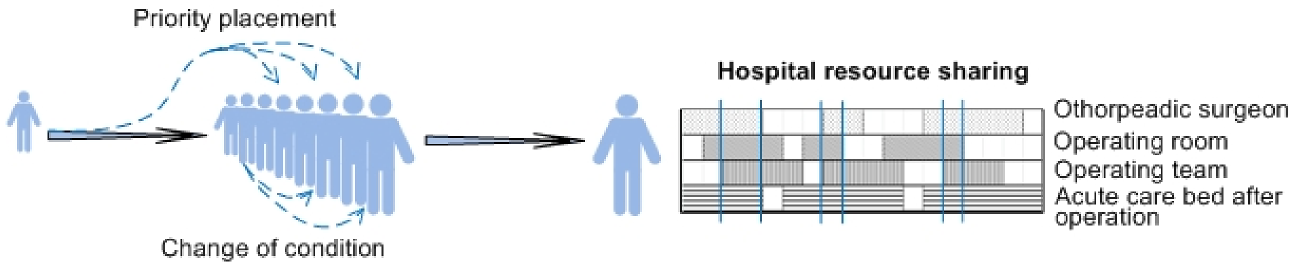

Figure 2. Queueing system component behaviour

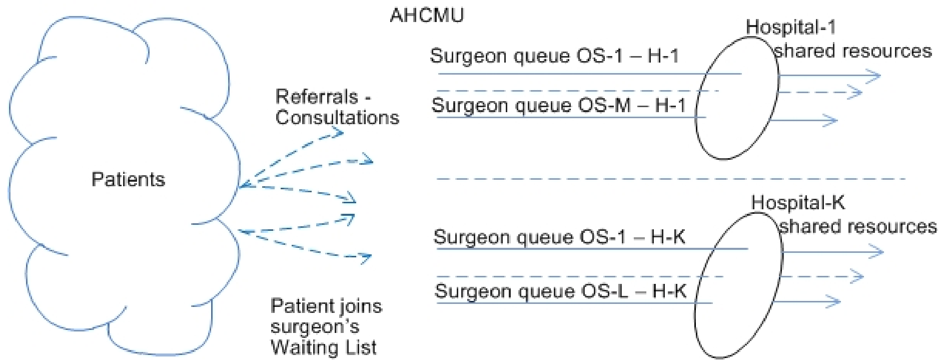

Figure 3. Integrated queueing system

4.2 Simulation Model Requirements

Simulation model development should be taking a pragmatic approach within a real-world framework of the health care system and HCA activities of funds allocation. This approach implies the following:

a) The model should be based on the current resources (i.e., the workforce of orthopaedic surgeons and nurses, number of operating rooms, etc.) which numerical parameters will be used as model inputs.
b) The total amount of funds allocated for the HKR operations within the available healthcare budget is known and used as a model input.
c) The model is aligned with the HCA allocation of HKR funds to the existing AHCMUs in an annual process.
d) The model should be based on the assumption that allocation of funds to each AHCMU conducted by the current method is "reasonably good" and took into account important factors such as ethical and equity considerations, AHCMUs' capacities, historical performance, etc. This allocation should be used as an initial condition for the simulation.
e) The model should be rather simple and practical with minimum number of parameters/variables to allow HCA employees to use it on a regular basis.
f) The model should be transparent and reproducible to avoid wrong perceptions.

The model should be able to simulate:

- Patient arrival process (at a patient-level).
- Distribution of patients by queueing components (individual surgeon waiting lines).
- Flow of patients through the waiting time period (using one day as a measuring unit).
- Service delivery process with hospital constraints on shared resources.
- Influence of funding constraints.

The model should be able to measure WT statistics of the individual patients progressing through the queueing system component. HCA-level WT should be calculated as output with 90th percentile as a performance measure. The model should not attempt mimicking hospital detailed models used to:

- workforce scheduling
- patient and appointment scheduling
- model length of inpatient stay
- operating room scheduling, etc.

