


Analyzing adverse drug reaction using statistical and machine learning methods

A systematic review

Hae Reong Kim, MS^a, MinDong Sung, MD^a, Ji Ae Park, MS^a, Kyeongseob Jeong, BS^a, Ho Heon Kim, MS^a, Suehyun Lee, PhD^b, Yu Rang Park, PhD^{a,*} 

Abstract

Background: Adverse drug reactions (ADRs) are unintended negative drug-induced responses. Determining the association between drugs and ADRs is crucial, and several methods have been proposed to demonstrate this association. This systematic review aimed to examine the analytical tools by considering original articles that utilized statistical and machine learning methods for detecting ADRs.

Methods: A systematic literature review was conducted based on articles published between 2015 and 2020. The keywords used were statistical, machine learning, and deep learning methods for detecting ADR signals. The study was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (PRISMA) guidelines.

Results: We reviewed 72 articles, of which 51 and 21 addressed statistical and machine learning methods, respectively. Electronic medical record (EMR) data were exclusively analyzed using the regression method. For FDA Adverse Event Reporting System (FAERS) data, components of the disproportionality method were preferable. DrugBank was the most used database for machine learning. Other methods accounted for the highest and supervised methods accounted for the second highest.

Conclusions: Using the 72 main articles, this review provides guidelines on which databases are frequently utilized and which analysis methods can be connected. For statistical analysis, >90% of the cases were analyzed by disproportionate or regression analysis with each spontaneous reporting system (SRS) data or electronic medical record (EMR) data; for machine learning research, however, there was a strong tendency to analyze various data combinations. Only half of the DrugBank database was occupied, and the k-nearest neighbor method accounted for the greatest proportion.

Abbreviations: ADR = adverse drug reaction, ARM = association rule mining, BCPNN = The Bayesian Confidence Propagation Neural Network, CDM = Common Data Model, CRR = combination risk ratio, EBGM = Empirical Bayes Geometric Mean, EMR = electronic medical record, FAERS = FDA Adverse Event Reporting System, KEGG = Kyoto Encyclopedia of Genes and Genomes, LRT = log-likelihood ratio test, PRR = Proportional Reporting Ratio, RF = Random Forest, ROR = reporting odds ratio, SRS = spontaneous reporting system, SVM = support vector machine, VAERS = Vaccine Adverse Event Reporting System.

Keywords: adverse drug reaction, drug safety, machine learning method, pharmacovigilance, statistical method, systematic review

Ethics approval and consent to participate: Not applicable.

Consent for publication: All the authors approved the final manuscript.

Availability of data and material: This manuscript has no associated data and materials.

The authors report no conflicts of interest.

Funding: This study was supported by the Foundational Technology Development Program (NRF - 2019M3E5D4064682) the Ministry of Science and ICT, Republic of Korea and this research was also supported by a grant from the Korea Institute of Drug Safety and Risk Management in 2015.

Data sharing not applicable to this article as no datasets were generated or analyzed during the present study.

Supplemental Digital Content is available for this article.

Funding: This study was supported by the Foundational Technology Development Program (NRF - 2019M3E5D4064682) the Ministry of Science and ICT, Republic of Korea and this research was also supported by a grant from the Korea Institute of Drug Safety and Risk Management in 2021

^aDepartment of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, South Korea, ^bDepartment of Biomedical Informatics, Konyang University College of Medicine, Daejeon, South Korea.

* Correspondence: Yu Rang Park, Assistant Professor, Department of Biomedical Systems Informatics, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, South Korea (e-mail: yurangpark@yuhs.ac).

Copyright © 2022 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Kim HR, Sung M, Park JA, Jeong K, Kim HH, Lee S, Park YR. Analyzing adverse drug reaction using statistical and machine learning methods: a systematic review. *Medicine* 2022;101:25(e29387).

