



Detecting Pharmacovigilance Signals Combining Electronic Medical Records With Spontaneous Reports: A Case Study of Conventional Disease-Modifying Antirheumatic Drugs for Rheumatoid Arthritis

Liwei Wang¹, Majid Rastegar-Mojarad¹, Zhiliang Ji², Sijia Liu¹, Ke Liu², Sungrim Moon¹, Feichen Shen¹, Yanshan Wang¹, Lixia Yao¹, John M. Davis III¹ and Hongfang Liu^{1*}

OPEN ACCESS

Edited by:

Cedric Bousquet,
Centre Hospitalier Universitaire (CHU)
de Saint-Étienne, France

Reviewed by:

Domenico Criscuolo,
Genovax S.r.l., Italy
Ludo Haazen,
Independent Researcher, Mechelen,
Belgium
Jinoos Yazdany,
University of California,
San Francisco, United States

*Correspondence:

Hongfang Liu
liu.hongfang@mayo.edu

Specialty section:

This article was submitted to
Pharmaceutical Medicine
and Outcomes Research,
a section of the journal
Frontiers in Pharmacology

Received: 16 December 2017

Accepted: 19 July 2018

Published: 07 August 2018

Citation:

Wang L, Rastegar-Mojarad M, Ji Z,
Liu S, Liu K, Moon S, Shen F, Wang Y,
Yao L, Davis JM III and Liu H (2018)
Detecting Pharmacovigilance Signals
Combining Electronic Medical
Records With Spontaneous Reports:
A Case Study of Conventional
Disease-Modifying Antirheumatic
Drugs for Rheumatoid Arthritis.
Front. Pharmacol. 9:875.
doi: 10.3389/fphar.2018.00875

¹ Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, MN, United States, ² State Key Laboratory of Cellular Stress Biology, School of Life Sciences, Xiamen University, Xiamen, China

Multiple data sources are preferred in adverse drug event (ADEs) surveillance owing to inadequacies of single source. However, analytic methods to monitor potential ADEs after prolonged drug exposure are still lacking. In this study we propose a method aiming to screen potential ADEs by combining FDA Adverse Event Reporting System (FAERS) and Electronic Medical Record (EMR). The proposed method uses natural language processing (NLP) techniques to extract treatment outcome information captured in unstructured text and adopts case-crossover design in EMR. Performances were evaluated using two ADE knowledge bases: Adverse Drug Reaction Classification System (ADReCS) and SIDER. We tested our method in ADE signal detection of conventional disease-modifying antirheumatic drugs (DMARDs) in rheumatoid arthritis patients. Findings showed that recall greatly increased when combining FAERS with EMR compared with FAERS alone and EMR alone, especially for flexible mapping strategy. Precision (FAERS + EMR) in detecting ADEs improved using ADReCS as gold standard compared with SIDER. In addition, signals detected from EMR have considerably overlapped with signals detected from FAERS or ADE knowledge bases, implying the importance of EMR for pharmacovigilance. ADE signals detected from EMR and/or FAERS but not in existing knowledge bases provide hypothesis for future study.

Keywords: adverse drug event, disease-modifying antirheumatic drug (DMARD), Electronic Medical Records (EMR), FDA Adverse Event Reporting System (FAERS), pharmacovigilance, natural language processing

INTRODUCTION

Adverse drug event (ADE), the untoward occurrence after exposure to a drug (Pirmohamed et al., 1998), has been an important public health concern. It is well recognized that pre-marketing randomized controlled trials (RCTs) may not detect all types of ADEs related to a particular drug in clinical practice (Schneeweiss and Avorn, 2005) due to some limitations (Stolberg et al., 2004). Therefore, FDA encourages the public to voluntarily

report any suspicious ADEs to facilitate pharmacovigilance. The FDA Adverse Event Reporting System (FAERS) has become an important resource in post-marketing surveillance for ADEs (Yang et al., 2011; Jiang et al., 2013; Wang et al., 2014, 2017a; Yu et al., 2016; European Medicines Agency, 2017). Commonly used methods to detect ADE signals from FAERS include Proportional Reporting Ratio (PRR), Reporting Odds Ratio (ROR), Information Component (IC) and Empirical Bayesian Geometric Mean (EBGM) (Bate and Evans, 2009). For example, a recent study discovered the higher risk of glioblastoma associated with tumor necrosis factor inhibitor by using ROR in FAERS (Guo et al., 2016). Since FAERS is a voluntary reporting system, some inherited disadvantages exist such as underreporting and no incidence of ADEs (Stephenson and Hauben, 2007). Other data sources have been studied to assist ADE detection further. Besides social media (Liu and Chen, 2015) and literature (Xu and Wang, 2014), Electronic Medical Record (EMR) has also been paid a lot of attention. E.g., Observational Health Data Sciences and Informatics (OHDSI) has been implementing the international collaboration to conduct clinical outcome research including ADE surveillance using EMR (Hripcsak et al., 2015). A majority of the valuable information about clinical outcomes and medication in EMR is captured in clinical narratives (Classen et al., 2011). To analyze the clinical narratives, natural language processing (NLP) methods have been employed to extract potential drug-ADE pairs. Signal detection methods such as χ^2 test or odds ratio have been utilized to rank the pairs (Wang et al., 2009, 2010; Harpaz et al., 2013). The feasibility of using EMR to validate potential drug-ADE pairs detected from FAERS has been demonstrated (Wang et al., 2017b). However, a majority of the existing studies used data-driven disproportionality analysis with little consideration of confounding variables such as race and gender. Analytic methods to monitor potential ADEs of a drug after prolonged drug exposure are still lacking (Schuemie et al., 2016). In this study, we propose a method aiming to screen potential ADEs from both FAERS and EMR. Specifically, this method incorporates NLP techniques to process EMR, where case-crossover study is employed with consideration of clinical context, including drug indication and temporal information. Subsequently, the potential ADEs from clinical notes and FAERS are evaluated by comparing against SIDER (Kuhn et al., 2015) and Adverse Drug Reaction Classification System (ADReCS) (Cai et al., 2015). We test our proposed method for screening potential ADEs for conventional disease-modifying antirheumatic drugs (DMARDs), the cornerstone treatment for Rheumatoid arthritis (RA) patients.

MATERIALS AND METHODS

Our method for detecting potential ADEs of a drug and indication pair (D, IND) is featured by NLP techniques and case-crossover study design. Specifically it includes three steps: data preparation, signal detection, and validation and hypothesis generation (**Figure 1**). In the data preparation step, we apply NLP techniques to extract drugs and clinical outcomes, i.e., problems from EMR. In the signal detection step, a cohort of patients taking

drug D for IND is identified from longitudinal EMR data and the corresponding spontaneous reports for the pair (D, IND) in FAERS are extracted. Then, we use case-crossover study design to detect potential ADE signals from EMR. Potential ADE signals from FAERS are detected using ROR. In the validation and hypothesis generation step, signals from EMR and FAERS are then normalized and evaluated using ADE knowledge bases, i.e., ADReCS and SIDER where exact and flexible mapping strategies are used. Existing records of ADEs in the knowledge bases serve as the gold standards and false positives (i.e., signals not present in the knowledge bases) are potential ADEs for further investigation, i.e., hypothesis. The study has been approved as a minimal risk study with Mayo IRB of 13-009317. Only the data of patients with research authorization for using their EMR records have been used.

