Evaluation of operational process variables in healthcare using process mining and data visualization techniques

Jimmy Armas Aguirre[1], André Coronado Torres [1], Misael Evangelista Pescoran[1], Santiago Aguirre Mayorga[2]

1 Universidad Peruana de Ciencias Aplicadas, Lima, Perú

jimmy.armas@upc.pe, u201112897@upc.edu.pe, u201111694@upc.edu.pe

[2] Pontificia Universidad Javeriana, Bogotá, Colombia

saguirre@javeriana.edu.co

Abstract– In this paper, a reference model is proposed for the evaluation of operational processes variables in healthcare using process mining and data visualization techniques. For this reason, the PM2 methodology is used as a reference to conduct projects oriented to the evaluation of data collected in business processes, including data visualization techniques, with the purpose of reducing the acquisition time of knowledge related to the processes of institutions of the healthcare sector. The proposed model is based on the application of a set of data visualization techniques to reduce the knowledge acquisition gap presented by process mining. The model consists of 5 stages: 1. Extraction, 2. Event processing, 3. Process mining, 4. Data visualization and 5. Evaluation of results. A testing scenario was defined in a Clinic network in Lima (Peru) to validate the proposed model and the surgery process was chosen, since it is critical for the organization. The results showed the existing bottleneck in the surgery process, between the activities of registering and preparing the patient. This allowed to take corrective measures between the activities to optimize the process cycle time. Likewise, a sequence was identified in the activities that had not been previously detected in the process documentation, these represented 2.6% difference, so the documented process was modified to achieve a 99.6% affinity.

Keywords- Process mining, data visualization, healthcare, business process.

I. INTRODUCTION

The Peruvian social security institution, Essalud, provides various types of insurance to the population, the number of insured rises to over 11 million Peruvians representing 35.2% of the total population [1]. During February 2015, reports were issued seeking to show that the insured generate 10,725 requests (consisting of questions, petitions and claims) of which 65% are focused on claims. Of these claims, it was observed that it is given in greater proportion by face-to-face attention channel and it is originated due to a deficient provision of health services and poor care quality that represents 30% of the total requests [2].

In the last decade, information technology has evolved in order to provide support for the development of different industries from different approaches, including the health sector, to provide quality and efficient services in their processes [3]. With the development of the technology, the information stored in a variety of information systems has increased exponentially, which generates a bias to analysts due to the large amount of information that need to be analyzed. At this point, process mining allows organizations to leverage this amount of information, as it is a discipline that allows detecting deviations, predicting delays, supporting decision making and recommending process redesigns. This discipline is used to measure the conformity of the traces with a prescribed (real) process model or to improve a known process model with additional information extracted from the event logs [4].

However, the information that results when applying the aforementioned discipline is represented in a process map or flowchart (the representation of a process) and through numerical data, which must be analyzed by specialists quickly, but it is difficult to understand its totality in a short term. Given this situation, we can see a gap between the resulting data and the analysis carried out by specialists. Therefore, there are two main reasons that motivate the development of this reference model for the evaluation of variables in the healthcare sector using process mining and data visualization.

On one hand, there is the aforementioned gap between the data resulting from applying process mining and the analysis of these, since it is considered a challenge to complement the aforementioned discipline with other types of techniques that allow an analysis of the data coming from data visualization.
It is easier to understand the graphics and images in a relatively intuitive way, since they will reveal a great variety of information [5], generate a better understanding of large and complex data, as well as to exploit the capabilities of humans to see patterns in unstructured data [6].

On the other hand, there are several developments, applications, cases, studies and implementation of tools that support the analysis of data extracted from business processes, considering a perspective of data mining and process mining, in various sectors and areas, standing out the industrial sector, however, no studies have been carried out at this level in the healthcare sector [7].

This article is structured as follows. First section presents this introduction, second section starts with a literature review, third section describes the phases of our proposal to outline a validation model for the surgery process. Finally, the conclusions and results obtained in the case study will be addressed.

17th LACCEI International Multi-Conference for Engineering, Education, and Technology: "Industry, Innovation, And Infrastructure for Sustainable Cities and Communities", 24-26 July 2019, Jamaica.

1

II. LITERATURE REVIEW

Customer service processes are of vital importance in any institution of the healthcare sector, if these are structured in an efficient way they can save the lives of a large population, due to this, previous studies that have applied Process Mining in various industries were revised with the purpose of observing the main results and their contribution to their respective industries.

A. Process mining algorithms

As can be seen in **Table I and II**, the algorithms used in studies of various industries with focus on financial, healthcare and retail respectively were identified. From these studies, it was possible to identify the behavior of customers to increase sales in retail, the loan processes in financial entities were evaluated to identify bottlenecks and in the healthcare sector, it was used to evaluate the attention of emergency areas and intensive care units in which the business processes could be discovered, but there were no conclusive results, due to which the health sector was selected as the focus of the project.

