# Supporting Capacity Management Decisions in Healthcare using Data-Driven Process Simulation

Peer-reviewed author version

VAN HULZEN, Gerard; MARTIN, Niels; DEPAIRE, Benoit & SOUVERIJNS, Geert (2022) Supporting Capacity Management Decisions in Healthcare using Data-Driven Process Simulation. In: Journal of biomedical informatics, 129 (Art N° 104060).

DOI: 10.1016/j.jbi.2022.104060
Handle: http://hdl.handle.net/1942/37267

# Supporting Capacity Management Decisions in Healthcare using Data-Driven Process Simulation

Gerhardus van Hulzen[a,*], Niels Martin[a,b], Benoît Depaire[a], Geert Souverijns[c]

[a] *Hasselt University, Research group Business Informatics, Martelarenlaan 42, 3500 Hasselt, Belgium*
[b] *Research Foundation Flanders (FWO), Egmontstraat 5, 1000 Brussels, Belgium*
[c] *Jessa Ziekenhuis, Stadsomvaart 11, 3500 Hasselt, Belgium*

**Abstract**

Healthcare managers are confronted with various Capacity Management decisions to determine appropriate levels of resources such as equipment and staff. Given the significant impact of these decisions, they should be taken with great care. The increasing amount of process execution data – i.e. event logs – stored in Hospital Information Systems (HIS) can be leveraged using Data-Driven Process Simulation (DDPS), an emerging field of Process Mining, to provide decision-support information to healthcare managers. While existing research on DDPS mainly focuses on the fully automated discovery of simulation models from event logs, the interaction between process execution data and domain expertise has received little attention. Nevertheless, data quality issues in real-life process execution data stored in HIS prevent the discovery of accurate and reliable models from this data. Therefore, complementary information from domain experts is necessary. In this paper, we describe the application of DDPS in healthcare by means of an extensive real-life case study at the radiology department of a Belgium hospital. In addition to formulating our recommendations towards the radiology management, we will elaborate on the experienced challenges and formulate recommendations to move research on DDPS within a healthcare context forward. In this respect, explicit attention

*Corresponding author
*Email address:* gerard.vanhulzen@uhasselt.be (Gerhardus van Hulzen)

is attributed to data quality assessment, as well as the interaction between the use of process execution data and domain expertise.

*Keywords:* Data-driven process simulation, Process mining, Capacity management, Healthcare processes, Domain knowledge

## 1. Introduction

Healthcare organisations are confronted with various challenges, including the increasing and ageing population, as well as the rapid evolution in medical equipment and technologies. At the same time, tightening government budgets put financial pressure on healthcare budgets [1]. Within this context, healthcare managers need to manage their scarce resources efficiently in order to safeguard high-quality healthcare services [2]. *Capacity Management (CM) decisions* in healthcare deal with establishing the suitable levels of resources, such as equipment and facilities (e.g. X-ray devices, beds, sterile instruments, operating theatres, etc.), and staff [3, 4]. Depending on the scale and time horizon of these decisions, correctly estimating the required resources is crucial, as wrong decisions could severely impact healthcare processes, patient and staff well-being, and the overall financial situation of the department or healthcare organisation [5].

Several CM decisions, e.g. regarding the acquisition of new medical devices, are of high strategic importance as they will have a long-lasting impact on the operations of a medical department and, given the close interdependency between departments [6], even for the hospital as a whole. While having too limited resource capacity will be detrimental for the quality of the healthcare services provided by the department (e.g. due to significant rises in waiting times), having too much idle capacity is equally undesirable given the significant investments required in, e.g. the purchase and maintenance of medical devices, and recurring personnel costs due to sub-optimal use of staff. Consequently, it is crucial to make well-informed decisions within a context where a multitude of choices can typically be made.

Against this background, *Process Simulation* can offer valuable decision-support information for healthcare managers confronted with CM decisions. In Process Simulation, a computer model of a process is simulated to imitate the behaviour of a process in a virtual setting. This has the advantage of being able to test process modifications without having to implement these changes in practice [7]. Using Process Simulation, one could, e.g., safely test whether the removal of an under-utilised medical device would not result in significantly higher waiting times without having to try this in reality, and thereby unnecessarily endangering patients' health.

In order to develop a simulation model, insights into the process behaviour should be gathered. This relates, amongst others, to the order of activities, their duration, and the availability of resources. Information sources such as consultations with domain experts and observations of the real-life process can be used to gather the required information. Moreover, as process execution data is increasingly being captured by Health Information Systems (HIS), this data is increasingly considered as an information source [8]. When process execution data – i.e. an event log – is extensively used during the development of a simulation model, this is referred to as *Data-Driven Process Simulation (DDPS)* [9].

Existing research on DDPS often focuses on the fully automated discovery of a simulation model from process execution events [10, 11]. However, in a real-life healthcare setting, it is widely recognised that process execution data suffers from a multitude of data quality issues, such as missing data (e.g. an activity is executed but does not lead to a registration in the system) and incorrect timestamps (e.g. an activity is recorded in the system some time after it has been executed) [12, 13]. The presence of these data quality issues limits the potential of solely relying on process execution data when developing a simulation model in a real-life setting. Consequently, the information that the data provides should be complemented with additional information sources such as domain expertise [14]. The interaction between process execution data and domain expertise when conducting a DDPS analysis received little attention in

existing research on DDPS.

In this paper, we describe a real-life DDPS project by means of an extensive case study at the radiology department of a Belgian hospital. Within the context of the construction of new facilities, the department is confronted with several CM decisions: "How much X-ray equipment is needed? How large should the waiting area be? How many receptionists are needed to ensure a smooth flow of patients? ...". Besides reporting on the simulation model's development and the analysis results, we will elaborate on the experienced challenges and formulate recommendations to move research on DDPS within a healthcare context forward. In particular, we draw explicit attention to data quality assessment, as well as the interaction between the use of process execution data and domain expertise, two elements that are often neglected in DDPS literature. In addition, we highlight some challenges concerning the operationalisation of DDPS and directions for future work.

The remainder of this paper is structured as follows. Section 2 gives an overview of the related work. The case study applying DDPS to support CM decisions is outlined in Section 3. The results of this case study and our recommendations towards the radiology management are presented in Section 4. An initial conceptualisation of a method for applying DDPS to provide hospital management with evidence-based CM recommendations is provided in Section 5. The paper ends with a conclusion in Section 6.

## 2. Related Work

This work is related to three key domains, i.e. (i) Capacity Management in healthcare, (ii) Data-Driven Process Simulation, and (iii) data quality issues in event logs. The following subsections give a brief overview of prior research in these domains.

### *2.1. Capacity Management in Healthcare*

*Capacity Management decisions in healthcare* deal with establishing the suitable resource levels in terms of equipment (e.g. sterile instruments), facilities

(e.g. operating theatres), and staff (e.g. nurses) [3, 4]. Depending on the time horizon, these decisions can be (i) strategical, (ii) tactical, or (iii) operational [4, 15]. Strategical decisions have a long planning horizon and typically require significant capital investment [4], such as resource capacity expansions (e.g. the acquisition of a new computed tomography (CT) radiology device [16], cardiotocography (CTG) machines [17], or even the design of entire facilities [18, 19]). Tactical decisions are situated between strategical and operational decisions concerning the planning horizon [4]. Typical decisions made on a tactical level are staff planning [17, 20, 21] or block planning for operating theatres [22]. Finally, operational planning involves short-term decisions. On this level, there is low flexibility as the resource levels are fixed [4]. Examples of operational decisions are patient appointment scheduling [23, 24], staff scheduling [25, 26], inventory replenishing [27], and emergency scheduling [28].

Several Operations Research (OR) techniques can be applied to support decisions makers confronted with CM assignments, such as Markov processes, queueing models, mathematical programming, and computer simulation [15, 17]. Due to the stochastic nature of healthcare processes, computer simulation – in particular *Discrete-Event Simulation (DES)* – is one of the preferred techniques to analyse these complex processes [29–31]. In a DES model, individual *entities* (e.g. patients) move through the process and wait in queues to be served by *resources* (e.g. nurses) [8, 30]. These entities can have attributes that describe them (e.g. age, patient type, diagnosis, etc.) and can be used to determine their pathway through the simulation model [30]. Unlike other OR techniques, such as queueing models or mathematical programming, DES enables measuring the individual characteristics of entities such as the waiting time or treatment outcome. This makes it possible to analyse the process in detail and investigate the effects of both small and substantial changes applied to the model. For further reference on the use of simulation in healthcare, the reader is referred to one of the existing review papers [32–36].

### 2.2. Data-Driven Process Simulation

In *Data-Driven Process Simulation* a simulation model is built by extensively using process execution data (i.e. event logs) [9], originating from, e.g. Health Information Systems (HIS). DDPS is an emerging field in *Process Mining*, which focuses on the retrieval of process-related insights from event logs [37]. Originally, PM research mainly focused on discovering the control-flow (e.g. patient pathways) from event logs. Over time, more and more algorithms have been developed to discover other components of simulation models from event logs as well, such as arrival patterns, queueing disciplines and resource schedules [8].

Despite the advances in process mining research, combining all these components (semi-)automatically into a single, simulation-ready model is still in its infancy. A few research efforts have been conducted on the discovery of simulation model components from event logs and combining them manually to derive a simulation-ready model [38, 39]. Only recently, the discovery and simulation of such models from event logs were fully automated by Camargo *et al.* [10] in their tool *Simod*, which was later extended to also take the presence of multitasking and resource availability into account [11]. While these studies demonstrate the feasibility of solely using event logs to construct simulation-ready models, many assumptions are usually made that oversimplify the model or are unrealistic in a healthcare setting, e.g. no resource constraints [39], or waiting times are explicitly modelled, instead of delays due to resource unavailability [38]. In addition, these studies assume the presence of clean and high-quality event logs [10, 11]; an assumption that often not holds in process execution data from healthcare processes [12, 13, 40]. Moreover, healthcare processes are knowledge-intensive and heavily rely on clinicians' expertise and experience. This rich knowledge is not contained in process execution data [41].

Instead of discovering and simulating models from event logs in a fully automated way, most research efforts leverage insights obtained from PM techniques and use these to construct simulation models manually. In this manner, control-flow discovery techniques are often applied to derive the patient's

journey through the care facility for specific diseases [42] and use Process Simulation to analyse and improve the performance of these care processes. For example, Augusto *et al.* [43] simulated the impact of different implantable cardioverter-defibrillator strategies for patients with cardiovascular diseases, Mans *et al.* [44] determined the impact of digitising the dental prosthesis process, Kovalchuk *et al.* [45] studied acute coronary syndrome patient flows, Tamburis and Esposito [46] analysed process execution data of a cataract process, and Johnson *et al.* [42] applied PM and simulation to alcohol-related emergency admissions, giant-cell arteritis, and functional neurological symptoms care.

While the aforementioned works focus on a particular disease, other studies analyse entire departments or care facilities to reduce patient waiting time and Length of Stay (LoS) using PM and simulation. Zhou *et al.* [47] altered the number of receptionists, nurses, and doctors to improve the performance of an outpatient clinic. A similar approach was used in the Emergency Department by Antunes *et al.* [26] to reduce patient waiting time by determining the optimal physician scheduling, and by Abohamad *et al.* [48] to identify performance bottlenecks and to explore improvement strategies to reduce LoS.

### *2.3. Data Quality Issues in Event Logs*

As mentioned in the previous section, event logs often contain *data quality issues*, especially in a healthcare context [12, 13, 40]. Various taxonomies and frameworks have been developed to identify, assess, and handle these issues. This section highlights some key research efforts on data quality in the Process Mining field. A more in-depth overview can be found in Martin [13].

A well-known taxonomy of event log quality issues is presented in Bose *et al.* [49], where 27 distinct issues are subdivided into four main categories: missing, incorrect, imprecise, and irrelevant data. These four main categories can apply to cases (e.g. missing cases that were not recorded), events (e.g. incorrect events which did not happen in reality), and attributes (e.g. imprecise timestamps recorded at a coarse level). A more applied taxonomy is proposed by Suriadi *et al.* [50], in which eleven fine-grained patterns of common data quality

issues encountered in event logs are distinguished, e.g. "elusive cases", which are events that are not linked to any case; or "distorted labels" which often indicate a typo.

Within the context of healthcare, Vanbrabant *et al.* [40] developed a synthesised taxonomy based on both general and healthcare-specific data quality issues. In total, fourteen specific event log quality issues are defined. A distinction was made between missing and non-missing data, where the latter category is further subdivided in wrong data on the one hand, which have to be corrected or removed in order to find meaningful insights from the data (e.g. timestamps violating the logical order of activities or values outside the domain range), and "not wrong but not directly usable" data on the other hand, which requires minor preprocessing (e.g. inconsistent formatting of timestamps or embedded values which contain aggregated information).

In addition to the taxonomies of various data quality issues encountered in event logs, several frameworks to assess event log quality have been proposed as well. A seven-step cyclical conceptual framework is provided by Andrews *et al.* [51] to detect quality issues such as coarse granularity (e.g. some timestamps were recorded at day-level, whereas others at millisecond-level granularity), null values, and the level of uniqueness of attributes. Kherbouche *et al.* [52] implemented an event log quality assessment framework in *ProM*[1] which considers complexity (e.g. number of events or average trace length), accuracy (e.g. rate of erroneous timestamps), consistency (e.g. free of outliers), and completeness (e.g. no events or attributes are missing). Other frameworks implemented in ProM by Fischer *et al.* [53] and Dixit *et al.* [54] focus specifically on identifying and assessing timestamp-related quality issues. Fox *et al.* [55] provided a framework for assessing Electronic Health Record (EHR) data called *Care Pathway Data Quality Framework (CP-DQF)* using a SQL database to maintain a data quality register. However, the different types of quality issues have to be manually provided by the analyst beforehand. A similar approach

[1] https://www.promtools.org/

using SQL was proposed by Andrews *et al.* [56], in which five event log imperfection patterns defined by Suriadi *et al.* [50] are detected. The R-package[2] *DaQAPO* by Martin *et al.* [57] supports ten imperfection patterns defined by Suriadi *et al.* [50] along with additional event log quality tests, such as the detection of batch registrations and mutual dependencies between activities in the process.

