Interactive Process Mining Applied in a Cardiology Outpatient Department

Juan José Lull[1(✉)], Adrián Cid-Menéndez[2], Gema Ibanez-Sanchez[1], Pedro Luis Sanchez[2], Jose Luis Bayo-Monton[1,2,3], Vicente Traver[1], and Carlos Fernandez-Llatas[1,3]

[1] Instituto Universitario de Tecnologías de la Información y de las Comunicaciones (ITACA), Universitat Politècnica de València, Valencia, Spain
{jualulno,geibsan,jobamon,vtraver,cfllatas}@itaca.upv.es
[2] Cardiology Department, University Hospital of Salamanca, Instituto de Investigación Biomédica de Salamanca (IBSAL), Salamanca, Spain
[3] Department of Clinical Science, Intervention and Technology (CLINTEC), Karolinska Institutet, Stockholm, Sweden
carlos.fernandezllatas@ki.se

Abstract. Cardiology departments receive many outpatients from primary care services and it is necessary to differentiate which patients need special attention. One-stop clinics were deployed in a hospital in Salamanca (Spain) to triage such patients, separating those who needed further examination and those who were discharged.

Data (covering December 2018—August 2020) was explored and there was an iterative process in which clinicians, process miners and technical staff at the hospital interacted in special interviews or Data Rodeos. Interactive Process Indicators (IPIs) were generated. During Data Rodeos data quality problems arose and were tackled, input data was cleaned and preconditioned, process activities were discovered and modelled.

The original assumption that the iterative implementation of the IPI would allow clinicians and managers to have a deeper understanding of the one-stop cardiology clinics process, was evaluated and validated by them. After each iteration, they found that the IPI was more useful and near to the reality they see everyday.

The final IPI was easy to interpret by the clinicians. In the end, many key indicators were extracted, but most importantly, clinicians had a comprehensive tool that they could use by themselves, without technical assistance, to extract and interpret different indicators at any time, providing a high-quality source of information to improve patient-centered daily medical care.

Keywords: Process mining · Cardiology · Interactive Process Mining · Healthcare system · Outpatient care

J. Munoz-Gama and X. Lu (Eds.): ICPM 2021 Workshops, LNBIP 433, pp. 340–351, 2022.
https://doi.org/10.1007/978-3-030-98581-3_25

1 Introduction

Hospitals have limited resources and they need to constantly evolve to attend patients in time and differentiate between patients that need extra attention and diagnosed patients and healthy subjects. Patients that need a heart diagnosis are usually referred from primary care to the cardiology service. In the normal pathway, when the patient arrives to the service, the doctor has a consultation and may require further tests in e.g., a month. Afterwards, the patient goes back to the cardiologist some weeks later and can be diagnosed as healthy and thus discharged. They can also be diagnosed with a specific known condition, and a treatment starts. Finally, the patient may also need extra tests because of a more complex malady. This process takes a long time and the number of consultations could be reduced with other approaches.

The model of one-stop clinic reduces the number of patients referred from primary care through a pathway that aims to provide a higher efficiency in the diagnosis of patients. Contrary to the normal pathway, the one-stop clinic aims to create more thorough consultations, in which the doctor has support from nursing staff and full access to specific extra test requests within the cardiology department. Many times, a basic echocardiogram (ECHO) or electrocardiogram is performed during the consultation and most patients are discharged, since they can be directly diagnosed and do not need any more tests. If extra tests or interventions need to be performed, they are carried out on a different (near) scheduled date. Tests generate a routine follow-up review. The way one-stop clinics are conducted makes it possible to quickly rule out most cardiac syndromes and, fundamentally, to discharge a great proportion of referred patients who do not meet the criteria for specific follow-up to primary care, saving resources for the sicker cohort of patients that benefit from a closer surveillance [2], and reducing the length of stay. One-stop clinics have been extended to European cardiology services in recent years as a solution for the management of first visits [7].

In the context of a European funded project, a cardiology department that had implemented one-stop clinics wanted to visualize and take decisions based on the processes related to these one-stop clinics, with the help of Process Mining (PM). The goal in the PM methodology is to provide solutions to the experts (in this case, the clinicians), that help them understand the behavior of the processes [11]. In the healthcare domain, process indicators need to be extracted from the data by analysts with the other stakeholders' help: Managers, clinicians, and Information Technology (IT) professionals.

