Building and representing the patient pathway from multidimensional and large datasets

Antoine LAMER[a,b], Fabio BOUDIS[b] , Niels MARTIGNENE[a], Grégoire FICHEUR[a], and Emmanuel CHAZARD[a]

[a] Univ. Lille, CHU Lille, ULR 2694 METRICS, F-59000, Lille, France
[b] Univ. Lille, Faculté Ingénierie et Management de la Santé, F-59000, Lille, France
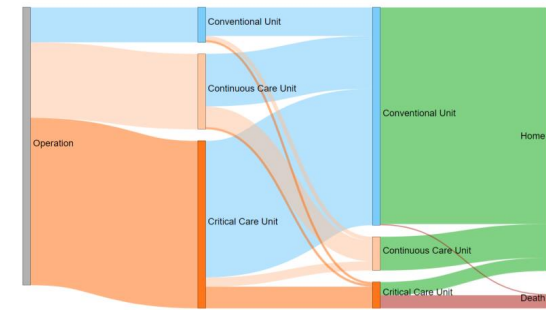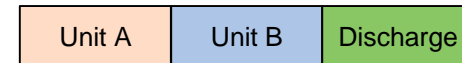
antoine.lamer@univ-lille.fr

Introduction

Clinical pathways

Clinical pathway : sequence of interventions from which the patients benefit during their encounters with health care structures.

Two contexts:

- an **individual-based context** when the visualization targets a single patient to provide appropriate care

- a **population-based context** when it relates to a whole population and addresses research or decision-making objectives

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur

Clinical pathways – Some examples

Clinical pathway : sequence of interventions from which the patients benefit during their encounters with health care structures.

Post-operation period
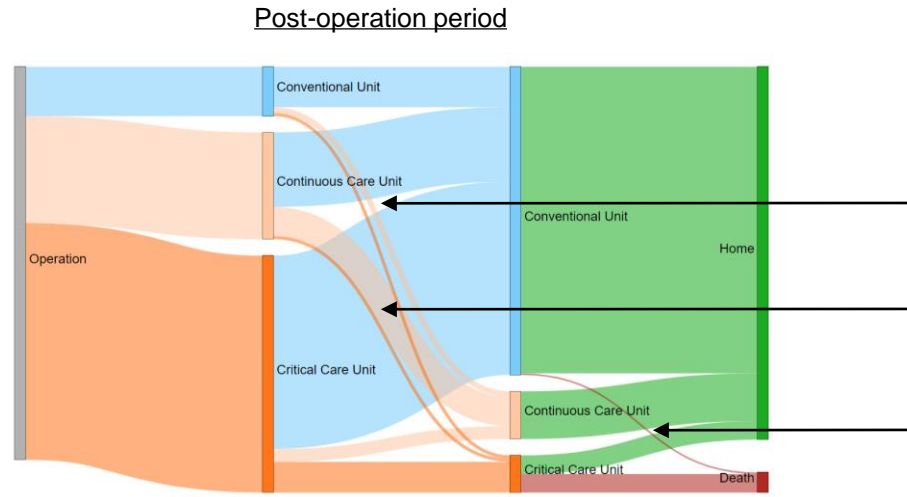
Atypical pathways :
- patients transfered too late to acute care units
- patients died in unmonitored units

Barriers

Clinical pathways - Limits

Different entities with different dimensions

patient

patient_id	age	sex	bmi
1	25	M	22.3
2	78	F	25.1
3	54	M	
4	65	M	
5	19	F	

hospital stay

stay_id	patient_id	admission	discharge
1	1	23/09/2019	30/09/2019
2	2	20/09/2019	10/10/2019
3	1	12/11/2019	20/11/2019
4	3	22/10/2019	25/10/2019
5	4	23/10/2019	27/10/2019

procedure

procedure_id	visit_id	procedure_date	procedure
1	1	24/09/2019	ADDA200
2	2	24/09/2019	DSZA300
3	3	12/11/2019	BJAD320
4	3	13/11/2019	NJMA421

unit stay

unit_id	stay_id	admission	discharge
1	1	23/09/2019	25/09/2019
2	1	25/09/2019	30/09/2019
3	2	23/09/2019	25/09/2019
4	2	23/09/2019	25/09/2019
5	2	23/09/2019	25/09/2019
6	3	12/11/2019	13/11/2019
7	3	13/15/2019	15/11/2019
8	3	15/11/2019	20/11/2019