Data Preparation

In this study we used the normalized knowledge-enhanced data set AERS-DM (Wang et al., 2014) derived from FAERS that is downloadable at our website¹. The details and summarization for the data processing can be found in our previous paper (Yang et al., 2011; Wang et al., 2017b).

The EMR data used in our study consists of all patient records at Mayo Clinic Rochester campus with Minnesota research permission over a period of 20 years (January 1995–October 2017). As a majority of the clinical information in EMR is captured in unstructured clinical notes, we process clinical notes using MedXN (Sohn et al., 2014), which extracts and normalizes medication mentions to RxNorm and MedTagger (Torii et al., 2011), which extracts and normalizes clinical concept mentions to the Unified Medical Language System (UMLS). We limit our analysis to the following UMLS semantic types: “Finding(T033),” “Laboratory or Test Result(T034),” “Sign or Symptom(T184),” “Disease or Syndrome(T047),” “Mental or Behavioral Dysfunction (T048),” “Neoplastic Process(T191)” and “Cell or Molecular Dysfunction(T049),” which are related to indications and clinical outcomes.

There are two ADE knowledge bases used in our study. One is SIDER, a resource that aggregates dispersed public information on side effects recorded in public documents and package inserts (Kuhn et al., 2015). In the downloadable files, drugs are coded using STITCH compound identifiers, adverse events and indications are coded using both UMLS and MedDRA lowest level term (LLT). We used SIDER 4.1 in our study.

The other is ADReCS, a more comprehensive ADE resource integrating multiple public medical repositories such as SIDER, DailyMed and ClinicalTrials (Cai et al., 2015). ADReCS adopts WHO-ART, MedDRA, and the UMLS as major references for standardization. In this study, ADReCS 1.6 was used.

Signal Detection From AERS-DM

The associated records of drug D and indication IND are extracted. ROR are computed based on a 2×2 contingency table (Evans et al., 2001; Rothman et al., 2004). More specifically the

¹<http://informatics.mayo.edu/adepedia/index.php/Download>

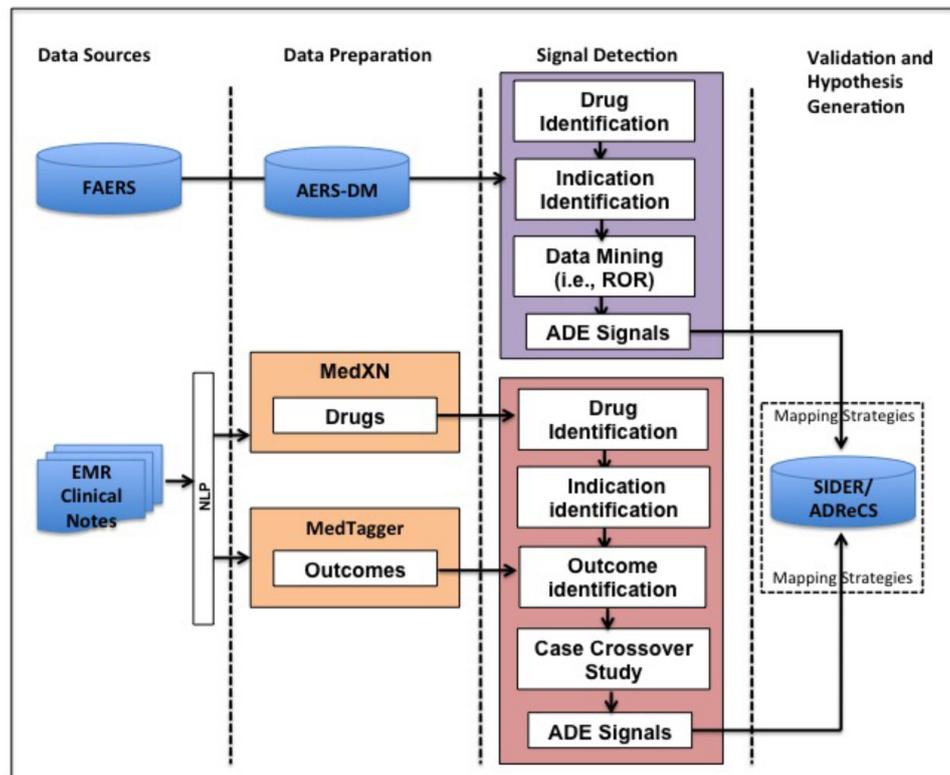


FIGURE 1 | Overview of our method for ADE mining from FAERS and EMR. FAERS, FDA Adverse Event Reporting System; EMR, Electronic Medical Record; ADE, Adverse Drug Event; ROR, Reporting Odds Ratio.

number of reports associated with the drug and ADE is defined as a . The number of reports with the drug and without ADE is defined as b . The number of reports with other drugs and with ADE is defined as c . The number of reports with other drugs and without ADE is defined as d . In this analysis, a signal is detected if the lower bound of 95% confidence interval of ROR exceeds 1 (van Puijenbroek et al., 2002). The analysis is implemented in R package PhViD 1.0.6 (Ahmed and Poncet, 2013).

$$ROR = \frac{a \times d}{b \times c} \quad (1)$$

Signal Detection From EMR

We use case-crossover design to detect signals of association between drug D and potential ADEs for patients with IND from EMR (Maclure and Mittleman, 2000). The case-crossover study is a popular epidemiological study design for which individuals act as their own controls, i.e., comparisons are made within individuals. It has been recommended for pharmacoepidemiological research to leverage EMR data as confounding factors, such as genetics and gender, are well controlled and they do not vary with time (Hallas and Pottegård, 2014).

In this study we conducted unconditional analysis using odds ratio (OR) to detect signals. A potential ADE is detected when the lower bound of the 95% confidence interval of the OR exceeds 1

and p -value is less than 0.05.

$$OR = \frac{a \times d}{b \times c} \quad (2)$$

Figure 2 defines a , b , c , and d for OR calculation in case-crossover study. We use the patient cohort during the observation period before the use of drug D and after the diagnosis of IND as the control, and the same patient cohort during the observation period after the first use until the last use of drugs of interest as the case. In the case group, if an outcome occurs, the patient is counted as “a”, otherwise as “b”. In the control group, if an outcome occurs, then the patient is counted as “c”, otherwise as “d”. OR value could be calculated for each outcome using the same patient cohort undergoing drug therapy. Therefore, the patient number in the case group is the same as that in the control group for each outcome.

