IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 18 October 2024, accepted 5 November 2024, date of publication 18 November 2024, date of current version 27 November 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3501157

RESEARCH ARTICLE

Combining Process Mining and Process Simulation in Healthcare: A Literature Review

EVELYN SALAS[1], MICHAEL ARIAS[2], SANTIAGO AGUIRRE[3], AND ERIC ROJAS[4,5], (Member, IEEE)

[1]School Industrial Engineering, Universidad de Costa Rica, San José 111-4250, Costa Rica
[2]Department of Engineering, Computer Science and Technology, Universidad de Costa Rica, San José 111-4250, Costa Rica
[3]Business Information Systems Department, Central Michigan University, Mount Pleasant, MI 48859, USA
[4]Institute for Biological and Medical Engineering, Pontificia Universidad Católica de Chile, Santiago 8320165, Chile
[5]Department of Clinical Laboratories, School of Medicine, Pontificia Universidad Católica de Chile, Santiago 8320165, Chile

Corresponding author: Evelyn Salas (evelynmaria.salas@ucr.ac.cr)

This work was supported in part by the Agencia Nacional de Investigación y Desarrollo de Chile under Grant ANID Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT) 11230708, and in part by FONDECYT Project through "Improving the Patient Journey in Telemedicine using Process Mining" under Grant 11230708.

ABSTRACT Organizations are increasingly incorporating new strategies to ensure that the complex processes involved in the healthcare sector are comprehensively understood and enhanced. This includes the adoption of process mining (PM) and process simulation (PS), which have been used separately or in combination with one another in the healthcare field to assist decision-makers in process optimization. Although both PM and PS have provided a number of valuable contributions to healthcare, analysis in existing literature is lacking regarding analysis of the benefits, limitations, and tools derived from their combination thereof. The present article conducts a literature review, based on the PRISMA methodology, in which both disciplines are analyzed in terms of their application to healthcare. By reviewing distinct scholarly databases, 31 research studies were selected for analysis, from which it was possible to characterize case studies, techniques, tools, perspectives and algorithms, as well as to identify key limitations. The results indicate a stronger focus on medical fields including cardiology and emergency departments, with a preference for software tools, such as ProM, Disco, Arena, and CPN Tools. The use of real data predominates across the research studies, and the two most commonly identified and detailed limitations in the analysis relate to data quality issues and the involvement of healthcare experts throughout the analysis. Moreover, there is an increasing interest in the publications of papers on these topics within Latin America. Finally, through the findings of the present article have led the authors to propose several opportunities for future research. For example, the compilation of case studies in relation to medical fields that have been historically overlooked, and the availability of software that integrates PM and PS, in addition to the extent to which the usefulness of this combination may improve as a result in the field of healthcare.

INDEX TERMS Process mining, process simulation, healthcare, process analysis, clinical processes, organizational processes, business process management.

I. INTRODUCTION

As technology evolves, organizations are continuously seeking for methods and tools with which to adapt to the resulting changes. While this progress often strengthens various organizational aspects, it also presents ongoing challenges. As a result, novel tools and business structures are increasingly implemented, and new operations are pursued to ensure effective optimization [1]. The context has changed in recent years and the so-called Industry 4.0 emphasizes

The associate editor coordinating the review of this manuscript and approving it for publication was Dominik Strzalka.

digital transformation and new types of technology [2]. One of the disciplines that has contributed significantly to the adaptation of organizations is business process management (BPM), which has generated new opportunities for improvement through the design, execution, automation, control, measurement and analysis of processes [1], [3]. BPM consists of several methods, techniques and tools for decision-making and helps to integrate technology and data. Included within this mix of techniques is process mining (PM), a discipline that enables the discovery and optimization of processes by facilitating data-driven decision-making across different application domains. PM serves as a bridge between data science and process science [3], [4], [5]. Based on real process execution data, known as event logs [6], PM can discover, monitor, analyze and, consequently, improve processes. PM algorithms applied to event logs enable the identification and analysis of real behaviors, thus allowing the detection of patterns, trends, bottlenecks and deviations in processes. This contributes to time and cost savings while simultaneously facilitating the identification of the root causes of inefficiencies and digressions, among other improvements [5], [7], [8]. Process mining has proven to be valuable not only in healthcare, but also in multiple sectors, including finance, public institutions, manufacturing and services, among others [3]. It should be noted that PM has been applied across several areas of the healthcare sphere, using diverse data, techniques and scenarios [9]. Its application has achieved positive results and continues to show enormous potential for the healthcare sector in terms of decision-making, identification of patient flows, possible treatments, quality control, cost reduction, resource management, determination of patterns and performance analysis, among others areas [9], [10], [11]. This is driving efficiency in the management and quality of healthcare services as well as an improved patient experience [12]. As an alternative for optimizing processes, PM has been combined with other tools and disciplines, including process simulation (PS). Simulation establishes scenarios with which to determine the state or behavior of a process or system, thereby allowing for its analysis and improvement [13].

In other words, PS seeks to replicate the behavior of a process in order to evaluate its performance and detect possible opportunities for improvement [14]. Accordingly, evaluations of process changes can be performed and decisions can be made prior to their implementation [11] and [15].

In the literature, PS has been applied by means of the use of synthetic and real data. Synthetic data refers to data that have been generated using a specifically designed mathematical model or algorithm, for the purpose of solving a series of data science tasks [16]. In the case of real data ([13], [17], [18]), information obtained directly from information systems is used. However, the latter approach remains a challenge, due to the different sources of information available with which to extract the data and data privacy implications [8]. In seeking to address this challenge, the combination with disciplines such as PM can prove useful in obtaining valuable knowledge from data derived from information systems [5]. In addition, PM can improve PS by creating more accurate models and more realistic behaviors [7].

Both PM and PS have been applied in the field of healthcare. According to [19], PS has been used in the sector since prior to the first simulation conference in 1952. Since then it has afforded the sphere a number of new opportunities. These include helping to determine the number of beds required in surgery and emergency departments [17], as well as measuring indicators in organizational departments or healthcare centers, for example, establishing the overall time a patient spends in a particular hospital department, defining how many times different activities are carried out for a patient, ascertaining the number of people receiving care, and the cost of the entire process, among others [11].

Although simulation may be costly and time-consuming, PM can help to leverage the use of real data [17]. Therefore, it directly contributes to a growing potential in the area whereby healthcare process owners can ensure more accurate, rapid and substantive decisions making [10], while simultaneously allowing them to optimize patient flow and plan scenarios prior to making their decision [14].

Several examples from the literature, as discussed in the present paper, present case studies in which PM and PS are integrated. One notable example is [14], where a Data-Driven Process Simulation (DDPS) was applied to a real-life case study in the radiology department of a Belgian hospital. The work provided concrete recommendations regarding the required number of radiology devices, the optimal size of the waiting area, and the staffing needs at the reception.

At least two further case studies identified in the literature review involved the development of designs for physical spaces through the combined use of PM and PS. One case focused on a cardiovascular and cardiac clinic [20], while the other centered on the emergency department of a hospital [21]. In the former, the authors sought to design a layout that aligned with patient demands and clinical operations, minimized patient travel distances, and upgraded space sizes. In the latter, the focus was placed on optimizing patient flow, staff numbers, and the presence of external individuals to identify the least invasive interventions in building design and space organization, thus enhancing the overall use of space.

Finally, in the case studies presented in [22], the aim was to assist healthcare organizations with the automatic generation of staff rotation plans by integrating optimization methods using PM techniques.

With that in mind, the present paper aims to characterize the case studies in which PM and PS have been combined in the field of healthcare, providing an overview of the current state of affairs with relation to how these two disciplines can be deployed in tandem. The objective is to help future researchers visualize limitations when applying

the techniques and enhance their understanding of the most prominent methodologies, algorithms and tools.

The article is organized as follows: Section II establishes the research methodology applied by the authors; Section III provides the results in relation to the research questions posed; Section IV presents the main discussion, the findings, and opportunities for future research; Section V outlines the limitations of the study; and, finally, Section VI sets out its main conclusions and Section VII list the Acknowledgements.

II. METHODOLOGY

A. PRISMA METHOD

There area several methods available for conducting a literature review exist, such as those proposed by Kitchenham and Charters [23] and Petersen [24]. For the present study, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method was selected, and the authors also incorporated the Covidence tool, on which this method that is based. PRISMA was primarily designed for health intervention studies [25], though it can also be applied in other fields. It dates back to 2009 and was most recently updated in 2022. It has been used to conduct literature reviews in a transparent and strict manner, with an emphasis on specifying how the process is being executed. Hence, it serves as a guide for presenting publications and elucidates the methods used in identifying, selecting, evaluating, and synthesizing studies. As a result,it enables researchers to select, evaluate and analyze studies and to report, in an in-depth manner, on how this was achieved [25].

The use of Covidence software has broadened knowledge in the area by offording users a wide range of new opportunities. It is considered a user-friendly resource that enhances the methodological process by increasing rigor as well as reducing bias thanks to the independent work of each person participating in the research. The software is compatible with Mendeley and the download formats of the virtual libraries used in the present study.

B. RESEARCH OBJECTIVES

The objectives of the present literature review are as follows:

Objective 1: To identify existing case studies in which PM and PS in healthcare are combined.

Objective 2: To characterize existing case studies, including a description of the most significant aspects, such as techniques applied, tools used and types of data, among others.

C. RESEARCH QUESTIONS

Based on the application of a rigorous literature review methodology, the following research questions are analyzed in the present paper:

1. Are there any published case studies in which process mining and process simulation have been combined in healthcare?

2. What are the most commonly used techniques and tools in the case studies involving both PM and PS in healthcare?

3. What are the main characteristics of the case studies in which PM and PS have been combined in healthcare and what were the main results?

4. What limitations and future trends can be derived from the case studies in which PM and PS have been combined in healthcare?

D. BIBLIOGRAPHIC SEARCH PROCESS

First, keywords were established. The authors conducted an exploratory search in PubMed, Google Scholar, Scopus and ScienceDirect, using the keywords: ''Process Mining'' & ''Process Simulation'', ''Process Simulation'' & ''Healthcare'', ''Process Mining'' & ''Healthcare'' and ''Process Mining'' & ''Process Simulation'' & ''Healthcare''. Subsequently, the decision was taken to prioritize the following scholar databases: Google Scholar, Scopus and PubMed, since they offered a greater number of studies for analysis purposes. In the case of ScienceDirect, the studies were already included in certain other libraries, whereas in the Google Scholar database, patents and citations were excluded. Therefore, the following search sequence was created:

''*Process Mining*'' & ''*Process Simulation*'' & ''*Healthcare*'' OR ''*Process Mining*'' & ''*Process Simulation*'' & ''*Health*'' *OR* ''*Process Mining*'' & ''*Simulation*'' & ''*Healthcare*'' OR ''*Process Mining*'' & ''*Simulation*'' &''*Health*''.

Using this search sequence, a total of 475 studies were extracted, of which 352 were from Google Scholar, 115 were from Scopus and 8 were from PubMed. The studies were entered into Covidence to perform the first analysis, corresponding to the abstract and title.

E. INCLUSION CRITERIA

The inclusion criteria used were as follows:

IC1. Only documents written in English.

IC2. Must contain the terms ''process mining'', ''process simulation'', ''healthcare'' or combinations thereof in the title and keywords.

IC3. Only journal articles, conference papers, books, theses and workshops.

IC4. Original and full-text papers, published between 2005 and February 20, 2023.

The search process and results have been summarized using a PRISMA diagram, shown in Figure 1.

F. EXCLUSION CRITERIA

The following types of papers were excluded:

EC1. Research that lacked a combination of PM and PS in healthcare.

EC2. Research that was not written in English.

EC3. Papers whose complete text was not available for public access or through digital libraries.

