Interactive Process Mining Applied in a Cardiology Outpatient Department

Juan José Lull[1 [0000−0002−4399−950X]], Adrián Cid-Menéndez[2 [0000−0002−5135−0305]], Gema Ibanez-Sanchez[1 [0000−0003−1824−281X]], Pedro Luis Sanchez[2], Jose Luis Bayo-Monton, Vicente Traver[1 [0000−0003−1806−8575]], and Carlos Fernandez-Llatas[1,3 [0000−0002−2819−5597]]

[1] Instituto Universitario de Tecnologías de la Información y de las Comunicaciones (ITACA), Universitat Politècnica de València, Camino de Vera S/N,Valencia, Spain
{jualulno, jobamon, geibsan, vtraver, cfllatas}@itaca.upv.es
[2] Cardiology Department, University Hospital of Salamanca, Instituto de Investigación Biomédica de Salamanca (IBSAL), Paseo S. Vicente, 58-182, 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 for 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. A final Interactive Process Indicator (IPI) was generated. During Data Rodeos data quality problems arose and were tackled, the data was modelled, process activities were discovered, etc.
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.
The results confirm that Interactive Process Mining helped the clinicians gather deeper information about the outpatient department, waiting lists, requested tests, the behavior of the clinic patient volume with a special focus on the impact of the first SARS-CoV2 pandemic wave.

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

1 Introduction

Salamanca University Hospital is a tertiary public centre that covers health care for a population of 350,000 patients and has 933 inpatient beds. The cardiology department has become, over the years, the referral service for specific cardiovascular diseases and has developed a high-volume outpatient clinic, with more than 25,000 consultations every year. Focusing on outpatient care, during the last decade new patient referrals were managed in a "one-stop" rapid assessment clinic. This clinic is essentially a connection between the Cardiology department and primary care and works as an entrance door for first-time patients to the Cardiology service.

The model of one-stop clinic has spread over the last years across European cardiology services as a solution for managing first visits [7]. This clinic usually assigns a higher amount of time per patient and the cardiologist running the clinic has the availability and necessary skillset for running simple, highly efficient tests such as electrocardiogram and echocardiography (ECHO) which allow the team to discriminate between patients relevant cardiac diseases and those who need to be scheduled for furthers tests or clinic reviews. The main goal of this specific model of clinic is efficiency, i.e., achieving a high percentage of direct discharges, avoiding unnecessary follow-ups and tests. As part of a public funded health system, with inherent limited resources, modern cardiology departments have adopted this model of outpatient care to improve patient safety, to reduce unnecessary tests and procedures that come with complications and side effects, and also to reduce workload and thus decrease the number of patients that will end unnecessarily in follow up clinics [2].

The one-stop clinic is led by cardiologists. The clinic is run most days of the week depending on space availability, period of the year and inpatient service requirements. Every doctor has the support of nursing staff and full access to specific extra test requests within the cardiology department. The extra tests are performed, most of the time, on a different scheduled date. The tests generate a routine follow up review.

The specific type of clinics 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 would benefit from a closer surveillance.

Shorter waitlists for first consultation are a quality parameter in healthcare and the result of one-stop clinics. Shorter waiting times increase patient satisfaction and improve relationship and communication with primary care and other specialists. This also has a direct impact on Emergency visits and admissions that can be prevented with an early referral to the clinic.

The cardiology department has, since 2015, dedicated resources to store and analyze the treatment processes carried out in the service. Electronic Health Records (EHR) are stored and accessed through Mediconnect(r), an ERP developed under an agreement with Medtronic (Minneapolis, MN, USA).

In the context of a European funded project, the cardiology department wanted to visualize and take decisions based on the processes related to 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. Even though clinicians are often the most interested in the processes, managers are the ones who can introduce big changes in the processes. Thus, they need to be included in the modelling of the process. IT professionals are the ones in charge of the data: They know how to access the EHR, the architectural design, the technical limitations, etc. They are thus key in the generation of Process Indicators.

PM has been applied to different Cardiology departments. In [8], it was applied in a Pakistan Cardiology Hospital without real event logs, but 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] is 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 [10]. In that study, they focused on what specific disease each paper had studied and which PM method had been used.

IPM lets the stakeholders define Interactive Process Indicators (IPI) from the data and the questions that the clinicians and managers have. IPM is a methodology that puts the healthcare professionals first, facilitating the understanding and easiness of exploration of the process indicators.

