Mining Safety Signals in Spontaneous Report Database using Concept Analysis

Amine Mohamed Rouane Hacene, Yannick Toussaint, Petko Valtchev

Mining Safety Signals in Spontaneous Reports Database using Concept Analysis

Mohamed Rouane-Hacene[1], Yannick Toussaint[2], and Petko Valtchev[1]

[1] Dépt. Informatique, UQÀM, CP 8888, succ. CV, Montréal, Canada, H3C 3P8
[2] LORIA, B.P. 239, F-54506 Vandœuvre-lès-Nancy, France

Abstract. In pharmacovigilance, linking the adverse reactions by patients to drugs they took is a key activity typically based on the analysis of patient reports. Yet generating potentially interesting pairs (drug, reaction) from a record database is a complex task, especially when many drugs are involved. To limit the generation effort, we exploit the frequently occurring patterns in the database and form *association rules* on top of them. Moreover, only rules of minimal premise are considered as output by concept analysis tools, which are then filtered through standard measures for statistical significance. We illustrate the process on a small database of anti-HIV drugs involved in the HAART therapy while larger-scope validation within the database of the French Medicines Agency is also reported.

1 Introduction

Pharmacovigilance (PV) aims at, first, studying and, then, preventing the adverse reactions to drugs (ADR) based on the data collected by spontaneous reporting systems (SRS) and stored in case report databases (DB). SRS DB comprises a collection of reports each capturing the patient characteristics including demographic data (age, race, gender, etc.), the suspected drugs and a description of the observed ADR. Table 1 depicts a set of case reports on AIDS patients and antiretroviral drugs, i.e., treating infection by retroviruses such as HIV.

In PV, drug-reaction combinations, known as *safety signals*, help devising a drug therapy, hence the importance of their detection. For instance, in HIV treatment, the caregivers are interested in the response of various classes of patients to the HAART therapy in order to adapt the overall anti-HIV therapy. Their prime target is an appropriate combination of antiretroviral drugs that, while effective, limits the ADR: e.g., older patients with HIV infection have robust responses to HAART with no increased risk of metabolic disorders or other ADR. Beside safety signals, i.e., (drug, ADR) pairs, further meaningful combinations from the SRS DB involve several drugs for a single ADR. These are potential *drug interactions* (higher-order signals).

Signal detection has been approached with a variety of analysis tools [7] including statistical methods for disproportionality assessment, deviation detection, etc. However, none of these proposes a way, both automated and feasible, for

generating all potential signals from the SRS DB. Moreover, even with an expert-provided potential signal, the underlying approaches would consider all drug-reaction combinations that can be derived from the signal, including many spurious ones. For instance, consider the anti-HIV drugs `Lopinavir` and `Tenofovir` in Table 1 and the ADR `HairLoss` and `Oedema`. A proportionate approach would suggest the study of signals (`Lopinavir`, `Oedema`), (`Lopinavir`, `HairLoss`), (`Tenofovir`, `Oedema`), and (`Tenofovir`, `HairLoss`). Yet the only sensible combination to study is ({`Lopinavir`, `Tenofovir`}, {`HairLoss`, `Oedema`}) as, given the dataset, the four combine to a *maximal* pattern. In summary, because of the large size of most SRS DB, the computation of all combinations is strongly combinatorial, hence their test may prove infeasible. Instead, a more careful approach would track the frequently occurring patterns in the records and use these as prototypes.

Patient	Age	Gender	Prescribed drugs	Observed adverse drug reactions
Daffy	24	`Female`	`Lopinavir`, `Efavirenz`	`Nausea`, `Hives` , `Vomiting`
Farley	63	`Male`	`Lopinavir`, `Tenofovir`	`Oedema`, `Hives`, `Headache`, `Nausea`, `Heart failure`, `Hair loss`
Lane	27	`Female`	`Maraviroc`, `Efavirenz`	`Fatigue`, `Oedema`, `Hives`, `Hair loss`, `Bleeding`
Shana	15	`Female`	`Tenofovir`, `Lopinavir`	`Fatigue`, `Oedema`, `Hair loss`
Trudy	41	`Male`	`Raltegravir`	`Fatigue`, `Breath disorder`, `Nausea`, `Heart failure`, `Bleeding`, `Vomiting`

Table 1. A fragment[1] of SRS DB.

