Toward automatic detection and prevention of adverse drug events

Nicolas LEROY [a], Emmanuel Chazard [b], Régis Beuscart [b,]
Marie Catherine BEUSCART-ZEPHIR [a] and the PSIP consortium[1]
[a] *Evalab, INSERM CIC-IT 807, Faculté de Médecine – CHU de Lille, France*
[b] *CERIM, EA 2694, Université de Lille II, Lille, France*

Abstract. Adverse Drug Events (ADE) due to medication errors and human factors are a major public health issue. They endanger the patients' safety and cause considerable extra healthcare costs. The European project PSIP (Patient Safety through Intelligent Procedures in Medication) aims at identifying and preventing ADE. Data mining of the structured hospital data bases will give a list of observed ADE, with frequencies and probabilities, thus giving a better understanding of potential risks. The main objective of the project is to develop innovative knowledge based on the mining results and to deliver professionals and patients a contextualized knowledge fitting the local risk parameters in the form of alerts and decision support functions.

Keywords: Adverse Drug Event (ADE), Data Mining, Computer-Decision Support System (CDSS), Knowledge Elicitation

1. INTRODUCTION

1.1. Adverse Drug Events in the hospital setting

In the last ten years, Adverse Drug Events (ADE) have become a major public health issue [1]. Healthcare Information and Communication Technology applications should help reducing the prevalence of preventable ADE but their efficiency is impeded by the lack of reliable knowledge about ADE, and the poor ability of ICT solutions to deliver contextualized knowledge. This is aggravated by a poor consideration of causative human factors [2].

In the hospital context, Adverse Drug Events occur during the course of the Medication Use Process that describes the typical flow of action related to drug therapy [3]. The main steps of the medical use process are the physician diagnosis and prescription, the pharmacist verification and dispensation, and the nurse control and administration. An Adverse Drug Event (ADE) is "an injury caused by medical management rather than the underlying condition of the patient" [4]. Non preventable ADE consecutive to a normal use of a drug is usually distinguished from preventable ADE consecutive to an error. A medication error is characterized as a distance to "what should have been done" in the therapeutic care process [5]. This normative definition

---

[1] CHRU Lille F, CHU Rouen F, CH Denain F, Region H Hospitals of Copenhagen DK, Oracle France, IBM Acure-DK, Medasys France, Vidal France, Ideea Romania, Kite Italy, Aalborg University DK, Aristotle University of Thessaloniki GR, Umit Austria

may generate some difficulties. Besides the fact that the "normal" usage is sometimes difficult to define, considering any distance to the "normal" usage of a drug as an "error" is probably a simplistic conception of the reality. In the field of the psychological analysis of work, the distinction between the standard (normative) procedure and the real activity of the users allows for a better understanding of the work situation, where the goals set by the organization cannot always be reached by the operators using the standard procedures. In this approach, in order to manage risk, a distinction has to be made between voluntary and involuntary risk-taking [5].

The National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP) adopts a more pragmatic definition where a preventable ADE is the result of one or several dysfunctions distributed across the socio-technical system of the Medication Use Process [6]. An ADE can be described along several dimensions, like the severity of its consequences, the stage of the medical use process in which it occurred and the type of cause. The NCC MERP created a taxonomy which purpose is to provide a standard language and structure of medication error-related data for use in developing databases analyzing medication error reports [7].

1.2. Detection and prevention of Adverse Drug Events

In order to efficiently prevent ADE, it is mandatory to have a proper knowledge of these ADE. Retrospective analyses methods consist in assessing events such as accidents, incidents or near-misses. The objectives are the identification of the fundamental reasons, facts and causes that fostered the accidents or incidents [8]. These methods are efficient but highly time consuming and intrusive and it is sometimes difficult to generalize the results. Therefore the most common method remains the voluntary report of ADE by the healthcare professional, where the operators fill a structured form including a narrative description of the incidents.

Unfortunately in healthcare the rate of incident reporting due to the use of drugs is extremely low. Although the medication accidents / incidents declarations are compulsory, the users often hesitate to fill them, due to the lack of time but also by fear of the possible blame [9].

However, another reason could explain the low-level of reporting of Adverse Drug Event: the difficulties to detect them. In fact, the detection of serious accident in the industrial or transport domain is quite easy. The technical systems are supposed to function correctly and any disturbance to the nominal functioning can be considered as an incident. For example, the crash of an airplane is necessarily an accident, whereas the death of a patient may not be so: it could be the consequence of the natural evolution of his disease. To distinguish a potential ADE from a "normal" patient' symptom is not so easy. The progressive computerization of medical records along with the development and installation of electronic prescribing functions or complete Computerized Physician Order Entry (CPOE) systems has opened interesting opportunities for new methods of ADE detection.

