Reuse of the French National Health Insurance data for patients suffering from rare diseases: the Dromos project challenge

Intelligence artificielle et santé : approches interdisciplinaires

Nantes, 29 juin 2022

Anne-Sophie Jannot
annesophie.jannot@aphp.fr

HeKA

Université Paris Cité

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

MINISTÈRE DES SOLIDARITÉS ET DE LA SANTÉ
Liberté
Égalité
Fraternité

ASSISTANCE PUBLIQUE HÔPITAUX DE PARIS

Outline

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

- Context
- The linkage challenges
- The methodological challenges

The Dromos project

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

- Final objectives:
  - Describe the typical care pathways of rare disease patients
  - Quantify the cost associated with each disease
  - Describe inequalities in care and their links with territorial and socio-economic inequalities and the practices of expert centers

- Sources of data
  - National Health Data System
  - Rare Disease National Registry

The National Health Data System (SNDS)

BNDMR
Banque Nationale de Données Maladies Rares

DCIR

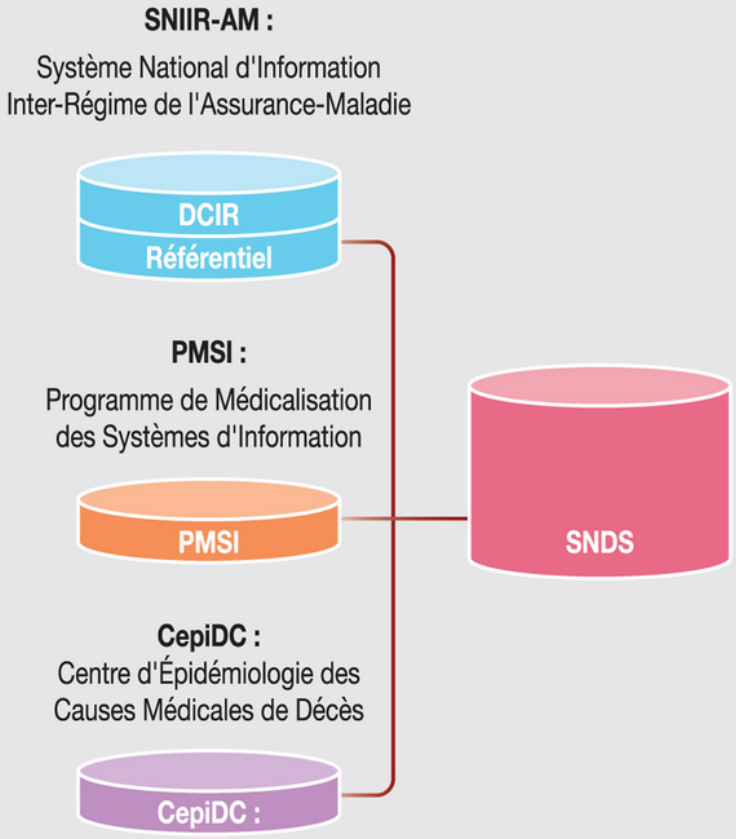
- Demographic area
- Date of death if applicable
- Ambulatory care reimbursements with dates of consultations
- Prescriptions
- Medical procedures
- Biological tests
- Medical devices
- Healthcare from other healthcare professionals

Référentiel

Records of chronic disease covered by the French reimbursement system

Uniform Hospital Discharge Data Set summarizing information about diagnoses, procedures and information about specific aspects of the hospital stay, e.g. a stay in an intensive care unit as well as the DRG assigned to the stay

Medical causes of death



- Offer the complete care pathways of all individuals in France over the last 10 years

How to identify patients suffering from a given rare disease in the SNDS?

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

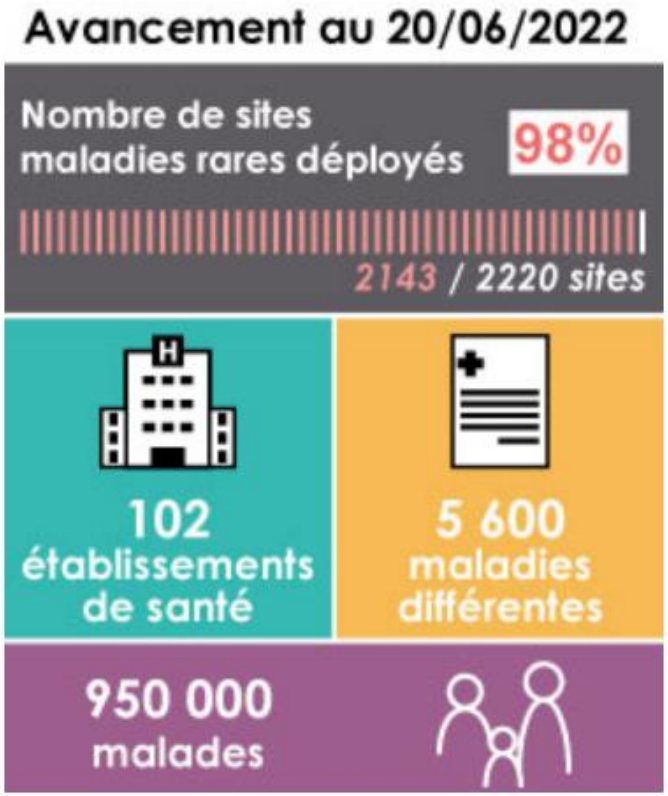
- Diagnoses are coded using the « international classification of disease » (ICD-10) published by the world health organization
- Few diagnoses available in the SNDS:
  - Declared chronic diseases
  - Diagnoses during hospitalizations
- Specificity of rare diseases:
  - Not always declared
  - Not always hospitalized
  - No ICD-10 code available for most rare disease

