New and Increasing Rates of Adverse Events Can be Found in Unstructured Text in Electronic Health Records using Natural Language Processing Tools

Roselie A Bright, ScD (OC/OCS/OHI)
Katherine Dowdy (Booz Allen Hamilton)
Summer K Rankin, PhD (Booz Allen Hamilton)
Serge Blok, PhD (Booz Allen Hamilton)
Lee Anne Palmer, VMD, MPH (CVM/OSC/VPS)
Susan J. Bright-Ponte, DVM, MPH (CVM/OSC/DS)

Supported by FDA contract. Presentation is not necessarily official view of FDA or Booz Allen Hamilton. We gratefully acknowledge our colleagues’ support.

INTRODUCTION

Our goal was to develop a new method to identify adverse event (AE) signals using the unstructured text of electronic health records (EHRs).
- Independent of vocabulary used.
- Even if AE is new.
- Even if writer does not attribute event as an AE.

We chose to explore clinical care notes from 2001-2008, to see if we could detect contaminated heparin AEs.
- Heparin contamination AEs in the US were recognized by FDA in early 2008, although some newspaper articles suggested contamination as early as 2005.
- Heparin is an anticoagulant used in surgeries, blood transfusions, and catheterizations.

METHODS

Clinical Notes
- MIMIC III
- 7/2001-6/2008
- Teaching hospital
- Critical care
- Dates within several weeks of real dates

Text Notes Preparation
- Concatenated notes in time order for each admission.
- De-duplicated notes using Bloatectomy.
- Vectorized each admission’s remaining notes.

Group Creation & Term Extraction
Separated admissions by time period:
- Period 1, 7/2001-6/2006 (14,441).

Used ensemble of supervised classification methods and statistical rules to filter notes to unusual terms that distinguish period 3 from period 2.

Topic Analysis of Period 3
Leveraged Latent Dirichlet Allocation topic modeling to organize notes into 20 topics.
Read admissions with top topic scores. Interpreted topics within context of frequency analysis across periods 1-3.

RESULTS

Period 3 topics (showed increases over periods 1-3) for conditions that are:

- Hypotension right after heparin exposure
- Attributed drug AEs
- Admissions for attributed therapy/ food AEs
- Rare conditions
- Continuing drainage after surgery

Uncommon (examples) Common (examples)
- • Long hospital stays
- • Heart attacks
- • Brain trauma
- • Brain bleed
- • Brain ischemia

Proportion of invasive cardiovascular procedure admissions (heparin presumed) with “hypotension” word:

SUCCESSES
- Could be a useful supplemental post-marketing surveillance method for generating hypotheses to be studied by finding unattributed AEs
- Unlocks clinical text notes
- Uses open source software
- Needs relatively few computing resources
- Adaptable to other settings

FUTURE
- Automated support of review step
- We invite new members to the Shakespeare Project Team!

The Shakespeare Project: Literature big data tools applied to clinical records
New and Increasing Rates of Adverse Events Can be Found in Unstructured Text in Electronic Health Records using Natural Language Processing Tools

REFERENCES

Introduction


Methods

MIMIC-III Critical Care Database. https://mimic.physionet.org/about/mimic/.


Altman DG, Bland JM. How to obtain the P value from a confidence interval. BMJ. 2011; 343.

https://doi.org/10.1136/bmj.d2304.

Invasive cardiovascular procedure codes were ICD9CM 3891, 3961, 3965-3966, from MIMIC-III definitions table.

Interpretation


Graphics