TABLE I
PROCESS MINING ALGORITHMS

Algorithms	Description	Reference
Alpha Miner	Discover a Petri network. Fast, although the results generated when applying the algorithm are not completely reliable due to overfitting problems.	[15]
Heuristic Miner	Discover a heuristic network quickly using a probabilistic approach, the results when applying this algorithm is aligned with the process when it comes to noise.	[16]
Inductive Miner	Discover a process tree or a Petri network, it generates good results when it comes to infrequent behavior and large event logs, it also has a guaranteed robustness.	[24]
Fuzzy Miner	Discover a fuzzy model, optimal when it comes to unstructured behavior, it is also fast and complete all the traces of the process.	[29]
Genetic Miner	Discover a heuristic network using genetic algorithms and it depends on the configuration of the parameters to generate reliable and efficient results.	[22]
Evolutionary Tree Miner	Discover a process tree using guided genetic algorithms based on the quality dimensions of the model. It guarantees solidity, but it can not represent all the possible behaviors due to its structured nature by blocks.	[20]
ILP Miner	Discover a Petri Network when solving problems of Integer Linear Programming (ILP), the results when applying this algorithm tend to present a perfect aptitude, but in general a poor precision, besides it is not so efficient when processing large quantities of event logs.	[23]

TABLE II
PROCESS MINING APPLICATION SECTORS

Algorithms	Sector	Reference	Fitness (f)	Precision (p)	Simplicity (s)	Generalization (g)
Alpha Miner	Retail	[17]	0.934	0.614	0.942	0.871
	Financial	[20]				
Heuristic Miner	Retail	[17]	0950	0.955	0.893	0.813
	Health	[7], [19]				
	Financial	[20]				
Inductive Miner	Health	[21]	0.810	0.810	0.500	0.500
Fuzzy Miner	Retail	[17]	0.950	0.900	0.893	0.950
Genetic Miner	Health	[7]	0.939	0.853	0.734	0.683
	Retail	[17]				
	Financial	[20]				
Evolutionary Tree Miner	Financial	[20]	0.934	0.992	0.962	0.917
ILP Miner	Financial	[20]	1	0.633	0.893	0.839

B. Tools

In the same way, process mining tools were evaluated with the purpose of identifying which ones provides better user experience and allows to use data visualization concepts in order to adequately analyze and interpret the variables of a business process. **Table III** shows the main tools of process mining, in which it is observed that ProM is the "open source" tool that allows modeling processes of different notations and has an unlimited capacity, although it lacks usability which generates a low user experience. On the other hand, Celonis, a tool that works with SAP ERP, allows usage of data visualization concepts with an excellent user experience and allows the import of data with an unlimited capacity of event logs. In addition, this tool uses the fuzzy miner algorithm, which is one of the best, as shown in **Table II**. The above are the main factors by which this tool was selected to support the proposed model. The licenses provided by Celonis correspond to a university alliances agreement that allows the nonproductive use for academic research.

TABLE III
COMPARISON OF PROCESS MINING TOOLS

TOOL	ProM	Disco	Celonis	QPR Process Analyzer
References	[10], [12], [13], [14]	[11], [12], [13]	[8], [9], [12], [13]	[13], [14]
Import Type Support	MXML, XES	CSV, XLS, MXMl, XES y FXL	CSV, XLSX, XES, DBF	XES
Capacity (Amount of Event Logs)	Unlimited	5 million events	Unlimited	10 million events
Output model notation	BPMN, Work Flow, Petri Nets, EPCs, Transition Systems, Heuristics	Fuzzy model	Fuzzy model	Work Flow

Monthly cost (Dollars)	Free	$ 1618	$ 598 - 1929	$ 1126
Number of Users per License	Unlimited	1	1	1
Support Time (Months)	-	Unlimited (Validity of the service)	Unlimited (Validity of the service)	1
License	-	X	X	X
Platform	On Premise	On Premise	On premise/ SaaS	On Premise
Filtering data	X	X	X	-
Process visualization	X	X	X	X
Process Comparison	X	-	-	-
Performance reports	X	X	X	X
Track traces	X	-	-	-
Usability	-	X	X	X

III. PROPOSAL: REFERENCE MODEL

Based on the research carried out, a reference model that takes the phases of the PM2 methodology is proposed for its design and was complemented with data visualization techniques as mentioned by Telea [27] and Chittaro [28], in which they delimit stages for the construction of representations graphics. This model is proposed to evaluate the variables used in the various processes of the healthcare sector. Also, the model focuses on using data visualization techniques to reduce the acquisition time of knowledge related to the execution of business processes.

