Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

A PRELIMINARY PREDICTIVE SIMULATION MODEL FOR HIP AND KNEE REPLACEMENT PROFILE-DEPENDENT PATHWAY STAGES

Ahmed Bakali El Kassimi
Xiaolan Xie

Mines Saint-Etienne, Univ Clermont Auvergne
INP Clermont Auvergne, CNRS
UMR 6158 LIMOS
Saint-Étienne, 42023, FRANCE

Marianne Sarazin
Pierre-Luc Fresard
Bertrand Semay

Équipe De Recherche
Clinique Médico-Chirurgicale Mutualiste
Aésio Santé
Saint-Étienne, 42100, FRANCE

ABSTRACT

Total hip and knee arthroplasty (THA/TKA) are commonly performed on elderly individuals, involving a lengthy clinical pathway that includes preoperative, operative, and rehabilitation phases. Despite efforts to improve patient satisfaction, there is a lack of personalized studies that optimize the THA/TKA pathway. Our aim is to address this gap by proposing a predictive simulation model that considers patient-specific factors to enhance patient satisfaction and organizational efficiency. To achieve this, we propose using process mining techniques to analyze the French national healthcare database and distinguish between standard care phases and patient-dependent phases. We then apply machine learning algorithms to predict specific stages of care. The insights gained from these analyses are used to compare and test predicted patient pathways and their performances using a hybrid simulation model.

1 INTRODUCTION

With the rapid evolution of healthcare information systems and the availability of vast medical databases, clinical pathways are garnering increasing attention from the medical and scientific communities. While several approaches are being explored, such as the implementation of standardized care pathways and the use of new technologies, these techniques often fall short of addressing and managing the considerable variability in patient requirements based on their profiles, and they typically focus on general care aspects, rarely encompassing the specific needs. In this context, the care pathway of patients undergoing hip or knee replacement surgery has received extensive attention due to the large demand and the extended pre and post-operative care periods. More than one million THA and TKA procedures are performed each year, and this number continues to grow annually in many countries, with 224,170 THA/TKA procedures performed in France in 2018. Despite several improvements, including enhanced recovery after surgery (ERAS) programs, no comprehensive study has examined the entire pathway organization to deliver care that caters to patient profiles and organizational constraints. Therefore, this study has two objectives: first, to provide a personalized simulation model that can predict a patient's trajectory based on their characteristics, and second, to propose new organizational solutions to enhance care pathways and meet patient expectations. The proposed simulation model can be used as a decision-making tool to assist practitioners in making decisions throughout the patient's care. This paper presents the initial analyses conducted within this framework and will be further expanded in future versions.

979-8-3503-6966-3/23/$31.00 ©2023 IEEE

2 LITERATURE REVIEW

This section provides a comprehensive review of the relevant literature. Specifically, the focus will be on studies that explore the application of hybrid modelling and simulation (M&S) approaches. The following paragraphs delve into the pertinent aspects of each concept, emphasizing the advantages of hybrid M&S approaches and drawing comparisons with our proposed approach.

Several studies have discussed hybrid M&S in the literature. First, Mustafee et al. (2020) provide a unified conceptual representation and classification of hybrid M&S, identifying six distinct model types. Among these, type "D" Hybrid operational research (OR)/M&S models combine computer simulation with hard OR techniques such as process mining and machine learning. This aligns with the approach presented in this paper. Approaches combining techniques such as process mining, modelling, prediction, and simulation aim to address the limitations of traditional simulation models, which often rely on a standardized representation of healthcare pathways and overlook the variability and uncertainties that can occur in real-life situations, leading to potential bias in the interpretation of results.

Several studies have emphasized the applications of process mining in healthcare. In their systematic review, Dallagassa et al. (2022) identified process discovery, resource analysis, and evaluation as common use cases for process mining in the healthcare domain. Additionally, De Roock and Martin (2022) presented an updated perspective on the state of the art that highlights the importance of converting process mining analysis results into specific actions to realize their true value. The combination of process mining with simulation in hybrid approaches offers the advantage of leveraging the strengths of both techniques to enhance the accuracy and reliability of the modelling and simulation process. Numerous studies have explored this aspect, Phan et al. (2019) analyzed the clinical pathway of patients with an incisional hernia and proposed a process mining and simulation approach to analyze potential complications and associated costs. Abohamad et al. (2017) proposed a hybrid process-mining framework that has been applied to identify performance bottlenecks, identified deviations from process guidelines and explore improvement strategies in an Emergency Department (ED). Augusto et al. (2016) proposed a methodology to automatically perform simulation analysis of clinical pathways of patient cohorts using process mining techniques with a combined agent-based and discrete-event simulation model. While the study considered only fixed probabilities, the authors proposed adding more parameters and including dynamic decision models as a future perspective of the work.

Machine learning (ML) algorithms have also been introduced to enhance M&S approaches. Elbattah (2019) explores the opportunities of integrating ML models into M&S practices to construct high-fidelity representation models. The study provides insights and recommendations for further research directions in this area. In addition, Gehlot et al. (2022) present a taxonomy and architecture for combining simulation models with AI/ML models. They illustrate the framework with various examples from the healthcare domain, demonstrating the potential and applicability of integrating simulation and ML techniques. Elbattah et al. (2018) proposed a framework that combines Simulation Modeling and Machine Learning to design care pathways and evaluate their return on investment, using unsupervised ML to extract knowledge from the Irish Hip Fracture Database and developing simulation models to optimize the quality and cost of care. Their work focuses on simulating the care pathways of patients with hip fractures using insights from unsupervised machine learning clustering. It should be noted that the pathway considered in their study differs from that of scheduled hip and knee surgeries, and the techniques used also vary. The aspect of discovering the pathway through process mining techniques is also not addressed.