diagnosis

drug

measurement
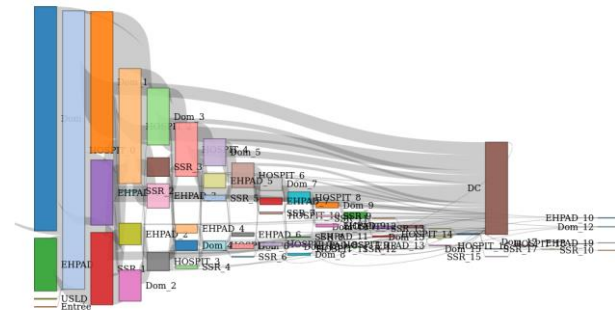
...

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur

Clinical pathways - Limits

High numbers of patients
High number of events and of modalities
Heterogeneity in duration

increase the complexity of datasets and representations

pattern explosion

| Event	Number of modalies
Diagnoses	CIM10 : 14 400 codes
Procedures	CCAM : 10675 codes
Drugs	ATC : 6,331 codes
Hospital units	> 500

Perer A, Wang F, Hu J. Mining and exploring care pathways from electronic medical records with visual analytics. J Biomed Inform. 2015 Aug;56:369-78. doi: 10.1016/j.jbi.2015.06.020. Epub 2015 Jul 2. PMID: 26146159.

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur

Clinical pathways - Limits

High numbers of patients
High number of events and of modalities
Heterogeneity in duration

increase the complexity of datasets and representations

pattern explosion

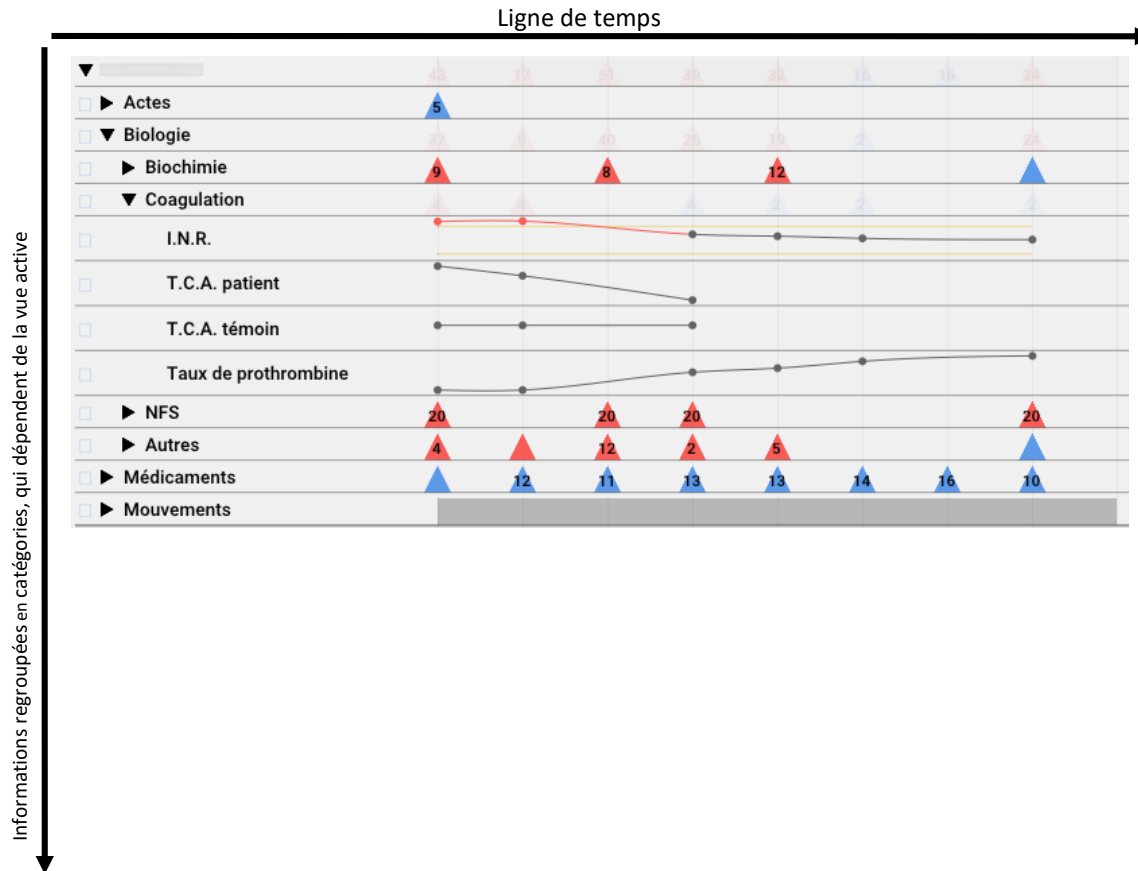Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur
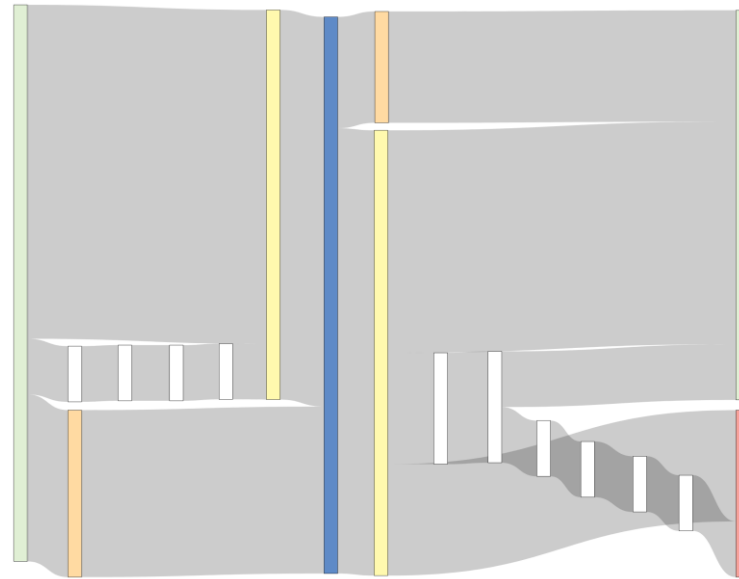
Clinical pathways - Limits

