UNIVERSITÀ DI PISA

DIPARTIMENTO DI INGEGNERIA DELL'ENERGIA DEI SISTEMI DEL TERRITORIO E DELLE COSTRUZIONI

RELAZIONE PER IL CONSEGUIMENTO DELLA LAUREA MAGISTRALE IN INGEGNERIA GESTIONALE

Comparative analysis of clinical processes through process mining: a COVID-19 case study of Aachen's Uniklinik

SINTESI

RELATORE

Prof. Ing. Davide Aloini,
Dipartimento di Ingegneria dell'Energia dei Sistemi del Territorio e delle Costruzioni

IL CANDIDATO

Mattia Antoniazzi
m.antoniazzi@studenti.unipi.it

SUPERVISORI

Prof. Dr. Wil van der Aalst,
RWTH Aachen University, Chair of Process and Data Science (PADS)

Dott.ssa Ing. Elisabetta Benevento,
Dipartimento di Ingegneria dell'Energia dei Sistemi del Territorio e delle Costruzioni

M. Sc. Marco Pegoraro,
RWTH Aachen University, Chair of Process and Data Science (PADS)

Sessione di Laurea Magistrale del 23/11/2022

Comparative analysis of clinical processes through process mining: a COVID-19 case study of Aachen's Uniklinik

Mattia Antoniazzi

Sommario

L'obiettivo principale di questo lavoro di tesi è stato quello di sviluppare una metodologia che potesse supportare la conduzione di un'analisi comparativa in ambito sanitario utilizzando strumenti di process mining. I dati analizzati nella tesi fanno riferimento al trattamento dei pazienti COVID-19 presso l'ospedale Uniklinik di Aachen. Il lavoro consiste in tre analisi principali. Gli algoritmi di process discovery mirano a ottenere modelli di processo per ogni ondata della pandemia, che possono poi essere confrontati tra loro per evidenziare le differenze nel flusso di lavoro. Attraverso tecniche di conformance checking è possibile confrontare i dati di ciascuna ondata con il modello prescrittivo che rappresenta le linee guida cliniche per il trattamento del COVID-19, al fine di identificare le deviazioni tra il comportamento previsto e quello effettivo del processo. Per analizzare le prestazioni dell'ospedale è stata creata un set di indicatori di prestazione. I principali cambiamenti sono stati riscontrati nella gestione dell'Unità di Terapia Intensiva, nell'uso di tecniche di supporto respiratorio e nella somministrazione di farmaci.

Abstract

The main goal of this thesis work was to develop a methodology that could support the conduct of a comparative analysis in the healthcare domain using process mining tools. The data analysed in the thesis refer to the treatment of COVID-19 patients at the Uniklinik hospital in Aachen. The work consists of three main analyses. Process discovery algorithms aim to obtain process models for each wave of the pandemic, which can then be compared with each other in order to highlight differences in the workflow. Through conformance checking techniques it is possible to compare the data of each wave with the normative model representing the clinical guidelines for COVID-19 treatment, in order to identify deviations between the expected behaviour and the actual behaviour of the process. A set of Key Performance Indicators is created to analyse the performances of the hospital. Major changes were found to be in the management of the Intensive Care Unit, the use of respiratory support techniques, and drug administration.

1.Objectives & Context

The objectives of this thesis are several. The first objective is to develop a methodology that is able to support the conduct of a comparative analysis in the healthcare domain. This framework is intended to obtain valuable information and insights into how the treatment of COVID-19 patients at the Aachen's Uniklinik has evolved over the three waves of the pandemic in terms of workflow, treatment executions, drug administration and clinical pathway. A second objective of this work is to show both strengths and limitations of process mining when applied in real-world contexts and in environments marked by high variability such as healthcare. One last goal of this work is to show how process mining can improve management and be a concrete tool in the hands of key managerial figures. This thesis is associated with the European Union-commissioned project ICU4COVID, designed and developed in order to hinder the development of further COVID-19 outbreaks or similar infectious diseases. The main goal is to provide EU citizens and healthcare workers with intensive medicine suitable for combating COVID-19 in an effective manner, leveraging the Cyber-Physical-system for Tele- and Intensive Care Medicine (CPS4TIC) platform. Data regarding COVID patients were provided in the context of the COVID-19 Aachen Study (COVAS).