Patterns comprised of two sets, a premise and a conclusion, called *associations*, have been successfully applied to a variety of practical problems involving co-occurrencies of phenomena and seem to fit well the PV context. Yet a notorious problem of association miners is the huge number of potentially useful associations that may be extracted from even a small DB. Formal concept analysis (FCA) [6] provides the theoretical foundation for association rule bases that only withhold a tiny proportion of all valid associations while keeping the total of the information. Hence we propose an FCA-based method for signal detection which, by examining a minimal set of association rules extracted from the SRS DB helps minimize the number of (drug, ADR) pairs to be statistically analyzed.

Here, we examine the detection of safety signals and drug-drug interactions by means of FCA and a set of disproportionality measures to discard statistically non significant associations. Our approach is illustrated on a set of case reports on AIDS patients and antiretroviral drugs. A validation thereof involving the SRS DB of the French Medicines Agency is also reported.

[1] Source : MedEffect, Canada vigilance online database.

The paper starts by a short presentation of concept lattices and association rules (Sect. 2). Follows the description of the proposed method (Sect. 3). Sect. 4 presents the results of the preliminary experiments. Related work is summarised in Sect. 5 while further research directions are given in Sect. 6.

2 Background on concept lattices and association rules

2.1 Concept lattices

Formal concept analysis (FCA)[6] is a method for designing concepts and conceptual hierarchies from collections of individuals (formal objects) described by properties (formal attributes). To apply FCA to PV data as presented in Table 1, the latter must first be encoded in standard format. The format, a binary context $\mathcal{K} = (O, A, I)$, (see Table 2) involves a set of objects O, a set of attributes A and an incidence relation $I \subseteq A \times O$ (oIa stand for "object o has the attribute a"). For instance, in Table 2, objects are patients and attributes demographic informations, drugs or reactions.

Demographic data						Adverse reactions										Drugs				
	Young	Adult	Senior	Male	Female	Fatigue	Oedema	BreathDisorder	Hives	Headache	Nausea	HeartFailure	HairLoss	Bleeding	Vomiting	`Raltegravir`	`Lopinavir`	`Tenofovir`	`Maraviroc`	`Efavirenz`
Daffy		×			×				×		×				×		×			×
Farley			×	×			×		×	×	×	×	×				×	×		
Lane		×			×	×	×		×				×	×					×	×
Shana	×				×	×	×						×				×	×		
Trudy		×		×		×		×			×	×		×	×	×				

Table 2. Binary context encoding AIDS patients with their drugs and ADR.

Two derivation operators, both denoted $'$ link objects and attributes [6]. Let $X \subseteq O$, $Y \subseteq A$: $X' = \{a \in A | \forall o \in X, oIa\}$, $Y' = \{o \in O | \forall a \in Y, oIa\}$. For example, following Table 2, {`Daffy`, `Trudy`}' = {`Adult`, `Nausea`, `Vomiting`}. The compound operators $''$ are *closure operators* over 2^O and 2^A, respectively. A set $Y \subseteq A$ is closed if $Y = Y''$ which means the objects sharing Y, i.e., Y', share no other attribute (i.e., from A/Y). A pair of sets corresponding to one-another through $'$ is called a (formal) *concept*: $c = (X, Y) \in \wp(O) \times \wp(A)$ is a concept of $\mathcal{K}$ iff $X' = Y$ and $Y' = X$ (here X and Y are called the *extent* and the *intent* of c, respectively). For instance, ({`Farley`, `Shana`}, {`HairLoss`, `Oedema`, `Lopinavir`, `Tenofovir`}) is a concept (c_6 in Fig. 1).

Furthermore, the set $\mathcal{C}_\mathcal{K}$ of all concepts of the context $\mathcal{K}$ is partially ordered by extent inclusion (intent containment). The structure $\mathcal{L} = \langle \mathcal{C}_\mathcal{K}, \leq_\mathcal{K} \rangle$ is a complete lattice, called the *concept lattice*. Fig. 1 shows the lattice of the context in Table 2, whereby a simplified labeling scheme is used where each object/attribute appears only once in the diagram. The extent of a concept is made of all objects whose labels can be reached from the concept on a downward-heading path while intent is recovered in a dual way. For example, the extent of the concept with the attribute label `Bleeding` is {`Lane`, `Trudy`} while its intent is {`Bleeding`, `Fatigue`}.

