Data Mining to Generate Adverse Drug Events Detection Rules

Emmanuel Chazard, Grégoire Ficheur, Stéphanie Bernonville, Michel Luyckx, and Régis Beuscart

***Abstract*—Adverse drug events (ADEs) are a public health issue. Their detection usually relies on voluntary reporting or medical chart reviews. The objective of this paper is to automatically detect cases of ADEs by data mining. 115 447 complete past hospital stays are extracted from six French, Danish, and Bulgarian hospitals using a common data model including diagnoses, drug administrations, laboratory results, and free-text records. Different kinds of outcomes are traced, and supervised rule induction methods (decision trees and association rules) are used to discover ADE detection rules, with respect to time constraints. The rules are then filtered, validated, and reorganized by a committee of experts. The rules are described in a rule repository, and several statistics are automatically computed in every medical department, such as the confidence, relative risk, and median delay of outcome appearance. 236 validated ADE-detection rules are discovered; they enable to detect 27 different kinds of outcomes. The rules use a various number of conditions related to laboratory results, diseases, drug administration, and demographics. Some rules involve innovative conditions, such as drug discontinuations.**

***Index Terms*—Adverse drug events (ADEs), data mining, decision trees, electronic health records, patient safety.**

I. Introduction

ADVERSE drug events (ADEs) endanger patients as they are the most common type of iatrogenic injury [1]. They can be defined as "injuries due to medication management rather than the underlying condition of the patient" [2]. ADEs can be split into two categories: preventable ADEs that are medication errors leading to patient harm, and nonpreventable ADEs that are called adverse drug reactions [3].

Different methods are used to identify ADEs [4]–[6], the most prominent ones being chart reviews and reporting systems. Retrospective medical chart reviews constitute the main source of reliable epidemiological knowledge on ADEs, but the method is extremely time and resource consuming. Reporting systems are the most ancient methods: they are useful for the analysis of contributing factors of ADEs, but all reporting systems suffer from important under-reporting biases [5], [7]. Another way to detect ADEs is to mine free-text reports by means of natural language processing [8]–[11], assuming that the ADEs are described in the reports, which is not frequent. Data mining is sometimes used in the field of ADE detection. But it was mainly used to analyze voluntary ADE reports [12]–[17] by means of supervised rule induction methods such as decision trees, association rules, or Bayesian neural networks, and not to analyze hospitalization records. As a consequence, the results can only be used to analyze other voluntary ADE reports.

In the literature, the automated detection of ADE cases in hospital records always relies on ADE detection rules. Whatever their origin, the ADE detection rules always consist of one or two conditions that lead to an outcome. Those conditions are simple: two drugs [18]–[25], a drug and a laboratory result [5], [18], [19], [26]–[28], [30], a drug alone [5], [18], [21], [22], a drug and one patient's characteristic [5], [18], [24], or a drug and a drug allergy [18], [24], [27]. Those works are not able to mix more complex patterns of conditions, and the effects of drug discontinuation are ignored. Those rules usually lead to overalerting, as they detect many potential cases that are not real ADE cases [29]. This is notably due to the absence of contextualization of the knowledge (the same rules are applied in every medical department) and the absence of segmentation of the population (the rules do not involve additional conditions that could increase the probability of the outcome).

II. Objectives

In order to improve the patient safety and avoid the under-reporting bias, this paper aims at automatically discovering ADEs that occurred in inpatients. This will be done by identifying situations at risk of ADE by data mining of routinely collected data of past hospitalizations. In those data, the ADEs are not explicitly flagged as no preliminary review is performed. Outpatients' ADEs leading to hospitalization will not be studied.

A list of outcomes will first be defined, and the link between those outcomes and prior drug administrations or discontinuations will be studied by means of supervised rule induction techniques applied on a training set. Rules will be obtained, in which an outcome is explained by a set of drugs in combination with a clinical background, in the form of ADE detection rules (e.g., *drug_A* & *background_B* $\rightarrow$ *outcome_C*). Then those rules will be applied onto past hospital stays of an evaluation set to get contextualized statistics such as the confidence (e.g., probability of *outcome_C* when *drug_A* and *background_B* are present).

Regarding data mining techniques, two issues have to be solved: 1) the temporal constraints have to be taken into account; and 2) we have to use supervised rule-induction methods, although the ADEs are not explicitly flagged in the routinely

Manuscript received December 25, 2010; revised May 14, 2011 and July 28, 2011; accepted July 29, 2011. Date of publication; date current version; This work was supported by the European Community's Seventh Framework Program (FP7/2007-2013) under Grant Agreement n° 216130: the PSIP Project.

The authors are with the Université Lille Nord de France, F-59000 Lille, France (e-mail: emmanuel.chazard@univ-lille2.fr; gregoire.ficheur@gmail.com; stephanie.bernonville@univ-lille2.fr; mluyckx@ch-denain.fr; regis.beuscart@univ-lille2.fr).

Digital Object Identifier 10.1109/TITB.2011.2165727

TABLE I
DESCRIPTION OF THE HOSPITALS AND STAYS USED

Hospital	Number of stays included	Age in years *mean (sd)*	Men proportion	Duration in days *mean (sd)*	Wards
French #1	50,072	52.8 (21.6)	29.2%	5.48 (6.10)	Medicine surgery obstetrics
French #2	1,367	71.4 (18.4)	42.1%	11.4 (15.1)	Geriatrics
French #3	7,846	45.4 (27.5)	51.6%	10.7 (15.3)	Geriatrics and Cardiology
Danish #1	26,245	55.6 (25.9)	40.4%	4.56 (11.8)	Medicine surgery obstetrics
Danish #2	23,067	53.1 (22.6)	44.8%	4.51 (8.41)	Medicine surgery obstetrics
Bulgarian	6,880	49.4 (16.1)	26.4%	6.96 (2.54)	Endocrinology

collected data, which are usually required in the classical rule induction method.

