HAL
open science

A visual approach of care pathways from the French nationwide SNDS database - from population to individual records: the ePEPS toolbox

André Happe, Erwan Drezen

To cite this version:

André Happe, Erwan Drezen. A visual approach of care pathways from the French nationwide SNDS database - from population to individual records: the ePEPS toolbox. Fundamental & Clinical Pharmacology, 2018, 32 (1), pp.81 - 84. 10.1111/fcp.12324 . hal-01697626

HAL Id: hal-01697626

https://univ-rennes.hal.science/hal-01697626

Submitted on 31 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DR. ANDRÉ HAPPE (Orcid ID : 0000-0001-9487-1533)

Article type : Short Communication

A visual approach of care pathways from the French nationwide SNDS database. From population to individual records: the ePEPS toolbox

Running head: Visualisation of care pathways from the French SNDS database

André HAPPE [a,b,*], Erwan DREZEN [a,c]

[a] Plateforme PEPS/ANSM, UPRES EA 7449 REPERES, Rennes, France

[b] Centre Hospitalier Régional et Universitaire (CHRU) de Brest, Centre de Données Cliniques, 2 Bld Tanguy, 29609 Brest, France

[c] Centre Hospitalier Universitaire (CHU) de Rennes, Rennes, France

***Corresponding author**
André HAPPE
andre.happe@chu-brest.fr

ABSTRACT

Secondary use of medical and administrative databases has become a powerful tool for epidemiological studies. In that respect, the recent access opening of French nationwide health record database or SNDS (Système National des Données de Santé) is a great opportunity to carry out comprehensive health studies at the country level. However, using this database is far from being straightforward for non-expert data scientists; so, dedicated tools needed to be developed. Our contribution is focused on data management and visualization tools able to query and cope with the complexity of care pathways of patients. The scope of the toolbox presented below goes from interactive tools for building groups of patients to timeline representations of individual patient healthcare trajectories.

Accepted Article

Keywords: SNDS , Care pathways , Datavisualization

Accepted Article

INTRODUCTION

The French healthcare system relies on universal coverage of health expenditures by different health care insurance plans. The SNDS database gathers anonymous information of reimbursed claims from most of these plans, linked to the national hospital-discharge summaries database (PMSI) and very soon the national death registry. It was opened on April 2017 and covers 98.8% of the French population, i.e. 66 million people, which constitutes one of the largest medical administrative database in the world.

The database includes demographic data, health care providers and prescribers, medical or paramedical visits, reimbursed drug deliveries, medical devices or lab tests, chronic diseases (ICD10 codes), hospital stays with ICD10 diagnoses codes, dates, medical or surgical procedures, diagnostic-related groups, rehabilitation stays, stays in psychiatric institutions. The power of the database relies on its representativeness while including almost the entire French population but its use stays delicate due the complexity of the data model [1].

Healthcare analytics is the mean by which large scale data like the SNDS database can be refined and interpreted into relevant information. While such massive datasets provide valuable insights and novel solutions, they are also challenging due to volume and variety. Visual representations take advantage of the human eye's ability to provide ways to users to see, explore, and understand large amounts of information at once. Visual analytics is the research field we are exploring to reduce both volume and variety in order to sharpen the analytic focus [2].

We present our approach of visualization tools developed to help SNDS users exploring those large datasets from the population level to the atomic individual patient pathway. Our goal was to build a toolbox (named ePEPS) to provide SNDS users with means to exploit this large and complex database. It encompasses global exploration tools to subset group of patients according to the presence of given events in their healthcare pathways. It is also possible to produce basic statistical indicators of the groups through visual outputs. Finally it provides interactive visual representation of the individual patient trajectories that can be customized and filtered by the end user.

Accepted Article

DEFINING GROUPS OF PATIENTS

The first module of the ePEPS toolbox provides a user interface to define groups of patients according to the presence of given events in their care pathways. Figure 1 (panel A) lists most of the medical classifications available in the SNDS database (ICD-10, CCAM, NABM, ATC, ...). They are presented in their hierarchical forms and the user can explore the hierarchical trees by unfolding or hiding the different levels of classification. Each leaf of the tree provides the number of patients concerned at least once by this event in their care pathway and the relative percentage from the whole population.