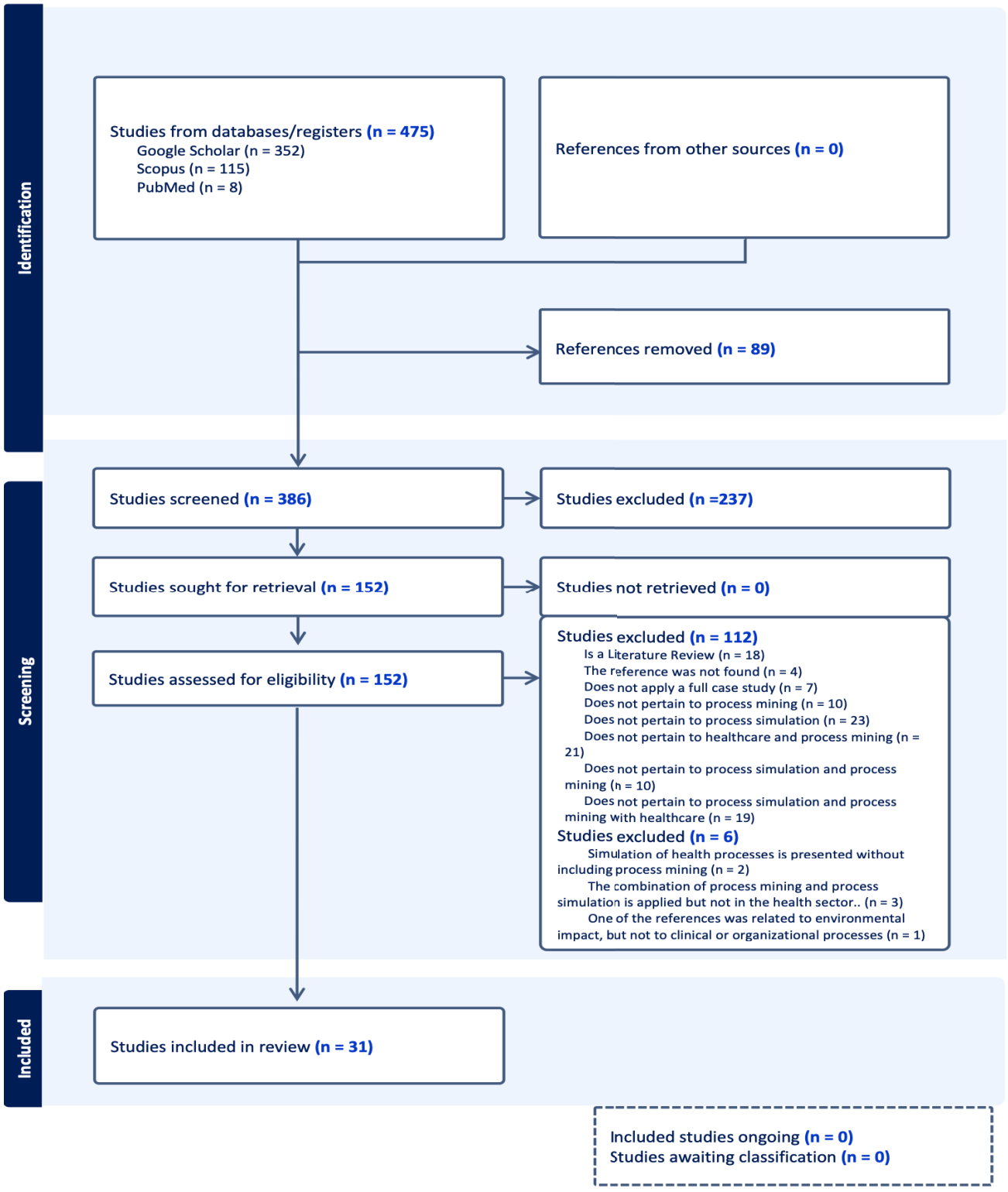

FIGURE 1. PRISMA protocol for the process mining and process simulation in healthcare systematic review.

EC4. Research that did not provide a review of the literature used.

EC5. Research that did not include a complete case study.

EC6. Research that did not pertain to PM.

EC7. Research that did not pertain to PS.

EC8. Research that did not pertain to healthcare and PM.

EC9. Research that did not pertain to PS and PM.

G. STUDY SELECTION PROCESS

The following subsection sets out to introduce the interdisciplinary team that participated in the present research. The team consists of four people, as follows: the main researcher, who has experience in PS; the second researcher, who is an expert in PM; the third researcher, who is an industrial engineer and, in conjunction with the main researcher, has extensive experience in BPM and PM and the fourth researcher, who is an expert in clinical informatics and conducts research into PM.

To reiterate, the methodology used in the present study is PRISMA [25]. This protocol consists of four main stages: I) First selection, by means of the review of key words and research document summaries; II) Second selection, by means of the review of additional sections of each document, for example the introduction, discussion and conclusions; III) Selection, consisting of a comprehensive review of the research studies included in the study, or primary studies; and finally, IV) Data extraction and data analysis.

In addition, the Covidence[1] software tool is used as a support. This tool ensures that analysis is strict because it helps to reduce bias, since individuals who are reviewing each document are unable to identify who exactly is analyzing each reference made; in addition, each reviewer works in an independent manner and remotely.

Using Covidence, reviewers can able to read the abstracts and titles of each paper and rate whether that study ''complies'', ''does not comply'' or ''maybe complies'' with the criteria. In the case of any differences in criteria, at least two reviewers review the studies in conflict, they discuss them and subsequently come to a joint decision. If the conflict persists between the reviewers, the entire team meets to decide whether to include or exclude the document.

With regard to the present paper, the first analysis to select the studies consisted of determining whether the document met the inclusion criteria by reading the title and abstract. When all the studies were entered into Covidence, it automatically discarded 68 documents from Scopus and 17 from Google Scholar due to being duplicates, while 3 documents were manually excluded, also for also being duplicates, resulting in a total of 89 duplicates being eliminated. The entire interdisciplinary team participated in this first step.

Once the studies had been selected, a the second analysis took place, whereby three researchers conducted a review of, and decided whether to include or exclude each study. The same classification option was used in Covidence. However, in this second step, the analysis of the studies was more extensive, since it covered the introduction, discussion and conclusions sections of the paper, as well as the abstract.

[1]https://app.covidence.org

After completing the second analysis and selecting the primary studies, a final analysis was conducted. This involved a full examination of studies and the extraction of relevant data. From this comprehension, it was determined that five of the documents failed to meet the inclusion criteria. This was due to the fact that only healthcare PS was addressed in two of the cases, i.e., it did not include PM, while in the other three, the combination of PM and PS is applied, but not in the healthcare sector. A total of 444 studies were excluded, 237 documents were eliminated from the first analysis and a further 112 from the second, plus the last 6 following the complete reading of the documents. Therefore, the remaining 31 documents form the basis of the present study. The references for these documents can be found in Appendix, Table 3.

H. DATA COLLECTION

One of the authors of the present paper extracted data from the studies used in this review. These were classified according to medical specialty, process, health facility, type of approach, country of publication, type of publication, data used, research objective, limitations encountered, date of data, type of PM used, type of PS used, type of PM algorithm used, and types of PM and PS computer tools used. This classification was subsequently reviewed by all authors.

I. QUALITY ASSESSMENT

For the quality assessment, an initial exploratory search was performed across several scholarly databases. The search was reviewed by all authors in order to determine the scope of the study. This approach allowed the review to be undertaken in line with a robust protocol with blinded analysis, whereby the reading of the titles and abstracts as well as the second analysis of the studies were performed by three of the authors. Any disagreement with the inclusion/exclusion of a case study was resolved via discussion. Full text reading and data extraction were conducted by one author and subsequent data classification was reviewed by three of the authors.

III. RESULTS

The following section outlines the key characteristics of the 31 publications included in the present literature review. Section III-A presents the objectives identified in the case studies. Sections III-B and III-C categorize the studies according to type of process involved and type of data collected. Section III-D identifies the types of PM used and provides a description of the PM tools and techniques deployed. Section III-E introduces the PS tools and techniques used, while Section III-F assesses the most commonly used tools in the combination of PM and PS in healthcare field. Section III-G establishes a geographical classification of the case studies, the predominant publication types and the most common years of publication and in section III-H details an analysis of the methods or methodologies used in the 31 research studies. Finally, Section III-I illustrates certain

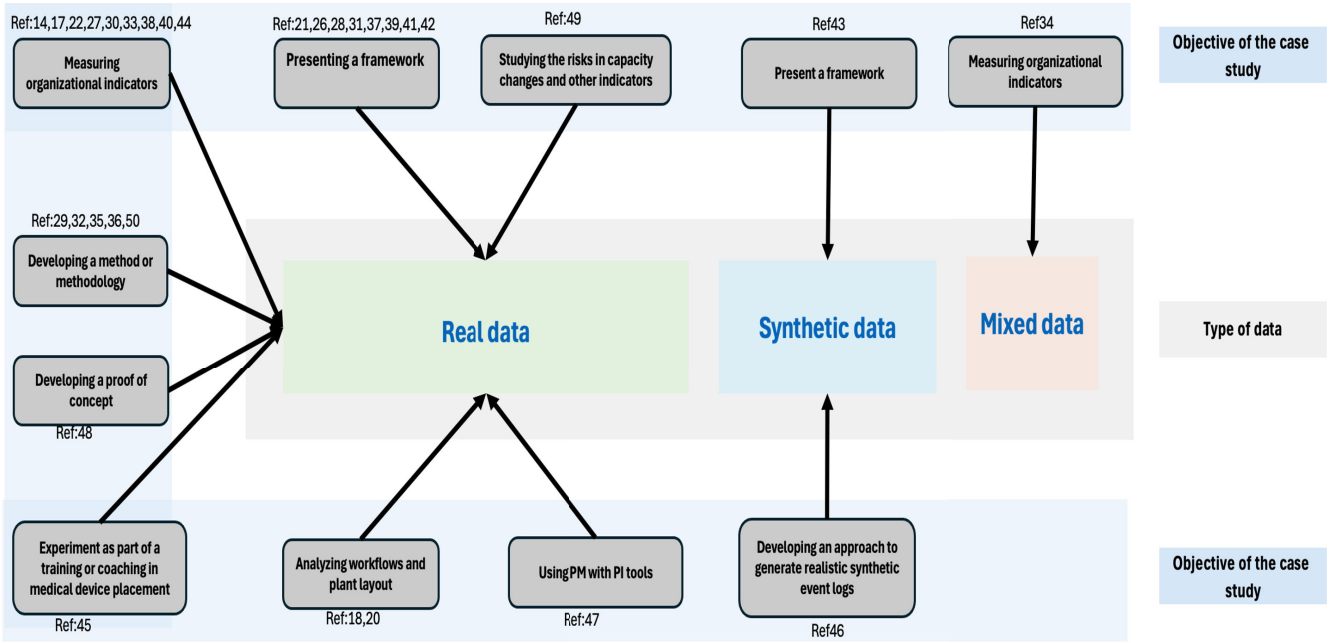

FIGURE 2. Association between the objective of the case study and the type of data used.

limitations and challenges encountered in several of the case reviewed.

A. OBJECTIVES POSED IN THE CASE STUDIES

The results show that a significant number of cases (10 of the papers, or 32%), the sources reviewed focused on combining PM and PS in healthcare in order to analyze, study and improve organizational indicators. For example, indicators to evaluate constraints to enhance the performance of operating theaters [18] and [26]; to address waiting times in wards pertaining of different medical specialties ([27], [28], [29], [30], [31], [32]); to determine total amounts of equipment and resources ([14], [17], [20], [30], [31], [33]) and how these amounts may affect patient treatment waiting times; to determine the number of staff required per day and per hour to ensure that the number of patient treated is unaffected ([14], [18], [21], [26], [28], [29], [30], [31], [32], [34], [35], [36]); and an indicator to study patient flow as a means to define changes in demand and capacity ([20], [22], [27], [28], [31], [32], [33], [35], [37], [38]). Conversely, a further five case studies (16% of the total) focus primarily on improving simulations by incorporating real data with PM ([39], [40], [41], [42], [43]). Finally, seven articles (22,6%) provide specific contributions across various healthcare fields, which are detailed in the following two paragraphs.

