Quels évènements après un infarctus du myocarde ?

Jessica Pinaire, Jérôme Azé, Sandra Bringay, Pascal Poncelet, Christophe Genolini, Paul Landais

▶ **To cite this version:**

Jessica Pinaire, Jérôme Azé, Sandra Bringay, Pascal Poncelet, Christophe Genolini, et al.. Quels évènements après un infarctus du myocarde ?. SIIM: Symposium Ingénierie de l'Information Médicale, Nov 2017, Toulouse, France. hal-01652800

HAL Id: hal-01652800
https://hal.science/hal-01652800

Submitted on 30 Nov 2017

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.

Quels évènements après un infarctus du myocarde ?

Jessica Pinaire[*,**,***], Jerôme Azé[**]
Sandra Bringay[**,****] Pascal Poncelet[**], Christophe Genolini[♯,♯♯], Paul Landais[**]

[*]Équipe d'accueil 2415, IURC, Université Montpellier, France
paul.landais@umontpellier.fr
[**] LIRMM, UMR 5506, Université Montpellier, France
prenom.nom@lirmm.fr
[***]D*IM*, CHU de Nîmes, France
jessica.pinaire@chu-nimes.fr,
[****]AMIS, Université Paul-Valéry Montpellier 3, Montpellier, France
sandra.bringay@univ-montp3.fr
[♯♯] CeRSM (EA 2931), UFR STAPS, Université Paris Ouest-Nanterre-La Défense, Nanterre, France
christophe.genolini@parisnanterre.fr
[♯] Start'up Zebrys, 5 place Jean Deschamps, 31100 Toulouse, France

Résumé. Connaître les flux de patients peut être décisif pour la planification sanitaire. Dans cet article, nous proposons une méthode pour caractériser les flux de patients mais aussi pour mettre en évidence des profils de délais et de tarifs des parcours de soins. À partir des données du PMSI (Programme Médicalisé des Systèmes d'Information), nous avons extrait des motifs spatio-temporels puis les avons intégrés dans un outil de visualisation. Ensuite, nous avons classé les trajectoires de délais et de tarifs afin d'identifier des profils en termes d'occurrence et d'évolution de tarifs des parcours de soins. Cette approche pourrait permettre d'aider à l'élaboration de stratégies de prises en charge des patients.

1 Introduction

Au cours de ces 20 dernières années, les études sur l'infarctus du myocarde (*IM*) (Dégano et al., 2015) soulignent un changement dans les tendances selon le sexe montrant une augmentation à la fois du risque cardiovasculaire mais aussi de la mortalité chez les femmes. On observe en parallèle, une augmentation du nombre d'hospitalisations. De plus, le tarif des consommations de soins après un *IM*, est multiplié par 3 sur l'année suivant l'*IM* comparativement à l'année le précédant (Philippe et al., 2017).

Dans un tel contexte, l'analyse des flux de patients apparaît opportune : 1) pour prévoir les admissions de patients ; 2) pour adapter le parcours du patient ; 3) pour gérer la disponibilité des ressources. Pour modéliser les flux de patients et construire des outils d'aide à la décision, les chercheurs utilisent aussi bien des méthodes de fouille de données (Pagnoni et al., 2001) que des modèles prédictifs basés sur des arbres de décision (Dart et al., 2003).

Dans cet article, nous proposons de caractériser les flux de patients, présentant un *IM*, par une approche innovante. À partir des données du PMSI (Programme Médicalisé des Systèmes d'Information), nous avons extrait des motifs caractérisant des trajectoires en adaptant une méthode de recherche de motifs spatio-temporels. L'originalité ici est de ne plus considérer l'espace dans ces motifs, mais la proximité des pathologies codées dans les bases du PMSI. Par ailleurs, nous avons classé ces trajectoires pour identifier des tendances dans les délais et l'évolution des tarifs. Les motifs spatio-temporels ont été intégrés dans un outil de visualisation et permettent de reconstituer les différentes évolutions possibles de cette maladie. De plus, l'association aux profils de délais et de tarifs pourrait permettre d'envisager des stratégies dans l'organisation des soins.

Dans la section 2, nous introduisons le vocabulaire spécifique au domaine de la recherche de motifs spatio-temporels et nous l'illustrons avec un exemple d'application à partir des données du PMSI. Ensuite, dans la section 3, nous présentons le processus de caractérisation des flux de patients. Dans la section 4, nous identifions à la fois des phénomènes de groupes, des profils d'évolution de tarifs et d'occurrence des hospitalisations chez les femmes de +65 ans. Enfin, nous discutons des résultats obtenus et des limites de cette approche dans la section 5.