IPIs are developed iteratively in sessions with the PM analyst. The implementation phases include an initial Shakedown, a Research phase and a final Production phase [3]. Data Rodeos are sessions that allow the analyst to create and refine IPIs that enable the experts to grasp the meaning of, assess and optimize the processes.

As shown in the following section, during the Shakedown and Research phases, three Data Rodeos were carried out, corresponding to three IPIs, each one replacing the previous. The final IPI was deemed as satisfactory for the clinicians, as they were able to find the answers to many questions about their service and were able to navigate the IPI without technical assistance, as we will see in the Results and the Conclusion sections.

This study shows how Data Rodeos were carried out in a Cardiology Service until the final IPI was achieved IPI. Finally, the findings about the process of the one-stop clinics are presented, along with a conclusion. To the best of our knowledge, this is the first study that analyzes cardiology one-stop clinics with PM, and more specifically with IPM.

2 Materials And Methods

For the study, EHR data collected from the Hospital Information System in the Cardiology Department was used. Clinical data had been manually introduced in a specific department software (MediConnect®; Fleischhacker, Schwerte, Germany), a clinical process management software tool.

The timespan for the data analysis was from December 2018 till August 2020. The records that were included are described in Table 1.

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	It identified a sub-type of process (e.g. Kind of outpatient visit: One-stop clinic, general hospital consultation, cardiotoxicity consultation... Or kind of ECHO: stress test ECHO, complex ECHO, basic ECHO).
Activity code	It identified the kind of test in a more specific way, e.g., for the ECHO test, there were different codes depending on the contrast applied, or no contrast.
Order Status	The action could be in one of the following statuses: Planned, Delayed, Running, Confirmed, Finished, Canceled
Order	Entry Date and Date and time when the clinician or nurse 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 test − Hemodynamic test − Structural intervention − Spirometry test − ECHO test − Consultation − Implant − MR test − CT test
Reason or symptom(s) for the request	Boolean values for one or more fields among the following: Asymptomatic, Dyspnea, Dizziness, Palpitations, Murmur, Syncope.

Diagnostic from the activity	Boolean 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 (Infection or myocarditis), 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.

Table 1: Records in the initial data provided.

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

2.1 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, and differentiate the time to treatment depending on specific treatments.

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 worse months of the first wave.

2.2 Data Rodeos

At the initial Phase of process modelling -Shakedown- rude cleaning of the data was needed to move away from the Spaghetti Effect (a process with so many nodes and connections that it is useless in terms of human interpretation) [5].

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. It could only be introduced 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).

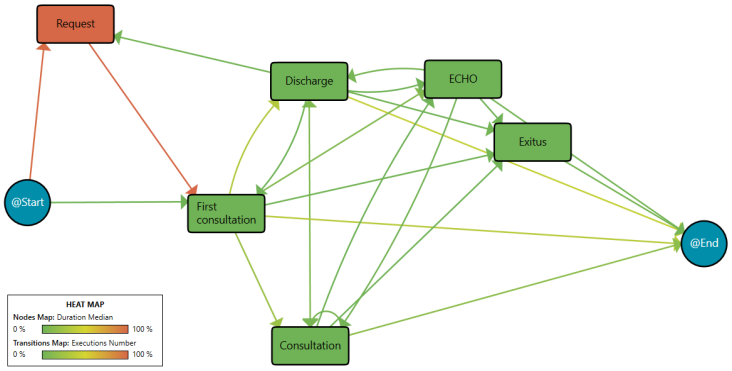


Fig. 1: Initial IPI from First Data Rodeo.

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 Figure 1. The following extra data was introduced with filters: symptoms and diagnostics for each patient.

Doctors were very interested in watching the different wait times, etc. depending on the symptoms. At the end of this phase, symptoms were discarded as most consultations did not have information about the patient's symptoms. A second Data Rodeo was appointed to further explore the data and obtain a better IPI.

During the **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 Figure 2.

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 ha been introduced, since one question was if registrars (junior clinicians) asked for more tests than consultant doctors (seniors). The IT team provided the data,

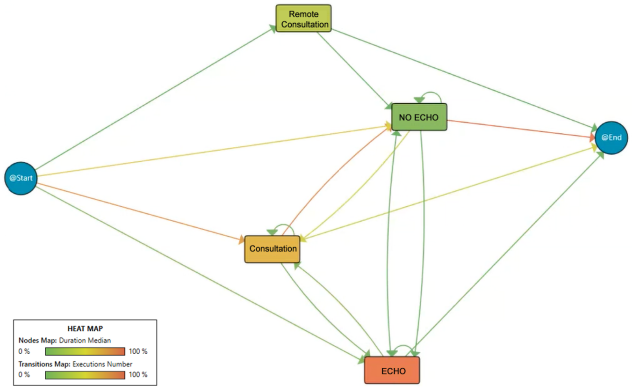


Fig. 2: Second IPI with First Consultation and two nodes for the extra ECHO test and the absence of it.

with the following fields, among others: – Mediconnect activity ID – Clinician name.