III. MATERIAL

A. Electronic Records of Past Hospital Stays

In order to analyze past hospital stays, data are extracted from several hospitals' electronic health records (EHRs) to feed a common repository with past fully anonymized hospital stays. The repository fits a common data model that has been designed within the PSIP Project (patient safety thought intelligent procedures in medication), a European project that aims at facilitating the development of knowledge on ADE, and improving the medication cycle in hospital environments [30], [31]. Only routinely collected data are used: no data have to be specifically recorded for the project. For each hospital stay, those data include the following.

1) Medical and administrative information (e.g., age, gender, admission date, medical department, etc.).
2) Diagnoses encoded using the International Classification of Diseases, tenth version (ICD10).
3) Medical procedures encoded using national classifications, including therapeutic and diagnostic procedures.
4) Drugs administered to the patient, encoded using the Anatomical Therapeutic Chemical classification (ATC).
5) Laboratory results encoded using the International Union of Pure and Applied Chemistry classification.
6) Anonymized free-text records, such as the discharge letter.

The data from EHRs are provided by six hospitals that are part of the PSIP Project. This study is performed using 115 447 records from six hospitals (see Table I). They allow for a four-year follow up (from 2007 to 2010).

B. ADE Detection Rules From the Summaries of Product Characteristics

In many countries, the summaries of product characteristics (SPCs) describe the official and exhaustive information about ADEs. They can be used to support the drug prescription process. They are available for healthcare professionals through websites and various supports, and for patients through patient information leaflets. In this paper, those SPCs are necessary 1) to get an exhaustive list of the possible outcomes that can be observed due to ADEs, irrespective of the causes and 2) to get a reliable set of rules to validate the results of data mining.

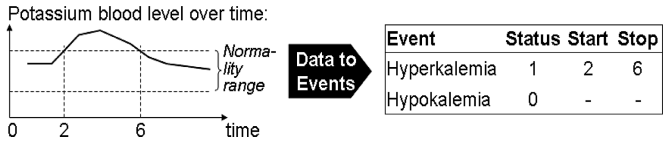

Fig. 1. Example of aggregation of laboratory results: several values of blood potassium over time enable to search for simple events.

In France, the SPCs are managed by the French drug agency (AFSSAPS, French agency for sanitary security of health products). A structured version of the SPCs is provided by the Vidal Company in the form of rules where one or two conditions lead to an outcome (e.g., Furosemide → hypokalemia). In those rules, the kind of outcome is described using free text. More than 500 000 rules are available, those rules lead to 228 different kinds of clinical or paraclinical outcomes. The paraclinical outcomes are mainly laboratory or electrocardiographic abnormal results.

IV. METHODS

A. Aggregation of the Complex Data of the Stays Into Simple Events

1) General Principles: The data described in the data repository are characterized by a complex data scheme, very numerous classes (about 17 000 codes for ICD10, about 5400 codes for the ATC, etc.) and repeated measurements throughout the hospitalization (e.g., laboratory parameters and drug administrations). Those characteristics make those data too complex to be mined using statistical methods. The aim of the data-to-event aggregation process is to automatically get a simpler representation of data for data mining purposes.

Aggregation engines are developed in order to transform the available data into information described as sets of events. For each kind of data (administrative information, diagnoses, drugs, and laboratory results), a specific aggregation engine is developed and fed with a mapping. Each mapping is described by means of extensible markup language (XML) files outside the engine. The aggregation engines enable to describe the events in terms of binary variables complemented by start and stop dates. Those engines are not static and can be adapted with respect to the context.

2) Example of Aggregation of Laboratory Results: In the example displayed in Fig. 1, for a given stay, several measures of potassium are available. Potassium is an electrolyte; its level in the blood should not reach too low or too high values; otherwise, it could lead to lethal heart arrhythmias. The aim of the aggregation process is to get simple information from those repeated measures. In the case displayed in Fig. 1, there is a hyperkalemia (too high potassium value) from day 2 to day 6 and no hypokalemia. Finally, the various measures can be summarized into two binary variables: hypokalemia = 0 and hyperkalemia = 1. In that case, a start and stop date can be added. Such variables are easier to mine using statistical methods.

3) New Variables Made Available by the Aggregation: The aggregation engines transform the data into several binary variables that can easily be mined.

1) Fifteen variables related to demographic and administrative information.
2) Forty-eight variables related to chronic diseases.
3) Five hundred variables related to drug administration or drug discontinuations. The classification considers pharmacodynamics and pharmacokinetics, although most of the existing drug classifications are based on indications.
4) Thirty-five variables related to laboratory value abnormalities.

B. Identification of the Outcomes in Relation With ADEs

As described in Section III, a list of outcomes is extracted from the summaries of product characteristics. The outcomes are traced in the data essentially by screening the laboratory results and administered drugs; this is possible through different ways depending on the category of outcome. For instance, the occurrence of a hyperkalemia (laboratory-related outcome) is directly traced using the potassium level in the blood. The occurrence of a hemorrhage under vitamin K antagonists (VKA) can be traced through different ways: 1) an increase of the international normalized ratio (INR), a laboratory parameter that rises up in case of VKA overdose; and 2) the vitamin K administration, an antidote which is prescribed in case of hemorrhage under VKA.

