# The Need for Interactive Data-Driven Process Simulation in Healthcare: A Case Study

Gerhardus van Hulzen[1(✉)] , Niels Martin[1,2,3] , and Benoît Depaire[1]

[1] Research group Business Informatics, Hasselt University, 3500 Hasselt, Belgium
{gerard.vanhulzen,niels.martin,benoit.depaire}@uhasselt.be
[2] Research Foundation Flanders (FWO), 1000 Brussels, Belgium
[3] Data Analytics Laboratory, Vrije Universiteit Brussel, 1050 Brussels, Belgium

**Abstract.** In healthcare, more and more process execution information is stored in Hospital Information Systems. This data, in conjunction with data-driven process simulation, can be used, e.g. to support hospital management with Capacity Management decisions. However, real-life event logs in healthcare often suffer from data quality issues, affecting the reliability of simulation results. In this work, we illustrate the effects of disregarding data quality issues on simulation outcomes and the importance of domain knowledge using a case study at the radiology department of a hospital. Current literature on data-driven process simulation acknowledges the need for domain expertise but does not provide a framework for conceptualising the involvement of domain experts. Therefore, we propose a novel conceptual framework which interactively involves experts during data-driven simulation model development.

**Keywords:** Data-driven process simulation · Data quality · Domain knowledge · Interactive modelling · Healthcare processes

## 1 Introduction

Worldwide, healthcare systems are under constant pressure. Increasing population numbers, lifestyle factors, ageing populations, and new technologies are the main drivers for increasing healthcare expenses. Simultaneously, healthcare budgets are under pressure due to national budget deficits and savings [14]. Healthcare managers have to improve their care processes to maintain high-quality care for all patients. One key aspect of ensuring this is efficient Capacity Management (CM), which is used to determine the suitable levels of resources, such as equipment, facilities, and staff size [28].

To support hospital management during CM decisions, *Business Process Simulation (BPS)* can be used to determine suitable resource levels objectively. BPS uses a (computer) model to imitate the process. This allows to evaluate the effect of various process modifications without actually implementing them into, nor

disrupting, the real process [21]. For instance, the effect on throughput rates and patient waiting times of installing an additional X-ray scanner can be simulated to determine suitable equipment levels.

Conducting a simulation study is often time-consuming and builds upon subjective inputs, such as interviews and observations. The emerging field of data-driven process simulation in Process Mining (PM) can overcome some of the limitations of "traditional" simulation model development by using data from Information Systems. Data-driven process simulation refers to the automated discovery of a simulation model from process execution data, i.e. an event log [9]. A key challenge in this field is data quality, given its strong impact on the reliability of the simulation results [31]. Because data quality issues are often encountered in healthcare event logs, it is imperative to assess these issues and correct them if needed. This will require domain knowledge. Current literature on data-driven simulation does not provide a clear framework to involve domain experts in model development.

This paper demonstrates the need for interactive data-driven process simulation in healthcare by assessing the impact of data quality issues on simulation results. To this end, a case study at the radiology department of a hospital is considered. In addition, we propose a novel conceptual framework which structures the integration of domain knowledge in the interactive development of data-driven simulation models.

The remainder of this paper is structured as follows. Section 2 gives an overview of the related work. The context of the case study is presented in Sect. 3. The experimental design, results, and discussion are presented in Sect. 4. Section 5 introduces our proposed framework for interactive data-driven process simulation. The paper ends with a conclusion in Sect. 6.

## 2   Related Work

This work relates to three key domains: (i) simulation for CM decisions in healthcare, (ii) data-driven process simulation, and (iii) data quality in process mining. The following paragraphs give a brief overview of these domains.

**Simulation for Capacity Management Decisions in Healthcare.** *Capacity Management decisions* in healthcare are concerned with determining the suitable levels of resources, such as staff size, equipment, and facilities [28]. In literature, simulation has been used to determine the required number of beds in general surgery [30]; the number of nurses, doctors, and buffer beds in an Emergency Department (ED) [7]; and the number of computed tomography (CT) scanners in a radiology department [27]. Within the radiology department, the context of our case study, Vieira et al. [32] gave an overview of Operations Research (OR) techniques – which includes simulation – for optimising resource levels and scheduling. For further reference on CM and the use of simulation in healthcare, the reader is referred to one of the existing review papers [26, 28, 33].