The first article [44] relates to an application used in the process of stroke treatment, where it is not feasible to perform real-life trials since such trials may place the lives of patients at risk. In this particular case, a hybrid PS-PM approach is proposed. The second article [45] discusses a control flow analysis using PM techniques in the training of ultrasound-guided central venous catheter placement in the internal jugular vein by means of a simulation. In the third article [46], the objective was to generate synthetic event logs from Petri nets. However, the authors indicate that it was not possible to fulfill this objective since synthetic event logs contain a number of unrealistic features.

The fourth article [47] focused on the integration points between PM and traditional process improvement (PI) tools, including value stream mapping, spaghetti diagrams, failure effects analyses and suppliers, inputs, process, outputs, customers (SIPOC), control charts and discrete-event simulations. In this particular case, the analysis focused on determining whether there is added value to the process improvement tool and whether the use of PM is necessary to implement these tools.

The fifth article that offers a specific contribution is [48], which focuses, in part, on using standards to access audit trails from healthcare information systems and identifying who provides automated mapping to an event log format that is suitable for PM. Regarding the final two case studies, paper [49] establishes the advantages and disadvantages of discrete-event and agent-based modeling approaches in healthcare using PM; while paper [50] defines its research objective as to investigate the effects of digital dentistry on the implant value chain using PM and discrete-event simulation.

Figure 2 outlines a classification of the objectives established in the case studies included in the present literature review and their association with the type of data used by the authors (real, synthetic or mixed). Overall, nine distinct categories are defined in Figure 2.

From left to right and from top to bottom, the aforementioned are as follows:1. Measuring organizational indicators, 2. Presenting a framework, 3. Studying the risks in capacity changes and other indicators, 4. Developing a method

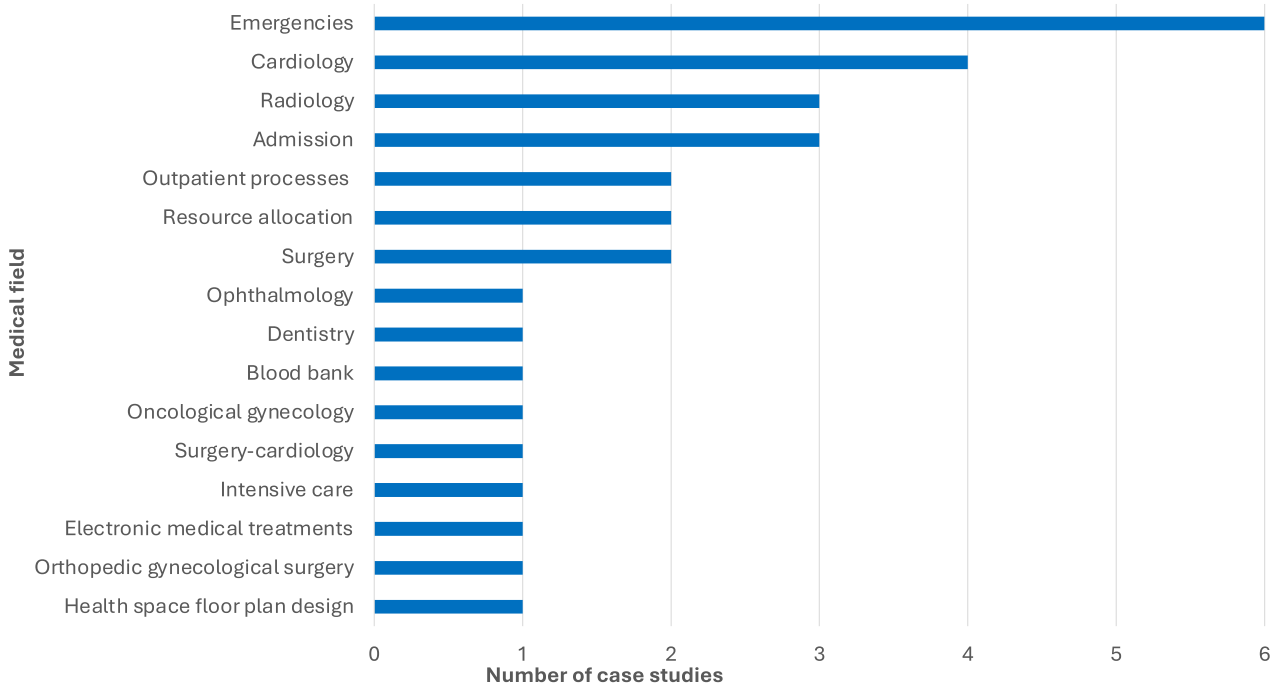

FIGURE 3. PM and PS case studies per Medical specialty / field.

or methodology, 5. Developing a proof of concept, 6. Experiment as part of a training or coaching in medical device placement, 7. Analyzing workflows and plant layout, 8. Using PM with PI tools and 9. Developing an approach to generate realistic synthetic event logs. Regarding the type of data used in these articles in relation to their objectives, two studies utilize synthetic data. Of these, one presents a framework and while the other develops an approach to generate realistic synthetic event logs. In the only publication which incorporates mixed data, the objective was to measure organizational indicators.

More detailed information regarding the objectives, results, and contributions made by the authors of each publication can be found in Figure 7.

B. TYPE OF PROCESS, MEDICAL FIELD AND HEALTHCARE FACILITIES

Figure 3 provides a complete list of medical specialties and fields covered by the case studies included in the present literature review. Processes related to medical or clinical treatments and organizational processes are differentiated. The medical fields that feature most prominently include emergencies (i.e.,in the emergency department) ([22], [28], [31], [39], [42], [47]) and cardiology ([20], [33], [35], [44]), which are present in 19% and 13% of the cases, respectively. Similarly, three case studies conduct investigations into the field of radiology ([14], [17], [48]), representing 10% of the papers reviewed. With respect to organizational processes, the admissions procedure is covered in four case studies ([27], [37], [38], [40]).

The processes of admissions, resource allocation and health space floor plan design [18] are categorized as organizational, whereas all others are classified as clinical. In one case study, the subject of research relates to a generic model [43], therefore, a specific related process is not identified therein. A majority of the case studies (77%) were conducted in hospitals, and one was conducted in a university hospital [29], thus highlighting its relevance not only in terms of healthcare, but also research and academia. A further study was performed in a dental clinic and laboratory [50]. In five of the cases (18%) the hospital in which the research took place was public ([33], [35], [39], [40], [49]) while one additional case study was undertaken in a private hospital [30].

C. TYPE OF DATA

Regarding the type of data used, real data is utilized in 28 of the 31 papers reviewed ([14], [17], [18], [20], [21], [22], [26], [27], [28], [29], [30], [31], [32], [33], [35], [36], [37], [38], [40], [41], [42], [44], [45], [47], [48], [49], and [50]), which represents 90% of all research papers studied. Two of the case studies (8%) use synthetic data [43] and [46], while one implements a mixed use, i.e., both synthetic and real data [34]. Regarding the case studies that use real data, 74% (24 publications) relate to research applied to medical processes and 14.5% to administrative processes in healthcare ([21], [27], [34], [40], and [41]). The study in [43] implements

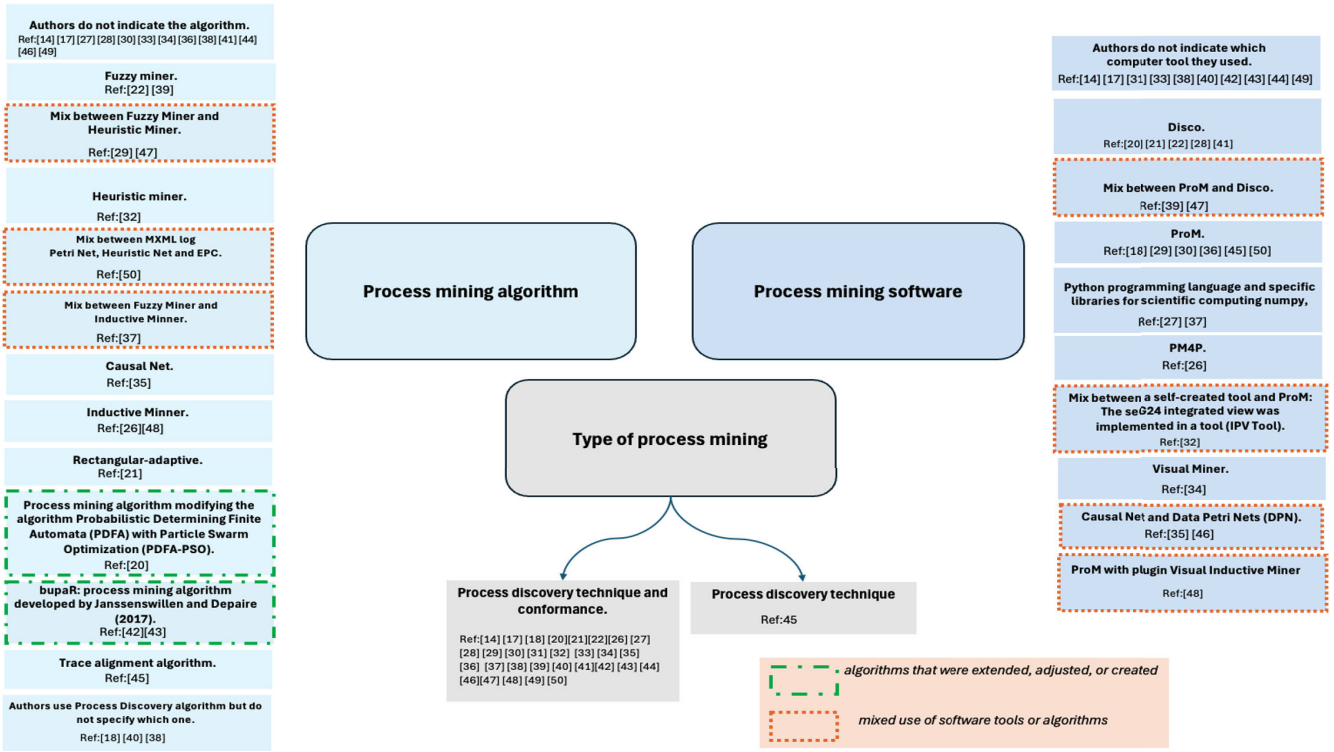

FIGURE 4. Process mining types, tools and techniques in PM and PS case studies.

a mixed use, also in an administrative process setting, while the remaining case relates to a study that aims to generate synthetic data [46].

D. PROCESS MINING TYPES, TOOLS AND TECHNIQUES

By analyzing the papers included in the present review, it is evident that the predominant type of PM applied in the integration of PM and PS in healthcare is process discovery. In 97% of the cases reviewed, this particular type of PM is used to determine activity flow, i.e., to work in conjunction with the PS. It should be noted that there is only one study that combines the discovery and compliance types [45].

Regarding the most frequently identified PM tools in the studies, 51% of the cases were found to use discovery algorithms for PM analysis. Nevertheless, none of these studies mentions the type of algorithm is deployed.

In contrast, regarding cases where the algorithm used is specified, the most frequently applied variants include Fuzzy Miner[2], Heuristic Miner[3], Petri Net[4], and Causal Net[5]. In three other cases ([29], [47], [50]) a mixture of algorithms was used (MXML Log Petri Net, Heuristic Net and Event-driven Process Chain [EPC]). In addition, in one case study [20], the authors modify the algorithm proposed by relates to a medical process. Regarding the two case studies in which synthetic data are used, one is applied farouq Halawa et al. [20], which is the Probabilistic Deterministic Finite Automata (PDFA) algorithm, and explain how they modify this with Particle Swarm Optimization (PDFA-PSO).