Validation and Hypothesis Generation

The ADE obtained from FAERS, EMR, SIDER and ADReCS are coded differently. In order to evaluate the performance of this method for ADE signal detection, we use two mapping strategies to map signals among various data sources. The first strategy is exact mapping. We unify all the codes from various data sources to the UMLS and then conduct direct mapping among sources. The second strategy is flexible mapping combining both exact

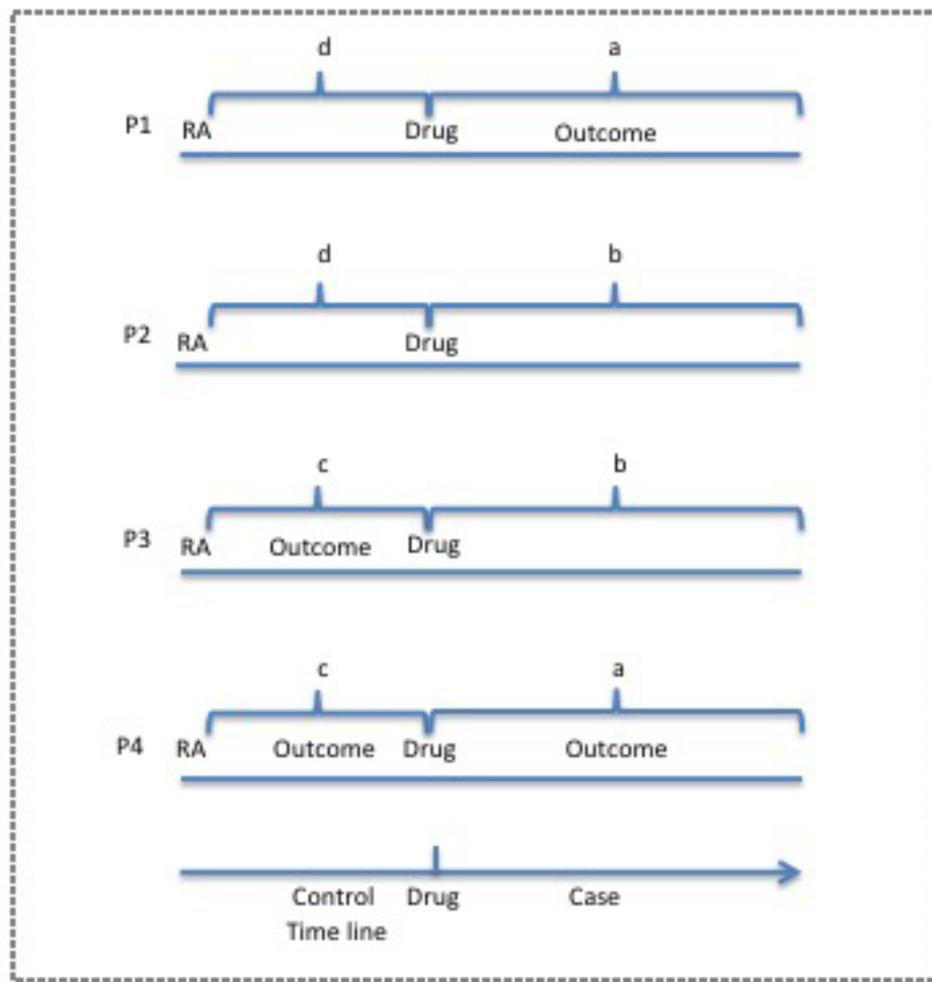


FIGURE 2 | Case-crossover study design for potential ADEs of drugs of interest from EMR clinical notes. The observation period before the use of drug is used as the control and the period after drug use is used as the case. P, Patient; RA, rheumatoid arthritis.

mapping and the mapping through hierarchical structure. The latter is conducted by leveraging the 5-level hierarchical structure of Medical Dictionary for Regulatory Activities (MedDRA) (Bousquet et al., 2005). Specifically, we expand high-level terms of MedDRA from knowledge bases, i.e., SIDER and ADReCS, to include all sub-level terms and then conduct mapping between knowledge bases and EMR/FAERS. This strategy considers semantic granularity of terms in various sources and tries to obtain as many mappings as possible. **Figure 3** shows the examples of these two mapping strategies, where A refers to exact mapping and B refers to flexible mapping. For exact mapping, the term “Abdominal discomfort” is exactly mapped between SIDER/ADReCS and EMR/FAERS as this term appears in those sources. For flexible mapping, if “Skin vasculitis NOS” is detected from EMR/FAERS but only “Vasculitides” exists in the knowledge bases, we first find the lower level term of Vasculitides, i.e., “skin vasculitis NOS” and then map to the signal from EMR/FAERS. Recall, precision, and F1 measures are calculated for FAERS and EMR using the following formula (Powers, 2011). F1-measure is

the harmonic mean of precision and recall (Rennie, 2004) and balances recall and precision (Yang, 1999).

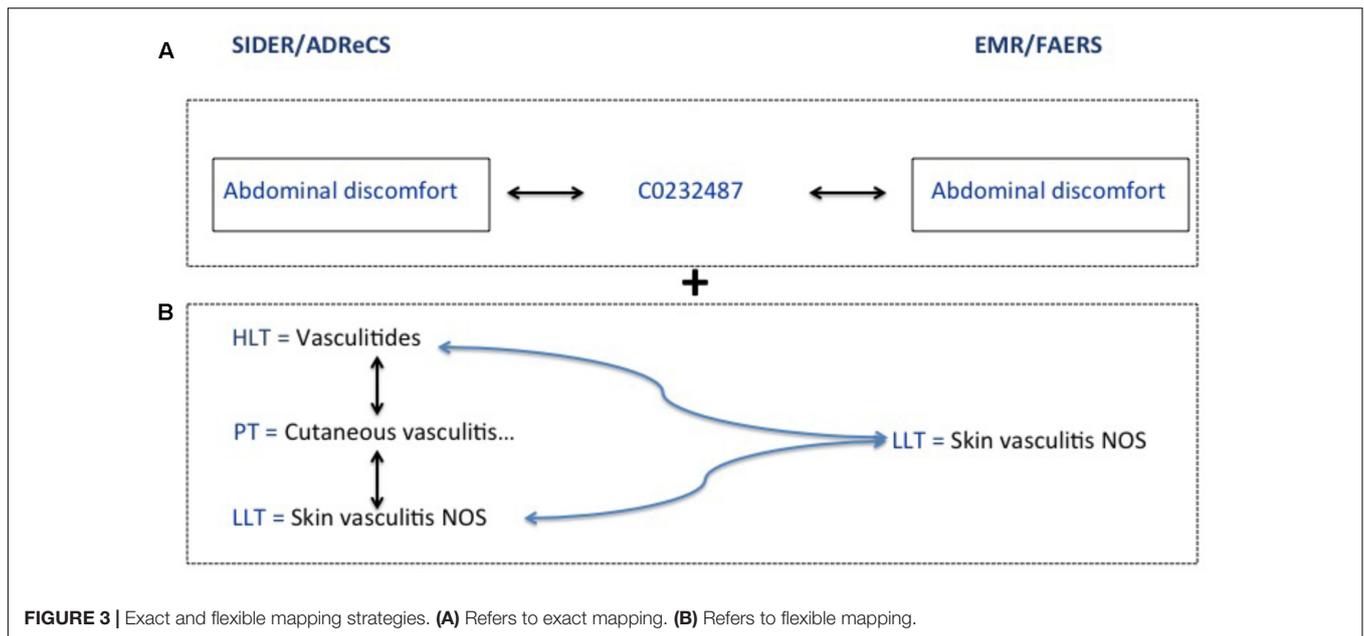
$$\text{Precision} = \frac{\text{Number of matched ADEs}}{\text{Number of discovered ADEs}} \quad (3)$$

$$\text{Recall} = \frac{\text{Number of matched ADEs}}{\text{Number of confirmed ADEs}} \quad (4)$$