PM has been applied to different Cardiology use cases. In [8], it was applied in a Pakistan Cardiology Hospital. It did not count on real event logs but was rather based on reports from the physicians. The aim of the research was to prove that PM could be applied to enhance the medical system in the country. In [1], Interactive Process Mining (IPM) [4] was applied to investigate how the time it takes to transfer the patient with myocardial infarction from their home to the percutaneous intervention center affects the survival rate. There is also a literature review about cardiovascular diseases studied with PM [9]. In that publication, they focused on what specific disease each paper had studied and

which PM method had been used. A deep inspection of the process and the way it was achieved was not apparent in the referred works.

Key performance indicators (KPIs) are commonly used in clinical settings. IPM, however, lets the stakeholders define Interactive Process Indicators (IPI) from the data and the questions that the clinicians and managers have. KPIs are single-dimension variables, e.g., the ratio of patients discharged per hospitalized patients, per week. Differently, IPIs contain a full visual description of the process, and also contain metadata about process traces and events. KPIs are just numbers, so errors in the data will not be easily caught, while IPIs show the process and errors are easily detected. Timing is included in the IPI and can be visualized, used for differentiating groups, etc. The reader is encouraged to consult the section Interactive Process Indicators by Example in [5]. IPM is a methodology that puts the healthcare professionals first, facilitating the understanding and easiness of exploration of the process indicators.

The structure of the rest of the paper is as follows: The application of IPM in the cardiology service at Salamanca Hospital, Spain, is explained in the Materials and Methods section. Afterwards, the results of applying IPM are shown in Results. Finally, next steps, limitations, and a comment on the COVID pandemics are shown in the Conclusion and Discussion section.

2 Materials and Methods

2.1 Data of Origin

For the study, EHR data collected from the Hospital Information Service (HIS) in the Cardiology Department was used.

Clinical data had been manually introduced in MediConnect®(Fleischhacker, Schwerte, Germany), a clinical process management software tool. The timespan for the data analysis was from December 2018 until August 2020. The records that were included are defined in Table 1.

Table 1. Records in the initial data provided.

Anonymized patient ID	Patient identifier
Mediconnect activity ID	Activity identifier
Activity Name	i.e. Nuclear medicine (NM) test, Magnetic Resonance (MR) test, Holter test, Computerized Tomography (CT) test, Structural intervention, Outpatient visit, etc.
Agenda	Sub-type of activity (e.g. Kind of outpatient visit: One-stop clinic, general hospital consultation).
Activity code	It identified the kind of test in a more specific way, e.g., for the ECHO test, the code referring to trans-thoracic test.

(*continued*)

Table 1. (*continued*)

Order Status	The action could be in one of the following statuses: Planned, Delayed, Running, Confirmed, Finished, Canceled
Order	Date and time when the clinician, nurse or administrative asked for a new activity.
Order scheduled Start	Date and time when the appointment was scheduled to start (the real start time was not available).
Order scheduled Finish	Date and time when the appointment was scheduled to finish.
Follow up request	It could be one of: – Rehabilitation – NM test – Holter – Hemodynamic test – Structural intervention – Spirometry – ECHO – Consultation – Implant – MR – CT
Reason or symptom(s) for the request	Logical values for one or more fields among the following: Asymptomatic, Dyspnea, Dizziness, Palpitations, Murmur, Syncope.
Diagnostic from the activity	Logical values for the following fields: Cardiac arrhythmia, Congenital cardiopathy, Ischemia, Valvular heart disease, Infectious Endocarditis, Aorta illness, Pericardiac illness, Structural damage, Lung hypertension, Heart failure, Myocardiopathy, Sudden death, Pulmonary embolism, Syncope.
Patient plan	The possible values were: Return to primary care, interconsultation, monographic consultation, request for tests, request for intervention, request for tests and results, etc.

Extra fields were also available but were already discarded at an initial data quality assessment stage, such as the logical field Patient discharge (in that case, Patient plan: return to primary care was more accurate).

2.2 Variables of Interest for the Clinician

The clinicians had in mind some aspects of the process that they wanted to measure and dive into. One aspect was the wait lists: They wanted to detect where and when bottlenecks occurred, along with time from primary care request for the cardiology department till the patient was attended. Another need was to discover long time-to-diagnosis and long time-to-treatment of patients with a cardiac disease.

