Title: Natural Language Processing for EHR-based Pharmacovigilance: A Structured Review

Journal: Drug Safety

Author: Yuan Luo, PhD; William K. Thompson, PhD; Timothy M. Herr, MS; Zexian Zeng, MS; Mark A. Berendsen MLIS; Siddhartha R. Jonnalagadda, PhD; Matthew B. Carson, PhD; Justin Starren MD PhD

Affiliation: Department of Preventive Medicine (YL, YT, TMH, ZZ, SJ, MC, JS), Galter Health Sciences Library (MAB), Northwestern University Feinberg School of Medicine, Chicago, IL, USA
Knowledge and Conversation Group (SJ), Microsoft, Redmond, WA, USA

Correspondence to: yuan.luo@northwestern.edu

Detailed Search Strategies for Bibliographic Databases

Ovid MEDLINE(R) 1946 to January Week 1 2016; Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations January 18, 2016
Date run: January 19, 2016
Results: 332

1. exp Product Surveillance, Postmarketing/
2. ((drug* or prescription or medication*) adj3 (surveillance or safety or interaction* or reaction* or monitor*)).tw.
3. (adverse adj3 (event* or reaction* or effect*)).tw.
4. pharmacovigilance.tw.
5. side effect*.tw.
6. exp Drug Information Services/
7. exp "Drug-Related Side Effects and Adverse Reactions"/
8. exp Drug Interactions/
9. or/1-8
10. exp Data Mining/
11. ((text or data) adj3 (mine or mining)).tw.
12. algorithms/
13. algorithm*.tw.
15. Artificial Intelligence/
16. artificial intelligence.tw.
17. natural language processing.tw.
18. or/10-17
19. 9 and 18
20. exp Medical Records/
21. ((health or patient or medical) adj record*).tw.
22. (electronic adj3 record*).tw.
23. narrative*.tw.
24. or/20-23
25. 19 and 24
26. limit 25 to yr="2000 -Current"
Embase 1974-present; Embase Classic 1947-1973; Medline 1966-present (embase.com)

Embase Session Results (19 Jan 2016)

<table>
<thead>
<tr>
<th>No.</th>
<th>Query</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>#27</td>
<td>#9 AND #20 AND #26 AND [2000-2016]/py</td>
<td>534</td>
</tr>
<tr>
<td>#26</td>
<td>#21 OR #22 OR #23 OR #24 OR #25</td>
<td>240697</td>
</tr>
<tr>
<td>#25</td>
<td>(clinical NEXT/1 text*):ab,ti</td>
<td>230</td>
</tr>
<tr>
<td>#24</td>
<td>narrative*:ab,ti</td>
<td>21976</td>
</tr>
<tr>
<td>#23</td>
<td>(electronic NEAR/3 record*):ab,ti</td>
<td>29727</td>
</tr>
<tr>
<td>#22</td>
<td>((health OR patient OR medical) NEXT/1 record*):ab,ti</td>
<td>127388</td>
</tr>
<tr>
<td>#21</td>
<td>'medical record'/exp</td>
<td>162259</td>
</tr>
<tr>
<td>#20</td>
<td>#10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19</td>
<td>322809</td>
</tr>
<tr>
<td>#19</td>
<td>'natural language processing':ab,ti</td>
<td>1323</td>
</tr>
<tr>
<td>#18</td>
<td>'natural language processing'/de</td>
<td>2356</td>
</tr>
<tr>
<td>#17</td>
<td>'artificial intelligence':ab,ti</td>
<td>2219</td>
</tr>
<tr>
<td>#16</td>
<td>'artificial intelligence'/de</td>
<td>14623</td>
</tr>
<tr>
<td>#15</td>
<td>'machine learning':ab,ti</td>
<td>7545</td>
</tr>
<tr>
<td>#14</td>
<td>'machine learning'/exp</td>
<td>84029</td>
</tr>
<tr>
<td>#13</td>
<td>algorithm*:ab,ti</td>
<td>185292</td>
</tr>
<tr>
<td>#12</td>
<td>'algorithm'/exp</td>
<td>206943</td>
</tr>
<tr>
<td>#11</td>
<td>((text OR data) NEAR/3 (mine OR mining)):ab,ti</td>
<td>8174</td>
</tr>
<tr>
<td>#10</td>
<td>'data mining'/de</td>
<td>6310</td>
</tr>
<tr>
<td>#9</td>
<td>#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8</td>
<td>1275247</td>
</tr>
<tr>
<td>#8</td>
<td>'drug interaction'/exp</td>
<td>309301</td>
</tr>
<tr>
<td>#7</td>
<td>'adverse drug reaction'/exp</td>
<td>410928</td>
</tr>
<tr>
<td>#6</td>
<td>'drug information'/de</td>
<td>21397</td>
</tr>
<tr>
<td>#5</td>
<td>(side NEXT/1 effect*):ab,ti</td>
<td>283654</td>
</tr>
<tr>
<td>#4</td>
<td>pharmacovigilance:ab,ti</td>
<td>5960</td>
</tr>
<tr>
<td>#3</td>
<td>(adverse NEAR/3 (event* OR reaction* OR effect*)):ab,ti</td>
<td>380700</td>
</tr>
<tr>
<td>#2</td>
<td>((drug* OR prescription OR medication*) NEAR/3 (surveillance OR safety OR interaction* OR reaction* OR monitor*)):ab,ti</td>
<td>102773</td>
</tr>
<tr>
<td>#1</td>
<td>'postmarketing surveillance'/exp</td>
<td>26672</td>
</tr>
</tbody>
</table>
Web of Science--Indexes=SCI-EXPANDED, CPCI-S Timespan=2000-2016
Date run: January 19, 2016
Results: 368

TS=(pharmacovigilance OR side effect* OR ((drug* OR prescription OR medication*) near/3 (surveillance OR safety OR interaction* OR reaction* OR monitor*)) OR (adverse near/3 (event* OR reaction* OR effect*))) AND TS=(text mine OR text mining OR data mine OR data mining OR artificial intelligence OR machine learning OR natural language processing OR algorithm*) AND TS=((electronic near/3 record*) OR patient record* OR health record*)
IEEE Xplore
Date run: January 19, 2016
Results: 188

(((MeSH Terms:"Product Surveillance, Postmarketing" OR MeSH Terms:"Drug-Related Side Effects and Adverse Reactions" OR MeSH Terms:"Drug Utilization Review" OR MeSH Terms:"Drug Information Services" OR MeSH Terms:"Drug Interactions" OR MeSH Terms:"Adverse Drug Reaction Reporting Systems" OR MeSH Terms:"Pharmacovigilance"))) OR (((((("drug surveillance") OR pharmacovigilance) OR "adverse drug event") OR "adverse event") OR "drug monitoring") OR "drug safety"))) AND ((((("health record") OR "patient record") OR "medical record") OR ehr) OR emr)) and refined by Year: 2000-2015

*IEEE Xplore search would only refine through 2015 since there were no 2016 articles retrieved.


Preliminary results from the EU-ADR project. *Basic and Clinical Pharmacology and Toxicology, 107*, 123-124. **Exclusion reason: Abstract Only;**