HAL open science

Découverte d'associations entre Evénements Indésirables Médicamenteux par les structures de patrons et les ontologies

Gabin Personeni, Emmanuel Bresso, Marie-Dominique Devignes, Michel Dumontier, Malika Smaïl-Tabbone, Adrien Coulet

To cite this version:

Gabin Personeni, Emmanuel Bresso, Marie-Dominique Devignes, Michel Dumontier, Malika Smaïl-Tabbone, et al.. Découverte d'associations entre Evénements Indésirables Médicamenteux par les structures de patrons et les ontologies. Journée I.A. et Santé, Jul 2018, Nancy, France. hal-01872312

HAL Id: hal-01872312

https://inria.hal.science/hal-01872312v1

Submitted on 11 Sep 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.

Découverte d'associations entre Evénements Indésirables Médicamenteux par les structures de patrons et les ontologies

Gabin Personeni[1], Emmanuel Bresso[1], Marie-Dominique Devignes[1], Michel Dumontier[2,3], Malika Smaïl-Tabbone[1], Adrien Coulet[1,3]

[1] Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
[2] Institute of Data Science, Maastricht University, MD Maastricht, Netherlands
[3] Stanford Center for Biomedical Informatics Research, Stanford, USA

Résumé : On présente dans cet article une méthode d'extraction d'associations entre événements indésirables médicamenteux (EIM) utilisant les structures de patrons. Cette méthode permet une comparaison d'EIM fondée sur des ontologies biomédicales et a déjà été présentée avec une application sur des dossiers médicaux électroniques. Une application sur un autre type de données utilisant d'autres ontologies biomédicales est présentée ici, ainsi qu'une comparaison et une évaluation plus détaillée des règles extraites. Cette méthode se révèle flexible puisque pouvant être appliquée à divers jeux de données et ontologies, et capable d'extraire des règles d'associations entre EIM avec une représentation expressive de ces EIM.

Mots-clés : Ontologies, Evénements Indésirables Médicamenteux, Analyse Formelle de Concepts, Structures de Patrons

1 Introduction

De nombreux travaux s'intéressent à la détection des Evénements Indésirables Médicamenteux (EIM) signalés après la commercialisation d'un médicament Harpaz *et al.* (2012). De tels événements peuvent être enregistrés dans des Dossier Médicaux Electroniques (DME) ou des systèmes de rapport d'EIM. Ces données peuvent alors être étudiées afin de démontrer ou non l'association entre un médicament et un effet indésirable. Ces EIM peuvent se manifester différemment, ou pas du tout, chez divers patients, et peuvent être causés par une combinaison de médicaments pris ensemble. Il peut également être intéressant d'étudier ces EIM au niveau d'une classe de médicaments ou de phénotypes, notamment à l'aide des ontologies biomédicales. Dans ce travail nous étudierons les associations possibles entre EIM.

Roitmann *et al.* (2014) proposent une représentation vectorielle des profils médicaments/effets indésirables de patients . Chaque patient est représenté par un vecteur où chaque dimension est un effet indésirable déclaré chez le patient. Cette représentation est utilisée par des algorithmes de clustering afin de former des groupes de patients ayant un profil similaire. Dans chacun de ces clusters sont ensuite étudiés les symptômes et médicaments les plus représentés. Cette méthode permet la découverte d'associations entre un ou plusieurs médicaments et un effet indésirable, ainsi que d'observer dans des clusters particuliers des phénotypes souvent associés. Cette méthode d'apprentissage non supervisé permet d'extraire des associations sans hypothèse *a priori*, mais ne prend pas en considération les similarités entre phénotypes ou entre médicaments. En particulier, lorsque plusieurs médicaments sont fortement représentés dans un cluster, cette méthode ne permet pas à elle seule de déterminer si c'est la combinaison de ces médicaments qui est responsable d'un effet indésirable, ou si ces médicaments sont associés car ils causent le même effet secondaire.