Clinicians also needed to detect low level of coordination with the general practitioner for derivation and follow-up of patients. Another question was whether there were differences in clinical decision making between junior and senior doctors, specially through the number of requested tests. Gender and age inequities in diagnosis and requested tests were of importance, along with the

impact of Covid-19 pandemic in the number of consultations over the worst months of the first wave.

2.3 Methods in Data Rodeos

Data rodeos are sessions with all the actors involved (managers, PM analysts, clinicians, technical staff), where an interactive analysis of the data and latest process indicators is performed. Each derived IPI helps better understand, quantify, and qualify the process that is being studied. Duration of data rodeos can range from hours up to one month. The result of each data rodeo, the IPI, must be validated by healthcare professionals. The process indicator should allow the clinician to check the representation against the HIS. This leads to an increase of confidence in the model by the professional, e.g., if privacy allows it, the doctor can see any patient identification and see that the patient follows the process as observable in the IPI.

Data rodeos are separated into three stages: Shake down, Research and Production [3]. The initial stage requires aggressive data cleaning and fast PM discovery algorithms: It corresponds to the initial interviews between all the parties and the iterations in the process model must keep doctors attentive. The second stage, Research, can introduce long processing research tools and must introduce a more respectful data cleaning strategy: In the medical domain, *outliers* may be related to patients that follow a different path that is especially important to the doctors. The production stage is carried out after the research stage and the identified IPI is deployed in a live environment and it incorporates the creation of a dashboard for the hospital, load tests and security and privacy implementations. In this study, Shake down and Research stages are presented.

Anonymized data files were created by the IT staff at the hospital containing the information that was described in the Subsect. 2.1. For each IPI, the data files were ingested, generating a PM log. The log was filtered and processed. Afterwards PM discovery was applied, and a model was discovered. The model was processed and, after enhancement and conformance, the IPI was generated. With each generated IPI, a report with invalid traces and other data quality problems was created. This let all the actors find any problem in the data or the IPI. Apart from the soundness of the IPI, it must also be compelling and easy-to-understand and interact with, by the health professional [10]. Different ways of achieving an augmented model with metadata have been described, creating maps where color, node or transition representation size, tags, transparency, etc. change to show information about the process. During this study color was used to represent duration in each activity (represented in nodes) and number of traces (coded in the transitions), both with gradients with value ranges represented in the legend in each figure.

3 Results

3.1 First Data Rodeo

During the first Data Rodeo, an initial IPI was generated, as a base to work on. The main nodes were: 1. Request for One-stop clinic 2. First consultation 3. Consultation (any further consultations, after the first one). 4. ECHO: an echocardiography had been performed for the patient. 5. Discharge 6. Exitus (deceased)

Since the time spent in each node was unknown, this could not be introduced in the IPI, but in the case of the time from request to scheduled consultation.

Figure 1 shows a representation of the IPI where node color implies time spent at the node and transition colors show the number of executions that go from node to node (see Heatmap legend).

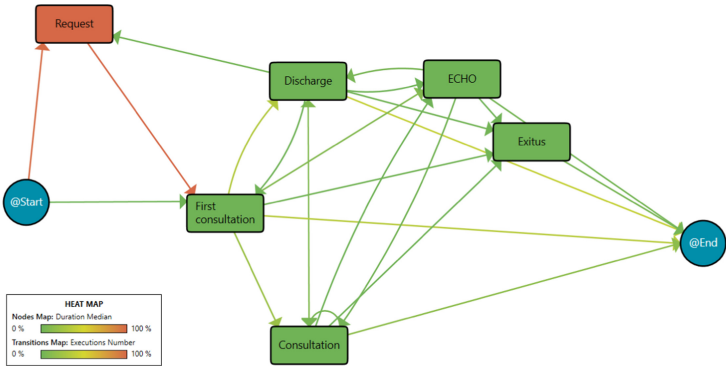


Fig. 1. Initial IPI from First Data Rodeo.

As mentioned earlier, the process at this stage needed further work, e.g. there are transitions that start at First Consultation and that cannot happen in reality. However, those details would be polished later.

In the IPI, most of the information remained in the model but was not visible at first glance, as in Fig. 1. The following extra data was introduced with filters: symptoms and diagnostics for each patient.

