Author's response to reviews

**Title:** A Pipeline To Extract Drug-Adverse Event Pairs From Multiple Data Sources

**Authors:**

SriJyothsna Yeleswarapu (srijyothsna@atc.tcs.com)
Aditya Rao (adityar.rao@tcs.com)
Thomas Joseph (thomas.joseph@tcs.com)
Vangala G Saipradeep (saipradeep@atc.tcs.com)
Rajgopal Srinivasan (raj@atc.tcs.com)

**Version:** 4  **Date:** 26 November 2013

**Author's response to reviews:** see over
Response to Reviewer 1

Major Revisions

1. Authors should pay carefully attention to the terminology being used. In particular, this study cannot in no way identify adverse drug reaction (implying causality assessment) but rather adverse event (based only on statistical association). I suggest reconsidering the tile and the objective of the pipeline. Moreover, correct definition of safety signal should be taken into account. Signal concerns a new or incompletely documented drug-adverse event association. Thus, by definition, a known-adverse event association should not be considered as signal if found through the pipeline.

The terminology has been corrected and we now use the term “adverse event”, implying only a statistical association between the concepts.


Based on all these ongoing initiatives, I suggest revising the statement “EHR are either not easily accessible due to privacy concerns or are very complex to understand”

Agree. The necessary changes have been made to describe some international initiatives which have been exploring electronic medical records for safety signal detection and have included the suggested references and others.

3. Objective of the study should be more clearly described

Done.

4. Methods: what is the rationale for choosing the study drugs apart from being cited in other papers? Are then results generalizable to other drugs belonging to different drug classes? Authors should comment on it. In addition, the types of medication should be revised: carbamazepine is antiepileptic drug and not antipsychotic either antidepressant;

In would distinguish between antidepressants (trazodone, and paroxetine) and antipsychotics (olanzapine and ziprasidone; why not including any old typical antipsychotics which are in the market since longer period?)

Rationale explained in the paper. Terminology used as per Wang et al. and Leaman et al. The pipeline has not been built for any particular drug or particular class of drugs.
5. Results: why results has been reported only for bupropion? Authors should describe results for all the study drugs trying to use visualization which may summarize results. The disagreement between different sources (as well as additional value of each source) should be better analysed and discussed. In general results should be described using more quantitative and measurable parameters.

The common trends observed from top scoring results results for the other drugs used in this study as listed in Table 4 and discussed. Also, the noticeable results for the top scoring adverse events across the three sources for anti-depressant/anti-psychotic class of drugs used in this study, which are prescribed for neurological problems are listed in Table 5. The supplementary material contains the comparative analysis of the pipeline results with the label information for some of the drugs, along with the top scoring results for the remaining drugs.

6. Table 1. Some important safety concern are not listed: a) stroke and all-cause mortality for olanzapine; b) sexual disorders for paroxetine. The list of safety concerns for each drug should be revised by experts in pharmacovigilance. Why no know AE has been reported for morphine?

We re-checked the known safety issues for all the drugs

Minor revisions

1. Introduction: replace “correlation between drugs...” with “causal association between drugs...”

2. Introduction: ADAMS does not fully relate to the topic


5. Methods: indicate from which section of label of bupropion the information on safety have been extracted

All Done.
**Response to Reviewer 2**

**Major Revisions**

1. In the state of art section the authors missed many projects and scientific papers quite relevant in this subject. The section starting with "Various studies have looked at mining potential AEs from unstructured text sources." lacks a deeper overview. The authors need to expand on the findings of other groups in this respect, for example the EU-ADR project, the Mini-Sentinel, the Observational Medical Outcomes Partnership (OMOP) and several others. The discussion section should highlight also in what way authors' methodology is able to achieve better results than previous solutions.

*Agree. The necessary changes have been made to include the projects and scientific papers relevant to the subject. The suggested references are now discussed.*

2. The unstructured data were obtained through web crawling of three health-related websites. However it is not described how drug-event associations were obtained. I guess that many symptom-drug associations will be gathered through this method. The same may occur in PV-TPX.