A. Foundation of the contribution

As mentioned above, after applying the techniques and algorithms of process mining, a gap is perceived when analyzing the resulting data. Likewise, there are several investigations related to the analysis of data that are extracted from business processes in various sectors and industries, but no studies have been done at this level in the healthcare sector [7], where agility and proper business process monitoring are of vital importance.

In response to this need, a reference model is proposed for the evaluation of variables in the healthcare sector using process mining and data visualization techniques, which will allow to discover the real process model, compare it with the documented process, and in turn, represent the resulting data through graphic or pictorial representations that will take advantage of the ability of human beings to recognize patterns and trends.

B. Specific View

The proposed platform contains the following roles and tools:

1) *Medical expert*: Professional in charge of the medical process to be evaluated.
2) *Process mining analyst*: Professional with studies related to Process Mining and with knowledge in the use of Celonis.
3) *Celonis*: Process mining tool.
4) *Graphics*: Representations by means of figures or signs that allow to understand in a fast way the information related to the processes.

C. Process View

The proposed model is composed of five related stages. On the first stage, data and process models are extracted from the information systems supported by the healthcare institution. The extracted data is processed to generate the event log, these event logs are used to apply the process mining and data visualization techniques. The choice of data visualization techniques depends on the number of dimensions and measurements that are needed based on what you want to analyze. That is, for example, if we want to analyze a single variable in a continuous time interval, we can use a histogram, but not a pie diagram. Finally, the results are analyzed through the Celonis tool. **Figure I** shows the stages of the proposed model.
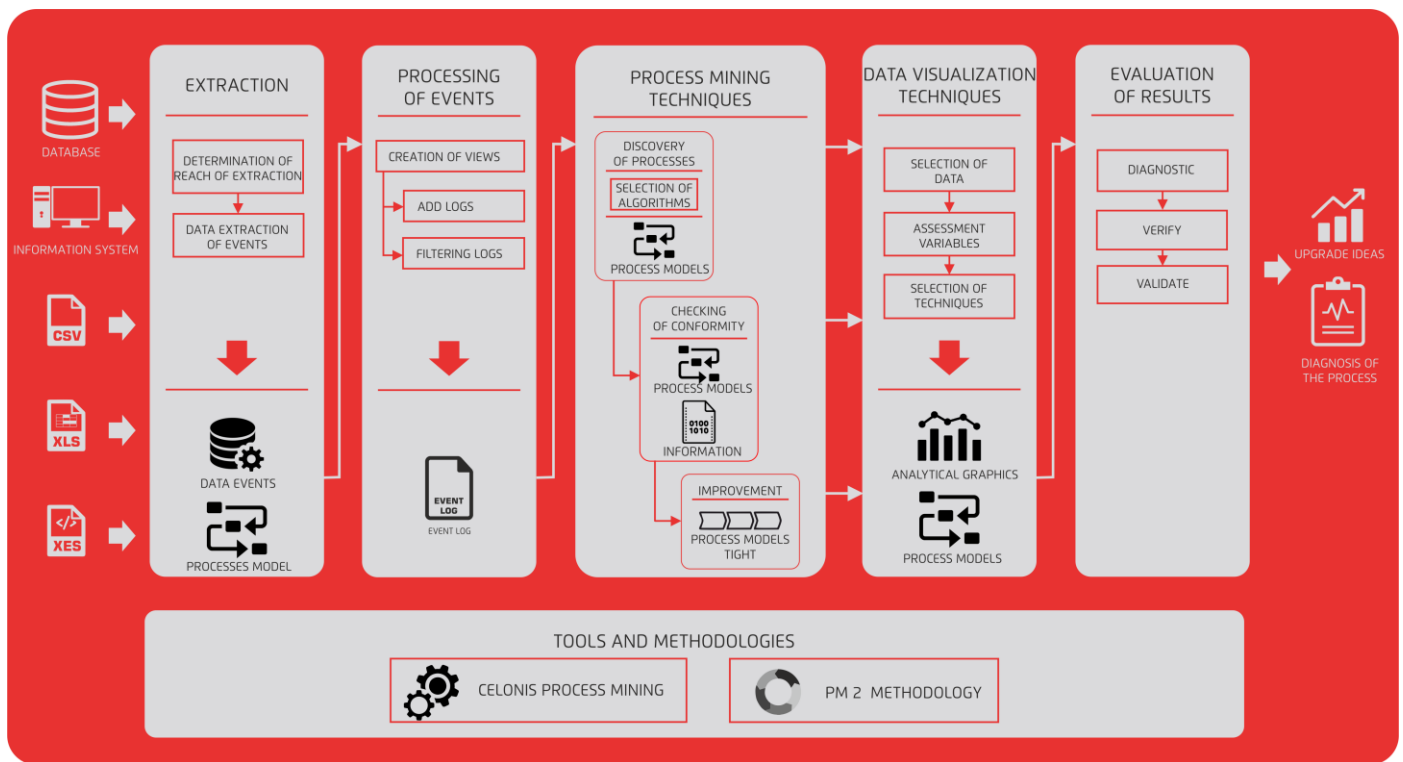
Fig. I Reference model for the evaluation of variables in the health sector using Process Mining and Data Visualization

1. Extraction

This stage consists of extracting the data of events and process models, although the latter are optional. The information systems that support the execution of the business processes is the main input to be able to execute this stage, resulting in data of the extracted events, which is a collection of events without a predefined case concept.