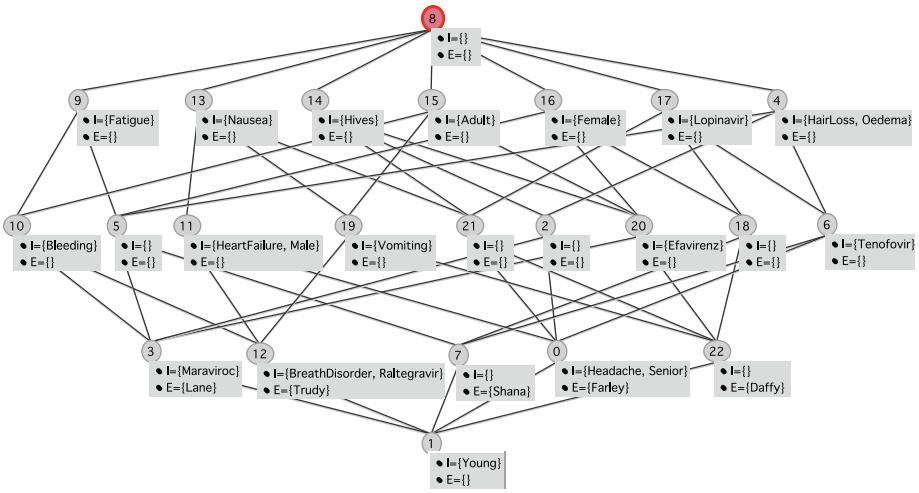


Fig. 1. Concept lattice of case reports given in Table 1.

Within the lattice $\mathcal{L} = \langle \mathcal{C}_\mathcal{K}, \leq_\mathcal{K} \rangle$, concepts have a unique greatest lower bound termed *meet* ($\bigwedge$) that is defined as follows: $\bigwedge_{i=1}^{k}(X_i, Y_i) = (\bigcap_{i=1}^{k} X_i, (\bigcup_{i=1}^{k} Y_i)'')$ For instance, in Fig. 1, the meet of $c_{\#19}$ = ({`Daffy`, `Trudy`}, {`Adult`, `Vomiting`, `Nausea`}) and concept $c_{\#20}$=({`Daffy`, `Lane`}, {`Adult`, `Female`, `Hives`, `Efavirenz`}) is $c_{\#22}$=({`Daffy`}, {`Adult`, `Female`, `Efavirenz`,`Vomiting`, `Hives`, `Nausea`, `Vomiting`, `Lopinavir`}). In addition, the function $\mu : A \rightarrow \mathcal{C}_\mathcal{K}$ maps an attribute a into the *maximal* concept in the lattice having that attribute ($\mu(a) = (a', a'')$). For instance, in Fig. 1, μ(`HeartFailure`)=$c_{\#11}$.

The lattice in Fig. 1 provides the analyst with a variety of insights into the data such as the profile of the AIDS patients under study, the different anti-HIV treatments and the respective most common ADR. For instance, the concept $c_{\#20}$ = ({`Daffy`, `Lane`}, {`Female`, `Adult`, `Hives`}) represents adult female patients under anti-HIV drug regimen containing NNRTIs[2], including `Efavirenz`, and experiencing `Hives`. In summary, the lattice of case reports provides an overview

[2] Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs) intervene in the early stages of the HIV replication cycle.

of drug-reaction combinations to be explored for pharmacological associations detection. In many cases, too specific concepts are not relevant. To only keep those having extents of certain size, the support of a concept is defined as its relative extent size, $\sigma(c) = \frac{\|X\|}{\|O\|}$. The corresponding sub-order of the lattice, i.e., its upper part induced by threshold α in $]0, 1]$, is $\bar{\mathcal{L}}^{\alpha} = \langle \bar{\mathcal{C}}^{\alpha}, \leq_{\mathcal{K}} \rangle$ where $\bar{\mathcal{C}}^{\alpha} = \{c \mid c \in \mathcal{C}, \sigma(c) \geq \alpha\}$. $\bar{\mathcal{L}}^{\alpha}$ is called the *iceberg* lattice [10].

