Author's response to reviews

Title: Adverse drug events with hyperkalemia during hospitalisation: Evaluation of an automated method for retrospective detection in hospital databases

Authors:

Grégoire Ficheur (gregoire.ficheur@univ-lille2.fr)
Emmanuel Chazard (emmanuel.chazard@univ-lille2.fr)
Béatrice Merlin (bmerlinsp@gmail.com)
Jean-Baptiste Beuscart (jean-baptiste.beuscart@univ-lille2.fr)
Michel Luyckx (mluyckx@ch-denain.fr)
Régis Beuscart (regis.beuscart@univ-lille2.fr)

Version: 2
Date: 23 May 2014

Author's response to reviews: see over
Dear Dr. Chap,

Please find enclosed the revised version of our article entitled: “Adverse drug events with hyperkalemia during hospitalization: Evaluation of an automated method for retrospective detection in hospital databases” for submission to BMC Medical Informatics and Decision Making. The revised version has been significantly modified.

In line with all of the reviewers’ comments, we have modified the manuscript in order to improve the quality of presentation, analysis of data and contents of the manuscript. These modifications are based on all specific comments to which we have responded.

Point-by-point responses to the Editor and Reviewers are given below.

Dr Grégoire Ficheur

Lille, Mai 15th 2014

Arlene Pura, PhD
Executive editor
BioMed Central
236 Gray’s Inn Road
London,
WC1X 8HB
UK
Responses to executive editor

The revised paper needs to be clearer about the novelty of what is achieved. This means both a better account of how the work relates to ADRs and ADEs and details of what is novel, is it the approach by which the rules are obtained, or the rules themselves?

We deeply apologize that our manuscript insufficiently separated the two problems as pointed out by the reviewer. In synthesis (these points are detailed below):

- Regarding the first point, our detection rules deal both with adverse drug events in general, i.e. the preventable ADEs (like medication errors) and the ADRs. Our rules do not distinguish these 2 types of ADEs as the dosage of drugs is not a field of the rules.
- Regarding the second point, the novelty of our work concerns primarily the rules themselves rather than the method of induction of these rules: our rules have 3 parameters including (1) cause conditions (contextual parameter AND a drug prescription), (2) an outcome (hyperkalemia) AND (3) the chronology between these cause conditions and the outcome. The set of rules we propose is built firstly by pharmacists and pharmacologists experts then secondarily optimized by testing them on anterior years of our partner general hospital. These rules are mainly expert rules.
MCR1: Definitions of ADRs and ADEs are provided as background information. It is not clear, however, what exactly the underlying method is capable of detecting. Is it rules for detecting preventable ADEs (i.e. not ADRs), is it ADEs in general? How can this be specified in rule-induced methods? If it is only on type of ADEs, it is suggested to shorten the relevant paragraphs, because they may be quite confusing for the reader. Citations to widely accepted definitions of these concepts shall be sufficient. Likewise, please, clarify and potentially revise appropriately the respective text (lines 70-72).

ADRs and ADRs' definition is a complicated topic, we tried to explicit each part of these definitions in our initial version, but we probably didn’t gain in clarity. We think that it is necessary to define both ADRs then ADEs as ADEs extend the initial definition to medication errors which are most of the time “preventable ADEs”. These "preventable ADEs"are particularly interesting for obvious reasons. Regarding the definition, we decided not to keep the part of the definition detailing the main types of ADRs.

On page 3 of the previous manuscript, we deleted the following paragraph:

“ADRs can be classified into two main types:

Type A reactions are dose-dependent events related to the exaggeration of a known pharmacological effect; they are frequent and predictable. Drug interactions are also included.

Type B reactions are stochastic, dose-independent events and non-reproducible from one individual to another, as opposed to type A ADRs. They represent rare ADRs and are unpredictable. It is considered that patients who develop these idiosyncratic events most often have a qualitative immunological (immunoallergic reactions) or genomic [4] (e.g. variability of cytochrome P450 [5, 6]) specificity.”