Recent studies have highlighted the benefits of combining machine learning and process mining approaches. Kempa-Liehr et al. (2020) proposed a process mining pipeline that combines healthcare pathway discovery and a probabilistic regression model to explore pathway features that influence patient recovery time. Mesabbah et al. (2019) present an extension to the Auto simulation Model Builder (ASMB) framework previously developed by authors (Mesabbah and McKeever 2018) adopted for healthcare systems to include resource handling and decision-making processes around hospital staff planning, using machine learning and real-time data-driven prediction to improve system performance.

In addition to their integration into hybrid M&S approaches, machine learning approaches have been widely utilized in healthcare systems for predicting patient outcomes. Several studies have examined machine learning techniques in healthcare, including predicting demand on healthcare institutions, resource requirements for patient transfers, inpatient treatment, length of stay and discharge timing (El-Bouri et al. 2021). Machine learning techniques have also found application in the field of orthopaedics. A systematic review conducted by Cabitza et al. (2018) explored the use of machine learning techniques in orthopaedics. Furthermore, Lopez et al. (2021) investigated the potential of AI/ML models in Total Joint Arthroplasty (TJA) for optimizing patient selection and accurately predicting postoperative outcomes, complications, and associated costs. Another study by Milella et al. (2022) applied machine learning to improve treatment appropriateness in an orthopaedic setting, comparing three XGBoost models to predict non-achievement of postoperative treatment based on changes in patient-reported outcome measures. Additionally, Hinterwimmer et al. (2023) proposed a standardized approach for machine learning in primary total knee arthroplasty, predicting complications and surgery duration using data from two German arthroplasty registries and achieved promising results using XGBoost.

Compared with the existing literature in the field of orthopaedics and hip and knee replacement surgeries. This paper represents the first study to explore the THA/TKA pathways using a hybrid M&S approach using process mining, machine learning, and simulation techniques, providing valuable insights into the entire care trajectory. This paper highlights the results of initial analyses conducted, elucidates the advantages and limitations of the approach used, and outlines the perspectives that we will incorporate in the subsequent stages of this work.

3 GENERAL FRAMEWORK

Our study utilizes a global methodology consisting of four main steps, which are outlined in Figure 1. Firstly, we collect and prepare data from the French national healthcare data system in Step 1. In Step 2, we convert patient trajectory data from Step 1 into an event log for process mining and use the disco process mining software to identify the most frequent paths and profile-dependent stages of care. Moving on to Step 3, we collaborate with medical experts to select and predict relevant profile-dependent stages and compare the performance of different machine learning algorithms. In Step 4, we introduce a simulation model that incorporates the outcomes of the previous steps and the experts' knowledge and compares different predicted patient pathways based on some of the patient characteristics.

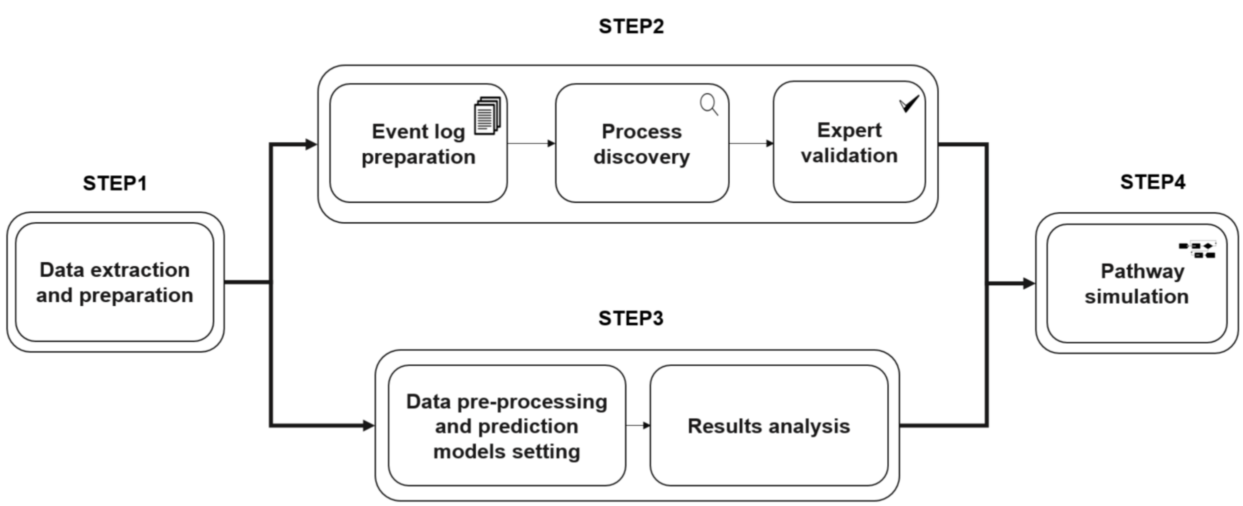

Figure 1: General framework.