**Data-Driven Process Simulation.** *Data-driven process simulation* aims to "discover" BPS models from event logs automatically [9]. While existing PM

research can support the discovery of individual BPS model components [19] – e.g. control-flow discovery, decision mining, or organisational mining – less work has been devoted to integrating all these components into a single, simulation-ready model. Rozinat et al. [24] made a first attempt by discovering Coloured Petri Nets (CPNs) to describe the control flow. In addition, gateway routing logic and resource pools were also included. Later, the authors extended their method with activity execution times and case inter-arrival times [25]. Khodyrev and Popova [16] described a similar approach. However, the resource perspective was not included, assuming no resource constraints [16]. Gawin and Marcinkowski [13] provided support for activity durations, control-flow, resources, gateway routing logic, resource schedules, and inter-arrival times. However, the latter two were not automatically derived from data and had to be defined by domain experts [13]. *ClearPath* [15] provides a methodology for discovering and simulating Care Pathways (CPs). Their approach follows an agile, iterative method which facilitates the interaction between the modeller and domain expert, but the obtained process models still have to be manually recreated in their simulation tool *NETIMIS* [15]. *Simod* was the first tool to automatically integrate all components into a single, simulation-ready model to support BPS [6]. In addition, *Simod* is also capable of measuring the accuracy of the derived model and improve it using hyperparameter optimisation [6].

**Data Quality in Process Mining.** Real-life event logs tend to suffer from *data quality issues*, especially when they originate from flexible environments with substantial manual recording, such as healthcare [5,23]. These issues include missing events and incorrect timestamps, where the latter is often caused by batched registrations by healthcare staff [18,31]. Given the potential impact of event log quality issues on the reliability of PM outcomes, research attention on this topic is increasing. Research efforts are centred around three key topics. Firstly, several frameworks are developed which define event log quality issues [5,29,31]. For instance, Bose et al. [5] define 27 event logs quality issues and group them in four broad classes (i.e. missing, incorrect, imprecise, and irrelevant data). Secondly, research is performed on data quality assessment, targeting the systematic identification of event log quality issues. In this respect, the R-package *DaQAPO* [20], the log query language *QUELI* [1], and the *CP-DQF* [12] for Electronic Health Records (EHRs) provide tools and frameworks to operationalise data quality assessment. They are based on the event log quality issues defined in Vanbrabant et al. [31], Suriadi et al. [29], and Bose et al. [5], respectively. Thirdly, heuristics have been developed which tackle specific data quality issues, e.g. adding missing events [10], imputing missing case identifiers [3], and handling event ordering issues [11].

## 3   Background: Capacity Management at the Radiology Department

To illustrate the impact of data quality issues in the context of data-driven simulation, a real-life case study is used. This section introduces the case study.

## 3.1 General Context

The case study relates to a project at the radiology department of a hospital. Hospital management is preparing plans to build new facilities and is requesting input from each department regarding the required capacity. For the radiology department, this relates to the number of examination rooms – i.e. scanners – and the size of the waiting rooms – i.e. the number of seats – for each examination room. The radiology department wants to approach this Capacity Management problem in a data-driven way.

To support this data-driven analysis, process execution data is obtained from the Radiology Information Systems (RIS). This system supports the entire process flow, of which a simplified representation is shown in Fig. 1. The process starts when a patient arrives at the registration desk, after which (s)he is registered. Afterwards, the patient will wait in the waiting room until (s)he is called into the examination room. A nurse helps the patient onto the scanning table and correctly positions the scanner. Next, the image is created. In case the patient needs an additional scan of the same type, e.g. an X-ray scan of both shoulder and neck, this image can be made without leaving the room. After all required scans have been made, the patient can leave the examination room, and the nurse will post-process the images. If the patient still requires additional scans – of a different kind than the previous (e.g. also a CT scan) – (s)he will go to the waiting room of the other examination room. After all scans have been made, the patient can leave the radiology department and return home. Note that the interpretation of the scans by a radiologist is out of scope as it does not impact the required scanner and waiting room capacity.
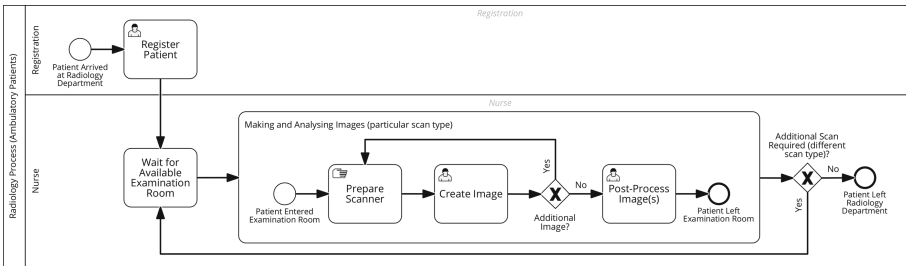