These models are very complex, require multiple inputs and performed by Master Surgical Scheduling software in hospitals (e.g., [52, 53]). Commonly used elements of queueing system should be defined:

- Patients – unlimited (open system)
- Arrival pattern – distribution of intervals between arrivals based on historical data.
- Maximum queue size (system capacity) – unlimited.
- Queueing discipline - priority queue. May be modelled as set of queues.

- Queueing system component – single channel.
- Integrated queueing system - multi channel.
- Number of servers – orthopedic surgeons.

Considering the nature of the processes revealed in the conceptual modelling and literature review, DES appears to be the most appropriate simulation method. Computational efficiency of the model is not of prime importance – allocation of funds is an annual procedure.

4.3 Optimization

In the previous section, we set approaches to build a simulation model which will allow us to imitate the health care processes based on the parameters of the health system (number of orthopedic surgeons, nurses, hospitals, operating rooms, etc.), input flows of patients, and a set of allocated funds for AHCMUs. We can run this simulation model and get WTs as outputs. WTs represent the objective function of the model.

The next step is to integrate the simulation component with an optimization part of the model. Figure 4 shows the logic of the simulation optimization. Optimization module generates SAF feasible solutions which are used as inputs to the simulation model, stores WT for each SAF, and determines the optimal SAF – the one that minimizes the HCA HKR WTs. Several considerations should be taken into account when selecting an appropriate approach to the optimization (e.g., [44]). The class of the problem we tackle can be defined as a discrete solution space, single output, single criteria (i.e., WT), with an objective function which cannot be explicitly stated mathematically.

To set an approach to optimization, three important actions must be performed:

- define solution space.
- select feasible solution generation technique.
- select optimal solution search method.

By defining solution space, we describe potential solutions to the problem. In our case, solution space includes possible SAFs. Funds may be distributed to AHCMUs in multiple ways. For example, all AHCMUs may get equal number of operations, or, hypothetically, one AHCMU may be assigned all hip operations and all other AHCMUs will get nothing, etc. Formally, number of SAFs equals to the number of combinations $C(n,k) = n!/(k!\,(n-k)!)$, where n is the number of hips operations and k is the number of AHCMUs. This problem is known in combinatorics as a problem of calculating the number of different ways to distribute n indistinguishable balls into k distinguishable boxes. The resulting number is usually huge. Obviously, this approach (in an environment with several thousand operations and a dozen of AHCMUs) is prohibitive by both the computational effort and the requirements to the memory. This example has been included as a convenient demonstration of the combinatorial nature of the problem and reason to use combinatorial optimization techniques later in the paper.

The above example is unrealistic not only number-wise, but also because it contradicts the pragmatic approach accepted in this research. First, we are not advocating a re-build of the health care system by firing orthopedic surgeons and/or relocating them (as some of the solutions would imply). Second, we need to reiterate one of our foundational assumptions which states that allocation of funds to each AHCMU conducted by the current method, i.e., initial SAF, is "reasonably good" and considered important population-based factors. Relying on this assumption, we can limit the solution space of the problem to the solutions that do not involve drastic redistribution of funds between AHCMUs, or in other words - to the initial SAF's "neighbourhood". Limiting the solution space to the initial SAF neighbourhood should make computational efforts feasible. The extent of the neighbourhood (size of the solution space) will be determined empirically. Similar approaches were used in combinatorial optimization to tackle computationally hard problems (e.g., [54-56]). To generate feasible solutions, we represent the solution space as a matrix where a central column equals initial SAF.

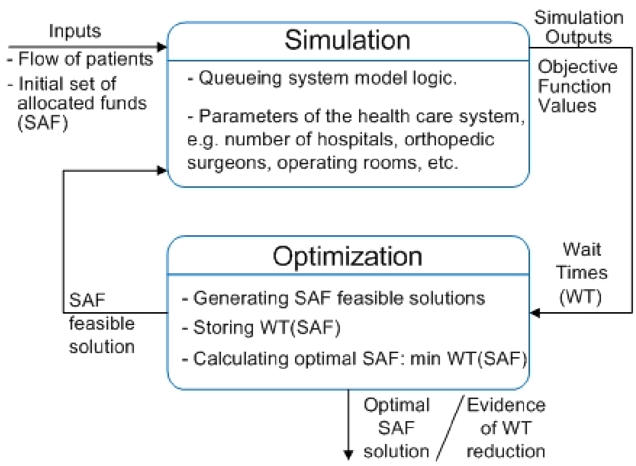

Figure 4. Simulation optimization logic

A sample matrix is shown as:

u1 - 2Δ	u1 - Δ	**u1**	u1 + Δ	u1 + 2Δ
u2 - 2Δ	u2 - Δ	**u2**	u2 + Δ	u2 + 2Δ
----------	----------	**----------**	----------	----------
uN4 - 2Δ	uN4 - Δ	**uN4**	uN4 + Δ	uN4 + 2Δ

where Δ - unit of deviation from the initial SAF. On the first step, the algorithm must generate all possible combinations from the elements of this matrix. Each combination must contain *N* elements – one from each row. Then, a check of the sum of elements of each combination is conducted. Only combinations with a sum of elements equal to the sum of the initial SAF will be added to a list of feasible solutions. Ranking and selection method is used to compare values of objective function and make a decision on the optimal solution (e.g., [45, 57]).

5. DISCUSSION

Conceptual modelling revealed several key findings which need to be considered during implementation of simulation modelling. Orthopedic surgeons are considered independent contractors with privileges at certain hospitals. Independency of surgeons translates into their productivity variations due to professional preferences, personal choices on the amount of work, etc. Individual surgeons are in full control of the HKR waiting lists (queues). That means that orthopedic surgeons make decisions on forming their wait lists (admitting patients to the queues), managing queues based on their assessments of the urgency of patients' needs, prioritization and scheduling patients for operations (taking into account hospital constraints), and delivering HKRs surgeries.

WTs are realized in the patient-level queueing system components (waiting lists of individual orthopedic surgeons) with hospital-level resource-sharing constraints. The study of WTs should focus on this level. The model should be able to measure WT statistics of the individual patients progressing through the queueing system component. The overall modelling framework should be able to consolidate lower-level individual queue components into higher (HCA) level integrated queueing system which is comprised of the individual queue components.

Optimization should start from the initial SAF, which is based on the HCA forecast of HKRs volumes by AHCMU and considered "reasonably good" in reflecting important population-based needs and HCA's commitment to providing equitable health care access. The search for an optimized SAF should be limited to the solution space in the initial SAF's neighbourhood and should not involve drastic redistribution of funds between AHCMUs.

Conceptual approaches, proposed in this study (e.g., modelling individual surgeons' queues and patient-level processes), differ from methods used in the literature. For example, Cipriano et al [34] used simulation modelling to investigate HKRs WTs and explored a system of 25 aggregated queues with monthly averages inflow and outflow of patients. We believe that using more detailed model, suggested in this study, will provide better representation of the variability of patient flows. As simulation modelling is an iterative process by nature, conceptual model should be revised on later phases of the project.

From the health economics perspective, the source of the queues is in the imbalance between demand and supply [5, 9]. The demand is characterized by the number of patients arriving to the queue. The supply or capacity is determined by the number of surgeons and their ability to serve patients per unit of time (not to mention other resources). The imbalance could be either long-term (or permanent), and in this case unmet demand will result in an ongoing build-up of the queue demonstrating lack of capacity; or it could be attributed to relatively short-term random variations of patients' arrivals and surgeons' productivity. The influence of both factors may be observed simultaneously. In this context, minimization of WTs does not imply the intention of zeroing WTs (e.g., [58]). It is more correct to discuss multi-criteria optimization with an objective of finding the best available solution simultaneously defined in at least three domains: clinically proven WT durations with the best patient outcomes (current WT target – 182 days – can be used), acceptable ranges of hospital resource utilizations, and satisfactory equity considerations among patients across local areas.

Certain modelling challenges should be noted. Establishing causality (i.e., the cause-effect relationship) between the HCA's allocation of funds and availability of funds in the service channel (individual orthopedic surgeon where WT is realized) is an ambiguous problem. As it was demonstrated in the conceptual model, after the HCA allocates funds to AHCMUs, there are two more rounds of allocations: AHCMUs to hospitals and hospitals/Surgical Departments to orthopedic surgeons. HCA has no power to affect these allocations. At best, the HCA may be aware about the AHCMUs to hospitals allocations. Allocations within the hospitals are less transparent for the HCA and may be driven by a different set of criteria applied by a private health provider.