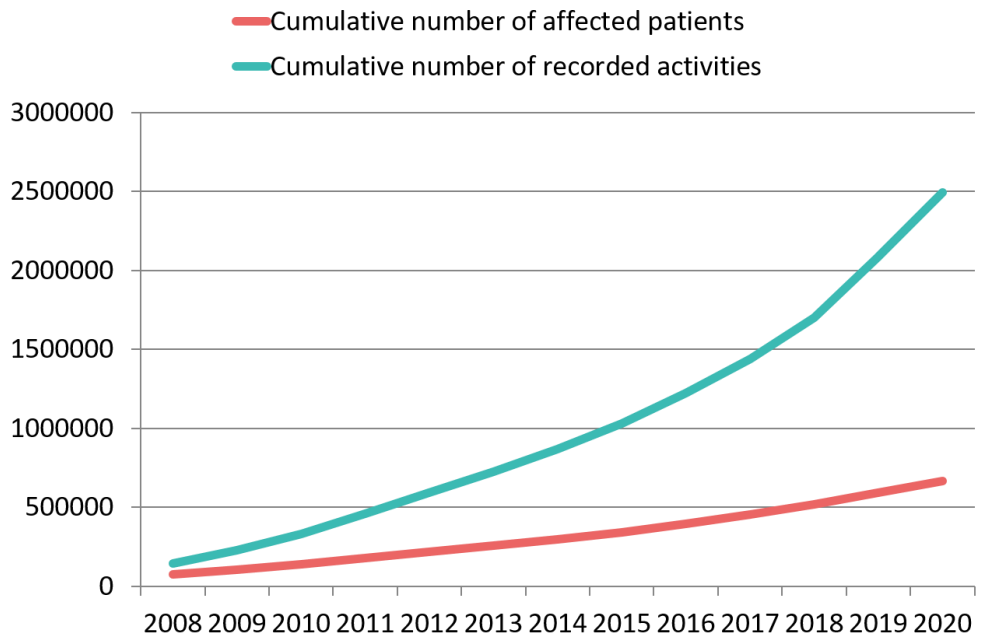
Patients suffering from a given rare disease cannot be identified in the SNDS

What is the National Rare Disease Registry (BNDMR)?

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

- Collects a common dataset for each patient followed in a rare disease expert center at each visit
- Financed by the National Plan for Rare Disease





What is the National Rare Disease Registry (BNDMR)?

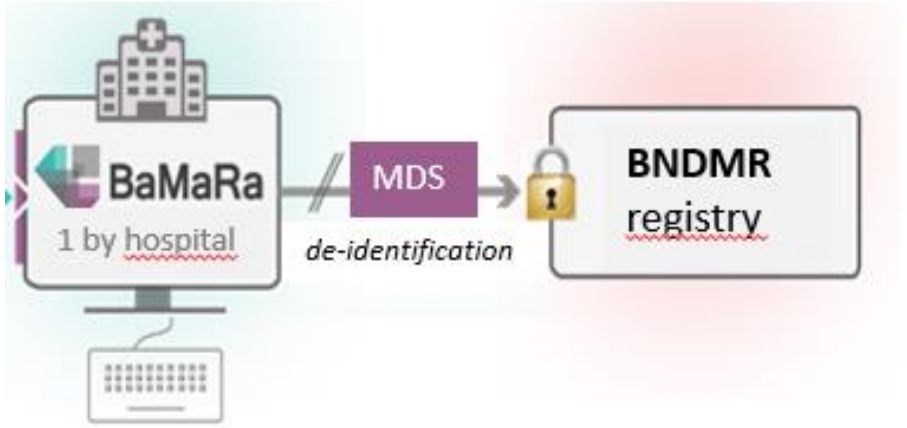
BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

The RD French minimum data set (MDS)

The French MDS contains about 60 items, including :

- Patient identification
- Vital status
- Care pathway & activities
- Ante and neonatal data
- Treatment (orphan drugs)
- Diagnosis (ORPHAcodes)
- Diagnostic history and accuracy
- Phenotype (HPO / ICD-10)
- Genotype (HGNC / HGVS)

Rare disease diagnoses classification



Local information systems

Clinical datawarehouse

Linking BNDMR and SNDS to identify care pathways for rare diseases

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

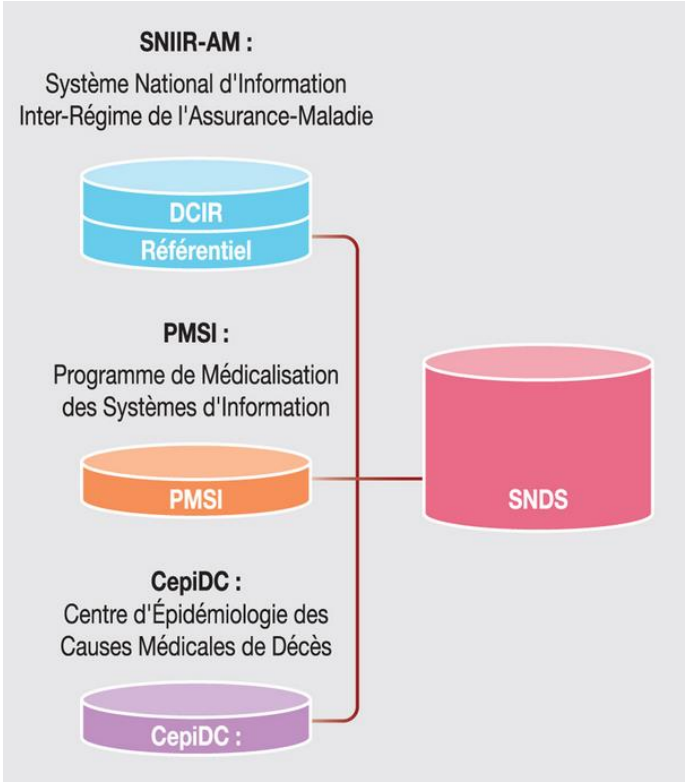
Patient identifier= name, surname, date of birth

The RD French minimum data set (MDS)

The French MDS contains about 60 items, including :