2.2 FCA-based association rule design

FCA framework is widely-used in mining patterns from DB, including association rules that express the co-ocurrences among attribute sets (called *itemsets*). An association rule is a pair of sets '*antecedent* $\longrightarrow$ *consequent*' with no claim of causality. A rule $B \longrightarrow D$ $(B, D \subseteq A)$ has a support $\bar{\sigma}(B \longrightarrow D) = \sigma(B \cup D)$ and a confidence that is the ratio of the rule support to the support of the antecedent ($\bar{\gamma}(B \longrightarrow D) = \frac{\bar{\sigma}(B \longrightarrow D)}{\sigma(B)}$).

	Support
`Tenofovir⟶HairLoss, Oedema`	0.4
`Maraviroc⟶Bleeding, Fatigue, HairLoss, Hives, Oedema`	0.2
`Efavirenz⟶Hives`	0.4
`Raltegravir⟶Bleeding, BreathDisorder, Fatigue, HeartFailure, Nausea,Vomiting`	0.2
`Lopinavir,Efavirenz⟶Hives, Nausea, Vomiting`	0.2
...	...

Table 3. Drug-reaction associations derived from the SRS data depicted in Table 1 with the corresponding support.

In FCA, mining association rules from a DB consists in: (i) extracting all frequent closed itemsets, i.e., concept intents from the DB, with support above α, (ii) generating all valid association rules, i.e., rules whose confidence exceeds a user-defined minimum threshold. The first step presents a greater challenge as the set of frequent itemsets may grow exponentially with the size of A while the second step is relatively straightforward. Moreover, several FCA-based algorithms [1] generate non-redundant bases of association rules. These bases are minimal with respect to the number of rules whereas the contained rules are informative, i.e., with minimal antecedents and maximal consequents. To extract a tractable number of association rules from PV data, we have used the Informative Generic Basis (IGB) [1] as it has been shown that this type of association rules conveys the maximum of useful knowledge, without information loss. Moreover, our IGB contains exact (versus approximative) associations rules, i.e., rules whose confidence is equal to 1 (as opposed to confidence < 1). Table 3 illustrates some of the drug-reaction associations from the IGB extracted out of data in Table 1.

3 Detecting safety signals using FCA

The outline of our mining method is as follows: First, SRS data is encoded into a binary context, where formal objects represent case reports while formal attributes are either taken drugs or the observed reactions (see Table 2). Then, FCA is used to derive both the lattice and the corresponding IGB. For instance, in the case of anti-HIV drugs `Lopinavir` and `Tenofovir` and the two ADR `HairLoss` and `Oedema`, the method will consider only the pair ({`Lopinavir`, `Tenofovir`}, {`HairLoss`, `Oedema`}) since it represents the only combination where the four elements appear (concept c_6 in Fig. 1).

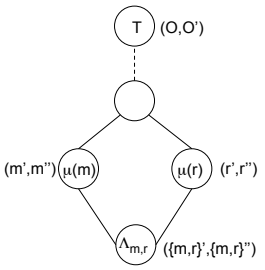
Rules of the basis are split into three groups. *Pure* association have both antecedent and consequent made exclusively of drugs and reactions respectively. *Semi-pure* associations, in contrast, admit only non-reaction items in their consequent that are further removed for analysis purposes. Finaly, *biaised* associations admit non-drug items in their antecedents as well. Their components are filtered to fit the *drugs* $\longrightarrow$ *reactions* rule scheme. Later, statistical filters are applied to detect statistically significant candidates for each of the two types of pharmacological associations, i.e., signals and drug-drug interactions.

In order to discard statistically non significant concepts, we use some of the measures of disproportionality [13] that are currently applied in various reporting centers, e.g., the British Medicines and Healthcare products Regulatory Agency (MHRA). Such measures for a suspected ADR of a drug of interest are calculated from the following variables: (a) reports including the drug of interest and the suspected reaction, (b) reports with the drug of interest and no reference to the suspected reaction, (c) reports where the suspected reaction appears without the drug of interest, (d) reports where neither the drug of interest nor the suspected reaction appear. The adopted measures are the proportional ADR reporting ratio (PRR), reporting odds ratio (ROR), and χ^2 test.