1.3. The PSIP project - (https://www.psip-project.eu/)

The European project PSIP (Patient Safety through Intelligent Procedures in medication) aims at overcoming the problem of ADE detection by searching huge repositories of electronic medical records and data in order to detect abnormal cases presenting typical ADE features. The objective of the PSIP project is (1) to facilitate

the systematic production of epidemiological knowledge on ADE and (2) to ameliorate the entire medication cycle in a hospital environment.

The first sub-objective of PSIP is to innovatively produce knowledge on ADE: Data mining of the structured hospital data bases will provide a list of observed ADE along with their frequency and probability and patterns of statistical associatons, thus giving a better understanding of potential risks. Data Mining, also called Knowledge-Discovery in Databases (KDD) or Knowledge-Discovery and Data Mining, is the process of automatically searching large volumes of data for patterns using tools such as classification, association rule mining, clustering, etc.

The second sub-objective of the PSIP project is to develop innovative knowledge based on the mining results and to deliver to professionals and patients a contextualized knowledge fitting the local risk parameters in the form of alerts and decision support functions. This knowledge will be implemented in a PSIP-platform independently of existing ICT applications. These applications will connect to the platform to access and integrate the knowledge in their local systems. Considering the complexity of the health care professional's activity, the design and development cycle of the PSIP solution will be human factor oriented.

Traditional approaches of the problem of ADE detection are usually knowledge oriented. For example the starting point of ADE reports is the knowledge that a potential ADE has occurred. The PSIP project addresses the problem of ADE detection the other way round, and attempts to track back potential ADEs from the manifestation of their outcomes identified via mining techniques. Then one of the most important challenges of the project is the validation and interpretation of the data and semantic mining results.

The relevance of the results provided by the data mining is critical for the proper functioning of the project. Indeed, results with a too large proportion of atypical cases turning out not to be actual ADE would make the development of alerting and decision support functions almost impossible. It is therefore necessary to closely monitor the validation and interpretation of data mining results and to set specific methods for this important knowledge elicitation phase. This will require the participation of groups of experts in charge of (1) assessing the adequacy of the rules for automatic selection of atypical records susceptible to be ADE-related (2) producing the necessary knowledge to characterize these ADE and to feed the decision support rules of the PSIP platform.

The human factors specialists participating in the PSIP project will both support and monitor the experts' activities while assessing the selection rules of potential ADE-related records and characterizing these ADE. The objective is to understand the experts' reasoning and the parameters or data they rely on while interpreting or validating the ADE cases. This information should help iteratively refining the data mining procedures and rules.

2. METHODS

2.1. Atypical medical records selection

2.1.1. Data model

One year of medical records' archives are extracted from different French and Danish hospitals repositories and analyzed by data and semantic mining techniques. The

atypical records are selected according to the characteristic of a data model specifically designed for extraction and mining purposes, characterized by 72 fields grouped into 7 main categories: (1) Administrative information (patient, flows), (2) Medical diagnosis, (3) Medical procedures, (4) Drug prescriptions (5) Biology results (6) Reports and letters.

2.1.2. Data-mining rules

Data mining techniques allow getting association rules describing the statistical link between several causes or contexts and an effect. Several different effects can be traced. The nature of some effects or the fact that some drugs appear as causes or contexts can often be interpreted as the possibility of an adverse drug event. For example the following descriptors should contribute to the characterization of a stay as abnormal:

- Specific sequences of steps of the stay, like a transfer from a standard medical unit to an intensive care or resuscitation unit in the middle of the stay without any surgical procedure before
- A duration of the stay longer than the expected duration when considering the patient's Diagnosis Related Group (DRG)
- Death of the patient while the probability of death of his DRG is low
- The fact that the stay crosses different medical specialties, etc.

The data mining process provides several decision rules that can be expressed under the following format:

{patient elder than 75} AND {vitamin K antagonist} AND {another drug having enzyme inhibition side effect} => higher probability of death.

Each rule is characterized by:

- its support (number of previous stays matching the conditions and having the effect)
- its confidence (probability of having the effect once the conditions are met).

An important point is that the support and the confidence may vary between two different medical departments and/or different hospitals. The contextualization of the statistical link appears as a very important feature. In each department those rules have to be filtered to make sense and to limit their number. Confidence thresholds have to be carefully tuned to obtain relevant and reliable rules.

As the results of the data mining process can be expressed under the form of rules, and as these rules can be weighed by confidence parameters, it is possible to use these rules as the basic foundation for the Decision Support System aimed at reducing the number of Adverse Drug events. The contextualization of the rules is obtained through the application of different weights to identical rules, or by the identification of specific rules.

2.2. Analysis and validation with the expert group

An expert group, composed of pharmacologists, pharmacists and physicians is asked to review the results obtained by the knowledge rules. They have to characterize two types of stays: (1) stays connected with knowledge rules, (2) stays not connected with knowledge rules. The experts have access to the medical record of the stays in order to infer the presence of Adverse Event, ADE and Preventable ADE. The main objective of this evaluation is to validate the accuracy of the knowledge rules for the detection of ADE. The experts are asked to analyze and interpret the atypical cases selected by the

data mining in order to (i) support the refinement of the data model and data mining rule (ii) issue usable knowledge to feed the decision support functions of the PSIP platform. Specialists in cognitive ergonomics provide methods to support this knowledge elicitation task, relying on the "think aloud" method to record the experts' reasoning processes.