The widespread use of free software throughout the case studies,is evident, since the most commonly used tools are ProM and Disco. This is likely due to the ease of their acquisition and the fact that many of these free packages allow a simple execution of multiples analysis, which can facilitate the identification of improvement opportunities. The details of these types, tools and techniques are shown in Figure 4. The most commonly used software was ProM[6], which was present in eight of the case studies (25.8%) ([18], [29], [30], [32], [36], [45], [48], [50]). The second most common was Disco[7], found in five of the studies (16%) ([20], [21], [22], [28], [41]). However,it should be noted that the number of case studies in which the tools used are not indicated is high (In 38% of the cases), with regard to both PM and PS analysis, this analysis type will be presented bellow.

In Figure 4 the case studies in which a mixture of algorithms were used can also be seen, highlighted with an orange border and in small dotted line. The algorithms that were extended, adjusted or created are highlighted with a green border and in larger dotted line. In the left-hand column, the algorithms used in the referenced studies are presented. In the right-hand column, the corresponding

[2]https://processmining.org/old-version/fuzzy.html
[3]https://mindzie.com/2022/12/15/types-of-process-mining-algorithms/
[4]https://www.sciencedirect.com/topics/physics-and-astronomy/petri-net
[5]https://causenet.org
[6]http://www.promtools.org
[7]https://www.processmining-software.com/tools/disco/

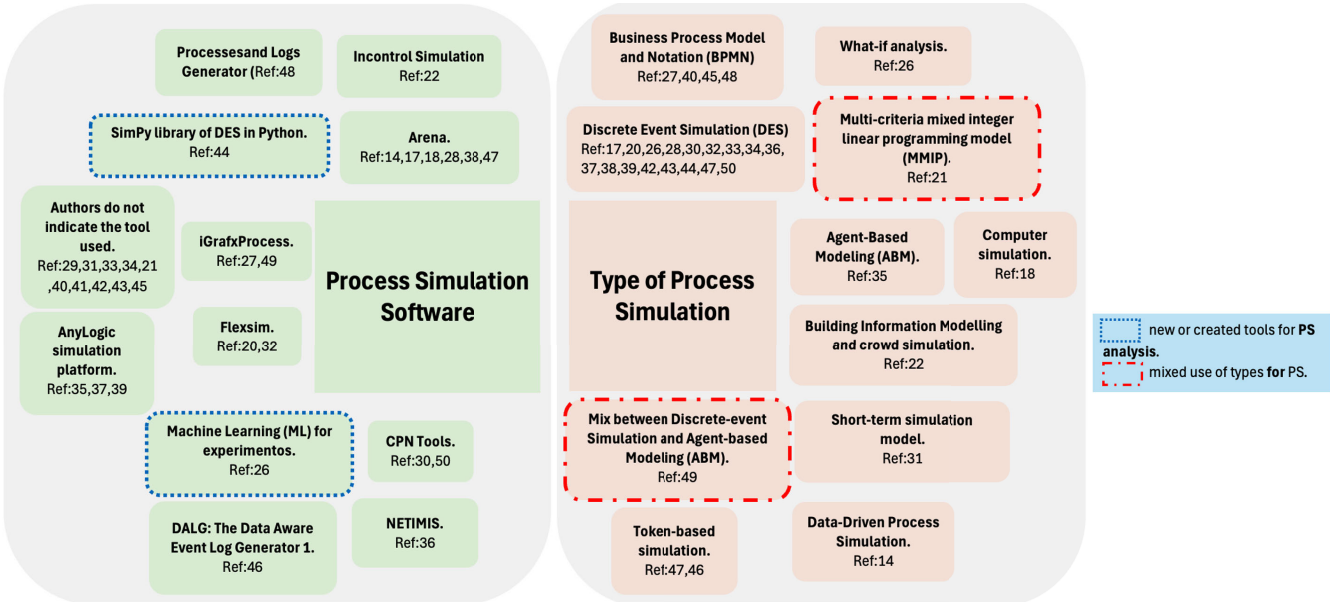

FIGURE 5. Process simulation tools and techniques applied in the PM and PS case studies.

computer tools or software packages employed are shown. In the lower-center of Figure 4, the types of PM techniques identified in the case studies are listed.

E. PROCESS SIMULATION TOOLS AND TECHNIQUES

Regarding tools tools and techniques used for PS analysis, 54% of the case studies apply discrete-event simulation (DES). In one paper [35], agent-based modeling (ABM) was used, while in another [49] a combination of these two types of simulation was employed. Furthermore, one case study uses Building Information Modeling and crowd simulation [22].

The most commonly used software for simulation analysis is Arena[8], which was applied in six of the case studies (19%) ([14], [17], [18], [28], [38], [47]). In certain other papers, CPN Tools[9] and the AnyLogic simulation platform[10] are used, albeit less frequently. Furthermore, 38% of the case studies reviewed do not indicate the software tool used.

As can be seen, specialized PS software is used to develop analyses in this discipline. The case studies use tools designed for this purpose and subsequently complement them with the implemented PM analysis separately. Therefore, indications are that PS software is able to adapt well to the use of real data provided from real case studies.

Figure 5 illustrates the range of PS tools used in the case studies. Those that employ a combination of simulation types are highlighted with a red border and in larger dotted line, while tools classified as new or custom-developed implementations are marked in blue border[11] and in small dotted line. On the righthand side, the types of simulation identified in the primary references are depicted. On the lefthand side, the tools or software used to develop the simulations are shown.

F. COMPARATIVE ANALYSIS OF PROCESS MINING AND PROCESS SIMULATION TOOLS

Following analysis, a clear pattern was identified in the use of PM and PS tools for the carrying out of specific studies. For example, regarding [30] and [50], in which process discovery is applied with DES, both cases use ProM and CPN Tools and the two sets of research are performed in the gynecological oncology care process and in dentistry, respectively. A similar pattern is observed with regard to studies ([20], [28], [32], [47]), in that all four cases apply process discovery with DES. In two, Disco is used [20] and [28], in one ProM is employed [32], and in the other, a mixture of Disco and ProM is applied [47]. Of these four cases, two were performed in the emergency department [28] and [47], one in ophthalmology and the other in a vascular and cardiac clinic. The final two cases in which similarities can be observed are [35] and [40], which apply process discovery by means of the AnyLogic simulation platform software in order to perform process simulation. In both cases, DES is used, although in study [35] it is combined with a multi-Agent simulation. Of these two latter cases, one is performed in the area of cardiology [35] while the other in the admissions process [40].

The aforementioned reflects, in addition to what can be seen in Figure 4 and 5, how recent research into the combination of PM and PS in the healthcare sector is, to a certain extent, favoring the use of process discovery in conjunction with discrete-event simulation, while also using ProM, Disco, or a mixture of both. In the case of software

[8]https://www.rockwellautomation.com/en-us/products/software/arena-simulation.html
[9]https://cpntools.org/
[10]https://www.anylogic.com/
[11]https://www.igrafx.com/products/process-design/

TABLE 1. Comparative analysis of PM and PS tools.

Article	PM Type	PS Type	PM Software	PS Software	Health Field	Ref
Simulation to analyze the impact of a schedule-aware workflow management system.	Discovery	Discrete Event Simulation (DES)	ProM	CPN Tools	Gynecological oncology healthcare process	[30]
A Process-oriented Methodology for Evaluating the Impact of IT: a Proposal and an Application in Healthcare.	Discovery	Discrete Event Simulation (DES)	ProM	CPN Tools	Dental	[50]
A Solution Framework Based on Process Mining, Optimization, and Discrete-Event Simulation to Improve Queue Performance in an Emergency Department.	Discovery	Discrete Event Simulation (DES)	Disco	Arena	Emergency department	[28]
Validation of Flexsim HC simulation models using an integrated view of process mining results.	Discovery	Discrete Event Simulation (DES)	ProM	Flexsim	Ophthalmology	[32]
Integrated framework of process mining and simulation–optimization for pod structured clinical layout design.	Discovery	Discrete Event Simulation (DES)	Disco	Flexsim	Vascular and cardiac clinic	[20]
Process mining to facilitate process improvement in a healthcare environment: An emergency department case study.	Discovery	Discrete Event Simulation (DES)	ProM and Disco	Arena	Emergency department	[47]
Evaluation of discovered clinical pathways using process mining and joint agent-based discrete-event simulation.	Discovery	Multi-Agente and DES	ProM and not indicated	AnyLogic simulation platform	Cardiology	[35]
A Process Mining Application for the Analysis of Hospital-at-Home Admissions.	Discovery	Discrete Event Simulation (DES)	ProM and Disco	AnyLogic simulation platform	Admission process	[40]

TABLE 2. Summary of primary studies by publishing vehicle and year of publication.

Vehicle	2009	2010	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Total	
Journal	1	1	1	-	-	-	1	1	-	2	1	2	1	**11**	35%
Conference	-	-	-	-	1	1	1	1	5	-	-	4	1	**14**	45%
Workshop	-	-	-	-	-	-	-	1	-	-	2	-	-	**3**	10%
Master thesis	-	-	1	1	-	-	-	-	-	-	-	1	-	**3**	10%
Articles per year	**1**	**1**	**2**	**1**	**1**	**1**	**2**	**3**	**5**	**2**	**3**	**7**	**2**	**31**	**100%**

for simulation, differences are found between the studies. However, the most commonly mentioned examples include Arena, CPN Tools, FlexSim[12] and the AnyLogic simulation platform. Table 1 outlines this comparison in further detail.

G. GEOGRAPHICAL ANALYSIS, TYPE AND YEAR OF PUBLICATION

A majority of the studies reviewed (20 out of 31, or 64%) were published in Europe. Italy is the country in which most primary studies have been published, with five papers in total ([27], [36], [40], [44], [49]. Notably, research combining these disciplines is increasing in the Americas, where seven of the reviewed paper were developed and published. Of these, five are from Latin America, with Brazil contributing four$ [26], [28], [31], [44]) and Chile one [32]. Figure 6 provides a breakdown of the articles evaluated by country of origin.

Regarding the type of publications reviewed, there was an inclination was found towards articles published in journals, with 11 cases in total (35%) ([14], [18], [20], [30], [33], [34], [40], [41], [44], [45], [50]), as well as conferences, with 14 (45 %) ([26], [27], [29], [31], [21], [22], [35], [36], [37], [38], [39], [42], [43], [49]). In addition, three relate to

[12]https://www.flexsim.com/es/

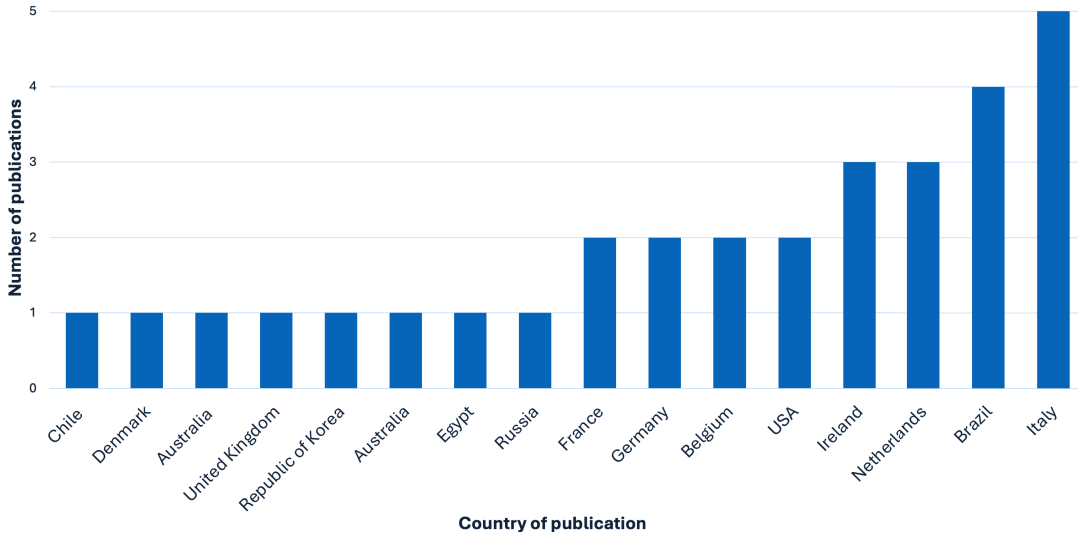

FIGURE 6. PM and PS case studies per country of publication.

research presented in workshops (10%) ([17], [28], [48]), and three more (10%) are Master's theses ([32], [46], [47]). Table 2 shows the number and types of publications used in the present paper.