3.1 Step 1: Data Description And Preparation

This study utilizes comprehensive anonymized data from the French national health hospitalization data system, which has been provided by our collaborating hospital, authorized to access the data system. The dataset encompasses information regarding patients' hospital stays, related medical visits, and reimbursement care details. It is important to note that these data do not include information on care received outside the hospital setting from independent healthcare practitioners. This limitation may impact the quality of the results, and therefore, we need to be cautious in interpreting the findings, and seeking guidance from domain experts.

To ensure the significance of care pathways, we consulted medical professionals for their expertise and focused on data from the years 2017, 2018, and 2019. The initial hospitalization for total knee or total hip replacements (THA), performed in 2018, was our reference event and the inclusion criterion. Hospitalization data were extracted using French Diagnosis Related Groups codes 08C24, 08C48 and ICD-10 codes for principal diagnoses M16 and M17. There were 214 308 TKA/THA procedures performed on 209 473 patients in 2018. To track patients' care pathways, we added to the data collected during the first THA/TKA stay in the hospital in 2018 all information about the activities carried out a year earlier and a full year after the reference event. This included information on subsequent hospitalizations and any related medical consultations. We aim at providing a comprehensive view of patient outcomes and care quality by examining this time period.

One challenge of utilizing these databases is the complexity of associating activities with their corresponding disciplines. To address this issue, we employed a specific standard called "Noemie" to indicate the specialities of the practitioners and developed distinctive labels for each discipline. Additionally, there are discrepancies in the consultation codes depending on the medical professional's status and practice, such as the inclusion of extra fees, which can lead to additional rows for reimbursement operations. Therefore, in collaboration with the practitioners, we decided to retain only the primary codes related to specialist consultation while grouping codes with the same meaning under a common label.

Furthermore, the hospital stays data tables contain interesting information such as the mode of hospital discharge, which varies for most cases between transfer to home or rehabilitation center. These discharges will be added to the patients' care events. The tables also contain information related to patient profiles (age, gender) and can be linked to associated diagnostic tables to obtain an idea of their characteristics and risk factors.

The rest of this article will detail the remaining data pre-processing organized by each work axis. Firstly, we will focus on the event log preparation for process discovery (Section 3.2). Then, we will use patient profile data such as associated diagnostics for predicting certain events in the patient care trajectory (Section 3.3).

3.2 Step 2: Process Discovery

3.2.1 Convert Data To An Event Log

Process mining algorithms rely on event logs, which contain information about the events that occur during a patient's care trajectory. However, these logs may have multiple events with the same label and occur at different times during the trajectory. To overcome this issue, we propose an enhanced approach that sequences events based on their relation to the principal date of the THA/TKA surgery. This allows us to distinguish between preoperative and postoperative events, providing a more precise understanding of the care trajectory. We establish an ordering of events using the sequence number of each event in the French healthcare database, representing the anonymized entry date for each patient. The temporal order is preserved as the difference between two consecutive events sequence numbers corresponds to the difference in days. Preoperative events are labelled with a 'PRE-' prefix, and postoperative events with a 'POST-' prefix. As the care trajectory data combines activities related to TKA/THA surgery and other types of care, a choice was made with experts to categorize the events based on their proximity to the main THA/TKA hospitalization. For this purpose, a suffix is added to the event label, as follows:

- Category 1 Events [-CAT1]: [-1 month, +1 month]
- Category 2 Events [-CAT2]: [-3 months, -1 month[$\cup$]+1 month, +3 months]
- Category 3 Events [-CAT3]: [-1 year, -3 months[$\cup$]+3 months, +1 year]

As the events are not precisely timestamped, with only the month and year, a proposal for event dating was made, considering that the hospital stay occurred on the "year-month-15" date. The other dates are then determined using the differences between the sequence numbers. These dates allow for a realistic interpretation of the patient care trajectory and have no impact on the results.

The following Figure 2 illustrates the event log conversion for a single patient:

ID	Sequence	Activity	Month	Year
Patient_1	21988	ORTHO	9	2017
Patient_1	22052	ANESTH	11	2017
Patient_1	22074	CARDIO	12	2017
Patient_1	22100	TKA SURGERY	1	2018
Patient_1	22107	REHABILITATION TRANSFER	1	2018
Patient_1	22145	ORTHO	3	2018
Patient_1	22237	ORTHO	6	2018

$\Downarrow$

ID	Sequence	Activity	Month	Year	Date
Patient_1	21988	PRE-ORTHO-CAT3	9	2017	2017-09-25
Patient_1	22052	PRE-ANESTH-CAT2	11	2017	2017-11-28
Patient_1	22074	PRE-CARDIO-CAT1	12	2017	2017-12-20
Patient_1	22100	TKA SURGERY	1	2018	2018-01-15
Patient_1	22107	POST-REHABILITATION TRANSFER-CAT1	1	2018	2018-01-22
Patient_1	22145	POST-ORTHO-CAT2	3	2018	2018-03-01
Patient_1	22237	POST-ORTHO-CAT3	6	2018	2018-06-01

Figure 2: Event log conversion.

The final event log consists of 1,139,107 events for 209,473 patients, and 397 types of activities (divided between pre- and post-operative activities). Despite the large size of the event log at our disposal, we would like to emphasize that it contains several incomplete pathways, which is due to data loss associated with extra-hospital activities carried out by independent practitioners. This necessitates caution and reliance on medical expertise in our analyses to avoid any bias in interpretation. The subsequent section focuses on the process of discovery of the hip and knee replacement pathway.