The structured SPC database describes 228 different kinds of outcomes. 83 (37%) of those outcomes are traceable in this paper, due to the available data. Duplicate entries are then removed; for instance, in the initial list, "hyperbilirubinemia" is also described using two synonyms, "bilirubinemia higher than twice the normal upper bound" and "jaundice." As a consequence, those 83 outcomes are traced through 56 different variables. Those outcomes correspond to life-threatening ADEs, such as hyperkalemia of hemorrhage hazard. Unfortunately, some outcomes cannot be traced in the data. This is the case especially for minor clinical incidents such as nausea or gastric pain cannot be traced. Those outcomes could correspond to ICD10 codes but in most hospitals, such codes are not flagged with a date.

C. Data-Mining-Based Induction of ADE Detection Rules

The knowledge about ADEs can be expressed using rules where some conditions lead to an outcome. Some of the variables computed by the aggregation process can be used as outcome (e.g., death) and some other ones can be used as conditions (e.g., chronic renal failure). In total, 588 variables can be used as conditions to explain 56 different outcomes. The objective here is to automatically link conditions with outcomes and then to discover ADE detection rules using data mining techniques.

After a complete review of the available data mining supervised and unsupervised techniques, and after several experiments, it was decided to use decision trees (with the CART method: Classification and Regression Trees) [32]–[39] and association rules [40]. Both methods enable to identify several decision rules containing 1 to K conditions such as

Rule A: drug_X & age≥70 → renal failure (probability= 15%)
Rule B: drug_X & age<70 → renal failure (probability= 3%)

Fig. 2. Example of two rules. A "segmentation" condition is underlined: it does not explain why the outcome occurs but deeply changes its probability.

IF(condition_1 &. . .& condition_K) THEN outcome.

The dataset (92 486 stays) is split into a learning set that is used for the rule induction (31 579 stays of the year 2007) and an evaluation set (60 907 stays of years 2008–2010). Decision trees and association rules are automatically launched for each outcome in each hospital and each medical department. The datasets are managed during the rule induction so that temporal constraints are taken into account. For a given outcome, only conditions that are compatible regarding time are tested: each condition must be an event that occurs before the outcome and is still active or has ended less than a fixed delay before the outcome occurs.

Both methods produce thousands of rules that must be filtered. Most of the outcomes are due to the patient's medical background rather than the drugs administered to him. For that reason, the rules are automatically filtered in order to keep the ones in which a drug is involved. Only the rules that increase the probability of the outcome and that have at least one of the following condition types are kept.

1) A drug administration.
2) A drug discontinuation.
3) A laboratory value that is implicitly due to a drug administration (e.g., lithium blood level > 0, INR > 1, etc.).

D. Expert Validation and Reorganization of the Rules

It is mandatory to filter, validate, and organize the rules that are obtained from the data mining: as the rules have to be used by physicians, they must provide simple, validated, and unquestionable knowledge. Several meetings are organized with external experts (physicians, pharmacologists, pharmacists, and statisticians) to filter and reorganize the set of rules. The rules are examined and validated against the SPCs and scientific references. During the review, the experts may ask for complementary queries on the potential ADE cases. At this step, the experts may manually add a few rules that are considered as mandatory although they were not discovered by the data mining process, for instance because the conditions of the rules never occur (e.g., absolute contraindication) or because the conditions occur but do not lead to any outcome.

In every rule, there is a set of conditions; the experts are asked to characterize each condition according to one of the following types.

1) *"Segmentation" conditions* are conditions that do not explain *why* an outcome occurs, but deeply change its probability. This kind of condition enables us to reduce overalerting. An example of "segmentation" condition is underlined in Fig. 2.
2) *"Subgroup" conditions* are fixed when, for some medical reasons, it does not make sense to consider the rules for all the patients in the same time. They are used to define

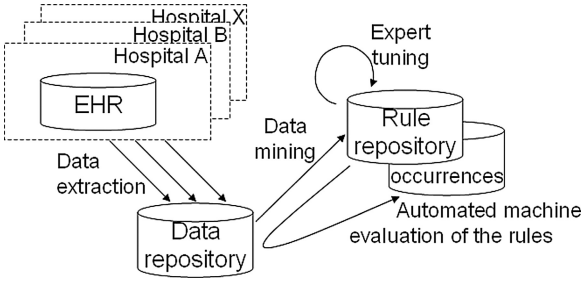

Fig. 3. Rules are stored in a rule repository. A machine evaluation automatically computes various statistics (occurrences) of the rules in every medical department.

the sample before computing the statistics. The following subgroups are systematically defined.

a) The INR deviations or vitamin K administrations are only explored for VKA-treated patients.
b) The increase of activated partial thromboplastin time is only explored for heparin-treated patients.
c) The hyperkalemia is explored separately for patients suffering from renal insufficiency or not.

3) *"Basic" conditions* group together all the other conditions.

E. Automated Computation of Contextualized Statistics About the Rules

Validated ADE detection rules are obtained from the phase of expert validation. Those rules are then evaluated from a statistical point of view to provide the users with the classical parameters such as confidence and relative risk. The statistics are contextualized, i.e., they are computed separately in each medical department or each hospital.

For that purpose, the validated rules are stored in a central rules repository [41], using an XML schema. Then, in a few minutes, an automated machine evaluation of the rules can be performed using the 60 907 stays of the evaluation set (see Fig. 3). A rule is a set of conditions leading to an outcome, such as C_1 & ... & C_k $=>$ O. Several statistics are computed for each rule, separately in every medical department.