Interpretations of the impact of the HCA allocation of funds may be subjective. If WTs decrease – is it because the HCA did a good job or because AHCMUs/hospitals improved initial allocation in their redistributions? If the WTs increase – is it due to the HCA or because AHCMUs/hospitals "distorted" initial allocation? Optimization would have impact only if there is a direct coordination in the HCA-AHCMU-hospital-OS framework. For the purposes of the study, we have to make an assumption that coordination has been established between HCA, AHCMUs, hospitals with an intention of a unified allocation of funds to the level of individual OS, i.e., using the same criteria and a tool. It should be also noted that the effort will be based on establishing association, not causality between the HCA's allocation of funds and availability of funds in the service channel.

Optimization caveat should be noted. In most studies, optimization exercise directly leads to the selection of the "best" possible solution (i.e., the SAF that minimizes WT). In our case, there could be situations with not so straightforward decision. For example, there are several SAFs (with quite different distributions of funds between AHCMUs) which lead to the same WTs (within the margin of error). Or there are several SAFs allowing to reach our target WT $\leq$ 182 days. In these situations, we have to select not "just the best" SAF, but the one which is closer to the initial SAF (i.e., demonstrates less redistribution of funds from the initial SAF). This approach emphasizes the importance of the criteria that were used in generating HKR forecasts. To some extent, this may be understood as a multi-criteria optimization. These considerations may be applied when using the model.

5.1 Limitations

This study is positioned as a first phase of ongoing research into optimization of funding. We are not modelling a complete continuum of care although it has been mentioned that the episode of HKR crosses all sectors of heath care (from primary to home care). Method that is used in this study is a single parameter optimization – WTs. The result might not represent the optimum, if several other parameters/factors are considered. Review of the results is needed.

Although the suggested model can be used (by multiple runs) to estimate what total amount of funds is needed to achieve certain level of WTs, (e.g., 182 days for 90 percentiles, 182 days for 100 percentiles, etc.), this is not an intended purpose of the model. The authors discourage to use the model for this purpose, as there are a number of other factors, e.g., types of resources, which would need to be accounted for in the calculations. It is known that by just increasing the funds, it is impossible to solve the WTs problem in the long run. Maybe other factors in the overarching framework could bring better results than pure funding increase. Multi-criteria decision model needs to be developed for this task.

5.2 Future Work

The second part of the project will involve building simulation model, validation, verification and getting quantitative results. Simulation model, when developed, will contribute to better understanding of the health system processes and may have a positive impact on policy decision making. It can be utilized by the HCA contributing to HKR WTs reductions. Patients' outcomes with HKR may be improved due to reduced WTs. Developed simulation model (with certain modifications) may be applied to a broader spectrum of interventions, i.e., to various elective procedures.

This project is a first step in a multiphase effort. Further elaboration will be performed in follow-up phases. We are planning to move from a single criteria optimization (reducing WTs) to performing multi-criteria optimization (i.e., simultaneous optimization based on a variety of health system resources: doctors, nurses, hospitals, and their geographic distribution) within a cross-sector health care model.

6. CONCLUSION

Effective investment of public funds is of prime concern for all HCAs. As a first step in exploring optimized allocations of funds for HKRs in using simulation modelling, this study developed a conceptual model and high-level requirements to the discrete event simulation tool. Implementation of the model will allow HCA to optimize allocation of funds to AHCMUs so that HKRs WTs can be minimized.

ACKNOWLEDGMENT

The views, opinions and conclusions expressed in this document are those of the author and do not necessarily represent the views of his current or former employers.

REFERENCES

[1] J. Callhoff et al., Disease burden of patients with osteoarthritis: Results of a cross-sectional survey linked to claims data, *Arthritis Care & Research*, 72(2), 2020, 193-200.

