Prediction of short-term mortality after the implementation of an Implantable Cardioverter Defibrillator using autoencoders neural networks on medical event logs

Martin Prodel (PhD), HEVA, Lyon, France
Julien Beisel (Msc), HEVA, Lyon, France
Hugo De Oliveira (PhD), HEVA, Lyon, France
Vincent Augusto (PhD), Mines Saint-Etienne, CNRS, UMR 6158 LIMOS, Centre CIS, Saint-Etienne France
Ludovic Lamarsalle (Pharm.D, Msc), HEVA, Lyon, France

December 2021 - International Congress on Healthcare and Medical Informatics

1

Problem Statement

Statement of the problem

Database :

SNDS: Exhaustive database of the French population covered by the National Health Insurance

- Approximately 68 million people insured
- True representation of the insured French population

Completeness of Medicare Health Insurance Expenditures

- Primary Care (consultations, biological tests, drug dispensing, DMI…)
- Hospitalizations (in public and private institutions)

→ Patient's pathways are complex, and every event is not meaningful

We need to use interpretable methods to understand which events increase the occurrence probability of the outcome

Contributions

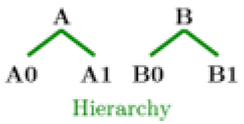
1. Modeling time in pathway sequences to make the extraction of relevant patterns possible

> We use enriched event-logs to model complex pathway sequences

timestamp	activity
0	A0
2	B1
3	A1
5	A0
7	B0
7	A1
8	B0

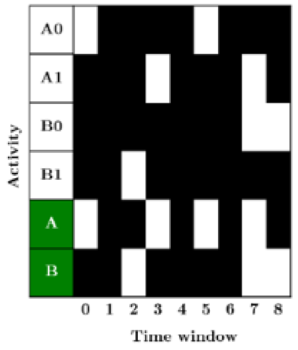
2. Dealing with the hierarchical structure of codes

> The enriched event-logs contain information about code hierarchy



3. Providing a visual explanation of what has been learned by the model

> By representing these event logs as visual matrices, we can explain the model



We present these contributions with a case study on heart failure mortality.

1

Case study

Case study

Available data

Data modeling

Algorithm Auto-encoder

Results

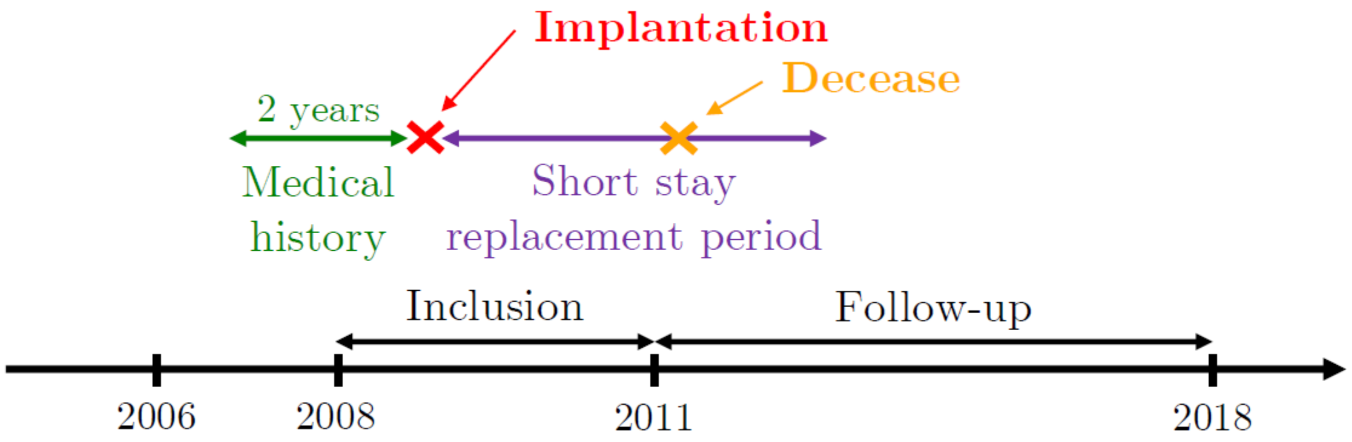
HEVA

Predicting short-term mortality after the implementation of an Implantable Cardioverter-Defibrillator