**Fig. 1.** Simplified process flow of (ambulatory) patients at the radiology department.

To solve the CM problem in this process, *Discrete-Event Simulation (DES)* is used due to the stochastic nature of the process. DES uses simulation to compare policy alternatives before implementing them in practice [33]. Arena v15 [2] was used to simulate the model.

In a DES model, *entities* are dynamic objects which move through the process and trigger the execution of activities [19]. In this case study, entities are patients visiting the radiology department. Four patient types are distinguished:

(i) *ambulatory patients (A)* which are outpatients, (ii) *day hospital patients (D)* which are admitted to the hospital for at most one day, (iii) *hospitalised patients (H)* which are inpatients, and (iv) *emergency patients (S)* which are transferred from the Emergency Department (ED).

The process flow depicted in Fig. 1 actually gives an overview of ambulatory patients. Nevertheless, the flow of the other patient types is, in essence, the same. Only the way patients arrive and where they wait are different. Hospitalised and day hospital patients will wait in their room until they are called in. Emergency patients will wait at the ED.

Depending on the type of scan, a different scanner – and thus a different examination room – is used. In this case study, there are six different types of scans of interest: *angiogram (ANGIO)*, *computed tomography (CT)*, *echocardiogram (ECHO)*, *mammogram (MAMMO)*, *magnetic resonance (MR)*, and *X-ray (RX)*. CT, ECHO, MAMMO, and MR all require separate rooms. ANGIO and RX are performed in RX rooms.

## 3.2   Data Description

To support the development of the DES model, two years of data from the RIS – from March 2017 until March 2019 – was available. The dataset includes various key timestamps for each patient visit, such as time of registration, and start and end time of scanning. Other attributes, such as the scan type (e.g. ECHO, RX, etc.) and patient type (e.g. ambulatory, emergency, etc.), were also recorded for each patient visit.

The dataset contains 404,750 individual patient visits. The proportions per patient type were 60%, 23%, 15%, and 2% for ambulant, hospitalised, emergency, and day hospital patients, respectively. In total, 464,053 scans were recorded, indicating that the majority of patients only needed one scan. Most scans were RX, i.e. 45%. ECHO represented 19%, followed by MR, 16%, 14% CT, and 5% MAMMO. A very small proportion, less than 0.001%, were ANGIO.

In the process, the activity "Create Image" (cf. Fig. 1) has the most considerable impact on waiting times and throughput rates because it generally takes longer than all other activities. Both start and end timestamps are available of this activity and are recorded when the nurse starts and stops the scanning device, respectively. We initially expected that this activity would not suffer much from quality issues because it is recorded automatically. However, this appeared not to be the case.

Table 1 gives an overview of the scan duration times per scan type. According to the data, some scans took over several years to complete. A few observations even had a negative duration, caused by the end timestamp being recorded before the start timestamp. Given its impact on capacity requirements, the scenario analysis will focus on the effect of scanning time data with data quality issues on simulation outcomes.

**Table 1.** Scan execution times (in mins).

| Scan Type | Min | Max | Mean | Median | SD | IQR |
|---|---|---|---|---|---|---|
| ANGIO | 0.00 | 323,258 | 14,372.11 | 26.05 | 57,336.29 | 87.83 |
| CT | −726.53 | 30,605 | 6.73 | 1.97 | 196.86 | 2.23 |
| ECHO | −79.00 | 116,685 | 71.36 | 23.38 | 636.37 | 28.48 |
| MAMMO | −6.48 | 40,780 | 16.41 | 2.98 | 531.00 | 1.35 |
| MR | 0.00 | 946,449 | 161.22 | 11.48 | 9,679.90 | 6.90 |
| RX | −1,031.63 | 2,109,457 | 22.69 | 0.55 | 5,111.25 | 1.20 |

## 4    Scenario Analysis: The Impact of Data Quality Issues

### 4.1    Experimental Design

To illustrate the impact of data quality issues w.r.t scanning times, we consider two scenarios:

– **Scenario 1 – Direct sampling:** In this scenario, actual observed data is sampled. This is useful when no theoretical distribution, such as the Gaussian, exponential, or gamma distribution, fits the data well. However, the disadvantage is that only the observed values can be used, which is problematic for smaller datasets [17].
– **Scenario 2 – Distribution fitting:** In this scenario, a distribution is fitted to the observed data. We used the distribution with the *least worst* fit because not a single distribution fitted the data well. With this approach, we follow the state-of-the-art of data-driven BPS techniques.