2 Définitions préliminaires

2.1 Motifs spatio-temporels

La fouille de données spatio-temporelles consiste à discerner des ensembles d'objets restant groupés spatialement sur une même période. Il existe de nombreux motifs spatio-temporels. Nous présentons uniquement ceux utilisés dans cet article. On note : $O = \{o_1, ..., o_n\}$ un groupe d'objets mobiles ; $T = \{t_1, ..., t_p\}$ un ensemble d'estampilles temporelles ; $x_{t_i}^{o_j}, y_{t_i}^{o_j}$ les informations spatiales de $o_j \in O$ au temps t_i ; min_o un support minimum, donné par l'utilisateur, correspondant au nombre d'objets minimum devant être ensemble ; min_t, le nombre minimum d'estampilles temporelles pendant lesquelles au moins min_o objets de O sont regroupés.

Définition 1 (Essaim) *Un couple (O, T) est un essaim (ou swarm en anglais) si :*

1) *Il y a au moins un cluster contenant tout objet de O à chaque estampille temporelle de T : $\forall\, t_i \in T, \exists\, c$ tel que $O \subseteq c$ où c est un cluster ;*
2) *Il y a au moins min_o objets : $|O| \geq min_o$;*
3) *Il y a au moins min_t estampilles temporelles : $|T| \geq min_t$.*

De manière intuitive, un essaim est un groupe d'objets contenant au moins min_o éléments qui sont proches les uns des autres pour au moins min_t estampilles de temps.

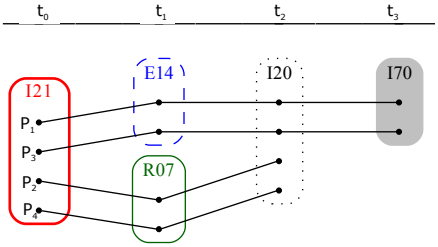
Définition 2 (Essaim clos) *Un couple (O, T) est un essaim clos (closed swarm) s'il vérifie :*

1) *(O, T) est un essaim ;*
2) *(O, T) est objet-fermé : $\nexists\, O'$ tel que (O', T) est un essaim et $O \subset O'$;*
3) *(O, T) est temporel-fermé : $\nexists\, T'$ tel que (O, T') est un essaim et $T \subset T'$.*

2.2 Exemple

Pour illustrer ces définitions, nous considérons les évènements de 4 patients. Ici, le temps est divisé en estampilles temporelles correspondant à la survenue d'une hospitalisation. La figure 1 décrit les différents DP (Diagnostic Principal) correspondants aux motifs d'hospitalisation (I21 : *IM* aigu ; R07 : Douleurs thoraciques ; E14 : Diabète ; I20 : Angine de poitrine ; I70 : Athérosclérose) de ces patients.

Si $\epsilon = 2$ et $min_t = 2$, nous identifions les **essaims** suivants : $\{(P_1, P_3), (0, 1)\}$ et $\{(P_1, P_3), (2, 3)\}$. Les patients P_1, P_3 ont le même DP I21 en t_0, et le même DP E14 en t_1. Ils ont également le même DP I20, en t_2 et aussi en t_3 I70. Notons que, ces groupes sont redondants et peuvent être réunis dans l'essaim $\{(P_1, P_3), (0, 2, 3)\}$ qui est un **essaim clos**.



	Estampilles de temps			
Patient	t_0	t_1	t_2	t_3
P_1	I21	E14	I20	I70
P_2	I21	R07	I20	
P_3	I21	E14	I20	I70
P_4	I21	R07	I20	

FIG. 1 – *Trajectoires de patients. Le temps est lié à la survenue d'une hospitalisation (de gauche à droite la visualisation des trajectoires et la base séquentielle).*

3 Protocole de caractérisation des flux de patients

La première étape de notre protocole consiste à extraire puis à trier des motifs spatio-temporels à partir des données du PMSI dans le but de caractériser les schémas de flux de patients. L'étape suivante consiste à classer les trajectoires de délais inter-séjours et de tarifs de ces évènements dans le but d'identifier des profils de délais et d'évolution de tarifs. Ces étapes sont expliquées dans la suite de cette section et sont schématisées dans la figure 2.