Population:
We extract patients with an ICD (Implantable Cardioverter-Defibrillator)

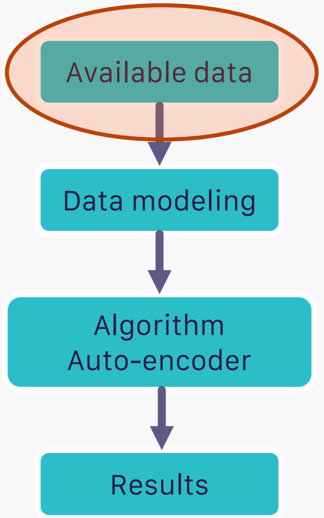
Goal:
We identify patients with a risk of post-implantation mortality within the short-term replacement period (during the period covered by the warranty).

> The goal is to identify **predictive factors in medical event logs** extracted from the SNIIRAM database, considering **time** and **hierarchy structure** in medical codes, and **without patient-centered information.**



Data available

HEVA



ICD-10 codes:

Diagnoses and comorbidities (2 levels of hierarchy)

CCAM codes:

Medical procedures (3 levels of hierarchy)

Medical devices:

(3 levels of hierarchy)

Biological tests:

(1 level of hierarchy)

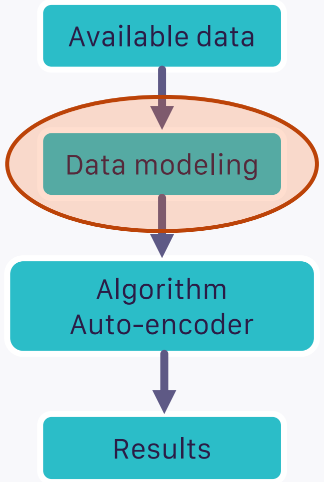
Consultations

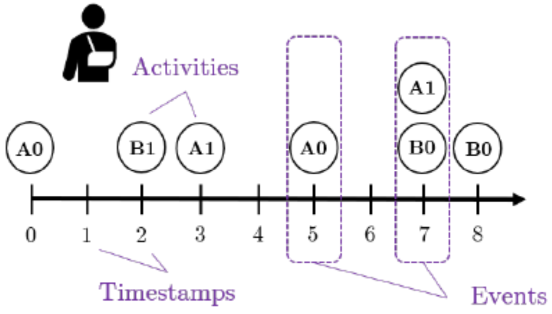
	#Patients	#Activities	#Events	#Labels
Train	14,942	9,215,335	493,863	18,758
after filtering	-	8,700,492	493,752	962
Test	3,736	2,344,217	126,423	12,282
after filtering	-	2,209,349	126,396	962

We need to extract relevant features

Data modeling

HEVA



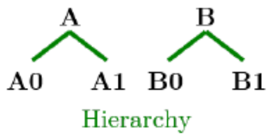


Patient pathway

1 – We extract the patient pathway from the SNDS database

ICD10, CCAM, MD, Bio, Consults

timestamp	activity
0	A0
2	B1
3	A1
5	A0
7	B0
7	A1
8	B0



Initial event log

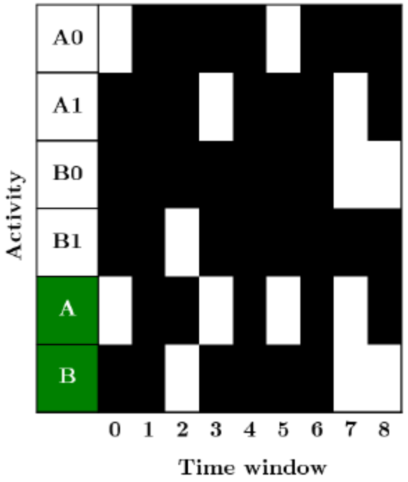
2 – We model the pathway as an event log

One row = One event

timestamp	activity
0	A0
0	A
2	B1
2	B
3	A1
3	A
5	A0
5	A
7	B0
7	B
7	A1
7	A
8	B0
8	B