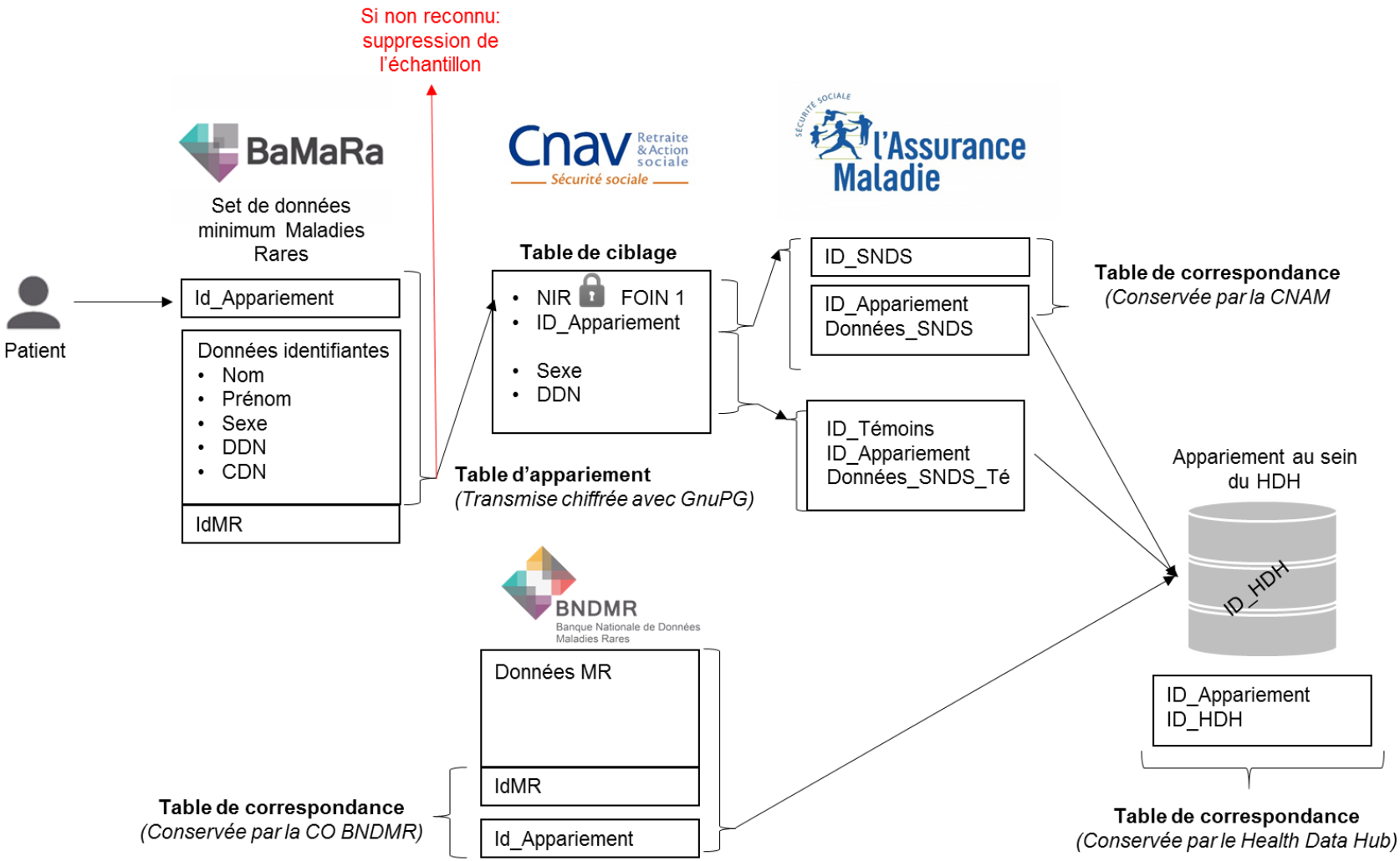
- Patient identification
- Vital status
- Care pathway & activities
- Ante and neonatal data
- Treatment (orphan drugs)
- Diagnosis (ORPHAcodes)
- Diagnostic history and accuracy
- Phenotype (HPO / ICD-10)
- Genotype (HGNC / HGVS)