This data rodeo let us find that, as observable in Fig. 1, there is a high number of patients who do not need extra tests and are directly discharged. This was observed in the transitions First consultation → @End, and First consultation → Discharge, and it was an approximate measure of efficiency with around 50% patients discharged after the initial consultation.

Doctors were very interested in watching wait times and other variables depending on the symptoms. However, they were discarded since most consultation did not have the information. A second data rodeo was appointed to further explore the data and obtain a better IPI.

3.2 Second Data Rodeo

The clinicians wanted to measure the ECHOs that had been solicited to be performed during the consultation, since the number of ECHOs impacted the efficiency of the clinic. A node was created for the case of non-request of an ECHO during the consultation. A view of the process can be seen in Fig. 2.

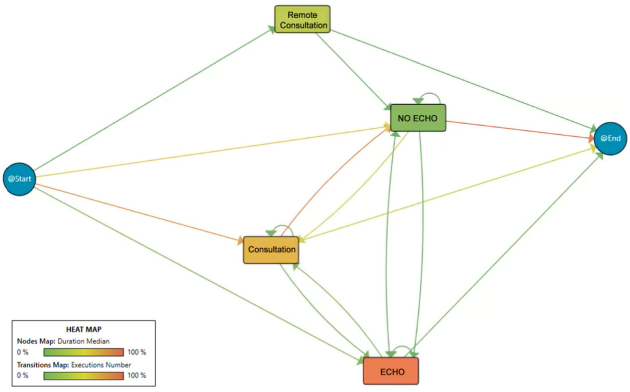


Fig. 2. Second IPI with Consultation and two nodes for the extra ECHO test and the absence of it.

Through the creation of trace metadata groups for doctor category (registrars, consultants), the proportion of the during-consultation ECHOs could be easily seen.

The data about clinicians who had attended the patients in the one-stop had been introduced, since one question was if registrars (junior clinicians) asked for more tests than consultant doctors (seniors). The IT team provided the data, with the following fields, among others: Mediconnect activity ID, Clinician name. The rest of the fields were discarded since they were not finally used to create the IPI. With this field, the clinicians involved in the generation of the IPI classified doctors between consultants and registrars. This information was included in each trace as Type of doctor.

The clinicians involved could see the percentage of patients that underwent an ECHO inside the consultation. However, it was estimated later that a relevant percentage of doctors did not fulfill the ECHO forms on the EHRs due to complexity and time allocation per patient in the clinics, which led to a lack of information in this regard. Thus, intra-consultation ECHOs had to be removed from the process.

3.3 Third Data Rodeo

Extra-consultation ECHO tests could be checked, along with other tests and interventions. Thus, if any test was requested, it was introduced in the IPI as

trace metadata. Age and gender were also included, as well as the type of doctor. This would let clinicians find out e.g., if registrars asked for more ECHO tests than consultants. Also, if there were more men than women with heart problems, etc.

It was also decided that, since the interest was in the first consultation for a referred patient from primary care, when there was more than one consultation, only the first consultation would be considered, along with the extra tests and the discharge. Further events in time would be discarded for the process model.

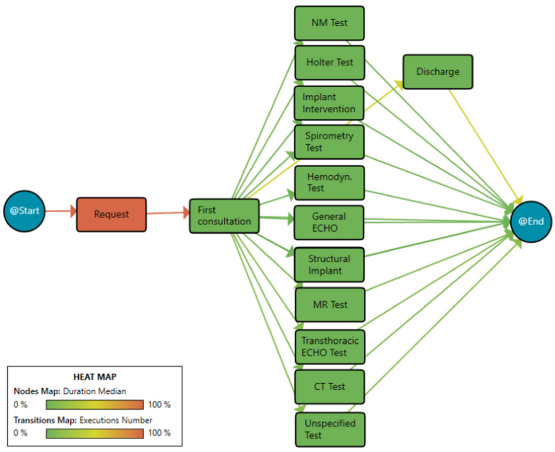


Fig. 3. Final IPI view.

Figure 3 shows a view of the final IPI. It included the following trace metadata: – Month of consultation – Year of consultation – Time to next consultation, in days (−1 if data was not available) – Gender – Age – Test/intervention required – Diagnostics (array with the different diagnostics) – Type of doctor.