Winnenburg *et al.* (2015) utilisent des EIM extraits de la littérature sous forme de paires médicament-phénotype pour explorer les relations entre médicaments, classes de médicaments et leur effets secondaires. Dans cette étude, la prévalence des effets secondaires est considérée tant au niveau d'un médicament que de sa classe ATC. Les auteurs mettent alors en évidence des associations entre une classe de médicament et des effets secondaires, et poursuivent leur investigation au niveau des médicaments de cette classe. Si l'association existe pour chacun des médicaments de la classe, alors cela permet d'imputer l'effet secondaire à la classe de médicaments elle-même, et on peut alors parler d'effet de classe.

L'Analyse Formelle de Concepts est une méthode pouvant être utilisée pour la détection de signaux pour la pharmacovigilance, par la mise en évidence d'attributs décrivant les patients

fréquemment associés. Villerd *et al.* (2010) utilisent un sous-ensemble des rapports d'un système d'auto-déclaration d'effets secondaires pour construire un contexte formel binaire dont les attributs sont des médicaments, phénotypes et données démographiques. Les auteurs identifient ensuite des concepts dans le treillis généré possédant comme attributs au moins un médicament et un phénotype, et vérifient la significativité statistique de leur association dans la population décrite par les attributs démographiques du concept. Cette approche permet alors d'extraire des associations entre un ensemble de médicaments et un phénotype.

Nous avons présenté une méthode permettant l'extraction d'associations entre deux EIM, fondée sur une extension de l'analyse formelle de concepts : les structures de patrons, et des ontologies biomédicales (Personeni *et al.* (2016)). Les règles d'associations extraites permettent de mettre en évidence des EIM fréquemment associés chez des patients, c'est-à-dire, où un premier EIM serait indicatif d'une susceptibilité à un second EIM causé par un autre médicament. Cette approche a déjà été appliquée à un ensemble de Dossiers Médicaux Electroniques (DME) de l'entrepôt STRIDE (Lowe *et al.* (2009)). Il s'agissait alors d'un ensemble de patients diagnostiqués avec un lupus érythémateux disséminé. Dans ce nouvel article, notre méthode est appliquée au système de rapport d'EIM américain FAERS [1] afin de démontrer sa flexibilité. De plus, une évaluation plus détaillée est présentée (Personeni *et al.* (2017)).

2 Méthodologie

2.1 Représentation des EIM

Un EIM est un évènement complexe qui peut impliquer plusieurs médicaments, et se manifester sous la forme de plusieurs phénotypes. Un EIM peut ainsi être caractérisé par un ensemble de médicaments et un ensemble de phénotypes. Afin de faciliter la comparaison entre des EIM, on considérera des ensembles d'ingrédients actifs de médicaments, plutôt que des ensembles de formes commerciales de médicaments. Ainsi, on utilisera ici le terme "médicament" pour désigner un ingrédient actif. Dans cet article, on représentera un EIM comme une paire (D_i, P_i), où D_i est un ensemble de médicaments, et P_i est un ensemble de phénotypes.

2.2 Analyse Formelle de Concepts et structures de patrons

L'Analyse Formelle de Concepts (ou *FCA – Formal Concept Analysis*) (Ganter & Wille (1997)) est un cadre mathématique pouvant être utilisé pour l'analyse de données et la découverte de connaissances. La FCA utilise des données représentées sous la forme d'un contexte formel, où des objets sont associés à des attributs par une relation binaire. Un contexte formel est un triplet (G, M, I) où G est un ensemble d'objets, M est un ensemble d'attributs et I est une relation binaire entre les objets de G et les attributs de M.