3.1 Étape 1. Extraction de motifs spatio-temporels

Le processus de fouille, s'effectue de la façon suivante : **a-** Nous sélectionnons les patients ayant eu un *IM*, sur la période 2009 à 2014, à l'aide d'une requête sur la base PMSI. Puis, nous construisons la base séquentielle en nous intéressant aux DP. Chaque patient a une séquence de DP, de longueur égale à son nombre de séjours. Nous procédons à des prétraitements de ces séquences : 1) filtrage des évènements en ne retenant que ceux liés à la cardiopathologie ; 2) recalage des séquences selon un temps relatif correspondant à la survenue d'un évènement hospitalier. Les séquences de DP ainsi triées constituent **la trajectoire du patient**. **b-** Nous extrayons des essaims clos à l'aide de l'algorithme Get_Move (Phan et al., 2016). **c-** Nous représentons les parcours de patients, identifiés dans l'étape b, à l'aide d'un diagramme de Sankey [1] : 1) les sommets sont rassemblés selon une cohérence médicale Abitbol (2005) ; 2)

1. Il s'agit d'un diagramme de flux dans lequel la largeur des bandes est proportionnelle au nombre de patients considérés.

des groupes de flux sont créés selon le premier évènement de la trajectoire.

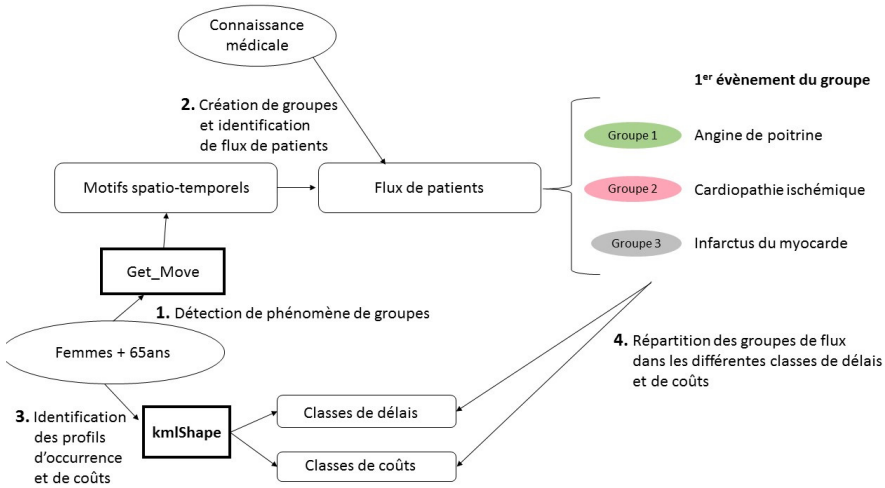


FIG. 2 – *Schéma du protocole de caractérisation des flux de patients.*

3.2 Étape 2. Identification de profils d'occurrence et d'évolution de tarifs

Le processus de classification, s'effectue en 2 étapes :

a. Le délai inter-séjours correspond au nombre de jours écoulés entre le dernier jour d'une hospitalisation et le premier jour de l'hospitalisation suivante. Le tarif d'un séjour a été déterminé selon les règles PMSI de valorisation des séjours.

b. Nous classons les trajectoires de délais puis celles de tarifs à l'aide de la méthode kmlShape (Genolini et al., 2016) [2]. Puis, nous établissons la répartition des groupes de patients, de l'étape 1, dans les classes de délais et de tarifs ainsi obtenues.

4 Expérimentations

Dans cette section, nous présentons les résultats dans le cas des femmes de +65 ans.

Étape 1. Dans la figure 3, nous identifions 3 flux de patients dont les évènements initiaux sont respectivement *Angine de poitrine*, *Cardiopathie ischémique* et *IM*. Les flux se séparent ensuite en plusieurs branches dont les évènements suivants sont ceux cités auxquels s'ajoute le *Décès*. À partir de la troisième hospitalisation apparaissent deux nouveaux évènements : les *Troubles du rythme* et l'*Insuffisance cardiaque*. Nous remarquons, par ailleurs, que les flux de patientes sont de plus en plus réduits au fur à mesure du temps.

2. Il s'agit d'une méthode dérivée de celle des k-moyennes pour données longitudinales (kml) qui prend en compte la forme des courbes, pour regrouper des courbes qui ont la même forme.