It should be noted that research into the combination of PM and PS in the healthcare sector has been gradually increasing in recent years, with 17 articles being published between 2019 and 2022. The data in Table 2 reflects the growth in research on these topics over the years. For example, in 2009, only one case study was published, compared to seven in 2022.

H. METHODS OR METHODOLOGIES

One of the key findings of the present study concerns the diversity observed among the methodologies used across the 31 case studies. While some exhibit similarities in their procedural stages, they tend to represent methodological frameworks proposed by the respective authors. Only two references explicitly cite a methodological model: [49] which employs the traditional life cycle of business processes, and also [50] which follows a process-oriented methodology to evaluate the impact of IT.

In a broader context, and in attempt to identify commonalities in methodological approaches, it can be seen that in 11 of the case studies begin with an analysis of the respective context or case in question ([14], [17], [18], [20], [27], [30], [32], [36], [40], [46], and [49]). Conversely, five studies propose, primarily, the development of a simulated process model ([14], [39], [47], [48], and [50]). However, the predominant approach, adopted by 24 of the case studies, involves the establishment of a process model through PM as the initial step, followed by the utilization of this data for the subsequent PS model ([17], [18], [20], [21], [22], [26], [27], [28], [29], [30], [31], [33], [34], [35], [36], [37], [40], [41], [42], [43], [44], [45], [46], and [49]). Regarding the stages of extraction, data analysis, and PM, 25 references explicitly detail the aforementioned step as part of their methodology([18], [21], [26], [27], [28], [29], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [48], [49] y [50]). On the other hand, only 10 studies explicitly mention the use of statistical analysis for PS.

Finally, 28 studies exhibit similarity in their final methodological stages, whereby parameters, scenarios, or performance indicators are defined for analysis within the simulation subsequent, to the simulation of the process. Subsequently, the simulation model is validated, and final analyses are conducted. It should be noted that this particular framework is not explicitly followed in three of the case studies reviewed, in which validation of the simulation is not mentioned ([22], [29], and [40]).

I. LIMITATIONS OR CHALLENGES IDENTIFIED IN THE CASE STUDIES

The main limitations identified in the case studies relate to problems with data quality (in 13% of the references used) ([14], [17], [36], [44]), the availability of data in the healthcare sector and the different sources from which this data is extracted (in 6,5%) ([14], [20]), as well as the fact that certain of these sources are unrelated to one another [34]. The main challenges identified include the lack of a PM algorithm with which to obtain the desired information [50]; the difficulties encountered in replicating or generalizing the methods used due to, for example, limited data, among other aspects [18]; a low number of patients [32], [37]; and how tools, such as FlexSim HC [32] were used, thus leading to the replication of results becoming limited to the use of that particular software.

Futhermore, analysis shows that the results may not be generalizable with regard to other countries, hospitals or departments due to data characteristics and cultural aspects,

among other elements [18]. Finally, one of the frequently mentioned limitations (in 21% of the case studies) is the need to incorporate stakeholders or experts in the field into the team managing the PS ([17], [21], [32], [33], [36], [39]) in order to ensure a more robust result validation process.

Reference [14] suggest that in order to address certain identified limitations, additional research is needed to develop tools capable of facilitating the way in which healthcare administrators are able to harness the benefits of Data-Driven Process Simulation (DDPS) analysis. Simultaneously, they indicate that data quality can be improved by facilitating accurate data collection by using intuitive and straightforward information system interfaces or by adopting technology that supports data capture. In [17], authors highlight the need to develop a framework that conceptualizes experience recognition in the healthcare sector, among other requirements, in order to validate simulation results.

On the other hand, [18] states that the limitation of generalizing the study in such a way such that it can be applied to other healthcare centers, departments or countries can be addressed by applying the described analytical processes to other departments. Subsequently, the paper argues that the results can be compared between several hospital settings, both within and outside the country of origin, which in this case is Germany. This case study also suggest the integration of analysis of other functions, such as support functions or additional ancillary departments.

Similarly, [35] the authors suggest certain potential contributions with which to address limitations and possible future research, including the introduction of more parameters to describe the health status of a patient, and additional decisions related to their care process. Reference [21] expresses a desire to implement the proposed approach, albeit by means of discrete-event simulation rather than the multi-criteria mixed-integer linear programming (MILP) model, in order to achieve schedule optimization during its execution. To improve data quality, [36] performed a complete audit of all data in the hospital systems and used clinical charts and discharge notes for each patient, by means of utilizing all available sources (including telephone interviews with treating physicians to fill in missing data). Also in relation to data quality, [39] state how errors or inaccuracies in data records can be reconciled by cross-checking traditional data sources, and that validation should be performed prior to any further analysis.

Finally, [46] discuss the challenges presented by errors in real event logs, such as false sensor readings. Their findings include the proposal to conduct further research on the generation of event logs with intentional errors. They also suggest investigating whether other types of PM, such as those based on BPMN notation, could provide the necessary accuracy to generate realistic data.

In order to achieve the planned objectives of the present research, a comprehensive literature review methodology was employed, resulting in the successful extraction of case studies illustrating the integration of PM and PS within healthcare contexts. The review meticulously elucidates the salient features of each case study assessed,both in terms their strengths and limitations. Notably, it sheds light on the challenges encountered by certain case studies and the strategies deployed to overcome them. As such this helps future researchers to collate a number of useful insights with which to pre-empt similar possible hurdles and ensure a seamless progression of their studies.

In addition, the literature review identifies notable gaps, delineating avenues for future investigation primarily from a methodological standpoint. These gaps include the underuse of tools to combine PM and PS, the need to consider expert health criteria and data quality issues in analyses, expanding studies to cover larger geographical or departmental areas, and adding dimensions of PM. These observations not only highlight areas for exploration but also furnish a roadmap for methodological refinement and innovation.

Moreover, the present review provides an in-depth examination of the prevalent PM and PS tools and techniques across the 31 cases, accompanied by comparative analysis of how they are applied. It identifies the most commonly researched healthcare domains the diverse typologies of processes—whether clinical or organizational—and the predominant data types employed. Such insights afford a vital framework for future researchers to optimize outcomes and refine existing methodologies.

All these contributions can be seen in Figure 7, which provides a concise summary of the objectives, results, and contributions presented by the authors in each of the 31 primary references included in the present literature review.

IV. DISCUSSION

Based on the literature review conducted by the present study and the nine aspects outlined in the Results section, the following discussion focuses on identifying the most significant findings from the 31 research studies included in the present paper. It also establishes several opportunities for future research with regards to the combination of process mining and process simulation in the healthcare sector.

A. MOST SIGNIFICANT FINDINGS

The literature review undertaken yielded a number of significant findings, all of which are described in the following paragraphs.

In general, certain studies mention that, at the research level, work is being conducted to generate new algorithms with which to perform PM analysis and that these are compatible with simulation tools. To validate this claim, efforts have been taken to design several simulation tools, such as FlexSim and Simio[13], which not only simulate discrete systems, but are also more closely related to the field of healthcare.

[13]https://www.simio-simulacion

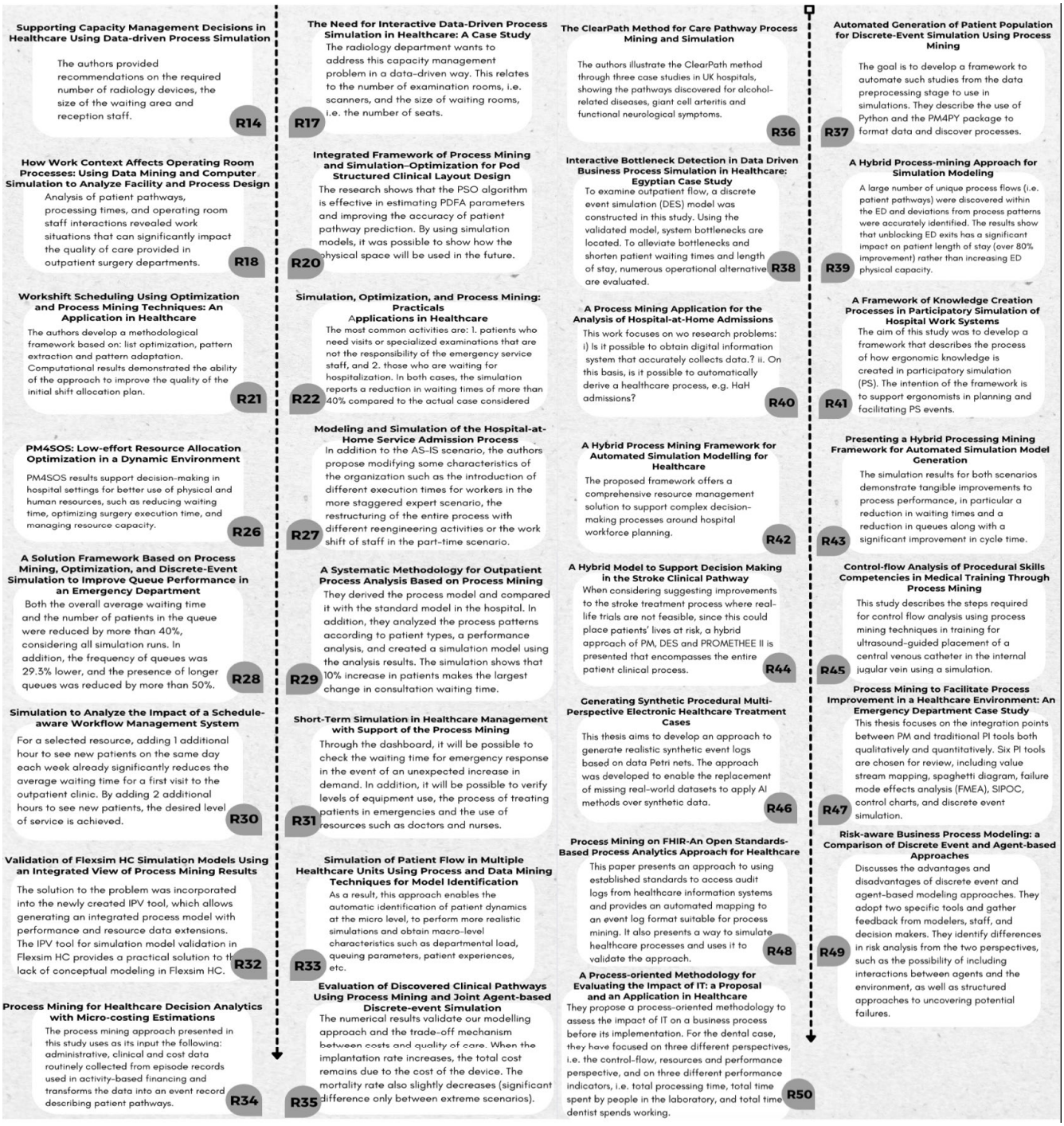

FIGURE 7. Summary of objectives, results, and contributions of the PM and PS case studies included in this literature review.