2. Background

The new industrial paradigm, known as Industry 4.0 (*I4.0*), was created with the aim of making companies even closer to the needs of their customers, which have become increasingly changing and varied. Companies have realised that their processes must be increasingly efficient and driven by data. The spectacular growth of the digital universe makes it theoretically possible to record, derive, and analyse events. An event may take place inside a machine, inside an enterprise information system, inside a hospital, inside a social network. In this context, data-driven techniques such as process mining has become more and more relevant. The stream of events that occur within a business process can be sequentially recorded to obtain an event log. Process mining aims to extract knowledge from event logs readily available in today's systems in order to discover, monitor, and improve real processes.

There are three types of process mining:

1. *Discovery:* these techniques take an event log and produce a process model without using any a-priori information. When modelling we tend to concentrate on the "normal" and "desirable" behaviour.
2. *Conformance:* these techniques compare an event log and a process model (the log is *replayed* on the top of the process model) to verify if the model is aligned to the real process, and vice versa. The model can be handmade or automatically discovered.
3. *Enhancement:* these techniques aim at changing or extending the a-priori model using the information about the actual process recorded in the log. By doing this it is possible *to repair* the model, so that this reflects better the reality, or *extend* the model adding performance data (bottlenecks, throughput times, frequencies) derived from the timestamps.

Process mining can be applied to any type of operational domain and is particularly useful for the analysis of healthcare processes. One of the biggest challenges in the healthcare domain in these last years is reducing costs while optimizing processes, providing better and better care. Execution data are a valuable information source to support the management and improvement of healthcare processes. Healthcare organisations make intensive use of Health Information Systems (HISs). During the execution of a clinical process, several entries in the HIS are recorded. These entries can be used to generate an event log describing the sequence of activities that were performed. However, the application of process mining in the healthcare domain is challenging for several reasons. Healthcare processes often present high variability, which results mainly from to the high numerosity and heterogeneity of the activities involved. Moreover, patients may respond very differently to treatments, affecting the sequence of the following activities. Despite the presence of Clinical Practice Guidelines (CPGs), which are recommendations on how diagnose and treat a certain medical condition, unknown co-morbities, unexpected complications, patient's preference or resource shortage may cause unplanned situations difficult to predict and create unexpected patterns in the data. Conversely, poor data quality is also a critical issue within the healthcare domain. Data quality problems such as missing events, imprecise timestamps, and incorrect timestamps can be related to fact that data and events are to a large extent still manually entered by physicians into the systems. This results in events that are often not registered or registered in the wrong order. Recently a new sub-branch of process mining is emerging. Comparative process mining aims to compare multiple logs from different perspectives by performing analysis and applying techniques of different nature. For example, these logs can be generated by the same process performed in different time intervals or the same process performed in two or more organizations of the same nature. Comparative process mining has several applications resulting in different outcomes:

- *Cross-conformance checking between logs*. With cross-conformance checking it is possible to verify how consistent the logs are with each other and capture the same process.
- *Conformance checking with prescriptive model.* Multiples logs are replied upon the same prescriptive model of the process.
- *Visual comparison between the model.* Comparing visually two or more process models has other benefits. It is possible to detect activities that are executed in one model and not in the others or that have different positions.

The goal is to understand how the process has changed in terms of activities executed, possible path, best practices, resources involved, why these changes occurred and the possible causes. Valuable information can be obtained on how a certain clinical process is changing over time in terms of process efficiency and quality of care.