The IPI was created with 15 nodes, as depicted in Fig. 3. The transition probability from consultation to discharge was high, 47%.

The number of traces was extracted for each month and year in the available data. This is shown in the column chart in Fig. 4. It can be observed that 2020 (January to August, amidst the pandemic) had a lower number of clinics than 2019.

Transition probability in 2019 from first consultation to discharge was 50% while it was reduced to 17% in January to August in 2020. This may be because of patients not going to the clinic due to the lockdown in the first pandemic wave. This would increase the percentage of patients that would need an extra test or intervention.

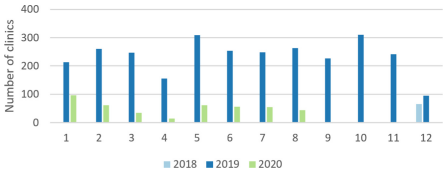
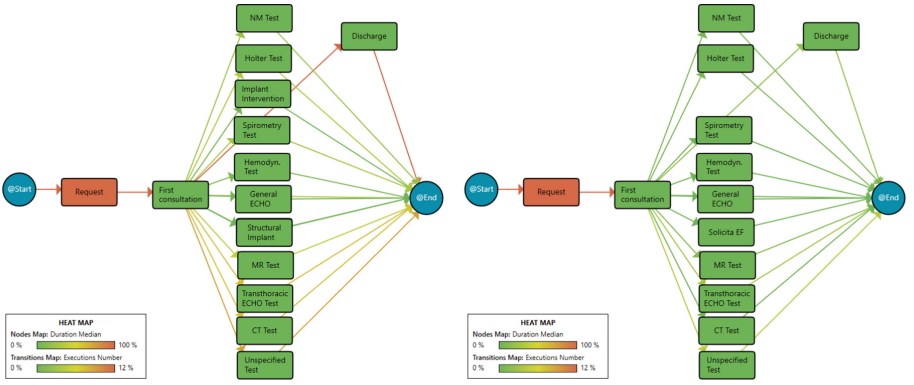


Fig. 4. Number of traces per month, equivalent to number of new patient clinics.

The percentage of unspecified tests/intervention requests went up from 10% in 2019 to 31% in 2020. Also, the most demanded test in 2020 was trans-thoracic ECHO (23%) compared to an 8% in 2019.

All these data were easy to interpret at one glance watching the IPI view (Fig. 5) by looking at the color-coded transition probabilities.



(a) 2019 IPI visualization (b) 2020 IPI visualization

Fig. 5. IPI view with same color gradient for 2019 and 2020.

It could also be seen that age groups had different number of clinics depending on the month. In Fig. 6, August and December 2019 are compared in local percentage (distribution of 100% between the age groups). Elderly people were more treated in December than in August, and adults were the ones that coped with that relative decrease in the same months.

Registrar-requested external ECHOs against those requested by consultants were compared. Although more clinics were performed by consultants (68.9%) than by registrars (37.1%), more than half of the ECHO tests were requested by registrars (specifically, 57,1%).

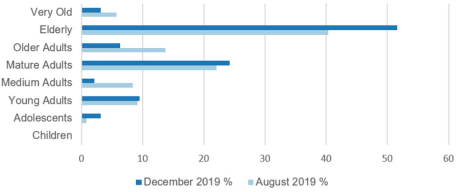


Fig. 6. Age group clinics, in relative percentage, depending on the month.

4 Conclusion and Discussion

After the initial IPM Shakedown, clinicians were provided with a comprehensive and easy-to-use tool that allowed them to answer most clinical questions proposed in the first meeting.

During the last years national and European cardiology societies have stablished metrics and benchmarks that every cardiology department should meet in the outpatient setting [6]. One of the key markers is average wait for a first consultation since the time a new referral is done. With IPM, clinicians observed the median waiting time was 19 days and 20 h, showing a clear improvement point in comparison with national standards. Overall view of gender and age distribution on the patient cohort was obtained, providing a better understanding of population that access a cardiology department and correlating with general population aging. These data were extracted by the clinicians by inspection into the IPI.

Through IPM analysis clinicians could define and classify the outcome of the clinic in big generic cardiac syndromes or the absence of a specific diagnosis in patients with a structural normal heart. The discharge rate from the clinic was 47% reaching the acceptable benchmark set by expert consensus mentioned before. However, improving communication with other specialists and primary care and the implementation of novel alternatives such us e-consultation could be an option to reduce even more unnecessary referrals.