The number of events variable from one individual to another, and each patient encounters the events at its own dateseach patient encounters the events at its own dates.
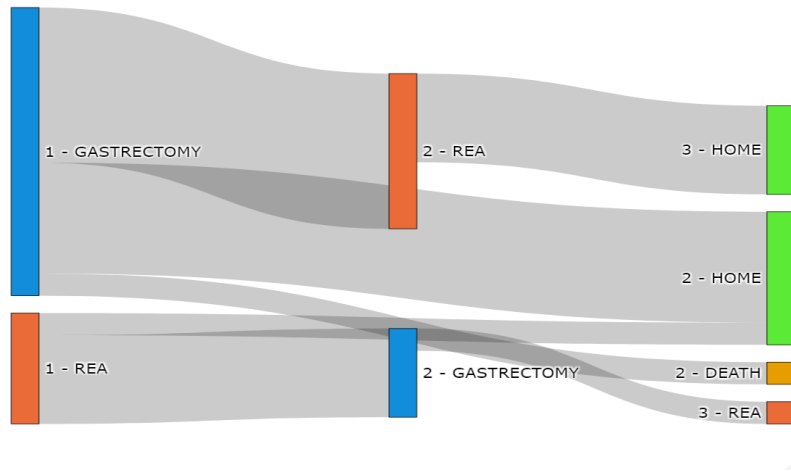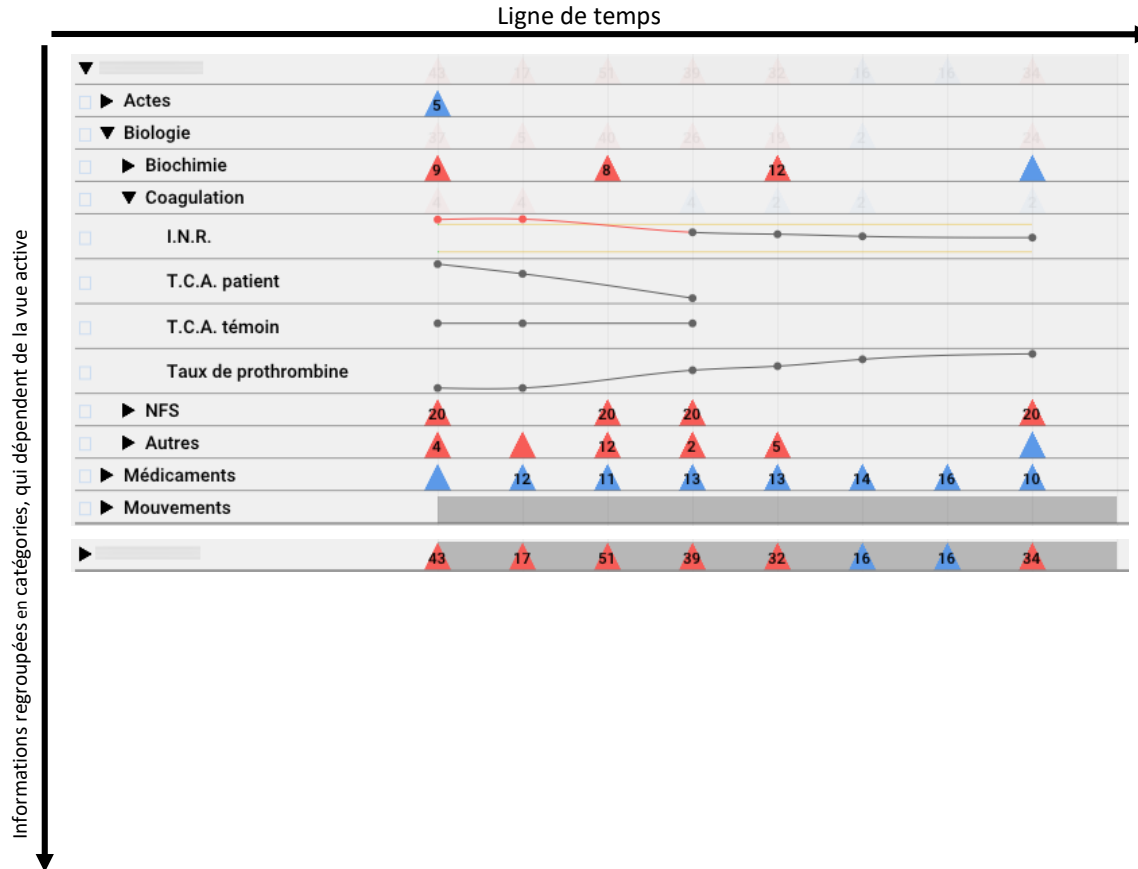