[2] A. Mobasheri, S. Saarakkala, M. Finnilä, M. A. Karsdal, A.-C. Bay-Jensen, and W. E. van Spil, Recent advances in understanding the phenotypes of osteoarthritis, *F1000Research*, 8, 2019.

[3] Projected volume of primary and revision total joint replacement in the U.S. 2030 to 2060. The American Academy of Orthopaedic Surgeons, 2018. https://aaos.new-media-release.com/2018_annual_meeting_pf/clinicals/Sloan_TJR.pdf

[4] C. Sanmartin, S. ED Shortt, M. L. Barer, S. Sheps, S. Lewis and P. W. McDonald, Waiting for medical services in Canada: lots of heat, but little light, *Canadian Medical Association Journal,* 162(9), 2000, 1305-1310.

[5] P. T. VanBerkel and J. T. Blake, A comprehensive simulation for wait time reduction and capacity planning applied in general surgery. *Health Care Management Science*, 10(4), 2007, 373-385.

[6] S. Robinson, Conceptual modelling for simulation: Progress and grand challenges, *Journal of Simulation*, 14(1), 2020, 1-20.

[7] G. T. Gabriel, A. T. Campos, F. Leal and J. A. Barra Montevechi, Good practices and deficiencies in conceptual modelling: A systematic literature review, *Journal of Simulation*, 2020, 1-17.

[8] L. Siciliani, V. Moran and M. Borowitz, Measuring and comparing health care waiting times in OECD countries, *Health Policy*, 118(3), 2014, 292-303.

[9] L. Siciliani, M. Borowitz, V. Moran, Waiting time policies in the health sector: What works? *OECD Health Policy Studies*, 2013, OECD Publishing.

[10] S. Dimakou, O. Dimakou and H. S. Basso, Waiting time distribution in public health care: Empirics and theory, *Health Economics Review*, 5(1), 2015, 25.

[11] Orthopaedic Practice in the U.S. 2018. AAOS Department of Clinical Quality and Value. January 2019. http://www.aaos.org/globalassets/quality-and-practice-resources/census/2018-census.pdf

[12] J. Schoenfelder, S. Kohl, M. Glaser, S. McRae, J. O. Brunner and T. Koperna, Simulation-based evaluation of operating room management policies, *BMC Health Services Research*, 21(1), 2021, 1-13.

[13] T. Zhu, L. Luo, W. Shen, X. Xu and R. Kou, Admission scheduling of inpatients by considering two inter-related resources: Beds and operating rooms, *Optimization*, 2020, DOI: 10.1080/02331934.2020.1829619.

[14] W. Chen, Z. G. Zhang and X. Chen, On two-tier healthcare system under capacity constraint, *International Journal of Production Research*, 58(12), 2020, 3744-3764.

[15] O. Mariana, V. Bélanger, I. Marques and A. Ruiz, Assessing the impact of patient prioritization on operating room schedules, *Operations Research for Health Care*, 24, 2020, 100232.

[16] M. Breton, et al., How the design and implementation of centralized waiting lists influence their use and effect on access to healthcare - A realist review, *Health Policy,* 124(8), 787-795.

[17] M. Antelo, F. R. Santias and A. M. Calvo, Bed capacity and surgical waiting lists: A simulation analysis, *European Journal of Government and Economics*, 4(2), 2015, 118-133.

[18] C. Van Riet and E. Demeulemeester, Trade-offs in operating room planning for electives and emergencies: A review, *Operations Research for Health Care*, 7, 2015, 52-69.

[19] M. M. Malik, M. Khan and S. Abdallah, Aggregate capacity planning for elective surgeries: A bi-objective optimization approach to balance patients waiting with healthcare costs, *Operations Research for Health Care*, 7, 2015, 3-13.

[20] S. N. Roy, B. J. Shah and H. Gajjar, Application of simulation in healthcare service operations: A review and research agenda, *ACM Transactions on Modeling and Computer Simulation,* 31(1), 2020, 1-23.

[21] S. Roy, S. P. Venkatesan and M. Goh, Healthcare services: A systematic review of patient-centric logistics issues using simulation, *Journal of the Operational Research Society,* 2020, 1-23.