All aforementioned frameworks – except Dixit *et al.* [54] – focus on event log quality assessment. It is up to the analyst to use these insights while cleaning the data. The tool implemented in ProM by Dixit *et al.* [54] is also capable of "repairing" the event log by inserting or removing events to make sure the ordering of events is correct. Other works also propose data cleaning heuristics, typically targeted on a particular event log quality issue. For instance, Bayomie *et al.* [58] developed an approach to handle missing case identifiers, while Di Francescomarino *et al.* [59] focused on imputing non-observable activities in an event log using a control-flow model created by domain experts. Martin *et al.* [60] proposed an interactive data cleaning approach with heuristics to solve missing timestamps, overlapping timestamps, time ordering violations, and invalid concurrent use of resources.

Even though more and more attention is devoted to the assessment and correction of data quality issues in the field of Process Mining, within the context of (automated) DDPS model discovery little explicit attention has been paid to these issues. Nevertheless, the quality of the input data has a significant impact on the reliability of the simulation results [40, 61]. Therefore, it is imperative to deal with these issues appropriately. Some of the discussed papers on DDPS briefly mention the presence of data quality issues [42, 45, 47] but do not describe the indispensable role of the domain experts. Only Johnson *et al.* [42] addresses the importance of the "Clinical Review Board" in their method *ClearPath*. We differentiate ourselves from the work of Johnson *et al.* [42] in the structure of the method. While ClearPath is an extension of the established Process

[2] https://www.r-project.org/

Mining project methodology PM$^2$ [62], we base our method on guidelines and best practices in the Process Simulation literature and describe explicit action points to be considered for every step.

## 3. Real-Life Case Study – Outline

This section outlines a real-life case study using DDPS to support healthcare management when confronted with CM decisions. In particular, the case study focuses on CM decisions at the radiology department of a Belgian hospital.

Firstly, the case study and its context is briefly introduced. Secondly, a high-level description of the simulation model is provided. Thirdly, the Capacity Management questions are formulated. Fourthly, a thorough description of the simulation model development is provided, including the data that were used, the configuration of the simulation model, and assumptions and simplifications that were made. Fifthly, the verification and validation of the model are discussed. Finally, the scenarios are formulated based on the CM questions.

### *3.1. Context*

A Belgian hospital is planning the construction of new facilities, which implies the centralisation of care activities to a single health campus. Within this context, hospital management is considering merging two of their radiology facilities into one location. Therefore, the management of the radiology department has to decide how many resources they will need in the new facility to serve all patients in a timely manner. These resources include radiology equipment, waiting room area for ambulatory patients (i.e. the number of seats), and receptionists.

Several discussions with the team of domain experts were held throughout the project, e.g. to specify the process and scope of the analysis, define the CM questions, discuss data quality issues, validate the simulation model, and report the results and evidence-based recommendations. This team of experts consisted of the head of the radiology department, the chief secretary of the

medical imaging department, two chief nurses of the radiology department, a medical imaging engineer, and the care manager laboratory medicine and imaging medicine.

### *3.2. Model Description*

The simulation model describes a standard medical imaging diagnostics process in a radiology department. A simplified process flow is depicted in Figure 1. A patient has to be registered by the reception upon arrival. For ambulatory and mobile emergency patients, this is performed by the front office at the counter, whereas hospitalised, day hospital, and trauma patients are registered by the back office. The back office will also assist the front office during busy periods. Outside the reception opening hours, the registration is performed by a nurse. Next, ambulatory patients will wait in the designated waiting room until they are called into the examination room. All other patients either wait in the emergency department or wait in their room in the ward they are admitted to. Some ambulatory patients scheduled for a CT examination are required to drink contrast fluid and wait one hour before being examined. When the patient is called into the examination room, one of two nurses will install the patient on the equipment and checks the patient's identity, while the other will use this information to plan and execute the examination. If a patient requires multiple examinations, e.g. an X-ray of both knee and ankle, the patient does not have to leave the examination room, as they are made right after each other. Examinations of other modalities cannot be made in the same room. Therefore, the patient will have to go to the waiting room of the other examination room so that the process can be repeated. When no further examinations are required, the patient can leave the radiology department and go home or back to the emergency department or wardroom.

There are four different patient types defined in the model: (i) *ambulatory patients (A)*, which are outpatients with a scheduled appointment; (ii) *hospitalised patients (H)*, which are inpatients; (iii) *day hospital patients (D)*, patients with a scheduled appointment, but need to be admitted to the hospital for at

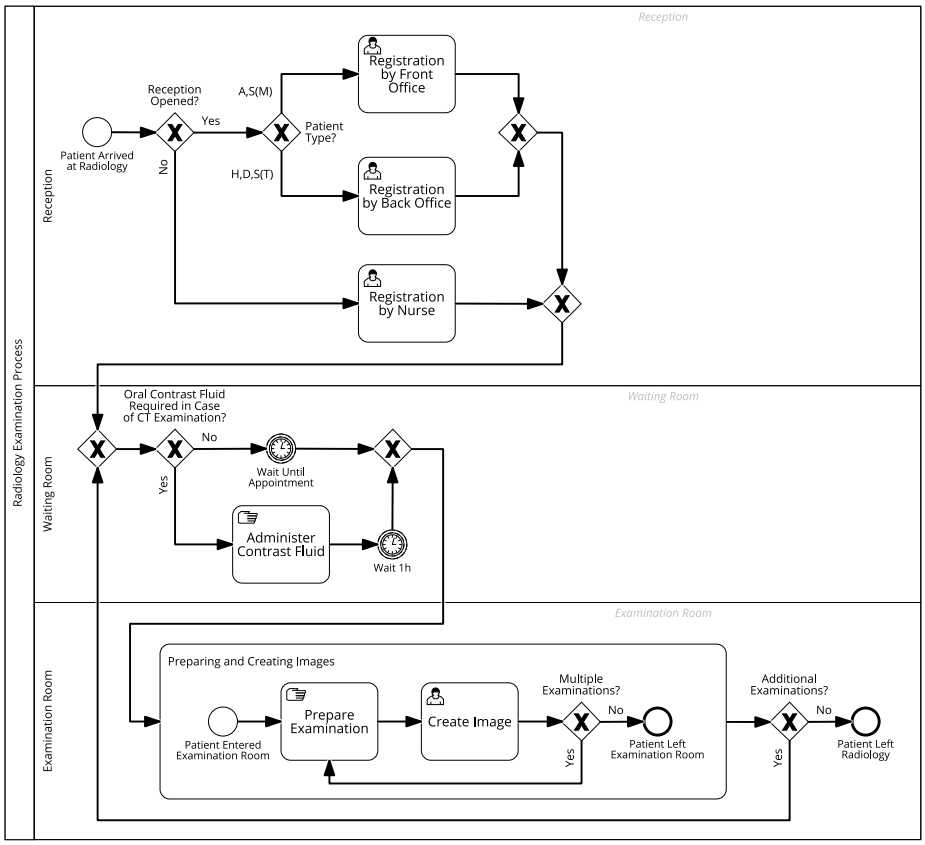


Figure 1: Simplified process flow of patients at the radiology department.

most one day; and (iv) *emergency patients (S)* transferred from the Emergency Department.

The patients can undergo various examination types. In this simulation study, there are six different modalities: *cone beam computed tomography (CBCT)*, *computed tomography (CT)*, *mammography (MAMMO)*, *magnetic resonance imaging (MRI)*, *ultrasound (US)*, and *X-ray*. Each examination type requires a separate examination room, and each examination room contains only one radiology device. The terms examination room, radiology device, and equipment are used interchangeably in this paper. In addition to the fixed X-ray equipment, a mobile X-ray device is also used to examine hospitalised and emergency patients who cannot be transferred to the examination room.

### *3.3. Capacity Management Questions*

The radiology department is considering merging two of their existing radiology facilities in the new hospital site. They have formulated three Capacity Management questions:

**CMQ 1.** *What is the impact of this centralisation on the utilisation of their resources, in particular, equipment, waiting area (i.e. the number of seats), and reception staff? They expect they would need less equipment than currently available, given the lower utilisation rates of some devices, and they would like to know whether this presentiment is correct, as this also has a major impact on personnel costs.*

**CMQ 2.** *To increase patient satisfaction, the radiology management is considering omitting the requirement for some ambulatory patients to drink contrast fluid one hour before having their CT examination. What is the impact on throughput time and required waiting area when this policy would be implemented?*

**CMQ 3.** *The radiology management is considering implementing an online registration system, allowing ambulatory patients to skip the reception and go straight to the waiting area of the examination room at the time of their appointment. The radiology department is interested in the effects of this new registration system on reception staff requirements and the required waiting area.*

### *3.4. Simulation Model Development*

The following subsections describe the development of the simulation model. First, we describe the data provided by the hospital and highlight the issues we encountered while working with this data. Next, we discuss some technical details about the configuration of the simulation model and give an overview of the assumptions and simplifications we applied to align the model's complexity with the desired scope.

### 3.4.1. Data

Two years of process execution data, from March 2017–March 2019, extracted from the hospital's Radiology Information System (RIS) was used to build the simulation model. This dataset contained various key timestamps, such as time of registration, and start and end time of examination, for each patient visit. In addition, the patient type and examination type were also recorded. Table 1 shows an overview of the attributes stored in the dataset with their description. Each row in the dataset represents a unique examination, and multiple rows could be linked to the same patient visit. An additional validation dataset from March 2019–March 2020 was used to validate the model.

| Label | Description |
|---|---|
| Accession number | Anonymised identifier for each individual examination |
| Patient number | Anonymised identification number for each patient visit |
| Entity | Radiology facility the patient visited |
| AHDS | Patient type: ambulatory (A), hospitalised (H), day hospital (D), and emergency (S, for "spoed" in Dutch) |
| Examination type | The performed type (e.g. CT, US, X-ray, etc.) |
| Examination code | Specific code of the examination (e.g. "CTKNIEL", a CT examination of the left knee) |
| Examination room | The room in which the examination was performed |
| Time booking appointment | Timestamp when the appointment was booked |
| Time start appointment | Timestamp of the appointment |
| Time creation request | Timestamp when the patient was registered by the reception |
| Time start examination | Timestamp when the patient was called into the examination room |
| Time start scan | Timestamp when the radiology device was started |
| Time end scan | Timestamp when the radiology device was finished |
| Time finish examination | Timestamp when the nurse finalised the images, including post-processing |
| Time creation report | Timestamp when the radiologist started working on the report |
| Time validation report | Timestamp when the radiologist finalised and validated the report |

Table 1: Overview of the attributes stored in the dataset with their description.

Table 2 gives an overview of the number of patient visits and examination types contained in the dataset used to build the model. The majority of patients were ambulatory, and the most commonly made examination was X-ray. A total of 478,342 examinations were recorded, which indicates that some patients required multiple examinations. Nevertheless, the majority of patients only needed a single examination. *Other* modalities included very infrequently made examinations, such as bone mineral content (BMC) and angiography, which were not of interest in the scope of this simulation study.

| Patient type | Proportion |
|---|---|
| Ambulatory (A) | 60% |
| Hospitalised (H) | 23% |
| Emergency (S) | 15% |
| Day hospital (D) | 2% |
| Total | 404,750 |

| Examination type | Proportion |
|---|---|
| X-ray | 44% |
| US | 18% |
| MRI | 16% |
| CT | 13% |
| MAMMO | 5% |
| *Other* | 3% |
| CBCT | 0.6% |
| Total | 478,342 |

Table 2: Proportion of patient and examination types in the dataset.

One of the biggest challenges of this project was dealing with data quality issues. Because the quality of the input data has a significant impact on the reliability of the simulation results [40, 61], it is imperative to deal with these issues appropriately. The two most prominent data quality issues were: (i) unavailable timestamps, such as the exact arrival time of the patient or no start timestamp of certain activities, and (ii) incorrectly stored timestamps, either caused by the system or "batch registration" by radiology personnel, in which administrative tasks are postponed and bundled to be dealt with later [40]. The first issue was resolved by using a different timestamp as a proxy of the arrival and asking the domain experts to identify the minimum, maximum, and most likely time required to complete a particular activity, which served as input parameters for a triangular distribution.

The second issue manifested itself mainly by large outliers of some activity durations, e.g. the scanning times. Both start and end timestamps were available in the dataset and recorded when the nurse starts and stops the radiology device, respectively. An overview of scan duration times per examination type is given in Table 3. Some scans took precisely zero seconds, while others took several years, and a few observations were even negative, indicating that the end timestamp occurred before the start timestamp. In our previous work [61], we highlighted the impact of data quality issues, specifically related to these scanning durations. Completely neglecting these issues, i.e. directly using the data without any pre-processing or filtering, would result in significantly higher throughput and waiting times due to congestion [61]. Therefore, it is necessary to exclude anomalies from the data. However, it is not trivial to determine which threshold values should be used to distinguish anomalies without domain knowledge. Therefore, domain knowledge is indispensable while analysing and processing data for simulation studies.