Received: 4 July 2021 / Received in final form: 12 April 2022 / Accepted: 12 April 2022

<http://dx.doi.org/10.1097/MD.00000000000029387>

1. Introduction

An adverse drug reaction (ADR) is an unintended negative response caused by the administration of a drug.^[1] In the United States, ADR accounted for almost 6% of all hospitalized patients in 2011, costing billions of dollars and generating significant morbidity and mortality. Studies on ADR are therefore relevant for improving patient safety. Spontaneous reporting system (SRS) data are the cornerstone of signal detection for patient safety. The ADR signal detection methods primarily exploit data from SRS using conventional statistical analysis methods.^[2] Statistical signal detection methods use a contingency table that relates the observed count of an adverse event of interest and a drug of interest in SRS data. However, SRS has several limitations and difficulties, such as under-reporting and bias, in detecting drug side effects.^[3,4] For instance, ADR reporting is influenced by a myriad of factors, including the severity of ADR, the duration of the drug's release on the market, the experience of medical professionals, and the qualifications of the doctors reporting it.^[5] Professional medical reports of adverse events often lack clarity regarding the diagnosis of adverse events. In fact, it is difficult to diagnose ADR even though most of them are included in the list of differential diagnoses available to doctors.^[6] In general, when the causal relationship is unclear, it is often not reported as an ADR.^[7] Therefore, many studies considering this limitation of SRS are in progress. Other sources of ADR study data include electronic medical record (EMR) data; these data are important for confirming clinical evidence. They provide more accurate temporal statistics on patients' experiences with health services, such as times of diagnosis, release of patients, and dates of start and completion of prescription orders.^[8] Research that relies on temporal data to examine the association between ADR induced by drugs can benefit from such information.^[8]

Typically, SRS data tend to center around signal detection using the reporting ratio of the statistical method. There are several methods, such as the reporting odds ratio (ROR), proportional reporting ratio (PRR), combination risk ratio (CRR), association rule mining (ARM) method, and the Bayesian statistical approach, which includes the Bayesian confidence propagation neural network (BCPNN) and the empirical Bayes geometric mean (EBGM).^[9,10] However, statistical methods are limited when it comes to analyzing free text or chemical structure data for signal detection.^[11] Therefore, machine learning techniques have emerged to make analysis of these forms of data feasible for ADR signal detection.^[2,12–16] Random forest (RF),^[17–20] adaboost,^[21] and neural network^[22] structures are actively used for these analyses. These 2 methods have provided clues regarding potential ADR and their mechanisms for further clinical verification of ADR.^[23] All data-driven methods for determining ADR depend on the quality of data sources and analytical methods involved.^[24] Although numerous studies have attempted to reveal ADR signals using different databases, only a few have focused on the methodology used. Thus, studies that concentrate on the methods to detect ADR signal are required using multiple databases. Our systematic review aimed to examine original articles that employed existing statistical and machine learning methods to detect ADR in humans.

2. Methods

2.1. Study selection and eligibility criteria

Our systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.^[25]

With the premise that a period of approximately 5 years is appropriate for establishing relatively recent research trends, we analyzed data obtained in the last five years based on the time of drafting this review article. The systematic literature search covered clinical research included in EMBASE and PubMed (including 100% of the MEDLINE database) and was conducted based on research published from January 2015 to March 2020, with an emphasis on original articles. Words that reflected adverse events, such as “adverse drug reactions,” “side effect,” and “drug safety,” were included. Since disproportionality analysis is widely used in the statistical analysis section, “statistical” and “disproportion” were included as additional search terms. For the machine learning component, both “machine learning” and “deep learning” were searched. Only cases in which the search term was included in the title and abstract were included (see Text Document and Graph, Supplemental Digital Content 1, <http://links.lww.com/MD2/B55>, <http://links.lww.com/MD2/B56> which details the search strategy and publication trends for each category by year). The aim of the present review was to determine the relationship between drugs and ADR using statistical or machine learning methods. This study is based on existing research; therefore, ethical approval was not necessary.

We reviewed the various analysis methods and the databases they corresponded to, with the purpose of detecting ADR. The analysis methods were divided into statistical and machine learning methods, and the method with the highest frequency was confirmed in detail. The statistical methods were divided into three categories: disproportionate analysis (eg, ROR and PRRs), regression (eg, survival, logistic, and Poisson), and Log-likelihood ratio test (LRT) (eg, LRT and zero-inflated Poisson-LRT). Machine learning methods were divided into three categories as follows: Bayesian methods (eg, Monte Carlo expectation maximization, MCEM), supervised methods (eg, random forest (RF), adaboost, support vector machine (SVM), and recurrent neural network), and other methods (eg, block matrices and matrix factorization).