Les structures de patrons (Ganter & Kuznetsov (2001)) généralisent la FCA de manière à pouvoir l'appliquer à des objets dotés de descriptions non seulement binaires, mais de nature plus complexe, comme par exemple des ensembles, graphes, intervalles ou des données annotées par des termes liés à des ontologies. Une structure de patrons est un triplet $(G, (\mathcal{D}, \sqcap), \delta)$ où :

- G est un ensemble d'objets
- $\mathcal{D}$ est l'ensemble des descriptions associées à ces objets,
- δ est une fonction qui associe une description à un objet,
- $\sqcap$ est un opérateur de comparaison de descriptions, tel que pour deux descriptions X et Y dans $\mathcal{D}$, $X \sqcap Y$ est la similarité entre les objets associés aux descriptions X et Y. Cet opérateur définit un ordre partiel $\leq_\sqcap$ sur les éléments de $\mathcal{D}$, tel que $X \leq_\sqcap Y \Leftrightarrow X \sqcap Y = X$.

Il existe, entre les objets G et les descriptions $\mathcal{D}$ une correspondance de Galois. Cette correspondance est établie via deux opérateurs de dérivation qui associent à un ensemble d'objets la description correspondante, et inversement. Pour une structure de patrons $(G, (\mathcal{D}, \sqcap), \delta)$, l'opérateur de dérivation $.^\square$ définit une correspondance de Galois entre $(2^G, \subseteq)$ et $(\mathcal{D}, \leq_\sqcap)$,

1. FDA (U.S. Food & Drug Administration) Adverse Event Reporting System, `http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/`

telle que :

$$A^{\square} = \bigsqcap_{g \in A} \delta(g) \text{ pour tout } A \subseteq G \qquad d^{\square} = \{g \in G \mid d \leq_{\sqcap} \delta(g)\} \text{ pour tout } d \in \mathcal{D}$$

Intuitivement, $A^{\square}$ est la description la plus précise (maximale selon $\leq_{\sqcap}$) pour l'ensemble d'objets A, et $d^{\square}$ est l'ensemble de tous les objets dont la description est plus spécifique ou égale à d.

Une structure de patrons peut être représentée sous la forme d'un treillis de Galois de concepts formels, c'est-à-dire une structure hiérarchique dans laquelle chaque nœud représente un concept formel, c'est-à-dire un ensemble d'objets partageant une description commune. Un concept formel est un couple (A, B), où A est un ensemble d'objets $A \subseteq G$, B est une description $B \in \mathcal{D}$, et $A^{\square} = B$; $B^{\square} = A$. On appelle alors A l'extension du concept et B son intention. Les concepts formels sont ordonnés par un ordre partiel défini sur l'inclusion de leur extensions, ou de façon duale sur l'ordre $\leq_{\sqcap}$ de leur intentions, tel que $(A_1, B_1) \leq (A_2, B_2)$ si et seulement si $A_1 \subseteq A_2$ et $B_2 \leq_{\sqcap} B_1$.

Luxenburger (1991) propose une méthode simple pour l'extraction de règles d'associations à partir d'un tel treillis de concepts. Soient deux concepts (A', A) et (B', B) directement liés dans un treillis de concepts tels que $(B', B) \leq (A', A)$, on peut extraire la règle d'association $A \rightarrow B \setminus A$ de support $|B'|$ et de confiance $|B'|/|A'|$.

3 Expériences

On présente ici une nouvelle expérience utilisant les structures de patrons pour extraire des règles d'association entre EIM, à partir du système de rapport d'EIM appelé FAERS.

3.1 Extraction des données

FAERS est une base de données comportant les EIM reportés par des patients, professionnels de la santé et laboratoires pharmaceutiques aux Etats-Unis d'Amérique. Une ressource récente, AEOLUS (Banda *et al.* (2016)) propose des outils pour mettre en correspondance les médicaments de FAERS avec les vocabulaire RxNorm et les phénotypes de FAERS avec le vocabulaire SNOMED CT. Nous avons utilisé ces outils pour reconstruire une base de données de rapports FAERS sur la période 2012 quatrième trimestre à 2016 second trimestre inclus, liée à RxNorm et SNOMED CT.