This stage is composed of two activities: scope determination and extraction of event data.

• Scope determination: This activity involves determining the scope of the data extraction, processes and variables are going to be analyzed since, based on this, the data of the event will be extracted.
• Some questions to consider are: With what level of granularity will the data be extracted? What period of history will be covered? What attributes will be used and how will they be selected?
• Extraction of event data: This activity is the consequence of the delimitation of the scope, the data related to the selected process of the relevant information systems is collected.

2. Event processing

This stage consists of creating the event logs using the different data events obtained and then processing them in order to be optimal for the process mining stage. In addition to the event data as the main input, process models can be used as additional input to filter the event data. The results are records of events to be used in the stage of process mining.

This stage is composed of three activities: create views, enrich records and filter records.

• Create views: Consists of creating the event records that are views generated by the event data, created when defining cases and event types. The case notion consolidates the events so that together they form a process instance, while the event classes distinguish different activities within a process instance. Variables to be transformed and grouped by cases need to be selected (example: CaseID, Activity Name, Start Date, End Date)
• Objective records: This activity consists of enriching the event records with additional variables, that can be achieved in two ways, the first one is to deduct, calculate additional variables or generate attributes based on the variables presented in the same record of events and the second is to add external data to the event logs.
• Record filtering: This data processing activity is often used to reduce complexity or focus the analysis on a specific

part. This activity is done several times in an iteration of analysis to obtain different perspectives on the same data. It is composed of three types of filtering techniques: segmented (filtering attributes), based on variance and based on compliance.

3. Process mining techniques

This stage consists of applying process mining techniques with the event logs and obtaining information about the processes and the fulfillment of these. Exploration techniques combined with process discovery can be applied in the event logs to get an overview of the business process.

The entries for this stage are event records. In addition to this, if there are existing process models in the company, they are also considered to verify compliance and improvement activities. The results of this stage are the real discovered process models and the knowledge about what happens with the process execution.

This stage is composed of three activities: process discovery, conformance checking, improvement.

• Process Discovery: Given an event log as input, we generally start with process discovery techniques, which return a fact-based process model as output.

• Conformance checking: Given a process model, discovered or documented, which describes the expected behavior, and a record of events of actual registered behavior, compliance verification techniques aimed to the detection of inconsistencies related to the compliance of business processes between a process model and its corresponding execution record. In this phase, different aspects can be evaluated, such as quality, time, resources and cost. In addition, the results of the evaluation can be used to improve processes.

• Improvement: The improvement activity is defined as extending or improving an existing process model using information about the actual process recorded in an event log, for example, extending the process model with performance information related to time or cost or repairing the process model according to the current executions shown by the corresponding event log. The results of the improvement are process models, improved with different aspects, while the results of the verification of compliance can be considered without any process model.

4. Data visualization techniques

This stage consists in applying data visualization techniques with the event log and obtaining necessary variables to understand the business processes. The visualization techniques depend on the variables to be evaluated. The inputs for this stage are the event logs, the results are dynamic graphs that explain the process performance.

This stage is composed of three activities: data selection, evaluation of variables, selection of techniques.

• Data selection: based on the information that has been loaded, it is necessary to define what information is to be represented and studied through graphic representations; in other words, the most important aspects or characteristics must be decided. To make this selection, it is important to consider the information that is relevant and how large the magnitude of the data is.

• Evaluation of variables: The data selection operation produces an enriched data set that must directly represent the characteristics of interest for a specific exploration task. Once this information is obtained, it is associated with different data visualization techniques, considering the type of information and the number of variables that it has.

• Selection of techniques: After having associated the selected information to the techniques of data visualization, it is necessary to choose some of these, since representing a group of data, one or more alternatives are available. In this phase, its purpose is to, not only represent the information through a graphic or pictorial format, but also to choose the best techniques considering details, such as the resolution of the screen (PC, mobile phone, tablet) on which the information will be presented to end users.

5. Evaluation of results

This stage consists of evaluating the findings of the analysis with improvement ideas that helps to achieve the project goals. This evaluation considers the models of documented processes, the results of the stages of process discovery, conformance checking, improvement and the graphic representations of each of the previous stages. The results are improvement actions and plans.

This stage is composed of three activities: Diagnosis, verification, validation.