For each scenario, three alternative data filtering approaches are compared:

– **Alternative 1 – Validated filtering (VF):** In this alternative, which is the baseline, we used filtered data validated by domain experts. For scenario 2, we had to use empirical distributions for this alternative as none of the theoretical distribution provided a good fit. In the other two alternatives, we always used theoretical distributions.
– **Alternative 2 – No filtering (NF):** Here, we used the unfiltered data directly. Only observations less than zero were filtered out because the simulation model cannot handle negative activity durations.
– **Alternative 3 – Context-agnostic filtering (CAF):** Even without any domain knowledge, one would immediately notice that the maximum values in Table 1 are unrealistic. Therefore, this alternative uses filtered data to exclude anomalies. We adopted the commonly used *box plot rule* to detect anomalies in the absence of domain knowledge. Any observation smaller than $Q_1 - 1.5IQR$ or larger than $Q_3 + 1.5IQR$ is removed [8]. If the lower limit was less than zero, zero was used instead.

The length of the simulation run was set at two years for each alternative in each scenario. Initial experimentation showed that outliers in Alternative 2 caused severe queue accumulation, which resulted in i.a. extreme waiting times. Therefore, we integrated a weekly "reset", which removed all patients from queues and ongoing scans. We will refer to this reset as "flushing" and kept track of the weekly number of flushed patients.

To compare the alternatives, we focused on patient throughput and waiting times. Moreover, we looked at the flush count mentioned above. To measure the true effect of the different distributions used in each alternative, *common random number streams (CRNs)* are used. Consequently, the same random numbers are sampled across all alternatives. To compare the difference between alternatives, we used the non-parametric *Wilcoxon-Mann-Whitney (WMW)* test. Instead of using the original observations, ranks are used to compare the difference between two samples. This has the advantage that no underlying distribution is assumed [22]. To control the *false discovery rate (FDR)* of the multiple testing problem, we used the *Benjamini–Yekutieli* procedure [4] to adjust the *p*-values.

### 4.2   Results

**Throughput Times.** The *throughput time* measures the elapsed time between the patient's arrival and departure. Because a patient could require multiple scans, the *average throughput time per examination* is considered by dividing the throughput time of a patient by the number of scans. Patients who were "flushed" did not complete all scans and are therefore excluded from this measure.

As shown in Table 2, the throughput times for NF are much higher than VF, e.g. in Scenario 2, the average throughput time per examination for hospitalised patients is almost 100 times longer. The differences between CAF and VF are also statistically significant, albeit much smaller. For day hospital patients, representing 0.5% of the observations for this measure, the differences between VF and CAF were not statistically significant. Nevertheless, important differences in mean throughput times are observed due to larger outlier values for CAF.

**Waiting Times.** The *waiting time* is the time a patient spends in a queue before undergoing a scan. Table 3 shows comparable differences as the throughput times. Again, large differences between VF and NF are observed, e.g. the average waiting time for hospitalised patients is more than 150 times longer in NF than VF for Scenario 2. For day hospital patients, only the difference between VF and CAF in Scenario 1 is not significant, even though the absolute difference between the means is, again, rather large, indicating the presence of outliers.