Patient identifier = NIR

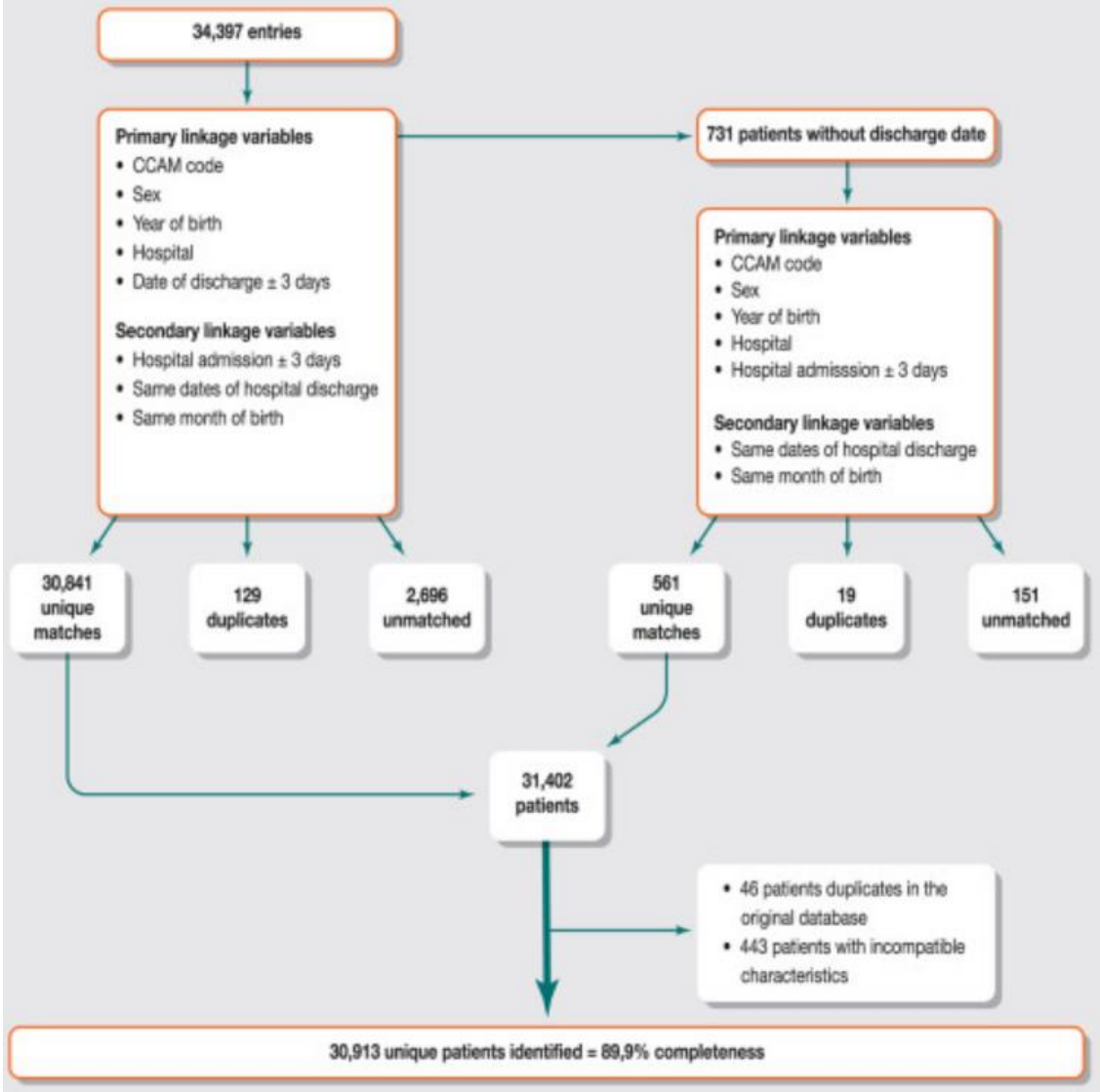


- No common identifier
- Two linkage strategies:
  - Direct: using patient identifier
  - Indirect: using

Direct linkage



Indirect linkage



Example of linkage with a registry

Many linkage variables are needed!

Didier, R., Gouysse, M., Eltchaninoff, H., Le Breton, H., Commeau, P., Cayla, G., ... & Gilard, M. (2020). Successful linkage of French large-scale national registry populations to national reimbursement data: improved data completeness and minimized loss to follow-up. *Archives of cardiovascular diseases*, *113*(8-9), 534-541.

Linkage strategies: limitations

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

- Direct linkage:
  - Need for strong identity quality
  - Multiple actors involved (CNAV, …)
- Indirect linkage:
  - Computationally intensive (+65 millions of people in SNDS)
  - Need for many common variables to obtain a high linkage rate

Modeling care pathways: Methodological challenges

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

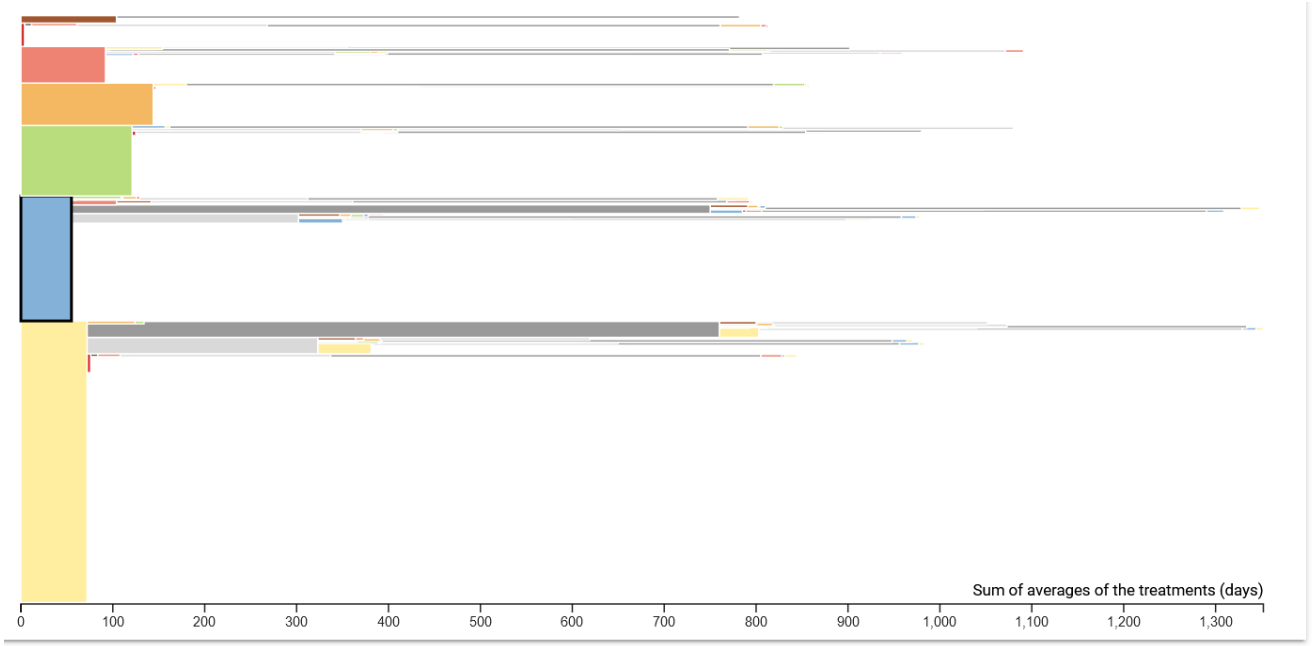
- Final objectives:
  - Describe the typical care pathways of rare disease patients
  - Quantify the cost associated with each disease
  - Describe inequalities in care and their links with territorial and socio-economic inequalities and the practices of expert centers

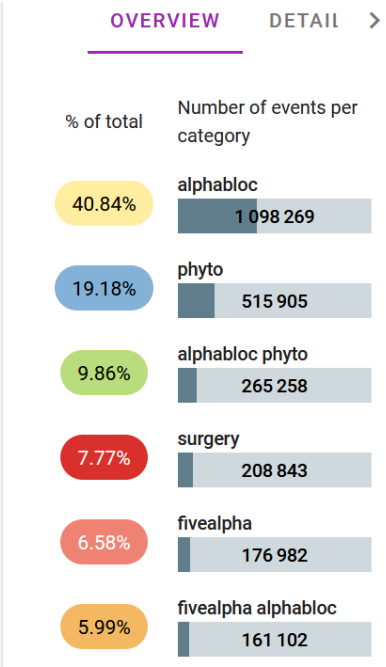
- Challenges using SNDS:
  - Thousands of variables
  - Each patient is followed at a different period of his/her life/disease: how to compare a child just diagnosed with an adult having been diagnosed at adult age 20 years ago?
  - Treatments evolve over time

How to describe the typical care pathways of rare disease patients?

