# Applying Process Mining in Population Health Management

Ricardo A. Quintano Neira[1] [0000-0001-9820-9096], Gert-Jan de Vries[1] [0000-0002-5180-4100], Ronny S. Mans[2], and Ioanna Sokoreli[1]

[1] Collaborative Care Solutions, Philips Research, High Tech Campus 34, 5656 AE Eindhoven, The Netherlands

[2] Philips VitalHealth, Zonneoordlaan 17, 6718 TK , Ede, The Netherlands

{ricardo.quintano, gj.de.vries, ronny.mans, ioanna.sokoreli}@philips.com

**Abstract.** Healthcare systems are facing challenges such as the increase in the number of chronically ill patients and the reduction in the availability of resources. This often leads to poor quality of clinical outcomes and increase of costs. One approach that contributes to minimizing the impacts of these challenges is to increase the adoption of preventive care. Population Health Management (PHM) develops and deploys healthcare programs aligned to the Quadruple Aim that promote the improvement of the population's health, while trying to contain or reduce costs and improve patient and clinician satisfaction. In this paper, we explore and discuss the use of process mining techniques to support the development and evaluation of PHM programs. In addition, we discuss possible challenges and recommend solutions, and we reflect upon using process mining to support addressing the Quadruple Aim in PHM.

**Keywords:** Population health · Quadruple aim · Process analytics

## 1 Introduction

Over the last years, healthcare systems across the world have constantly been affected by aging populations [32] as elderly people are more vulnerable to diseases and disabilities [33]. According to the United Nations [28], in 2017, the global population of elderly people (60+) was 962 million and it is expected to reach 2.1 billion by 2050.

Some of the challenges that healthcare systems face due to aging populations include the increasing number of chronically ill patients and a reduction in the availability of resources, resulting in inadequate measures to cover the populations' needs. This often leads to poor quality of clinical outcomes and increase of costs. One approach that contributes to minimize the impacts of these challenges is to increase the adoption of preventive care. This implies a proactive attitude, considering the overall population needs instead of a reactive attitude focused on the treatment of already diagnosed (ill) individuals and offering early interventions to those who are at risk of being diagnosed in the future.

Population Health Management (PHM) aims to transform care delivery and payment models for the purposes of improving the health status of a group of patients whilst containing costs, using better the available resources and being considerate for patient and clinician satisfaction [5, 21]. The Quadruple Aim is a key framework [3] that supports PHM aiming at the optimization of the healthcare system performance by simultaneously considering the following four dimensions:

- Enhance patients experience and satisfaction
- Improve the populations' health
- Decrease healthcare costs
- Enhance healthcare staff work life experience and satisfaction

PHM programs are often developed and deployed in order to promote the improvement of populations' health while taking into account the Quadruple Aim. Some examples of PHM programs deployed in healthcare organizations can be care management programs for chronic disease patients, looking into individual patient's needs, addressing them and following up on them in a timely manner to prevent patient deterioration. Other PHM programs or initiatives might have a more preventive goal such as smoking cessation programs.

The common iterative process of rolling out a PHM program includes the following steps:

- Identification of population, understanding their needs and associated risk factors and health goals
- Design and deployment of the program considering clinical interventions and/or process (re)design
- Evaluation and program adaptation

Process mining (PM) is a relatively new discipline that supports the analyses of business processes using data extracted from Information Systems [1]. PM has a panoply of techniques that allows to automatically discover processes (process discovery), check the adherence in the execution of a normative/prescriptive process (conformance checking), identify performance measures (performance checking), and optimize processes (process enhancement). The raw material to execute most of PM algorithms is the event log, i.e. a repository of retrospective events from an executed process. The minimal elements of an event log are: the case id that represents a unique instance of the process, the step from the process that was executed, and the date and time the step was performed. In most cases, the event log is created using data extracted from Information Systems. PM has been widely applied in the healthcare area as described in previous studies [26] [19] [30] [12] [10].

Our hypothesis is that process mining techniques can support PHM in the adoption of the Quadruple Aim. The objective of this paper is to explore and discuss the use of PM techniques to support the development and evaluation of PHM programs to accomplish the Quadruple Aim. In this paper, we focus mainly in PM techniques. We acknowledge that other existing techniques and methods can also contribute in the management of PHM programs and that PM is not the exclusive way to address the open challenges.