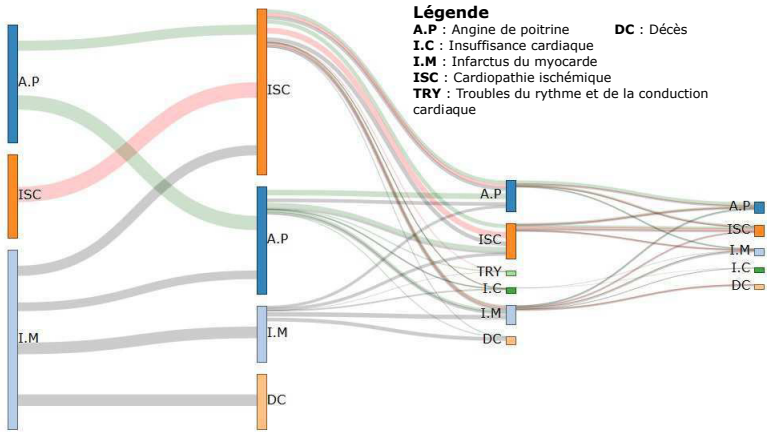


FIG. 3 – *Flux de patientes de +65 ans dans les trajectoires de DP dans un diagramme de Sankey.*

Étape 2. La classification des trajectoires de délais donne 3 classes (voir partie droite de la figure 4). Ces classes sont des courbes différenciées par les couleurs et l'épaisseur du trait : bleue et trait épais (classe B), magenta et trait fin (classe M) et verte avec trait très épais (classe V). Elles caractérisent des profils différents de délais de ré-hospitalisations. En effet, nous observons dans la figure 4, que les courbes ont des formes très différentes [3]. La classe B représente 3 070 patientes ayant des délais courts (inférieurs à 4 mois) qui augmentent puis diminuent. La classe M concerne 1 480 patientes ayant des évènements espacés en début de parcours puis plus rapprochés. La classe V caractérise 930 patientes ayant des délais courts en début de parcours, un évènement très éloigné en milieu de parcours et des délais courts à nouveau. La répartition des groupes de flux dans les classes de délais (voir tableau 1) montre qu'une majorité de la population a des délais qui diminuent en fin de parcours.

La classification des trajectoires de tarifs donne également 3 classes équi-réparties (voir partie gauche de la figure 4). La classe B représente 1 910 patientes ayant des séjours avec des tarifs augmentant au fur à mesure du parcours et diminuant légèrement. La classe M caractérise 1 840 patientes ayant des séjours avec des tarifs en forte diminution, puis ré-augmentant fortement en fin de parcours. Pour finir, la classe V caractérise 1 730 patientes avec des tarifs de séjours qui augmentent en début de parcours puis diminuent. La répartition des groupes de flux dans les classes de tarifs montre, dans le cas du groupe *IM*, qu'il y a une majorité de profils avec des tarifs élevés en début de parcours mais qui diminuent ensuite.

3. Le graphique se lit de la façon suivante, en t_0 est affiché le temps qui sépare la première hospitalisation de la seconde. Par exemple, la classe M, représentée par la courbe magenta (trait fin), a une ordonnée égale à 500 en t_0 : le délai entre la première hospitalisation et la seconde, pour les femmes de cette classe, est de 500 jours.

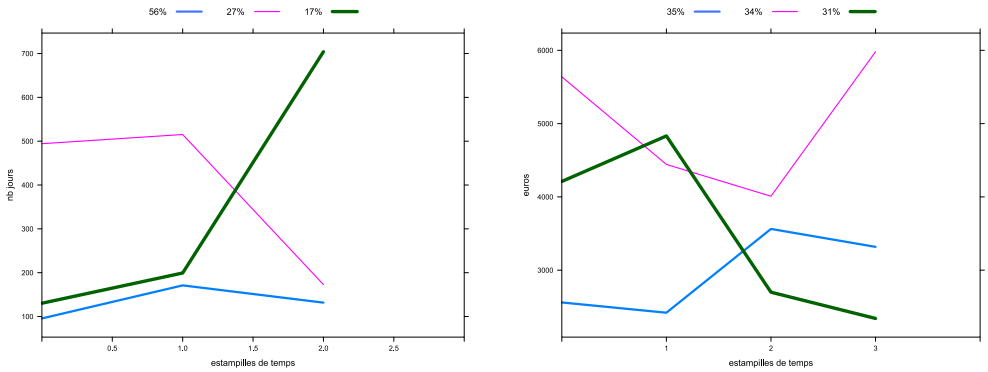


FIG. 4 – *Classes de délais (à gauche) et classes de tarifs (à droite) obtenues par kmlShape. Les courbes bleues, magentas et vertes correspondent resp. aux clusters B, M, V du tableau 1.*