Chaque rapport d'EIM dans FAERS liste l'ensemble des médicaments pris avant l'EIM D_i et la liste de phénotypes caractérisant cet EIM P_i. Ainsi chaque rapport d'EIM peut être formalisé sous la forme (D_i, P_i). Ces rapports d'EIM sont groupés par "cas", un cas se constituant de plusieurs rapports d'EIM d'un même patient. Nous avons sélectionné dans la base de données FAERS les cas présentant plusieurs EIM, en excluant les EIM où l'ensemble des médicaments est déjà inclus dans un autre EIM. On conserve uniquement les patients avec 2 ou plus EIM, puisque l'on s'intéresse à des associations entre EIM. Avec ces contraintes, nous avons pu extraire 570 cas de patients avec un total de 1148 EIM.

3.2 Opérateur de comparaison

On définit ici une structure de patrons $(G, (\mathcal{D}, \sqcap), \delta)$ où : G est un ensemble de patients, $\mathcal{D}$ est un ensemble de descriptions représentant chacune un ensemble d'EIM, et δ est la fonction associant à un patient la description de ses EIM. On définira cette structure de patrons de manière à permettre l'utilisation de l'ontologie ATC pour la décrire les médicaments ainsi qu'une ontologie de phénotypes pour décrire les effets secondaires.

Dans notre cas, afin d'éviter une sur-généralisation des descriptions d'EIM, on exclut les 3 niveaux les plus généraux de SNOMED CT. La Table 1 présente un exemple de la représentation de données utilisées avec la structure de patrons décrite ici, exprimé avec les ontologies ATC et ICD.

TABLE *1: Exemple de représentation d'EIM pour 3 patients*

Patient	Description
P1	{⟨{H02AB07},{ICD 599.8}⟩, ⟨{N02BE01},{ICD 599.9}⟩}
P2	{⟨{H02AB07},{ICD 599.9}⟩, ⟨{H02AB07},{ICD 719.4}⟩}
P3	{⟨{H02AB07, N02BE01},{ICD 599.9}⟩, ⟨{N02BE01},{ICD 719.4}⟩}

Ici, un EIM est représenté comme un vecteur $\langle D_i, P_i \rangle$ à deux dimensions : un ensemble de médicaments D_i pour la première, associé à un ensemble de phénotypes P_i pour la deuxième. La description d'un patient est alors un ensemble de tels vecteurs. On définit d'abord un opérateur permettant la comparaison de deux ensembles de classes d'une ontologie $\sqcap_{\mathcal{O}}$ tel que, pour x et y deux ensembles de classes d'ontologies :

$$x \sqcap_{\mathcal{O}} y = \max(\sqsubseteq, \{\text{LCA}(c_x, c_y) \mid (c_x, c_y) \in x \times y\})$$

où $\max(\leq_i, S)$ est l'unique ensemble des élements maximaux appartenant à un ensemble S pour un ordre partiel $\leq_i$, tel que $\max(\leq_i, S) = \{s \mid \nexists x.(s \leq_i x)\}$.

On définit ensuite un opérateur permettant la comparaison deux à deux d'EIM $\sqcap_{EIM}$ tel que :

$$\langle D_x, P_x \rangle \sqcap_{EIM} \langle D_y, P_y \rangle = \begin{cases} \langle D_x \sqcap_{\mathcal{O}} D_y, P_x \sqcap_{\mathcal{O}} P_y \rangle & \text{si les deux dimensions sont} \neq \varnothing \\ \langle \varnothing, \varnothing \rangle & \text{sinon.} \end{cases}$$

L'opérateur $\sqcap_{EIM}$ applique l'opérateur de comparaison $\sqcap_{\mathcal{O}}$ sur les deux dimensions du vecteur représentant l'EIM, en utilisant l'ontologie correspondant aux données. Si au moins une des deux dimensions du résultat devait être égale à $\varnothing$, alors le résultat est remplacé par $\langle \varnothing, \varnothing \rangle$ afin de l'ignorer dans les prochaines généralisations. En effet, on ne veut pas considérer la similarité de deux EIM si ils n'ont aucun médicament ou phénotype commun.