The importance of this manuscript lies in the dissemination of traditional and well known PM techniques and approaches to healthcare organizations, such as Accountable Care Organization (ACO) and Integrated Delivery Network (IDN), that are developing, deploying and evaluating PHM programs. To the best of our knowledge there is no PM literature directed to guide the healthcare community in the evaluation and optimization of PHM programs, even though there are studies applying PM techniques to evaluate chronic diseases such as diabetes and asthma.

This paper is organized as follows: first, we describe how process mining analyses can support PHM programs; second, we discuss possible challenges and recommend solutions; then we reflect upon using PM to support addressing the Quadruple Aim in PHM; finally, we conclude the paper and present recommendations and future steps.

## 2 How Can PM Support PHM Programs?

In this section, we present different application opportunities on how PM can support the design, deployment and evaluation of PHM programs. Subsections 2.1 and 2.2 are part of the "Design and deployment of the program considering clinical interventions and/or process (re)design" step of the iterative process of rolling out a PHM program. Subsections 2.3 and 2.4 are executed in the "Evaluation and program adaptation" step.

We will present examples from a fictitious IDN titled "Good Health" (GH) that wants to improve the management of their population of individuals older than 17 years diagnosed with Asthma. They want to reduce the emergency department utilization rate and improve their population satisfaction. GH is composed of 7 hospitals, 6 laboratories, 50 primary care facilities distributed in 5 cities (A, B, C, D and E). We intentionally kept the examples simple to support the reader to better understand how PM can be applied in the process of rolling out PHM programs.

### 2.1 Identify How a Population of Interest is Currently Being Managed in Healthcare Organizations

There are guidelines written by medical associations which give directions on how best to treat a number of clinical conditions (such as asthma and diabetes). However, even when guidelines do exist, the way to manage the populations in actual practice can vary due to factors such as resources availability, social circumstances and individual behaviours (e.g. sleep and diet patterns, anxiety level).

PM can be applied to identify how a population of interest is currently being managed in healthcare organizations taking into consideration different process perspectives [17], such as control flow, resource, time performance and data. The control flow perspective provides information regarding the order in which the activities of the process are executed. The resource perspective (also known as organization perspective) describes the required resources in the process and their interactions. The time performance perspective describes time-related characteristics of the executed process (e.g. waiting time, time to execute an activity). The data perspective shows how data objects are managed and used as input for decision making during the execution of the process. The analysis of different process perspectives is useful to make healthcare organizations aware about their current practice, to identify problems in their routine, identify gaps in care and optimization opportunities and to find out differences in the management of a population in different contexts (such as location).

For the control flow perspective, PM algorithms that aim to discover processes (such as fuzzy [9], heuristic [29], inductive miners [14]) can be used to identify the current process that healthcare organizations are acting upon a given population.

*In our example, IDN GH has little insight into their asthma management process and they would like to identify and understand their currently executed processes. By applying process discovery individually to each of the 5 cities, GH identified that city B executes a different management process comparing to the other cities. They identified that 40% of city B population skips the lung function tests (pulmonary function tests) after the yearly outpatient visit. In addition, city B does not provide the yearly patient education activity, in which a specialist checks if the patient is administrating correctly their medication and provides proper guidance.*

For the resource perspective, the use of organization mining techniques (such as social network algorithms) [34] can support the identification of professionals or units that are overloaded, the relationship between professionals and bottlenecks in social networks. It is important to note that, for the identification of overloaded professionals, all activities executed by a given healthcare organization should be used since a healthcare professional may act in different treatment processes.

*With respect to our example, IDN GH identified applying organization mining techniques that two physicians from city D are currently overloaded in their activities. In addition, GH identified that only one nurse in city C executes the patient education activity, which would affect the correct execution of the program in case the nurse is not able to work.*

For the performance perspective, performance checking algorithms can be applied to get time indicators and to identify potential bottlenecks in the current practice. These indicators can support the identification of activities that should be optimized to provide a better patient and/or professional experience (e.g. reduce the waiting time to execute an exam or outpatient visit). Identified bottlenecks should be analyzed to find and treat their root causes.