For instance, the PRR is the proportion of the suspected ADR versus all ADR reported for the drug of interest devided by the corresponding proportion for other drugs. It can be expressed as PRR $=\frac{a\times(c+d)}{c\times(a+b)}$. Fig. 2 shows how the various cells of the drug-ADR contingency table are calculated using a drug-reaction concept lattice. Hence, every meet concept $\wedge_{m,r}$ in the lattice $\mathcal{L}_P$, for a given pair of a medicine m and a reaction r is the source of a drug-reaction contingency table.

For instance, the calculation of PRR for the anti-HIV drug `Lopinavir` and the suspected ADR `HairLoss` using the concept lattice of Fig. 1 is as follows: $a = |Ext((\bigwedge_{\texttt{Lopinavir,HairLoss}})| = |Ext(c_{\#6})| = 2$, $b = |Ext(\mu(\texttt{Lopinavir}))|$-$|Ext(\bigwedge_{\texttt{Lopinavir,HairLoss}})| = |Ext(c_{\#17})|$-$|Ext(c_{\#6})| = 1$, c=$|Ext(\mu(\texttt{HairLoss}))|$-$|Ext(\bigwedge_{\texttt{Lopinavir,HairLoss}})|$=$|Ext(c_{\#4})|$ -$|Ext(c_{\#6})|$=1, d=$|O|$-($|Ext(\mu(\texttt{Lopinavir}))$ $\cup$ $Ext(\mu(\texttt{HairLoss}))|$) = 5-$|Ext(c_{\#17}) \cup Ext(c_{\#4})|$=1, PRR $= \frac{2\times(1+1)}{1\times(2+1)}$. The obtained value of PRR is $1.33 \leq 2$. Hence, the assumption stating that `Lopinavir` causes `HairLoss` is statistically non significant.

The detection of higher-order drug-reaction associations, such as drug interactions, has been carried out so far by logistic regression modelling [7] where



	Adverse reaction r	$\neg r$	
Drug m	a	b	$a+b$
$\neg m$	c	d	$c+d$
	$a+c$	$b+d$	$\|O\|$

$a = |Ext(\wedge_{m,r})|$
$b = |Ext(\mu(m))| - |Ext(\wedge_{m,r})|$
$c = |Ext(\mu(r))| - |Ext(\wedge_{m,r})|$
$d = |O| - |Ext(\mu(m) \cup Ext(\mu(r))|$

Fig. 2. Left: Drug-concept and ADR-concept in the drug-reaction concept lattice $\mathcal{L}_P$. **Right**: The two-by-two contingency table for target drug m and a suspected ADR r.

concomitant drugs (resp. reactions) are considered as covariates and the suspected reaction (resp. drug) as dependent variable. For instance, in our running PV data, the logistic model predicting whether `Nausea` reaction is a result of possible interaction between `Lopinavir` and `Efavirenz` would look like:

$$N = \beta_0 + \beta_1 \times L + \beta_2 \times E + \beta_3 \times L^*E$$

The variables L and E are exposure variables (or predictors) representing risk factors associated with concomitant drugs `Lopinavir` and `Efavirenz`, respectively, while L^*E is the interaction term. The intercept β_0 represents the value of the dependent variable `Nausea` (N) in a patient with no risk factors, while logistic (or logit) coefficients β_1, β_2, and β_3 basically quantify the expected variation in N associated with a unit change in the binary predictor variables `Lopinavir`, `Efavirenz`, and the interaction term, respectively.

Maximum likelihood estimation (MLE) can be used to calculate logistic coefficients. In the case of `Nausea`, `Lopinavir` and `Efavirenz`, calculating logit coefficients using the hypothetical contingency table depicted in left-hand side of Fig. 3 and R package yields $N = -1.609 - 0.993 \times L + 0.226 \times E + 2.337 \times L \times E$ with the p-values depicted in right-hand side of Fig. 3. The interpretation would be that the interaction is statistically significant as the p-value for the interaction term is 0.0381, a value that is less than the usually accepted threshold of 0.05.