1) Support $= \mathrm{P}(\mathrm{O} \cap \mathrm{C}_1 \cap \ldots \cap \mathrm{C}_k)$.
2) Confidence $= \mathrm{P}(\mathrm{O} \mid \mathrm{C}_1 \cap \ldots \cap \mathrm{C}_k)$.
3) Relative risk $\frac{RR=P(O/C_1 \cap \ldots \cap C_k)}{P(O/(C_1 \cap \ldots \cap C_k))}$.
4) p value of the Fisher's exact test for independency between the outcome (O) and the set of conditions$(\mathrm{C}_1 \cap \ldots \cap \mathrm{C}_k)$.
5) Median delay between t1 (all the conditions are met) and t2 (the outcome occurs).

The cases that match the conditions of a rule and have the outcome are considered as potential ADE cases. Additional statistics are computed to describe them: number, average age, death rate, average length of stay, proportion of renal insufficiency, etc.

F. Preliminary Evaluation

An independent medical expert is asked to review the cases detected by the rules, i.e., the cases that match the conditions and the outcome of the rules with respect to temporal constraints. He is asked to assess whether each case is an ADE (the conditions are responsible from the occurrence of the outcome) or not (there is another explanation). He is also asked to review all the cases that do not match any rule but present the outcome on the track of false negatives.

V. Results

A. Overview of the Rules Obtained in This Paper

In this paper, 56 different outcomes enable to trace the potential consequences of ADEs. The supervised rule induction generates rules that predict each outcome. The rules are always filtered, validated, and tuned by the expert committee. 236 validated rules are obtained. The experts also add some rules that appear to be important in the academic knowledge and are not discovered by the data mining (e.g., the conditions never occur, or occur but not lead to the outcome). Over the 56 outcomes, we have the following.

1) Twenty-seven kinds of outcomes are observed and enable to discover ADE detection rules.
2) Ten outcomes are never or too rarely observed in the data, so that no rule is discovered. Data mining will be performed on larger datasets to get results.
3) Eighteen outcomes are observed but cannot be explained by the use of drugs in the available dataset: the medical background of the patient is a sufficient explanation, so that no rule is discovered.

The 236 rules that are obtained can be classified through the outcome they enable to predict (see Table II). Those rules can also be classified into several categories.

1) One hundred and twenty-seven rules have been discovered by data mining and confirmed by the SPCs and they bring new knowledge such as additional segmentation conditions.
2) Forty-four rules have been discovered by data mining and are not present in the SPCs but can be indirectly explained using academic knowledge. Those rules bring new knowledge in ADE detection.
3) Twenty-five rules have been discovered by data mining and already exist as is in the SPCs.
4) Forty rules have not been discovered by data mining but are important in the SPCs. They have been enforced by the experts and do not produce significant statistics. The contribution of this paper is to compute statistics about those rules and quantify their usefulness.

B. Example of an ADE Detection Rule and Related Statistics

The following rule is generated by data mining:
Vitamin K antagonist & anti-diarrheal drug $\rightarrow$ *INR* $\geq$ *5.*

The rule can be explained as follows: in case of diarrhea, the VKA absorption is decreased, which is probably balanced by an increase of the VKA dose. Once an antidiarrheal drug is administered, the VKA absorption is restored. In the absence of dose adjustment, this leads to a VKA overdose detected by an INR value over 5, which can lead to a hemorrhage. The statistics that are related to that rule for the year 2009 are described

TABLE II
OUTCOMES AND NUMBER OF ADE DETECTION RULES

Outcome	Rules
Coagulation disorders	
Hemorrhage (detected by the administration of haemostatic)	7
Heparin overdose (activated partial thromboplastin time>1.23)	5
VKA overdose (INR>4.9)	57
VKA overdose (detected by the administration of vitamin K)	2
VKA underdose (INR<1.6)	18
Thrombocytosis (count>600,000)	5
Thrombopenia (count<75,000)	24
Nosocomial infections	
Bacterial infection (detected by the administration of antibiotic)	4
Fungal infection (detected by the administration of a systemic antifungal)	8
Fungal infection (detected by the admin. of local antifungal)	2
Ionic and renal disorders	
Hyperkalemia (K^+>5.3 mmol/l)	63
Hypocalcemia (Ca^{++} <2.2 mmol/l)	1
Hypokalemia (K^+<3.0 mmol/l)	1
Hyponatremia (Na^+<130 mmol/l)	2
Renal failure (creatinine>135 µmol/l or urea>8 mmol/l)	8
Others	
Acetaminophen overdose (detected by the administration of N-acetyl-cystein)	1
Anemia (Hb<10g/dl)	2
Diarrhea (detected by the administration of an anti-diarrheal)	1
Diarrhea (detected by the administration of an antipropulsive)	1
Hepatic cholestasis (alkaline phosphatase>240 UI/l or bilirubins>22 µmol/l)	3
Hepatic cytolysis (alanine transaminase>110 UI/l or aspartate transaminase>110 UI/l)	4
High CPK level (CPK>195 UI/l)	2
Hypereosinophilia (eosinophilocytes>10^9/l)	4
High level of pancreatic enzymes (amylase>90 UI/l or lipase>90 UI/l)	7
Lithium overdose (to high a lithium rate)	1
Neutropenia (count<1,500/mm^3)	2
Pancytopenia	1
Total	*236*