Enriched event log

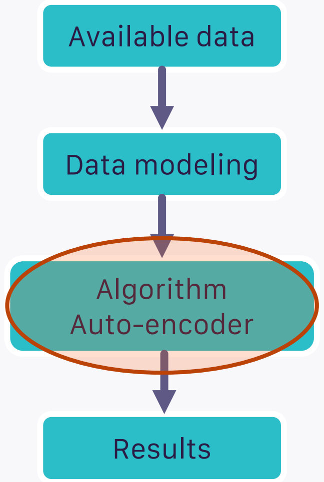
3 – We enrich the event-log with knowledge about the code hierarchy

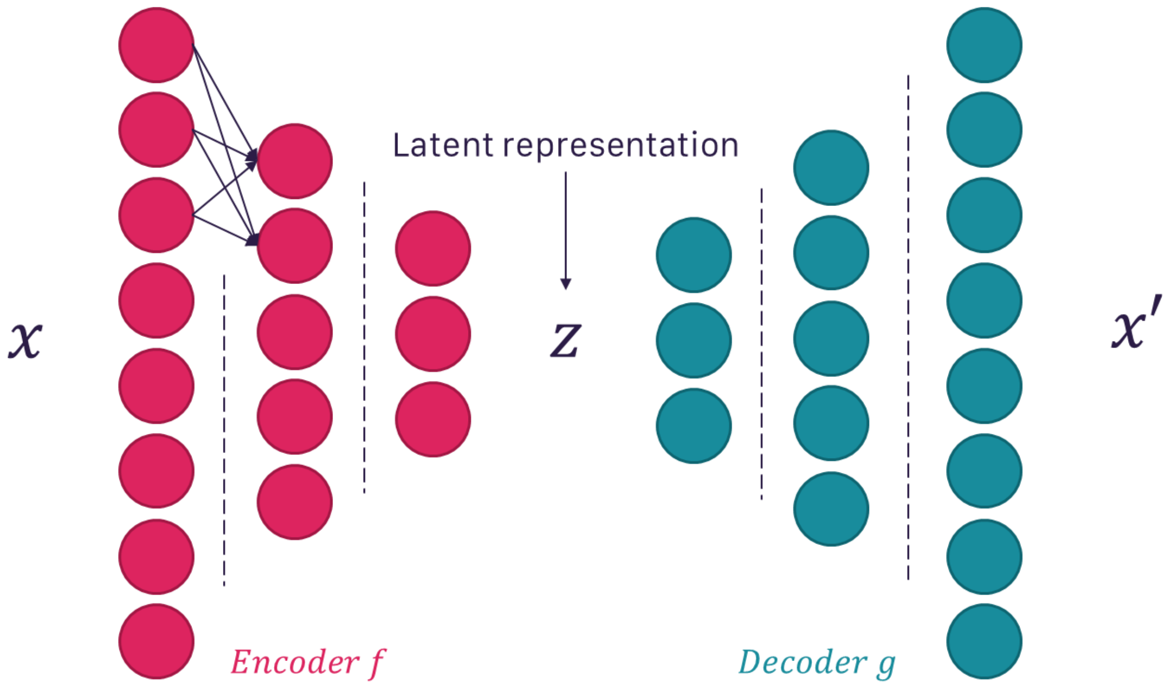


Matrix representation

4 – We represent the enriched event log as a matrix (patient as image)