Regarding the ability of our rules, they were built to detect ADEs in general (i.e. preventable ADEs or ADRs). These rules do not try to distinguish the 2 types of ADEs. If we would have had this goal, the dosage of drugs could have been an interesting value to specify rules targeting medication errors. After having defined ADRs then ADEs, we decided to use the label “ADEs” for the rest of the paper.

On page 4 line 72 of the previous manuscript, we completed the sentence to obtain: "For this reason, the present work deals with ADEs in general (including preventable ADEs and ADRs)."
MCR2: It is considered essential to explicitly state what the novelty of the proposed detection rules is. For example, can the 18 rules presented in the study be found in commercial pharmacovigilance databases?

The novelty of our work concerns primarily the rules themselves rather than the method of induction of these rules: our “complex detection rules” go further than those described in literature as they include all the following conditions:

1. cause conditions (contextual parameter AND a drug prescription)
2. outcome (hyperkalemia)
3. the chronology between these cause conditions and the outcome

The set of rules was built by pharmacists and pharmacologists experts then secondarily optimized by testing them on anterior years of our partner general hospital. These rules are mainly expert rules. All contextual parameters used to build rules are known as favoring the occurrence of hyperkalemia and the 6 retained drugs are well known as potential cause of adverse drug events with hyperkalemia. So, individually the parameters could be in pharmacovigilance database but the combination of all these parameters including respect of chronology are not issued of a pharmacovigilance database, it is the result of expert consensus.

According to the reviewer’s comment, we added on page 9 on the previous manuscript line 236: "The set of rules has been elaborated by experts including pharmacists and pharmacologists who aggregated different sources of knowledge. These rules were then optimized by testing them on previous dataset of inpatient stays."

MCR3: Please clarify whether there is any relation between the data used for evaluation and the data used to extract the detection rules

As explained above, these rules were mainly built by expert consensus. In a second phase, these rules were optimized by testing them on previous dataset (year 2009).

MCR4: The structure of the employed rules (presented in line 100) has to be explained. What is a Cause in the general case? Why only AND is foreseen as an operator linking Causes?

We agree with the reviewer that this point has to be specified as it is the main point of our work. Both a contextual context (diabetes, age>70years or renal failure) AND a drug favoring hyperkalemia have to be present for the cause of the rule. Then, the outcome has also to be present (in these retrospective rules) respecting a delay of 5 days after the presence of cause.

According to the reviewer’s comment, we modified on page 9 on the previous manuscript line 236, we added: "The presence of "cause conditions" means that a drug AND a context favoring the hyperkalemia have to be fulfilled in the 5 days time period prior to the outcome."
**MCR5:** The selection of the “5 days” threshold in the “chronology of the rules” is not self-evident. A justification has to be provided.

This threshold was an expert consensus, we share the reviewer’s opinion: It is necessary to specify this choice. The experts chose this delay to take into account the drugs having the longest half-time for elimination. The half-time for elimination (which impacts the delay to obtain the target dose in blood and which impacts the delay for elimination of the drug when the drug is stopped) is heterogeneous among the six drugs we retained:

- renin angiotensin system inhibitor (half life: 6-24 hours)
- potassium sparing diuretic (half life: 5.6-14.8 hours)
- beta blocker (half life: 4-5 hours)
- high molecular weight heparin (half life: 1.5 hours)
- non-steroidal anti-inflammatory (ibuprofen half life: 1.3-3 hours)
- (potassium chloride)

Further developments could implement a specific delay for each ATC code but we did not easily obtain this information for all drugs.

According to the reviewer’s comment, we modified on page 9 on the previous manuscript line 227, we added: "This delay seems to be coherent to take into account the drugs having the longest half-life like some renin angiotensin system inhibitors or some potassium sparing diuretics."

**MCR6:** It is important to refer to recent studies presenting and evaluating methods for drug safety risk detection exploiting observational healthcare data. Also, as the authors concentrate on the "strength" of laboratory examination results for detecting ADEs, reference to relevant studies could be provided as well (e.g. [Liu M, et al. Comparative analysis of pharmacovigilance methods in the detection of adverse drug reactions using electronic medical records. JAMIA (2013) 20(3):420-426]). Overall, the bibliography of the paper shall be updated with more recent works.