*Regarding our example, IDN GH identified that the patients from city C have to wait 1 month to execute the lung function test after the yearly outpatient visit, which is a potential bottleneck in the program execution.*

For the data perspective, PM can be applied for the identification of decision rules (i.e. rules that support the process path selection taking into consideration the available attributes) or contextual factors that can affect the performance (i.e. data characteristics such as patient or organization attributes that affect time indicators) [11].

*In our example, GH could find out that if a medication of a patient is changed, an educational activity is always executed (decision rule); and that the age of the patient affects the follow-up visit time duration (contextual factor that affects performance).*

### 2.2 Design a New PHM Program for a Population of Interest

Often healthcare organizations are asked to improve their Key Performance Indicators (KPI) or the way they are managing a specific population. Payers, such as health plans and health insurance organizations, are moving from fee-for-service payment model (the healthcare organization is paid based on the quantity of services it provides) to value-based care model (the healthcare organization is paid based on the outcomes and it is accountable for both quality and cost of care). As stated in subsection 2.1, PM techniques such as process discovery, organization mining, performance checking can be used to identify the current management process of a given population. The output from these analyses provides valuable insights and input to design new PHM programs allowing healthcare organizations to tailor the implementation plan according to the context (such as different locations and population characteristics).

*With respect to the GH example, the network would like to deploy their new designed Asthma program. Using process discovery techniques, the IDN GH identifies that the current asthma management practice of city B is quite different from the designed program which would possibly cause negative impacts to clinicians during its deploy. In this way, GH decided to create a different deployment plan to city B, in which their current practice would be iteratively updated towards the designed program.*

The utilization of process simulation models can support healthcare organizations to execute what-if analyses to test the newly designed PHM program before deploying it, helping them to better evaluate its outcomes and impacts (such as bottlenecks, clinical outcomes and costs) [2] [27]. Mans et al. [18] present a process-oriented methodology in which PM techniques support the creation and validation of these simulation models. Martin et al. [20] describe how PM can support each of the process simulation modeling tasks. The designed program can be updated until the outcomes from the simulation are aligned to the healthcare organization's expectations. The combination of PM and simulation techniques promote a smoother and easier program deployment, which possibly contributes to better healthcare staff work-life experience and satisfaction.

*With the utilization of simulation techniques, GH discovered that they should review the current professional allocation in city E as the new designed program would increase the waiting time for the execution of the yearly lung function tests.*

### 2.3 Program Evaluation

The application of the PM techniques described in subsection 2.1 provide the different perspectives (such as the control flow, organizational, performance) of the execution of the designed program, supporting the healthcare organizations in its evaluation.

Conformance checking algorithms confront the designed program (process model) with the event log (created using data from Information Systems). These algorithms can be applied to identify deviations in the execution of the program. The analysis of these deviations may support healthcare organizations to adjust their program execution to be closer to the defined/designed program. Alternatively, these deviations can identify some populations or healthcare organization constraints indicating the necessity to review and adapt the program for a specific context.

*After 1 year of the implementation of their new asthma program, IDN GH wants to check the adherence of each city's current practice to the designed program. They found out that city A had a lower adherence to the program since some patients did not execute the lung function test as prescribed. The reason for low adherence was that the test was mainly scheduled during day time hours when the patients were working. As a solution, GH decided to extend some of the clinic hours to early morning and early evening.*

### 2.4 Program Optimization

PM can support the optimization of PHM programs. The analysis of frequent deviations and/or behaviours (e.g. the execution of certain activities or their associated order) in the execution of a program combined with their outcomes (e.g. clinical outcomes, cost, length of stay, patient and professional satisfaction) can provide the identification of behaviors that can contribute to the improvement of the program. The PHM program can be updated to incorporate these positive behaviors, or to remove the negative ones.

*Regarding our example, IDN GH identified that the execution of an extra activity that was not present in the original program (a.k.a., deviation), in which they provide remote education to their patients by e-mail, was associated with a lower emergency department utilization rate. GH decided to optimize the program by including this extra activity in the overall program for all sites.*

Another application of PM for optimization is to identify redundant or unnecessary activities in patient care, especially for multimorbid patients. Process discovery techniques can support healthcare organizations to find out that patients are undergoing to the same tests and services (prescribed from different specialists) twice within a short time while that could have been only done once.