L	E	L*E	N	¬N
1	1	1	9	9
0	1	0	6	23
1	0	0	2	27
0	0	0	4	20

Logit coefficient	p-value
$\beta_0 = -1.609$	0.0033
$\beta_1 = -0.993$	0.2776
$\beta_2 = 0.226$	0.7099
$\beta_3 = 2.337$	0.0381

Fig. 3. Left : $2 \times 2 \times 2$ contingency table of reports for the regression of `Nausea` (`N`) on two exposure level `Lopinavir` (`L`), `Efavirenz` (`E`) and their interaction term `L*E`. **Right :** The corresponding logit coefficients provided by the R package.

4 Tools and experiments

SignalMiner[3], is an open source tool dedicated to mining significant drug-reaction associations. The tool is coupled with, on the one hand, Galicia open-source platform[4] for handling FCA data including the input contexts, concept/iceberg lattices and rule basis, and on the other hand, the open-source statistical computing and graphics environment R[5] for data pre-processing and multivariate statistics including logistic regression analysis. In addition, SignalMiner performs a wide range of standard calculations, e.g., PRR, ROR, χ^2 (with Yates correction), etc.

The SRS DB of the French Medicines Agency (Afssaps) was used for the validating experiments. We have tested the proposed method on several moderate-size subsets of the dataset. For instance, for a pool of 3249 case reports containing 527 drugs and 639 ADR. The obtained lattice comprises 13178 concepts while the corresponding rule basis contains 28117 rules among them only 1165 represent candidates for pharmacological associations. These candidates are further distilled by SignalMiner to identify pure, semi-pure or biaised associations as illustrated in Table 4. Thus, the 1165 suggested association candidates (Table 4) are further filtered, on the one hand, by focusing potential safety signals satisfying the above MHRA '*interestingness*' criteria, and on the other hand, by focusing drug interactions that have been revealed significant using regression analysis. The minimum criteria for raising hypotheses regarding safety signals are as follows: number of reports (patients) ≥ 3, PRR ≥ 2, and $\chi^2 \geq 4$ (with Yates correction).

Among 834 candidates representing safety signals (Table 4), we have found that 63 candidates are statisticaly significant safety signals including 36 known signals (57%), e.g., {`Abciximab`, `Thrombopenia`}, 16 new signals warranting further investigations, e.g., {`Lamivudine`, `Arthralgia`}, while the remaining potential signals are either association where the drug appears as an innocent bystander, e.g., {`Ritonavir`, `Hypophosphatemia`}, or non-interpretable association, e.g., {`Bupivacaine`, `decrease of the therapeutic effect`}. In addition, among 331 associations representing candidates for drug interactions (Table 4), 10 candidates are revealed to be statistically interesting. In a previous work [3, 4], disproportinality measures extracted 523 and 360 statistically significant $\{drug, \text{ADR}\}$ couples, respectively. Our approach returns a smaller set of drug-reaction associations to be further investigated.

5 Related work

Several studies from the literature address the use of DMA to identify drug-reaction associations. In [4], the use of FCA in signal detection is briefely addressed. To assess the strength of the association between a target drug and

[3] `http://safetyseer.cvs.sourceforge.net/signalminer/`
[4] `http://www.iro.umontreal.ca/~galicia`
[5] `http://www.r-project.org/`

	# Pure	# Semi-pure	# Biaised	
Signals	1	88	745	834
Interactions	1	260	70	331
	2	348	815	1165

	# Pure	# Semi-pure	# Biaised	
Signals	0	4	59	63
Interactions	0	0	10	10
	0	4	69	73

Table 4. Left: Candidates for pharmacological associations obtained from 3249 case reports containing 527 drugs and 639 ADR. **Right:** Statistically significant candidates.

suspected ADR, disproportionality approach introduces several parameters such as, the PRR [5], χ^2 that is often coupled with the PRR, and the ROR [13], whereas Bayesian approach consists of the Multi-item Gamma Poisson Shrinker (MGPS) algorithm [11] and the Bayesian Confidence Propagation Neural Network (BCPNN) [2].

In [9], an interpretation of mathematical structures from FCA into epidemiology is described. A comprehensive survey of state-of-the-art in statistical modelling used by a various DMA of PV data is proposed in [7]. However, to the best of our knowledge, none of them supports automatic detection of pharmacological associations involving several drugs and/or reactions.