[22] K. M. Long, F. McDermott and G. N. Meadows, Factors affecting the implementation of simulation modelling in healthcare: A longitudinal case study evaluation, *Journal of the Operational Research Society,* 71(12), 2020, 1927-1939.

[23] M. M. Gunal, A guide for building hospital simulation models, *Health Systems*, 1(1), 2012, 17-25.

[24] M. M. Gunal and M. Pidd, Discrete event simulation for performance modelling in health care: A review of the literature, *Journal of Simulation*, 4(1), 2010, 42-51.

[25] K. Katsaliaki and N. Mustafee, Applications of simulation within the healthcare context, *Journal of the Operational Research Society*, 62(8), 2011, 1431-1451.

[26] B. Mielczarek, Review of modelling approaches for healthcare simulation, *Operations Research and Decisions*, 26(1), 2016, 55-72.

[27] B. Mielczarek and J. Zabawa, Modeling healthcare demand using a hybrid simulation approach, In *Proceedings of the 2016 Winter Simulation Conference,* Virginia, USA, 2016, 1535-1546.

[28] M. Shoaib and V. Ramamohan, Simulation modelling and analysis of primary health centre operations, *arXiv preprint arXiv:2104.12492* (2021).

[29] C. Cubukcuoglu, P. Nourian, I. S. Sariyildiz and M. F. Tasgetiren, A discrete event simulation procedure for validating programs of requirements: The case of hospital space planning, *SoftwareX*, 12, 2020, 100539.

[30] U. Yakutcan, E. Demir, J. R. Hurst and P. C. Taylor, Patient pathway modelling using discrete event simulation to improve the management of COPD, *Journal of the Operational Research Society,* 2020, 1-25.

[31] S. Singla, Demand and capacity modelling in healthcare using discrete event simulation, *Open Journal of Modelling and Simulation,* 8(4), 2020, 88-107.

[32] S. A. Vanderby, M. W. Carter, T. Noseworthy and D. A. Marshall, Modelling the complete continuum of care using system dynamics: The case of osteoarthritis in Alberta, *Journal of Simulation*, 9(2), 2015, 156–169.

[33] M. Q. Corpuz, C. F. Rusnock, V. V. Valencia and K. F. Oyama, Reducing wait-time of a system of clinics using discrete-event simulation, In *Proceedings of IIE Annual Conference*, Nashville, USA, 2015, 969.

[34] L. E. Cipriano, B. M. Chesworth, C. K. Anderson and G. S. Zaric, Predicting joint replacement waiting times, *Health Care Management Science*, 10(2), 2007, 195-215.

[35] M. R. Davahli, W. Karwowski and R. Taiar, A system dynamics simulation applied to healthcare: A systematic review, *International Journal of Environmental Research and Public Health,* 17(6), 2020, 5741.

[36] N. J. Zhu, R. Ahmad, A. Holmes, J. V. Robotham, R. Lebcir and R. Atun, System dynamics modelling to formulate policy interventions to optimise antibiotic prescribing in hospitals, *Journal of the Operational Research Society,* 2020, 1-13.

[37] A. A. Alzu'bi, S. Ibrahim A. Alasal and V. J. M. Watzlaf, A simulation study of coronavirus as an epidemic disease using agent-based modeling, *Perspectives in Health Information Management* 18, 2021.

[38] M. Comis, Martin, C. Cleophas and C. Büsing, Patients, primary care, and policy: Agent-based simulation modeling for health care decision support, *Health Care Management Science,* 2021, 1-28.

[39] dos Santos, V. Horsti, K. Kotiadis and M. P. Scaparra, A review of hybrid simulation in healthcare, In *2020 Winter Simulation Conference (WSC)*, 2020, 1004-1015.

[40] C. Kittipittayakorn and K. -C. Ying, Using the integration of discrete event and agent-based simulation to enhance outpatient service quality in an orthopedic department, *Journal of Healthcare Engineering*, 2016, 4189206.

[41] L. Keshtkar, W. Rashwan, W. Abo-Hamad and A. Arisha, A hybrid system dynamics, discrete event simulation and data envelopment analysis to investigate boarding patients in acute hospitals, *Operations Research for Health Care*, 26, 2020, 100266.