As presented in subsection 2.2, simulation methods can be applied to perform what-if analyses to check the expected outcomes and impacts of the proposed adaptations in the PHM program.

## 3 Challenges

In this section, we discuss some challenges regarding the use of PM applied for PHM programs. For the identification of these challenges, we considered our own experience and the literature (i.e. looking for PM case studies associated to chronic diseases such as diabetes, asthma).

Extracting the system data and creating the event log can be burdensome [25]. One challenge is mapping the attributes from the activities of a given PHM program to the system database. An incorrect mapping will result in the creation of an erroneous event log. In addition, the extraction and creation of the event log can consume more time than expected. The healthcare organization (or network) should have close interaction with the systems (e.g. hospital information system) development teams to get further information regarding the database model. Alternatively, each system could have a functionality to better facilitate extraction of data. To make this challenge even more complex, healthcare organizations from an IDN may use disparate systems from different vendors. This scenario requires different attributes mapping for each system (e.g. the execution time for activity "lung function test" will have a mapping in system A that is different in system B).

*To illustrate the extraction of system data we will use the IDN GH example. One activity of their asthma program is the "lung function test". For this activity, GH needs the patient identification and the start and end time. To extract and create this activity in the eventlog, GH needs to know in which table from their system this information is stored. They identified that the table TB_EXEC_PROCEDURE stores the date and times of the execution of the exam in the fields dt_start and dt_end, that the exam code is stored in the field id_procedure, and the patient identification is stored in the field id_person.*

A second challenge is the aggregation of data to create the event log using data originating from disparate systems and organizations. This issue can be quite common for PHM programs since they are performed across different healthcare organizations and locations (such as the patient home). These organizations commonly use disparate systems with different patients and procedure/medicines identification codes. The adoption of interoperability standards such as Fast Healthcare Interoperability Resources (FHIR) [8], OpenEHR [22] and terminology services (that provide the mapping of different terminologies) may contribute to support the aggregation of data from disparate sources.

*Taking in consideration the IDN GH example, the asthma program is executed in different locations, like the General Practitioner (GP) office to perform the annually follow up visit, in a hospital to perform the lung function test, in an educational center to teach the patient in how to correctly use their medication. Patient John Smith (one of the individuals from the Asthma population) has the identification code 54278 in his GP office, the identification 80008 in hospital A, and the code 90921 in the educational center. The lung function test has the code 267364 in the GP office, and the code PC78768 in hospital A. The medication Montelucast Sodium has the code 90897 in the GP office, and the code MS0982 in the educational center B. To extract and aggregate the data regarding the* *execution of the program by the asthmatic population, IDN GH would need to map all their patients, medicines and procedure codes from all their systems.*

Some initiatives have been started to collect all data in one place regarding care management (i.e. Personal Health Records - PHR) but were not yet successfully deployed. For example, when there is no interoperability between healthcare systems (such Electronic Health Record and Laboratory Information System) and a PHR, the clinical information needs to be registered manually in the PHR which increases the likelihood of errors and missing data. Alternatively, systems that support healthcare organizations for population health management focus in general on a particular condition of the chronic disease and other relevant information associated to the individual are not collected (such comorbidities or other acute care events).

A typical challenge in process mining analyses is guaranteeing the quality of the extracted data. Some issues regarding data quality can be: missing events (e.g. an activity of the program is not recorded in the system like education efforts performed in workshops or medication adherence [6]); the timestamp of an executed event is not properly recorded (e.g. the event was registered in the system 20 minutes later than the real action occurred); missing event attribute (e.g. the age or gender of patients is not recorded); how to guarantee that the execution of a given event is directly related to a PHM program (e.g. a visit to an ophthalmologist may not be associated to the diabetes treatment [15]). Jagadeesh Chandra Bose et al. [13] present a complete list of issues regarding data quality, and Mans et al. [19] discuss a set of logging guidelines that contribute to tackle this challenge. PHM programs in general happen across different environments such as hospitals, GP office, and even inside of the patient home. Consequently, there is the potential for high availability of data, which from one side can be positive as it can provide a richer view on each individual and a better identification of their needs. On the other hand, some challenges like lower data quality due to more noise factors and ethical considerations (such as the privacy of the information) may be introduced. This highlights the need for proper documentation of the source of the data and context in which the measurement was performed and ideally also a quantification of the quality of the measurement. Chen et al.[4] provide an overview of measures of quality of data in the neighbouring field of Public Health that are also relevant for PHM.