Gastrectomy, post-operative stay in intensive care unit and home stay occur at different ranks of the pathway while they represent the same steps
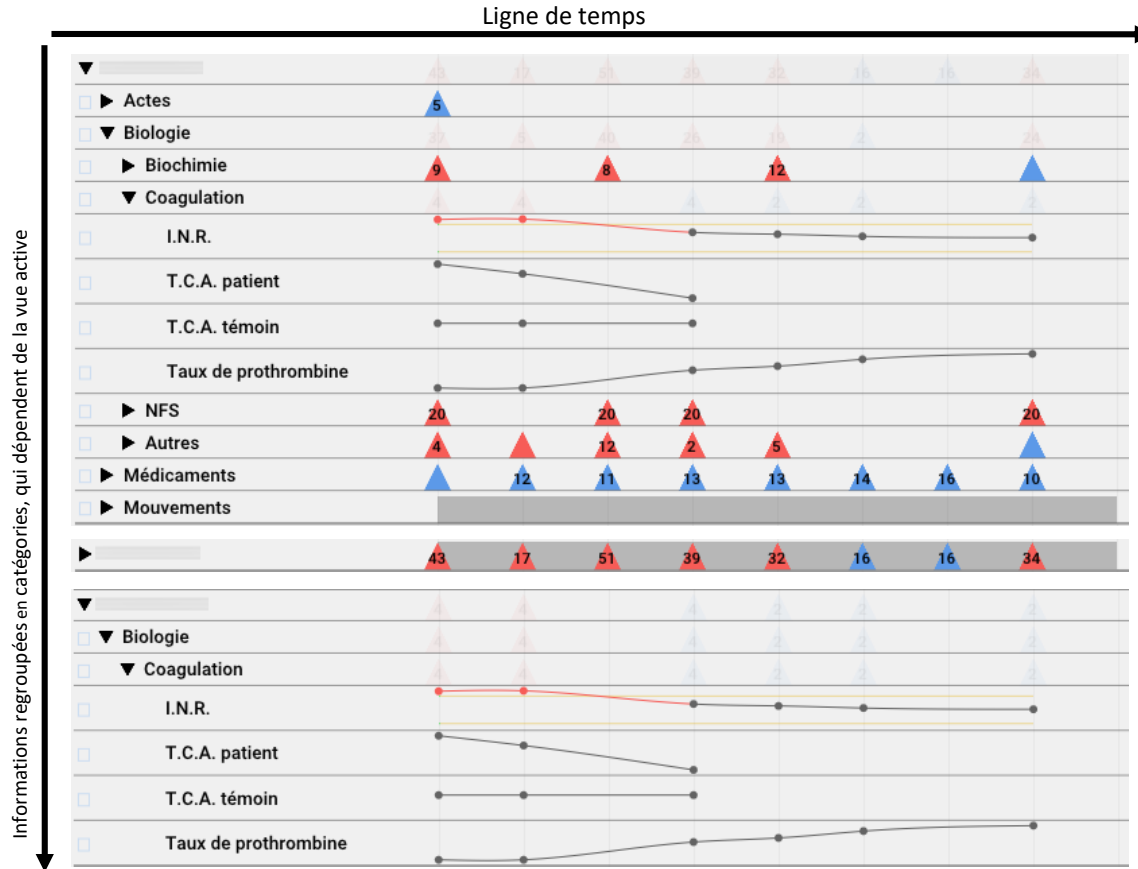
Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur

Individual-based context

Individual-based context

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur

Individual-based context

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur

Individual-based context

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur

Population-based context

Objective

Build patient pathway in preserving essential information and in a intelligible way:

- Extract and gather relevent data

- Address the complexity of pathway (heterogeneity, temporality, …)

- Evaluate the final information

- Identify clusters

- Compare clusters

Implementation

Implementation

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur
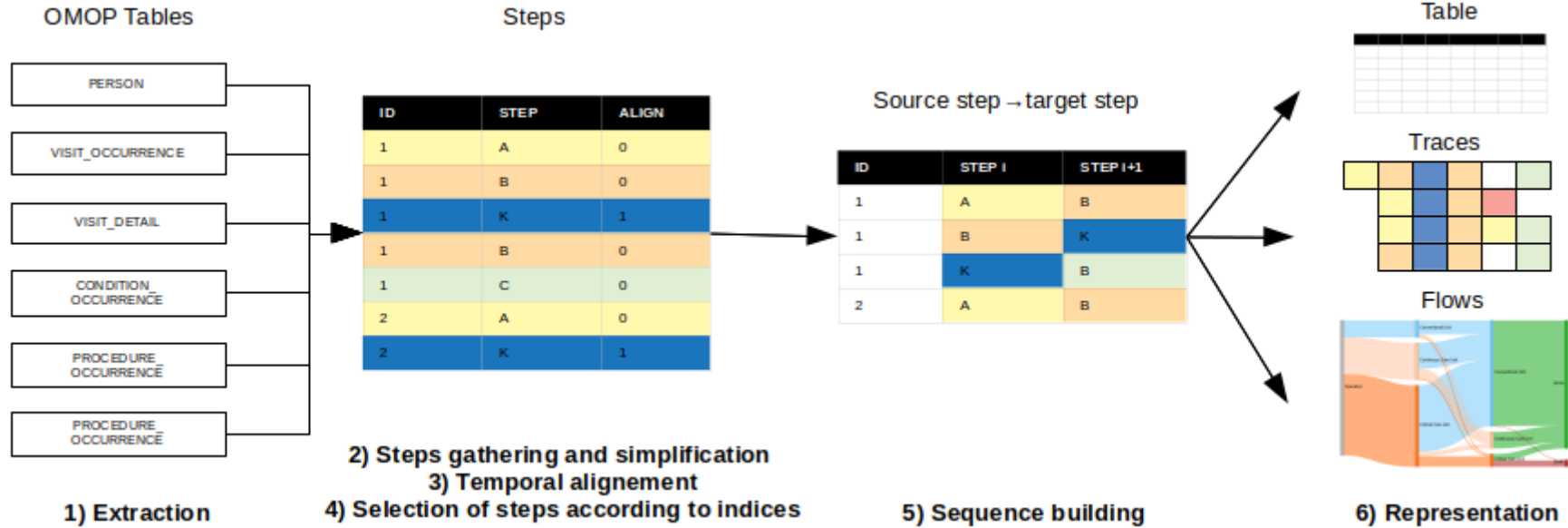
Data transformation

2) Steps gathering and simplification
3) Temporal alignement
4) Selection of steps according to indices

Three main transformations :

- Filter

- Simplification

- Time-oriented transformations

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur

Data transformation

Filter

<u>Events selection :</u>

- identification and extraction of useful information in the source database. take into account the hierarchy, and extracting event based on the hierarcal level (e.g. a class of drugs)