Finalement on définit l'opérateur de comparaison de notre structure de patron $\sqcap$ tel que, pour toute paire de descriptions (X, Y) :

$$X \sqcap Y = \max(\leq_{EIM}, \{v_x \sqcap_{EIM} v_y \mid (v_x, v_y) \in X \times Y\})$$

Cet opérateur peut s'appliquer aux données extraites des DME ou de FAERS, exprimées à l'aide de classes provenant de différentes ontologies de phénotypes. Cet opérateur est ici décrit de manière générique, et est utilisable avec n'importe quelle ontologie.

4 Résultats

4.1 Comparaison entre STRIDE et FAERS

La structure de patrons présentée dans la section précédente permet de construire un treillis de concepts duquel on extrait des règles d'associations, avec un support et une confiance d'au moins 5 et 0.75, respectivement. La Table 2 présente quelques statistiques sur ce processus.

TABLE *2: Statistiques sur le processus de FCA et d'extraction de règles d'associations.*

Données	STRIDE (DME)	FAERS
#patients	548	570
#EIM	3,286	1,148
#concepts dans le treillis	≈2.5 millions	≈22,700
#règles extraites	≈9 millions	≈18,500
#règles conservées après filtrage	913	493
#règles avec un support > 8	15	151
Support maximal	10	27

On observe d'abord que l'expérience sur FAERS produit un nombre de règles beaucoup moins important qu'avec les DME de STRIDE. Cependant, comme dans l'expérience sur les DME, beaucoup de ces règles contiennent des associations triviales, notamment lorsque le ou les EIM de la partie droite ne sont qu'une spécialisation de celui ou ceux de la partie gauche (par exemple l'association $\{\langle\{d_1\}, \{p_1\}\rangle\} \rightarrow \{\langle\{d_1, d_2\}, \{p_1\}\rangle\}$ semble triviale dans le cas de l'étude des EIM). On ne conserve alors que les règles possédant dans leur partie droite un EIM (D_R, P_R) telle que il n'existe aucun EIM (D_L, P_L) dans la partie gauche de la règle tel que D_R et D_L ou P_R et P_L ne soient comparable par $\leq_{\mathcal{O}}$.

On observe également que l'expérience utilisant les données de FAERS génère un treillis de concepts beaucoup plus petit que celle utilisant les données de DME (100 fois moins de concepts), et ce malgré un nombre de patients comparable. Cependant après filtrage des règles, leur nombre est dans le même ordre de grandeur (seulement 2 fois moins). Cette différence peut être expliquée par les différences entre les deux jeux de données : le jeu de

données de DME ne concerne que des patients atteints de lupus érythémateux disséminé, tandis que le jeu de données de FAERS concerne une population plus générale. De plus, le plus grand nombre d'EIM extraits des DME tends à augmenter la similarité entre patients, augmentant alors la taille du treillis de concepts généré. L'ensemble de règles filtrées extraites de chaque expérience est disponible à l'adresse : `https://github.com/g-a-perso/ADE-associations/`.

4.2 Analyse statistique des associations entre EIM

Pour chaque paire de classes de médicaments ATC (l, r), on cherche l'ensembles des règles extraites des DME et de FAERS de la forme $L \rightarrow R$ telles que l (ou un de ses descendants) apparaît dans L et r (ou un de ses descendants) apparaît dans R. On calcule ensuite le support de cet ensemble de règles comme étant le nombre de patients vérifiant au moins une de ces règles. On calcule ensuite pour chaque (l, r) le ratio entre *(i)* le support des règles telles que l apparaît dans L et r apparaît dans R, *(ii)* le support des règles telles que l apparaît dans L. Ce ratio exprime la fréquence à laquelle les règles associent un EIM impliquant un médicament de la classe r à un EIM impliquant un médicament de la classe l.