Although, as explained previously, high quality log information is essential for PM, one should always keep in mind that the most important aspect of healthcare is the interaction between caregiver and patient. The challenge here is to identify a way to better collect high quality events without adding extra burden to healthcare professionals.

The execution of healthcare processes stores highly sensitive personal data, thus, another challenge is related to protecting the privacy of individuals. Over the last years, privacy regulations had been inserted such as the California Consumer Privacy Act and the General Data Protection Regulation. According to Pika et al. [24] "ensuring high levels of privacy protection for such data while also preserving data utility for process mining purposes remains an open challenge for the healthcare domain". Pika et al. propose a privacy-preserving framework that supports dealing with this challenge.

The Spaghetti effect [1] is a common and challenging problem faced by PM. This effect is associated with the identification of unstructured process models that contain many connections, often presenting as long alternative (parallel) task-sequences, making it difficult to be understood by humans. In the healthcare domain this problem is common since many factors can affect the process execution (e.g. the severity of diseases, drug interactions, allergies and comorbidities) thereby creating such alternative process execution sequences. Fernandez-Llatas et al. [7] present some strategies they applied for the Diabetes treatment process that contributed to the reduction of the Spaghetti effect, such as the selection of the workflow model notation (e.g. Classical Finite Deterministic Automatons and Timed Parallel Automaton produced better results) or, the usage of rendering algorithms for highlighting specific situations (e.g. frequent executed paths). Litchfield et al.[16] display other strategies like aggregating and clustering the events to create more abstracted events, and clustering traces to reduce the number of multiple process variants (unique sequence of activities). PHM programs developed for chronic diseases such as Chronic Obstructive Pulmonary Disease (COPD), diabetes and Alzheimer usually do not have a clear end point in the process, contributing to the Spaghetti effect. This behavior happens since this type of disease needs a frequent follow up and typically, the process presents an infinite loop of specific activities. This provides a challenging environment for PM but dividing and analyzing the event log based on the different stages of the program (e.g. start of the program, diagnosis, treatment set up, maintenance, acute event intervention) may support and simplify the analysis. This strategy leads to a reduction in the number of activities and, as a consequence, may reduce the number of connections in the discovered models.

Conformance checking algorithms can support the identification of deviations in the execution of a PHM program, but the use of such type of technique may be challenging when it is required (or accepted) to adapt care plans to the patient's needs (personalized care). We believe that two solutions can overcome this challenge: 1. considering in the conformance checking analysis a process with a level of abstraction that only presents the behaviours expected to be executed correctly by most of the individuals (e.g. 80% or more of the population) or; 2. not considering in the analysis deviations associated to personalized care.

The selection of individuals that compose a population (case-mix) is another challenge. In other words, data from which individuals should be selected in the event log for the PM analyses? One approach used in healthcare research studies is selecting individuals based on their diagnostic codes (e.g. International Statistical Classification of Diseases and Related Health Problems - ICD). Considering only the diagnosis is not the most appropriate method since the diagnosis information is not always present or correctly updated. In addition, some populations may be composed of healthy individuals, that means that no diagnosis code will be present. Moreover, the diagnostic information alone does not give a complete picture of the health status of an individual. For that reason, we suggest to combine this information with other sources, such as clinical parameters including vital signs and laboratory and imaging exam results, and socio-demographic information. Finally, we can not ignore the fact of the existence of individuals that avoid healthcare treatments. They are known as "care avoiders". This fact poses another challenge: how to increase the likelihood of considering care avoiders in the case-mix definition as they probably have insufficient clinical data recorded in systems?

There are several challenges associated on the use of PM techniques to support the creation of simulation models. Martin et al. [20] presents a broad list of challenges identified from the literature, such as, determining the appropriate entity attributes and their abstraction level, finding the entity arrival rate in the presence of queues, identifying the duration of the activities taking into consideration the different determinants and the availability of timestamp data. In addition, they provide suggestions to tackle some of these challenges.

## 4 Discussion

We have identified the potential for PM to contribute to PHM as well as challenges that need to be overcome for successful application of PM in this field. In this section we will reflect upon our aim of using PM to support addressing the quadruple aim in PHM.