3.2.2 Process Discovery

In this section, we aim to gain insights into the TKA/THA patient care process from the national health data system using the sequenced event log. To achieve this goal, various solutions are available, ranging from commercial software like Disco, Prom and PM4PY Python library, to more sophisticated options such as designing a novel process discovery approach. The literature highlights two commonly used process discovery techniques in healthcare, Fuzzy Miner and Heuristic Miner. Viewed with experts in our collaborating hospital, conducting an in-depth analysis of the event log may not be useful due to the linear nature of the pathway and the incomplete path traces of some cases. Therefore, we select Disco 3.5.7 to analyze the event log, as it provides various analytical tools and enables all project stakeholders to interpret the results. By doing so, we can validate practitioners' feedback on the pathway and establish a definitive pattern for the pathway model, which can be simulated more clearly. As shown in Figure 3 we display the discovered pathway using the case coverage metric, which presents the percentage of cases where an activity or path in the process map has been observed.

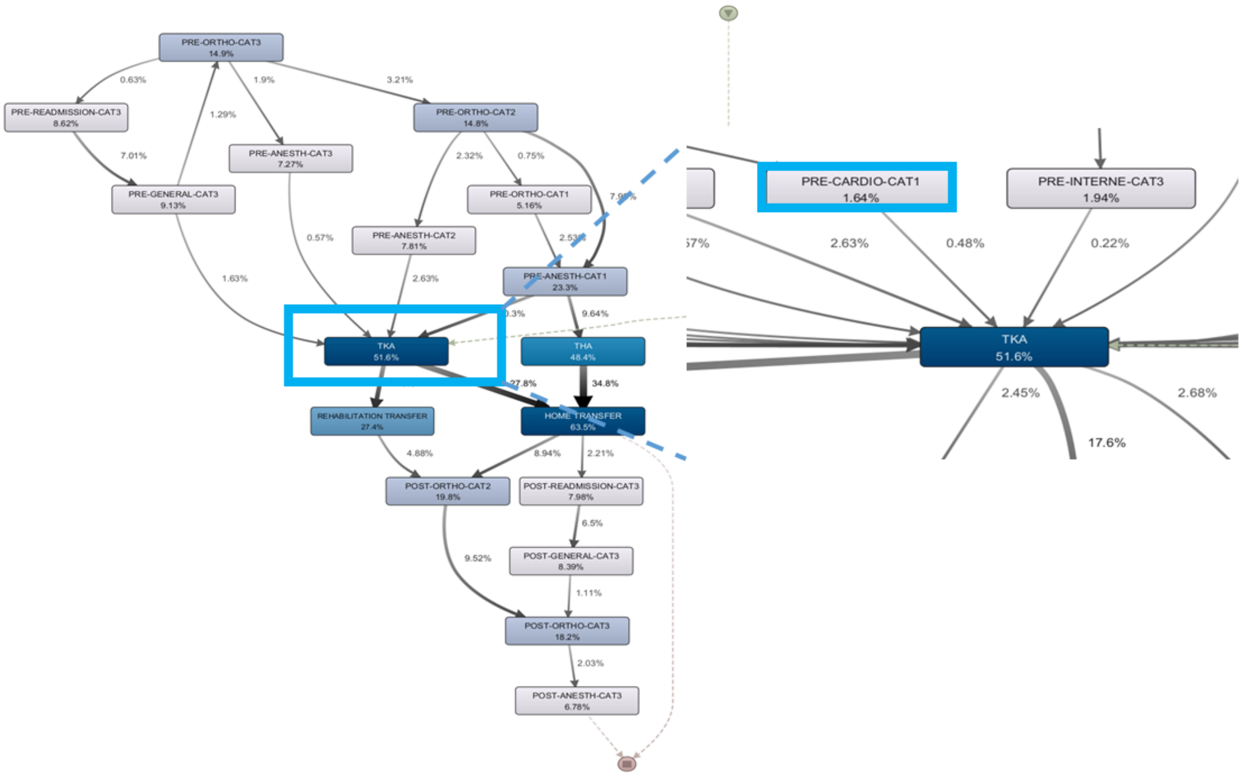

Figure 3: Discovered process model using Disco (Version 3.5.7).

The simplified model on the left side of Figure 3 represents only 4% of the most frequent events labels. This model reflects the standard organization of the patient pathway based on the feedback from physicians in our collaborating hospital. It confirms that the patient care process is fairly standardized across the French territory with more or less similar practices.

Despite the common practices, we can detect certain less-frequent events depending on the patients' profiles, as shown in the Disco model diagram (15% of the frequently occurring event labels) (Figure 3 on the right). These events are highly relevant, as they may have an impact on the standard of care, the use of resources, and patient satisfaction. It is therefore important to predict and anticipate the occurrence of such events from the very first consultation, to ensure an optimal organization tailored to the specific needs of each patient.

In this preliminary study, we focused on predicting three specific events, selected in collaboration with experts. This choice was validated based on the degree of importance of these events and their specific impact on the organization. The first event is the pre-operative cardiology consultation, which is crucial when patients with cardiac risk factors have not been properly monitored prior to the decision for hip or Knee replacement. Neglecting this consultation can jeopardize the smooth progression of the hospital stay and the overall success of the procedure. The second event is the transfer to a rehabilitation center, as this decision takes into account multiple patient characteristics and requires reserving a place in the rehabilitation center upon discharge. It also involves the mobilization of specific resources, such as social workers, to support the patient throughout the rehabilitation process. Lastly, we are interested in predicting postoperative readmissions. Analyzing the causes behind these readmissions will allow us to identify areas for improvement and implement actions to reduce their occurrence in the future.