Our bibliographic work has been completed and references have been added according to the remarks of the reviewer.

We can distinguish 3 aspects:

1) enrichment of references describing the use of observational databases including datamining methods

- references added:


• on page 4 on the previous manuscript line 90, we replaced the text "For this reason, it becomes valuable to obtain objective data, such as computerized hospital data including data relative to drug administration, laboratory results, administrative data, etc." by "For this reason, it becomes valuable to obtain objective data, such as computerized electronic medical records including drug administration, laboratory results, and administrative data to explore alternative signal discovery approaches such as data-mining methods [13–17] among EMR which seems to be a major resource for observational post marketing analyses."

2) justification of recent studies on these observational databases confirming the difficulty to have a low rate of false positive

• reference added:


• on page 6 on the previous manuscript line 130 we added:"A series of empirical assessments that have been conducted [26] revealed that high false-positive rates are still a major limitation of all methods."

3) justification of use of laboratory results as trigger of ADEs

• reference added:
• We added this reference in discussion section to complete the paragraph explaining the interest of laboratory results on page 14 of the previous manuscript line 345: "Our work illustrates the role of laboratory results in the automated detection of ADEs. Laboratory results were used as conditions and as outcome of these “complex detection rules”: they seem to be pertinent indicators for detecting ADEs and this is particularly true for hyperkalemia as it is by itself a complication. Moreover, cardiac symptoms occur secondarily to this “abnormal laboratory test”, meaning that this last approach may be more sensitive. This is in agreement with the review conducted by Handler et al.[17] who identifies 36 unique ADE signals, including 10 medication levels, 19 laboratory values and 7 antidotes. Laboratory results are indeed available during inpatient stays and they constitute structured data."
**ANSWERS to Reviewer 2, Shobha Phansalkar.**

**CR1:** This section seems incomplete with only one line and one reference that too from a study in the US when the current study was conducted in France. Authors are encouraged to draw a stronger background for conducting the study.

We share the opinion it is necessary to enrich this part, the two following references were added:

- A Swedish study which seems to be one of the largest evaluation on hospital data in Europe [ref]
- The main French survey (realised on large sample of French hospitals) "ENEIS"

It is important to note that these two references deal with the incidence of ADEs during hospitalisation (it doesn't include the ADEs as the reason for hospitalisation).

According to the reviewer's comment, we modified on page 3 of the previous manuscript line 47, we added: "(..) and could account for 5% of hospital deaths [3] in a Swedish hospital. Incidence of ADEs occurring during hospitalisation (according to ENEIS national French survey) finds an incidence of 7.6 severe ADEs per 1000 inpatient-days and 3.0 severe ADEs per 1000 inpatient-days are preventable ADEs."

**CR2:** Clinico-Biological context is an interesting term and I understand how it is used in this context. Since its use is prevalent throughout the text and because I have not seen it used commonly, at least in the US literature would the authors provide a definition/description of what it means?

Our rules are able to take into account the context in which the drug is prescribed and the parameters allowing to describe the patients associated clinical data (diagnoses or symptoms detected with ICD-10 codes of the database) and lab test results (detected with C-NPU codes of the database). We think that ADEs occur in a context favouring this kind of outcomes.

According to the reviewer's comment, we added on page 6 of the previous manuscript line 148: "(..) complex clinical-biological context, associating clinical data (mainly diagnoses and symptoms) and lab test results. Our hypothesis is that this context could favour the occurrence of ADEs."
CR3: The three main approaches [28] are: “probabilistic approaches”, “expert judgment”, and “algorithm” (what does this mean, building algorithms is not an approach but a means to a computational method?): The probabilistic approach is the most reproducible, though it is not usable on a routine (basis?/computational routine?)