2.2. Assessment for risk of bias

The first author (HRK) assessed the risk of bias in all included studies. The risk of bias was evaluated using the “Risk of Bias in Systematic Review (ROBIS)” tool.^[26] The existing systematic review section is limited because we could not perform quantitative analysis since our candidate papers were method-oriented articles. To supplement this, the ROBIS tool was considered suitable in terms of qualitative analysis and was applied and prepared. In particular, the ROBIS method is a systematic review-specific evaluation method and is the most commonly used method in qualitative analysis.^[27]

It evaluates the bias in 5 domains: Domain 1 is randomization, Domain 2 is the deviation from the intended intervention, Domain 3 is risk of missing, Domain 4 is the outcome measurement, and Domain 5 is configured to evaluate the bias against the selection of reported results. We conducted the evaluation according to statistical methods (see Tables and Graph, Supplemental Digital Content 2, <http://links.lww.com/MD2/B57>, <http://links.lww.com/MD2/B58>, <http://links.lww.com/MD2/B59>, <http://links.lww.com/MD2/B60>, <http://links.lww.com/MD2/B61>, <http://links.lww.com/MD2/B68> which details the assessment of the risk of bias for research with statistical methods) and machine learning methods (see Tables and Graph, Supplemental Digital

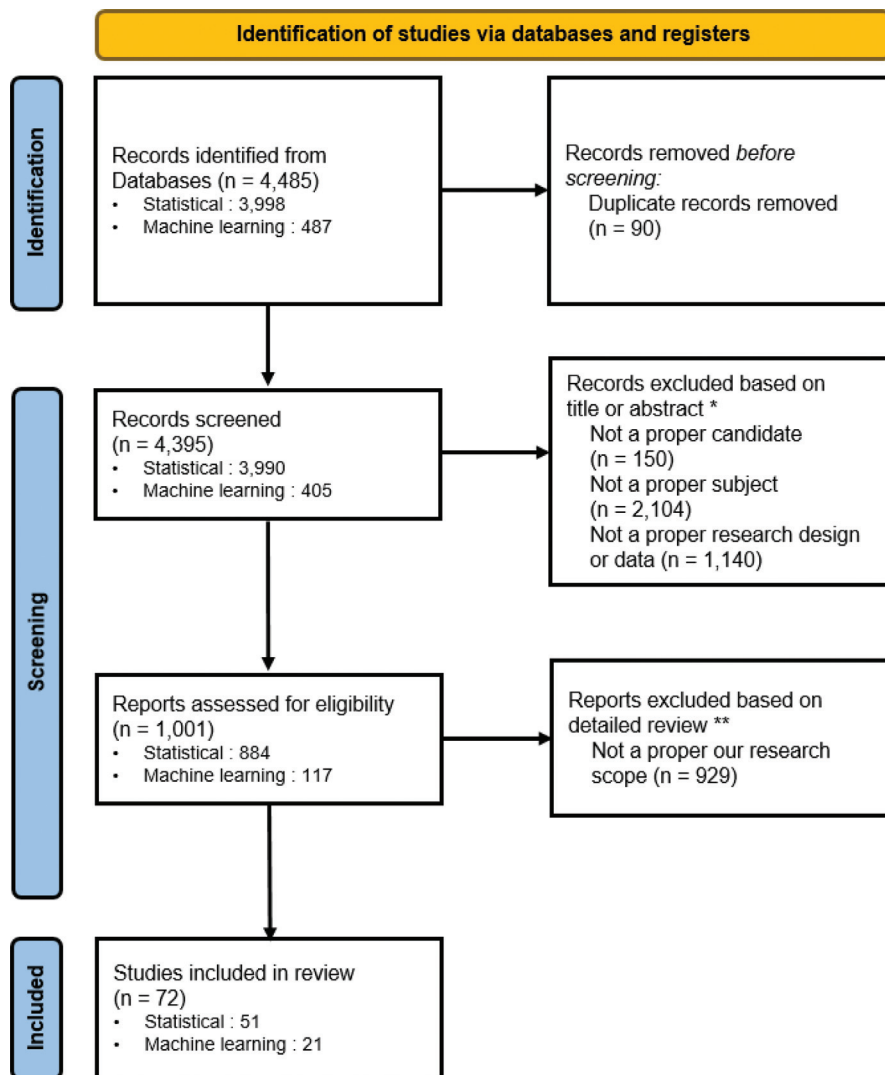


Figure 1. Publications removed based on title or abstract*; improper study subject (eg, bird, mouse, cats, and dog), improper candidate (eg, biology, genetic, gene expression, stem cell, HER2, DNA, biologics, β 1, beta blockers, mutation, inhibitor, genotype, chemical, pathway, T-cell, surgical, surgery, image, MRI, alcohol, smoke, marijuana, and diet), and improper research design (eg, randomized clinical trial, RCT, clinical trial, meta-analysis, pilot, systematic review, Delphi, and social media). Full-text articles excluded, based on manual reviews^{**}; In case of lack of a clear goal, improper candidate or research design that is not filtered out of search terms, and lack of drug-induced adverse event. Figure 1. PRISMA Flow Chart.

