KAKOPANDE YAKETE Lyna – M2 MCI EUR
My level 1: Maëlle Bozo Berrios

Process mining and data mining applications in the domain of chronic diseases:
A systematic review

Source : Google Scholar [Process mining and data mining applications in the domain of chronic diseases: A systematic review - ScienceDirect](#)

> **KEYWORDS :**
> **Patterns, pathway, risk factor, decision trees, "random forest", discovery, performance, enhancement, software, application, statistical analysis**

This research work is a study carried out with the help of scientific articles selected on the basis of criteria at the crossroads of healthcare and process mining.

This document explores the advances made by data and process mining in the field of healthcare, and more specifically in the detection of chronic illnesses and the streamlining of patient care. Indeed, a third of adults are affected by and live with at least one chronic disease. Data analysis through process mining and its methods, combined with artificial intelligence, can prove extremely useful. We can summarize this paper in three main areas: the objectives of process mining in the field of healthcare and chronic diseases, the methods used, and the challenges and limitations of using process mining.

I. Objectives of process mining in healthcare

In the field of patient care, detection or anticipation of (chronic) illnesses, a massive amount of data is collected. This data contains **vital information** which, if properly exploited, has the potential to shed light on patients' health trajectories (risk factors, adaptation and management of care, anticipation of disease aggravation, etc.), and to influence and adapt medical decisions. The aim of process mining is to **take advantage of the vast amount of data collected** through technology, in order to improve patients' quality of life. The document explains that the application of process mining combined with artificial intelligence plays a predictive analysis role.

The document also establishes **a slight difference between data and process mining** in this field. Data mining **analyzes patterns** (disease trajectory and give predictions). Process mining **emphasizes clinical data and pathway to optimize the care pathway and subsequently relieve the patient**. For example in the case of cancer treatment, which is very complex due to the multiplicity of substantial challenges (trying to get an early detection, the treatment, or the disease management). Process mining is a good method to optimize the care of patient suffering from chronic disease, because it identifies people who are the most prone to chronic disease with the collect and the treatment of their data.

II. Process mining methods in healthcare

The main process mining methods associated with artificial intelligence, used in the healthcare field, are :

- "Random Forrest", a process mining technique based on machine learning algorithms. It is an output of **multiple decision trees**, with different data for one given theme (for example genome , protein archetypes, etc.), to reach **a single result**.

- Process mining method is through **3** steps : Process **discovery**, **conformance** and **enhancement**. **Process discovery** is the **most use process** mining method to figure out hidden/underlying process in healthcare. The process discovery step is performed using PALLA, Inductive Minter-

Infrequent (Imf) and Careflow Miner (CFM) techniques, and it should be noted that the use of these process mining techniques **is generally always accompanied by specific applications/software**. For example, the "PALLA" process mining technique is used with "PMApp" software, the Inductive Miner-Infrequent is used with "ProM" and the "Careflow Miner" technique is used with "pMineR" software.

- When Process Mining is used in conjunction with traditional statistical analysis methods, the following methods are employed: "The Fisher Exact Test" and "The Cox regression".

III.   The challenges of process mining in healthcare

In terms of healthcare optimization, it is difficult to rely **solely on** process mining in data analysis, as **this approach is very recent** in this field. For the time being, traditional statistics are preferred. However, the combination of both methods - process mining with AI and traditional statistics - is a major step forward.

There are also many process mining methods, and so there is also variability in results if the emphasis is not placed on a single way of proceeding. We need to **find a guideline**, a **reference process mining method**, so that comparisons can be made around the same methods, so that results are consistent, and so that optimization is effective.