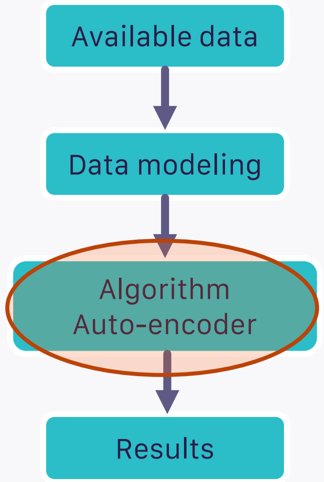
Auto-encoders

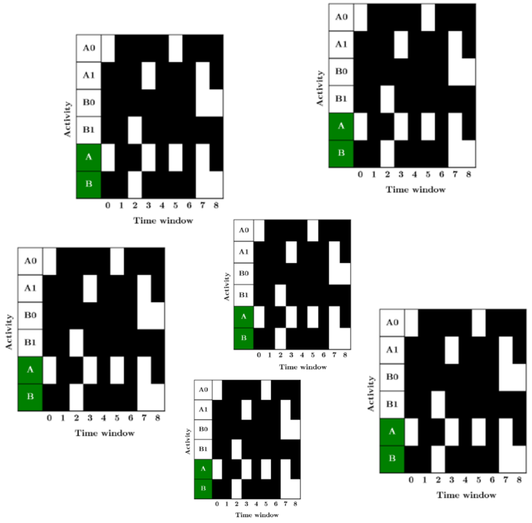


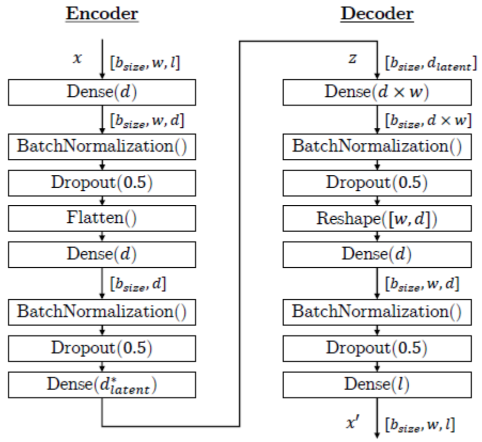


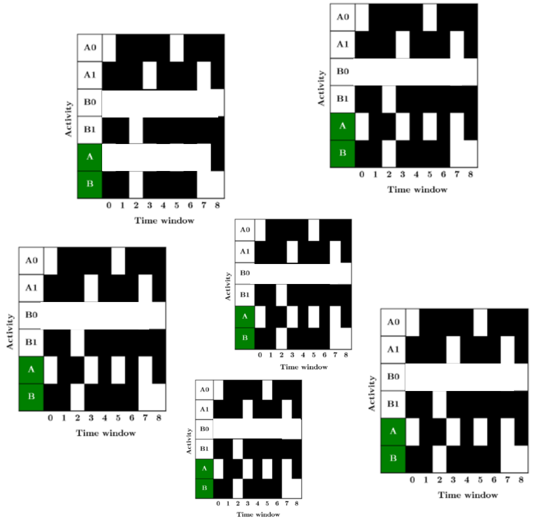
Objective: Minimize the difference between x and x'.