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

- Visual approach focusing on relevant variables defined by expert





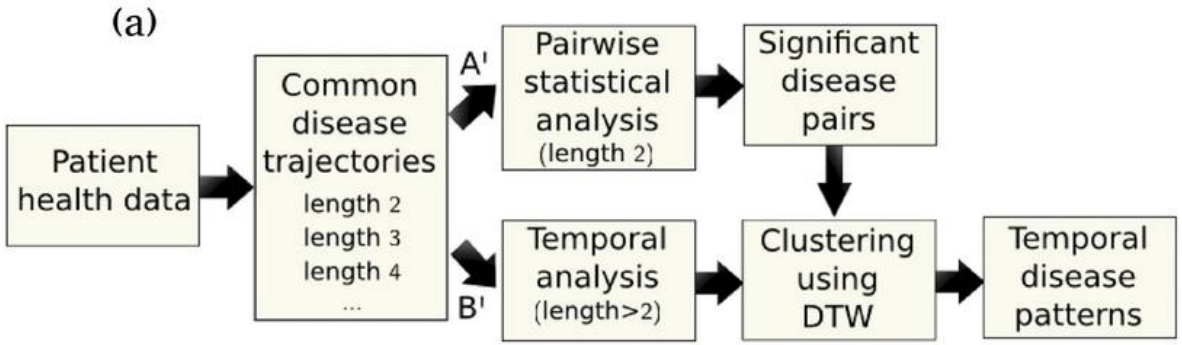
http://parcoursvis.lisn.upsaclay.fr/

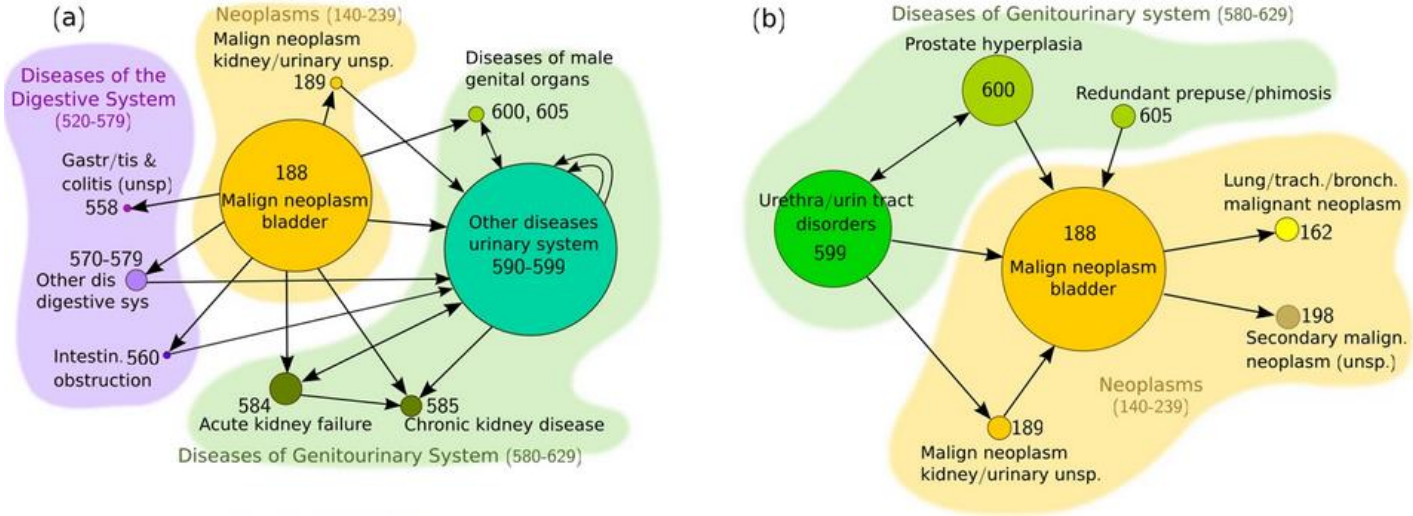
- Advantage: easy to interpret
- Limitation: need to define relevant variables, no subgroup definition, limited number of considered variables

How to describe the typical care pathways of rare disease patients?

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

- Process mining strategies





Can only be applied to events -> what to do with drugs?

No clear subgroups
Difficult to interpret

Giannoula, A., Gutierrez-Sacristan, A., Bravo, A., Sanz, F., Furlong, I. L. (2017). Identifying temporal patterns in patient disease trajectories using dynamic time warping : a population-based study. *Nature Scientific reports*

How to describe the typical care pathways of rare disease patients?