The papers reviewed do not clearly establish a methodology or framework for understanding how these PM and PS tools interact or how the use of event logs in simulation analysis can be achieved. This finding may suggest the lack of a clear methodological approach taken in the case studies with regards to the application of PM and PS in health processes, at least in some instances.

Moreover, there is a lack of clarity in the studies regarding the evaluation of simulation models generated through PM. It is unclear whether the simulated process accurately reflects the observed behavior, as assessed by both domain experts and analysts conducting the evaluations. Rectifying this issue would help to discern the strengths and weaknesses in simulation modeling. It would also aid in pinpointing

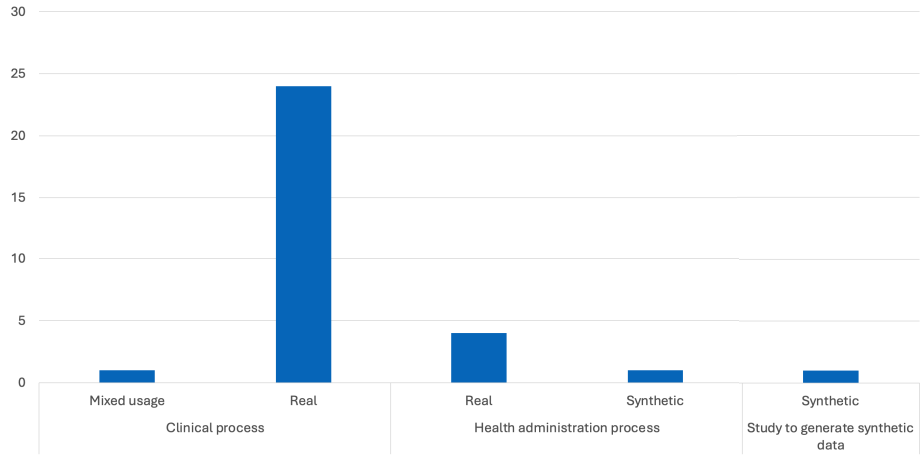

FIGURE 8. Relationship between process types and data types in PM and PS case studies.

sources of variation, therefore enhancing possible future research techniques. Even these aspects could be addressed by applying the conformity analysis of PM, which is only addressed in one of the case studies. There is potential to include different PM perspectives in studies of this kind. In this way, future research could be conceived of in terms of using resources that execute several activities in the simulations. Once one task has been completed, the subsequent one is performed, thereby more closely resembling the reality experienced in many medical and organizational processes. Such aspects were not identified in the case studies evaluated in the present review.

In general, research into instances in which PM and PS are combined in health centers, departments, medical areas or clinics in a simultaneous manner is generally scarce. It is unusual for several processes to be considered at the same time in the analysis, which implies that the studies are limited to single processes and spaces (both physical and geographic) and specific medical specialties. Consequently, there is an opportunity to focus research on the combination of PM and PS in healthcare from the conformance and improvement types of PM. This is because the type of PM analysis that is most commonly performed is the discovery type. It is also possible to propose case studies in which PM types are mixed. Indeed, only one study reviewed for the present paper deployed a methodology based on a mixture of discovery and conformance PM types.

Regarding the most commonly used PM algorithms in the case studies reviewed, Fuzzy Miner and Heuristics Miner were identified. This may be due to the fact that these two algorithms enable less complex process models to be defined for healthcare decision-makers. However, there are opportunities for improvement at the methodological level among the studies, since 14 of the 31 publications fail to indicate the algorithm used and only refer to the fact that they are discovery algorithms. There is a further gap in the studies whereby authors do not indicate either the PM algorithms or the PS software used to perform their respective analyses.

To conclude the discussion on PM tools, the most notable example identified is the ProM software. This is likely attributable to its open-source nature. In contrast, commercial tools, such as, Celonis are noticeably absent from all cases. Therefore, the case studies included in the present paper constitute a broad sample of tools that combine PM and PS tools, in which distinct software is used for each analysis. Consequently, there is not a single paper in which the tools used applied both types of analysis.

A further important aspect that emerges from the literature review is that the case studies show that the type of PS that seemingly best fits the combination with PM in healthcare is the simulation of discrete events, which is the most commonly used analysis type (in 55% of case studies). This may be because this particular type of simulation allows pre-defined systems to be modeled with predictable interactions. The most widely used software tool for these analyses is Arena, since it is widely considered to be user-friendly, according to the principal researcher of the present paper. However, further research in this area may wish to use more state-of-the-art tools, such as AnyLogic, which is utilized in only two of the cases reviewed herein. Or, to reiterate, such research could focus on cases in which the analyses are executed in software that integrates both disciplines.

Regarding tool combination, it should be noted that one of the most frequent interactions found in the literature review was the use of discrete-event simulation with discovery PM, applying ProM or Disco and using Arena.

Additionally, most of this research relates to medical processes rather than administrative or organizational processes.

In terms of the type of data used, the prevalence of the use of real data for PM analysis, and subsequent use in PS, may be the result of the difficulties involved in locating or processing synthetic data in the healthcare field. This use of

real data focuses mainly on medical processes rather than on organizational processes (in 77% of the studies). Indeed, the only case in which mixed data was used also involved a medical process. In the case of organizational processes, both real and synthetic data are used, as shown in Figure 8.

As mentioned in Section III-A, the declared objectives of the majority of the case studies focus on the analysis and improvement of indicators related to the organizational management of healthcare centers, without explicitly referencing the PM perspectives used. In fact, only two of the cases indicate the perspective(s) under which they are working. In other cases, the emphasis is placed more on improving the simulation of processes with the use of PM, as well as in presenting a framework, in addition to specific objectives across certain medical fields (see Figure 2).

In terms of the geographical predominance of research into the combination of PM and PS in healthcare, Europe leads the way. However, the Americas have gradually been increasing their contribution and the recent growth in research in these healthcare domains, has been significant. For example, between 2019 and 2023, a total of 19 papers were identified. Regarding the main medical fields present in the case studies, the areas of the emergency department, cardiology and radiology are the most common. It should be noted that certain other significant health areas were entirely absent from the case studies reviewed, including pediatrics and telemedicine, among others.

From among the results reported, it is possible to identify certain limitations from the case studies. The most frequently mentioned of these relates to the need to incorporate experts in the field to ensure the validation of the simulation results. In addition, there are challenges concerning data quality and a lack of integration of the sources from which the data is obtained. Moreover, the replicability of the methodologies reviewed is sometimes difficult owing to numerous factors such as the availability of data, the use of specific software tools, and cultural aspects that vary between countries, healthcare centers or health departments.

According to the present literature review, research that combines PM and PS in healthcare has been increasing since 2018. That particular year is notable because three studies were published in that 12-month period through distinct vehicles, with a further 19 publications identified as having been published from between 2019 and 2023. This gradual growth of research into PM and PS in healthcare is driving substantial changes in terms of analysis, as well as an improvement and optimization. This is primarily because PM allows for the use of real stored data. Such data helps to provide concrete knowledge about the execution of the processes and, thus, improve the representativeness and scope of the simulations. In turn, this facilitates the identification of improvement opportunities and decision-making.

Considering all the findings derived from the analysis of primary references within this literature review, it becomes evident that the integration of PM and PS holds significant implications for health management, decision-making processes, and even health policies, thereby influencing organizational efficiency, cost reduction, and patient satisfaction.

Several key implications can be underscored:

-The combination of PM and PS facilitates the identification of areas for improvement within healthcare processes, including waiting times, bottlenecks, and other associated inefficiencies.

-Simulation models aid in determining the optimal allocation of resources, both material and human, which are necessary for meeting patient demand, as well as optimizing the physical size of waiting areas. This contributes to enhanced capacity management and more efficient resource allocation.

-Through process optimization and resource management, healthcare institutions can mitigate organizational costs linked to several important factors such as excessive staffing, prolonged waiting times, or inefficient utilization of medical equipment.

-PS enables the testing of various scenarios and scheduling strategies in multiple clinical settings including surgical procedures, thereby enhancing resource utilization and reducing costs.

-By reducing waiting times and streamlining processes, patients can benefit from expedited and more efficient care delivery, thereby enhancing overall patient satisfaction.

-The integration of PM and PS contributes more robust, data-driven evidence with which to inform decision-making processes, thereby potentially shaping the development of health policies.

-Simulation models empower decision-makers to assess the impact of different strategies, scenarios, or policies prior to their implementation, thereby aiding in risk mitigation.

B. FUTURE RESEARCH OPPORTUNITIES

A number of opportunities for future research were identified in the case studies reviewed, which are shown in Figure 9 and further elaborated below.

First, the importance of involving experts in the development of simulations based on PM, in order to improve the representativeness and modeling of the simulation.

Second, the need to consider additional variables in the analysis of organizational indicators from the results stemming from simulations. This includes, for example, the chance to cover further aspects in the patient trajectory, include more variables in the processes, study the scheduling of appointments in multiple processes, and be able to incorporate more PM perspectives in interactions with the simulation,or be able to perform analysis on multiple processes at the same time.

Third, to broaden the scope of the simulation and process engineering analyses so that departmental, hospital, geographical and process environments can be compared. Likewise, more medical specialties could be included in this area of research. There is also the possibility of making comparative analyses between different medical specialties.

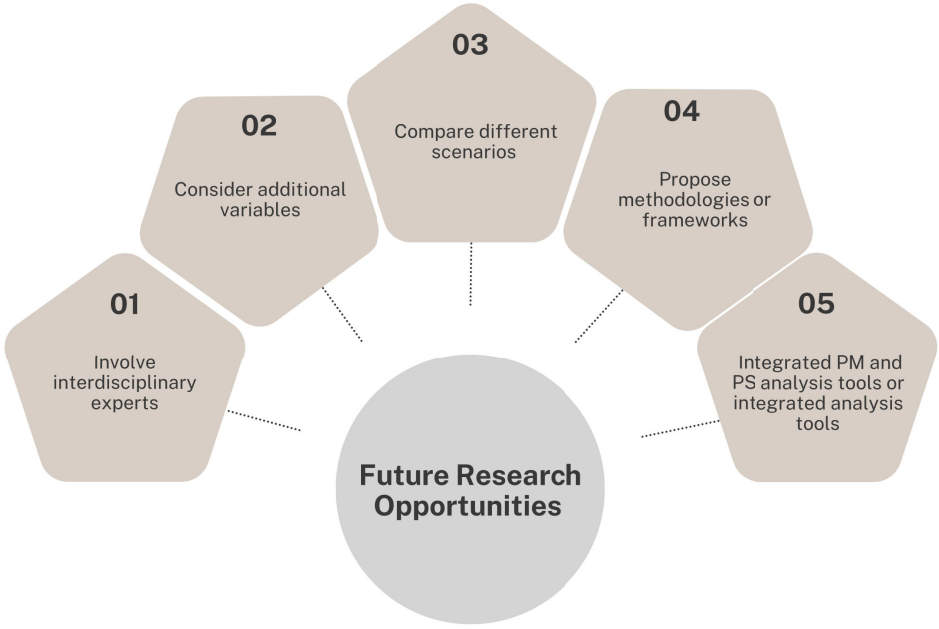

FIGURE 9. Future research opportunities.

Fourth, it would be possible to propose a specific methodology or framework with which to define the complementarity of PM and PS in the healthcare field. Currently, there is no singular approach uniformly adopted by the case studies within this research domain. Rather, each author or group of authors is conducting their research on the basis of their own individual preferences, or at least, there is no explicit framework guiding them in the case studies reviewed.

Finally, the inclusion of computer tools or software that integrate both PM and PS analyses, such as Celonis[14] or Apromore[15], should be considered in future studies. A detailed evaluation of the advantages and disadvantages of these integrated tools, compared to the independent tools discussed in the case studies included in this literature review, is necessary to assess the potential benefits thereof.