Use of auto-encoders







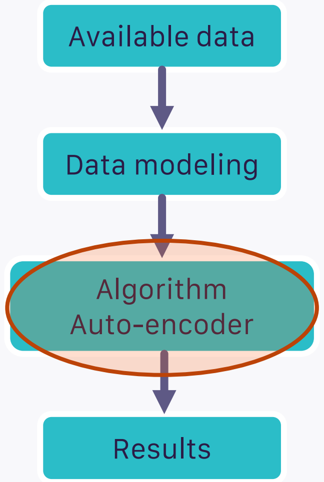


Input data :
Visual representation of enriched event logs

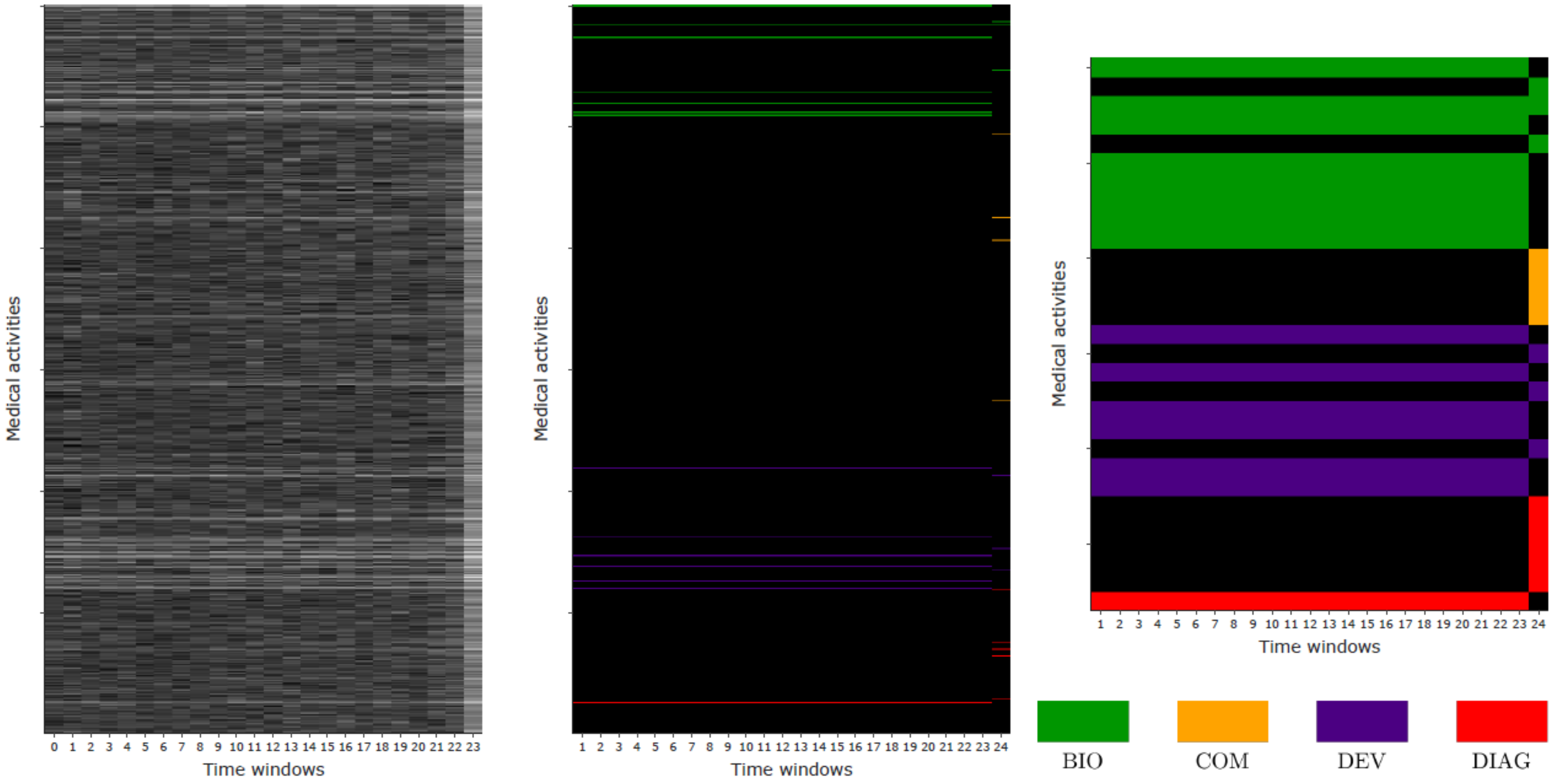
Algorithm :
Auto-encoder – We train the algorithm to only reconstruct « target » patients

Output data :
The relevant parts of the matrices are reconstructed

Results



Visual results



(a) Explanation element $\hat{x}$ (b) $\hat{x}$ after selection of top activities (c) Minimal representation of $\hat{x}$