• Process diagnosis: The diagnosis of the findings obtained through process mining includes the following: (1) correctly interpreting the results (for example, understanding the discovered process model), (2) distinguishing interesting or unusual results from the expected results (for example, large number of abnormal executions) and (3) identify or refine research questions for possible additional iterations.

• Verify - validate: The accuracy of the (unexpected) findings is investigated. The verification compares the findings obtained with the documented data, while the validation compares the findings with what the stakeholders of the process expect, interviewing the resources involved in the processes. Both verification and validation help to identify the underlying causes and design ideas for possible process improvements.

D. Indicators view:

Indicators were defined, as can be seen in **Table IV**, to measure the results that will be used in the case study.

TABLE IV
INDICATORS

N°	Indicator	Description
1	% Process transparency	It measures the quantity of variations that identifies the discovered process, considering nodes and arcs, among the quantity of flow variations recorded in the system.
2	% Described process tasks	It measures the number of tasks that are identified when applying the model among the tasks that are executed in the system for the process to be studied.
3	Average evaluation time	Time of evaluation of the resulting data when applying Process Mining.

IV. CASE STUDY

A. Organization

A Peruvian network of health centers was the medical care entity where the model was validated. Its mission is to be a health institution that provides the best medical services, transforming the patient experience in Peru. Currently, this entity has five specialized health facilities. Before the implementation of the model, there were no evaluation records between the processes documented by the organization and what was implemented in the information system. Given this, the institution allowed a pilot plan to be implemented for the application and validation of this proposed model. Data ranged from January to August 2018, from two clinics located in Lima, Peru.

B. Implementation

To validate the model, we proceeded to perform tests with the information of the surgery block process. This implementation was worked done with the collaboration of the process expert and with the support of the plant doctor.

1) Extraction.

This phase of the model was developed by the IT area of the company due to the data confidentiality policies that the network of clinics needs to comply. IT extracted the information from the event data, for this they reconciled the information from different master tables and included the time intervals for each activity, considering the base structure of an event log, finally provided a file containing such data information. The output format chosen for this phase, was an .xlsx file for each clinic.

2) Event processing

On this stage, the information provided by the process specialist was taken into account, which needed to be refined because some records were not complete due to the operational error generated by the assistants and the nurses, these traces were eliminated and the event log was elaborated for each clinic, mainly considering 3 attributes, which are "encuentro" (trace identifier), activity (name of the activity) and start activity time stamp (date and time of the activity). Finally, the records of cancelled surgeries were excluded.

3) Process mining techniques

For the application of process mining techniques, the previously mentioned Celonis tool was used. In the first stage (Process discovery), we proceeded to elaborate and load the event logs in the tool, and then use the Fuzzy Miner algorithm to discover the surgery process that was executed in the system. **Fig. II** and **Fig. III** show the process discovered in each clinic and its respective variants, which were reviewed by the process specialist, who decided which nodes and arcs should be considered. This review is detailed in the results evaluation stage.
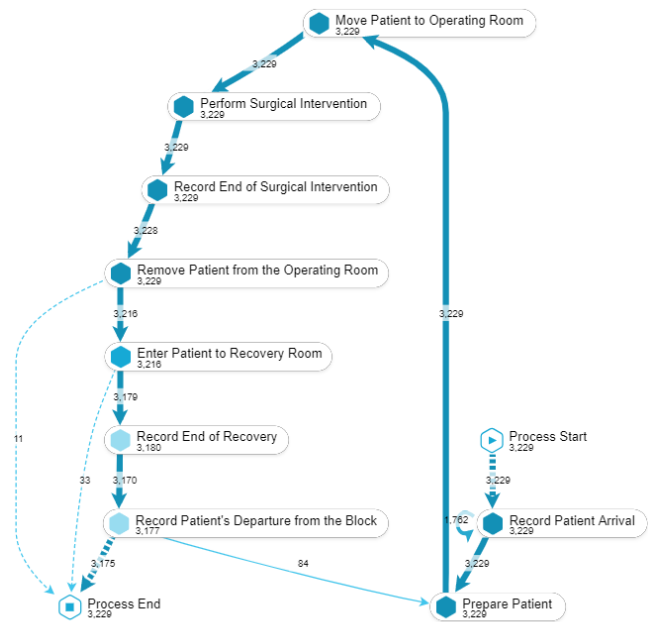