3.3 Step 3: Events Prediction

The aim is to compare several models for predicting the following targets: scheduling of a pre-operative cardiology consultation, transfer to a rehabilitation center, and the risk of postoperative readmission.

3.3.1 Data Pre-processing And Prediction Models Setting

To predict specific events, we consider three categories of predictive feature data: general patient information (such as age and gender), associated diagnoses (including conditions such as diabetes and obesity), and medical history (which includes all the treatments and care provided to the patient in the past). We evaluate the impact of associated diagnoses on our predictions by comparing two sets of ICD-10 codes' hierarchical levels. Specifically, we use codes grouped into general chapters (Hiera-CH), and sub-chapters (Hiera-SUB-CH).

Regarding patient information data, we first categorize age into four groups using expert knowledge. We classify adults aged 0-64 as pre-elderly, those aged 65-74 as youngest-old, those aged 75-84 as middle-old, and those over 85 as oldest-old. Since all variables are categorical, we perform one-hot encoding to represent them as binary features. An inter-correlation analysis is then conducted using a chi-square test with a threshold value set at 0.7 to remove redundant variables. Refer to Table 1 for a comprehensive overview of the features.

Table 1: Features description.

Features categories	Type	Categories	Proportion	Sets	Targets
General patient information					
Surgery type	Categorical	Hip	0.48	Set 1, Set 2	All
		Knee	0.52		
Age	Categorical	Pre-elderly	0.27	Set 1, Set 2	All
		Youngest-old	0.39		
		Middle-old	0.28		
		Oldest-old	0.06		
Sexe	Categorical	Men	0.42	Set 1, Set 2	All
		Women	0.58		
Associated diagnosis					
Hiera CH	Categorical	CH-01..CH-22		Set 1	All
Hiera SCH	Categorical	CH-01.01...		Set 2	All
Medical History					
pre-activities CAT 3 & CAT 2	Categorical			Set 2	Cardiology & transfer
All preopeative activities	Categorical			Set 2	Readmission

Initially, we select the feature set and the target for prediction. We then divide the data into training (80%) and test (20%) data sets. Since all feature variables are binary, there is no need to perform variable scaling. As the three targets to be predicted have minority classes (refer to Table 2), we propose to compare several oversampling and under-sampling methods, or a combination of both, on the training data for each prediction.

Table 2: Targets variables.

	Frequency of targets variables		
	Cardiology consultation	Rehabilitation transfer	Postoperative readmission
Negative	206043 (98,36%)	152107 (72,61%)	206702 (98,70%)
Positive	3430 (1,64%)	57366 (27,39%)	2771 (1,30%)

To balance the minority classes we compare the following methods:

- **Oversampling Methods:** RandomOverSampler **(ROS)** and **SMOTE**.
- **Undersampling Methods:** RandomUnderSampler **(RUS)**, NearMiss **(NM)**, and TomekLinks **(TL)**.
- **Combined Method: SMOTETomek**.

We aimed to predict target variables based on available information at the time of decision-making, and thus proposed to study and compare the following feature sets:

- **Set 1 (Hiera-CH):** This set includes ICD- codes grouped into general chapters, this patient profile-related information can be collected at the start of the pathway, and used to predict all targets.
- **Set 2 (Medical history + Hiera-SUB-CH):** In this set we incorporated medical history data into the Hiera-sub-chapter (Hiera-SUB-CH) diagnosis information. To predict preoperative cardiology consultations (PRE-CARDIO-CAT1) and rehabilitation transfers, we only considered medical activities that took place between one month and one year prior to surgery (Pre-activities CAT3 and CAT2), excluding category 1 pre-activities. For predicting postoperative readmission, we can expand the medical history activities to include the entire information until the end of the hospital stay.

In this study, we compare the performance of three machine learning algorithms: logistic regression **(LR)**, decision tree **(DT)**, and random forest **(RF)**. To evaluate the performance of each method, we use multiple metrics, including precision, recall, F1 score, accuracy, and ROC-AUC. Furthermore, each method is tested with various hyperparameters, optimized through the HalvingGridSearch method to achieve an optimal configuration.

3.3.2 Events Prediction Results

The Table 3 presented below summarizes the results obtained after testing different combinations of feature sets, machine learning models, and data balancing methods. Due to the imbalance in the data, the results were compared using the F1-score. We observed that the use of feature set 2, which includes subchapter diagnosis data and patient medical history, led to a slight improvement in the results. However, none of the F1 scores showed good performance. The prediction of rehabilitation transfers still needs improvement, with an F1-score of 0.5, only 41% of true positive predicted classes, and 65% of identified rehabilitation transfers. The prediction models for cardiology consultations and readmission showed very poor performance, with a high number of false positive predicted classes (low precision).

To improve the prediction results, we will introduce a new feature set that represents associated diagnoses using the first three characters of the ICD-10 codes in the following of this work. Additionally, we have requested access to the private practice French healthcare database (activity of independent practitioners), to include more examples in the minority classes and improve the prediction results.

Table 3: Prediction results.