Nonetheless, if no domain expertise is available, statistical techniques could be applied to detect anomalies, such as the commonly used *box plot rule*, which marks any observation smaller than $Q_1 - 1.5IQR$ or larger than $Q_3 + 1.5IQR$ as an anomaly [63]. However, this technique is not guaranteed to mark anomalies correctly, as some of these observations might be exceptional, i.e. unlikely, but not impossible [61]. In this study, we validated the scanning durations with the domain experts by verifying the durations for each examination type. For instance, the domain experts pointed out that scanning durations that took exactly zero seconds were most likely so-called "one-shot" images, and the duration could be substituted by one minute instead. For each examination type, the domain experts also determined a maximum duration, e.g. CT examinations are unlikely to take longer than twenty minutes.

#### *3.4.2. Configuration*

The hospital is opened 24/7, although regular "opening hours" from 8AM–6PM are used to indicate the time period which is usually the busiest. Hos-

| Examination type | Min | Max | Mean | Median | SD | IQR |
|---|---|---|---|---|---|---|
| CBCT | 0.00 | 32.07 | 0.06 | 0.00 | 1.09 | 0.00 |
| CT | -726.53 | 30,605 | 7.02 | 2.07 | 201.14 | 2.22 |
| MAMMO | -6.48 | 40,780 | 16.41 | 2.98 | 531.00 | 1.35 |
| MRI | 0.00 | 946,449 | 161.22 | 11.48 | 9,680 | 6.90 |
| US | -79.00 | 116,685 | 71.36 | 23.38 | 636.37 | 28.48 |
| X-ray | -1,031.63 | 2,109,457 | 22.69 | 0.55 | 5,111 | 1.20 |

Table 3: Scanning duration per examination type (in mins).

pitalised and emergency patients can arrive any time of the day. Ambulatory patients only arrive from 7AM–10PM during the week and from 7AM–9PM during weekends. Day hospital patients arrive between 7AM–3PM during the week only.

The number of patients arriving is different throughout the day, which is also shown in Figure 2 for ambulatory patients. The morning starts quietly, and gradually more and more ambulatory patients arrive throughout the morning. At noon it gets a little quieter, after which it gets busier again in the course of the afternoon. Starting from the early evening, the number of patients arriving gradually decreases until the last ambulatory patients arrive before 11PM. Therefore, we modelled all patient arrivals as non-stationary Poisson processes with varying hourly rates obtained from data. Non-stationary Poisson processes are used to model many real-world systems, where the arrival rate $\lambda(t)$ varies over time $t$ [64].

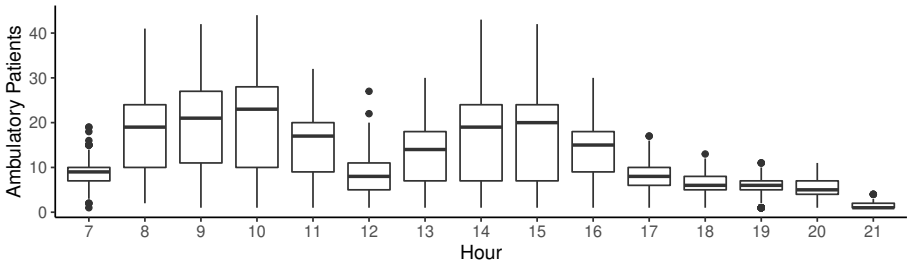


Figure 2: Number of ambulatory patients arriving throughout the day.

Simulation depends on the concept of random numbers for input, e.g., drawing samples from the specified probability distributions to generate patient arrivals or determining the required scanning durations. This input randomness induces randomness in the output as well [65]. Consequently, the output measures of the simulation model might differ each time the same model is simulated. By extending the simulation run length, the variance in the output measures decreases, leading to similar results each time the same model is simulated [64, 65]. Therefore, one long simulation run of 730 days was used to limit the impact of randomness.

In addition, common random number (CRN) streams with a fixed seed were used to mitigate the effects of randomness when comparing different scenarios. CRNs are the most commonly used variance reduction technique (VRT). The basic idea is that the same random numbers are used to draw samples from distributions across scenarios, ensuring that different results cannot be attributed to different random samples, but because of the changes to the process purposefully introduced by the modeller [64]. The simulation model was build using Arena 16.10 [66] which has all the required features to create and run simulation models.

### *3.4.3. Assumptions & Simplifications*

The following assumptions and simplifications have been drawn up in consultation with the domain experts to simplify the model in correspondence with the desired scope:

- Only ambulatory patients wait in the designated waiting area. All other patient types, i.e. hospitalised, day hospital, and emergency patients, wait in their wardroom or at the emergency department (which are not included in the model) and are assumed to be present in the radiology department whenever they are called into the examination room;
- Each examination room has a separate waiting area. If, in reality, multiple examination rooms would share a common waiting area, the required

capacity of the shared waiting area can be derived afterwards from the separate waiting areas;

- All queues follow the same priority logic in which emergency patients have the highest priority, followed by ambulatory, day hospital, and finally, hospitalised patients. Ongoing examinations are never aborted when a patient with a higher priority is presented. If a patient requires multiple examinations of the same type, these examinations are all performed on the same radiology device right after each other. This flow is not aborted in the presence of a high priority patient;

- Any examination type can occur at any time of the day, i.e. no time blocks are reserved for specific examination types, while the latter is common practice in a typical radiology department. However, including an appointment system with reserved time blocks would significantly increase the complexity of the model and is, therefore, not implemented;

- Any patient can undergo a maximum of four examinations, as more than four examinations per patient is highly exceptional, i.e. less than 0.15% of the recorded patient visits in the dataset required more than four examinations. These four examinations do not have to be all of the same modality;

- The duration of the examination is independent of the patient type, health condition, radiology device, time of the day, etc.;

- The movement of patients from the reception to the examination rooms and between examination rooms is not included in the model, as the impact of these movements on the process flow is deemed minimal;

- Patients arrive alone. This can have an impact on the number of seats used in the waiting rooms, as patients are often accompanied by their partner, a parent, or child(ren). However, no information was available on how frequent patients are accompanied. Moreover, including this would make

the model very complex, as not all "visits" would require examinations and leave when the patient they accompany has undergone all examinations;

- In reality, each examination room has two dressing rooms allowing the next patient to change while the previous patient is being examined. This way, no time is lost when the radiology device becomes available again. Therefore, dressing rooms and time lost due to (un)dressing are not included in the model;

- The back office reception is also responsible for handling incoming calls of the call centre. However, we did not include the call centre operations in the model because this would increase the model's complexity. Instead, the utilisation rates of the back office can be used to determine whether there is enough capacity left for handling incoming calls;

- When an examination room is opened, it is assumed to be staffed by the required number of nurses. Therefore, nurses are not modelled as resources. Neither for their duties in the reception;

- The interpretation of the images by a radiologist does not impact the utilisation of radiology equipment and is, therefore, not included in the model.

### *3.5. Model Verification & Validation*

To enhance the credibility of simulation models, each model should be verified to make sure that it contains no errors, and validated to guarantee that the output corresponds to the reality [64, 67–71]. A combination of statistical validation and face validation by the domain experts was carried out. For validation purposes, the operations of the largest facility were modelled as it constitutes the starting point for the scenario analysis.

Especially the number of patients arriving, the throughput times, and waiting times were the key statistics to compare. Figure 3 shows a comparison of weekly patient visits by patient type between historical data from the validation dataset and the simulation model's output. Both mean and median weekly

patient visits are very close to one another, but the historical data has a higher variation than the simulation data. Nevertheless, as shown in Table 4, the non-parametric *Wilcoxon-Mann-Whitney (WMW)* test could not detect statistically significant differences between the means at a 95% significance level. Unlike the $t$-test, which assumes that normality is satisfied, the *WMW*-test does not assume any underlying distribution [72]. A similar assessment of the correspondence between historical data and the simulation model's output for throughput times and waiting times resulted in similar findings.

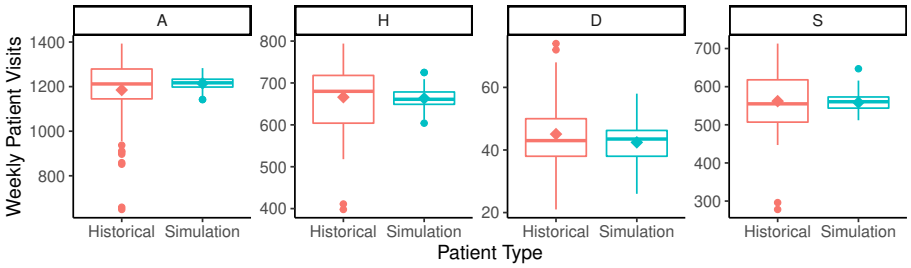


Figure 3: Comparison of weekly patient visits by patient type between historical and simulation model. The diamond represents the average weekly patient visits.

| Patient type | Historical mean | Simulation mean | $p$-value | Significance |
|---|---|---|---|---|
| A | 1,184.8 | 1,215.0 | 0.702 | ns |
| H | 666.0 | 666.2 | 0.128 | ns |
| D | 45.1 | 42.8 | 0.129 | ns |
| S | 562.0 | 557.0 | 0.806 | ns |

Table 4: Mean weekly patient visits per patient type of the historical data versus the simulation model (ns = not significant at 95% significance level)

.

In addition to the statistical validation, we also presented and discussed the results of the validation model with the team of domain experts. During the validation, the domain experts identified some issues with the opening hours of the examination rooms. These were extracted from the process execution data of the RIS. However, due to data quality issues, these were modelled too broad.

They also pointed out that the receptionist staff work with alternating roles in the front office, back office, and call centre. As this way of working was highly appreciated by the reception staff, they wanted to preserve this practice. After two iterations of validating and adjusting the model, the domain experts were convinced that the model accurately reflected reality.

### *3.6. Scenario Analysis*

Consistent with the Capacity Management questions formulated in Section 3.3 we defined the following scenarios in consultation with the domain experts:

**S 1.** *"Facility Merger": What is the required capacity for radiology devices, waiting area (i.e. the number of seats), and reception staff when merging the two largest radiology facilities to one location to accommodate all patients in a timely manner?*

**S 2.** *"Contrast Fluid": Some ambulatory patients requiring a CT examination need to drink contrast fluid one hour before they can be examined. What is the impact on throughput time and required waiting area if this policy is nullified?*

**S 3.** *"Online Registration": The hospital is considering letting (a proportion of) the ambulatory patients register themselves online before they visit the hospital. This allows them to go straight to the waiting area of the examination room instead of having to pass by the reception. What will be the impact on waiting time, waiting area dimensions, and reception staff requirements if this new registration system were to be adopted? We will evaluate the impact when* 30% *of ambulatory patients would register online upfront.*

## 4. Real-Life Case Study – Results & Recommendations

In this section, we describe the approach, present the numerical results, and provide our recommendations towards the radiology management for each scenario of the real-life case study.

### *4.1. Scenario 1 – Facility Merger*

Scenario 1 starts with the same number of examination rooms as the validated model. However, the patient visits of the second facility are added as this resembles the future situation at the new facilities. The number of weekly patient visits is increased by 71%, 30%, 117% for ambulatory, hospitalised, and day hospital patients, respectively. There is no increase in weekly emergency patient visits, as the second facility does not have an Emergency Department. Because the available capacity of the reception is already almost fully utilised in the validated model – which was also confirmed by the domain experts – we decided to add the receptionists of the second facility from the start, so no severe delays due to congestion at the reception would propagate to the examination rooms.

To investigate the first scenario, three alternatives are considered besides the *Base* configuration. The number of resources – i.e. reception and examination rooms – with their opening hours for each alternative are shown in Table 5. The *Base* configuration corresponds with the validated model of the largest facility and serves as a baseline for comparison. Note that CBCT examinations are only performed in the second facility, so this modality is not included in the *Base* configuration. Each alternative configuration has the same resources as the previous configuration and some additional resources. In configuration *S1a*, the patients of the second facility are added, the reception is expanded, and a CBCT radiology device is included. Next, in configuration *S1b*, an additional US, MRI, and X-ray device are added, which is repeated once more in configuration *S1c*.

#### *4.1.1. Numerical Results*

*Examination Rooms.* The results clearly show that the available examination rooms are insufficient to accommodate all patients on time for examination types US, MRI, and X-ray (albeit less problematic for X-ray). This manifests itself on the one hand by the very high utilisation rates of these examination rooms and on high waiting times on the other hand.

Figure 4 shows the hourly utilisation rates for each examination room for

| Configuration | Resources | |
|---|---|---|
| Base | Reception | FO: 1 Mon–Fri (7AM–5PM); BO: 3 Mon–Fri (8AM–6PM), 1 Mon–Fri (6PM–10PM), 1 Sat–Sun (7AM–4PM) |
| | CBCT | - |
| | CT | $CT_1$: Mon–Fri (8AM–6PM); $CT_2$: 24/7 |
| | MAMMO | $MAMMO_1$: Mon–Fri (8AM–6PM) |
| | MRI | $MRI_{1,2}$: Mon–Fri (7AM–10PM), Sat–Sun (7AM–9PM) |
| | US | $US_1$: 24/7; $US_{2,3}$: Mon–Fri (8AM–5PM) |
| | X-ray | X-$ray_{1,3}$: 24/7; X-$ray_2$: 24/5; MX-$ray_1$: 24/7 |
| S1a | Base + | |
| | Reception | FO: 1 Mon–Fri (7AM–5PM); BO: 2 Mon–Fri (8AM–5PM), 1 Mon–Fri (5PM–8PM), 1 Sat–Sun (7AM–1PM) |
| | CBCT | $CBCT_1$: Mon–Fri (10AM–10PM), Sat–Sun (8AM–11AM & 1PM–5PM) |
| S1b | S1a + | |
| | MRI | $MRI_3$: Mon–Fri (7AM–10PM), Sat–Sun (7AM–9PM) |
| | US | $US_4$: 24/7 |
| | X-ray | X-$ray_4$: Mon–Fri (8AM–6PM) |
| S1c | S1b + | |
| | MRI | $MRI_4$: Mon–Fri (7AM–10PM), Sat–Sun (7AM–9PM) |
| | US | $US_5$: Mon–Fri (8AM–6PM) |
| | X-ray | X-$ray_5$: Mon–Fri (8AM–6PM) |

Table 5: Overview of the evaluated alternatives. Each alternative configuration has the same resources as the previous configuration with some additional resources (FO = front office; BO = back office, including call centre; MX-ray = mobile X-ray device).

each configuration. A utilisation rate of 1 (or 100%) means that the device was used non-stop for the entire hour, whereas a utilisation rate of 0 means that the device was idle for the entire hour.