**Flush Counts.** The more patients are flushed at the end of a week, the more this indicates that queues have accumulated throughout that week. Especially in the NF alternative, many patients have to be flushed to "reset" the process at the end of a week, in some cases even more than a thousand patients in total. The differences between VF and CAF are much smaller, i.e. on average less than one patient more was flushed in CAF. However, it should be noted that

**Table 2.** Throughput times per examination (in min) per patient type (A, D, H, S) and alternative (VF, NF, CAF).

| ADHS | Model 1 | Model 2 | Mean Model 1 | Mean Model 2 | Adj. $p$-value | Significance |
|---|---|---|---|---|---|---|
| | | | Scenario 1 | | | |
| A | VF | NF | 25.8493 | 556.622 | <0.0001 | **** |
| A | VF | CAF | 25.8493 | 45.8470 | <0.0001 | **** |
| D | VF | NF | 25.0905 | 129.1714 | 0.0077 | ** |
| D | VF | CAF | 25.0905 | 42.2703 | 0.6525 | ns |
| H | VF | NF | 29.6365 | 606.8702 | <0.0001 | **** |
| H | VF | CAF | 29.6365 | 34.1812 | <0.0001 | **** |
| S | VF | NF | 13.4890 | 92.7292 | <0.0001 | **** |
| S | VF | CAF | 13.4890 | 16.3655 | <0.0001 | **** |
| | | | Scenario 2 | | | |
| A | VF | NF | 25.8622 | 938.1933 | <0.0001 | **** |
| A | VF | CAF | 25.8622 | 48.8646 | <0.0001 | **** |
| D | VF | NF | 24.6727 | 363.3950 | <0.0001 | **** |
| D | VF | CAF | 24.6727 | 118.0489 | 0.1479 | ns |
| H | VF | NF | 29.6764 | 2,740.0265 | <0.0001 | **** |
| H | VF | CAF | 29.6764 | 253.3411 | <0.0001 | **** |
| S | VF | NF | 13.5586 | 76.9503 | <0.0001 | **** |
| S | VF | CAF | 13.5586 | 17.5182 | <0.0001 | **** |

****: $p$-value < 0.0001, ***: $p$-value < 0.001, **: $p$-value < 0.01, *: $p$-value < 0.05, ns: not signif.

**Table 3.** Waiting times (in min) per patient type (A, D, H, S) and alternative (VF, NF, CAF).

| ADHS | Model 1 | Model 2 | Mean Model 1 | Mean Model 2 | Adj. $p$-value | Significance |
|---|---|---|---|---|---|---|
| | | | Scenario 1 | | | |
| A | VF | NF | 8.7470 | 536.3071 | <0.0001 | **** |
| A | VF | CAF | 8.7470 | 26.4227 | <0.0001 | **** |
| D | VF | NF | 10.8402 | 112.1685 | <0.0001 | **** |
| D | VF | CAF | 10.8402 | 27.4978 | 1.0000 | ns |
| H | VF | NF | 17.3111 | 592.3354 | <0.0001 | **** |
| H | VF | CAF | 17.3111 | 20.7817 | <0.0001 | **** |
| S | VF | NF | 1.7663 | 76.4688 | <0.0001 | **** |
| S | VF | CAF | 1.7663 | 3.6703 | <0.0001 | **** |
| | | | Scenario 2 | | | |
| A | VF | NF | 8.7252 | 912.6071 | <0.0001 | **** |
| A | VF | CAF | 8.7252 | 31.3267 | <0.0001 | **** |
| D | VF | NF | 10.3187 | 337.0036 | <0.0001 | **** |
| D | VF | CAF | 10.3187 | 99.9275 | 0.0140 | * |
| H | VF | NF | 17.2705 | 2,737.3727 | <0.0001 | **** |
| H | VF | CAF | 17.2705 | 243.9300 | 0.0140 | * |
| S | VF | NF | 1.7545 | 47.0165 | <0.0001 | **** |
| S | VF | CAF | 1.7545 | 5.3972 | <0.0001 | **** |

****: $p$-value < 0.0001, ***: $p$-value < 0.001, **: $p$-value < 0.01, *: $p$-value < 0.05, ns: not signif.

sometimes the maximum number of flushed patients in CAF was much higher than in VF, e.g. for Scenario 2, VF flushed at most two hospitalised patients, whereas in CAF this was at most 46. For ambulatory patients, this was smaller, i.e. nine and seventeen, respectively.

### 4.3   Discussion

The results illustrate the need to consider data quality issues seriously. The unfiltered alternative – which completely neglects these issues – exhibits much higher throughput times, waiting times, and flush counts than the validated baseline. The difference between context-agnostic and validated filtering is smaller but still highly relevant. For instance, waiting times for hospitalised patients are up to eight times longer in CAF. However, for other performance metrics, such as flush counts, the differences between VF and CAF are smaller.