6 Discussion

FCA has been applied in combination with statistical metrics to the detection of several types of statistically significant pharmacological associations, e.g., safety signals and drug interactions. Compared to the classical DMA-based detection, the proposed FCA method improves the quantity and quality of extracted pharmacological associations, including those involving several drugs and/or reactions. Indeed, the amount of extracted associations is reduced by targeting basis of association rules using FCA framework, yet relevant associations with respect to the referred population of case reports, thereby saving investments in time and money that would be spent in further clinial trials.

In the future, we intent to reformulate drug-reaction analysis so that detecting pharmacological association is mapped to a relational data mining problem [8]. Moreover, because drug-reaction analysis deals with a dynamic DB that comprises high volume of data, the reconstruction –from scratch– of a new concept lattice for every change in the SRS DB is so computationally expensive that it is prohibitive. We shall address the on-line analysis of pharmacovigilance data using the incremental maintenance of concept lattice [12] and the respective association basis.

7 Acknowledgments

This work was supported partially by project grant (Ref. ANR-07-TecSan) from the French *Agence Nationale de la Recherche, Biologie et Santé* and *Caisse Nationale de Solidarité pour l'Autonomie*, and a discovery grant of the Natural Sciences and Engineering Research Council of Canada.

References

1. Y. Bastide, N. Pasquier, R. Taouil, G. Stumme, and L. Lakhal. Mining minimal non-redundant association rules using fci. *LNCS*, 1861:972–986, 2000.
2. A. Bate, M. Lindquist, I.R. Edwards, S. Olsson, R. Orre, A. Lansner, and R.M. De Freitas. A bayesian neural network method for adverse drug reaction signal generation. *European Journal of Clinical Pharmacology*, 54(4):315–321, July 1998.
3. C. Bousquet, C. Sadakhom, C. Le Beller, M.-C. Jaulen, and A. L. Louet. A review of potential signals generated by an automated method on 3324 pharmacovigilance case reports. *Therapie*, 61(1):39–47, 2006.
4. A. Estacio-Moreno, Y. Toussaint, and C. Bousquet. Mining for adverse drug events with formal concept analysis. In *Proceedings of MIE'08*, volume 136 of *Studies in Health Technology and Informatics*, pages 803–808. IOS Press, 2008.
5. S. J. Evans, P. C. Waller, and S. Davis. Use of proportional reporting ratios (prrs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiology and Drug Safety*, 10(6):483–486, December 2001.
6. B. Ganter and R. Wille. *Formal Concept Analysis, Mathematical Foundations*. Springer-Verlag, 1999.
7. M. Hauben, D. Madigan, C. M. Gerrits, L. Walsh, and E. P. van Puijenbroek. The role of data mining in pharmacovigilance. *Expert Opinion on Drug Safety*, 4(5):929–948, September 2005.
8. M. Huchard, M. Rouane-Hacene, C. Roume, and P. Valtchev. Relational concept discovery in structured datasets. *Annals of Mathematics and Artificial Intelligence*, 49(1-4):39–76, 2007.
9. A. Pogel and D. Ozonoff. Contingency structures and concept analysis. In *ICFCA*, volume 4933 of *LNCS*, pages 305–320. Springer, 2008.
10. G. Stumme, R. Taouil, Y. Bastide, N. Pasquier, and L. Lakhal. Computing iceberg concept lattices with TITANIC. *Data Knowledge Engineering*, 42(2):189–222, 2002.
11. A. Szarfman, S. G. Machado, and R. T. ONeill. Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the us fdas reports database. *Drug Safety*, 25:381–392(12), 2002.
12. P. Valtchev, M. Rouane-Hacene, and R. Missaoui. A generic scheme for the design of efficient on-line algorithms for lattices. In A. de Moor, W. Lex, and B. Ganter, editors, *In Proceedings of the 11th Intl. Conference on Conceptual Structures (ICCS'03)*, volume 2746 of *LNCS*, pages 282–295, Berlin (DE), 2003. Springer--Verlag.
13. E. P. van Puijenbroek, W. E. Diemont, and K. van Grootheest. Application of quantitative signal detection in the dutch spontaneous reporting system for adverse drug reactions. *Drug Safety*, 26:293–301, 2003.