These methods are intended to validate cases individually. Probabilistic approach can be for example the result of a multivariate model given Odds Ratio for each parameter identified during the inpatient stay. The most classical approach is the “algorithmic-based approaches” as algorithm of Naranjo, algorithm of Kramer or algorithm of Bégaud, these 3 algorithms are similar. We could maybe have used the label "scale approach".

To clarify the labels we changed on page 7 of the previous manuscript line 157 "The three main approaches [28] are: “probabilistic approaches”, “expert judgment”, and “algorithm”: by “The three main approaches to validate cases individually [34] are: “probabilistic approaches”, “expert judgment”, and “algorithmic-based approaches according to Naranjo [35], Kramer [36] or Bégaud [37]”:"

CR4: (...) "non-steroidal anti-inflammatory" and "high molecular weight heparin" are six drugs retained among rules—some of these are drug classes and not single drugs. I imagine that both generic and brand names were taken into account and that some medication knowledge base was utilized in order to identify synonymy of medication names/ memberships within classes.

We think we explained it in the paragraph "Types of data available for building rules and analysis" page 8 of the previous manuscript on line 192. This paragraph (including a modification of formulation) is "The conditions used for building detection rules are aggregated variables, i.e. groups of codes built by an expert. For example for drugs, the variable “Potassium sparing diuretic” includes all the ATC codes that are compatible with this kind of drug. Similarly for laboratories, the variable “Hepatic cytolysis” includes all the abnormal laboratory tests (using C-NPU) that are compatible with this syndrome (Alanine transaminase OR Aspartate transaminase > 3 times the upper normal limit), and “urinary infection” includes all the ICD-10 codes that are compatible with this diagnosis."

All the variables used in the rules are the result of aggregation of several parameters. Thus, each one of the 6 drugs retained among rules are essentially part of drug classes which have similar pharmacodynamics properties, and the variable is present if at least one of the corresponding ATC codes is found.

According to the reviewer’s comment, we added on page 9 line 212 of the previous manuscript: "These drugs or drug classes are built by aggregation of list of ATC codes as explained above".

9
In keeping with the comment above, what medication knowledge base was used? Was it a vendor KB or one that the authors build in-house for the purposes of this study?

The set of rules we propose was firstly built by pharmacists and pharmacologists experts and secondarily optimized by testing them on anterior years of our partner general hospital. These rules are mainly expert rules. All contextual parameters used to build rules are known as favoring the occurrence of hyperkalemia and the 6 retained drugs are well known as potential cause of adverse drug events with hyperkalemia. So, individually the parameters could be in pharmacovigilance database but the combination of all these parameters including respect of chronology are not issued of a pharmacovigilance database, it is the result of expert consensus.

According to the reviewer’s comment, we modified on page 9 on the previous manuscript line 236, we added: "The set of rules has been elaborated by experts including pharmacists and pharmacologists who aggregated different sources of knowledge. These rules were then optimized by testing them on previous dataset of inpatient stays."

Could the authors comment on the EHR that was being utilized?

In a first step, the data are extracted from the EHR of the partner hospital. The EHR of the partner hospital is not used for the review. The computerized tool used for the review is the "ADE scorecards” [33. Chazard E, Băceanu A, Ferret L, Ficheur G: The ADE scorecards: a tool for adverse drug event detection in electronic health records. Stud. Health Technol. Inform. 2011, 166:169–179.], the tool contains all the data used for the study and facilitates the review of inpatient stays as shown on figure 2. The main point of this tool is its ability to make a visual link between the day of occurrence of the outcome and the day of prescription of the drugs.

![Figure 1. Scorecard of hyperkalemia (K^+>5.3)](image1)

![Figure 2. Main screen of the stay review facility](image2)
CR7: How was the time period of 5 days chosen for the development of the rule? Expert consensus, if so, please mention this?