F1 =

$$\frac{2 \times (\text{Number of matched ADEs})}{2 \times (\text{Number of matched ADEs}) + \text{number of discovered unmatched ADEs} + \text{number of undiscovered matched ADEs}} \quad (5)$$

ADE signals detected from FAERS and/or EMR can be validated if they are also captured in SIDER/ADReCS. ADE signals detected from EMR and/or FAERS that are not validated by gold standards could very well likely be false positives but also could be potential novel ADEs, providing hypothesis for future study.



EXPERIMENTS

We tested the proposed method by detecting potential ADEs for conventional disease-modifying antirheumatic drugs (DMARDs). DMARDs include methotrexate (MTX), sulfasalazine, hydroxychloroquine and leflunomide, which are the cornerstone treatment for Rheumatoid arthritis (RA). An alternative drug choice for RA is biological agents (biologics) such as etanercept, but they are more expensive (Lee and Bae, 2016). DMARDs remain as the primary drug choice for RA treatment (Schmitz et al., 2011).

After data preparation, we extracted RA patients who have been on DMARDs for the treatment of RA. We also extracted associated records of those drugs and indication RA from AERS-DM. We then conducted signal detection for FAERS and EMR.

In validation, two files, meddra_all_label_indications.tsv.gz and meddra_all_label_se.tsv.gz from SIDER 4.1 were used in this study as gold standard for the evaluation of FAERS signals and clinical notes results. Note that not all package inserts with the same ingredient have the same adverse events. Therefore, it's important to extract adverse events from package inserts with indications of "Rheumatoid

TABLE 1 | Clinical characteristics of patients in various sources.

Source	Gender	Sulfasalazine [report number (%)]	Methotrexate [report number (%)]	Leflunomide [report number (%)]	Hydroxychloroquine [report number (%)]
FAERS	Female (%)	2,312 (67.6%)	29,786 (74.2%)	5,807 (74.4%)	5,635 (79.1%)
	Male (%)	894 (26.1)	7,518 (18.7%)	1,617 (20.7%)	1,040 (14.6%)
	F:M ratio	2.6:1	4.0:1	3.6:1	5.4:1
	Total	3,420	40,161	7,820	7,123
	Female (%)	1,264 (64.8%)	4,193 (68.6%)	1,619 (70.2%)	3,985 (73.0%)
	Male (%)	687 (35.2%)	1,921 (31.4%)	686 (29.8%)	1,471 (27.0%)
	F:M ratio	1.8:1	2.2:1	2.4:1	2.0:1
	Race				
	American Indian/Alaskan Native	16	33	20	32
	Asian	22	56	16	56
	Black or African American	19	51	16	74
	White	1,729	5,385	2,084	4,873
	Other	34	94	38	117
	Unknown	121	467	121	278
Choose not to disclose	10	28	10	26	
Total	1,951	6,114	2,305	5,456	

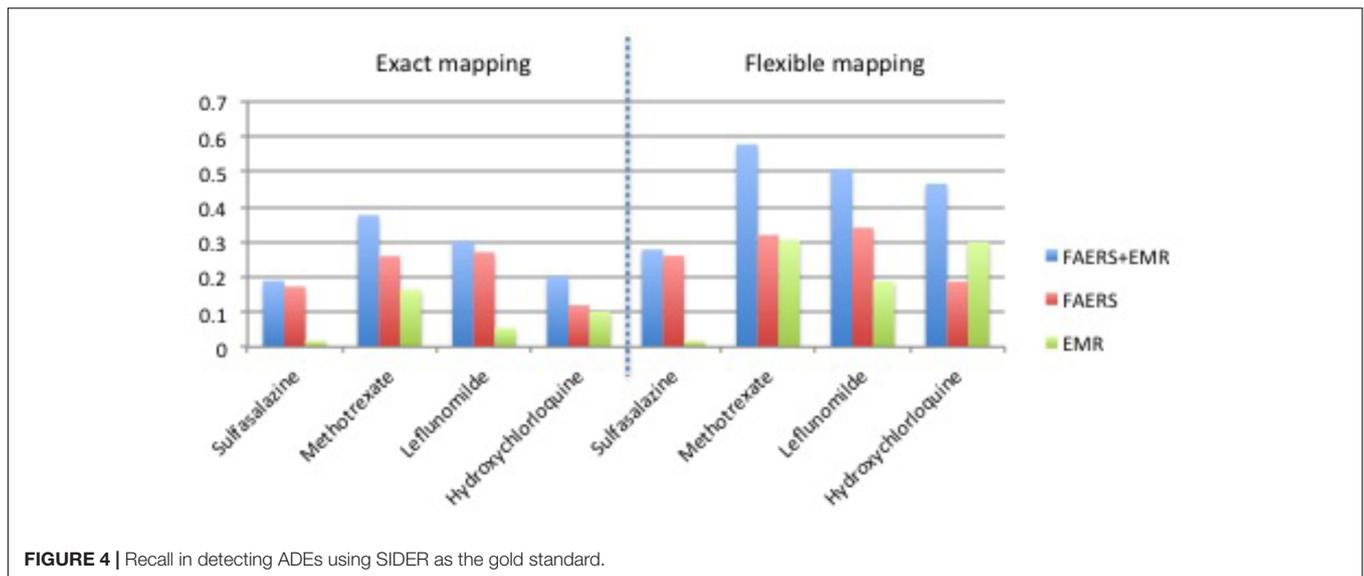


FIGURE 4 | Recall in detecting ADEs using SIDER as the gold standard.