From Figure 4, it follows that the CBCT device has plenty of spare capacity for accommodating additional patients, which is confirmed by the low average utilisation rates shown in Table 6 of 9% during the regular "opening hours" (8AM–6PM) on weekdays and 7% outside these opening hours and during the weekends. For CT and MAMMO, the radiology devices are more frequently used, but the utilisation rates remain acceptable to also serve the patients from

the second facility, with still some room for unplanned emergency examinations. The situation is slightly different for X-ray. Especially during the opening hours, the utilisation rates are fairly high, and the available devices can barely keep up with the flow of patients. Lastly, for the US and MRI devices, the results clearly show that these devices are overburdened during the opening hours, with average utilisation rates approximating 100% and overdue examinations must be made up during the evening and night.

| Configuration | Week | | Weekend | |
|---|---|---|---|---|
| | 8AM–6PM | <> | 8AM–6PM | <> |
| Reception (Front Office) | | | | |
| Base | 0.82 | 0.70 | - | - |
| S1a | 0.84 | 0.69 | - | - |
| Reception (Back Office) | | | | |
| Base | 0.73 | 0.67 | 0.74 | 0.66 |
| S1a | 0.70 | 0.69 | 0.60 | 0.61 |
| Reception (Back Office – Call Centre) | | | | |
| Base | 0.32 | - | - | - |
| S1a | 0.25 | - | - | - |
| CBCT | | | | |
| Base | - | - | - | - |
| S1a | 0.09 | 0.07 | 0.07 | 0.07 |
| CT | | | | |
| Base | 0.32 | 0.23 | 0.23 | 0.16 |
| S1a | 0.45 | 0.29 | 0.25 | 0.18 |
| MAMMO | | | | |
| Base | 0.23 | 0.16 | - | - |
| S1a | 0.33 | 0.20 | - | - |
| MRI | | | | |
| Base | 0.81 | 0.40 | 0.75 | 0.42 |
| S1a | 0.98 | 0.76 | 0.97 | 0.76 |
| S1b | 0.83 | 0.41 | 0.73 | 0.53 |
| S1c | 0.66 | 0.35 | 0.60 | 0.44 |
| US | | | | |
| Base | 0.75 | 0.75 | 0.50 | 0.43 |
| S1a | 0.94 | 0.95 | 0.78 | 0.81 |
| S1b | 0.85 | 0.69 | 0.36 | 0.32 |
| S1c | 0.74 | 0.57 | 0.36 | 0.32 |
| X-ray | | | | |
| Base | 0.55 | 0.29 | 0.40 | 0.25 |
| S1a | 0.79 | 0.34 | 0.43 | 0.26 |
| S1b | 0.64 | 0.33 | 0.43 | 0.26 |
| S1c | 0.54 | 0.33 | 0.43 | 0.26 |

Table 6: Average hourly utilisation rates of examination rooms per configuration. Regular opening hours are from 8AM–6PM, outside opening hours is denoted by "<>".

The average waiting times, shown in Table 7, confirm this shortage of X-ray, US, and MRI devices. In configuration *S1a*, ambulatory patients have to wait on average four times longer for US examinations than in the *Base*

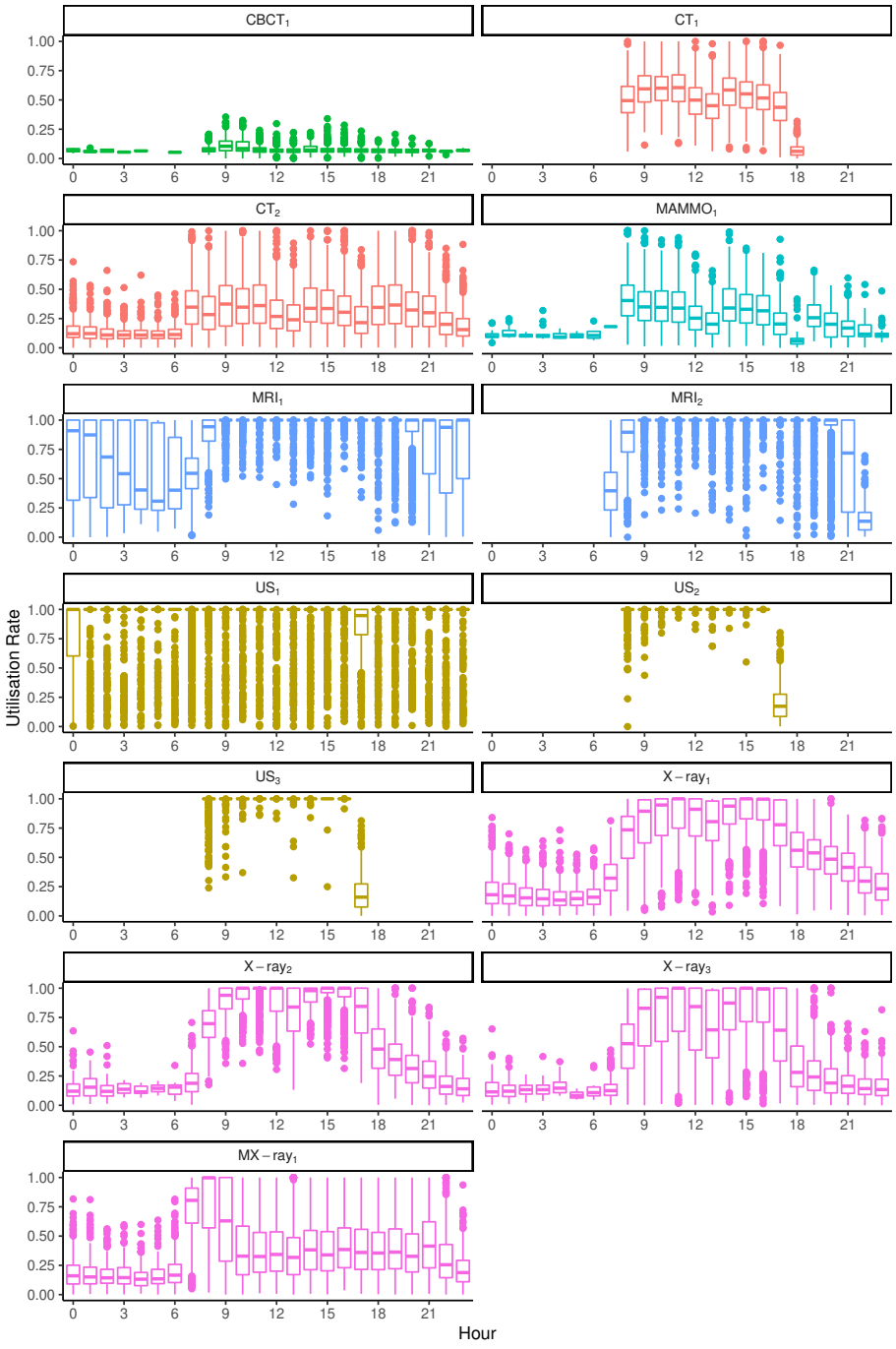


Figure 4: Hourly utilisation rates of examination rooms in configuration *S1a*.

configuration. For hospitalised patients, the situation is even worse because they have the lowest priority. This results in an average waiting time of over 30 hours, which is more than 26 times longer than in the *Base* configuration. The same observation can be made for MRI and X-ray, albeit to a less extreme degree.

By adding an additional radiology device for US, MRI, and X-ray in configuration *S1b*, the utilisation rates, and the waiting times in particular, decrease significantly to approximately the same levels as the *Base* configuration, except for X-ray during opening hours in the week. Moreover, for US and MRI devices, the *Base* configuration already showed relatively high utilisation rates. Therefore, we evaluated the effect of an additional US, MRI, and X-ray device in configuration *S1c*. This results in a further reduction of utilisation rates and waiting times to accommodate all patients in a timely manner, and creates more room for unplanned emergency examinations.

*Reception.* In configuration *S1a* we added the receptionists who are currently already scheduled in the second facility. As shown in Table 6, the average utilisation rates of the front office remain at the same levels as the *Base* configuration when the patients from the second facility are added. However, the back office has a slightly lower workload on average, especially during the weekends and for the receptionists handling the call centre. The reduced utilisation rates also result in lower waiting times for registration, as shown in Table 7. Another analysis regarding the reception is conducted in Scenario 3 and described in Section 4.3.

*Waiting Area.* The radiology management also wants to know how many seats are required for each waiting area for the examination rooms. For the sake of brevity, we only show the results for MRI examination rooms. The results for the other modalities are analogous. Table 8 shows an overview of the number of seats taken by ambulatory patients waiting for an MRI examination. Due to the high waiting times in configuration *S1a*, the required number of seats is significantly larger than in the *Base* configuration. An additional MRI device in

| Configuration | A | | H | | D | | S | |
|---|---|---|---|---|---|---|---|---|
| | Mean | ± | Mean | ± | Mean | ± | Mean | ± |
| Reception | | | | | | | | |
| Base | 7.46 | 0.06 | 1.99 | 0.03 | 0.49 | 0.04 | 1.27 | 0.02 |
| S1a | 6.08 | 0.06 | 1.21 | 0.03 | 0.09 | 0.01 | 0.94 | 0.01 |
| CBCT | | | | | | | | |
| Base | - | - | - | - | - | - | - | - |
| S1a | 11.43 | 0.55 | - | - | - | - | - | - |
| CT | | | | | | | | |
| Base | 16.79 | 0.30 | 0.91 | 0.06 | 0.52 | 0.09 | 0.70 | 0.04 |
| S1a | 19.85 | 0.25 | 2.21 | 0.10 | 1.92 | 0.17 | 0.96 | 0.04 |
| MAMMO | | | | | | | | |
| Base | 9.63 | 0.22 | - | - | 4.04 | 1.80 | - | - |
| S1a | 11.24 | 0.19 | - | - | 5.13 | 1.89 | - | - |
| MRI | | | | | | | | |
| Base | 24.31 | 0.20 | 50.94 | 2.37 | 47.87 | 7.48 | 7.26 | 2.13 |
| S1a | 81.79 | 0.51 | 356.7 | 7.32 | 429.71 | 17.64 | 7.52 | 1.55 |
| S1b | 21.81 | 0.17 | 45.23 | 2.07 | 53.26 | 5.10 | 5.87 | 2.05 |
| S1c | 13.82 | 0.11 | 7.27 | 0.52 | 8.30 | 1.18 | 4.76 | 2.08 |
| US | | | | | | | | |
| Base | 25.59 | 0.51 | 76.81 | 2.47 | 16.97 | 2.24 | 8.15 | 0.23 |
| S1a | 110.70 | 1.30 | 2,025.18 | 21.65 | 525.18 | 27.86 | 10.84 | 0.23 |
| S1b | 23.19 | 0.22 | 83.34 | 2.16 | 43.41 | 3.53 | 3.15 | 0.09 |
| S1c | 13.41 | 0.13 | 17.65 | 0.72 | 9.39 | 0.94 | 1.91 | 0.07 |
| X-ray | | | | | | | | |
| Base | 9.13 | 0.09 | 3.64 | 0.08 | 1.80 | 0.25 | 0.84 | 0.02 |
| S1a | 12.96 | 0.08 | 24.29 | 0.33 | 14.96 | 0.79 | 1.32 | 0.03 |
| S1b | 10.00 | 0.07 | 8.21 | 0.14 | 2.70 | 0.19 | 0.92 | 0.02 |
| S1c | 9.18 | 0.07 | 6.59 | 0.13 | 0.69 | 0.07 | 0.77 | 0.02 |

Table 7: Mean waiting times in minutes per configuration with 95% confidence intervals.

configuration *S1b* reduces the number of waiting ambulatory patients to around the same level as the *Base* configuration. Adding yet another MRI device further decreases the required number of seats in configuration *S1c* to two seats to ensure that all patients can wait seated in 95% of the cases.

| Configuration | Exam. room | Mean | Median | Q25 | Q75 | Q90 | Q95 |
|---|---|---|---|---|---|---|---|
| Base | $MRI_1$ | 1.34 | 1 | 0 | 2 | 3 | 4 |
| | $MRI_2$ | 1.45 | 1 | 1 | 2 | 3 | 4 |
| S1a | $MRI_1$ | 4.99 | 4 | 2 | 7 | 11 | 13 |
| | $MRI_2$ | 4.51 | 4 | 2 | 7 | 9 | 11 |
| S1b | $MRI_1$ | 1.23 | 1 | 0 | 2 | 3 | 3 |
| | $MRI_2$ | 1.27 | 1 | 0 | 2 | 3 | 4 |
| | $MRI_3$ | 1.34 | 1 | 0 | 2 | 3 | 4 |
| S1c | $MRI_1$ | 0.77 | 1 | 0 | 1 | 2 | 2 |
| | $MRI_2$ | 0.78 | 1 | 0 | 1 | 2 | 2 |
| | $MRI_3$ | 0.80 | 1 | 0 | 1 | 2 | 2 |
| | $MRI_4$ | 0.85 | 1 | 0 | 1 | 2 | 2 |

Table 8: Number of seats taken by ambulatory patients waiting for MRI examination rooms per configuration.