One of the primary purposes of PM is to assess to which extent processes are being followed in practice. Given that protocolized care has been shown to lead to better patient outcomes [31], there is a direct link between process adherence and the aim to improve the health of individuals and therewith the health of the population.

Whilst protocolized care does not directly reduce healthcare costs, it provides narrower bound to the costs and it makes deviating choices more explicit. Hence making, often costly, deviations from the protocol a conscious decision, thereby likely reducing the rate of occurrence for deviation. As a consequence, a reduction in healthcare costs can be anticipated. In addition, the analysis of the combination of process deviations/behaviours and outcomes may support the optimization of PHM programs to reduce costs.

By taking the learnings of PM (e.g. identification of bottlenecks and overloaded professionals) and optimizing the processes accordingly, processes and daily practice have the potential to align better and create a smoother way of working, which could lead to improved staff satisfaction. This, in turn will lead to more motivated staff, which is likely to have a positive influence not only on health outcomes but also importantly on patient experience.

PM can support healthcare organizations to increase the patient satisfaction. One example is to support a better fit of necessary care into and around an individual's life (such as extend some of the clinic hours to early morning and early evening to give more time options for patients that work). A second example is associated with the identification and optimization of high waiting times in the execution of one of the activities of the program.

Nevertheless, patient and professional satisfaction might prove difficult to be identified and/or quantified using exclusively data extracted from Information Systems. In addition, even if the satisfaction could be well measured, it may be difficult to link it to the process events. We suggest to add in the PHM programs specific activities to collect the patient and professional satisfaction using questionnaires. The answer from such questionnaires would enhance the event log and as consequence, provide better PM analyses and process optimization aligned to the Quadruple Aim.

In summary, we see opportunities for direct and indirect impacts of PM in all of the four dimensions of the Quadruple Aim.

## 5 Conclusions and Future Steps

In this paper, we provided a set of PM analyses that may be applied in PHM and we discussed the associated challenges. We understand that PM is able to support organizations in the development and evaluation of population health programs to fulfill the Quadruple Aim.

Given that the healthcare area is mostly organized per disease, we believe that a good starting point in applying PM to PHM is for managing and optimizing chronic disease programs. Ultimately, the end-goal could be to follow any person (healthy or not) throughout their life to reduce the likelihood of getting ill. This might have severe ethical considerations, though, which go beyond the scope of this paper.

We have recently started a PM feasibility study using data from a system that supports healthcare providers in population health management (Philips Engage [23]) in which so far we have gotten positive results (see more details in the Appendix). As future steps, we foresee to expand our analyses to cover the specificities of the different PHM programs related to chronic diseases such as pre-diabetes, COPD and CHF. In doing so, the challenges identified need to be faced and might lead to further insight into how to overcome these and how to make optimal use of PM for the execution and improvement of PHM.

**Acknowledgements.** The authors thank Aleksandra Tesanovic, Helen Schonenberg and Jennifer Caffarel for their support in the development of this manuscript.

## Appendix - Preliminary Results of the Feasibility Study

We have recently started a feasibility study to check if and how we can apply PM to support the development and evaluation of PHM programs. We are conducting our experiments using test/simulated data from Philips Engage (PE)[3]. In this section we present some preliminary/initial results.

[3] Philips Engage is a healthcare platform that supports healthcare providers to work remotely with their patients. It supports providers to monitor their patients, engage patients in their treatment, and offer health and care programs [23].

As a starting point, we are working on one part of the remote patient monitor process in which a patient provides their health data (such as weight and vitals), and based on some rules PE generates tasks for the healthcare professionals to act upon. When working on a task, the professional checks the patient's health data and if is necessary they can take actions such as interacting with the patient.