<u>Records filtering :</u>

- at the individual level or population level

- at the start or at the end of the process

Both events and records may be filtereg based on the presence or absence of an event (e.g. administration of drug A), but also based on the value of an attribute of the event (e.g. administration of drug A with posology over 50 µg, the posology being the attribute)

Mica L, Niggli C, Bak P, Yaeli A, McClain M, Lawrie CM, Pape HC. Development of a Visual Analytics Tool for Polytrauma Patients: Proof of Concept for a New Assessment Tool Using a Multiple Layer Sankey Diagram in a Single-Center Database. World J Surg. 2020 Mar;44(3):764-772

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur

Data transformation

Filter

Events selection : identification and extraction of useful information in the source database

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur
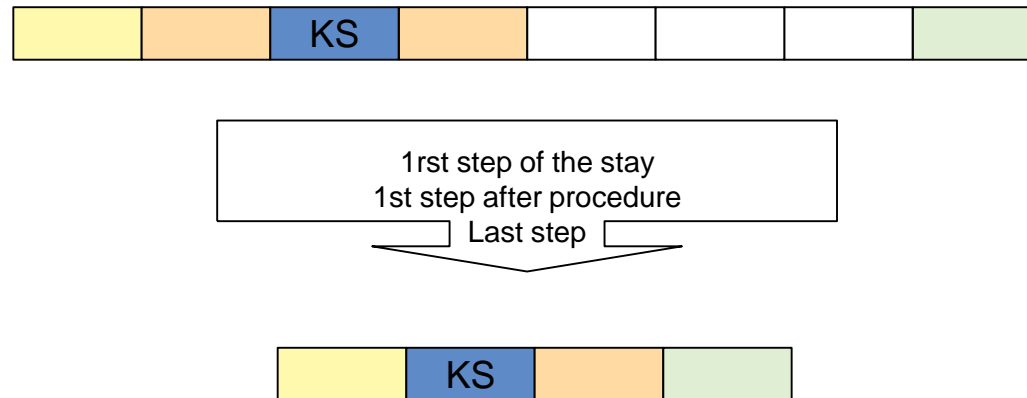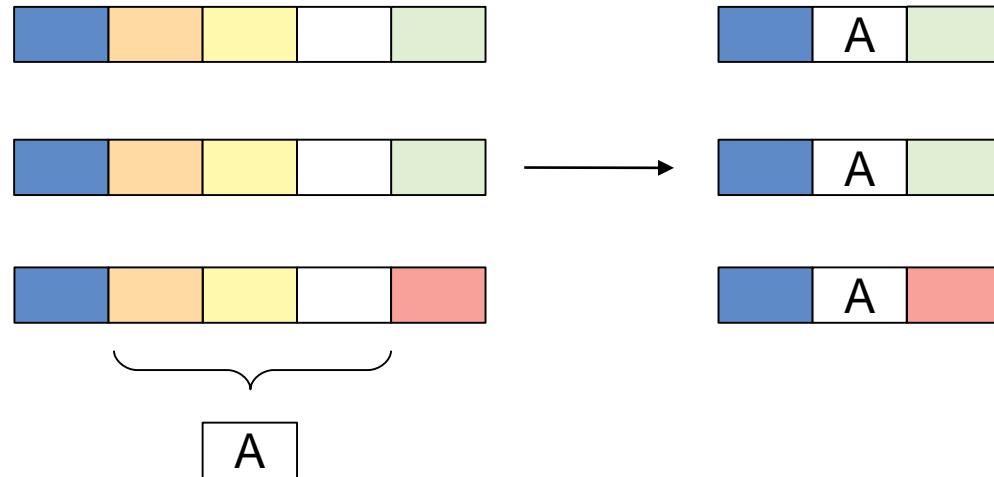
Data transformation

Simplication

Consecutive event merging : two consecutive events represent the same state (e.g. a diagnosis) and are merged into a single event to reduce information overload, and the complexity and length of the trajectories.
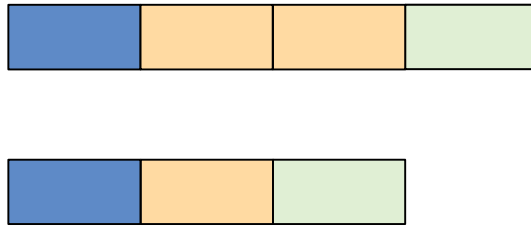
Data transformation

Simplication

Frequency consecutive merging : Merging / grouping in a super-event based on the frequency of consecutive events

Perer A, Wang F, Hu J. Mining and exploring care pathways from electronic medical records with visual analytics. J Biomed Inform. 2015 Aug;56:369-78.