Groupe		**Classes de délais**			**Classes de tarifs**		
Nom*	Effectif	*B*	*M*	*V*	*B*	*M*	*V*
Angine de poitrine	1 711	47	28	25	41	23	36
Cardiopathie ischémique	1 201	52	32	16	52	19	29
IM	2 651	45	34	21	25	36	39

*Le nom du groupe correspond au premier évènement du parcours de soins.

TAB. 1 – *Répartition (en pourcentages) des groupes de flux dans les classes de délais et de tarifs (B, M, V correspondent resp. aux courbes bleues, magentas et vertes de la figure 4).*

5 Discussion

Flux de patients. Nous avons identifié 3 étapes clés dans les schémas de flux de patients : l'*IM*, l'*Angine de poitrine* et la *Cardiopathie ischémique*. La majorité des patientes manifeste des signes de récidives de l'insuffisance coronarienne sous la forme d'une angine de poitrine. Beaucoup d'entre-elles font, par ailleurs, de nouveau un *IM* et/ou développent une cardiopathie ischémique. Nous retrouvons ici une partie des évolutions possibles de cette pathologie (Dujardin et Fabre, 2008). Les résultats obtenus par cette approche pourraient être intégrés dans un outil d'aide à la décision pour le clinicien. Il pourrait ainsi orienter ses recommandations et mettre en garde son patient sur ce qu'il risque, en comparant son profil à celui de patients similaires. Ce type d'application est par exemple utilisée dans le cas de la prescription d'anti-thrombotiques afin d'éviter des risques d'interactions médicamenteuses (Wang et Bajorek, 2016).

Profils de délais et de tarifs. Les profils de délais nous renseignent sur les occurrences des hospitalisations à venir, en lien avec la pathologie cardiaque. Dans la majeure partie des cas, à la suite d'un *IM*, les hospitalisations sont de plus en plus rapprochées dans le temps (3 mois en moyenne). Les flux de patientes ayant comme premier évènement *Angine de poitrine* ou *Cardiopathie ischémique* ont, majoritairement, une évolution des tarifs à la hausse. Toutefois, les tarifs moyens n'excèdent pas 4 000€. Au contraire, les flux de patientes initialisés par *IM*

ont, majoritairement, une évolution à la baisse de leurs tarifs avec un pic à presque 5 000€ en moyenne en début de parcours. En outre, la synthèse de ces travaux soulèvent des interrogations au sujet des rythmes de ré-hospitalisation et notamment sur le lien entre le rythme des occurrences d'hospitalisation et l'évolution des tarifs de séjour.

Limites. Nous pouvons formuler quelques critiques au regard de cette méthode : 1) *le choix de la base de données :* car le PMSI est un outil d'allocation budgétaire, il a donc des limites dans le domaine épidémiologique (Lombrail et al., 1994) ; 2) *la comparaison à d'autres études :* les travaux sur les délais sont nombreux (Miller et al., 2017), toutefois, il existe peu d'études sur les délais de la ré-hospitalisation en général, liée à un évènement de la cardiopathologie. Par ailleurs, dans la plupart des études de coûts directs (Blin et al., 2016), les frais d'hospitalisation ainsi que ceux de la consommation de médicaments sont pris en compte. L'accès aux bases du Sniiram (Système national d'information inter-régimes de l'assurance maladie) permettrait de mener des analyses similaires.

6 Conclusion

Nous avons recherché des motifs spatio-temporels dans les trajectoires des patientes ayant eu un *IM*. L'originalité de l'approche est qu'ici la spatialité a été assimilée à celle de proximité de pathologies WHO (2004). De plus, l'aspect temporel est lié à la survenue d'une hospitalisation. Ensuite, nous avons classé leurs trajectoires de délais et de tarifs afin de déterminer des profils types en termes d'occurrence et de facturation de ces évènements. La visualisation des schémas de flux nous renseigne sur les évolutions de la pathologie cardiaque. Nous avons retrouvé des éléments connus du domaine de la cardiologie. Ce type d'approche peut permettre d'améliorer la prise en charge en prévoyant, par exemple, un parcours de soins avec des visites programmées comme cela a été fait dans le cas du cancer (Defossez et al., 2014). Dans la suite de ces travaux, nous envisageons de prendre en compte la co-évolution des trajectoires de délais et de tarifs avec la méthode proposée par Genolini et al. (2013) afin d'établir si elles sont liées.