TABLE III
EXAMPLE OF A RULE: VKA & ANTI-DIARRHEAL → VKA OVERDOSE (INR≥5)

Hospital	Confidence	Support	Median delay	Relative risk	Fisher's exact test
N°1	9/41=22%	9/6110=1.5‰	3 days	16.45	p=0
N°2	2/9=22.2%	2/11923=0.2‰	2 days	75.64	p=0.0003
N°3	0/2=0%	0/1022=0‰		0	p=1
N°4	0/8=0%	0/7685=0‰		0	p=1
N°5	0/1=0%	0/1816=0‰		0	P=1
N°6	0/2=0%	0/6880=0‰		0	p=1

Confidence	Year 2007	Year 2008	Year 2009	Year 2010
Hospital N°1	6/43=14.0%	8/46=17.4%	9/41=22.0%	6/36=16.7%

in Table III, as well as the follow-up of one hospital during four years. Each line of the table displays the results obtained in each of the studied hospitals. In fact, the results are available for each medical department of those hospitals. In both hospitals 1 and 2, the probability of VKA overdose once the conditions are met is around 20%, with a significant increase of the risk, and a median delay of two or three days. In hospitals 3, 4, and 5, no case of VKA overdose is observed when the same conditions are matched. It is interesting to notice that even for a validated rule, its confidence may vary a lot with respect to the hospital or medical department. This is probably due to the fact that the patients (demographic and medical background), the medication processes, and the monitoring policies are different. For instance, in some departments, the nurses are quite self-powered for "comfort" drug administration. In other departments, the INR is not frequently monitored even in case of change in the medication. An immediate consequence of the results of Table III is that the use of such a rule in a clinical decision support system (CDSS) would lead to about 78% of false alerts in hospitals 1 and 2, which is acceptable, but also to 100% of false alerts in hospitals 3, 4, and 5, which is not acceptable.

TABLE IV
EVALUATION OF ADE DETECTION IN THE FIELD OF HYPERKALEMIA

Measure	Value
Number of stays	14,747
Number of hyperkalemia cases	117 (7.93‰)
Recall	39/41=95.1%
Precision	39/75=52.0%
F-Measure	67.2%
Number of cases reported	0 (0%)
Cases above 6 mmol/l	11/41=26.8%
Administration of Kayexalate	12/41=29.3%

C. Preliminary Evaluation

A preliminary evaluation has been conducted exhaustively on the hyperkalemia cases of hospital n°1 during the year 2010 (14 747 stays). The results are reported in Table IV. None of the cases detected in the review had been reported to the patient safety unit or to official agencies, although the potassium was above 6 mmol/l in 26.8% cases, and there was an administration of Kayexalate, a potassium chelator, in 29.3% cases. This review is being continued in all the fields covered by the rules.

VI. DISCUSSION

A. Overview

In this study, data about 115 447 past hospital stays are collected and prepared for data mining. A list of potential outcomes is obtained from the SPCs and 56 of them are traced in the data. By means of decision trees and association rules, decision rules are extracted from a training set (34% of the stay). An expert committee filters and validates those rules: 236 validated ADE detection rules are obtained, and statistics are automatically computed in an evaluation set (66% of the stays).

B. Discussion of the Method

This method is able to automatically discover ADE detection rules. Some are already known and validated. In addition, the method enables to discover new knowledge, such as segmentation conditions or unknown rules. The academic knowledge does not provide any probability of the ADEs. In this paper, we are able to sort the rules by confidence and to prioritize the knowledge. Each one of the 236 ADE detection rules is automatically complemented with contextualized statistics, i.e., statistics computed separately in every hospital or medical department. As shown in the example in Table III, the confidence

often varies a lot with respect to the place a rule is applied. Those differences might be due to latent variables that are not observed in the data, such as the risk monitoring policies or the medical background of the patient.

A drawback of the method is that only the data that are recorded can be mined. In this paper, we are not able to detect clinical events that are not registered in routinely collected data, e.g., rash, nausea, stomach pain, etc. The patient's weight and known drug allergies could have been used, but this information was not sufficiently present in the dataset. The drugs prescribed shortly before the hospitalization were not available and could not be analyzed. Finally, as the rule induction is data mining based, events that never or too rarely occur do not enable to discover rules, which is the case here for 11 outcomes. For that reason, the experts were allowed to add some important rules that never occur into the rule base, such as absolute contraindications. In that case, this paper contributes to compute the statistics about those academic rules. The same method could provide interesting results using other data where the outcomes occur as soon as they are available in databases, such as electrocardiographic records or oxymetry records.

For the data mining phase, the data have to be simplified. For instance, the duration and dose of medications have been ignored, as well as the numeric value of the laboratory results. However, the rules so obtained can be enriched by such parameters later, for instance, in a CDSS, for prospective ADE prevention.

Producing ADE detection rules by data mining is complex. Indeed, the ADE cases are not flagged in the data: when hyperkalemia can be observed, we do not simply know if it is an ADE or not. However, the supervised rule induction methods are used to get some rules that predict hyperkalemia, and in this paper, we try to obtain rules that predict hyperkalemia *in the frame of an ADE*. Yet most of the outcomes are principally due to the patients' diseases, and occasionally due to drugs. For that reason, an automated filtering and an expert filtering and reorganization of the rules are performed. As the decision trees are launched several times in different department and on different periods, their instability is not a problem and provides experts with several partially redundant rules, as the association rules do. Once the rules have been filtered and modified by the experts, they are automatically evaluated in all the medical departments using the evaluation set.