Predicted target	Features set	ML model	Balancing method	Precision	Recall	F1-Score	Accuracy	ROC-AUC
Cardiology consultation	S1	DT	RUS	0.02	0.47	0.04	0.64	0.56
	S2	LR	TL	0.04	0.56	0.08	0.77	0.72
Rehabilitation transfer	S1	RF	RUS	0.39	0.69	0.50	0.62	0.69
	S2	LR	TL	0.41	0.65	0.50	0.65	0.71
Postoperative readmission	S1	DT	RUS	0.02	0.20	0.03	0.85	0.53
	S2	RF	ROS	0.02	0.51	0.04	0.70	0.64

3.4 Step 4: Process Simulation

Based on insights from the previous steps and the experts' opinions, we propose to model the clinical pathway for the simulation model as a combination of two types of stages, patient common stages and profile-dependent stages (predicted activities) as illustrated in the following Figure 4:

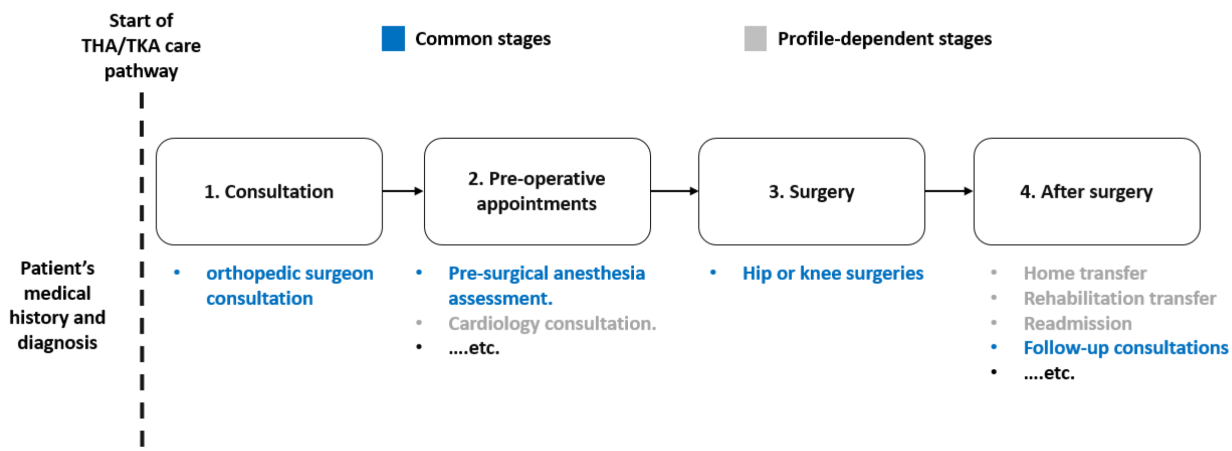

Figure 4: Commons and profile-depedent care-stage.

Figure 4 depicts the essential stages involved in hip or knee surgery, including the initial consultation with the surgeon, anesthesia consultation, and follow-up consultations. In addition to these standard stages, there are profile-dependent care stages, such as cardiology consultation, rehabilitation transfer, and postoperative readmission, which correspond to predicted targets. At the onset of the simulation, we predict these three profile-dependent stages based on patients' characteristics. Subsequently, we employ various precedence rules within our pathway model to generate personalized pathway traces that serve as inputs for our simulation model. The simulation model itself was created using Anylogic 8.4.0 software, as shown in Figure 5.

The simulation model is designed as a hybrid simulation model, where each patient is represented as an agent characterized by a set of parameters. These agents are added to a population and injected into the discrete event simulation model to trace their individual care pathways. This approach allows for the inclusion of distributions and indicators related to wait times, resource utilization, and average durations of care.

We employed our simulation model to compare the predicted patient paths with their actual paths based on their age group. To do this, we generated a sample of 1000 observations for each age category from our initial data set. The simulation findings are presented in Table 4. The patient-dependent care stages demand comparison has confirmed for each age group that predictive models are not effective in forecasting profile-dependent events, particularly in cardiology and postoperative readmission. Hence, they need to be improved in the future through the incorporation of new factors, additional observations, and the testing of new algorithms and methodologies.

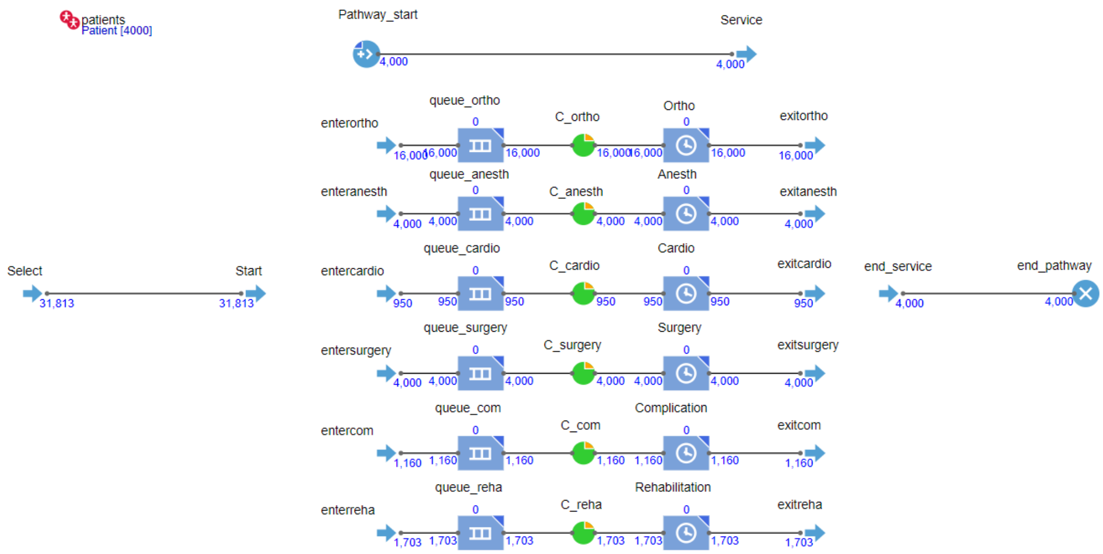

Figure 5: Process simulation model.