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

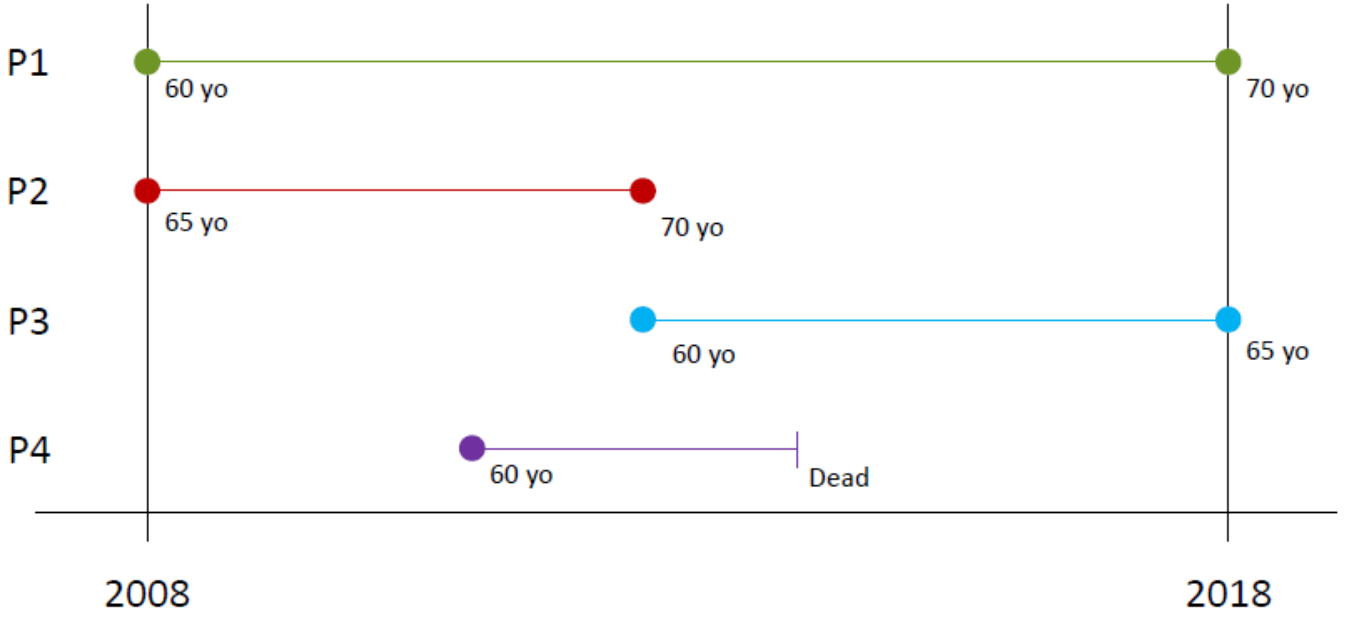
- Other approaches to handle longitudinal data: methods to cluster longitudinal data
  - Raw-based clustering
  - Feature based clustering (exemple: tsfresh)
  - Model-based clustering (mixture models)

- Limits:
  - Number of subgroups to specify a priori
  - High dimensional longitudinal modelling is challenging
  - Missing data are not handled

How to describe the typical care pathways of rare disease patients?

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

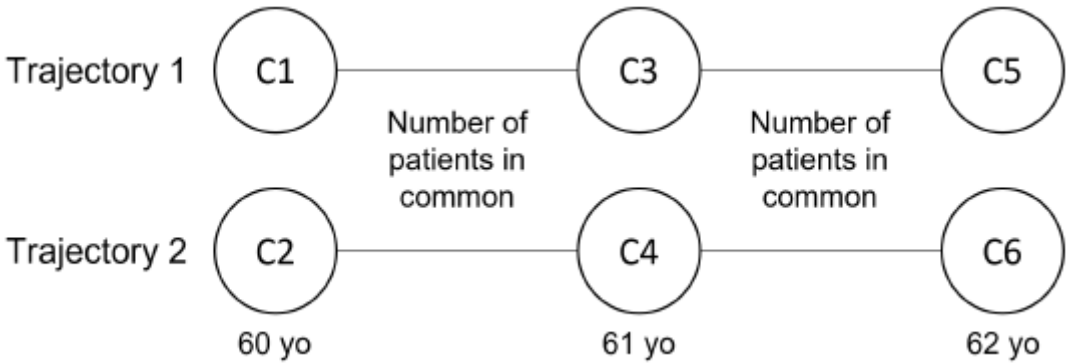
- Specific issue for SNDS data : different lengths of follow-up for each patient → truncated data



How to describe the typical care pathways of rare disease patients?

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

- Cluster tracking strategy:
  - Align patients (same age, start of the disease)
  - Cluster at each time step
  - Track clusters



Tracking temporal clusters from patient networks, Judith Lambert, Anne Louise Leutenegger, Anne Sophie Jannot, Anaïs Baudot, ongoing work

How to describe the typical care pathways of rare disease patients?

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

Age : 60

Patient	Aspirin	Combinations	Clopidogrel	Dabigatran	Enoxaparin	Fluindione	Fondaparinux	Tinzaparin
A	12	0	0	0	0	0	0	0
B	12	0	0	0	0	0	0	0
C	0	0	0	0	8	0	0	0
D	10	0	0	0	0	0	0	0
E	11	0	0	0	0	0	0	0
F	1	1	1	1	0	0	0	0
G	0	0	0	0	0	10	0	3
H	0	0	0	0	0	0	2	2