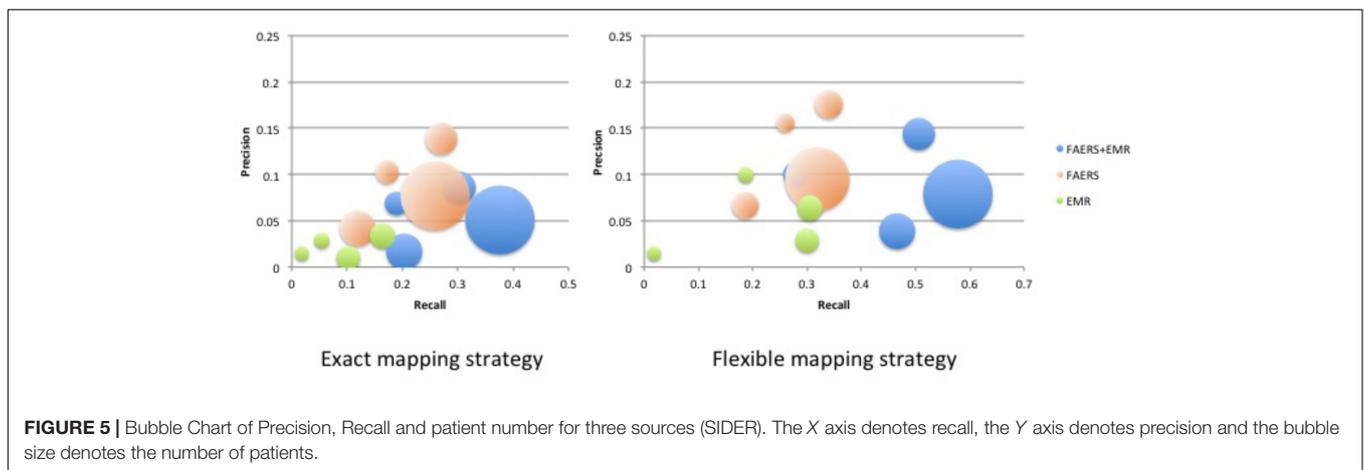


FIGURE 5 | Bubble Chart of Precision, Recall and patient number for three sources (SIDER). The X axis denotes recall, the Y axis denotes precision and the bubble size denotes the number of patients.

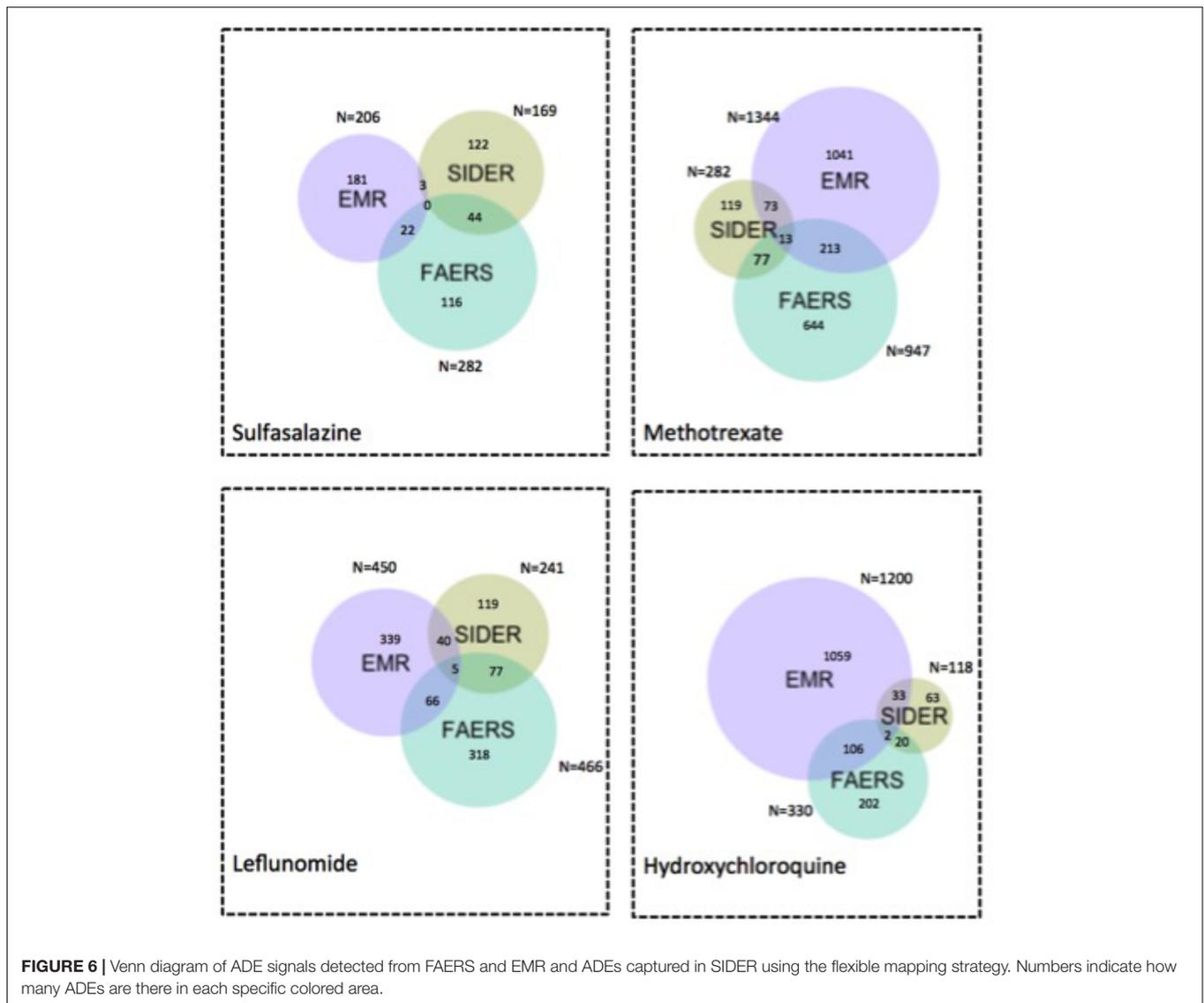
arthritis” for drugs of interest. Using the STITCH compound identifiers of methotrexate (CID100004112), leflunomide (CID100003899), hydroxychloroquine (CID100003652) and sulfasalazine (CID105353980), package inserts with indications of “Rheumatoid arthritis” were extracted from meddra_all_label_indications.tsv.gz, and then those package inserts were used to identify recorded adverse events of these drugs.

RESULTS

Table 1 shows clinical characteristics of RA patients on drugs of interest with ADE reports in FAERS and RA patients on drugs of interest in EMRs. **Supplementary Table 1** shows the recall, precision, and F1 in detecting ADEs when using ADReCS and SIDER as the gold standards. **Figure 4** shows recall (SIDER) greatly increased when combining FAERS with EMR compared with FAERS alone and EMR alone, especially for flexible mapping strategy. For example, recall for

detecting ADEs associated with Methotrexate increased to 0.578 (FAERS + EMR) from 0.319 (FAERS) and 0.305 (EMR) using flexible mapping strategy. **Figure 5** (SIDER) and **Supplementary Figure S1** (ADReCS) show the three-dimensional bubble charts of precision, recall and patient number for three sources, where the bubble sizes represent the patient number. This figure demonstrates that flexible mapping strategy helps to enhance both precision and recall for three sources. **Supplementary Figure S2** shows precision (FAERS + EMR) in detecting ADEs improved using ADReCS as gold standard compared with SIDER.