3. Comparative Methodology

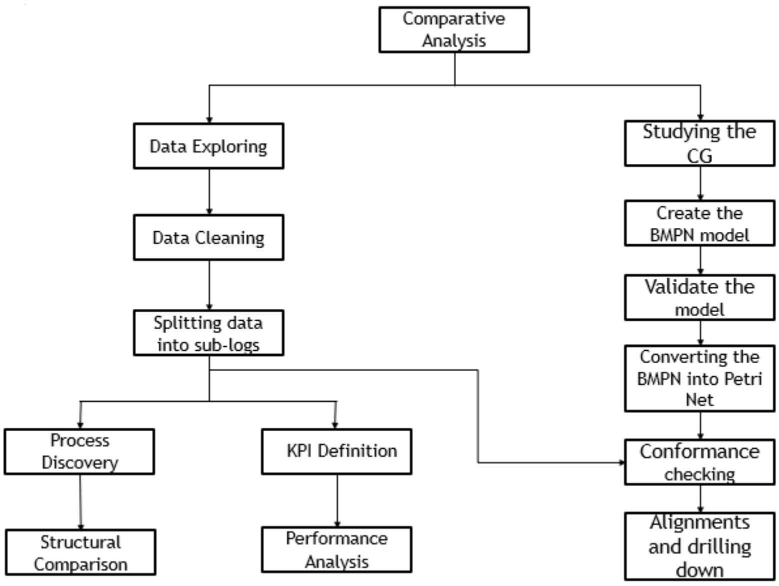

Figure 1 Overview of the comparative methodology.

In Fig. 1, the comparative methodology that has been developed is shown. The methodology consists of two main parts. The one on the left has as its main input the event log, which is first filtered by outliers and then splitted into sub-logs. Process discovery algorithms are then applied in order to obtain process models that can be compared in the Structural

Comparison phase, and a set of KPIs are defined in order to analyse hospital performances. The section on the right has as its main actor the normative handmade model representing the guidelines. This is constructed by hand, then validated using the opinion of a physician, and finally converted to Petri net. At this point, conformance checking techniques are applied using the three sub logs obtained previously. Alignments for each wave are calculated and their deviations are investigated.

4. Case study: Aachen's Uniklinik

The data provided by the COVAS (Covid Aachen Study). This Project contain records regarding patients affected by SARS-CoV-2 and refer to the entire course of the disease from the onset of symptoms through the beginning of the hospitalization, until the patient's discharge.

4.1 Data Exploration, Filtering and Splitting

The COVAS event log originally presented 3542 events, 269 cases, 210 variants and 33 activities. 192 variants appear only once in the log and very few have a higher frequency, showing the high variability of the process. Several steps of filtering have been carried out. The first one concerned the *hospitalization time* – the time from the beginning of the hospital stay until the patient's discharge – in order to detect possible patients with an extremely long hospitalization time and eliminate them from the analysis. The second one consisted of filtering out on less frequent start activities, end activities and paths. The resulting log presents 2397 events, 187 cases, 135 variants and 32 activities. Based on the scientific literature and through a visual analysis of hospitalization events, the event log has been divided into three sub-logs representing the three waves of the pandemic. The event log of the first wave contains 106 cases and 1410 events, the log of the second wave contains 59 cases and 892 events, and the log of the third wave only 22 cases and 892 events.

4.2 Process Discovery

The first step of this phase is the choice of the discovery algorithm. The choice fell on the Inductive Miner and on its variant, the Infrequent Inductive Miner. The Inductive Miner lends itself well to real-life logs, and more importantly, produces models that are surely sound. The inductive frequent miner (*IMf*) gives the ability to filter out the noise and infrequent behaviour present in the log. The second phase is the choice of the discovered model. A methodology is proposed which aims to choose the best model among those obtained for multiple threshold levels of the Inductive Miner Infrequent. The model has then been validated with widely-adopted quality metrics, ensuring a balanced trade-off between adherence to historical data, generalisability to unseen process instances, and readability of the model.

