

SEPSIS COHORT CLUSTERING

Sepsis is a serious syndrome, characterized by organ dysfunction triggered by infection. It can be life-threatening, or it can affect permanently survivors' functional and psychological abilities. This complex syndrome involves all defense mechanisms which, when deregulated, can have deleterious effects which amplify those of the pathogen. Multiple genetic, microbiological, biological and environmental factors influence the patient's response to infection. This complexity can only be grasped by a transverse, translational and rigorous approach when developing treatment suitable for each type of patient.

Aside from early and appropriate antibiotic therapy, no other treatment of sepsis is currently available, despite often promising experimental work. One of the main obstacles in the development of new treatments is the difficulty in obtaining homogeneous groups of patients on which clinical trials can be reliably run. Indeed, the treatments tested in some experimental work are too general to act on defense mechanisms, which vary from one patient to another depending on the level of infection, co-morbidities or the presence of genetic susceptibility factors.

This capstone project will aim at clustering sepsis-affected patients in the MIMIC freely available dataset into coherent subpopulations that could be better suited for the testing of tailored treatments. Starting from the available clinical and demographic data this study will aim at learning in an unsupervised way a new representation of sepsis affected-patients that would allow their separation into homogeneous groups. This represent a first step towards rethinking experimental models and pathophysiological paradigms in order to develop a research program whose ultimate goal is a personalized therapeutic approach, based both on the clinical-biological phenotype and the patient's genotype.

To achieve the main objective, this study will be subdivided into:

- Identification of patients with Sepsis
- Carry out an exploratory analysis to understand the disease and relevant feature
- Explore different clustering techniques to identify the solution