A sandbox window (Figure 1, B) provides a way to define groups of patients. By dragging and dropping a chosen item from the top left panel to the sandbox window, the user is building step by step its group (i.e. patients that have at least one drug reimbursement from angiotensin II receptor antagonist - C09C ATC class). The different items can be combined with boolean operators (union, intersection, minus) in order to elaborate complex selection criteria. The number of patients concerned by the new settings is calculated dynamically.

VIEWING CHARACTERISTICS OF A GROUP OF PATIENTS

Once a group of patients has been built, the user can get raw characteristics of the group through several visualization views such as (Figure 1, blue frames):

1. Age pyramid: classical age pyramid view; possible use to compare a reference population (plain bars in blue/pink colors) to one or more population displayed with lines; Here the reference (ARA) and compared (IEC) populations presents similar distribution for men while rather different distribution for women.
2. Geographical distribution: possibility to zoom (up to city level) by clicking on the map.
3. Classification distribution: represents the distribution of the group over a given hierarchical classification (here ICD-10).
4. Venn diagram: Venn diagram of sub-groups for the selected population defined by drug consumption.
5. Patients list: displays information about the patients of the group (sensitive information was blanked).

By default, the user can update all these views for a given group by double-clicking on the item representing it; it is also possible to update a single view by drag-and-dropping an item to the view. For instance, the Venn diagram on figure 1-(4) corresponds to the "IEC" item of the sandbox (n = 23.062), all the other views correspond to the sandbox item previously double-clicked by the user (n = 4.163).

VIEWING TRAJECTORIES OF PATIENTS

Once a group of patients has been selected and characterized, users may want to observe more precisely the care pathway events for a given patient. In order to summarize all these events in a single view, the ePEPS toolbox provides a timeline representation of a patient care pathway. Invoking a timeline for a patient can be done by double-clicking on one item from the patient list (Figure 1, part 5) or by inputting their identifier.

The timeline (Figure 2) piles up several views selected by the user, each view corresponding to an information source available in the SNDS: medical visits, drug reimbursements, medical acts, biological analysis, hospital stays, etc. A view is made of 1) a left part defining how to organize the different temporal events through attributes of the information source, 2) the events displayed as a rectangle along a time axis. In the example, the green view shows hospital stays, organized on the left by two columns, one column for the first ICD-10 level of the principal diagnosis and one column for the third ICD-10 level; the pink view shows rehabilitation stays with two columns for two different ICD-10 diagnosis; the yellow view shows drug deliveries organized by the first and third ATC levels. Users can customize the timeline by zooming in/out the time axis, filtering out event types on the columns, adding/removing columns, renaming columns, etc.

Such a timeline representation allows embracing complementary sources of information from the SNDS in a single view; in the example, we can notice a "hole" in the drug deliveries from March to May 2015. It can be explained by the rehabilitation stay (second line of the pink view) during which the patient did not go the pharmacy. It should be noticed that achieving a similar understanding would have been quite difficult by merely analyzing tabular query results from the SNDS SQL database.

Accepted Article

Accepted Article

CONCLUSION

As stated in the Care Pathway Explorer project [3], an interactive hierarchical information exploration system is useful for analyzing longitudinal medical data. By providing an overview of relevant patterns visually mined from reimbursement patient traces in the SNDS database, the ePEPS toolbox supports interactive exploration for researchers. Moreover, instead of visualizing all details at once, it allows end users to scan the level of detail relevant for their use cases by leveraging medical hierarchical classifications.

We are currently working on extensions of the group building process. In particular, we investigate how to search efficiently all patients matching a temporal request like "Hospital stay for reason A after a drug B delivery occurring between 0 and 30 days before".