4.3 Steps for creating the prescriptive model

The main source from which the information for creating the prescriptive model was extracted was The STAKOB guidelines, which contain recommendations not only for the treatment but also the diagnosis of COVID-19 disease. These were supplemented with other sources such as scientific articles or papers in order to fill in the missing information. The creation of the BPMN model was supported through the Signavio tool. The whole process is composed of 3 sub-processes, 23 activities and approximately 36 gateways (AND XOR and OR). The model is composed of a multitude of varied pathways reflecting the generality of the guidelines, and goes from the onset of the patient's symptoms until the patient's discharge. The three sub-processes refer to the outpatient care, Intensive Care Unit (ICU) and respiratory support. The validation of the BPMN model was made using a qualitative approach. Three online meetings with a physician from the Uniklinik's ICU were held in which the model has been reviewed and discussed, until the final version is obtained. The labels were unified, i.e., the task names in the BPMN were changed to match those in the log, and the model converted into a Petri net using the ProM software.

4.4 Conformance Checking

Conformance checking is applied between the prescriptive model of the clinical guidelines and each sub-log, with the goal to acquire information about how much and in what way data are aligned to the model. The discrepancies between model and event log can be seen as a chance to improve the process and quality of care. It is possible to understand which activities are being skipped and how many times, activities that are executed in a different order to that in the model, activities that are in the log and not in the model and vice versa. The technique used was Alignments which is unaffected by the model notation used. The ProM plug-in "*Replay a Log on Petri Net for Conformance Analysis*" was used to calculate alignments restricted only to control-flow perspective.

4.5 KPIs Definition

This phase deals with the definition of a set of KPIs to analyse the hospital's performances. The indicators are being grouped by the sub-process which they belong, except for activities concerning drug administration which have been grouped all together for a matter of simplicity. In Fig. 2, Fig.3 and Fig.4 some examples of indicators are shown.

KPIs	First wave	Second wave	Third wave
ICU			
OUTPUT			
ICU mortality rate	37.0 %	62.0%	50.0%
Number of patients admitted to ICU (% on hospitalized)	54 (50.9)	24 (40.67)	6 (27.27)

Figure 2 This table contains KPIs regarding the ICU sub-process.

KPIs	First wave	Second wave	Third wave
RESPIRATORY SUPPORT			
OUTPUT			
Number of patients treated with HiFlo (% on ICUs)	1 (1.85)	2 (8.3)	3 (50)
Number of patients treated with NIV (% on ICUs)	13 (24.07)	14 (58.33)	6 (100)
Number of patients treated with ventilation (%on ICUs)	43 (79.62)	19 (79.16)	3 (50)
Ventilation mortality rate	47.0 %	74.0 %	67.0 %
Number of patients treated with ECMO (% on ICUs)	10 (18.51)	8 (33.3)	-
Number of patients treated with Pronation (% on ICUs)	35 (64.81)	24 (66.66)	2 (33.33)
Ventilated patients requiring ECMO	23.0%	42.0%	-
Patients treated with ECMO and then died	50.0%	88.0%	-

Figure 3 This table contains KPIs regarding the respiratory support sub-process.

KPIs	First wave	Second wave	Third wave
DRUGS ADMINISTRATION			
OUTPUT			
Number of patients treated with antibiotics (% on hospitalized patients)	60 (56.6)	29 (49.15)	8 (36.36)
Number of patients treated with Remdesivir (% on hospitalized patients)	3 (2.83)	21 (35.59)	-
Number of patients treated with Dexamethasone (% on hospitalized patients)	-	40 (67.79)	17 (77.27)

Figure 4 This table contains KPIs regarding the drug administration.