The other 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.

In the **third Data Rodeo**, however, it was discovered that not all clinicians used to enter the request of during-consultation ECHOs in the system, because of the difficulty of the process for the request. Thus, the "ECHO / NO ECHO" events apparent at Figure 2, were removed. Extra-consultation ECHO tests could be checked, along with other tests and interventions. Thus, if any test was requested after the consultation, it was introduced in the IPI as trace metadata. Age and gender were also included, as well as the doctor's category (registrar or consultant cardiologist). 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.

Furthermore, the events were being extracted as one per line. However, to obtain information about specific tests or interventions that were performed, the request for the test/intervention and the scheduled date (in different rows) had to be mixed, so traces were no more "one or multiple events in one row" but "one or multiple events in multiple rows" [3] in the data file. There was not an exact field that joined the test/intervention with the consultation, though. Both the consultation and test rows coincided in two fields: the anonymized patient ID and the date of the scheduled test/intervention was the same as the consultation date.

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.

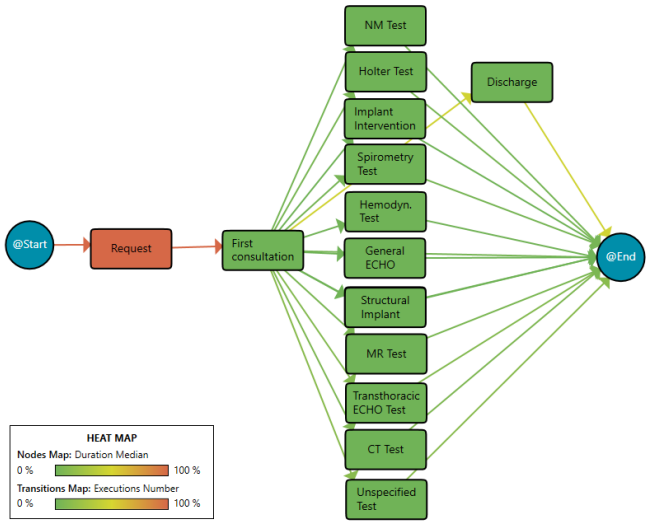


Fig. 3: Final IPI view.

Figure 3 shows a view of the final IPI.

The IPI 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.

Clinicians started working with the IPI using the PMApp tool [9]. The results of the Data Rodeos and clinician investigations are shown in the following section.

3 Results

3.1 General results

The number of traces, after processing the input data, was around 3,211. They matched with 3,211 patients that had had a one-stop clinic. There were some traces that were discarded due to lack of information or wrong data. This number changed slightly as the IPI was enhanced through the data rodeos.

3.2 Data Rodeos

The **first Data Rodeo** let us find that, as observable in Figure 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 directly discharged.

In the **second Data Rodeo**, 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 physicians 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.

In the **third Data Rodeo**, an IPI with 15 nodes was created. As depicted in Figure 3, it included: – Start and End nodes – Initial consultation request – First consultation – Tests: ECHO, MR, CT, NM, normal ECHO, holter, spirometry, hemodynamics – Interventions: structural, implant – Discharge.

As observable in Figure 3, the transition probability from consultation to discharge was high. In fact, it was 47%.