Reducing the number of unnecessary requested tests is of key importance for a public funded healthcare system. Prior to rationalisation of diagnostic and interventions, it is fundamental to know the exact volume and statistics of requests generated by the one-stop clinic. This task was successfully achieved with the analysis of the process. The IPI is effectively an audit of the outpatient service that will promote the update of clinical protocols, and refreshment educational sessions, reducing unnecessary and costly tests, benefiting both patient safety and the heath system economy.

Time allocated per patient for a one-stop clinic is usually enough for a general cardiologist to perform an external ECHO if deemed necessary after formal clinical interview and physical examination. The existence of too many early requests of extra ECHOs is perceived as a failure of the main goal of this kind of clinics. The indicator of 10% of ECHOs directly requested from one-stop clinic shows an improvement opportunity.

The SARS-Cov2 pandemic had a major impact in the healthcare system during 2020, lock-downs and resource reallocation to an over-saturated inpatient care dropped the number of first consultations [12]. Currently, keeping up with missed appointments is a struggle in outpatients services. IPM analysis has helped the cardiology department to quantify in an accurate way the damage made to the outpatient service during the worst months of Covid spread.

In conclusion, clinicians were provided with a useful tool for data analysis. The results through IPM were used as a complete audit of outpatient service deriving into clinical protocol changes and exposing improvement opportunities. Developments in the IPI are still to come (such as introducing the distance between the patient and the hospital, and other data that will help clinicians with new questions), but it is mature enough to answer the initial queries.

Acknowledgment. This activity has received funding from EIT Health (www.eithealth.eu) ID 20328, the innovation community on Health of the European Institute of Innovation and Technology (EIT, eit.europa.eu), a body of the EU, under Horizon 2020, the EU FP for Research & Innovation. https://www.eithealth.eu.

References

1. Borges-Rosa, J., Oliveira-Santos, M., Simoes, M., et al.: Process mining tools: where should we build another PCI centre to reduce STEMI mortality? Eur. Heart J. **41**(Supplement 2) (2020)
2. Falces, C., Sadurní, J., Monell, J., et al.: One-stop outpatient cardiology clinics: 10 years' experience. Rev. Esp. Cardiol. **61**(5), 530–533 (2008)
3. Fernandez-Llatas, C.: Bringing interactive process mining to health professionals: interactive data rodeos. In: Fernandez-Llatas, C. (ed.) Interactive Process Mining in Healthcare. HI, pp. 119–140. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-53993-1_8
4. Fernandez-Llatas, C.: Interactive Process Mining in Healthcare. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-53993-1
5. Fernandez-Llatas, C.: Interactive process mining in practice: interactive process indicators. In: Fernandez-Llatas, C. (ed.) Interactive Process Mining in Healthcare. HI, pp. 141–162. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-53993-1_9
6. González-Juanatey, J.R., Virgós Lamela, A., García-Acuña, J.M., Pais Iglesias, B.: Clinical management in cardiology. Measurement as a means to improvement. Rev. Esp. Cardiol. **74**(1), 8–14 (2021)
7. González-Llopis, F., Palazón-Bru, A., et al.: Clinical effectiveness of a cardiology outpatient management plan to reduce inefficiency in consultations. Postgrad. Med. **133**(2), 166–172 (2021)
8. Hakim, A.: Improving healthcare systems through process mining. In: 2020 IEEE 23rd International Multitopic Conference (INMIC), pp. 1–4 (2020)
9. Kusuma, G., Hall, M., Johnson, O.: Process mining in cardiology: a literature review. IJBBB **8**, 226–236 (2018)

10. Mendling, J., Djurica, D., Malinova, M.: Cognitive effectiveness of representations for process mining. In: Polyvyanyy, A., Wynn, M.T., Van Looy, A., Reichert, M. (eds.) BPM 2021. LNCS, vol. 12875, pp. 17–22. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-85469-0_2
11. Van der Aalst, W.: Process Mining: Data Science in Action. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49851-4
12. Wosik, J., Clowse, M.E.B., Overton, R., et al.: Impact of the COVID-19 pandemic on patterns of outpatient cardiovascular care. Am. Heart J. **231**, 1–5 (2021)