Results

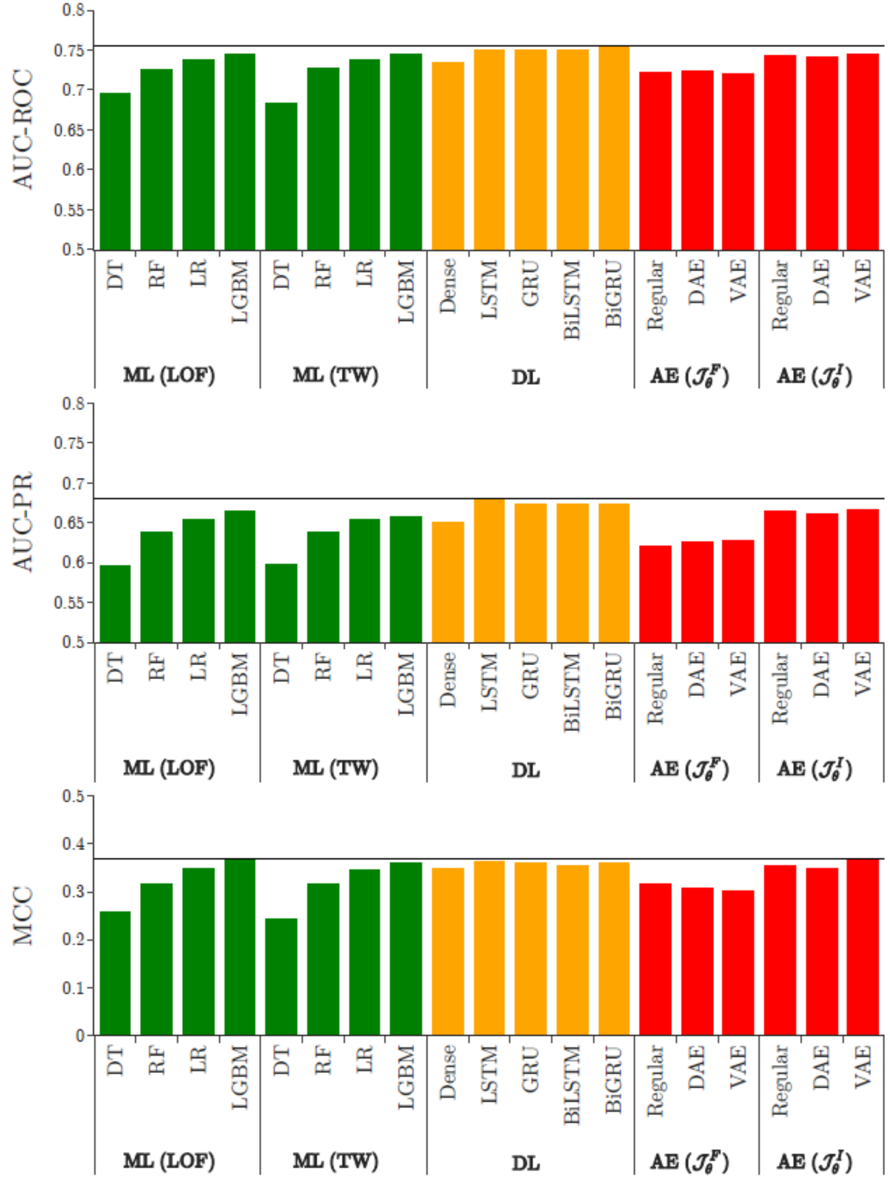
Performance

Available data

Data modeling

Algorithm
Auto-encoder

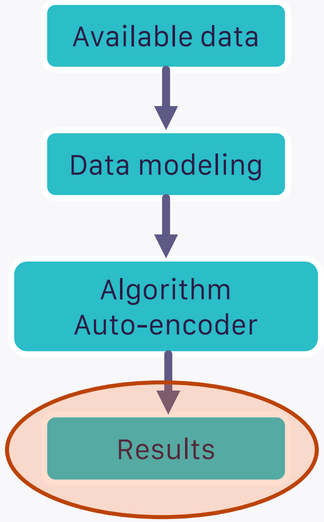
Results

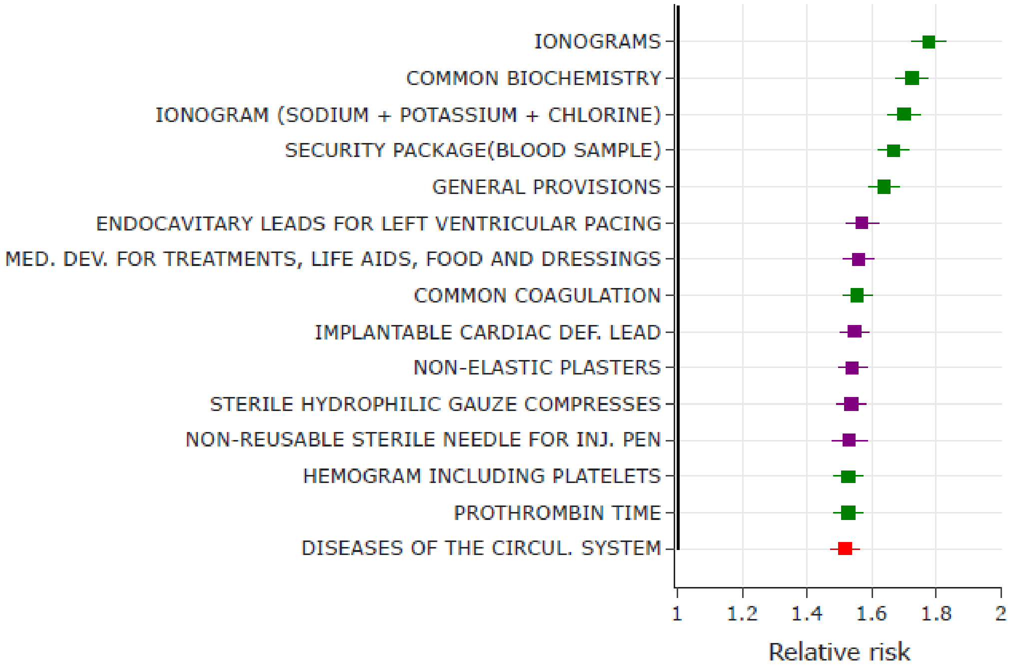


HEVA

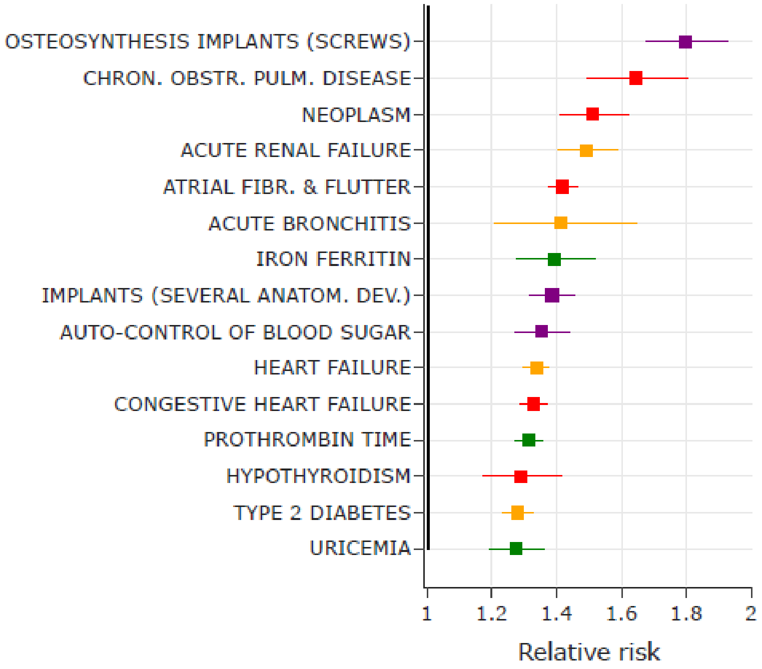
Results

Extraction of relevant events & relative risks linked to the presence of events



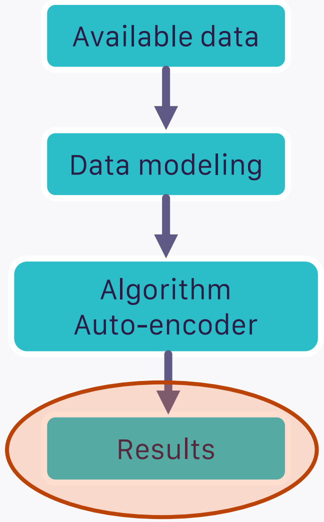


(d) Relative risks deduced from $\hat{x}$ (frequent events)



(e) Relative risks deduced from $\hat{x}$ (last time window)

Conclusion



Improvements:

- **We could improve performance:**
  - Try different AE architectures
  - Try attention layers

  Note: Recurrent layers were implemented and improved performance but did not reveal any particular predictive factors

- **Adding more data (images, vital signs, free text…) might give more interesting insights**

Conclusion: The explainability in this context could produce knowledge directly from patient pathway data and help performing early at-risk patient detection