Content 3, <http://links.lww.com/MD2/B62>, <http://links.lww.com/MD2/B63>, <http://links.lww.com/MD2/B64>, <http://links.lww.com/MD2/B65>, <http://links.lww.com/MD2/B66>, <http://links.lww.com/MD2/B67> which details the assessment of the risk of bias for research with machine learning methods). Each figure was obtained using the “robvis” and “ggplot2” packages in the R software (version 3.6.3).

2.3. Visualization tools

The Sankey diagram is the most accessible tool for expressing all kinds of flows, and the width of each flow is determined based on its respective quantity. For each statistical and machine learning analysis, we began with the database and connected it to the method. Each figure was obtained using the “networkD3” package in the R software (version 3.6.3).

3. Results

Ninety duplicate articles were excluded (Fig. 1). The criteria for improper candidates, subjects, or designs led to the exclusion of 3394 articles. If the full text was not available for an article, or its goal was not clear, it was excluded. In total, we manually reviewed 72 articles, of which 51 and 21 were categorized as addressing statistical and machine learning methods, respectively (Fig. 2).

The databases listed in Table 1 were included as the data sources.^[12–17,20,22,28–60] This table shows database such as FAERS, SIDER, VigiBase, and other national specific database or web, app data that are associated with detecting ADR. In addition, we also included EMR data and other databases such as DrugBank, which contains information on drug targets, enzymes, and proteins related to metabolism. PubChem and KEGG DRUG also contain chemical information on drugs. The

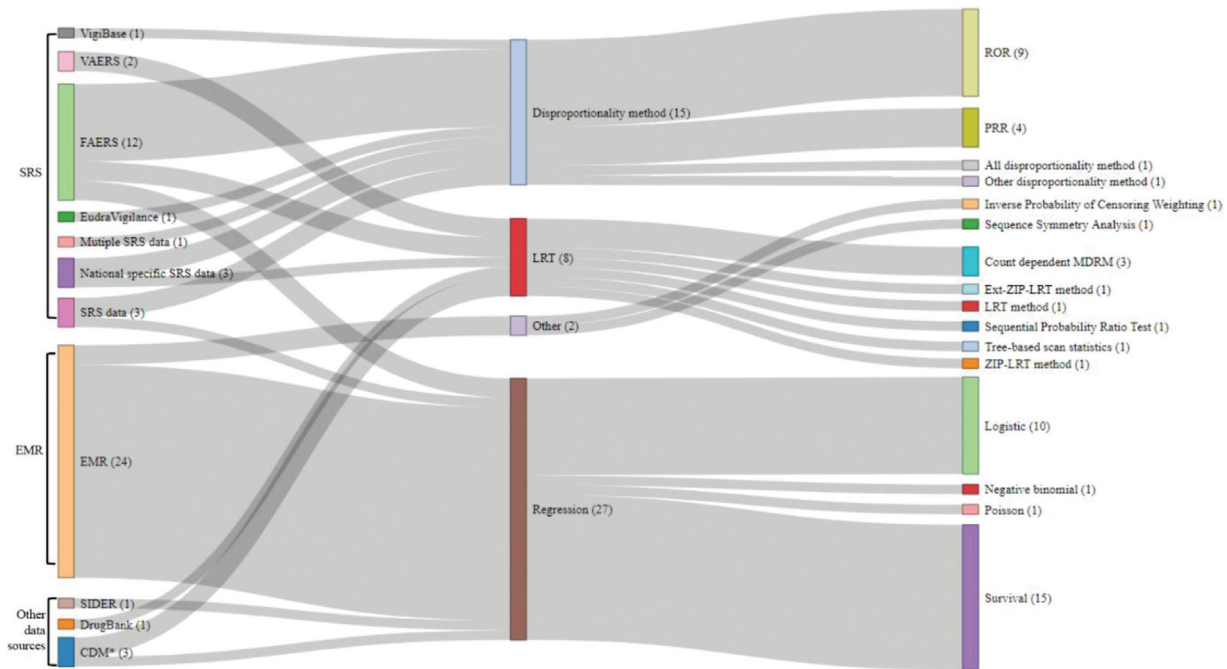


Figure 2. CDM* = Common data model. The use of multiple algorithms within one study may result in duplicate inclusions. Figure 2. Sankey diagram for statistical methods.

data sources in this review were categorized as follows: SRS data (eg, FAERS and VAERS), EMR, and other data sources (eg, DrugBank and PubChem).