#### *4.1.2. Recommendations*

The results indicate that the CT and MAMMO equipment of the first facility have sufficient spare capacity to accommodate the increased patient volumes, and some margin remains for additional patients. Therefore, the CT and MAMMO equipment of the second facility become redundant after the merge. As there are no CBCT examinations performed in the first facility, the utilisation of the CBCT device of the second facility remains the same after the merge with a large residual capacity.

In contrast, the US, MRI and X-ray equipment cannot keep up with the higher flow of patients. Moreover, the addition of the equipment of the second facility is insufficient to increase the available examination time for US and MRI examinations, especially during the week. Therefore, the simulation results indicate the need for an extra US device to be used during opening hours in

the week, and an MRI device with the same opening schedule as the others, in addition to the equipment the hospital currently owns. The effects of the instalment of a fifth X-ray device, on the other hand, are relatively insignificant, as the utilisation rates and waiting times are acceptable with the four already available X-ray devices.

Regarding the reception, it was clear that the staff schedule of the first facility would be insufficient to serve all patients after the merge, as the utilisation rates were already relatively high. With the addition of the receptionists already scheduled in the second facility, the reception is capable of serving all patients in a timely manner. In fact, the utilisation rates indicate a reduced workload during the weekends and for the call centre. However, the difference in utilisation rates is considered too low to reduce the schedule with one receptionist.

Finally, we analysed the required number of seats in the waiting areas for each examination room. For all rooms, two or three seats are sufficient to allow all waiting ambulatory patients to take a seat in 95% of the cases. Only in exceptional cases, this number of seats would be insufficient, but even then, patients only have to stand for a few minutes at most. However, we should note that in the simulation model, we assumed that patients arrive alone (cf. supra Section 3.4.3). As patients are often accompanied by their partner, a parent, or child(ren), the number of occupied seats would be higher in reality. Therefore, the number of required seats needs to be multiplied by the number of people usually accompanying a patient.

### *4.2. Scenario 2 – Contrast Fluid*

For Scenario 2 we continue with configuration *S1c* of Scenario 1 and evaluate the effect of omitting the requirement for some ambulatory patients to drink contrast fluid one hour before they can have their CT examination.

#### *4.2.1. Numerical Results*

The average throughput times of ambulatory patients requiring at least one CT examination are shown in Table 9. In configuration *S2*, none of the am-

bulatory patients has to wait one hour before their CT examinations(s) can be performed, which results in an average reduction of 9.08 minutes, or −22%, in throughput time for all ambulatory patients requiring at least one CT examination.

| Configuration | CF | | No CF | | All CT | |
|---|---|---|---|---|---|---|
| | Mean | ± | Mean | ± | Mean | ± |
| S1c | 76.77 | 0.24 | 32.93 | 0.24 | 41.91 | 0.28 |
| S2 | - | - | 32.85 | 0.21 | 32.85 | 0.21 |

Table 9: Average throughput times (in mins) with 95% confidence intervals of ambulatory patients requiring at least one CT examination (CF = required to drink contrast fluid and wait one hour before being examined).

Patients who currently need to drink contrast fluid have to wait this hour in the waiting area of the examination room where they will undergo their examination(s). Therefore, changing this policy also impacts the required capacity of the waiting areas, i.e. seats. However, as Table 10 shows, the impact is relatively limited in absolute terms, typically requiring one seat less.

| Configuration | Exam. room | Mean | Median | Q25 | Q75 | Q90 | Q95 |
|---|---|---|---|---|---|---|---|
| S1c | $CT_1$ | 1.39 | 1 | 1 | 2 | 3 | 3 |
| S1c | $CT_2$ | 1.15 | 1 | 0 | 2 | 2 | 3 |
| S2 | $CT_1$ | 0.94 | 1 | 0 | 1 | 2 | 2 |
| S2 | $CT_2$ | 0.85 | 1 | 0 | 1 | 2 | 2 |

Table 10: Number of seats taken by ambulatory patients waiting for CT examination rooms per configuration.

*4.2.2. Recommendations*

Even though some of the ambulatory patients had to wait a full hour less, the average reduction in the throughput time for all ambulatory patients is only 9.08 minutes – or 22% – for all ambulatory patients requiring at least one CT examination. Two elements that contribute to this explanation are: (i) only 21% of all ambulatory patients requiring a CT examination have to drink the

contrast fluid, and (ii) when these patients need multiple examinations, the other examinations are made first such that the patient loses less time due to waiting. Therefore, the required number of seats in the CT waiting areas is only marginally impacted by this change as well, typically requiring one seat less. From the previous, it follows that the decision regarding the contrast fluid policy has a fairly limited impact on the CM decisions that need to be made. Hence, this impact can be disregarded when developing the plans for the new facility.

### *4.3. Scenario 3 – Online Registration*

Scenario 3 evaluates the impact of allowing 30% of the ambulatory patients to register themselves online before they visit the hospital. This allows them to skip the reception and go straight to the waiting area of the examination room for their appointment. We compare this to configuration *S1c* of Scenario 1.

First, we maintained the staffing schedule of the reception from configuration *S1c* and evaluated the impact on utilisation rates of the reduced number of ambulatory patients that were served by the reception in configuration *S3a*. Next, we slightly adjusted the number of scheduled receptionists in configuration *S3b* to anticipate the reduced workload.

#### *4.3.1. Numerical Results*

As expected, the utilisation rates of both the front and back office, as well as the waiting time before registration, decreased in configuration *S3a*. Table 11 provides an overview of the average hourly utilisation rates per reception. In configuration *S1c*, the utilisation rate of the front office during opening hours was relatively high at 84% on average, which indicates little margin for an increased workload. When 30% of the ambulatory patients would register themselves online in configuration *S3a*, the average utilisation rate of the front office drops to 70%.

As the back office helps the front office during busy periods, a sufficiently large workload reduction for the front office will also result in a reduction of

workload for the back office, as shown in Table 11. Consequently, the average waiting time before registration decreases for all patient types, albeit most significantly for ambulatory patients, as shown in Table 12. The latter have to wait on average less than a third of the time they used to wait in configuration *S1c*.

| Configuration | Week | | Weekend | |
|---|---|---|---|---|
| | 8AM–6PM | <> | 8AM–6PM | <> |
| Reception (Front Office) | | | | |
| S1c | 0.84 | 0.69 | - | - |
| S3a | 0.70 | 0.51 | - | - |
| S3b | 0.82 | 0.80 | - | - |
| Reception (Back Office) | | | | |
| S1c | 0.70 | 0.69 | 0.60 | 0.61 |
| S3a | 0.46 | 0.58 | 0.53 | 0.47 |
| S3b | 0.68 | 0.58 | 0.69 | 0.47 |
| Reception (Back Office – Call Centre) | | | | |
| S1c | 0.25 | - | - | - |
| S3a | 0.24 | - | - | - |
| S3b | 0.25 | - | - | - |

Table 11: Average hourly utilisation rates of the reception per configuration. Regular opening hours are from 8AM–6PM, outside opening hours is denoted by "<>".

| Configuration | A | | H | | D | | S | |
|---|---|---|---|---|---|---|---|---|
| | Mean | ± | Mean | ± | Mean | ± | Mean | ± |
| Reception | | | | | | | | |
| S1a | 6.08 | 0.06 | 1.21 | 0.03 | 0.09 | 0.01 | 0.94 | 0.01 |
| S3a | 1.85 | 0.03 | 0.85 | 0.02 | 0.05 | 0.02 | 0.66 | 0.01 |
| S3b | 5.27 | 0.04 | 0.88 | 0.02 | 0.08 | 0.01 | 0.86 | 0.01 |

Table 12: Mean waiting times before registration (in mins) per configuration with 95% confidence intervals.

The reduced utilisation rates of configuration *S3a* allow adjusting the number of scheduled receptionists slightly. The front office only registers ambulatory and mobile emergency patients, while the back office can register any patient

type. Therefore, it is most apparent to reduce the number of front office receptionists first. Configuration *S3b* evaluates the impact of scheduling one instead of two front office receptionists. The utilisation rates and waiting times increase again to more or less the same levels as configuration *S1c*.

Finally, the required waiting area for the reception has to be sufficiently large to accommodate the waiting ambulatory patients. This is summarised in Table 13, e.g., in configuration *S3b*, wherein one receptionist is scheduled less, nine seats are required to make sure that in 90% of the cases everyone can take a seat, which is slightly less than the ten seats needed in configuration *S1c*.

| Configuration | Mean | Median | Q25 | Q75 | Q90 | Q95 |
|---|---|---|---|---|---|---|
| S1c | 4.15 | 3 | 1 | 6 | 10 | 13 |
| S3a | 1.77 | 1 | 1 | 2 | 4 | 5 |
| S3b | 3.55 | 2 | 1 | 5 | 9 | 11 |

Table 13: Number of seats taken by ambulatory patients while waiting in the reception.

### *4.3.2. Recommendations*

The reduction in workload is sufficient to schedule one front office receptionist instead of two. Even though the average waiting times are slightly lower with the new registration system, the utilisation rates do not allow a further reduction of receptionists while maintaining an acceptable service level. Concerning the waiting area for the reception, the required seats are comparable to the configuration where none of the ambulatory patients registers online. A median of two seats is required, and nine seats would suffice to allow everyone to take a seat in 90% of the cases. However, as mentioned earlier, this does not take into account persons accompanying the patient.

## 5. Towards a Method for Data-Driven Process Simulation for Capacity Management in Healthcare

As demonstrated in our real-life case study, a DDPS analysis generates valuable insights for healthcare managers to support decision-making in an evidence-

based way. This provides them with a solid objective foundation to better understand the operational effects when intervening in the process, instead of relying mainly on intuition and experience. Even though several research efforts that harness the power of Process Mining and Process Simulation exist in literature [26, 42–48], the methodologies used to conduct the studies are rather ad-hoc defined according to the problem at hand. Only Johnson *et al.* [42] and Kovalchuk *et al.* [45] provide an impetus to a generalised conceptual framework.

In an effort to pave the way for conducting DDPS studies to support hospital management with evidence-based CM recommendations, the remainder of this section introduces an initial conceptualisation of such a method, based on our experienced challenges during the case study and considerations from simulation literature. This framework differentiates itself from existing works such as Johnson *et al.* [42] and Kovalchuk *et al.* [45] on three essential points. Firstly, our method gives explicit consideration to data quality assessment, as well as the interaction between the use of process execution data and domain expertise. These additional perspectives – which are often overlooked in literature – are indispensable when using DDPS in a real-life healthcare context, where the assumption of having clean and high-quality event logs is unlikely to hold [13, 40]. Secondly, whereas ClearPath by Johnson *et al.* [42] is an extension of the established Process Mining project methodology $PM^2$ [62], we base our method on guidelines and best practices in the Process Simulation literature, describe explicit action points to be considered for every step, and provide references to relevant literature. Finally, we extend existing work by identifying specific challenges with respect to the operationalisation of DDPS, providing valuable directions for future research.

Sections 5.1–5.3 discuss the proposed method visualised in Figure 5. Subsequently, in Section 5.4, we provide some general reflections to move research on DDPS within a healthcare context forward. Although this paper focuses on CM problems, we believe that this framework can also be valuable for other problems in the healthcare domain due to its generic character. DDPS analysis can provide a valuable tool to evaluate policy alternatives before implementing

them in practice whenever a healthcare organisation has policy questions that would impact an operational process.

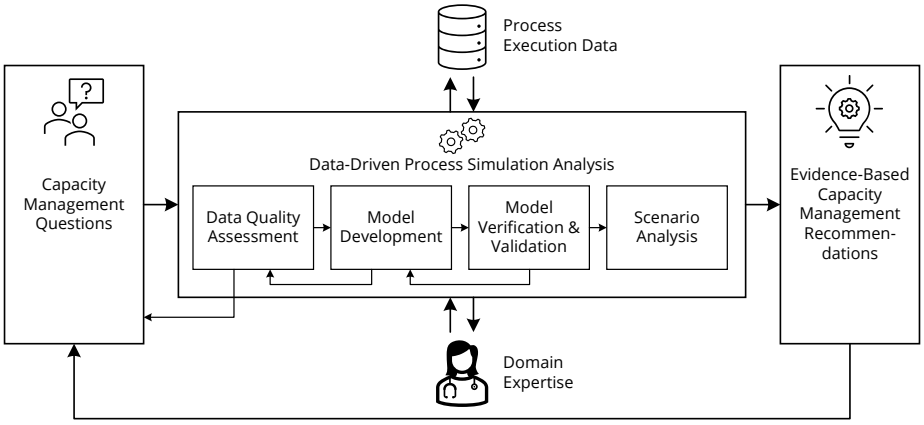


Figure 5: Data-Driven Process Simulation (DDPS) Analysis for evidence-based Capacity Management (CM) recommendations.