In this case study, the cut-off points for outliers in VF and CAF happened to be reasonably close to each other, except for echocardiograms. The domain experts indicated a maximum of 30 mins, whereas the box plot rule returned 84.64 min. However, this does not give any guarantee for other cases as context-agnostic filtering does not take into account the specificities of a particular domain in any way. Therefore, domain knowledge is always required to achieve accurate simulation results.

When comparing the differences between the two scenarios for each alternative (i.e. comparing the outcomes under direct sampling with their counterpart under distribution fitting), large differences are often observed between throughput and waiting times, even though the same input data was used. A possible explanation is that the theoretical distributions did not fit the data well. Therefore, we highlight the need to report goodness-of-fit (GoF) statistics in state-of-the-art data-driven BPS discovery algorithms and use direct sampling or empirical distributions in case no theoretical distribution fits the data well.

## 5   Interactive Data-Driven Process Simulation

As illustrated in the case study, data quality issues can have a profound impact on the reliability of simulation results. Moreover, domain knowledge plays a vital role in the development of a simulation model. Without domain knowledge, it is, e.g. challenging to determine whether particular observations are exceptional – but plausible – or data errors. Even though current literature on data-driven process simulation acknowledges the need for domain expertise for i.a. validation purposes, no framework conceptualises how this knowledge should be incorporated.

To enhance the integration of domain knowledge in the development of data-driven simulation models, we propose a novel conceptual framework which interactively involves experts during model building. This framework, which is visualised in Fig. 2, distinguishes three interaction cycles. In the *first cycle*, the initial model is constructed. For each required modelling task (e.g. entity arrival rate,

activity durations, resource roles, etc.) – of which an overview is presented in Martin et al. [19] – the data requirements are verified. For instance, mining resource roles requires the presence of a resource attribute. If these requirements are not fulfilled, the domain expert is asked for additional input to perform this modelling task. Conversely, if the requirements are fulfilled, the quality of the data is assessed, and an applicable discovery algorithm is employed. Next, the results of the discovery algorithm and detected data quality issues are presented for a check by the domain expert. (S)he can then solve any data quality-related issues and tweak the discovery parameters until the results are satisfactory.

The *second cycle* integrates all discovered model components from the first cycle into a single, simulation-ready model. The entire model is simulated, and the domain expert checks the preliminary results. If the simulation outputs do not satisfactorily reflect reality, the model can be "calibrated" by altering the simulation parameters. An estimation of the impact of the altered parameter on simulation outcomes is delivered in real-time, so the expert does not have to wait until the entire simulation has been completed before receiving an indication whether the altered parameter results in the desired change.

The final and *third cycle* is concerned with the validation of the model. The calibrated model from the second cycle is simulated comprehensively and validated by the domain expert. In addition, a validation dataset – which was not used to discover the model – can be used as well. If the desired accuracy level is not achieved, the domain expert can modify the simulation parameters again. The final validated model can be used for the evaluation of various scenarios and further analyses.
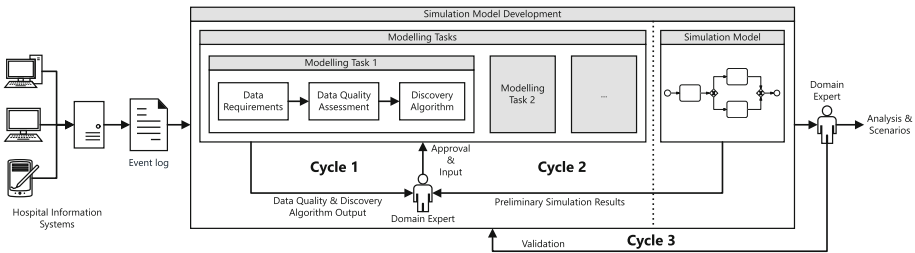


**Fig. 2.** Interactive data-driven process simulation framework.

## 6   Conclusion

Data-driven process simulation has great potential within a healthcare context, e.g. to support hospital management with Capacity Management decisions. However, real-life data extracted from Hospital Information Systems tend to suffer from data quality issues, which affects the reliability of simulation results. The presented case study at the radiology department of a hospital illustrates the impact of these issues, as well as the importance of domain knowledge.