V. LIMITATIONS

The limitations of this literature review are related to the following two aspects:

1. Sample and Selection

The present article is a literature review on the combination of PM and PS in healthcare. Consequently, it does not entail an in-depth analysis of each document included or a comparison of the results thereof. The searches performed were carried out by establishing the keywords defined by the authors and, as such, some relevant publications may have been overlooked.

Nevertheless, more digital libraries could be included and the search years could be extended in each library used. The documents that constitute the subject of the present study relate to those that fulfilled the inclusion criteria defined in the methodology (Sections II-E and II-F).

The search strategy, along with the inclusion and exclusion criteria, were defined by the authors to ensure that included references addressed pertinent topics exclusively within the healthcare sector. Despite this, five references were ultimately discarded during the comprehensive analysis phase of the documents. Further examination revealed that three of these studies were not conducted in the healthcare domain, while two references discussed PS without incorporating PM or employing alternative techniques or tools.

The use of Covidence software helped to mitigate potential biases during the review stages (inclusion and exclusion) of documents. However, it is important to acknowledge inherent limitations in the search strategies, which could be influenced by human error or occasional difficulties in navigating search libraries. Such errors may have resulted in the exclusion of documents that met the inclusion criteria, thereby warranting cautious interpretation of the findings.

2. Time Restrictions

The present paper examined studies from between 2009 and 2023. Therefore, the period of analysis could be increased in order to, potentially, increase the number of references generated.

Although documents published prior to 2009 were excluded from the methodological process, it would be possible to extend the search timeframe to encompass a broader range of years, spanning approximately 20 to 25 years. This consideration is particularly significant given the relatively recent emergence of PM applications within the healthcare sector. Indeed, the combination of PM with

[14]https://www.celonis.com
[15]https://apromore.com

TABLE 3. Primary studies published in conferences, journals, master's thesis and workshops.

Publishing vehicle	Article Title	Authors	REF
Conferences	PM4SOS:Low-effort Resource Allocation Optimization in a Dynamic Environment.	Ferronato et al. (2022)	[26]
	Modeling and Simulation of the Hospital-at-Home Service Admission Process.	Amantea et al. (2019)	[27]
	A Solution Framework Based on Process Mining, Optimization, and Discrete-Event Simulation to Improve Queue Performance in an Emergency Department.	Antunes et al. (2019	[28]
	A Systematic Methodology for Outpatient Process Analysis Based on Process Mining.	Cho et al. (2015)	[29]
	Evaluation of Discovered Clinical Pathways Using Process Mining and Joint Agent-based Discrete-event Simulation.	Augusto et al. (2016)	[35]
	Workshift Scheduling Using Optimization and Process Mining Techniques: An Application in Healthcare.	Guastalla et al. (2022)	[21]
	The ClearPath Method for Care Pathway Process Mining and Simulation.	Johnson et al. (2019)	[36]
	Simulation, Optimization, and Process Mining: Practicals Applications in Healthcare.	Aringhieri et al. (2023)	[22]
	Automated Generation of Patient Population for Discrete-event Simulation Using Process Mining.	Le et al.(2022).	[37]
	Interactive Bottleneck Detection in Data Driven Business Process Simulation in Healthcare: Egyptian Case Study.	Albakary et al.(2022).	[38]
	A Hybrid Process-Mining Approach for Simulation Modeling.	Abohamad et al. (2017)	[39]
	A Hybrid Process Mining Framework for Automated Simulation Modelling for Healthcare.	Mesabbah et al. (2019)	[42]
	Presenting a Hybrid Processing Mining Framework for Automated Simulation Model Generation.	Mesabbah et al. (2018)	[43]
	Risk-aware Business Process Modeling: A Comparison of Discrete Event and Agent-based Approaches.	Sulis et al. (2019)	[49]
Journals	Supporting Capacity Management Decisions in Healthcare Using Data-driven Process Simulation.	van Hulzen et al. (2022)	[14]
	How Work Context Affects Operating Room Processes: Using Data Mining and Computer Simulation to Analyze Facility and Process Design.	Baumgart et al. (2009)	[18]
	Simulation to Analyze the Impact of a Schedule-aware Workflow Management System.	Mans et al. (2010)	[30]
	Integrated Framework of Process Mining and Simulation–Optimization for Pod Structured Clinical Layout Design.	Halawa et al. (2021)	[20]
	Simulation of Patient Flow in Multiple Healthcare Units Using Process and Data Mining Techniques for Model Identification.	Kovalchuk et al. (2018)	[33]
	Process Mining for Healthcare Decision Analytics with Micro-costing Estimations.	Leemans et al. (2023)	[34]
	A Process Mining Application for the Analysis of Hospital-at-Home Admissions.	Amatea et al. (2020)	[40]
	A Framework of Knowledge Creation Processes in Participatory Simulation of Hospital Work Systems.	Andersen et al. (2017)	[41]
	A Hybrid Model to Support Decision Making in the Stroke Clinical Pathway.	Boareto et al. (2022)	[44]
	Control-flow Analysis of Procedural Skills Competencies in Medical Training Through Process Mining.	De la Fuente et al. (2020)	[45]
	A Process-oriented Methodology for Evaluating the Impact of IT: a Proposal and an Application in Healthcare	Mans et al. (2013)	[50]
Master's Thesis	Validation of Flexsim HC Simulation Models Using an Integrated View of Process Mining Results.	Wijgergangs (2014)	[32]
	Generating Synthetic Procedural Multi-Perspective Electronic Healthcare Treatment Cases.	Jilg (2022)	[46]
	Process mining to Facilitate Process Improvement in a Healthcare Environment: An Emergency Department Case Study.	Matthews (2013)	[47]
Workshops	The Need for Interactive Data-Driven Process Simulation in Healthcare: A Case Study.	van Hulzen et al. (2021)	[17]
	Short-Term Simulation in Healthcare Management with Support of the Process Mining.	Pegoraro et al. (2018)	[31]
	Process Mining on FHIR-An Open Standards-Based Process Analytics Approach for Healthcare.	Helm et al. (2021)	[48]

techniques such as PS remains a more recent development, suggesting potential insights from a more extensive historical perspective.

VI. CONCLUSION

The present paper describes the results of a literature review conducted using the PRISMA methodology that combines the areas of PM and PS in the healthcare field. The analysis covers 31 primary research studies.

The investigation found that no previous study has conducted a literature review of research which combines PM and PS in the field of healthcare. While there are several literature reviews that provide related information on how these two disciplines can work together, they focus on other sectors or do not include an analysis of the current state of research. Such findings demonstrate the potential significance of the present paper in relation to possible future research.

The literature revealed certain limitations, such as the need to involve healthcare experts in developing data-based simulations. Concurrently, this particular finding highlights the limitations of the PM results themselves, since the lack of involvement of such experts can limit knowledge regarding whether the data are behaving in accordance with the process being analyzed.

In addition, several gaps were identified in the case studies. For example, a number of manuscripts do not make clear reference to the tools that were applied in relation to both PM and PS. In turn, this affects how the methodological approaches used can be understood and, moreover, their future replicability or implementation in other settings.

The results show that the study of scenarios and variations of processes from simulation and PM facilitate the analysis of organizational indicators, such as waiting times, allocation of operating rooms and resources, among others. Furthermore, these analyses help to improve the health services being accessed by the general public. In addition, the results reflect a growing interest in working on these areas of research.

There was a clear preference for the use of certain software tools, both for PS analysis (Arena) and PM (ProM or Disco). Similarly, there is an opportunity to include more medical fields in future related research, since a general preference was identified in the present review for the areas of emergency healthcare, cardiology and radiology.

Consequently, the present work provides a series of contributions that go hand in hand with the achievement of the objectives defined herein: it identifies the existing case studies in which PM and PS in healthcare are combined and, simultaneously, characterizes these case studies by highlighting the PM and PS techniques and tools used, the types of data collected, and the types of PM and PS performed. Finally, it offers a summary of the limitations encountered in the literature and a description of how some of them have been dealt with.

The overall conclusion is that PM is an effective technique for supporting the construction of PS models in healthcare.

APPENDIX

PRIMARY STUDIES PUBLISHED IN CONFERENCES, JOURNALS, MASTER'S THESIS AND WORKSHOPS

See Table 3.

ACKNOWLEDGMENT

The authors would like to thank the Universidad de Costa Rica and the Pontificia Universidad Católica de Chile for their support to this research and the academia. Additionally, they wish to thank FONDECYT for their support of this research and their ongoing assistance to the scientists of Chile.

REFERENCES

[1] M. Silva, M. J. Rosa, and L. Teixeira, "Supporting digital transformation using business process management: Proposal of a digital tool," in *Proc. Int. Conf. Ind. Eng. Oper. Manag.*, 2021, pp. 1–13.

[2] S. Castro and L. Teixeira, "Industry 4.0 and business process management: An exploratory study on the bilateral effects," in *Proc. Int. Conf. Ind. Eng. Oper. Manage.*, Mar. 2021, pp. 4840–4847.

[3] M. Thiede, D. Fuerstenau, and A. P. B. Barquet, "How is process mining technology used by organizations? A systematic literature review of empirical studies," *Bus. Process Manage. J.*, vol. 24, no. 4, pp. 900–922, Jun. 2018.

[4] M. Arias, E. Rojas, S. Aguirre, F. Cornejo, J. Munoz-Gama, M. Sepúlveda, and D. Capurro, "Mapping the patient's journey in healthcare through process mining," *Int. J. Environ. Res. Public Health*, vol. 17, no. 18, p. 6586, Sep. 2020.

[5] W. M. P. van der Aalst, *Process Mining: Data Science in Action*, 2nd ed., Cham, Switzerland: Springer, 2016, doi: 10.1007/978-3-662-49851-4.

[6] W. Van Der Aalst et al., "Process mining manifesto," in *Proc. Bus. Process Manage. Workshops (BPM)*, in Lecture Notes in Business Information Processing, vol. 99, Clermont-Ferrand, France, F. Daniel, K. Barkaoui, and S. Dustdar, Eds., Cham, Switzerland: Springer, 2011, pp. 169–194.

[7] W. M. P. van der Aalst, "Process mining and simulation: A match made in heaven!" in *Proc. 50th Comput. Simulation Conf., SummerSim*, Bordeaux, France, Jul. 2018, pp. 4:1–4:12.

[8] A. H. Ter Hofstede, W. M. Van der Aalst, M. Adams, and N. Russell, *Modern Business Process Automation: YAWL and Its Support Environment*. Cham, Switzerland: Springer, 2009.

[9] E. Rojas, J. Munoz-Gama, M. Sepúlveda, and D. Capurro, "Process mining in healthcare: A literature review," *J. Biomed. Informat.*, vol. 61, pp. 224–236, Jun. 2016.

[10] M. R. Dallagassa, C. dos Santos Garcia, E. E. Scalabrin, S. O. Ioshii, and D. R. Carvalho, "Opportunities and challenges for applying process mining in healthcare: A systematic mapping study," *J. Ambient Intell. Humanized Comput.*, vol. 13, no. 1, pp. 165–182, Jan. 2022.

[11] A. Guzzo, A. Rullo, and E. Vocaturo, "Process mining applications in the healthcare domain: A comprehensive review," *Wiley Interdiscipl. Rev., Data Mining Knowl. Discovery*, vol. 12, no. 2, p. e1442, Mar. 2022.

[12] J. Munoz-Gama et al., "Process mining for healthcare: Characteristics and challenges," *J. Biomed. Informat.*, vol. 127, Mar. 2022, Art. no. 103994.

[13] D. Chapela-Campa, I. Benchekroun, O. Baron, M. Dumas, D. Krass, and A. Senderovich, "Can I trust my simulation model? Measuring the quality of business process simulation models," 2023, *arXiv:2303.17463*.