*A detailed explanation of how the drug-event associations were obtained is now provided in the Methods section.*

3. Despite the paper is well written and organized, it includes many references to websites, Wikipedia (e.g. REST, SOAP, ..). Even Wikipedia is included. I recommend a more scientific approach for references.

*The generic references are replaces with more scientific publications, wherever possible. However, the references to certain website hyperlinks are still retained, wherever necessary.*

4. The pipeline is well described and it seems adequate. However, besides the NER part, which is a previous work from authors, it uses some known data and methods (FAERS, BCPNN algorithm). The authors should emphasize the novelty of their approach.

*The previous work on NER (TPX) described in this paper, does not use FAERS or BCPNN. It contains the modules for concept identification and pairwise concept associations and was previously implemented and used only for biomedical literature in MEDLINE database. The novelty of this study lies in resuing these components for the other sources like user comments from health blogs. The statistical methods like BCPNN used in this study to analyze and bring out the differences in the data from various sources are also novel. The necessary changes have been made to emphasize the novelty of the approach.*

5. The manuscript could be improved by performed a thoroughly validation of the obtained results. As it is now,
the results are presented in a descriptive way showing events from the several methods (table 2).

The common trends observed from top scoring results results for the other drugs used in this study as listed in Table 4 and discussed. Also, the noticeable results for the top scoring adverse events across the three sources for anti-depressant/anti-psychotic class of drugs used in this study, which are prescribed for neurological problems are listed in Table 5. The supplementary material contains the comparative analysis of the pipeline results with the label information for some of the drugs, along with the top scoring results for the remaining drugs.

6. Finally, despite part of the system is available for registered users, I would suggest to have a set of guess accounts for this kind of test.

The system requires a login, which cannot be bypassed. However, for convenience and review purposes, three guest accounts have already been created and provided to the publishers with the following login credentials:
guest1/guest1; guest2/guest2; guest3/guest3

Response to Reviewer 3

Major Revisions

1. Relation to previous work and to state-of-the-art: The work does not appropriately take into consideration the published literature on the topic of automated detection of adverse drug reactions from text (e.g. PMID:23935003, 23256479) and on existing corpus of adverse drug reactions (e.g. PMID: 22554702; LREC2010-2nd Workshop on Building and Evaluating Resources for Biomedical Text Mining). Furthermore they also cited only Leaman 2010 paper for analysis of ADR on social media, although the published literature in the field is much larger (e.g. Ahmed Abbassi from Virginia University grant http://news.virginia.edu/content/research-sift-social-media-early-signs-adverse-drug-reactions; the european WEBAE initiative http://www.imi.europa.eu/webfm_send/912). It is essential that the authors reconsider their work in view of all of the above, and provide clear description on how they go beyond/contribute to state of the art.

The necessary changes have been made to include the projects and scientific papers relevant to the subject.

2. Precision/recall of their tagging technology for adverse drug reactions The publication that is used as reference for the NER tagging technology (TCS PubMed Explorer) does not provide quantitative measures of the precision/recall for any of the entities involved.

The complete TPX paper will have a quantitative measures and evaluations against known corpora.
3. Legal implications in webscraping The authors do not even address the legal implications of web scraping techniques (e.g. use of web-harvest) and particularly on patient / health information. Furthermore they do not elaborate on the EU and USA normatives on the topic.

Has been answered with a completed paragraph devoted to this.

4. Lack of quantitative benchmark

The authors used the BPCNN algorithm to detect co-occurrence of drug/problem pairs, since their text mining technology is not able to detect relationships (although the TEMIS MER technology is able to do so); however they do not compare their results quantitatively and therefore it is not possible to know the precision/recall of their method.

A quantitative measure, in the form of the IC variance values obtained as a result from the BCPNN algorithm, is provided in the results section. Table 4 and 5 include these values for the top scoring events for the drugs used in this study across the various sources, in order to bring out the trends in drug safety signals.

5. automatic dictionary creation: There is very extensive literature on how to create automatically a dictionary of specific terms: statistical and machine learning methods are commonly adopted. I wonder the robustness of their patent.

NA