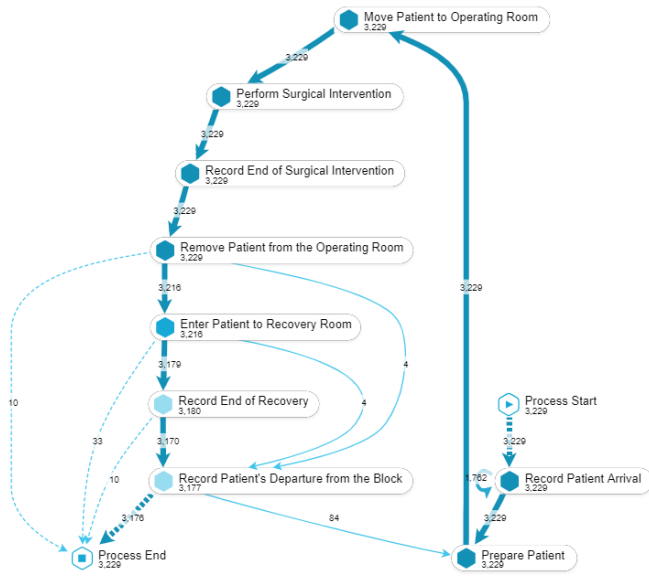Fig. II Discovered Process of the Clinic 1

Fig. III Discovered Process of the Clinic 2

Subsequently, the second stage of process mining, conformance checking, was executed, where, the documented process was loaded and compared with the discovered process, considering all the traces in the event log, in order to determine the percentage of affinity between the processes. In **Fig. IV** the activities, described in the documented process, are shown in a linear and sequential manner, which can also be seen in **Fig. II** and **Fig. III**.
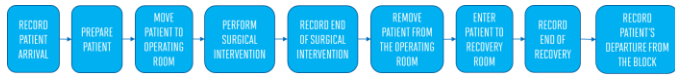

Fig. IV Activities of the Documented Process of Surgery Block

4) Data visualization techniques

On the first stage of process mining, the first interaction with data visualization techniques was started, this interaction was executed after having applied the fuzzy miner algorithm. In this phase, the information related to the process was taken, then the number of dimensions and measures that would be necessary to show what the company wants to know was recognized, and thus the graphic representations that could be used were identified.
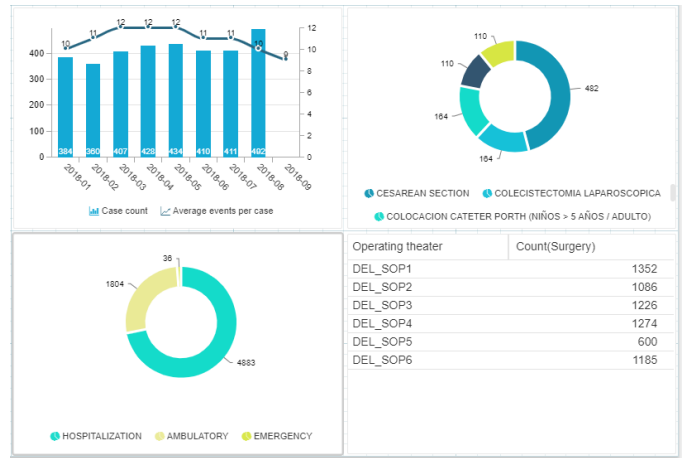

Fig. V Graphs Elaborated Based on the Information of the Process - Clinic 1

Fig. V shows the data visualization techniques that were used for the case study, first, to evaluate the number of surgical encounters that are performed each month, where we chose to use the bar graph due to that the dimension "event month" and the measure "number of encounters" are needed. Secondly, to analyze the trends by type of surgery, it was decided to use a donut diagram, since the dimension "type of surgery" consists of only 3 categories and each one is substantially different from the other in scale when considered as measure the amount of surgeries. In third place, to represent the most recurrent surgeries, the donuts diagram was chosen, since a very similar case is presented when analyzing the trends by type of surgery. Finally, to know the amount of surgeries performed in each operating room, an OLAP table was selected given that the number of surgeries measured has very close values.

For the second stage of process mining, conformance checking, we sought to compare the documented process with the process discovered in the first stage, and it was observed that there is a large percentage of affinity between them, as shown in **Fig. VI** and **Fig. VII**, where a data visualization technique called area graph was applied, which shows the percentage of similarity of the processes between the months of January to August.
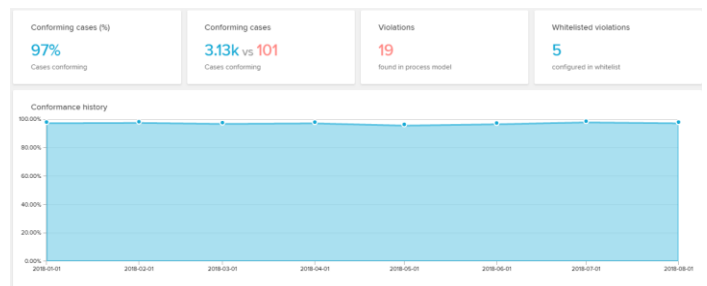