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

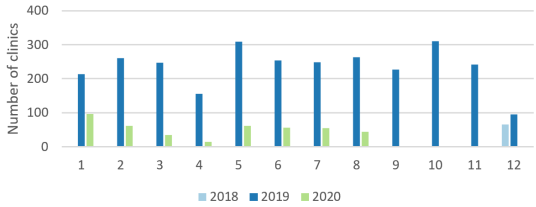


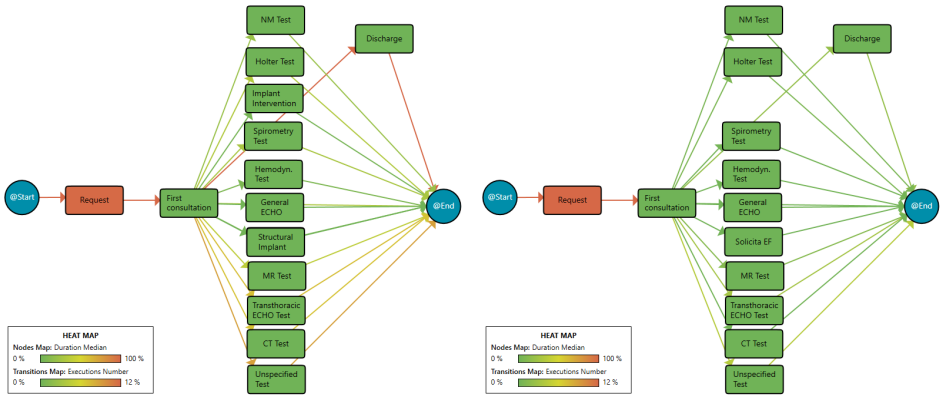
Fig. 4: Number of traces per month, equivalent to number of new patient clinics.

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.

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 transthoracic ECHO (23%) compared to an 8% in 2019.

All these data were easy to interpret at one glance watching the IPI graphical representation (Figure 5) by looking at the color-coded transition probabilities. In 2020, the transthoracic ECHO had a higher number than the discharge.

It could also be seen that age groups had different number of clinics depending on the month. In Figure 6, August and December 2019 are compared in local percentage (distribution of 100% between the age groups). Elderly people were



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

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

more treated in December than in August, and adults were the ones that coped with that relative decrease.

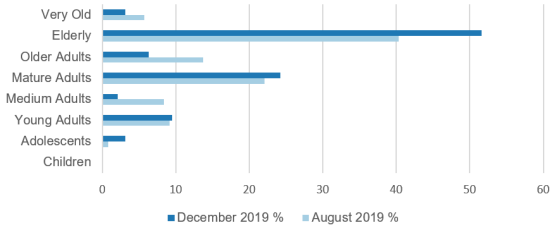


Fig. 6: Age group clinics, in relative percentage, depending on the month.

There was also interest in finding if the registrar cardiologists asked for more tests than consultants, specifically ECHO tests. The intuition was that the juniors would ask for more extra tests, due to lack of experience and the need to have extra confirmation of their suppositions. This was indeed the case, with most clinics performed by consultants (68.9%) than by registrars (37.1%), but more than half of the ECHO tests were requested by registrars (specifically, 57,1%).

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 hours, 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, with the objective of reducing unnecessary and costly tests, benefiting both patient safety and heath system economic efficiency.

Time allocated per patient for a one-stop clinic is generally enough for a general cardiologist to perform a departmental (protocolised) ECHO if considered 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 our healthcare system during 2020, lockdowns and resource reallocation to an oversaturated inpatient care dropped the number of first consultations [12]. Measuring the impact of Covid in a comprehensive way allowed the team to understand the consequences after the pandemic and this data could be used in a preliminary way to predict the number of consultations that would need to be rescheduled and to understand the approximate number of patients that missed a necessary appointment due to an overwhelmed health system. 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 worse 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 European Union, 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? European Heart Journal **41**(Supplement 2) (11 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 (May 2008)
3. Fernandez-Llatas, C. (ed.): Bringing Interactive Process Mining to Health Professionals: Interactive Data Rodeos, pp. 119–140. Springer, Cham (2021)
4. Fernandez-Llatas, C.: Interactive Process Mining in Healthcare. Springer (2021)
5. Fernandez-Llatas, C., Martinez-Millana, A., Martinez-Romero, A., Benedí, 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 (2015)
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 (Jan 2021)
7. González-Llopis, F., Palazón-Bru, A., Mares-García, E., Carbonell-Torregrosa, M.L.A., Bertomeu-Martínez, V., Gil-Guillén, V.F.: Clinical effectiveness of a cardiology outpatient management plan to reduce inefficiency in consultations. Postgrad Med **133**(2), 166–172 (Mar 2021)
8. Hakim, A.: Improving healthcare systems through process mining. In: 2020 IEEE 23rd International Multitopic Conference (INMIC). pp. 1–4 (2020)
9. Ibanez-Sanchez, G., Fernandez-Llatas, et al.: Toward value-based healthcare through interactive process mining in emergency rooms: The stroke case. Int. J. of Env. Research and Public Health **16**(10) (2019)
10. Kusuma, G., Hall, M., Johnson, O.: Process mining in cardiology: A literature review. International Journal of Bioscience, Biochemistry and Bioinformatics **8**, 226–236 (10 2018)
11. Van der Aalst, W.: Process mining: Data science in action. Springer (Jan 2016)
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 (Jan 2021)