[42] M. Yousefi, M. Yousefi and F. S. Fogliatto, Simulation-based optimization methods applied in hospital emergency departments: A systematic review, *Simulation*, 96(10), 2020, 791-806.

[43] B. R. P. e Oliveira, J. A. de Vasconcelos, J. F. F. Almeida and L. R. Pinto, A simulation-optimisation approach for hospital beds allocation, *International Journal of Medical Informatics*, 141, 2020, 104174.

[44] S. Amaran, N. V. Sahinidis, B. Sharda and S. J. Bury, Simulation optimization: A review of algorithms and applications, *Annals of Operations Research*, 240, 2016, 351-380.

[45] J. Xu, E. Huang, C. H. Chen and L. H. Lee, Simulation optimization: A review and exploration in the new era of cloud computing and big data, *Asia-Pacific Journal of Operational Research*, 32(3), 2015, 1550019.

[46] C. Banditori, P. Cappanera and F. Visintin, A combined optimization–simulation approach to the master surgical scheduling problem, *IMA Journal of Management Mathematics*, 24(2), 2013, 155-187.

[47] K. D. Allen et al., Osteoarthritis: Models for appropriate care across the disease continuum, *Best Practice & Research Clinical Rheumatology*, 30(3), 2016, 503-535.

[48] J. Cox, C. Cormack, M. Prendergast, H. Celestino, S. Willis and M. Witteveen, Patient and provider experience with a new model of care for primary hip and knee arthroplasties, *International Journal of Orthopaedic and Trauma Nursing*, 20, 2016, 3-27.

[49] S. M. Mahajan, C. Nguyen, J. Bui, E. Kunde, B. T. Abbott and A. S. Mahajan, Risk factors for readmission after knee arthroplasty based on predictive models: A systematic review, *Arthroplasty Today*, 6(3), 2020, 390-404.

[50] L. Pradelli, M. Pinciroli, H. Houshmand, B. Grassi, F. Bonelli, M. Calleri and M. Ruscio, Comparative cost and effectiveness of a new algorithm for early lyme disease diagnosis: Evaluation in US, Germany, and Italy, *ClinicoEconomics and Outcomes Research*, 13, 2021, 437.

[51] K. Degeling, H. Koffijberg and M. J. IJzerman, A systematic review and checklist presenting the main challenges for health economic modeling in personalized medicine: Towards implementing patient-level models, *Expert Review of Pharmacoeconomics & Outcomes Research*, 17(1), 2017, 17-25.

[52] J. Britt, M. Fazle Baki, A. Azab, A. Chaouch and X. Li, A stochastic hierarchical approach for the master surgical scheduling problem, *Computers & Industrial Engineering,* 158, 2021, 107385.

[53] V. Kayvanfar, M. R. Akbari Jokar, M. Rafiee, S. Sheikh and R. Iranzad, A new model for operating room scheduling with elective patient strategy, *INFOR: Information Systems and Operational Research,* 2021, 1-24.

[54] R. K. Ahuja, O. Ergun, J. B. Orlin and A. P. Punnen, Very large-scale neighborhood search: Theory, algorithms, and applications, *Handbook of Approximation Algorithms and Metaheuristics*, Chapman and Hall/CRC, 2018.

[55] P. Hansen and N. Mladenović. Variable neighborhood search, In *Search Methodologies,* Springer, 2014, 313-337.

[56] A. Sifaleras, I. Konstantaras and N. Mladenović, Variable neighborhood search for the economic lot sizing problem with product returns and recovery, *International Journal of Production Economics*, 160, 2015, 133-143.

[57] M. C. Fu, F. W. Glover and J. April, Simulation optimization: A review, new developments, and applications, In *Proceedings of the 37th Conference on Winter Simulation,* Orlando, USA, 2005, 83-95.

[58] L. Siciliani and J. Hurst, Tackling excessive waiting times for elective surgery: A comparative analysis of policies in 12 OECD countries, *Health Policy*, 72(2), 2005, 201-215.