Figure 6 and **Supplementary Figure S3** show the Venn diagrams of ADE signals detected from FAERS and EMR as well as ADEs captured in SIDER, using the flexible mapping strategy and the exact mapping strategy, respectively. In **Figure 6**, the overlaps represent the typical ADEs in SIDER that are also detected from FAERS and EMR, e.g., bacterial infection, Leukocytoclastic vasculitis and sepsis for methotrexate and retinopathy for hydroxychloroquine. **Supplementary Figures S4, S5** show the Venn diagram of signals detected



from FAERS and EMR with ADEs captured in ADReCS, using the exact mapping strategy and the flexible mapping strategy, respectively. From the Venn diagrams, it can be observed that both FAERS and EMR have a considerable number of detected signals overlapping with known ADEs captured in knowledge bases. However, there is a significant overlap between signals detected from FAERS and signals detected from EMR but not in existing knowledge bases, suggesting potential ADEs for further investigation. The top 10 potential ADEs based on the OR score are listed in **Table 2**. The disorders shown in bold letters may represent potential confounding factors due to indications. For example, diskitis is likely to be the manifestations of the underlying inflammatory arthritis for which sulfasalazine is being used. Other disorders may be potential ADEs associated with the drug. For example, methotrexate associated malignant melanoma has been confirmed in a recent study (Polesie et al., 2017). Additionally, signals detected independently by either FAERS or EMR deserve further

investigation. For example, FAERS alone detected ADE signal associated with sulfasalazine, i.e., pleural effusion, which has been reported in literature to occur after sulfasalazine use (Gupta et al., 2012) but not recorded in knowledge bases. Hyperparathyroidism associated with methotrexate detected by using EMR alone has also been reported in literature (Levy et al., 2006).

DISCUSSION

Drug safety surveillance in clinical practice is important for each drug after introduction on the market. In this study we propose a method for ADE detection from EMR and FAERS and used DMARDs for the treatment of RA as the example case. Some detected ADEs can be confirmed by knowledge bases, and some are potentially new ADEs detected from EHR and/or FAERS but not in

TABLE 2 | Top 10 potential new ADEs detected from FAERS and EMR but not recorded in ADReCS.

Drugs	ADE AAD ADE signals	OR	P-value	Bonferroni-corrected P-value
Sulfasalazine	Postoperative wound infection	8.02	3.60E-14	1.44E-12
	Diskitis	8.02	3.60E-14	1.44E-12
	Delayed healing	6.02	2.47E-18	9.88E-17
	Organizing pneumonia	4.01	3.82E-13	1.53E-11
	Lymph node metastasis	4.01	3.82E-13	1.53E-11
	Finger infection	4.01	3.82E-13	1.53E-11
	Staphylococcus	3.01	8.19E-05	0.003276
	Foot ulcer	2.67	0.00	0.00
	Interstitial fibrosis	2.51	6.93E-11	2.77E-09
	Septic joint	2.50	0.00	0.00
Leflunomide	Hospital acquired pneumonia	22.11	1.23E-24	4.92E-23
	Polymicrobial infection	14.04	1.23E-24	4.92E-23
	Foot cellulitis	14.04	0.01	0.40
	Septic joint	12.06	7.74E-23	3.10E-21
	Klebsiella pneumoniae	12.03	7.74E-23	3.10E-21
	Furunculosis	12.03	1.69E-20	6.76E-19
	Pseudomonas pneumonia	10.02	1.72E-26	6.88E-25
	Hip dislocation	9.03	7.05E-18	2.82E-16
	Subtrochanteric fracture	8.01	7.05E-18	2.82E-16
	Perforated gastric ulcer	8.01	0.02	0.80
Methotrexate	Lymphoproliferative disorder	34.09	1.05E-12	4.20E-11
	Shoulder fracture	16.02	2.19E-32	8.76E-31
	Malignant melanoma <i>in situ</i>	16.02	2.19E-32	8.76E-31
	Bacterial pneumonia	16.02	2.19E-32	8.76E-31
	Hospital acquired pneumonia	14.03	5.39E-07	2.16E-05
	Varicella zoster	14.02	2.05E-30	8.20E-29
	B cell lymphoma	12.04	0.01	0.40
	Osteoporotic fracture	12.02	0.00	0.00
	Chest congestion	12.02	0.00	0.00
	Wrist injury	12.01	1.47E-28	5.88E-27
Hydroxychloroquine	Postoperative infection	12.01	7.51E-40	3.00E-38
	Rheumatoid arthropathy	8.01	5.13E-43	2.05E-41
	Metastatic carcinoma	8.01	5.13E-43	2.05E-41
	Hip dislocation	8.01	5.13E-43	2.05E-41
	SLE flare	8.01	5.47E-34	2.19E-32
	Lumbar fracture	8.01	5.47E-34	2.19E-32
	Injection site reactions	8.01	5.47E-34	2.19E-32
	Ghon complex	8.01	5.47E-34	2.19E-32
	Upper extremity edema	5.00	2.30E-35	9.20E-34
	Shoulder dislocation	5.00	2.30E-35	9.20E-34

existing knowledge bases, warranting further studies for validation.

Researchers have tried to build comprehensive ADE knowledge bases but currently there are no well established ones. One of reasons lies in the fact that post-marketing surveillance is an ongoing process. Though SIDER is included in ADReCS, the comparison of SIDER and ADReCS strengthens our method as some false positives detected from SIDER are true ADEs captured in ADReCS.

The findings from **Supplementary Table S1** demonstrated that flexible mapping strategy helped to improve recall, precision and F1 on the whole. One possible reason maybe because EMR clinical narratives contain dense ADE signals in various

semantic granularities where exact mapping strategy may miss some mappings. ADE signals detected from EMR and/or FAERS that are not validated by gold standards could very well likely be false positives but also could be potential novel ADEs.

In our previous study, we demonstrated the feasibility of EMR in validating potential ADEs from FAERS (Wang et al., 2017b). The present study focuses on a method for ADE detection from EMR and FAERS. Our method offers multiple contributions to the community. First, we use NLP techniques to extract treatment outcome information captured in unstructured text and standardize the information using UMLS codes which enables us to detect potential ADEs not captured in structured EMR data and at the same time, allows us to adopt the flexible

mapping strategy to combine signals detected from diverse data sources. Secondly, our method adopts case-crossover design with drug use as exposure. This method can automatically eliminate many underlying confounding factors and false positive signals, especially after the incorporation of temporal information of treatment outcomes and drug uses.

A previous study investigated a large-scale signal boosting approach combining FAERS and MEDLINE articles (Xu and Wang, 2014), where pairs appearing in both FAERS and MEDLINE sentences had the highest precision of 0.140 compared to 0.025 (FAERS) and 0.111 (FAERS + abstract), recall varied from 0.507 (FAERS), 0.138 (FAERS and sentences) to 0.234 (FAERS + abstract), and F1 varied from 0.045 (FAERS), 0.139 (FAERS and sentences) to 0.151 (FAERS + abstract). The study used SIDER as the gold standard, and the findings represented the performance level for large-scale databases like FAERS and MEDLINE as roughly from 0.025 to 0.507. Our results showed similar performances using another large-scale database, i.e., EMR.