FIGURE LEGENDS

Figure 1: Building group of patients

Figure 2: Timeline of a patient care pathway

REFERENCES

[1] Bezin J., Duong M., Lassalle R. et al. The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. Pharmacoepidemiol. Drug Saf. (2017) **26** 954-962

[2] Du F., Shneiderman B., Plaisant C., Malik S., Perer A. Coping with Volume and Variety in Temporal Event Sequences: Strategies for Sharpening Analytic Focus. IEEE Trans. Vis. Comp. Graph. (2017) **23** 1636-1649

[3] Perer A., Wang F., Hu J. Mining and exploring care pathways from electronic medical records with visual analytics. J. Biomed. Inform. (2015) **56** 369-378.

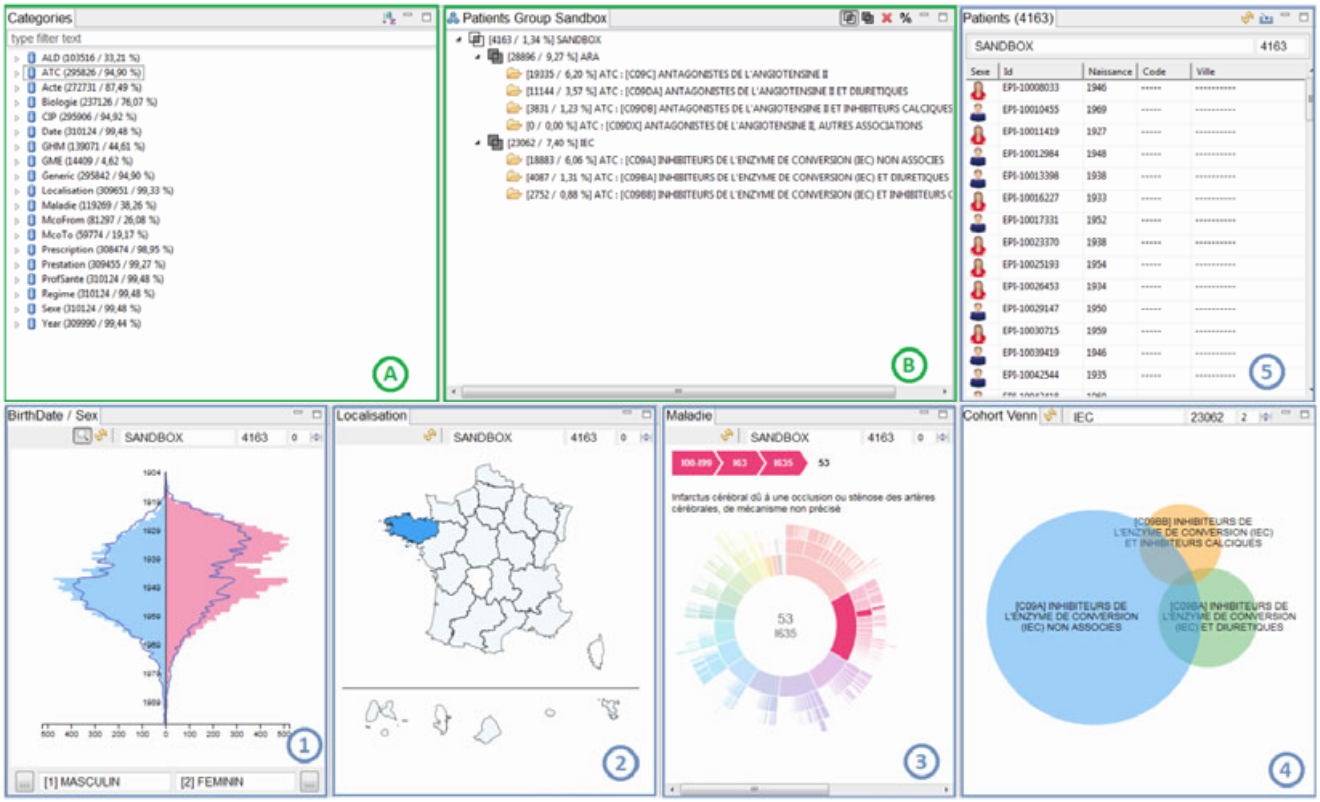

Figure 1: Building group of patients

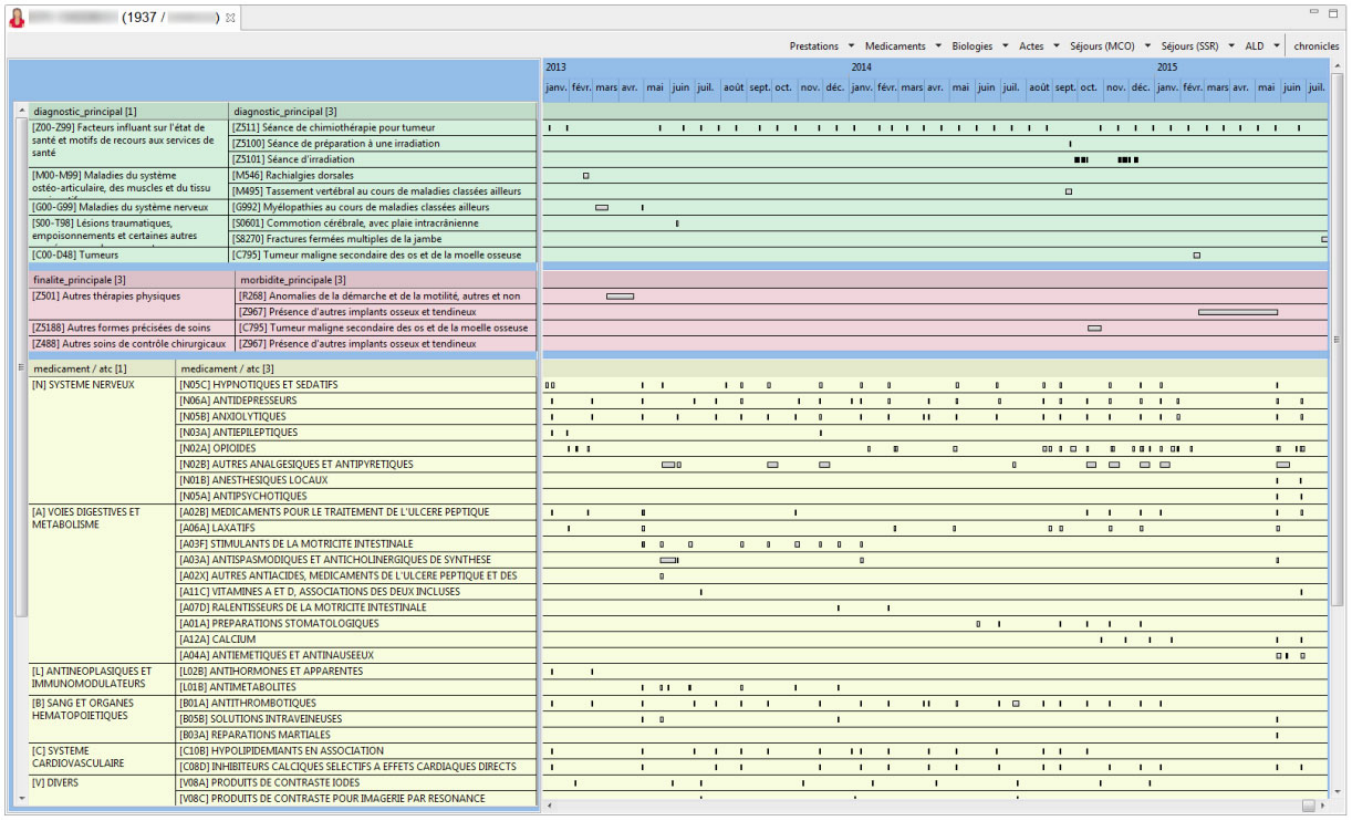

Figure 2: Timeline of a patient care pathway

Accepted Article