Références

Abitbol, E. (2005). Complications mécaniques de l'infarctus du myocarde récent. *Médecine Thérapeutique Cardio 1*(6), 559–569.

Blin, P., C. Dureau-Pournin, R. Lassalle, J. Jové, F. Thomas-Delecourt, C. Droz-Perroteau, N. Danchin, et N. Moore (2016). Outcomes, health care resources use, and costs in patients with post-myocardial infarction : The horus cohort study in the egb french claims and hospital database. *Value in Health 19*(3), A16.

Dart, T., Y. Cui, G. Chatellier, et P. Degoulet (2003). Analysis of hospitalised patient flows using data-mining. *Studies in Health Technology and Informatics 95*, 263–268.

Defossez, G., A. Rollet, O. Dameron, et P. Ingrand (2014). Temporal representation of care trajectories of cancer patients using data from a regional information system : an application in breast cancer. *BioMed Central Medical Informatics and Decision Making 14*(1), 24.

Dégano, I. R., V. Salomaa, G. Veronesi, J. Ferriéres, I. Kirchberger, T. Laks, A. S. Havulinna, J.-B. Ruidavets, M. M. Ferrario, C. Meisinger, et al. (2015). Twenty-five-year trends in myocardial infarction attack and mortality rates, and case-fatality, in six european populations. *Heart 101*(17), 1413–1421.

Dujardin, J.-J. et O. Fabre (2008). Complications de l'infarctus du myocarde. Évolution et pronostic. *Cardiologie 3*(1), 1–13.

Genolini, C., R. Ecochard, M. Benghezal, T. Driss, S. Andrieu, et F. Subtil (2016). kmlShape : An Efficient Method to Cluster Longitudinal Data (Time-Series) According to Their Shapes. *Plos one 11*(6), e0150738.

Genolini, C., J.-B. Pingault, T. Driss, S. Côté, R. E. Tremblay, F. Vitaro, C. Arnaud, et B. Falissard (2013). Kml3d : a non-parametric algorithm for clustering joint trajectories. *Computer Methods and Programs in Biomedicine 109*(1), 104–111.

Lombrail, P., E. Minvielle, L. Comar, et S. Gottot (1994). Programme de médicalisation des systèmes d'information et épidémiologie : une liaison qui ne va pas de soi. *Revue d'épidémiologie et de santé publique 42*(4), 334–344.

Miller, A. L., D. Simon, M. T. Roe, M. C. Kontos, D. Diercks, E. Amsterdam, et D. L. Bhatt (2017). Comparison of Delay Times from Symptom Onset to Medical Contact in Blacks Versus Whites With Acute Myocardial Infarction. *The American Journal of Cardiology 119*(8), 1127–1134.

Pagnoni, A., S. Parisi, et S. Lombardo (2001). Analysis of patient flows via data mining. *Studies in Health Technology and Informatics* (2), 1379–1383.

Phan, N., P. Poncelet, et M. Teisseire (2016). All in one : mining multiple movement patterns. *International Journal of Information Technology & Decision Making 15*(05), 1115–1156.

Philippe, F., P. Blin, S. Bouée, C. Laurendeau, E. Torreton, J. Gourmelin, M. Velkovski-Rouyer, L. Levy-Bachelot, et G. Steg (2017). Coût des consommations de soins après un infarctus du myocarde en france : une estimation à partir de l'échantillon généraliste des bénéficiaires (egb). In *Annales de Cardiologie et d'Angéiologie*, Volume 66, pp. 74–80.

Wang, Y. et B. Bajorek (2016). Selecting antithrombotic therapy for stroke prevention in atrial fibrillation : Health professionals' feedback on a decision support tool. *Health Informatics Journal*, 1460458216675498.

WHO (2004). *International statistical classification of diseases and related health problems*, Volume 1.

Summary

A better knowledge of patient flows can be decisive for health planning. In this article, we propose a method to characterise patients flow but also to highlight profiles of care pathways times and costs. From the medical-administrative data we extracted spatio-temporal patterns and integrated them into a visualization tool. Then, we clustered time between hospitalisation and cost trajectories in order to identify profiles profiles of change over time. This approach may support renewed management strategies.