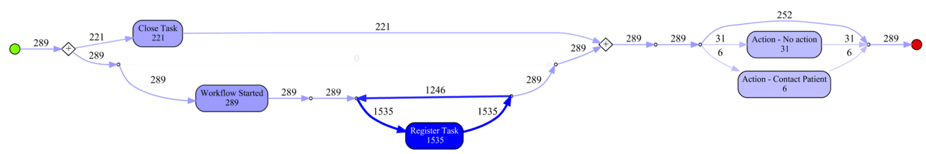


**Fig. 1.** Current executed process (AS-IS) of part of the remote patient monitor process. ProM plugin used: inductive visual miner

**Table 1.** Time performance analysis of part of the remote patient monitor process. The time is in hours. Empty cells indicate that there is no link between the 2 activities. ProM plugin used: replay a log on petri net for performance/conformance analysis

| From\To | Action Contact Patient | Action No Action | Close Task | Open Task | Register Task | Workflow Finished | Workflow Started |
|---|---|---|---|---|---|---|---|
| **Action Contact Patient** | 0.00 | | | | | | |
| **Action No Action** | | 0.00 | | | | | |
| **Close Task** | 0.00 | 0.00 | 0.00 | | | 0.00 | |
| **Open Task** | 0.21 | 0.08 | 3.59 | 0.00 | | 5.12 | |
| **Register Task** | 0.23 | 0.09 | 20.17 | 19.62 | 0.00 | 25.20 | |
| **Workflow Finished** | | | | | | 0.00 | |
| **Workflow Started** | 0.23 | 0.09 | 20.27 | 19.76 | 0.07 | 25.33 | 0.00 |

We created an event log composed of 289 cases and applied a set of PM techniques. We could identify the executed process (AS-IS) (see Fig. 1), identify the adherence in the execution of a normative process and related deviations, identify time performance indicators (see Table 1) and potential bottlenecks. We used ProM[4] for the execution of the PM analyses.

We consider these initial results promising and the execution of this study is supporting us to identify improvement areas to enhance the PM analyses (e.g., inclusion of complementary data related to a created task). As future steps, we expect to expand our analyses to cover the different PHM programs specificities (related to different diseases).

## References


1. van der Aalst, W.M.P.: Process Mining - Data Science in Action, Second Edition. Springer (2016). https://doi.org/10.1007/978-3-662-49851-4
2. Augusto, V., Xie, X., Prodel, M., Jouaneton, B., Lamarsalle, L.: Evaluation of discovered clinical pathways using process mining and joint agent-based discrete-event simulation. In: Proceedings of the 2016 Winter Simulation Conference. pp. 2135–2146. IEEE Press (2016)
3. Bodenheimer, T., Sinsky, C.: From triple to quadruple aim: care of the patient requires care of the provider. The Annals of Family Medicine **12**(6), 573–576 (2014)
4. Chen, H., Hailey, D., Wang, N., Yu, P.: A review of data quality assessment methods for public health information systems. International Journal of Environmental Research and Public Health **11**(5), 5170–5207 (2014)
5. Chilmark Research: Population Health Management: 2018 Market Trends Report (2018), https://www.researchandmarkets.com/reports/4714909/population-health-management-2018-market-trends
6. Conca, T., et al.: Multidisciplinary collaboration in the treatment of patients with type 2 diabetes in primary care: Analysis using process mining. Journal of Medical Internet Research **20**(4), e127 (2018)
7. Fernandez-Llatas, C., Martinez-Millana, A., Martinez-Romero, A., Benedi, J.M., Traver, V.: Diabetes care related process modelling using process mining techniques. Lessons learned in the application of Interactive Pattern Recognition: coping with the Spaghetti Effect. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). pp. 2127–2130. IEEE (2015)
8. FHIR - Fast Healthcare Interoperability Resources (2020), https://www.hl7.org/fhir/summary.html
9. Günther, C.W., Van Der Aalst, W.M.: Fuzzy mining–adaptive process simplification based on multi-perspective metrics. In: International Conference on Business Process Management. pp. 328–343. Springer (2007)
10. Helm, E., Baumgartner, D., Lin, A.C., Küng: Adopting standard clinical descriptors for process mining case studies in healthcare. Process-Oriented Data Science for Healthcare 2019 (PODS4H19), International Conference on Business Process Management (BPM 2019) (2019)
11. Hompes, B.F.A., Buijs, J.C.A.M., van der Aalst, W.M.P.: A generic framework for context-aware process performance analysis. In: On the Move to Meaningful Internet Systems: CoopIS, Rhodes, Greece. pp. 300–317 (10 2016). https://doi.org/10.1007/978-3-319-48472-3_17


