The implementation of clinical pathways is used to ensure a well-coordinated care offer. It allows for a standardization of practices.

The objective of the study is to reveal the potential of the French national hospital PMSI database to describe clinical pathways. It will bring new knowledge about care management and it will find underlying correlations between patients’ profiles and their pathways. The illustrative case study is the 8-year follow-up of patients implanted with a cardiac defibrillator.

**METHOD FOR DATA MANAGEMENT**

**DATA EXTRACTION**

**EXCLUSION:**

**BACKGROUND:**

We extracted all 2008 hospital stays with the implantation of a CRT-D (defibrillator allowing resynchronization). The source databases are the PMSI (French Medical Information System) and PICHOMP (medical devices register). Using the anonymous ID of the 1 602 obtained patients, we then extracted their stays for 2 years after (2006-2008) and 5 years afterwards (2008-2013), being a total of 18 304 stays.

**MORE DETAILS ABOUT DATA EXTRACTION**

The extraction of the hospital stays from 2006 to 2013 was initially including all kinds of stays, being 23 930. Then, to label each stay with the main medical diagnosis, 549 different diagnosis were used. A filter was applied to remove “noisy” stays and to keep only the most relevant: a 79% more diagnoses represent 80% of stays (being 18 304). This filter reduces data complexity and allows to focus on the main pathways.

**SCIENTIFIC BACKGROUND**

**PROCESS MINING FOR PATHWAYS ANALYSIS**

Clinical pathway model was discovered from data using a Process Mining approach and an Integer Linear Programming model. It searches for similarities in the sequence of stays of all the patients, each stay being labeled by the medical diagnosis.

**SPECIFICATIONS OF THE MATHEMATICAL MODEL**

Process Mining is a scientific field which is focused on studying end-to-end processes through databases. Time is key feature of the approach. Clinical pathways can be analyzed over several years, and not only locally.

We developed our own optimization algorithm to detect similarities in the patients’ sequences. It is based on both frequencies of stays and of transitions. The mathematical model is an Integer Linear Programming. To solve this combinatorial problem, we used IBM CPLEX solver.

The final result is a synthetic graphical view of the most representative pathways of the cohort. Users can choose the level of detail to print more or less different paths.

**CONCLUSION**

This study shows that clinical pathways can be analyzed by mining the French PMSI database.

The pathway models are made of three components: the key hospital stays; the delay of transition between two stays and the probability of following each of the paths. Such analysis can be performed for any sub-group of patients (age, gender, comorbidities).

The resulting model provide insights in the care management disease and help identifying improvement drivers.

Finally, the strength of this methodology and of the underlying algorithm is to be generic. It can be readily applied on other diseases.

![Diagram of clinical pathways](image-url)

**FIG. 1 - INTEGER LINEAR PROGRAMMING MODEL**

**DECISION VARIABLES:**

Assign hospital stays to nodes in the model: $S(x) = 1$ if stay $x$; 0 else.

Assign edges between nodes: $C_{\text{ex}} = 1$ if transition between $x$ and $y$; 0 else.

**CONSTRAINTS:**

Size: Number of nodes and edges is bounded: $|E| + |N| \leq \text{THRESHOLD}$.

Evidency: Assign hospital stays to nodes at most once.

Consistency: Model must be compact (no isolated node allowed).

**OBJECTIVE FUNCTION:**

Maximize (Model's quality)

**VARIABLES DESCRIPTION:**

Resulting model = $(E, N)$

$N$: set of assigned stays in nodes

$E$: set of assigned edges in model

$C_{\text{ex}}$: all existing nodes in the log

$l_{\text{max}}$: size of the log

$S_{\text{in}}$: all existing edges in the log

$S_{\text{end}}$: significance of $S_{\text{in}}$ (relative importance)

$\alpha, \beta, \gamma, \theta$: weighted factors in [0-1]

**FIG. 2 - RAW DATA USED AS INPUT OF THE ILP MODEL**

**FIG. 3 - SYNTHETIC PATHWAY MODEL OF IMPLANTED PATIENTS**

**HEART ISSUES**

2581 - Left ventricular failure

148 - Atrial fibrillation

2390 - Unstable angina

1212-2] - Cardiomyopathy

1372 - Ventricular tachycardia

2351 - Atherosclerotic heart disease

**COMPLICATIONS, RELAPSES & OTHERS**

2450 - Adjustment of cardiac devices

2514 - Preparatory care

R570 - Cardiogenic shock

Z098 - Follow-up examination

**HISTORY PRIOR TO IMPLANTATION (2006-2008)**

- 51% of patients were hospitalized for heart failure (code 9511), on average 5 months before implantation.

- 73% of patients had at least 1 stay in the last 6 months before implantation.

- 11% of patients did not have any history over the 2 previous years.

**FOLLOW-UP AFTER IMPLANTATION (2006-2013)**

- 8% had complications related to their defibrillator (on average 2 years and 1 month after). They were readmitted to adjust their device (code Z450) on average 2 years and 2 months after.

- 28% died at the hospital within 5 years (average: 2 years and 4 months after).

- 29% died at the hospital within 6 years (average: 2 years and 4 months after).

**NEXT STEP: DRIVERS TO IMPROVE THE CARE**

- a. Anticipate failure by frequent follow-up.

- b. Early implantation of the device in the absence of heart failure risk.

- c. Search for deeper explanations of the early replacements of some devices (< 3 years).

**8-YEAR FOLLOW-UP (FROM 2006 TO 2013)**

- ~ 4-6 years

- Lifetime of a defibrillator

- Stays on average per patient

- Different types of stays