Current literature on data-driven process simulation acknowledges the need for domain expertise but does not provide a framework to conceptualise the involvement of domain experts. Therefore, we propose a novel conceptual framework which interactively involves experts during data-driven simulation model building. In this framework, a distinction is made between three cycles: an initial development cycle, a calibration cycle, and a validation cycle.

Future work will focus on how the interaction between the domain expert and the framework will occur more specifically. Ultimately, our goal is to implement our framework into a tool to support the integration of domain knowledge into the development of data-driven process simulation models. In addition, this case study highlights the need for further research on identifying and remedying data quality issues in a healthcare context.

# References

1. Andrews, R., Suriadi, S., Ouyang, C., Poppe, E.: Towards event log querying for data quality. In: Panetto, H., Debruyne, C., Proper, H.A., Ardagna, C.A., Roman, D., Meersman, R. (eds.) OTM 2018. LNCS, vol. 11229, pp. 116–134. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-02610-3_7
2. Arena: Rockwell Automation, Inc. (2016). https://www.arenasimulation.com/
3. Bayomie, D., Awad, A., Ezat, E.: Correlating unlabeled events from cyclic business processes execution. In: Nurcan, S., Soffer, P., Bajec, M., Eder, J. (eds.) CAiSE 2016. LNCS, vol. 9694, pp. 274–289. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-39696-5_17
4. Benjamini, Y., Yekutieli, D.: The control of the false discovery rate in multiple testing under dependency. Ann. Stat. **29**(4), 1165–1188 (2001). https://doi.org/10.1214/aos/1013699998
5. Bose, R.P.J.C., Mans, R.S., van der Aalst, W.M.P.: Wanna improve process mining results? It's high time we consider data quality issues seriously. In: Proceedings of the 2013 IEEE Symposium on Computational Intelligence and Data Mining, pp. 127–134 (2013). https://doi.org/10.1109/CIDM.2013.6597227
6. Camargo, M., Dumas, M., González-Rojas, O.: Automated discovery of business process simulation models from event logs. Decis. Support Syst. **134**, 113284 (2020). https://doi.org/10.1016/j.dss.2020.113284
7. Carmen, R., Defraeye, M., Van Nieuwenhuyse, I.: A decision support system for capacity planning in emergency departments. Int. J. Simul. Model **14**(2), 299–312 (2015). https://doi.org/10.2507/ijsimm14(2)10.308
8. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: a survey. ACM Comput. Surv. **41**(3), 15:1–15:58 (2009). https://doi.org/10.1145/1541880.1541882
9. Depaire, B., Martin, N.: Data-driven process simulation. In: Sakr, S., Zomaya, A. (eds.) Encyclopedia of Big Data Technologies. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-63962-8_102-1
10. Di Francescomarino, C., Ghidini, C., Tessaris, S., Sandoval, I.V.: Completing workflow traces using action languages. In: Zdravkovic, J., Kirikova, M., Johannesson, P. (eds.) CAiSE 2015. LNCS, vol. 9097, pp. 314–330. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-19069-3_20
11. Dixit, P.M., et al.: Detection and interactive repair of event ordering imperfection in process logs. In: Krogstie, J., Reijers, H.A. (eds.) CAiSE 2018. LNCS, vol.

10816, pp. 274–290. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-91563-0_17

12. Fox, F., Aggarwal, V.R., Whelton, H., Johnson, O.A.: A data quality framework for process mining of electronic health record data. In: Proceedings of the 2018 IEEE International Conference on Healthcare Informatics, pp. 12–21 (2018). https://doi.org/10.1109/ICHI.2018.00009

13. Gawin, B., Marcinkowski, B.: How close to reality is the "as-is" business process simulation model? Organizacija **48**(3), 155–175 (2015). https://doi.org/10.1515/orga-2015-0013

14. Hicks, C., McGovern, T., Prior, G., Smith, I.: Applying lean principles to the design of healthcare facilities. Int. J. Prod. Econ. **170**, 677–686 (2015). https://doi.org/10.1016/j.ijpe.2015.05.029

15. Johnson, O.A., Ba Dhafari, T., Kurniati, A., Fox, F., Rojas, E.: The ClearPath method for care pathway process mining and simulation. In: Daniel, F., Sheng, Q.Z., Motahari, H. (eds.) BPM 2018. LNBIP, vol. 342, pp. 239–250. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11641-5_19