Various sources of ADEs could exert different effects with unique advantages. FAERS includes voluntary spontaneous reports that may not capture the true incidence information. However, it contains millions of ADE-relevant records. Meanwhile, as a real world data source with the capability to yield valuable insights, EMR contains abundant information on treatment outcome and drug exposure. Therefore, a higher number of false positives by combining EMR and FAERS implies more potential novel ADEs detected. In addition, ADE signals detected by both FAERS and EMR can be prioritized for follow-up epidemiological or laboratory studies.

Limitations of our study include the following. First, our experiment used EMR data from a single institution. Secondly, the design of the case crossover studies did not exclude complex drug exposure situation and drug–drug interactions, and the carry over effects of previous drugs during the “control” period and other time-varying confounders that could result in ADE such as new comorbidities of patients. Additionally, the semantic granularity difference across diverse data sources has not been fully addressed even through the flexible mapping strategy. We plan to tackle those limitations by using EMR data from multiple institutions and investigate strategies to exclude complex drug exposure situation. Also, we can adopt semantic similarity to address the semantic granularity difference issue.

CONCLUSION

We proposed a method for detecting potential ADEs associated with drugs by combining two data sources: FAERS and EMR, and evaluated their performances using two ADE knowledge bases: ADReCS and SIDER. The use of NLP and the innovative adoption of case-crossover study design enable us to detect signals from large-scale EMR data. Findings showed that Recall

greatly increased when combining FAERS with EMR compared with FAERS alone and EMR alone, especially for flexible mapping strategy. Precision (FAERS + EMR) in detecting ADEs improved using ADReCS as gold standard compared with SIDER. In addition, signals detected from EMR have considerably overlapped with signals detected from FAERS or ADE knowledge bases, implying the importance of EMR for pharmacovigilance. ADE signals detected from EMR and/or FAERS but not in existing knowledge bases provide hypothesis for future study.

AUTHOR CONTRIBUTIONS

All co-authors are justifiably credited with authorship, according to the authorship criteria. Final approval is given by each co-author. LW: conception, design, development, data collection, analysis of data, interpretation of results, drafting and revision of manuscript. MR-M: design, data collection, analysis of data, interpretation of results and revision of manuscript. ZJ, SL, KL, SM, FS, and YW: analysis of data. LY: revision of manuscript. JD: interpretation of results and critical revision of manuscript. HL: conception, design, development, data collection, analysis of data, interpretation of results and critical revision of manuscript.

FUNDING

This work was made possible by joint funding from FDA FD05938-01 and National Institute of Health, R01GM102282, R01LM11369, R01LM011829, R01LM011934, and 1U24CA194215-01A1.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphar.2018.00875/full#supplementary-material>

FIGURE S1 | Bubble chart of precision, recall and patient number for three sources (ADReCS). The X axis denotes recall, the Y axis denotes precision and the bubble size denotes the number of patients.

FIGURE S2 | Precision of FAERS + EMR in detecting ADEs when using ADReCS and SIDER as the gold standards.

FIGURE S3 | Venn diagram of ADE signals detected from FAERS and EMR and ADEs captured in SIDER using the exact mapping strategy. Numbers indicate how many ADEs are there in each specific colored area.

FIGURE S4 | Venn diagram of ADE signals detected from FAERS and EMR and ADEs captured in ADReCS using the exact mapping strategy. Numbers indicate how many ADEs are there in each specific colored area.

FIGURE S5 | Venn diagram of ADE signals detected from FAERS and EMR and ADEs captured in ADReCS using the flexible mapping strategy. Numbers indicate how many ADEs are there in each specific colored area.

TABLE S1 | Performance of FAERS, EMR and FAERS + EMR in detecting ADEs when using ADReCS and SIDER as the gold standards.

