

## Early detection of adverse drug events using the full text of letters to the editor

Chao Yang\*, Jacob Simmering, Padmini Srinivasan, Linnea Polgreen and Philip Polgreen

The University of Iowa, Iowa City, IA, USA

### Objective

Our objective was to explore the intuition that letters to the editor in leading medical journals contain early signals about adverse drug events. We explored this with letters in two leading journals.

### Introduction

Adverse drug events (ADEs) are a significant source of morbidity and mortality. The majority of postmarketing surveillance for ADEs is passive. Information regarding ADEs is reported to the medical community in peer-reviewed journals. However, in most cases, there is significant lag in the publication of peer-reviewed articles concerning ADEs. Within medical journals, our intuition is that letters to the editor may provide the earliest reports of ADEs. They often report single case reports or a collection of cases and usually precede more formal investigations and reports.

Although these letters may contain useful and timely information, the challenge is that letters to the editor may be 'buried' inside print journals. Furthermore, they may be more difficult to find and access even when using electronic searches because unlike other published reports, there is no corresponding abstract to view. Due to the lack of an abstract, detection depends almost exclusively upon words in a title or manually applied Medical Subject Headings (MeSH). We propose that searching the full text of letters to the editor can provide a faster and perhaps more complete detection of ADEs compared to searches based on MeSH terms or titles alone.

### Methods

We first identified a list of the most commonly used 179 drugs in 2008 based on the Agency for Healthcare Research and Quality compiled Medical Expenditures Survey. We then used MicroMedex, a commercial drug information service, to find a list of key publications describing ADEs for these drugs. Next, we obtained the text for the majority of letters to the editor published in *The Lancet* (6558 from 1967 to date; 82% of total) and *The New England Journal of Medicine* (3524 from 1966 to date; 75% of total). We restricted the letters to those that had the MeSH term 'adverse effect' in the indexing data. We also eliminated the letters with label 'Comment' to avoid searching letters specifically referencing a previously published paper in each journal, respectively. The resulting dataset contained 2166 letters for *The Lancet* and 1449 for *The New England Journal of Medicine*.

We then compared the results from two different search strategies. In the first, an emulation of a PubMed search, we only examined the MeSH terms and the title for the letter. Our second approach included a search of the full text of the letters in addition to the title. Using these two strategies, we were able to determine the 'earliest' letter in these two journals for a given drug/ADE pair. We compared this date against the date of the

citation referenced by MicroMedex to determine which search method provided the earliest detection.

### Results

Both search strategies, with and without full text, were able to find the particular drug/ADE pair mentioned in letters before the corresponding Micromedex reference. However, using full-text search outperformed title/MeSH-based search, not only based on the number of drug/ADE pairs found but also on the time of detection. The percentage of letters in the dataset that are not related to specific articles is 0.6% using title/MeSH, 2.1% using title/full-text. Furthermore, we found that MeSH terms are not always reliable. For example, some of the letters had MeSH terms like "Adverse Effect" but no mention of adverse effects in the letter. Not surprisingly, since MeSH is a controlled vocabulary, some of these terms do not appear in the full text of the letters. These findings are shown in Table 1.

Table 1. The results of two search strategies (title/mesh vs. title/full-text)

	Title/ MeSH	Title/full- text
No. specific drug/ADE letters found	23	77
Distinct No. drug/ADEs pairs found	10	35
Average time prior to the year cited by MicroMedex	5.291	6.901
Standard Deviation (year)	4.846	5.291

### Conclusions

Our results suggest that the full texts of the letters to the editor provide a potential stream of information regarding early warnings for ADEs. Future work will need to expand the number of journals considered and, furthermore, consider the potential for 'false positive' warnings.

### Keywords

Adverse drug events; letters to the editor; signal detection

### References

1. Chapman WW, Chu D, Dowling JN. ConText: an algorithm for identifying contextual features from clinical text. *BioNLP Prague*. 2007;2007:81-88.
2. Reynar JC, Ratnaparkhi A. A maximum entropy approach to identifying sentence boundaries. *Proc. ANLP97*. Washington, DC; 1997: 16-9.

\*Chao Yang

E-mail: chao-yang@uiowa.edu