Some authors have developed specific rule-induction methods that deal with temporal aspects [42]–[45]. These methods try to discover some events that, in a given order, lead to an outcome. Regarding ADEs, those methods appear not to be relevant because the order of appearance of the conditions is not overriding, but the conditions have to be active simultaneously. It is not a problem of order of appearance, but a problem of concomitant presence and delay up to the condition. In addition, the discontinuation of a drug itself is a kind of event. For all these reasons, the temporal conditions are analyzed and filtered before the rule induction to ensure that all the events that are candidate to explain an outcome are compatible with the outcome regarding time. Then, the same constraints are applied for the rule automated evaluation.

C. Discussion of the ADE Detection Rules Discovered

This study enables us to automatically discover ADE detection rules by means of data mining techniques; the rules are then filtered and validated by experts. The rules consist of a set of conditions that lead to an effect, those conditions being related to demographic characteristics, drug administrations (without dose), laboratory results, or diagnoses. The number of conditions is not constrained by the method, and the output provides more complex rules than in other studies. In addition, this study takes into account the effects of drug discontinuation. Some previous works have involved segmentation conditions, such as the age, the renal function, the hepatic function, and the patient's weight [18], [24]. This study does so (53% of the 236 rules), except that the patient's weight is not available in the data.

One of the main risks of the systems developed to detect or prevent ADEs is overalerting. This is easily understandable when the official SPCs describe situations at risk of ADEs by means of thousands of rules. This leads to low positive predictive values and makes the system unreliable. Other authors describe sets of rules [29] but most of them lead to overalerting, notably because the rules are too simple and rarely involve segmentation conditions, i.e., conditions that are not directly responsible from the outcome but change its probability. In addition, those works do not support contextualization, i.e., the fact that the confidence of a rule varies deeply with respect to the place.

The rules discovered in this study mainly deal with the effects of anticoagulant drugs (35% of the rules) and hyperkalemia (27% of the rules). This paper also highlights the importance of pharmacokinetic drug-to-drug interactions (25% of the rules) that are often underestimated. Contrary to the accepted wisdom, many "important" ADE detection rules are not discovered by data mining because either the conditions never occur or, when the conditions are present, the outcome never occurs. This is probably because those rules are well known and, consequently, the risk is well monitored. However, it is possible to input the corresponding rules and enforce their automated evaluation.

D. Exploitation of the Results

Except the rules filtering performed by the experts, the whole process is fully automated. In order to analyze the data of a new hospital, about 1 h is required for 10^5 stays. The 236 rules and the related files (description of the mappings, lexicon, statistics computed in the automated evaluation, and textual explanations) consist of a set of XML files. The format of those files is well documented, so that it is easy to use them in any ADE detection or prevention application.

VII. Conclusion

This paper brings innovative and semiautomated solutions for ADE detection. The method is quite generic and could be applied to other kinds of data as soon as they are available in the EHR, such as structured results of electrocardiograms. The results of the method used here bring an important contribution to ADE knowledge. The rules that are obtained are versatile and can be used either as detection rules on past hospital stays, or

as prevention rules in a CDSS context. Those rules are already loaded in several prototypes that are developed in the frame of the PSIP Project.

1) A tool designed for retrospective ADE detection and follow-up in past hospitalizations: the Scorecards [46].
2) A knowledge-based system for prospective ADE prevention during the medication process, which is used by three CDSS: one embedded in a computerized physician order entry, another embedded in an EHR, and a prescription simulation tool that is available even without any Hospital information system.

REFERENCES

[1] L. T. Kohn, J. M. Corrigan, and M. S. Donaldson, *To Err Is Human: Building a Safer Health System*. Washington, DC: Natl. Acad., 1999.

[2] Institute of Medicine, *Preventing Medication Errors*. Washington, DC: Natl. Acad., 2007.

[3] T. K. Gandhi, D. L. Seger, and D. W. Bates, "Identifying drug safety issues: From research to practice," *Int. J. Qual. Health Care*, vol. 12, no. 1, pp. 69–76, Feb. 2000.

[4] D. W. Bates, R. S. Evans, H. Murff, P. D. Stetson, L. Pizziferri, and G. Hripcsak, "Detecting adverse events using information technology," *J. Am. Med. Inf.*, vol. 10, no. 2, pp. 115–128, Mar./Apr. 2003.

[5] T. Morimoto, T. K. Gandhi, A. C. Seger, T. C. Hsieh, and D. W. Bates, "Adverse drug events and medication errors: Detection and classification methods," *Qual. Saf. Health Care*, vol. 13, no. 4, pp. 306–314, Aug. 2004.

[6] R. Amalberti, C. Gremion, Y. Auroy, P. Michel, R. Salmi, P. Parneix, J. L. Quenon, and B. Hubert, "Typologie et méthode d'évaluation des systèmes de signalement des accidents médicaux et des événements indésirables," DRESS, Paris, France, Report, 2006.

[7] H. J. Murff, V. L. Patel, G. Hripcsak, and D. W. Bates, "Detecting adverse events for patient safety research: A review of current methodologies," *J. Biomed. Inf.*, vol. 36, no. 1–2, pp. 131–143, Feb./Apr. 2003.

[8] M. N. Cantor, H. J. Feldman, and M. M. Triola, "Using trigger phrases to detect adverse drug reactions in ambulatory care notes," *Qual. Saf. Health Care*, vol. 16, no. 2, pp. 132–134, Apr. 2007.

