Digital Personalized Health and Medicine
L.B. Pape-Haugaard et al. (Eds.)

doi:10.3233/SHTI200154

Exploring Patient Path Through Sankey Diagram: A Proof of Concept

Antoine LAMER[a,b,1], Gery LAURENT[a], Sylvia PELAYO[a], Mehdi EL AMRANI[c], Emmanuel CHAZARD[b] and Romaric MARCILLY[a]
[a] *Univ. Lille, INSERM, CHU Lille, CIC-IT/Evalab 1403 - Centre d'Investigation clinique, EA 2694, F-59000 Lille, France*
[b] *Univ. Lille, CHU Lille, Public Health Department, EA 2694, F-59000 Lille, France*
[c] *Univ. Lille, CHU Lille, Department of Digestive Surgery and Transplantation, F- 59000 Lille, France*

Abstract. Managers, physicians and researchers need to study patient's path for purposes of management, quality of care and research. We present the proof of concept of the use of a flow diagram, the Sankey diagram, to visualize the trajectory of a population that experienced an event. This representation was tested with two case studies in populations from the anesthesia data warehouse of Lille University Hospital. For the 551 patients undergoing a pancreaticoduodenectomy, Sankey diagram helped us identify atypical care paths of patient being transferred too late in an intensive care unit. For 473953 patients who have had anesthesia procedure, Sankey diagram highlighted that mortality and re-operation rates increase with the number of operations. This preliminary work has been well received by end-users and allowed managers, physicians and researchers to visualize the paths of patients and to provide visualization support for research questions. This work will be followed by generalization.

Keywords. Data Visualization, Data Reuse, Patient Path, Sankey Diagram

1. Introduction

To display complex information in a most effective way, graphical representations are more appropriate than tables and figures [1, 2]. In healthcare facilities, to have access to visual representations of patients' path would be of great value for assessing quality of care (e.g. to detect atypical care paths) and for management purpose [3]. From a research perspective, studying patient flows may generate hypotheses about patient care that can be assessed later with classical biostatistics methods.

Designing and providing clear and efficient visualization of patients' paths faces several issues:

- Hospital Information Systems (HIS) generate a huge amount of data and the patient path is documented by many variables (admission and discharges modes, type of units, succession of interventions and re-interventions, medical acts, diagnostics, etc). This makes it difficult to extract relevant information;

---

[1] Corresponding Author, Antoine Lamer, 2 avenue Oscar Lambret, 59037 Lille cedex, France; E-mail: antoine.lamer@chru-lille.fr.

- Patients' stays differ in complexity and may represent a large number of patterns (e.g. one patient went through two care units and underwent one operation while another went through 6 units and underwent two operations).

Several kinds of diagrams are commonly used to represent flows. From a practical point of view, a flow is a succession of steps, each one defined by an initial state, a final state, and a quantity. From a graphical point of view, it is represented by links carrying information about quantity from start nodes to end nodes. Table 1 compares 8 types of flows diagrams [4]. Most of these representations enable to visualize only one step of a process. On the contrary, the Sankey diagram enables to represent several steps along with the information about flows' volumes, and to manage the complexity of the patient paths.

This work aims to assess whether the Sankey diagram is feasible to represent patient paths, in a way that would be readable and understandable by clinicians. With this goal in mind, two cases were used: (i) paths of patients with the pancreaticoduodenectomy and (ii) association between re-operation and mortality.

Table 1. Comparison of flow diagrams

Flow diagram	Advantage	Limit
Network Diagram highlights connections between nodes at a given step.	The connections may be directed and weighted.	It addresses only one step of the whole path.
Chord Diagram is similar to network diagram, except that nodes are represented in a circle.	The same as Network diagram.	It addresses only one step of the whole path and presents reading difficulties.
Dendrogram is a tree displaying paths from the first to last step and considering the intermediate states.	It displays successive steps of the path with detail about the count.	It displays as many final nodes as paths, which overwhelms the diagram.
Sunburst Diagram is a circular representation of the dendrogram: the inner circle represents the first step of the process and each additional layer represents a new step.	It displays successive steps of the path with detail about the count, and aggregate links at every steps.	The number of steps must be limited. At each step, patients with the same criteria are displayed on different parts of the circle, which makes the reading harder.
Sankey Diagram displays flows between successive steps through arrows whose thickness is proportional to flow quantity.	It displays successive steps of the path with detail about the count.	For clarity sake, the number of steps must be limited.
In the **Agenda Diagram**, each row represents an individual and each cell represents a time unit.	It displays every detail for each patient.	It does not synthesize information on the whole population.
Connection Map is connecting locations placed on a map.	It highlights flows between distant places.	It represents only one step of the path and does not consider the volume of flows and the temporal sequence.
Funnel Chart represents the successive proportions of a given population that meet given criteria.	Readability.	Only one final sub-population.