### *5.1. Defining the Capacity Management Questions*

Before starting the analysis, the problem at hand should be thoroughly formulated, involving the specification of the process under consideration and the policy questions at hand [67, 68, 73–75]. Accurately defining and fully understanding the prevailing *CM questions* of the healthcare organisation requires close consultation between the team responsible for conducting the analysis and the domain experts within the hospital [75, 76]. A healthcare organisation can be confronted with a wide variety of questions: "What would be the operational impact of investing in an additional medical device? What is the effect on waiting times when the number of patients with a particular profile would increase? What is the required adjustment in staffing when a change in a clinical pathway is implemented? ..." Such CM questions constitute the starting point of the DDPS analysis, and define its goals [68, 69]. Several meetings with the domain experts of the healthcare organisation might be required to formulate the CM questions [77]. In addition, the boundaries of the modelled process need to be set [65, 67, 69, 73, 78], e.g. by discussing which activities are important for

the analysis and which are out of scope given the defined CM questions. These questions also guide decisions regarding the required level of detail of the simulation model [64, 65, 67, 79]. The formulated CM questions might be strategical, tactical, operations, or a combination of these decision levels (cf. Section 2.1). Typically, strategical decisions can be modelled on a higher level of abstraction, whereas operational decisions require a high level of detail. Finally, the project planning can be made by estimating the required time to perform the DDPS analysis [65, 74, 75]. This also entails setting project milestones [65].

### *5.2. Data-Driven Process Simulation Analysis*

The *Data-Driven Process Simulation analysis* consists of four key stages: (i) data quality assessment, (ii) model development, (iii) model validation, and (iv) scenario analysis.

#### *5.2.1. Data Quality Assessment*

In the *data quality assessment* stage, the process execution data is explored and inspected for data quality issues. Because the quality of the input data has a profound impact on the reliability of the output results [40, 61], the data quality has to be assessed at the earliest possible stage. While some issues can be easily detected, e.g. negative activity durations or missing data, other problems are much harder to find without specific domain knowledge, e.g. the range within an activity duration is feasible. As the quality of the input data is typically assessed in function of the intended use, a close relationship exists between data quality assessment and the development of the model, i.e. the next stage. As the development of the model progresses, the suitability of the available process execution data for the modelling task at hand (e.g. activity durations, resource availabilities, or the arrival rate of new cases) will need to be (re-)assessed. The various tools and frameworks on data quality assessment (cf. Section 2.3) can support this stage in order to detect a variety of data quality issues which are relevant for the construction of the simulation model.

Besides detecting data quality issues, decisions need to be made on how

to handle them appropriately. This requires domain knowledge as well, which highlights, once again, the indispensable character of involving domain experts during this stage. Before reaching an acceptable quality level, several iterations of data quality assessments and corrections might be required. In addition, these newly uncovered insights from the data might trigger additional policy questions or, due to the severity of data quality issues, some policy questions might need to be reformulated. Moreover, as the development of the simulation model progresses, new issues might come to light. Therefore, several of these iterations of data quality assessment and data cleaning might be required throughout the DDPS analysis. Consequently, this stage, in which the process execution data is collected, analysed, and prepared, typically requires a significant amount of time [68, 69].

While literature on the emerging branch of DDPS in Process Mining often refers to the fully automated discovery of simulation models from event logs [10, 11, 38], the data quality issues of hospital process execution data impede this fully automated derivation, as many important details are not recorded (e.g. the exact arrival time of patients or when a particular activity started) or are inaccurately recorded (e.g. inverted timestamps, system errors, or batch registration). Notwithstanding the increasing research attention devoted to the development of taxonomies and frameworks to identify, assess, and handle data quality issues (cf. Section 2.3), there are still substantial improvements to be made to the field of DDPS by extending and integrating data quality assessment and correction algorithms into techniques and tools.

#### *5.2.2. Model Development*

In the *model development* stage, the simulation model – which represents the operational behaviour of the real-life process in a computer model – is developed [7, 64, 73, 74]. To this end, a wide variety of modelling tasks needs to be conducted, such as the specification of the control-flow (i.e. the order of activities), the arrival rate of new cases, activity durations, and resource availabilities [8]. Ideally, these model components should be modelled based on

insights from process execution data as much as possible as this data reflects the real-life execution of the process. However, as this might not always be possible due to, e.g. data quality issues, the input of domain experts is likely to be required alongside the data for some modelling tasks [75]. A structured overview of these modelling tasks and how PM can be used to support simulation model construction is provided by Martin *et al.* [8].

During the development of a simulation model, abstractions of reality will need to be made [64, 68, 74, 79]. In this respect, it is essential that all assumptions and simplifications that were made are thoroughly formulated and documented [64, 65, 67, 69, 74, 79].

While state-of-the-art DDPS model discovery [10] demonstrates the feasibility of using process execution data to construct simulation-ready models, many assumptions are still made, entailing the risk of generating oversimplified models and unrealistic behaviour in a healthcare setting. Especially the resource perspective, such as resource requirements, roles, and schedules, are indispensable when making accurate CM decisions. Current state-of-the-art DDPS model discovery offers limited support for the resource perspective [80].

#### *5.2.3. Model Verification & Validation*

The *model verification and validation* stage involves assessing whether the simulation model operates as intended and contains no errors (i.e. *verification*), and whether the model's behaviour and output represent the real-life process in a sufficiently accurate way (i.e. *validation*) [64, 67–71].

In this stage, a combination of statistical and domain expert verification and validation can be used [76]. Some output measures, such as throughput or waiting times, can typically be statistically validated using a validation dataset that has not been used during the model development phase. Other aspects of the model, especially the parts that extensively required the input of domain experts, cannot be statistically validated as limited or no data was available and have to be validated by the domain experts. This validation by domain experts can, for instance, be operationalised by means of a structured walk-through, in

which the correctness of the model is inspected by discussing it with domain experts [64, 75]. In addition, the assumptions and simplifications that were applied during model development should be discussed with and understood by the domain experts to determine whether these abstractions were valid to be made [67, 71].

If the simulation model is not reaching the desired accuracy level, either statistically or deemed by the domain experts, the modeller needs to return to the model development stage to improve the simulation model [64, 71, 76]. Several iterations of verification and validation, and simulation model corrections might be required to attain the desired accuracy level.

When modelling complex operational systems, such as healthcare processes, domain expertise is indispensable to give meaning to patterns arising from the data and provide additional information regarding the process. Simultaneously, specific expertise is required when conducting DDPS [64, 74], which is not widespread in many hospitals. Hence, domain experts need to rely on simulation modellers to build the simulation model and conduct the analysis. While this third party has the required competencies to conduct the DDPS, they typically do not have the same expertise about the process under study. This creates an interdependency between the simulation modellers and the domain experts. Considering that the interaction between the simulation modellers and domain experts tends to occur asynchronously, the project duration is often significantly extended, particularly when several iterations are needed, e.g. to handle data quality issues or validate the model.

#### *5.2.4. Scenario Analysis*

Fourthly and finally, once a validated model is obtained, various *scenarios* are defined in an effort to answer the CM questions under investigation. Each scenario describes a potential future situation that will be simulated under a specific set of conditions [64, 67, 69, 75], e.g. the addition of a new medical device, an increase of the number of patients with a particular profile, or an adjustment in scheduled staff. The scenarios should also be explained to and

discussed with the domain experts to determine whether these scenarios are suitable for determining answers to the CM questions [67].

In addition to the scenarios, the run parameters of the model need to be specified. These include the length of each simulation run, in terms of time units in the simulation model [64, 65, 67, 68, 74, 78, 79], the warm-up period needed to attain a stable state in the model [64, 65, 67, 74, 79], and the number of replications needed to reduce the variability of the outputs [64, 65, 67, 74].

Each scenario is then simulated, which generates output. Typically, a statistical output analysis is performed to analyse the performance metrics of each scenario [64, 65, 68, 69, 73, 74, 78].

### *5.3. Evidence-Based Capacity Management Recommendations*

The results from the scenario analysis are used to formulate CM recommendations as decision-support information for the hospital management [75]. As shown in Figure 5, these recommendations can spawn additional CM questions, which could instigate another iteration through the DDPS method.

The scenario analysis results should be presented to the domain experts, along with evidence-based recommendations and suggestions for improvement regarding the defined CM questions [64, 67, 73, 75, 77, 79]. The actual decision-making and implementation of the evidence-based CM recommendations is the competence of the domain experts and, therefore, outside the scope of the proposed method [75].

### *5.4. Reflections on Data-Driven Process Simulation Analysis*

DDPS analysis is a promising way to support healthcare managers in their CM decisions. Nevertheless, the aforementioned challenges – (i) data quality issues, (ii) limited support for the resource perspective, and (iii) asynchronous interaction – highlight the need for further research to develop tools enabling healthcare managers to reap the benefits of DDPS analysis fully. An interesting perspective to explore in future work is *interactive Data-Driven Process Simulation* [61]. This implies that the modelling, validation, and analysis

in DDPS is conducted *synchronously* as the domain expert directly interacts with a simulation tool, guiding the expert during the data-driven development and validation of a simulation model. Only limited research has been devoted to interactively involving domain experts in developing data-driven simulation models. Martin *et al.* [60] proposed, at a conceptual level, an interactive data cleaning approach to identify and correct data quality issues in event logs.

Similarly, Benevento *et al.* [81] proposed an Interactive Process Discovery technique in which the modeller has complete control over discovering a process model from an event log. However, the resulting process models are Petri Nets, which are typically rather difficult to understand for domain experts without prior knowledge about this modelling technique [82, 83]. Even though algorithms exist to convert Petri Nets to more comprehensible notations, such as Business Process Model and Notation (BPMN) [84, 85], these algorithms often create large and complex models [84, 85].

The interactive involvement of domain experts during DDPS analyses will require the development of new tools which operationalise an interactive approach throughout the various stages of a simulation study. In the first stage, tools could be developed that support enhanced interaction between modeller and domain expert, e.g., facilitating joint data quality assessment and data cleaning. This also requires additional research on interactively involving domain experts at a conceptual level [86], i.e. best practices when visualising statistics for validation to domain experts and eliciting input to build, correct, and enhance data-driven simulation models. In parallel, data quality can be improved by facilitating accurate data registration using intuitive and straightforward information system interfaces or by using technologies that support data capturing, such as Real-Time Location Systems (RTLS), which can enrich and validate event logs [48, 87–89]. In a second stage, such tools could be refined and combined into an integrated simulation environment that allows domain experts to conduct DDPS analyses independently.

## 6. Conclusion

Data-Driven Process Simulation focuses on developing simulation models by extensively using process execution data – i.e. event logs – originating from, e.g. Health Information Systems. This paper outlined a DDPS analysis to support Capacity Management decisions using an extensive real-life case study at the radiology department of a Belgian hospital. In this case study, we provided recommendations regarding the required number of radiology devices, waiting area size, and reception staffing.

Based on our experienced challenges during the case study, and guidelines and best practices from simulation literature, we provide an effort to pave the way for conducting DDPS studies to support hospital management with evidence-based CM recommendations. This method attributes explicit consideration to data quality assessment, as well as the interaction between the use of process execution data and domain expertise, two elements that are often overlooked in literature.

As highlighted in the discussion, future research could focus on the interactive involvement of domain experts during the different steps of a DDPS analysis, the assessment and correction of data quality issues encountered in healthcare process execution data, and extending the support for the resource perspective in DDPS model discovery. In addition, the applicability of the introduced method for other CM decisions and other types of policy questions in healthcare processes can be assessed.

*Funding.* This study was supported by the Special Research Fund (BOF) of Hasselt University under Grant No. BOF19OWB20.

## References

[1] C. Hicks, T. McGovern, G. Prior, I. Smith, Applying Lean Principles to the Design of Healthcare Facilities, International Journal of Production Economics 170 (2015) 677–686. doi:10.1016/j.ijpe.2015.05.029.

[2] D. McCaughey, C. O. Erwin, J. L. DelliFraine, Improving Capacity Management in the Emergency Department: A Review of the Literature, 2000-2012, Journal of Healthcare Management 60 (1) (2015) 63. doi:10.1097/00115514-201501000-00011.

[3] V. L. Smith-Daniels, S. B. Schweikhart, D. E. Smith-Daniels, Capacity Management in Health Care Services: Review and Future Research Directions, Decision Sciences 19 (4) (1988) 889–919. doi:10.1111/j.1540-5915.1988.tb00310.x.

[4] E. W. Hans, M. Van Houdenhoven, P. J. H. Hulshof, A Framework for Healthcare Planning and Control, in: R. W. Hall, F. S. Hillier (Eds.), Handbook of Healthcare System Scheduling, Vol. 168 of International Series in Operations Research & Management Science, Springer, Boston, MA, USA, 2012, pp. 303–320. doi:10.1007/978-1-4614-1734-7_12.

[5] D. B. McLaughlin, J. M. Hays, Healthcare Operations Management, Health Administration Press, Chicago, IL, USA, 2008.

[6] A. Kolker, Interdependency of Hospital Departments and Hospital-Wide Patient Flows, in: R. Hall, F. S. Hillier (Eds.), Patient Flow: Reducing Delay in Healthcare Delivery, 2nd Edition, Vol. 206 of International Series in Operations Research & Management Science, Springer US, Boston, MA, USA, 2013, pp. 43–63. doi:10.1007/978-1-4614-9512-3_2.

[7] N. Melão, M. Pidd, Use of Business Process Simulation: A Survey of Practitioners, Journal of the Operational Research Society 54 (1) (2003) 2–10. doi:10.1057/palgrave.jors.2601477.

[8] N. Martin, B. Depaire, A. Caris, The Use of Process Mining in Business Process Simulation Model Construction, Business & Information Systems Engineering 58 (1) (2016) 73–87. doi:10.1007/s12599-015-0410-4.

[9] B. Depaire, N. Martin, Data-Driven Process Simulation, Encyclopedia of Big Data Technologiesdoi:10.1007/978-3-319-63962-8_102-1.