[9] M. Gysbers, R. Reichley, P. M. Kilbridge, L. Noirot, R. Nagarajan, W. C. Dunagan, and T. C. Bailey, "Natural language processing to identify adverse drug events," in *Proc. AMIA Annu. Symp. Proc.*, 2008, p. 961.

[10] G. B. Melton and G. Hripcsak, "Automated detection of adverse events using natural language processing of discharge summaries," *J. Am. Med. Inf. Assoc.*, vol. 12, no. 4, pp. 448–457, Jul./Aug. 2005.

[11] E. Aramaki, Y. Miura, M. Tonoike, T. Ohkuma, H. Masuichi, K. Waki, and K. Ohe, "Extraction of adverse drug effects from clinical records," *Stud. Health Technol. Inf.*, vol. 160 (Pt 1), pp. 739–743, 2010.

[12] J. Almenoff, J. M. Tonning, A. L. Gould, A. Szarfman, M. Hauben, R. Ouellet-Hellstrom, R. Ball, K. Hornbuckle, L. Walsh, C. Yee, S. T. Sacks, N. Yuen, V. Patadia, M. Blum, M. Johnston, C. Gerrits, H. Seifert, and K. Lacroix, "Perspectives on the use of data mining in pharmacovigilance," *Drug Safety*, vol. 28, no. 11, pp. 981–1007, 2005.

[13] J. S. Almenoff, E. N. Pattishall, T. G. Gibbs, W. DuMouchel, S. J. Evans, and N. Yuen, "Novel statistical tools for monitoring the safety of marketed drugs," *Clin. Pharmacol. Ther.*, vol. 82, no. 2, pp. 157–166, Aug. 2007.

[14] A. Bate and I. R. Edwards, "Data mining in spontaneous reports," *Basic Clin. Pharmacol. Toxicol.*, vol. 98, no. 3, pp. 324–330, Mar. 2006.

[15] C. L. Bennett, J. R. Nebeker, P. R. Yarnold, C. C. Tigue, D. A. Dorr, J. M. McKoy, B. J. Edwards, J. F. Hurdle, D. P. West, D. T. Lau, C. Angelotta, S. A. Weitzman, S. M. Belknap, B. Djulbegovic, M. S. Tallman, T. M. Kuzel, A. B. Benson, A. Evens, S. M. Trifilio, D. M. Courtney, and D. W. Raisch, "Evaluation of serious adverse drug reactions: A proactive pharmacovigilance program (RADAR) vs safety activities conducted by the Food and Drug Administration and pharmaceutical manufacturers," *Arch. Int. Med.*, vol. 167, no. 10, pp. 1041–1049, May 2007.

[16] D. M. Coulter, A. Bate, R. H. Meyboom, M. Lindquist, and I. R. Edwards, "Antipsychotic drugs and heart muscle disorder in international pharmacovigilance: Data mining study," *BioMed. J.*, vol. 322, no. 7296, pp. 1207–1209, May 2001.

[17] M. Hauben, V. Patadia, C. Gerrits, L. Walsh, and L. Reich, "Data mining in pharmacovigilance: The need for a balanced perspective," *Drug Safety*, vol. 28, no. 10, pp. 835–842, 2005.

[18] D. W. Bates, A. C. O'Neil, D. Boyle, J. Teich, G. M. Chertow, A. L. Komaroff, and T. A. Brennan, "Potential identifiability and preventability of adverse events using information systems," *J. Am. Med. Inf. Assoc.*, vol. 1, no. 5, pp. 404–411, 1994.

[19] A. K. Jha, G. J. Kuperman, J. M. Teich, L. Leape, B. Shea, E. Rittenberg, E. Burdick, D. L. Seger, M. Vander Vliet, and D. W. Bates, "Identifying adverse drug events: development of a computer-based monitor and comparison with chart review and stimulated voluntary report," *J. Am. Med. Inf. Assoc.*, vol. 5, no. 3, pp. 305–314, 1998.

[20] G. D. Fiol, B. H. Rocha, G. J. Kuperman, D. W. Bates, and P. Nohama, "Comparison of two knowledge bases on the detection of drug-drug interactions," in *Proc AMIA Symp.*, 2000, pp. 171–175.

[21] T. K. Gandhi, S. N. Weingart, A. C. Seger, J. Borus, E. Burdick, E. G. Poon, L. L. Leape, and D. W. Bates, "Outpatient prescribing errors and the impact of computerized prescribing," *J. Gen. Int. Med.*, vol. 20, no. 9, pp. 837–841, 2005.

[22] J. Judge, T. S. Field, M. DeFlorio, J. Laprino, J. Auger, P. Rochon, D. W. Bates, and J. H. Gurwitz, "Prescribers' responses to alerts during medication ordering in the long term care setting," *J. Am. Med. Inf. Assoc.*, vol. 13, no. 4, pp. 385–390, 2006.

[23] M. D. Paterno, S. M. Maviglia, P. N. Gorman, D. L. Seger, E. Yoshida, A. C. Seger, D. W. Bates, and T. K. Gandhi, "Tiering drug-drug interaction alerts by severity increases compliance rates," *J. Am. Med. Inf. Assoc.*, vol. 16, no. 1, pp. 40–46, 2009.

[24] A. Schedlbauer, V. Prasad, C. Mulvaney, S. Phansalkar, W. Stanton, D. W. Bates, and A. J. Avery, "What evidence supports the use of computerized alerts and prompts to improve clinicians' prescribing behavior?," *J. Am. Med. Inf. Assoc.*, vol. 16, no. 4, pp. 531–538, 2009.