2. Materiel and methods

We worked on two different cases to illustrate the design of the Sankey diagram used to represent patient paths. Each case aimed to answer a question: (i) in the digestive surgery department, what is the patient volume in each care unit after surgical operation for pancreaticoduodenectomy and what are the atypical patient paths? (ii) how mortality is related to the number of surgical operations during an hospital stay?

To design the Sankey diagrams and representing answers to those questions, we used data from the anesthesia data warehouse of Lille University Hospital. This data warehouse stores and organizes data about surgery and anesthesia procedures (drugs, vital signs, steps of surgery) with administrative data (hospital stay characteristics, medical unit, entry and discharge modes, medical and surgical acts, diagnoses) [5,6].

The Sankey diagrams were designed following three stages. The first stage is crucial and consists in determining information needed to answer the question. The information are summarized in Table 2.

Table 2. Aggregated data to upload in the graphical library

Source node	Case study 1	Case study 2
Population	Patients who undergo a duodenopancreatectomy (HNFA007 code)	Patients who undergo an operation
Event under scrutiny	Admission to unplanned critical care unit after operation	Death and re-operation rates
Steps of the path	Successive medical units after operation	Discharge status after each operation

The second stage consists in aggregating the data and compute the count of patients in each step between two successive states of the path. For each link, we had to provide a source node, a target node and the quantity between the two nodes. Finally, these aggregated data are loaded into graphical libraries (D3.js and D3sankey [7]).

The resulting Sankey diagrams were presented to clinicians to get their feedback on the readability, the understandability of this visual representation and its impact.

3. Results

3.1 Case study 1 : Patient flow for pancreaticoduodenectomy

Operations of pancreaticoduodenectomy were selected based on the code HNFA007 of the French medical act classification (CCAM, Classification Commune des Actes Médicaux). For operations between 2010 and 2018, 551 patients were included. We selected two significant states of the hospital stay for pancreaticoduodenectomy: the post-operative unit and the last unit before discharge from the hospital. These two states were characterized with the type of unit (conventional, continuous care or intensive care unit) and final discharge status (mortality) was reported. The Figure 2 is the Sankey representation for this population.

Two managers and two physicians examined the Sankey diagram. They were able to describe patient flows in each unit after operation. Moreover, physicians highlighted atypical paths: some patients were transferred from conventional care unit to critical care after operation, and some patient died even thought they did not go through critical care beforehand. They reported that this Sankey diagram was intuitive to read, more understandable than a table but that they needed contextual information when the graph is accessed for the first time. Physicians will now investigate why these patients were not treated earlier in an appropriate care unit.