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur

Data transformation

Simplication

<u>Label events merging :</u> multiple events or categories to be combined into a single meta-event or meta-category

⇨ With a business rule
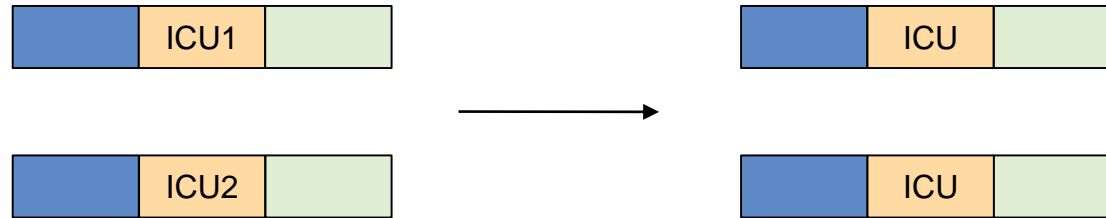
ICU : Intensive Care Unit

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur

Data transformation

Simplication

Label events merging : multiple events, categories or clusters to be combined into a single meta-event, meta-category or meta-cluster

⟹ With a similarity distance between two nodes (Here, on the outcome).
Minimum distance being defined by the user

$$d(node_A, node_B) = |node_A.outcome - node_B.outcome|$$

$$d(cluster_X, cluster_Y) = \frac{\sum_{m \in cluster_X \& n \in cluster_Y} d(m,n)}{size(cluster_X) * size(cluster_Y)}$$

Wongsuphasawat and D. Gotz, "Exploring Flow, Factors, and Outcomes of Temporal Event Sequences with the Outflow Visualization," in IEEE Transactions on Visualization and Computer Graphics, vol. 18, no. 12, pp. 2659-2668, Dec. 2012, doi: 10.1109/TVCG.2012.225.

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur

Data transformation

Time-oriented transformations

Alignment according to event of interest : the surgery day, a diagnostic, the start of a treatment, each day of the pathway the admission day
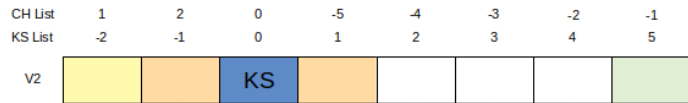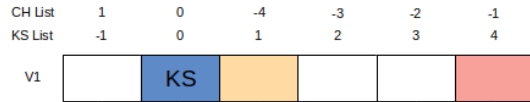

Alignment point

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur
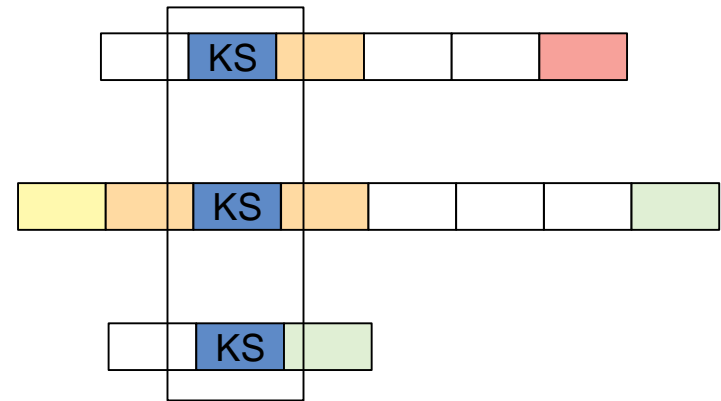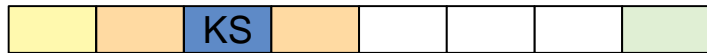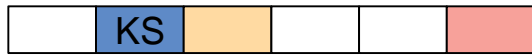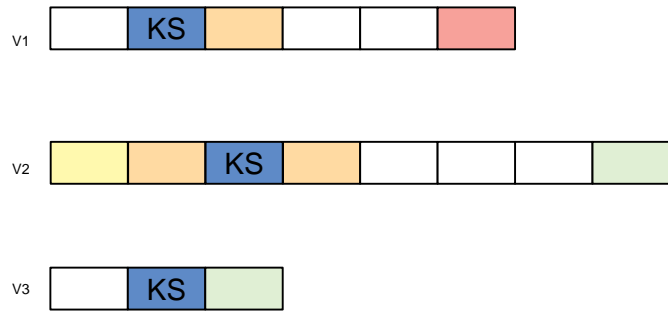
In practice