3. RESULTS

In the PSIP project, the data mining is currently in progress but some preliminary results demonstrate the feasibility of the method and its potential to deliver a contextualized knowledge on Adverse Drug Events. In this section, we present an example of knowledge discovered from the analysis of medical records by means of decision trees methods. These first results have been obtained from the data mining of 2700 records from cardiologic units of the Region H Hospitals (Copenhagen, DK). The results are expressed under the form of association rules. We give here two rules as examples.

Rule 1:
{Drug: Vitamin K antagonist} AND {Drug: Prokinetic} => Appearance of a too low INR
Rule1 characteristics: Support: 4; Confidence: 67%
This means that 6 stays match the conditions and four of them present the effect (67% = 4/6)
Outcomes:
Unexpected death 16.67%
Average duration of the stay: 15 days (the ordinary mean duration of stay for this type of patient is 6.5 days)

Rule 2:
{Drug: Vitamin K antagonist} and {Drug: antibiotic = betalactamin} and {age < 76 years} => Appearance of a too low INR
Rule 2 characteristics: Support: 3; Confidence: 60%
This means that 5 stays match the conditions, 3 of them present the effect (60%=3/5)
Outcomes:
Death: 0%
Average duration of the stay: 12.6 days (ordinary mean duration: 6.0 days)

At this stage of the project over 150 rules have been obtained by mining 2 different data bases from Danish and French hospitals. The rules are under validation process, and about 95% of the already reviewed rules have been validated. The experts' review of the stays attached to the rules is in progress.

4. DISCUSSION AND CONCLUSION

The current identification rate of ADE through reporting systems is too low to support an efficient prevention of these ADE. Computerized-based screening of electronic medical records is considered an interesting alternative method to identify

ADE [10] but current research suffers from low specificity in the identification of ADE and would therefore issue too general, non context-related potential alerts or DSS rules. The present research project PSIP is oriented by the strong hypothesis that data and semantic mining may allow to identify a significant proportion of abnormal cases potentially due to ADE, along with the characteristics of their context of occurrence. The preliminary results of the data mining performed on two groups of hospitals from two different countries look promising, as the association rules seem able to catch the context of occurrence of the identified ADE. However, in order to turn these retrospective data into prospective CDSS functions aiming at preventing those ADE, it is necessary to:

- Properly review and validate the association rules elicited by the data mining procedures
- Review the abnormal stays attached to these rules and validate their ADE status, as compared to a sample of "normal" stays not positively screened by the data mining process
- Analyze the corresponding work system relying on a Human Factors (HF) approach in order to identify HF potential root causes of the identified ADE. This analysis is necessary to design acceptable and usable alerts or DSS functions aiming at preventing the ADE.

The objectives of the PSIP project are ambitious, but the success of such a project would significantly contribute to patient safety by detecting and preventing a significant part of potential ADE. Human factors and ergonomics competencies are critical to enhance the chances of the project to reach its objectives.

5. [2]Reference List

[1] Kohn LT, Corrigan JM, and Donaldson MS To Err is Human: Building a Safer Health System, 1999.
[2] J.W.Gosbee, Conclusion: You need human factors engineering expertise to see design hazards that are hiding in "plain sight!", Jt.Comm J Qual Saf. 30 (2004) 696-700.
[3] N.H.a.N.Stacy Ackroyd-Stolarz, Approaches to Improving the Safety of the Medication Use System, Healthcare Quarterly 8 (2005) 59-64.
[4] Aspden P., Wolcott J.A., Bootman J.L., Cronenwett L.R. Eds, Preventing Medication Errors, IOM, The national Academic Press, Washington DC, 2007
[5] ANAES Principles of risk management in healthcare organisations, 2003. downloaded Nov 2008 at:
http://www.has-sante.fr/portail/upload/docs/application/pdf/vvjun05-guidegrisques-entofinal_2007_03_07__10_02_13_247.pdf
[6] National Coordinating Council for Medication Error Reporting and Prevention. What is a Medication Error? 2008. 4-9-2008.
[7] NCC MERP. NCC MERP Taxonomy of Medication Errors1. 2008. 4-10-2008.
[8] P.C.Cacciabue Guide to Applying Human Factors Methods, SpringerVerlag, 2004.
[9] A.Figueiras, F.Tato, J.Fontainas, B.Takkouche, and J.J.Gestal-Otero, Physicians' attitudes towards voluntary reporting of adverse drug events, J Eval.Clin.Pract. 7 (2001) 347-354.
[10] A.C.Seger, A.K.Jha, and D.W.Bates, Adverse drug event detection in a community hospital utilising computerised medication and laboratory data, Drug Saf. 30 (2007) 817-824.