[4] See http://www.promtools.org/doku.php

12. Hompes, B., Dixit, P., Buijs, J.: Using process analytics to improve healthcare processes. In: Data Science for Healthcare, pp. 305–325. Springer (2019)
13. Jagadeesh Chandra Bose, R., Mans, R., van der Aalst, W.M.: Wanna improve process mining results?: it's high time we consider data quality issues seriously. BPM Reports **1302** (2013)
14. Leemans, S.J., Fahland, D., van der Aalst, W.M.: Scalable process discovery with guarantees. In: Enterprise, Business-Process and Information Systems Modeling, pp. 85–101. Springer (2015)
15. Lismont, J., Janssens, A.S., Odnoletkova, I., vanden Broucke, S., Caron, F., Vanthienen, J.: A guide for the application of analytics on healthcare processes: a dynamic view on patient pathways. Computers in Biology and Medicine **77**, 125–134 (2016)
16. Litchfield, I., et al.: Can process mining automatically describe care pathways of patients with long-term conditions in uk primary care? a study protocol. BMJ open **8**(12), e019947 (2018)
17. Mannhardt, F.: Multi-perspective process mining. Ph.D. thesis, Department of Mathematics and Computer Science, Technische Universiteit Eindhoven (2018)
18. Mans, R., Reijers, H., Wismeijer, D., Van Genuchten, M.: A process-oriented methodology for evaluating the impact of IT: a proposal and an application in healthcare. Information Systems **38**(8), 1097–1115 (2013)
19. Mans, R.S., Van der Aalst, W.M., Vanwersch, R.J.: Process mining in healthcare: evaluating and exploiting operational healthcare processes. Springer (2015)
20. Martin, N., Depaire, B., Caris, A.: The use of process mining in business process simulation model construction. Business & Information Systems Engineering **58**(1), 73–87 (2016)
21. National Health Service (NHS): Population health and the population health management programme (2020), https://www.england.nhs.uk/integratedcare/building-blocks/phm/
22. OpenEHR - Open industry specifications, models and software for e-health (2020), https://www.openehr.org/
23. Philips: Philips Engage (2020), https://www.usa.philips.com/healthcare/services/population-health-management/patient-engagement/philips-engage
24. Pika, A., Wynn, M.T., Budiono, S., ter Hofstede, A.H., van der Aalst, W.M., Reijers, H.A.: Towards privacy-preserving process mining in healthcare. Process-Oriented Data Science for Healthcare 2019 (PODS4H19), International Conference on Business Process Management (BPM 2019) (2019)
25. Quintano Neira, R.A., de Vries, G., Caffarel, J., Stretton, E.: Extraction of data from a hospital information system to perform process mining. In: Proceedings of the 16th World Congress on Medical and Health Informatics (MEDINFO), Hangzhou, China. pp. 554–558 (2017). https://doi.org/10.3233/978-1-61499-830-3-554
26. Rojas, E., Munoz-Gama, J., Sepúlveda, M., Capurro, D.: Process mining in healthcare: A literature review. Journal of Biomedical Informatics **61**, 224–236 (2016)
27. Rozinat, A., Mans, R.S., Song, M., van der Aalst, W.M.: Discovering simulation models. Information systems **34**(3), 305–327 (2009)
28. United Nations: World Population Ageing. United Nations (2017)
29. Weijters, A., Ribeiro, J.: Flexible heuristics miner (fhm). In: 2011 IEEE Symposium on Computational Intelligence and Data Mining (CIDM). pp. 310–317. IEEE (2011)

30. Williams, R., Rojas, E., Peek, N., Johnson, O.A.: Process mining in primary care: A literature review. Studies in Health Technology and Informatics **247**, 376–380 (2018)
31. Winslow, J., et al.: North Carolina College of Emergency Physicians' guidance document on emergency medical services. North Carolina Medical Journal **76**(4), 256–262 (2015)
32. World Health Organization: Global Health and Aging. World Health Organization (2011)
33. World Health Organization: Risk factors of ill health among older people (2020), http://www.euro.who.int/en/health-topics/Life-stages/healthy-ageing/data-and-statistics/risk-factors-of-ill-health-among-older-people
34. Zhao, W., Zhao, X.: Process mining from the organizational perspective. In: Foundations of Intelligent Systems, pp. 701–708. Springer (2014)