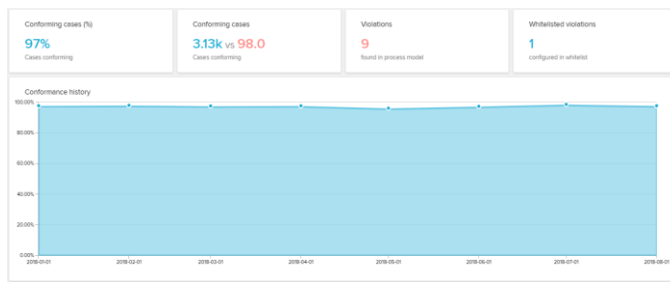Fig. VI Graph of the Conformance Checking Stage
Results Evaluation - Clinic 1

Fig. VII Graph of the Conformance Checking Stage
Results Evaluation - Clinic 2

5) Evaluation of results

On the discovery phase, the process of the actual surgery executed in the system was obtained, and when evaluating the variables mentioned in the event processing stage, the following indicators were identified: average process time, average time between each activity, the most commonly requested surgeries, tendency by surgery type, number of surgeries per operating room and operational capacity. The value of the first indicator, average time of the process, was 26 hours for clinic 1 and 27 hours for clinic 2, of which, in the two health centers, approximately 14 hours were among the activities: record patient arrival and prepare patient. Consequently, these results led the specialist to verify the tasks that were performed between these activities. Similarly, in regard to the 5 most recurrent surgeries, it was identified that the most requested surgery in the clinics was cesarean, which represented 7.17% of the total surgeries recorded.

In the conformance check phase, it was found that approximately 97% of the traces of the event record were performed according to the documented procedures (Fig. VI and Fig. VII). However, 3% of the traces presented different flows, so these cases were analyzed deeply. Afterwards, the specialist of the health institution, a Systems Engineer with a specialization in BPM, indicated that one of the different flows, which represented 2.60% of the total traces, should be considered and the documented process was adapted. Likewise, it was decided to implement a greater control system, to regulate the cases with flows other than those considered.

As a result, the company improved its documented process, after recognizing in the discovered process that in some cases after the activity "record patient's departure from the block" followed the activity "prepare the patient". This sequence was not registered in the documented process; however, it is possible since in some cases the patient receives several surgeries on different days and although he leaves the surgical unit, he does not leave the clinic. On the other hand, it has been recognized the existence of a bottleneck between the activities "record patient arrival" and "patient preparation", where the process specialist found that this was due to the execution of too many manual tasks, so it began to develop new functionalities in the system, to automate these tasks and thus be able to optimize the process cycle time.

V. CONCLUSIONS

In this research, a reference model was proposed for the evaluation of variables in operational processes of the healthcare sector using process mining and data visualization. This model was tested in the surgery process of a network of clinics, where 6,458 surgical encounters where analyzed, in which it was possible to demonstrate that the joint interaction of process mining and data visualization allows to identify patterns, trends, deviations, rework, inefficiencies and bottlenecks...

Likewise, the application of the model allowed to reduce man hours invested in the elaboration of graphs and in the analysis of business processes with the purpose of validating if what was executed in the operation is aligned with what was established by the institution.

Finally, the use of the model allowed to, not only, understand the process performance, but also the relation with key variables through the use of graphic representations with data visualization concepts in order to understand more effectively the variables that influence the business process results.

As a future study it is proposed to find, merge and clean the records of events in an automated way, since there were difficulties in extracting and cleaning the data coming from different information systems.

REFERENCES

[1] EsSalud, « Seguro Social de Salud, » 04 February 2015. [Online]. Available: http://www.essalud.gob.pe/downloads/estadistica/pres_princ_result_enc_nac_soc_ec_acces_salud_aseg_essalud.pdf

[2] EsSalud, « Seguro Social de Salud, » 05 February 2015. [Online]. Available: http://www.essalud.gob.pe/transparencia/pdf/consultas/estad_febrero2015.pdf

[3] Rojas, E., Munoz-Gama, J., Sepúlveda, M., & Capurro, D. (2016). Process mining in healthcare: A literature review. Journal of Biomedical Informatics, 61, 224–236.

[4] Appice, A., & Malerba, D. (2016). A Co-Training Strategy for Multiple View Clustering in Process Mining. IEEE Transactions on Services Computing, 9(6), 832–845.

[5] Liu, Q., Guo, X., Fan, H., & Zhu, H. (2017). A novel data visualization approach and scheme for supporting heterogeneous data. In 2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). https://doi.org/10.1109/itnec.2017.8284978

[6] Van Der Aalst, W. (2012). Process mining. Communications of the ACM, 55(8), 76.

[7] Orellana Garcia, A., Garcia, A. O., Armenteros, O. U. L., Ramirez, Y. E. P., & Alfonso, D. P. (2016). Inductive Visual Miner Plugin Customization for the Detection of Eventualities in the Processes of a

Hospital Information System. IEEE Latin America Transactions, 14(4), 1930–1936.