This delay of 5 days is the result of an expert consensus according to the following idea: They chose this delay to take into account the drugs having the longest half-time for elimination. The half-time for elimination (which impacts the delay to obtain the target dose in blood and which impacts the delay for elimination of the drug when the drug is stopped) is heterogeneous among the six drugs we retained:

- renin angiotensin system inhibitor (half life: 6-24 hours)
- potassium sparing diuretic (half life: 5.6-14.8 hours)
- beta blocker (half life: 4-5 hours)
- high molecular weight heparin (half life: 1.5 hours)
- non-steroidal anti-inflammatory (ibuprofen half life: 1.3-3 hours)
- (potassium chloride)

Further developments could implement a delay specific for each ATC code but we didn’t have easily this information for all drugs.

According to the reviewer’s comment, on page 9 on the previous manuscript line 227, we added:

“This delay seems to be coherent to take into account the drugs having the longest half-life like some renin angiotensin system inhibitors or some potassium sparing diuretics.”

CR8: Would all conditions (causes as the authors call them) have to be fulfilled in the 5 day time period prior to the outcome for the rule to be fulfilled? What if one or more conditions were not met? I see below the authors say 3 conditions but is each “bullet” a separate condition? What is patient were on 2 drugs and was > 70yrs old?

If a patient is more than 70 years old AND receives two drugs (among the six drug classes) AND AFTER presents the outcome so two rules are present. If the patient receives only one (or two) drug(s) and does not present a favoring context AND AFTER presents the outcome, then no rule is present.

CR8: If the death is the ADE and cannot be directly attributed to the drug then how can we link the cause and the condition? Perhaps precision is even lower than 3.7% for the serious ADEs since out of 3, 2 were deaths that did not have direct causality.

We wanted to be extremely prudent concerning the responsibility of our partner hospital for these ADEs, but they are “probable” or “definite” according to Kramer’s algorithm.

According to the reviewer’s comment, we deleted on page 13 on the previous manuscript line 320, we deleted a part of this paragraph “the drug played a role, but death cannot be directly attributed to the drug.”
CR9: How was causality determined for the chart review? Was causality determined using the Naranjo scale?

The scale (algorithm) used in our paper is "Kramer algorithm". We think this part of our methodology is detailed in method section on page 10 of our previous manuscript on line 250: "The task is carried out by a physician expert. This work is performed while ignoring the results obtained by an automated analysis of inpatient stays. The drug causality assessment algorithm used in this study is Kramer’s algorithm [36]."

CR10: The precision of 63% maybe better than then 3 studies that the authors chose to cite but is still not great. This means that 4 out of 10 cases that the system would identify would be false positives. That is a very high false positive rate and may not garner much trust in the system over time. Could the authors comment on this?

The main study allowing to compare our result is the study of Dormann as he is the only one who calculated the recall and precision for the dyskalemia [23. Dormann H, Criegee-Rieck M, Neubert A, Egger T, Levy M, Hahn EG, Brune K: Implementation of a computer-assisted monitoring system for the detection of adverse drug reactions in gastroenterology. Aliment Pharmacol Ther 2004, 19:303–309.]. In method 1, for a similar recall (91%), the precision is 36%. In method 2, for a similar precision (67%), the recall is 40%. For us, these results could illustrate the necessity to take into account recall and precision to compare results.

We agree with the idea that it is still not great but the main point for us was to have great recall as it is dangerous to miss hyperkalemia cases which are dangerous situations for patients.

CR11: Could the authors suggest what else could be done to improve performance of a computational approach for detection of hyperkalemia so the precision and recall can be improved?

According to the review we conducted, our detection rules have a high level of complexity and we think that a higher level of complexity could improve the precision, giving a better focus, but could decrease the recall. However, two suggest can be done to improve performance of computational approach:

1. The delay of 5 days between the causes and the outcome could be specified for each ATC code, it would permit to fit the biochemical behaviour of the drug
2. Most of the cases that were not confirmed by experts are cases in which a diagnosis or clinical symptom is sufficient to explain the outcome. Some rules could try to implement this information excluding cases having evident clinical cause. However,
the ADEs often occurs in cases where clinical aspects modify the pharmacokinetics of the drug, so it is not easy to distinguish it in the rule a priori.