3.1. Statistical methods for ADR detection

This ADR signal detection study aimed to reveal the association between drugs and ADR. We unified the statistical methods and expressed them as a single graph. As observed from the results, >80% of the total results were linked to statistical analyses using SRS and EMR data. Various methods have been applied to examine FAERS data, such as the disproportionality, LRT, and regression methods. The EMR data were analyzed exclusively with the regression method. In the disproportionality method, the ROR and PRR were mainly used. In the regression method, survival and logistic regression analyses were mostly used to determine the degree of risk.

3.1.1. Statistical methods for ADR detection in SRS.

Spontaneous adverse event reports collected under voluntary reporting systems were the major sources of structured data (Table 2).^[11,28–40,42,43,61–67] Some of the prominent SRS include the adverse event reporting system (maintained by the FDA) and VigiBase (maintained by the World Health Organization).^[11] The post-marketing phase is needed to monitor high-priority adverse events and gain insights into actual drug safety profiles by reflecting on concrete clinical practice. SRS represents a primary source of information for detecting safety signals, especially for newly marketed drugs and rare events with drug-related components.^[29]

Studies have used disproportionate analysis through FAERS. Raschi et al^[29] assessed the hepatic safety of novel oral anticoagulants. Rahman et al and Alatawi et al^[31,32] explored methods for brand versus generic ADR reports. Hoffman et al^[33] constructed a list of drug-induced adverse event signals. Takada

et al^[34] found that the use of sodium channel-blocking antiepileptic drugs is inversely associated with cancer development. Yu et al^[35] identified drugs that showed significant sex differences with RORs.

Other studies have used PRRs from pharmacovigilance data. Monaco et al^[28] found suspected ADR of drug products using the EudraVigilance data. Yue et al^[36] investigated reports of acute kidney injury events associated with nonsteroidal anti-inflammatory drugs. Chandler et al^[42] explored global reporting patterns of human papillomavirus vaccines. Sugawara et al^[43] evaluated the incidence of respiratory depression using opioids. Tan et al^[61] explored drugs related to injection-related ADR in children. Trinh et al^[62] optimized signal detection by investigating interest in time-series analysis with PRRs.

Log-likelihood ratio test (LRT) has not been used as frequently as disproportionate analysis. However, some studies employed LRT based on characteristics of the SRS. Cai et al^[37] proposed a testing procedure for the signal detection of temporal variation in ADR reporting using VAERS. Tong et al^[38] analyzed the vaccine FLU4, which can protect against four influenza viruses using VAERS. The authors assumed a zero-inflation-based Poisson model and performed an LRT to detect vaccine safety signals by testing the zero proportion and heterogeneity of reporting rates of vaccine-event combinations. Zhao et al^[39] identified ADR signals that have disproportionately high reporting rates compared with other ADRs and drug signals that have disproportionately high reporting rates for a group of ADR using extended LRT methods based on Poisson (Ext-LRT) and zero-inflated Poisson models (Ext-ZIP-LRT). Wang et al^[40] suggested a method with a fixed and count-dependent probability using mixture drug-count response models based on the number of combination drugs with a maximum risk threshold model. Chan et al^[63] explored the behavior of the sequential probability ratio test and its ability to detect signals of

Table 1
Including databases in this research.

Category	Database	Information from the database
SRS	EudraVigilance ^[28]	The database for adverse reactions to drug which have been authorised in clinical trials in the European Economic Area
	FAERS (including VAERS) ^[12–16,28–41]	Drug and ADR association for postmarketing drug safety surveillance from the Food and Drug Administration's
	VigiBase ^[42]	Individual Case Safety Reports of suspected ADRs
EMR	Other type of SRS ^[43]	National specific SRS database
	Medical records ^[17,44]	Institution specific standardized data (eg, diagnosis, medication)
Other data sources	Medical note ^[20,22,45]	Unstructured text data (eg, nursing records, surgery, and hospitalization records)
	SIDER ^[46–52]	The information of side effects and indication for marketed drugs
	DrugBank ^[14,17,41,46–49,51–55]	Non-redundant protein (drug target, enzyme, transporter, carrier, thus informing on drugs' mechanism of action and metabolism) sequences
	PubChem ^[47,49,52]	The chemical information of drugs, unique chemical structures, and biological activity data of chemical substances tested in assay experiments
	KEGG ^[47–49,52]	Drug, Compound and Disease databases providing chemical structures, targets, metabolizing enzymes
	Common Data Model ^[56]	A uniform set of metadata, allowing data and its meaning to be shared across applications (eg, OMOP CDM)
	Health Insurance system ^[57]	National specific health insurance system data (eg, NHIS)
	App, web data ^[58,90]	Data generated and collected through the app or web (eg, MedHelp)
	Registry ^[59]	National Data Registry (eg, cardiovascular disease)
	Simulation data ^[60]	Fake data created for specific situations for algorithm verification