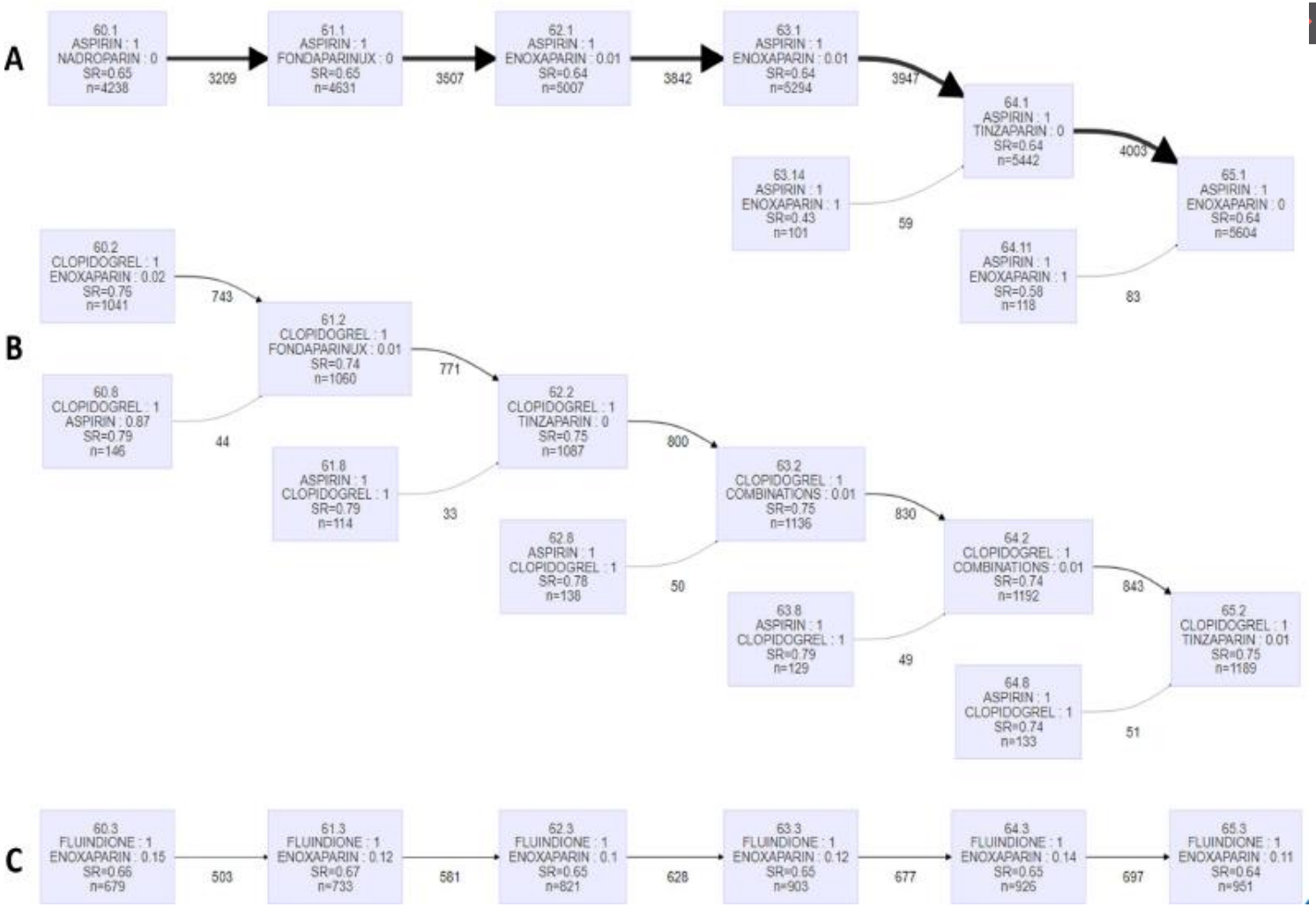
Combinations : Combinations of platelet aggregation inhibitors

	Cosine similarity	Jaccard index	Euclidean distance
A/B	1	1	1
C/D	0	0	0.54
E/F	0.5	0.25	0.64
G/H	0.20	0.33	0.63

How to describe the typical care pathways of rare disease patients?

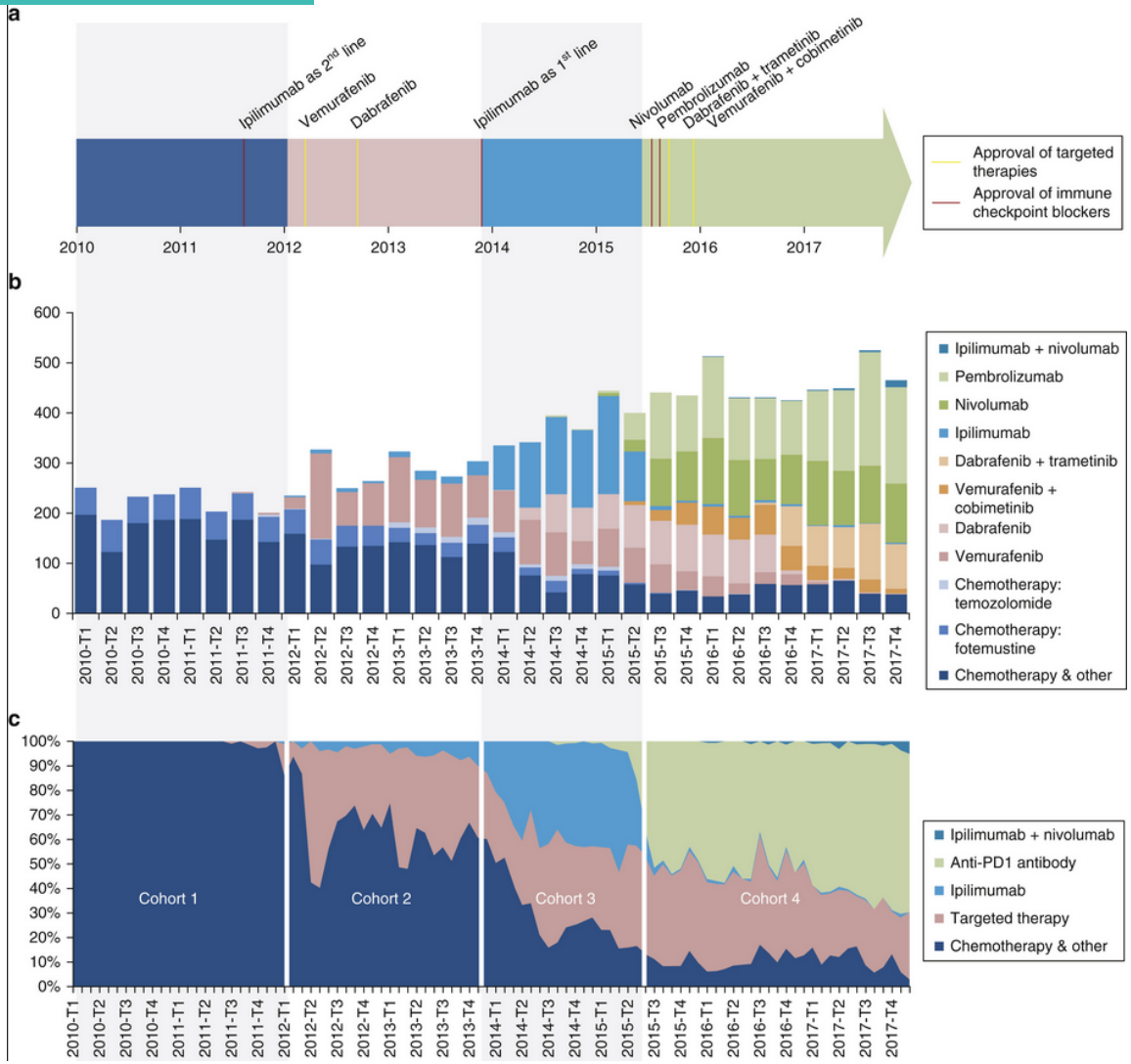
BNDMR
Banque Nationale de Données Maladies Rares