5. Results

The models obtained in the process discovery phase are poorly interpretable, and it is difficult to conduct a satisfactory structural comparison analysis. This is mainly related to the fact that timestamps are recorded at the daily level, and within a single day multiple activities may be performed within a hospital. Process mining algorithms interpret these activities as concurrent and fail to produce proper sequencing of them. Expect for the first wave, where the workflow is slightly more sequential and constrained, the other models present mainly concurrent activities allowing a high number of possible pathways. This limits the understanding of the actual sequencing of activities and leads to executions that are obviously wrong. Regarding the *"Alignments and drilling down"* analysis, for each wave activities performed even if not prescribed by the clinical guidelines have been detected. For example, the admission to the ICU may have been executed more than required because of critical course not foreseen by physician. Some activities present the opposite situation, e.g., they are prescribed by the clinical guidelines but not actually performed. The high-flow oxygenation activities (*HiFlo start* and *HiFlo end*) have this kind of mismatches. Indeed, especially at the start of the pandemic, high flow oxygenation was thought to increase the likelihood of contracting the virus which is why it was not performed frequently. Lastly, other activities present both situations. Activities like *Discharge ICU* (*the* discharge from ICU) and *EndOfFever* (the end of the patient's fever) in some cases are present in the model but not in reality, in others vice-versa. The exact point in which such activities should be executed or must be executed differs widely from patient to patient, outlining the high variability of the process. From the analysis of KPIs several insights emerged. There were fewer patients admitted to the ICU overtime, despite high mortality rates in the latter. This may suggest better management of this ward. At the beginning of the pandemic all procedures that were thought to have the potential to produce viral particles, such as high-flow oxygenation and non-invasive ventilation, were avoided. Subsequently these were performed to try to reduce the high mortality rates of invasive ventilation and facilitate extubation of patients. ECMO (*Extra Corporal Membrane Oxygenation*) was not performed during the third wave because of the lower severity of the virus and the high costs associated with this practice. Patients received increasingly rapid hospitalization, probably because of the improved ability to recognize the symptoms of the disease. The use of antibiotics has decreased over time. This is probably to avoid AMR (*AntiMicrobial Resistance*) and prevent the virus from strengthening. Dexamethasone (immunomodulatory drug) gradually supplanted Remdesivir (antiviral drug) as it was less expensive and better suited to improve the condition of patients in the later part of the pandemic who often presented severe pneumonias and more lungs damages.

6. Final results & Managerial implications

The major differences that were found were in the management of the ICU, the use of respiratory support techniques, and drug administration. Regarding the ICU, its use gradually decreased over the course of the pandemic, even though mortality levels of the ICU recorded considerable levels. Utilization of the ICU has refined over time, also allowing better use of the resources within it, such as nurses, doctors, and respiratory support equipment. Invasive respiratory support techniques have been progressively joined by less invasive ones, which were not initially performed in contrast to what the guidelines suggested. This also had a considerable positive impact on resource allocation and costs incurred in the care of COVID patients in ICU. In addition, performing treatments suggested by guidelines which had never been performed in reality can strengthen care pathways and reduce hospitalization time. Drug administration has changed, allowing cost savings. The comparative methodology represents a supporting tool that can enable the application of process mining techniques in order to compare clinical processes. This framework aims to guide physicians through a step-by-step approach in carrying out comparative analysis in healthcare domain. The comparative methodology, as structured, is able to give precise and data-driven information about various aspects of the hospital process. It is possible to analyse bottlenecks, service times, waiting times, deviations from the prescriptive model, and other process-centric information. Treatments that are performed with a wrong timing or that are often skipped can be identified and their execution corrected. The framework is changeable in both content and structure to any type of hospital and non-hospital context and can even be expanded and enriched as needed.

7.Conclusions

This work aimed at developing a methodology that can support the conduct of a comparative analysis in the healthcare domain using process mining tools. This work presents the following limitations. The data present timestamps recorded at the day level that make the recording sequence of events unreliable, and it is an obstacle for process discovery algorithms. Another limitation is the one related to the numerosity of observations. The low numerosity of the third wave greatly affects a variety of aspects, especially the values of some indicators. Regarding possible future developments of the thesis, certainly the numerosity of the dataset needs to be increased in order to have a balanced situation between the first, second and third waves. Another aspect that can be strengthened are the indicators. Indicators could be added to cover the parts left uncovered, e.g., resource-related indicators and cost indicators. The handmade model can be enriched and updated to be consistent with the new guidelines. An important goal would be to obtain data with timestamps with a greater level of detail in order to produce more accurate and interpretable process models. Lastly, it would be possible to gather data from other hospitals of the same geographical area and compare their processes using the framework proposed, also in a non-COVID healthcare context.