[10] M. Camargo, M. Dumas, O. González-Rojas, Automated Discovery of Business Process Simulation Models from Event Logs, Decision Support Systems 134 (2020) 113284. doi:10.1016/j.dss.2020.113284.

[11] B. Estrada-Torres, M. Camargo, M. Dumas, L. García-Bañuelos, I. Mahdy, M. Yerokhin, Discovering Business Process Simulation Models in the Presence of Multitasking and Availability Constraints, Data & Knowledge Engineering 134 (2021) 101897. doi:10.1016/j.datak.2021.101897.

[12] R. S. Mans, W. M. P. van der Aalst, R. J. B. Vanwersch, Process Mining in Healthcare: Evaluating and Exploiting Operational Healthcare Processes, no. 13170 in SpringerBriefs in Business Process Management, Springer International Publishing, Cham, Switzerland, 2015. doi:10.1007/978-3-319-16071-9.

[13] N. Martin, Data Quality in Process Mining, in: C. Fernández-Llatas (Ed.), Interactive Process Mining in Healthcare, no. 1114 in Health Informatics, Springer International Publishing, Cham, Switzerland, 2021, pp. 53–79. doi:10.1007/978-3-030-53993-1_5.

[14] N. Martin, B. Benoît, A. Caris, Event Log Knowledge as a Complementary Simulation Model Construction Input, in: M. S. Obaidat, J. Kacprzyk, T. I. Ören (Eds.), Proceedings of the 4th International Conference On Simulation And Modeling Methodologies, Technologies And Applications, Vol. 1 of SIMULTECH '14, SciTePress, Vienna, Austria, 2014, pp. 456–462. doi:10.5220/0005100404560462.

[15] P. J. H. Hulshof, N. Kortbeek, R. J. Boucherie, E. W. Hans, P. J. M. Bakker, Taxonomic Classification of Planning Decisions in Health Care: A Structured Review of the State of the Art in OR/MS, Health Systems 1 (2) (2012) 129–175. doi:10.1057/hs.2012.18.

[16] M. Shakoor, Using Discrete Event Simulation Approach to Reduce Waiting Times in Computed Tomography Radiology Department, International

Journal of Industrial and Manufacturing Engineering 9 (1) (2015) 177–181. doi:10.5281/zenodo.1338044.

[17] B. Vieira, E. W. Hans, C. van Vliet-Vroegindeweij, J. van de Kamer, W. van Harten, Operations Research for Resource Planning and -Use in Radiotherapy: A Literature Review, BMC Medical Informatics and Decision Making 16 (1) (2016) 149. doi:10.1186/s12911-016-0390-4.

[18] H. Cai, J. Jia, Using Discrete Event Simulation (DES) To Support Performance-Driven Healthcare Design, HERD: Health Environments Research & Design Journal 12 (3) (2019) 89–106. doi:10.1177/1937586718801910.

[19] S. T. Ponis, A. Delis, S. P. Gayialis, P. Kasimatis, J. Tan, Applying Discrete Event Simulation (DES) in Healthcare: The Case for Outpatient Facility Capacity Planning, International Journal of Healthcare Information Systems and Informatics 8 (3) (2013) 58–79. doi:10.4018/jhisi.2013070104.

[20] R. Carmen, M. Defraeye, I. Van Nieuwenhuyse, A Decision Support System for Capacity Planning in Emergency Departments, International Journal of Simulation Modelling 14 (2) (2015) 299–312. doi:10.2507/ijsimm14(2)10.308.

[21] I. M. Ibrahim, C.-Y. Liong, S. A. Bakar, N. Ahmad, A. F. Najmuddin, Minimizing Patient Waiting Time in Emergency Department of Public Hospital Using Simulation Optimization Approach, in: Z. H. Zamzuri (Ed.), Proceedings of the 4th International Conference on Mathematical Sciences, Vol. 1830 of AIP Conference Proceedings, AIP Publishing, Putrajaya, Malaysia, 2017, pp. 060005–1–8. doi:10.1063/1.4980949.

[22] B. B. Saleh, G. B. Saleh, O. Barakat, Operating Theater Management System: Block-Scheduling, in: M. Masmoudi, B. Jarboui, P. Siarry (Eds.),

Artificial Intelligence and Data Mining in Healthcare, Springer International Publishing, Cham, Switzerland, 2021, pp. 83–98. doi:10.1007/978-3-030-45240-7_5.

[23] A. Legrain, M.-A. Fortin, N. Lahrichi, L.-M. Rousseau, Online Stochastic Optimization of Radiotherapy Patient Scheduling, Health Care Management Science 18 (2) (2015) 110–123. doi:10.1007/s10729-014-9270-6.

[24] D. Petrovic, E. Castro, S. Petrovic, T. Kapamara, Radiotherapy Scheduling, in: A. Ş. Etaner-Uyar, E. Özcan, N. Urquhart, J. Kacprzyk (Eds.), Automated Scheduling and Planning: From Theory to Practice, Vol. 505 of Studies in Computational Intelligence, Springer, Berlin, Germany, 2013, pp. 155–189. doi:10.1007/978-3-642-39304-4_7.

[25] A. Ganguly, S. Nandi, Using Statistical Forecasting to Optimize Staff Scheduling in Healthcare Organizations, Journal of Health Management 18 (1) (2016) 172–181. doi:10.1177/0972063415625575.

[26] B. B. P. Antunes, A. Manresa, L. S. L. Bastos, J. F. Marchesi, S. Hamacher, A Solution Framework Based on Process Mining, Optimization, and Discrete-Event Simulation to Improve Queue Performance in an Emergency Department, in: C. Di Francescomarino, R. M. Dijkman, U. Zdun, W. M. P. van der Aalst, J. Mylopoulos, M. Rosemann, M. J. Shaw, C. Szyperski (Eds.), Proceedings of the Business Process Management Workshops, Vol. 362 of Lecture Notes in Business Information Processing, Springer International Publishing, Vienna, Austria, 2019, pp. 583–594. doi:10.1007/978-3-030-37453-2_47.

[27] S. Saedi, O. E. Kundakcioglu, A. C. Henry, Mitigating the Impact of Drug Shortages for a Healthcare Facility: An Inventory Management Approach, European Journal of Operational Research 251 (1) (2016) 107–123. doi:10.1016/j.ejor.2015.11.017.

[28] D. Daldoul, I. Nouaouri, H. Bouchriha, H. Allaoui, Scheduling Patients in Emergency Department: A Case Study, in: Proceedings of the 2017

IEEE International Conference on Industrial Engineering and Engineering Management, IEEM '17, IEEE, Singapore, Singapore, 2017, pp. 870–874. doi:10.1109/IEEM.2017.8290016.

[29] K. Katsaliaki, N. Mustafee, Applications of Simulation Within the Healthcare Context, Journal of the Operational Research Society 62 (8) (2011) 1431–1451. doi:10.1057/jors.2010.20.

[30] S. C. Brailsford, Tutorial: Advances and Challenges in Healthcare Simulation Modeling, in: S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, R. R. Barton (Eds.), Proceedings of the 2007 Winter Simulation Conference, IEEE, Washington, DC, USA, 2007, pp. 1436–1448. doi:10.1109/WSC.2007.4419754.

[31] S. Robinson, Discrete-Event Simulation: From the Pioneers to the Present, What Next?, Journal of the Operational Research Society 56 (6) (2005) 619–629. doi:10.1057/palgrave.jors.2601864.

[32] J. B. Jun, S. H. Jacobson, J. R. Swisher, Application of Discrete-Event Simulation in Health Care Clinics: A Survey, Journal of the Operational Research Society 50 (2) (1999) 109–123. doi:10.1057/palgrave.jors.2600669.

[33] S. Salleh, P. Thokala, A. Brennan, R. Hughes, A. Booth, Simulation Modelling in Healthcare: An Umbrella Review of Systematic Literature Reviews, PharmacoEconomics 35 (9) (2017) 937–949. doi:10.1007/s40273-017-0523-3.

[34] M. M. Günal, M. Pidd, Discrete Event Simulation for Performance Modelling in Health Care: A Review of the Literature, Journal of Simulation 4 (1) (2010) 42–51. doi:10.1057/jos.2009.25.

[35] S. C. Brailsford, P. R. Harper, B. Patel, M. Pitt, An Analysis of the Academic Literature on Simulation and Modeling in Health Care, Journal of Simulation 3 (2009) 130–140. doi:10.1057/jos.2009.10.

[36] X. Zhang, Application of Discrete Event Simulation in Health Care: A Systematic Review, BMC Health Services Research 18 (1) (2018) 687. doi:10.1186/s12913-018-3456-4.

[37] W. M. P. van der Aalst, Process Mining: Data Science in Action, 2nd Edition, Springer, Berlin, Germany, 2016. doi:10.1007/978-3-662-49851-4.

[38] A. Rozinat, R. S. Mans, M. Song, W. M. P. van der Aalst, Discovering Simulation Models, Information Systems 34 (3) (2009) 305–327. doi:10.1016/j.is.2008.09.002.

[39] I. Khodyrev, S. Popova, Discrete Modeling and Simulation of Business Processes Using Event Logs, in: D. Abramson, M. Lees, V. Krzhizhanovskaya, J. Dongarra, P. M. A. Sloot (Eds.), Proceedings of the 14th International Conference on Computational Science, Vol. 29 of Procedia Computer Science, Elsevier, Cairns, QLD, Australia, 2014, pp. 322–331. doi:10.1016/j.procs.2014.05.029.

[40] L. Vanbrabant, N. Martin, K. Ramaekers, K. Braekers, Quality of Input Data in Emergency Department Simulations: Framework and Assessment Techniques, Simulation Modelling Practice and Theory 91 (2019) 83–101. doi:10.1016/j.simpat.2018.12.002.

[41] C. Di Ciccio, A. Marrella, A. Russo, Knowledge-Intensive Processes: Characteristics, Requirements and Analysis of Contemporary Approaches, Journal on Data Semantics 4 (1) (2015) 29–57. doi:10.1007/s13740-014-0038-4.

[42] O. A. Johnson, T. Ba Dhafari, A. Kurniati, F. Fox, E. Rojas, The ClearPath Method for Care Pathway Process Mining and Simulation, in: F. Daniel, Q. Z. Sheng, H. Motahari (Eds.), Proceedings of the Business Process Management International Workshops, Vol. 342 of Lecture Notes in Business Information Processing, Springer International Publishing, Sydney, NSW, Australia, 2019, pp. 239–250. doi:10.1007/978-3-030-11641-5_19.

[43] V. Augusto, X. Xie, M. Prodel, B. Jouaneton, L. Lamarsalle, Evaluation of Discovered Clinical Pathways Using Process Mining and Joint Agent-Based Discrete-Event Simulation, in: T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. R. Huschka, S. E. Chick (Eds.), Proceedings of the 2016 Winter Simulation Conference, WSC '16, IEEE, Washington, DC, USA, 2016, pp. 2135–2146. doi:10.1109/WSC.2016.7822256.

[44] M. Mans, Ronny S., H. A. Reijers, D. Wismeijer, M. van Genuchten, A Process-Oriented Methodology for Evaluating the Impact of IT: A Proposal and an Application in Healthcare, Information Systems 38 (8) (2013) 1097–1115. doi:10.1016/j.is.2013.06.005.

[45] S. V. Kovalchuk, A. A. Funkner, O. G. Metsker, A. N. Yakovlev, Simulation of Patient Flow in Multiple Healthcare Units Using Process and Data Mining Techniques for Model Identification, Journal of Biomedical Informatics 82 (2018) 128–142. doi:10.1016/j.jbi.2018.05.004.

[46] O. Tamburis, C. Esposito, Process Mining as Support to Simulation Modeling: A Hospital-Based Case Study, Simulation Modelling Practice and Theory 104 (2020) 102149. doi:10.1016/j.simpat.2020.102149.

[47] Z. Zhou, Y. Wang, L. Li, Process Mining Based Modeling and Analysis of Workflows in Clinical Care: A Case Study in a Chicago Outpatient Clinic, in: M. Pan, W. Wu (Eds.), Proceedings of the 11th IEEE International Conference on Networking, Sensing and Control, ICNSC '14, IEEE, Miami, FL, USA, 2014, pp. 590–595. doi:10.1109/icnsc.2014.6819692.

[48] W. Abohamad, A. Ramy, A. Arisha, A Hybrid Process-Mining Approach for Simulation Modeling, in: W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, E. H. Page (Eds.), Proceedings of the 2017 Winter Simulation Conference, WSC '17, IEEE, Las Vegas, NV, USA, 2017, pp. 1527–1538. doi:10.1109/WSC.2017.8247894.

[49] R. P. J. C. Bose, R. S. Mans, W. M. P. van der Aalst, Wanna Improve Process Mining Results? It's High Time We Consider Data Quality Issues

Seriously, in: B. Hammer, Z.-H. Zhou, L. Wang, N. Chawla (Eds.), Proceedings of the 2013 IEEE Symposium on Computational Intelligence and Data Mining, CIDM'13, IEEE, Singapore, Singapore, 2013, pp. 127–134. doi:10.1109/CIDM.2013.6597227.

[50] S. Suriadi, R. Andrews, A. H. M. ter Hofstede, M. T. Wynn, Event Log Imperfection Patterns for Process Mining: Towards a Systematic Approach to Cleaning Event Logs, Information Systems 64 (2017) 132–150. doi:10.1016/j.is.2016.07.011.

[51] R. Andrews, M. T. Wynn, K. Vallmuur, A. H. M. ter Hofstede, E. Bosley, M. Elcock, S. Rashford, Leveraging Data Quality to Better Prepare for Process Mining: An Approach Illustrated Through Analysing Road Trauma Pre-Hospital Retrieval and Transport Processes in Queensland, International Journal of Environmental Research and Public Health 16 (7) (2019) 1138. doi:10.3390/ijerph16071138.