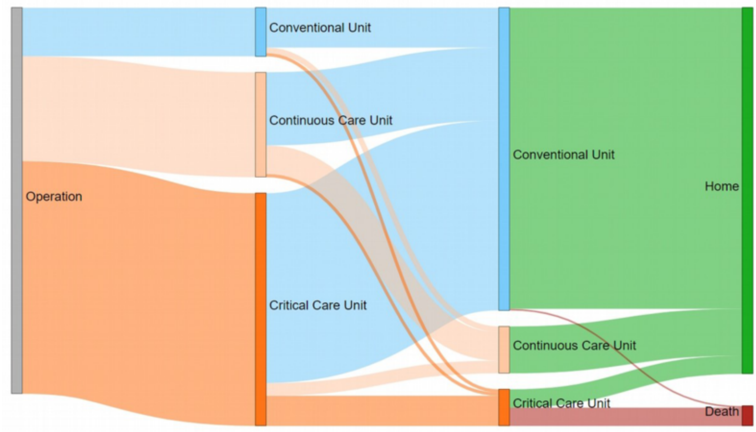

Figure 1. Patient flow for pancreaticoduodenectomy

3.2 Case study 2: Mortality rates after re-operation

Between 2010 and 2018, 473953 patients underwent operation. For each patient, we computed the number of operations, and final discharge status (death). In order to get the most out of the diagram, two key points were relevant. First, we had to define a maximum number of operations to be covered. Secondly, we had to display results in percentages to compare modalities with small numbers (3rd and 4th surgical procedures) with larger modalities (1st and 2nd surgical procedures). With the Sankey diagram in Figure 2, we can see that mortality and re-operation rates increase with the number of operations performed during the same hospital stay. This diagram has been shown to 2 managers and 2 physicians, and they all agreed on the intuitiveness of the information displayed. They only asked for contextual information to know the source of the data. According to them, this result will be the starting point of research program about how going through multiple operations could impact mortality rates.

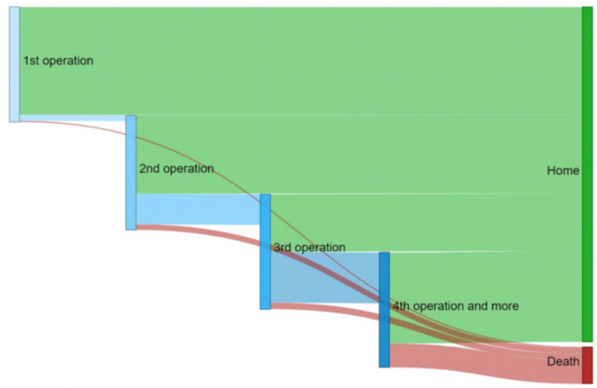

Figure 2. Sankey diagram representing the mortality and re-operation rates according to the number of operations during a hospital stay

4. Discussion

In this paper, we offer a way of visualizing patients' path with the Sankey diagram. This representation allowed managers, physicians and researchers to visualize the paths of patients, to detect atypical patient flows between units and to provide visualization support for research questions. We will now investigate which signals in operating room and post-anesthesia care unit could be use to detect patients prone to complications and refer them to an intensive care unit as soon as possible.

Usually, end-users are provided with data in a table format, on a case-by-case basis, or for a single step of the patient path. With the Sankey diagram, the benefits reported by users are that this representation is easily understandable, gives a useful overview of patient path and offers the opportunity to identify atypical paths that they may not have seen otherwise. It could also be used as a way to validate their hypotheses and further investigate.

However, the implementation of such diagrams might be challenging as patient paths have more than two steps and are complex. Thus, the data aggregation must be adapted to the specific issue to deal with. Morevover, the steps displayed must be selected to highlight important facts and avoid overload of information.

A limitation of the Sankey diagram, with its current implementation, is that it is not possible to follow an individual path from the first to the last node. The next step is to improve the current Sankey representation by adding a traceable multi-level feature. This would represent a more precise way to highlight a specific path.

Another perspective would be to offer an online Sankey visualisation tool, in which the users would select the targeted population (e.g. a surgical act) and predefined steps (e.g. surgical complication) to have a display of the corresponding patients' path.

References

[1] West VL, Borland D, Hammond WE. Innovative information visualization of electronic health record data: a systematic review. J Am Med Inform Assoc JAMIA. mars 2015;22(2):330-9.
[2] Few S. Information Dashboard Design : The Effective Visual Communication of Data / S. Few. 1 janv 2006;
[3] Chittaro L. Information visualization and its application to medicine. Artif Intell Med. mai 2001;22(2):81-8.
[4] Holtz Y, Healy C. From data to Viz | Find the graphic you need [Internet]. [cited on 24 th sept 2019].
[5] Available on : data-to-viz.com
[6] Lamer A, Jeanne M, Vallet B, Ditilyeu G, Delaby F, Tavernier B, et al. Development of an anesthesia data warehouse: Preliminary results. IRBM. 2013 Dec;34(6):376–8.
[7] A. Lamer, G. Ficheur, L. Rousselet, M. van Berleere, E. Chazard, A. Caron, From Data Extraction to Analysis: Proposal of a Methodology to Optimize Hospital Data Reuse Process, Stud. Health Technol. Inform 247 (2018), 41–45.
[8] Bostock M. D3.js - Data-Driven Documents [Internet]. [cited on 9th July 2019]. Available on: https://d3js.org/
[9] CCAM en ligne - CCAM [Internet]. [cited on 8th January 2019]. Available on: http://www.ameli.fr/accueil-de-la-ccam/index.php