REFERENCES

- Ahmed, I., and Poncet, A. (2013). *PhViD: An R package for Pharmacovigilance signal Detection. R package version 1.0. 6.2014.*
- Bate, A., and Evans, S. (2009). Quantitative signal detection using spontaneous ADR reporting. *Pharmacoepidemiol. Drug Saf.* 18, 427–436. doi: 10.1002/pds.1742
- Bousquet, C., Lagier, G., Lillo-Le Louët, A., Le Beller, C., Venot, A., and Jaulent, M.-C. (2005). Appraisal of the MedDRA conceptual structure for describing and grouping adverse drug reactions. *Drug Saf.* 28, 19–34. doi: 10.2165/00002018-200528010-00002
- Cai, M. C., Xu, Q., Pan, Y. J., Pan, W., Ji, N., Li, Y. B., et al. (2015). ADReCS: an ontology database for aiding standardization and hierarchical classification of adverse drug reaction terms. *Nucleic Acids Res.* 43, D907–D913. doi: 10.1093/nar/gku1066
- Classen, D. C., Resar, R., Griffin, F., Federico, F., Frankel, T., Kimmel, N., et al. (2011). 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. *Health Aff.* 30, 581–589. doi: 10.1377/hlthaff.2011.0190
- European Medicines Agency (2017). *Guideline on Good Pharmacovigilance Practices (GVP)*. London: European Medicines Agency. doi: 10.1111/apt.12385
- Evans, S. J., Waller, P. C., and Davis, S. (2001). Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiol. Drug Saf.* 10, 483–486. doi: 10.1002/pds.677
- Guo, M., Luo, H., Samii, A., and Etminan, M. (2016). The risk of glioblastoma with TNF inhibitors. *Pharmacotherapy* 36, 449–454. doi: 10.1002/phar.1731
- Gupta, R., Hennebray, T. A., and Kinasewitz, G. T. (2012). Pleuropericardial effusion after 37 years of sulfasalazine therapy. *J. Cardiovasc. Med.* 13, 541–543. doi: 10.2459/JCM.0b013e328343595d
- Hallas, J., and Pottegård, A. (2014). Use of self-controlled designs in pharmacoepidemiology. *J. Intern. Med.* 275, 581–589. doi: 10.1111/joim.12186
- Harpaz, R., Vilar, S., Dumouchel, W., Salmasian, H., Haerian, K., Shah, N. H., et al. (2013). Combing signals from spontaneous reports and electronic health records for detection of adverse drug reactions. *J. Am. Med. Inform. Assoc.* 20, 413–419. doi: 10.1136/amiainl-2012-000930
- Hripcsak, G., Duke, J. D., Shah, N. H., Reich, C. G., Huser, V., Schuemie, M. J., et al. (2015). Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers. *Stud. Health Technol. Inform.* 216, 574–578.
- Jiang, G., Wang, L., Liu, H., Solbrig, H. R., and Chute, C. G. (2013). Building a knowledge base of severe adverse drug events based on AERS reporting data using semantic web technologies. *Stud. Health Technol. Inform.* 2013, 496–500.
- Kuhn, M., Letunic, I., Jensen, L. J., and Bork, P. (2015). The SIDER database of drugs and side effects. *Nucleic Acids Res.* 44, D1075–D1079. doi: 10.1093/nar/gkv1075
- Lee, Y., and Bae, S.-C. (2016). Efficacy and safety of methotrexate plus certolizumab pegol or placebo in active rheumatoid arthritis. *Z. Rheumatol.* 76, 528–534. doi: 10.1007/s00393-016-0133-z
- Levy, B., Nisar, A., and Karanjia, N. (2006). Cholangiocarcinoma, renal cell carcinoma and parathyroid adenoma found synchronously in a patient on long-term methotrexate. *HPB* 8, 151–153. doi: 10.1080/13651820410016705
- Liu, X., and Chen, H. (2015). A research framework for pharmacovigilance in health social media: identification and evaluation of patient adverse drug event reports. *J. Biomed. Inform.* 58, 268–279. doi: 10.1016/j.jbi.2015.10.011
- Maclure, M., and Mittleman, M. A. (2000). Should we use a case-crossover design? *Annu. Rev. Public Health* 21, 193–221. doi: 10.1146/annurev.publhealth.21.1.193
- Pirmohamed, M., Breckenridge, A. M., Kitteringham, N. R., and Park, B. K. (1998). Adverse drug reactions. *Br. Med. J.* 316, 1295–1298. doi: 10.1136/bmj.316.7140.1295
- Polesie, S., Gillstedt, M., Sönnergren, H., Osmancevic, A., and Paoli, J. (2017). Methotrexate treatment and risk for cutaneous malignant melanoma: a retrospective comparative registry-based cohort study. *Br. J. Dermatol.* 176, 1492–1499. doi: 10.1111/bjd.15170
- Powers, D. M. (2011). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *Int. J. Mach. Learn. Technol.* 2, 37–63.
- Rennie, J. D. (2004). Derivation of the F-Measure. *Other Words* 1.
- Rothman, K. J., Lanes, S., and Sacks, S. T. (2004). The reporting odds ratio and its advantages over the proportional reporting ratio. *Pharmacoepidemiol. Drug Saf.* 13, 519–523. doi: 10.1002/pds.1001
- Schmitz, S., Adams, R., Walsh, C. D., Barry, M., and FitzGerald, O. (2011). A mixed treatment comparison of the efficacy of anti-TNF agents in rheumatoid arthritis for methotrexate non-responders demonstrates differences between treatments: a Bayesian approach. *Ann. Rheum. Dis.* 71, 225–230. doi: 10.1136/annrheumdis-2011-200228
- Schneeweiss, S., and Avorn, J. (2005). A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J. Clin. Epidemiol.* 58, 323–337. doi: 10.1016/j.jclinepi.2004.10.012
- Schuemie, M. J., Trifiro, G., Coloma, P. M., Ryan, P. B., and Madigan, D. (2016). Detecting adverse drug reactions following long-term exposure in longitudinal observational data: The exposure-adjusted self-controlled case series. *Stat. Methods Med. Res.* 25, 2577–2592. doi: 10.1177/0962280214527531
- Sohn, S., Clark, C., Halgrim, S. R., Murphy, S. P., Chute, C. G., and Liu, H. (2014). MedXN: an open source medication extraction and normalization tool for clinical text. *J. Am. Med. Inform. Assoc.* 21, 858–865. doi: 10.1136/amiainl-2013-002190
- Stephenson, W. P., and Hauben, M. (2007). Data mining for signals in spontaneous reporting databases: proceed with caution. *Pharmacoepidemiol. Drug Saf.* 16, 359–365. doi: 10.1002/pds.1323
- Stolberg, H. O., Norman, G., and Trop, I. (2004). Randomized controlled trials. *Am. J. Roentgenol.* 183, 1539–1544. doi: 10.2214/ajr.183.6.01831539
- Torii, M., Waghlikar, K., and Liu, H. (2011). Using machine learning for concept extraction on clinical documents from multiple data sources. *J. Am. Med. Inform. Assoc.* 18, 580–587. doi: 10.1136/amiainl-2011-000155
- van Puijenbroek, E. P., Bate, A., Leufkens, H. G., Lindquist, M., Orre, R., and Egberts, A. C. (2002). A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. *Pharmacoepidemiol. Drug Saf.* 11, 3–10. doi: 10.1002/pds.668
- Wang, L., Jiang, G., Li, D., and Liu, H. (2014). Standardizing adverse drug event reporting data. *J. Biomed. Semantics* 5:36. doi: 10.1186/2041-1480-5-36
- Wang, L., Li, M., Cao, Y., Han, Z., Wang, X., Atkinson, E. J., et al. (2017a). Proton pump inhibitors and the risk for fracture at specific sites: data mining of the FDA adverse event reporting system. *Sci. Rep.* 7:5527. doi: 10.1038/s41598-017-05552-1
- Wang, L., Rastegar-Mojarad, M., Liu, S., Zhang, H., and Liu, H. (2017b). Discovering adverse drug events combining spontaneous reports with electronic medical records: a case study of conventional DMARDs and biologics for rheumatoid arthritis. *AMA Jt. Summits Transl. Sci. Proc.* 2017, 95–103.
- Wang, X., Chase, H., Markatou, M., Hripcsak, G., and Friedman, C. (2010). Selecting information in electronic health records for knowledge acquisition. *J. Biomed. Inform.* 43, 595–601. doi: 10.1016/j.jbi.2010.03.011
- Wang, X., Hripcsak, G., Markatou, M., and Friedman, C. (2009). Active computerized pharmacovigilance using natural language processing, statistics, and electronic health records: a feasibility study. *J. Am. Med. Inform. Assoc.* 16, 328–337. doi: 10.1197/jamia.M3028
- Xu, R., and Wang, Q. (2014). Large-scale combining signals from both biomedical literature and the FDA Adverse Event Reporting System (FAERS) to improve

- post-marketing drug safety signal detection. *BMC Bioinformatics* 15:17. doi: 10.1186/1471-2105-15-17
- Yang, H., Xu, H., Liu, H., Ouyang, C., and Xu, W. (2011). A novel heparin release system based on blends of biomedical polyurethane and native silk fibroin powder. *J. Control. Release* 152(Suppl. 1), e106–e108. doi: 10.1016/j.jconrel.2011.08.155
- Yang, Y. (1999). An evaluation of statistical approaches to text categorization. *Inf. Retr.* 1, 69–90. doi: 10.1023/A:1009982220290
- Yu, Y., Chen, J., Li, D., Wang, L., Wang, W., and Liu, H. (2016). Systematic analysis of adverse event reports for sex differences in adverse drug events. *Sci. Rep.* 6:24955. doi: 10.1038/srep24955
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Wang, Rastegar-Mojarad, Ji, Liu, Liu, Moon, Shen, Wang, Yao, Davis and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.