On utilise un Z-test pour évaluer la significativité de l'écart obtenu à l'écart attendu si les associations entre EIM étaient extraites aléatoirement. On observe notamment quelques associations d'intérêt entre classes ATC (avec $p < 0.001$). Par exemple, on constate parmi les règles extraites des DME, que les EIM impliquant des agents bêta-bloquants (classe ATC C07A) sont fortement associés à des EIM impliquant des diurétiques de l'anse (classe ATC C03C). Ces deux classes de médicaments sont impliquées dans des thérapies contre l'hypertension, parfois en combinaison, ce qui peut expliquer la forte association entre des EIM causés par ces deux types de médicaments. On observe également que les EIM impliquant des agents antithrombotiques (classe ATC B01A) sont associés à d'autres EIM impliquant d'autres médicaments de la même classe. Ainsi, il semble que l'approche proposée ici permet de révéler des associations significatives entres EIM causés par des médicaments d'une même classe ou de plusieurs classes différentes.

4.3 Evaluation des règles extraites sur STRIDE

Le jeu de données extrait des DME de STRIDE représente une petite partie de la base de données STRIDE qui contient environ 2 millions de DME. Pour les 15 règles ayant le plus grand support (entre 8 et 10) parmi les règles extraites des DME, on a calculé le support sur la totalité de la base de données STRIDE des pour évaluer la généralité de ces règles. Chacune de ces règles a été transformée en une requête SQL permettant d'identifier les DME de STRIDE vérifiant chaque règle. On observe dans que pour ces règles, le nombre de DME les vérifiant varie entre 33 à 326 pour une moyenne de 86.2 Cela illustre que des règles d'association extraites à partir de DME de patients atteints de lupus érythémateux disséminé peuvent être pertinentes en dehors du jeu de données initial.

5 Discussion et conclusion

Les structures de patrons permettent de fouiller des descriptions détaillées d'EIM. Cependant il est nécessaire de considérer la complexité des algorithmes de FCA combinés avec les opérateurs de comparaison proposés. Les ontologies biomédicales permettent une meilleure comparaison des EIM d'un patient, notamment en étant capable de généraliser plusieurs EIM causés par des médicaments d'une même classe. Cela augmente néanmoins le nombre de concepts générés, puisque l'on augmente les possibilités de trouver une description générale commune à de nombreux patients. On observe une grande différence entre les résultats obtenus par les expériences sur les DME de STRIDE et sur les rapports de FAERS. Cela peut s'expliquer par la différence de nature des deux jeux de données : le jeux de données extrait de FAERS est constitué d'EIM rapportés d'abord par des patients, tandis que le jeu de données provenant des DME est constitué d'EIM automatiquement extraits de ces dossiers. Le jeu de données issu de STRIDE ne concerne que des patients diagnostiqués avec un lupus érythémateux disséminé, tandis que FAERS concerne la population générale, avec nécessairement une plus grande diversité des traitements et des phénotypes.

Interprétation des associations entres EIM Une limitation de la représentation des EIM décrite dans cette étude est l'absence de temporalité. Il est par exemple possible que les EIM

présents dans la partie gauche d'une règle d'association se produisent avant les EIM de la partie droite. Ainsi les règles d'associations n'expriment qu'une association fréquente d'EIM chez les patients, sans information de temporalité. Dans les expériences utilisant les données extraites des DME, seuls les EIM exprimés dans une fenêtre de 14 jours après la prescription ont été considérés. Ainsi, seuls des EIM à court terme ont été considérés. La représentation des EIM pourrait être enrichie avec une information sur le délai entre la prescription et l'apparition du phénotype indésirable. Cela permettrait d'extraire des associations dans un jeu de données d'EIM à court et long terme, tout en étant capable de discriminer au besoin entre ces différentes manifestations. Cela permettrait par exemple l'extraction d'associations entre un EIM à court terme et un EIM à long terme : la toxicité d'un médicament à court terme pourrait dans ce cas être utilisée comme un prédicteur de la toxicité à long terme d'un autre. Une autre limitation de l'extraction de règles d'association est que les règles extraites n'expriment pas une relation causale entre les EIM associés. En effet, il semble plus approprié de rechercher une cause biologique commune à deux EIM associés par une règle, que de chercher une relation causale directe entre ces deux EIM.