[14] G. van Hulzen, N. Martin, B. Depaire, and G. Souverijns, "Supporting capacity management decisions in healthcare using data-driven process simulation," *J. Biomed. Informat.*, vol. 129, May 2022, Art. no. 104060.

[15] M. Arias, R. Saavedra, M. R. Marques, J. Munoz-Gama, and M. Sepúlveda, "Human resource allocation in business process management and process mining: A systematic mapping study," *Manage. Decis.*, vol. 56, no. 2, pp. 376–405, Feb. 2018.

[16] J. Jordon, L. Szpruch, F. Houssiau, M. Bottarelli, G. Cherubin, C. Maple, S. N. Cohen, and A. Weller, "Synthetic data—What, why and how?" 2022, *arXiv:2205.03257*.

[17] G. van Hulzen, N. Martin, and B. Depaire, "The need for interactive data-driven process simulation in healthcare: A case study," in *Proc. Int. Workshops, Process Mining (ICPM)*, Padua, Italy. Cham, Switzerland: Springer, Oct. 2020, pp. 317–329.

[18] A. Baumgart, C. Denz, H.-J. Bender, and A. Schleppers, "How work context affects operating room processes: Using data mining and computer simulation to analyze facility and process design," *Quality Manage. Health Care*, vol. 18, no. 4, pp. 305–314, 2009.

[19] S. C. Brailsford, M. W. Carter, and S. H. Jacobson, "Five decades of healthcare simulation," in *Proc. Winter Simulation Conf. (WSC)*, Dec. 2017, pp. 365–384.

[20] F. Halawa, S. C. Madathil, and M. T. Khasawneh, "Integrated framework of process mining and simulation–optimization for pod structured clinical layout design," *Expert Syst. Appl.*, vol. 185, Dec. 2021, Art. no. 115696.

[21] A. Guastalla, E. Sulis, R. Aringhieri, S. Branchi, C. Di Francescomarino, and C. Ghidini, "Workshift scheduling using optimization and process mining techniques: An application in healthcare," in *Proc. Winter Simulation Conf. (WSC)*, Dec. 2022, pp. 1116–1127.

[22] R. Aringhieri, M. Di Cunzolo, M. Dutto, L. Genga, A. Guastalla, M. Locatelli, L. Pellegrini, G. Ruffini, E. Sulis, L. C. Tagliabue, and A. Boccuzzi, "Simulation, optimization, and process mining: Practicals applications in healthcare," in *Proc. 3rd Nat. Conf. Artif. Intell. (Ital-IA)*, 2023, pp. 1–6.

[23] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Softw. Eng. Group School Comput. Sci. Math. Keele Univ. Keele, Staffs, U.K., Dept. Comput. Sci., Univ. Durham, Durham, U.K., Tech. Rep., 2007.

[24] K. Petersen, S. Vakkalanka, and L. Kuzniarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," *Inf. Softw. Technol.*, vol. 64, pp. 1–18, Aug. 2015.

[25] M. J. Page et al., "Declaración PRISMA 2020: Una guía actualizada para la publicación de," *Revista Española de Cardiología*, vol. 74, no. 9, pp. 790–799, 2021.

[26] J. J. Ferronato, E. E. Scalabrin, and D. R. Carvalho, "PM4SOS: Low-effort resource allocation optimization in a dynamic environment," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2022, pp. 1742–1747.

[27] I. Amantea, M. Arnone, A. Di Leva, E. Sulis, D. Bianca, E. Brunetti, and R. Marinello, "Modeling and simulation of the hospital-at-home service admission process," in *Proc. SIMULTECH*, 2019, pp. 293–300.

[28] B. B. Antunes, A. Manresa, L. S. Bastos, J. F. Marchesi, and S. Hamacher, "A solution framework based on process mining, optimization, and discrete-event simulation to improve queue performance in an emergency department," in *Proc. Int. Workshops, Bus. Process Manage. (BPM)*, Vienna, Austria. Cham, Switzerland: Springer, Sep. 2019, pp. 583–594.

[29] M. Cho, M. Song, and S. Yoo, "A systematic methodology for outpatient process analysis based on process mining," in *Proc. Asia Pacific Bus. Process Manag., 2nd Asia Pacific Conf. (AP-BPM)*, Brisbane, QLD, Australia. Cham, Switzerland: Springer, Jul. 2014, pp. 31–42.

[30] R. S. Mans, N. C. Russell, W. van der Aalst, P. J. M. Bakker, and A. J. Moleman, "Simulation to analyze the impact of a schedule-aware workflow management system," *Simulation*, vol. 86, nos. 8–9, pp. 519–541, Aug. 2010.

[31] F. Pegoraro, E. A. P. Santos, E. de Freitas Rocha Loures, G. da Silva Dias, L. M. dos Santos, and R. O. Coelho, "Short-term simulation in healthcare management with support of the process mining," in *Trends and Advances in Information Systems and Technologies*, vol. 2. Cham, Switzerland: Springer, 2018, pp. 724–735.

[32] G. Wijgergangs, "Validation of Flexsim HC simulation models using an integrated view of process mining results," Master thesis, TUE, School Ind. Eng.. Eindhoven Univ. Technol., 2014.

[33] S. V. Kovalchuk, A. A. Funkner, O. G. Metsker, and A. N. Yakovlev, "Simulation of patient flow in multiple healthcare units using process and data mining techniques for model identification," *J. Biomed. Informat.*, vol. 82, pp. 128–142, Jun. 2018.

[34] S. J. J. Leemans, A. Partington, J. Karnon, and M. T. Wynn, "Process mining for healthcare decision analytics with micro-costing estimations," *Artif. Intell. Med.*, vol. 135, Jan. 2023, Art. no. 102473.

[35] V. Augusto, X. Xie, M. Prodel, B. Jouaneton, and L. Lamarsalle, "Evaluation of discovered clinical pathways using process mining and joint agent-based discrete-event simulation," in *Proc. Winter Simulation Conf. (WSC)*, Dec. 2016, pp. 2135–2146.

[36] O. A. Johnson, T. B. Dhafari, A. Kurniati, F. Fox, and E. Rojas, "The ClearPath method for care pathway process mining and simulation," in *Proc. Int. Workshops, Bus. Process Manage. (BPM)*, Sydney, NSW, Australia. Cham, Switzerland: Springer, Sep. 2018, pp. 239–250.

[37] J. L. Lay, J. Neveu, B. Dalmas, and V. Augusto, "Automated generation of patient population for discrete-event simulation using process mining," in *Proc. Annu. Modeling Simulation Conf. (ANNSIM)*, Jul. 2022, pp. 42–53.

[38] W. A. AlBakary, A. A. H. Sedky, and W. Abdelmoez, "Interactive bottleneck detection in data driven business process simulation in healthcare: Egyptian case study," in *Proc. 32nd Int. Conf. Comput. Theory Appl. (ICCTA)*, Dec. 2022, pp. 234–240.

[39] W. Abohamad, A. Ramy, and A. Arisha, "A hybrid process-mining approach for simulation modeling," in *Proc. Winter Simulation Conf. (WSC)*, Dec. 2017, pp. 1527–1538.

[40] I. A. Amantea, E. Sulis, G. Boella, R. Marinello, D. Bianca, E. Brunetti, M. Bo, and C. Fernandez-Llatas, "A process mining application for the analysis of hospital-at-home admissions," *Stud. health Technol. Informat.*, vol. 270, pp. 522–526, Jun. 2020.

[41] S. N. Andersen and O. Broberg, "A framework of knowledge creation processes in participatory simulation of hospital work systems," *Ergonomics*, vol. 60, no. 4, pp. 487–503, Apr. 2017.
[42] M. Mesabbah, W. Abo-Hamad, and S. McKeever, "A hybrid process mining framework for automated simulation modelling for healthcare," in *Proc. Winter Simulation Conf. (WSC)*, Dec. 2019, pp. 1094–1102.
[43] M. Mesabbah and S. McKeever, "Presenting a hybrid processing mining framework for automated simulation model generation," in *Proc. Winter Simulation Conf. (WSC)*, Dec. 2018, pp. 1370–1381.
[44] P. A. Boareto, J. Safanelli, R. B. Liberato, C. H. C. Moro, J. E. Pécora Jr., C. Moro, E. de Freitas Rocha Loures, and E. A. P. Santos, "A hybrid model to support decision making in the stroke clinical pathway," *Simul. Model. Pract. Theory*, vol. 120, Nov. 2022, Art. no. 102602.
[45] R. de la Fuente, R. Fuentes, J. Munoz-Gama, A. Riquelme, F. R. Altermatt, J. Pedemonte, M. Corvetto, and M. Sepúlveda, "Control-flow analysis of procedural skills competencies in medical training through process mining," *Postgraduate Med. J.*, vol. 96, no. 1135, pp. 250–256, May 2020.
[46] D. Jilg, "Generating synthetic procedural multi-perspective electronic healthcare treatment cases," Master thesis, Trier Univ. Dept. IV Comput. Sci. Chair Bus. Inform. II, 2022.
[47] L. Matthews, *Process Mining to Facilitate Process Improvement in a Healthcare Environment: An Emergency Department Case Study*. Binghamton, NY, USA: State University of New York, 2013.
[48] E. Helm, O. Krauss, A. Lin, A. Pointner, A. Schuler, and J. Küng, "Process mining on FHIR—An open standards-based process analytics approach for healthcare," in *Proc. Int. Workshops, Process Mining (ICPM)*, Padua, Italy. Cham, Switzerland: Springer, Oct. 2020, pp. 343–355.
[49] E. Sulis, I. A. Amantea, and G. Fornero, "Risk-aware business process modeling: A comparison of discrete event and agent-based approaches," in *Proc. Winter Simulation Conf. (WSC)*, Dec. 2019, pp. 3152–3159.
[50] R. Mans, H. Reijers, D. Wismeijer, and M. van Genuchten, "A process-oriented methodology for evaluating the impact of IT: A proposal and an application in healthcare," *Inf. Syst.*, vol. 38, no. 8, pp. 1097–1115, Nov. 2013.

EVELYN SALAS received the master's degree in industrial engineering from the Universidad de Costa Rica. She is currently pursuing the Ph.D. degree in process mining and process simulation. Her professional focus is on the area of organizational development, process management, and process simulation. She is currently a Professor of industrial engineering with the Universidad de Costa Rica.

MICHAEL ARIAS received the Ph.D. degree in engineering from the Pontificia Universidad Católica de Chile (PUC), with a focus on computing. He is a Full Professor and a Researcher with the Universidad de Costa Rica (UCR). He is currently with the Human and Process Research Laboratory (HAPLAB), Department of Computer Science (DCC), PUC, and leads the Celonis Process Mining Center of Excellence, UCR. His research interests include process mining, business process management, process-oriented data science, project management, human resource management, IT governance, and gender in computer science.

SANTIAGO AGUIRRE received the Ph.D. degree in engineering from Pontificia Universidad Javeriana, Colombia. He has over two decades of teaching and research experience in fields related to process mining, information systems, data analytics, process automation, and business process management (BPM). He is currently a Faculty Member with the Business Information Systems Department, Central Michigan University, USA. He previously held the position of a Professor with the Industrial Engineering Department, Pontificia Universidad Javeriana. He has published record 14 peer-reviewed articles, 16 conference papers, and one book.

ERIC ROJAS (Member, IEEE) received the Ph.D. degree in engineering from the Pontificia Universidad Católica de Chile (PUC), with a focus on computing. His research interests include clinical informatics since his Ph.D. thesis in process mining in healthcare, in 2017. He is currently an academic with the Department of Clinical Laboratories, Faculty of Medicine, PUC, and the Leader of the Quality Assurance Area, Chilean National Center for Health Information Systems (CENS). In addition to being a member of the Institute of Biological and Medical Engineering, he is also an academic at the Department of Clinical Laboratories of the PUC Faculty of Medicine and Leader of the Quality Area at the National Center for Health Information Systems.