In practice

Raw data

CH List	1	0	-4	-3	-2	-1
KS List	-1	0	1	2	3	4

V1: [ ] [KS] [ ] [ ] [ ] [ ]

CH List	1	2	0	-5	-4	-3	-2	-1
KS List	-2	-1	0	1	2	3	4	5

V2: [ ] [ ] [KS] [ ] [ ] [ ] [ ] [ ]

CH List	1	0	-1
KS List	-1	0	1

V3: [ ] [KS] [ ]
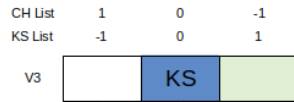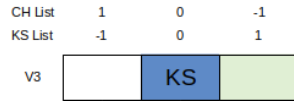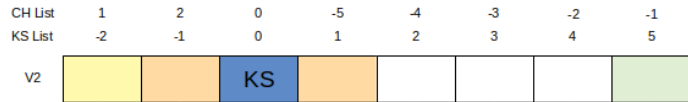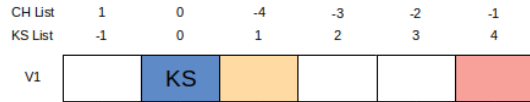
CH List : chronological order

KS List : key step order

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur

In practice

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur
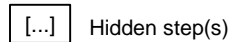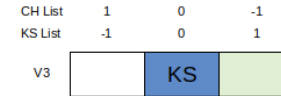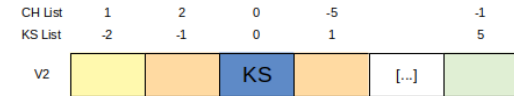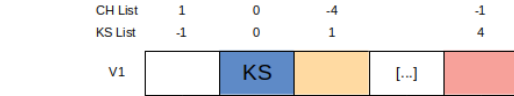
In practice

In practice

Indivicual traces
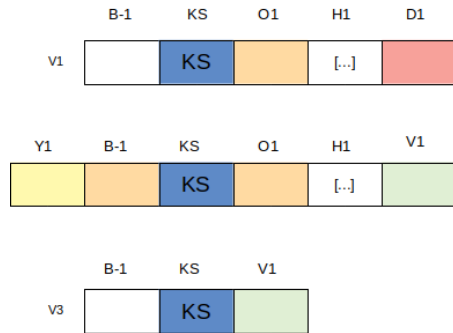
Population flows

Population aggregation

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur
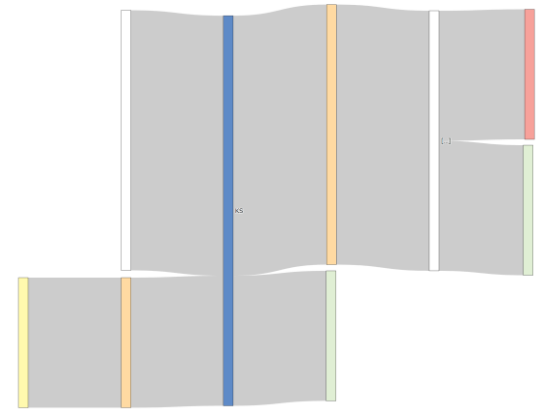
Next steps

Integrating :

- Statistical modeling

- Cluster identification

- Clusters / steps merging (based on outcome?)

- Prediction of the next steps / outcome

- Awareness of temporal dimension in model

Wongsuphasawat K, Plaisant C, Taieb-Maimon M, Shneiderman B. Querying Event Sequences by Exact Match or Similarity Search: Design and Empirical Evaluation. *Interact Comput*. 2012;24(2):55-68. doi:10.1016/j.intcom.2012.01.003
K. Wongsuphasawat and B. Shneiderman, "Finding comparable temporal categorical records: A similarity measure with an interactive visualization," 2009 IEEE Symposium on Visual Analytics Science and Technology, 2009, pp. 27-34, doi: 10.1109/VAST.2009.5332595.

Antoine Lamer, Emmanuel Chazard, Grégoire Ficheur

Discussion and perspective

Discussion and perspective

- Visualize the pathway is an help in generating hypotheses

- Need to select relevant steps and to focus on a segment of the pathway

- Current limit with reccurent and concurrent steps, and in the consideration of time