16. Khodyrev, I., Popova, S.: Discrete modeling and simulation of business processes using event logs. In: Proceedings of the 14th International Conference on Computational Science. Procedia Comput. Sci. **29**, 322–331 (2014). https://doi.org/10.1016/j.procs.2014.05.029

17. Law, A.M.: Simulation Modeling and Analysis, 5th edn. McGraw-Hill Education, New York (2014)

18. Mans, R.S., van der Aalst, W.M.P., Vanwersch, R.J.B.: Process Mining in Healthcare: Evaluating and Exploiting Operational Healthcare Processes. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16071-9

19. Martin, N., Depaire, B., Caris, A.: The use of process mining in business process simulation model construction. Bus. Inf. Syst. Eng. **58**(1), 73–87 (2015). https://doi.org/10.1007/s12599-015-0410-4

20. Martin, N., Van Houdt, G., Janssenswillen, G.: Towards more structured data quality assessment in the process mining field: the DaQAPO package. In: Proceedings of the European R Users Meeting (2020)

21. Melão, N., Pidd, M.: Use of business process simulation: a survey of practitioners. J. Oper. Res. Soc. **54**(1), 2–10 (2003). https://doi.org/10.1057/palgrave.jors.2601477

22. Neuhäuser, M.: Wilcoxon-Mann-Whitney test. In: Lovric, M. (ed.) International Encyclopedia of Statistical Science, pp. 1656–1658. Springer, Berlin (2011). https://doi.org/10.1007/978-3-642-04898-2_615

23. Rebuge, Á., Ferreira, D.M.R.: Business process analysis in healthcare environments: a methodology based on process mining. Inf. Syst. **37**(2), 99–116 (2012). https://doi.org/10.1016/j.is.2011.01.003

24. Rozinat, A., Mans, R.S., van der Aalst, W.M.P.: Mining CPN models: discovering process models with data from event logs. In: Proceedings of the 7th Workshop and Tutorial on Practical Use of Coloured Petri Nets and the CPN Tools. DAIMI PB, vol. 579, pp. 57–76 (2006)

25. Rozinat, A., Mans, R.S., Song, M., van der Aalst, W.M.P.: Discovering simulation models. Inf. Syst. **34**(3), 305–327 (2009). https://doi.org/10.1016/j.is.2008.09.002

26. Salleh, S., Thokala, P., Brennan, A., Hughes, R., Booth, A.: Simulation modelling in healthcare: an umbrella review of systematic literature reviews. Pharmacoeconomics **35**(9), 937–949 (2017). https://doi.org/10.1007/s40273-017-0523-3

27. Shakoor, M.: Using discrete event simulation approach to reduce waiting times in computed tomography radiology department. Int. J. Ind. Manuf. Eng. **9**(1), 177–181 (2015). https://doi.org/10.5281/zenodo.1338044

28. Smith-Daniels, V.L., Schweikhart, S.B., Smith-Daniels, D.E.: Capacity management in health care services: review and future research directions. Decis. Sci. **19**(4), 889–919 (1988). https://doi.org/10.1111/j.1540-5915.1988.tb00310.x

29. Suriadi, S., Andrews, R., ter Hofstede, A.H.M., Wynn, M.T.: Event log imperfection patterns for process mining: towards a systematic approach to cleaning event logs. Inf. Syst. **64**, 132–150 (2017). https://doi.org/10.1016/j.is.2016.07.011

30. VanBerkel, P.T., Blake, J.T.: A comprehensive simulation for wait time reduction and capacity planning applied in general surgery. Health Care Manag. Sci. **10**(4), 373–385 (2007). https://doi.org/10.1007/s10729-007-9035-6

31. Vanbrabant, L., Martin, N., Ramaekers, K., Braekers, K.: Quality of input data in emergency department simulations: framework and assessment techniques. Simul. Model. Pract. Theory **91**, 83–101 (2019). https://doi.org/10.1016/j.simpat.2018.12.002

32. Vieira, B., Hans, E.W., van Vliet-Vroegindeweij, C., van de Kamer, J., van Harten, W.: Operations research for resource planning and-use in radiotherapy: a literature review. BMC Med. Inform. Decis. Mak. **16**(1) (2016). Article number: 149. https://doi.org/10.1186/s12911-016-0390-4

33. Zhang, X.: Application of discrete event simulation in health care: a systematic review. BMC Health Serv. Res. **18**(1), 687 (2018). https://doi.org/10.1186/s12913-018-3456-4