Table 4: Patient flow for specific stages of care for each age group.

		Cardiology consultation	Rehabilitation transfer	Postoperative readmission
pre-elderly	Predicted	22,90%	41,40%	31,10%
	Real	1,90%	27,10%	1,40%
youngest-old	Predicted	25,40%	42,30%	30,50%
	Real	2%	27,60%	0,80%
middle-old	Predicted	22,10%	41,10%	26,60%
	Real	1,60%	26,50%	1,70%
oldest-old	Predicted	24,60%	45,50%	27,80%
	Real	1,40%	29,70%	1,30%

4 DISCUSSION AND CONCLUDING REMARKS

This paper introduces a patient-centred approach to simulate the clinical pathway of hip and knee replacements, based on electronic health records (EHR) from the French national health data system. The study validates the standard pathway while identifying stages that depend on patient profiles. The primary goal of predicting these profile-dependent stages is to personalize treatment approaches and support healthcare practitioners' decision-making. The simulation model proposed in this paper served as an initial evaluation of prediction outcomes for different patient age categories, comparing them with the actual patient journeys. This will enable testing various patient profiles in the future and obtaining precise estimates of the necessary resources to be deployed.

However, this paper has some limitations related to both the data used and the resulting models. Regarding the data, one limitation is the factors considered for prediction and their level of detail, which could be improved. In this regard, we propose primarily using ICD-10 codes instead of codes grouped by chapter or sub-chapter to enhance prediction outcomes. An analysis of the importance of these factors and their influence on the patient pathway is also being considered, supported by the expertise of medical professionals.

The second limitation pertains to the inclusion of other types of data typically not well captured in national databases. Specifically, social and environmental data play a crucial role in several medical decisions and will contribute to the improvement of the decision support model. Efforts have been made to collect these data through mobile tracking applications.

The third limitation involves addressing all stages dependent on the patient's profile, particularly those related to the postoperative phase and rehabilitation protocols. This can be achieved by incorporating data on activities performed by independent practitioners. An authorization request for access to this data has been submitted to the relevant authorities.

Regarding the model, we aim to enhance it by introducing a dynamic aspect to our prediction models, allowing them to adapt to changes in the patient's condition and the impact of certain events on the subsequent care trajectory. Additionally, a multi-output prediction model and the testing of new prediction algorithms could be proposed and integrated into the simulation environment.

ACKNOWLEDGMENTS

The authors would like to thank the Aésio Santé group for their support for this project, Mr Remi BOUVIER, General Manager, Mr Guillaume GARDIN, Director of Research and Innovation, Mrs Karima TAHTAH, Director of the clinique mutualiste de Saint-Etienne healthcare institution of the group, Mrs Françoise MORAND, Mission leader for the healthcare division of the group as well as all the teams of the clinique mutualiste de Saint-Etienne, and its research department, for their involvement throughout the project.

REFERENCES

Abohamad, W., A. Ramy, and A. Arisha. 2017. "A Hybrid Process-mining Approach For Simulation Modeling". In *Proceedings of the 2017 Winter Simulation Conference*, edited by W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, 1527–1538. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Augusto, V., X. Xie, M. Prodel, B. Jouaneton, and L. Lamarsalle. 2016. "Evaluation Of Discovered Clinical Pathways Using Process Mining And Joint Agent-Based Discrete-Event Simulation". In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, 2135–2146. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Cabitza, F., A. Locoro, and G. Banfi. 2018. "Machine Learning In Orthopedics: A Literature Review". *Frontiers in Bioengineering and Biotechnology* 6:75.

Dallagassa, M. R., C. Dos Santos Garcia, E. E. Scalabrin, S. O. Ioshii, and D. R. Carvalho. 2022. "Opportunities And Challenges For Applying Process Mining In Healthcare: A Systematic Mapping Study". *Journal of Ambient Intelligence and Humanized Computing* 13(1):165–182.

De Roock, E., and N. Martin. 2022. "Process Mining In Healthcare: An Updated Perspective On The State Of The Art". *Journal of Biomedical Informatics* 127:103995.

El-Bouri, R., T. Taylor, A. Youssef, T. Zhu, and D. A. Clifton. 2021. "Machine Learning In Patient Flow: A Review". *Progress in Biomedical Engineering* 3(2):022002.

Elbattah, M. 2019. "How Can Machine Learning Support The Practice Of Modeling And Simulation?—A Review And Directions For Future Research". In *Proceedings of the 23rd International Symposium on Distributed Simulation and Real Time Applications*, 1–7. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Elbattah, M., O. Molloy, and B. P. Zeigler. 2018. "Designing Care Pathways Using Simulation Modeling An Machine Learning". In *Proceedings of the 2018 Winter Simulation Conference*, edited by M. Rabe, A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, 1452–1463. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Gehlot, V., P. Rokowski, E. B. Sloane, and N. Wickramasinghe. 2022. "Taxonomy, Tools, And A Framework For Combining Simulation Models With AI/ML Models". In *Proceedings of the 2022 Annual Modeling and Simulation Conference*, 18–29. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Hinterwimmer, F., I. Lazic, S. Langer, C. Suren, F. Charitou, M. T. Hirschmann, G. Matziolis, F. Seidl, F. Pohlig, D. Rueckert, R. Burgkart, and R. Von Eisenhart-Rothe. 2023. "Prediction Of Complications And Surgery Duration In Primary TKA With High Accuracy Using Machine Learning With Arthroplasty-specific Data". *Knee Surgery, Sports Traumatology, Arthroscopy* 31(4):1323–1333.