bndmr.fr



How to describe the typical care pathways of rare disease patients?

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

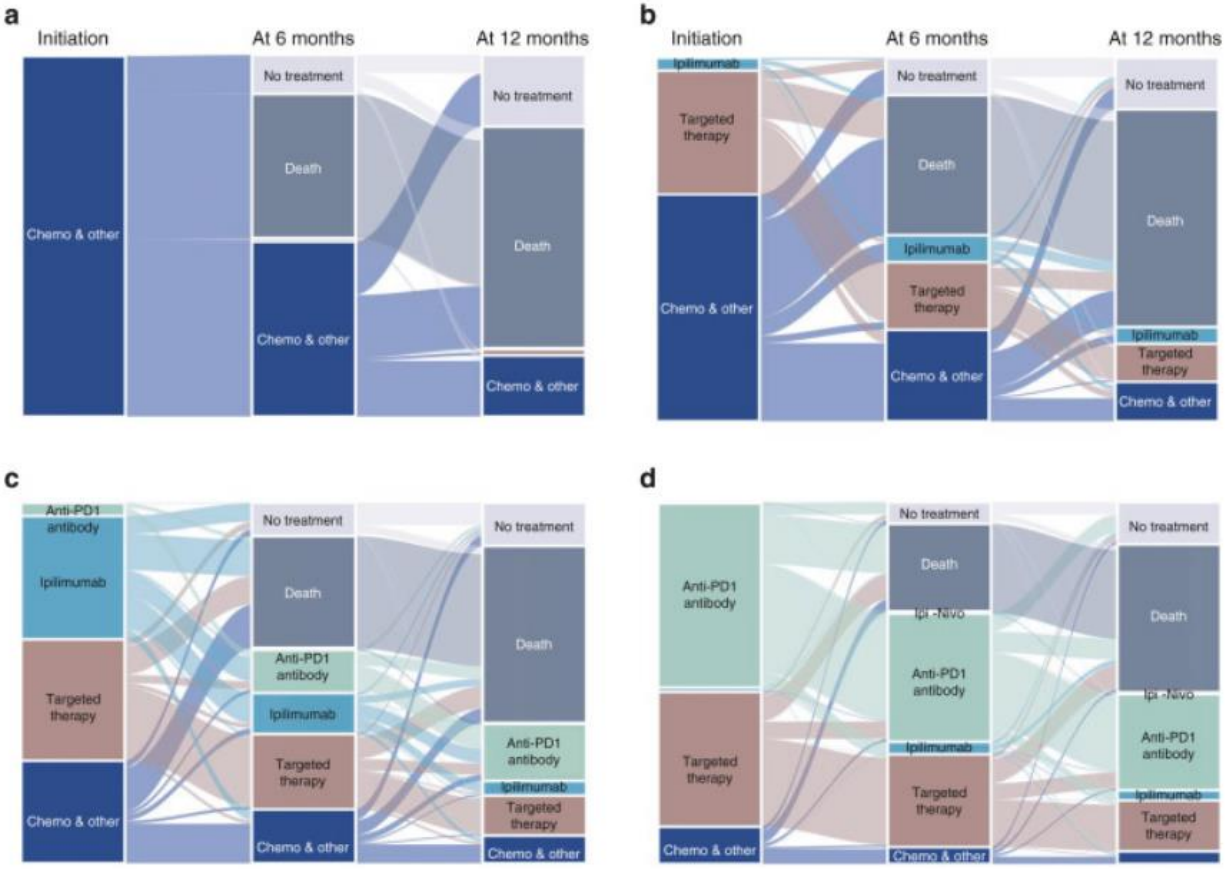


Need to define subgroups resulting in small sample sizes

Poizeau, F., Kerbrat, S., Happe, A., Rault, C., Drezen, E., Balusson, F., ... & Dupuy, A. (2021). Patients with metastatic melanoma receiving anticancer drugs: Changes in overall survival, 2010–2017. *Journal of Investigative Dermatology*, *141*(4), 830-839.

How to describe the typical care pathways of rare disease patients?

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr



Poizeau, F., Kerbrat, S., Happe, A., Rault, C., Drezen, E., Balusson, F., ... & Dupuy, A. (2021). Patients with metastatic melanoma receiving anticancer drugs: Changes in overall survival, 2010–2017. *Journal of Investigative Dermatology*, *141*(4), 830-839.

Project schedule

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

- July 2020: agreement of the scientific comittee of French Data Protection Agency (CESRESS)
- May 2021: agreement from the French Data Protection Agency (CNIL)
- Access denied for direct linkage: new agreement needed for indirect linkage
- New agreement still in process….

Reuse of the French National Health Insurance data for patients suffering from rare diseases: the Dromos project challenge

Intelligence artificielle et santé : approches interdisciplinaires

Nantes, 29 juin 2022

Anne-Sophie Jannot
annesophie.jannot@aphp.fr

HeKA

Université Paris Cité

BNDMR
Banque Nationale de Données Maladies Rares
bndmr.fr

MINISTÈRE DES SOLIDARITÉS ET DE LA SANTÉ
Liberté
Égalité
Fraternité

ASSISTANCE PUBLIQUE HÔPITAUX DE PARIS