Conclusion Cet article développe et complémente une approche fondée sur les structures de patrons pour l'extraction d'EIM fréquemment associés dans des DME et dans FAERS. Une grande quantité de règles d'association peut être extraite des treillis de concepts générés. Les règles sont filtrées automatiquement de manière à exclure les règles qui ne répondent pas à la problématique étudiée. L'approche proposée est flexible et applicable à différents jeux de données, il est toutefois difficile de comparer des règles d'association extraites de jeux de données différents décrits par différentes ontologies. On a pu tester les règles sélectionnées obtenues sur le jeu de données des DME de STRIDE sur la base de données STRIDE entière. Ces tests indiquent que les règles extraites d'un sous-ensemble de DME de patients atteint de lupus érythémateux disséminé peuvent être vérifiées dans un plus grand ensemble de patients. Une autre perspective de ce travail consisterait à classer ces règles d'association par rapport à leur importance en terme de risque et de coût des EIM décrits.

Références

BANDA J. M., EVANS L., VANGURI R. S., TATONETTI N. P., RYAN P. B. & SHAH N. H. (2016). A curated and standardized adverse drug event resource to accelerate drug safety research. *Scientific data*, **3**.

GANTER B. & KUZNETSOV S. O. (2001). Pattern structures and their projections. In *International Conference on Conceptual Structures*, p. 129–142 : Springer.

GANTER B. & WILLE R. (1997). *Formal Concept Analysis : Mathematical Foundations*. Secaucus, NJ, USA : Springer-Verlag New York, Inc., 1st edition.

HARPAZ R., DUMOUCHEL W., SHAH N. H., MADIGAN D., RYAN P. & FRIEDMAN C. (2012). Novel data-mining methodologies for adverse drug event discovery and analysis. *Clinical Pharmacology & Therapeutics*, **91**(6), 1010–1021.

LOWE H. J., FERRIS T. A., HERNANDEZ P. M., WEBER S. C. *et al.* (2009). STRIDE-an integrated standards-based translational research informatics platform. In *AMIA*.

LUXENBURGER M. (1991). Implications partielles dans un contexte. *Mathématiques, informatique et sciences humaines*, **29**(113), 35–55.

PERSONENI G., BRESSO E., DEVIGNES M.-D., DUMONTIER M., SMAÏL-TABBONE M. & COULET A. (2017). Discovering associations between adverse drug events using pattern structures and ontologies. *Journal of biomedical semantics*, **8**(1), 29.

PERSONENI G., DEVIGNES M.-D., DUMONTIER M., SMAÏL-TABBONE M. & COULET A. (2016). Extraction d'association d'eim à partir de dossiers patients : expérimentation avec les structures de patrons et les ontologies. In *Deuxième Atelier sur l'Intelligence Artificielle et la Santé*.

ROITMANN E., ERIKSSON R. & BRUNAK S. (2014). Patient stratification and identification of adverse event correlations in the space of 1190 drug related adverse events. *Frontiers in physiology*.

VILLERD J., TOUSSAINT Y. & LILLO-LE LOUËT A. (2010). Adverse drug reaction mining in pharmacovigilance data using formal concept analysis. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, p. 386–401 : Springer.

WINNENBURG R., SORBELLO A. & BODENREIDER O. (2015). Exploring adverse drug events at the class level. *Journal of Biomedical Semantics*, **6**(1), 18.