Kempa-Liehr, A. W., C. Y.-C. Lin, R. Britten, D. Armstrong, J. Wallace, D. Mordaunt, and M. O'Sullivan. 2020. "Healthcare Pathway Discovery And Probabilistic Machine Learning". *International Journal of Medical Informatics* 137:104087.

Lopez, C. D., A. Gazgalis, V. Boddapati, R. P. Shah, H. J. Cooper, and J. A. Geller. 2021. "Artificial Learning And Machine Learning Decision Guidance Applications In Total Hip And Knee Arthroplasty: A Systematic Review". *Arthroplasty Today* 11:103–112.

Mesabbah, M., W. Abo-Hamad, and S. McKeever. 2019. "A Hybrid Process Mining Framework For Automated Simulation Modelling For Healthcare". In *Proceedings of the 2019 Winter Simulation Conference*, edited by N. Mustafee, K.-H. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, 1094–1102. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Mesabbah, M., and S. McKeever. 2018. "Presenting A Hybrid Processing Mining Framework For Automated Simulation Model Generation". In *Proceedings of the 2018 Winter Simulation Conference*, edited by M. Rabe, A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, 1370–1381. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Milella, F., L. Famiglini, G. Banfi, and F. Cabitza. 2022. "Application Of Machine Learning To Improve Appropriateness Of Treatment In An Orthopaedic Setting Of Personalized Medicine". *Journal of Personalized Medicine* 12(10):1706.

Mustafee, N., A. Harper, and B. S. Onggo. 2020. "Hybrid Modelling And Simulation (M&S): Driving Innovation In The Theory And Practice Of M&S". In *Proceedings of the 2020 Winter Simulation Conference*, edited by K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, 3140–3151. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Phan, R., V. Augusto, D. Martin, and M. Sarazin. 2019. "Clinical Pathway Analysis Using Process Mining And Discrete-Event Simulation: An Application To Incisional Hernia". In *Proceedings of the 2019 Winter Simulation Conference*, edited by N. Mustafee, K.-H. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, 1172–1183. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

AUTHOR BIOGRAPHIES

AHMED BAKALI EL KASSIMI is currently a Ph.D student at the Center for Health Engineering at Mines Saint-Etienne, FRANCE. He received a master's degree in advanced methods in industrial engineering from Ecole Centrale de Lyon, Auvergne-Rhone-Alpes in 2021 and an Industrial engineer degree from the Académie internationale Mohammed VI d'aviation civile in Casablanca, Morocco. His research interests include modelling, simulation and optimization of health-care systems. ahmed.bakali@emse.fr.

XIAOLAN XIE is a professor and the head of the department of Healthcare Engineering of the Center for Biomedical and Healthcare Engineering, Ecole Nationale Superieure des Mines (EMSE), Saint Etienne, France. His research interests include data analytics, design, planning and scheduling, supply chain optimization, and performance evaluation of healthcare and manufacturing systems. He is the author/coauthor of 300+ publications including 110+ journal articles and six books. He has rich industrial application experience with European industries and is PI for various national and international projects. Pr. Xie is a fellow of IEEE, the founding chair of the IEEE RAS-Technical Committee on Automation in Health Care Management and an IEEE RAS 2017 Distinguished Lecturer. He has been an editor/associate editor for IEEE Transactions on Automation Science & Engineering, IEEE Transaction on Automatic Control, IEEE Transactions on Robotics & Automation and International Journal of Production Research. He has been a Guest Editor of various special issues on healthcare engineering and manufacturing systems. He is a general chair of the 2021 IEEE Conference on Automation Science and Engineering to be held in Lyon, France. His email address is xie@emse.fr.

MARIANNE SARAZIN is a public health doctor with a doctorate in Life Sciences from Mines Saint-Etienne, France. She is the head of the Medical Information Department of the Mutualiste Sanitary Group of Saint-Etienne and a collaborator in the Center for Health Care Engineering at Mines Saint-Etienne. She is also with the UMRS 1136 Inserm lab, which specialized in the modelling of epidemics. Her email address is marianne.sarazin@iplesp.upmc.fr.

PIERRE LUC FREESARD is an orthopaedic surgeon at the "Clinique Mutualiste de Saint-Etienne" of the Aésio santé group and a member of its research team. He actively manages and launches projects to improve orthopaedic care pathways, including the Enhanced Recovery After Surgery program, the episode-based financing project, and the thesis project on modelling, simulation, and optimization of care pathways for patients undergoing hip or knee replacement surgery. His email address is pfresard@mutualite-loire.com.

BERTRAND SEMAY is an orthopedic surgeon at the "Clinique mutualiste de Saint-Etienne" of the Aésio Santé group, and a member of its research team. He leads several studies aimed at improving orthopaedic care pathways, developing personalized care pathways, and implementing value-based care. He actively participates in communicating and publishing research findings in these areas. His email address is bsemay@mutualite-loire.com.