[8] Celonis Process Mining, 01 August 2018. [Online]. Available: https://www.pwc.dk/da/arrangementer/2018/web-finance-digital%20day-celonis.pdf

[9] Celonis Open Guide, June 2018. [Online]. Available: https://help.sap.com/http.svc/rc/6163e388c6ce431a893758775e8dca41/4.3/en-US/SAP-Process-Mining-by-Celonis-4.3_Operation-Guide-1.6.pdf

[10] prom67, «The Process Mining Toolkit, » 01 July 2018. [Online]. Available: http://www.promtools.org/doku.php?id=prom67

[11] Fluxicon Disco, «Process Mining and Automated Process Discovery Software for Professionals, » 2018. [Online]. Available: https://fluxicon.com/disco/

[12] Kebede, Musie., Comparative Evaluation of Process Mining Tools 2015. Master's Thesis.

[13] Process Mining Software - Process Mining: Data Science in Action (van der Aalst)

[14] Ailenei, Irina-Maria., Process mining tools a comparative analysis 2015. Eindhoven University of Technology - Department of Mathematics and Computer Science

[15] van der Aalst, W., Weijters, T., & Maruster, L. (2004). Workflow mining: discovering process models from event logs. IEEE Transactions On Knowledge And Data Engineering, 16(9), 1128-1142. doi: 10.1109/tkde.2004.47

[16] van der Aalst, W., Weijters, T., & Maruster, L. (2003). Rediscovering workflow models from event-based data using little thumb. Integrated Computer-Aided Engineering, 10(2), 151-162.

[17] Chamorro Ahumada, M., & Maturana Valderrama, S. (2014). Método para Aplicar Minería de Procesos a la Distribución de Bebestibles No Alcohólicos. JCC, 1(1), 1-9.

[18] Orellana Garcia, A., Perez Ramirez, Y., & Armenteros Larrea, O. (2015). Process Mining in Healthcare: Analysis and Modeling of Processes in the Emergency Area. IEEE Latin America Transactions, 13(5), 1612-1618. doi: 10.1109/tla.2015.7112022

[19] Erdogan, Tugba, and Ayca Tarhan. 2016. "Process Mining for Healthcare Process Analytics." In 2016 Joint Conference of the International Workshop on Software Measurement and the International Conference on Software Process and Product Measurement (IWSM-MENSURA), 125–30. IEEE.

[20] Buijs, J., van Dongen, B., & van der Aalst, W. (2012). On the Role of Fitness, Precision, Generalization and Simplicity in Process Discovery. On The Move To Meaningful Internet Systems: OTM 2012, 305-322. doi: 10.1007/978-3-642-33606-5_19

[21] Bolt, Alfredo, Massimiliano de Leoni, and Wil M. P. van der Aalst. 2016. "Scientific Workflows for Process Mining: Building Blocks, Scenarios, and Implementation." International Journal on Software Tools for Technology Transfer 18 (6). Springer Berlin Heidelberg: 607–28.

[22] van der Aalst, W., de Medeiros, A., & Weijters, A. (2005). Genetic Process Mining. Applications And Theory Of Petri Nets 2005, 48-69. doi: 10.1007/11494744_5

[23] van der Werf, J., van Dongen, B., Hurkens, C., & Serebrenik, A. (2008). Process Discovery Using Integer Linear Programming. Applications And Theory Of Petri Nets, 368-387. doi: 10.1007/978-3-540-68746-7_24

[24] Leemans, S., Fahland, D., & van der Aalst, W. (2013). Discovering Block-Structured Process Models from Event Logs - A Constructive Approach. Application and Theory Of Petri Nets And Concurrency, 311-329. doi: 10.1007/978-3-642-38697-8_17

[25] Van Der Heijden. (2012). Process Mining Project Methodology: Developing a General Approach to Apply Process Mining in Practice

[26] M. L. van Eck, X. Lu, , S. J. J. Leemans, and W. M. P. van der Aalst. , (2015). PM 2: A Process Mining Project Methodology, Advanced Information Systems Engineering, Lecture Notes in Computer Science, vol 9097, pp. 297–313.

[27] Telea, A. C. (2014). Data Visualization: Principles and Practice, Second Edition. CRC Press.

[28] Chittaro, L. (2006). Visualizing Information on Mobile Devices. Computer, 39(3), 40–45.

[29] Günther, C., & van der Aalst, W. (2007). Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics. Lecture Notes In Computer Science, 328-343. doi: 10.1007/978-3-540-75183-0_24

[30] Rodrigo Garcia Oliva, Jesus Santos Barrenechea (2017) Control Metrics Evaluation Model for Business Processes using Process Mining.