[52] M. O. Kherbouche, N. Laga, P.-A. Masse, Towards a Better Assessment of Event Logs Quality, in: Y. Jin, S. Kollias (Eds.), Proceedings of the 2016 IEEE Symposium Series on Computational Intelligence, SSCI '16, IEEE, Athens, Greece, 2016, pp. 1–8. doi:10.1109/SSCI.2016.7849946.

[53] D. A. Fischer, K. Goel, R. Andrews, C. G. J. van Dun, M. T. Wynn, M. Röglinger, Enhancing Event Log Quality: Detecting and Quantifying Timestamp Imperfections, in: D. Fahland, C. Ghidini, J. Becker, M. Dumas (Eds.), Proceedings of the 18th International Conference on Business Process Management, Vol. 12168 of Lecture Notes in Computer Science, Springer International Publishing, Seville, Spain, 2020, pp. 309–326. doi:10.1007/978-3-030-58666-9_18.

[54] P. M. Dixit, S. Suriadi, R. Andrews, M. T. Wynn, A. H. M. ter Hofstede, J. C. A. M. Buijs, W. M. P. van der Aalst, Detection and Interactive Repair of Event Ordering Imperfection in Process Logs, in: J. Krogstie, H. A. Reijers (Eds.), Proceedings of the 30th International Conference on

Advanced Information Systems Engineering, Vol. 10816 of Lecture Notes in Computer Science, Springer International Publishing, Tallinn, Estonia, 2018, pp. 274–290. doi:10.1007/978-3-319-91563-0_17.

[55] F. Fox, V. R. Aggarwal, H. Whelton, O. A. Johnson, A Data Quality Framework for Process Mining of Electronic Health Record Data, in: J. Pathak, F. Wang, R. Moskovitch, J. Ancker, C. Weng (Eds.), Proceedings of the 2018 IEEE International Conference on Healthcare Informatics, ICHI '18, IEEE, New York, NY, USA, 2018, pp. 12–21. doi:10.1109/ICHI.2018.00009.

[56] R. Andrews, S. Suriadi, C. Ouyang, E. Poppe, Towards Event Log Querying for Data Quality, in: H. Panetto, C. Debruyne, H. A. Proper, C. A. Ardagna, D. Roman, R. Meersman (Eds.), On the Move to Meaningful Internet Systems, Vol. 11229 of Lecture Notes in Computer Science, Springer International Publishing, Valletta, Malta, 2018, pp. 116–134. doi:10.1007/978-3-030-02610-3_7.

[57] N. Martin, G. Van Houdt, G. Janssenswillen, DaQAPO: Supporting Flexible and Fine-Grained Event Log Quality Assessment, Expert Systems with Applications 191 (2022) 116274. doi:10.1016/j.eswa.2021.116274.

[58] D. Bayomie, A. Awad, E. Ezat, Correlating Unlabeled Events from Cyclic Business Processes Execution, in: S. Nurcan, P. Soffer, M. Bajec, J. Eder (Eds.), Proceedings of the 28th International Conference on Advanced Information Systems Engineering, Vol. 9694 of Lecture Notes in Computer Science, Springer International Publishing, Ljubljana, Slovenia, 2016, pp. 274–289. doi:10.1007/978-3-319-39696-5_17.

[59] C. Di Francescomarino, C. Ghidini, S. Tessaris, I. V. Sandoval, Completing Workflow Traces Using Action Languages, in: J. Zdravkovic, M. Kirikova, P. Johannesson (Eds.), Proceedings of the 27th International Conference on Advanced Information Systems Engineering, Vol. 9097 of Lecture Notes in

Computer Science, Springer International Publishing, Stockholm, Sweden, 2015, pp. 314–330. doi:10.1007/978-3-319-19069-3_20.

[60] N. Martin, A. Martinez-Millana, B. Valdivieso, C. Fernández-Llatas, Interactive Data Cleaning for Process Mining: A Case Study of an Outpatient Clinic's Appointment System, in: C. Di Francescomarino, R. M. Dijkman, U. Zdun, W. M. P. van der Aalst, J. Mylopoulos, M. Rosemann, M. J. Shaw, C. Szyperski (Eds.), Proceedings of the Business Process Management International Workshops, Vol. 362 of Lecture Notes in Business Information Processing, Springer International Publishing, Vienna, Austria, 2019, pp. 532–544. doi:10.1007/978-3-030-37453-2_43.

[61] G. van Hulzen, N. Martin, B. Depaire, The Need for Interactive Data-Driven Process Simulation in Healthcare: A Case Study, in: S. J. J. Leemans, H. Leopold (Eds.), Proceedings of the ICPM 2020 International Workshops, Vol. 406 of Lecture Notes in Business Information Processing, Springer, Padua, Italy, 2021, pp. 317–329. doi:10.1007/978-3-030-72693-5_24.

[62] M. L. van Eck, X. Lu, S. J. J. Leemans, W. M. P. van der Aalst, PM$^2$: A Process Mining Project Methodology, in: J. Zdravkovic, M. Kirikova, P. Johannesson (Eds.), Proceedings of the 27th International Conference on Advanced Information Systems Engineering, Vol. 9097 of Lecture Notes in Computer Science, Springer International Publishing, Stockholm, Sweden, 2015, pp. 297–313. doi:10.1007/978-3-319-19069-3_19.

[63] V. Chandola, A. Banerjee, V. Kumar, Anomaly Detection: A Survey, ACM Computing Surveys 41 (3) (2009) 15:1–15:58. doi:10.1145/1541880.1541882.

[64] A. M. Law, Simulation Modeling and Analysis, 5th Edition, McGraw-Hill Education, New York, NY, USA, 2014.

[65] W. D. Kelton, R. P. Sadowski, N. B. Zupick, Simulation with Arena, sixth Edition, McGraw-Hill Education, New York, NY, USA, 2015.

[66] Arena, Rockwell Automation Technologies, Inc. (2020).
URL https://www.arenasimulation.com/

[67] J. S. Carson II, Introduction to Modeling and Simulation, in: M. E. Kuhl, N. M. Steiger, F. B. Armstrong, J. A. Joines (Eds.), Proceedings of the 2005 Winter Simulation Conference, WSC '05, IEEE, Orlando, FL, USA, 2005, pp. 16–23. doi:10.1109/WSC.2005.1574235.

[68] R. E. Shannon, Introduction to the Art and Science of Simulation, in: D. J. Medeiros, E. F. Watson, J. S. Carson II, M. S. Manivannan (Eds.), Proceedings of the 1998 Winter Simulation Conference, Vol. 1 of WSC '98, IEEE, Washington, DC, USA, 1998, pp. 7–14. doi:10.1109/WSC.1998.744892.

[69] W. B. Nordgren, Steps for Proper Simulation Project Management, in: C. Alexopoulos, K. Kang, W. R. Lilegdon, D. Goldsman (Eds.), Proceedings of the 1995 Winter Simulation Conference, WSC '95, IEEE Computer Society, Arlington, VA, USA, 1995, pp. 68–73. doi:10.1145/224401.224425.

[70] J. Banks, J. S. Carson II, B. L. Nelson, D. M. Nicol, Discrete-Event System Simulation, 5th Edition, Pearson Education Ltd., Edinburgh Gate, Harlow, Essex, UK, 2014.

[71] R. G. Sargent, Verification and Validation of Simulation Models, Journal of Simulation 7 (1) (2013) 12–24. doi:10.1057/jos.2012.20.

[72] M. Neuhäuser, Wilcoxon–Mann–Whitney Test, International Encyclopedia of Statistical Science (2011) 1656–1658doi:10.1007/978-3-642-04898-2_615.

[73] V. Hlupic, S. Robinson, Business Process Modelling and Analysis Using Discrete-Event Simulation, in: D. J. Medeiros, E. F. Watson, J. S. Carson II, M. S. Manivannan (Eds.), Proceedings of the 1998 Winter Simula-

tion Conference, Vol. 2 of WSC '98, IEEE, Washington, DC, USA, 1998, pp. 1363–1369. doi:10.1109/WSC.1998.746003.

[74] S. Robinson, Simulation: The Practice of Model Development and Use, John Wiley & Sons, Ltd., Chichester, West Sussex, England, 2004.

[75] N. Martin, B. Depaire, A. Caris, A Synthesized Method for Conducting a Business Process Simulation Study, in: M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, B. J. I. Johansson (Eds.), Proceedings of the 2018 Winter Simulation Conference, WSC '18, IEEE, Gothenburg, Sweden, 2018, pp. 276–290. doi:10.1109/WSC.2018.8632298.

[76] M. D. Rossetti, Simulation Modeling and Arena, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2010.

[77] O. Balci, Guidelines for Successful Simulation Studies, in: O. Balci, R. P. Sadowski, R. E. Nance (Eds.), Proceedings of the 1990 Winter Simulation Conference, WSC '90, IEEE, New Orleans, LA, USA, 1990, pp. 25–32. doi:10.1109/WSC.1990.129482.

[78] W. M. P. van der Aalst, Business Process Simulation Survival Guide, in: J. vom Brocke, M. Rosemann (Eds.), Handbook on Business Process Management: Introduction, Methods, and Information Systems 1, 2nd Edition, no. 3795 in International Handbooks on Information Systems, Springer-Verlag, Berlin, Germany, 2015, pp. 337–370. doi:10.1007/978-3-642-45100-3_15.

[79] A. M. Law, How to Build Valid and Credible Simulation Models, in: N. Mustafee, K.-H. G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. J. Haas, Y.-J. Son (Eds.), Proceedings of the 2019 Winter Simulation Conference, WSC '19, IEEE, National Harbor, MD, USA, 2019, pp. 1402–1414. doi:10.1109/WSC40007.2019.9004789.

[80] G. van Hulzen, Interactive Data-Driven Business Process Simulation (Extended Abstract), in: C. Di Ciccio, B. Depaire, J. De Weerdt,

C. Di Francescomarino, J. Munoz-Gama (Eds.), Proceedings of the ICPM 2020 Doctoral Consortium and Tool Demonstration Track, Vol. 2703 of ICPM'20, CEUR Workshop Proceedings, Padua, Italy, 2020, pp. 7–8.

[81] E. Benevento, P. M. Dixit, M. F. Sani, D. Aloini, W. M. P. van der Aalst, Evaluating the Effectiveness of Interactive Process Discovery in Healthcare: A Case Study, in: C. Di Francescomarino, R. M. Dijkman, U. Zdun, W. M. P. van der Aalst, J. Mylopoulos, M. Rosemann, M. J. Shaw, C. Szyperski (Eds.), Proceedings of the Business Process Management International Workshops, Vol. 362 of Lecture Notes in Business Information Processing, Springer International Publishing, Vienna, Austria, 2019, pp. 508–519. doi:10.1007/978-3-030-37453-2_41.

[82] S. Aguirre, C. Parra, J. Alvarado, Combination of Process Mining and Simulation Techniques for Business Process Redesign: A Methodological Approach, in: P. Cudre-Mauroux, P. Ceravolo, D. Gašević (Eds.), Proceedings of the 2nd International Symposium on Data-Driven Process Discovery and Analysis, Vol. 162 of Lecture Notes in Business Information Processing, Springer, Campione d'Italia, Italy, 2013, pp. 24–43. doi:10.1007/978-3-642-40919-6_2.

[83] K. Rosenthal, B. Ternes, S. Strecker, Business Process Simulation on Procedural Graphical Process Models, Business & Information Systems Engineeringdoi:10.1007/s12599-021-00690-3.

[84] C. Favre, D. Fahland, H. Völzer, The Relationship Between Workflow Graphs and Free-Choice Workflow Nets, Information Systems 47 (2015) 197–219. doi:10.1016/j.is.2013.12.004.

[85] A. A. Kalenkova, W. M. P. van der Aalst, I. A. Lomazova, V. A. Rubin, Process Mining Using BPMN: Relating Event Logs and Process Models, Software & Systems Modeling 16 (4) (2017) 1019–1048. doi:10.1007/s10270-015-0502-0.

[86] C. Fernández-Llatas (Ed.), Interactive Process Mining in Healthcare, no. 1114 in Health Informatics, Springer International Publishing, Cham, Switzerland, 2021. doi:10.1007/978-3-030-53993-1.

[87] R. Miclo, F. Fontanili, G. Marquès, P. Bomert, M. Lauras, RTLS-Based Process Mining: Towards an Automatic Process Diagnosis in Healthcare, in: Proceedings of the 2015 IEEE International Conference on Automation Science and Engineering, CASE '15, IEEE, Gothenburg, Sweden, 2015, pp. 1397–1402. doi:10.1109/CoASE.2015.7294294.

[88] S. N. Araghi, F. Fontanili, E. Lamine, L. Tancerel, F. Benaben, Monitoring and Analyzing Patients' Pathways by the Application of Process Mining, SPC, and I-RTLS, IFAC-PapersOnLine 51 (11) (2018) 980–985. doi:10.1016/j.ifacol.2018.08.480.

[89] C. Fernández-Llatas, J.-M. Benedí Ruíz, J. M. Gama, M. Sepúlveda, E. Rojas, S. Vera, V. Traver, Interactive Process Mining in Surgery with Real Time Location Systems: Interactive Trace Correction, in: C. Fernández-Llatas (Ed.), Interactive Process Mining in Healthcare, no. 1114 in Health Informatics, Springer International Publishing, Cham, Switzerland, 2021, pp. 181–202. doi:10.1007/978-3-030-53993-1_11.