[25] J. M. Teich, J. P. Glaser, R. F. Beckley, M. Aranow, D. W. Bates, G. J. Kuperman, M. E. Ward, and C. D. Spurr, "The Brigham integrated computing system (BICS): Advanced clinical systems in an academic hospital environment," *Int. J. Med. Inf.*, vol. 54, no. 3, pp. 197–208, 1999.

[26] G. J. Kuperman, J. M. Teich, D. W. Bates, F. L. Hiltz, J. M. Hurley, R. Y. Lee, and M. D. Paterno, "Detecting alerts, notifying the physician, and offering action items: A comprehensive alerting system," in *Proc. AMIA Annu. Fall Symp.*, 1996, pp. 704–708.

[27] B. Honigman, J. Lee, J. Rothschild, P. Light, R. M. Pulling, T. Yu, and D. W. Bates, "Using computerized data to identify adverse drug events in outpatients," *J. Am. Med. Inf. Assoc.*, vol. 8, no. 3, pp. 254–266, 2001.

[28] T. S. Field, J. H. Gurwitz, L. R. Harrold, J. M. Rothschild, K. Debellis, A. C. Seger, L. S. Fish, L. Garber, M. Kelleher, and D. W. Bates, "Strategies for detecting adverse drug events among older persons in the ambulatory setting," *J. Am. Med. Inf. Assoc.*, vol. 11, no. 6, pp. 492–498, 2004.

[29] S. M. Handler, R. L. Altman, S. Perera, J. T. Hanlon, S. A. Studenski, J. E. Bost, M. I. Saul, and D. B. Fridsma, "A systematic review of the performance characteristics of clinical event monitor signals used to detect adverse drug events in the hospital setting," *J. Am. Med. Inf. Assoc.*, vol. 14, no. 4, pp. 451–458, Jul./Aug. 2007.

[30] E. Chazard, B. Merlin, G. Ficheur, J. C. Sarfati, and R. Beuscart, "Detection of adverse drug events: Proposal of a data model," *Stud. Health Technol. Inf.*, vol. 148, pp. 63–74, 2009.

[31] *Patient Safety by Intelligent Procedures in Medication*. (Dec. 2010). [Online]. Available at http://www.psip-project.eu

[32] L. Breiman, *Classification and Regression Trees*. Belmont, CA: Wadsworth, 1984.

[33] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From data mining to knowledge discovery: An overview," in *Knowledge Discovery and Data Mining*. Menlo Park, CA: Amer. Assoc. Artif. Intell., 1996.

[34] N. Lavrac, "Selected techniques for data mining in medicine," *Artif. Intell. Med.*, vol. 16, no. 1, pp. 3–23, May 1999.

[35] R. Quilan, *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kauffman, 1993.

[36] B. D. Ripley, *Pattern Recognition and Neural Networks*. Cambridge, U.K.: Cambridge Univ. Press, 1996.

[37] H. P. Zhang, J. Crowley, H. Sox, and R. A. Olshen, *Tree Structural Statistical Methods. Encyclopedia of Biostatistics*. Chichester, U.K: Wiley, 2001, pp. 4561–4573.

[38] T. M. Therneau, B. Atkinson, and B. Ripley, *Rpart: Recursive Partitioning*, 2007.

[39] R_Development_Core_Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation Statistical Comput., 2008.

[40] R. Agrawal, T. Imielinski, and A. Swami, "Mining association rules between sets of items in large databases," in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, Washington, DC, May1993, pp. 207–216.

[41] E. Chazard, G. Ficheur, B. Merlin, E. Serrot, and R. Beuscart, "Adverse drug events prevention rules: Multi-site evaluation of rules from various sources," *Stud. Health Technol. Informat.*, vol. 148, pp. 102–111, 2009.
[42] M. Campos, J. Palma, and R. Marìn, "Temporal data mining with temporal constraints," in *Proceedings of the 11th Conference on Artificial Intelligence in Medicine*. Amsterdam, The Netherlands: Springer-Verlag, 2007, pp. 67–76.
[43] L. Sacchi, C. Larizza, C. Combi, and R. Bellazzi, "Data mining with temporal abstractions: Learning rules from time series," *Data Mining Knowl. Discov.*, vol. 15, pp. 217–247, 2007.
[44] J. M. Ale and G. H. Rossi, "An approach to discovering temporal association rules," in *Proc. ACM Symp. Appl. Comput.*, 2000, pp. 294–300.
[45] M. Stacey and C. McGregor, "Temporal abstraction in intelligent clinical data analysis: A survey," *Arif. Intell. Med.*, vol. 39, pp. 1–24, 2007.
[46] R. Marcilly, E. Chazard, M. C. Beuscart-Zephir, W. Hackl, A. Baceanu, A. Kushniruk, and E. Borycki, "Design of adverse drug events-scorecards," *Stud. Health Technol. Informat.*, vol. 164, pp. 377–381, 2011.

Emmanuel Chazard was born in Agen, France, in 1977. He received the M.D. and Ph.D. degrees in 2007 and 2011, respectively, from the Lille University of Law and Health Sciences, Lille, France.

He is currently a Physician at the Lille University Hospital, Lille, France. He is also a Teacher and Researcher at the Lille University, Lille, in the fields of public health, biostatistics, and medical informatics.

Grégoire Ficheur, photograph and biography not available at the time of publication.

Stéphanie Bernonville, photograph and biography not available at the time of publication.

Michel Luyckx, photograph and biography not available at the time of publication.

Régis Beuscart, photograph and biography not available at the time of publication.