disproportionate reporting with hypothesized relative risks. Assessments of the risk of bias for Table 2 were performed (Supplemental Digital Content 2, <http://links.lww.com/MD2/B57>, <http://links.lww.com/MD2/B58>, <http://links.lww.com/MD2/B59>, <http://links.lww.com/MD2/B60>, <http://links.lww.com/MD2/B61>, <http://links.lww.com/MD2/B68>).

3.1.2. Statistical methods for ADR detection in EMR data.

The EMR datasets that include the medical records of patients have proven to be useful materials in clinical research and have become an essential source for the analysis of patient medication in healthcare-related big data (Table 3).^[44,50,68–88] Proper analytical tools and EMR data are required for medication surveillance. Several studies have shown the value of pharmacovigilance research using EMR data as decision support tools; EMR include passive or active referential information, alerts, and guidelines related to ADR. Thus, EMR data may have considerable potential in pharmacovigilance research and can be used for rapid identification of patients in observational studies.^[89]

Examples of EMR include clinical data (eg, patient admission and discharge summaries and medications) and para-clinical data (eg, laboratory test results, radiographs, and diagnostic images).^[2] In contrast to existing surveillance data, there are various variables. Thus, ADR detection can be classified based on the goal of the study: reporting only drug and ADR information, correcting baseline information and ADR (patient-level prediction), and analyzing multiple drugs and ADR. Regression methods (eg, linear, logistic, Poisson, and survival) were used to determine the risk of independent variables affecting dependent variables. We summarized methods used to analyze which adverse events occur (dependent variable) depending on a specific drug (independent variable).

Regression methods were primarily used to analyze EMR data. Using survival analysis, the risk of ADR was calculated using single-center EMR.^[44,68–76] Using EMR, these studies examined the detection of ADR at multiple centers.^[77–79] Using

logistic regression,^[80–85] Khong et al analyzed multivariate negative binomial models to confirm that interleukin-2 therapy for metastatic melanoma and renal cell carcinoma affects rigors, which are significant ADR in a single center.^[86] Daley et al examined the safety of the live-attenuated influenza vaccine in a large multicenter cohort using conditional Poisson regression.^[87] The risk of bias for Table 3 was also assessed (see Supplemental Digital Content 2, <http://links.lww.com/MD2/B57>, <http://links.lww.com/MD2/B58>, <http://links.lww.com/MD2/B59>, <http://links.lww.com/MD2/B60>, <http://links.lww.com/MD2/B61>, <http://links.lww.com/MD2/B68>).

3.1.3. Statistical methods for ADR detection in other data sources.

In the statistical method, databases other than SRS and EMR were rarely used. Studies using the LRT with common data model (CDM) data are summarized in Table 4.^[56,57] Assessments of the risk of bias in Table 4 are presented in Supplemental Digital Content 2, <http://links.lww.com/MD2/B57>, <http://links.lww.com/MD2/B58>, <http://links.lww.com/MD2/B59>, <http://links.lww.com/MD2/B60>, <http://links.lww.com/MD2/B61>, <http://links.lww.com/MD2/B68>. Wang et al^[56] implemented tree-based scan statistics with propensity score-matched analyses using sentinel CDM. Tree-based scan statistics were defined as unconditional tree scan statistics that used the maximum log-likelihood ratio. In this report, exposure to a DPP4 inhibitor was analyzed, with sulfonylurea exposure serving as a comparator. The variables used were age, sex, chronic kidney disease, hypoglycemia, and diabetic nephropathy.

In addition, there is a method for measuring the risk by calculating the ratio according to the order of drug use and outcome. Maura et al^[57] used sequence symmetry analysis to classify patients according to their temporal sequence (outcome → oral anticoagulant (OAC) → outcome). This concept is used to evaluate the association between OAC initiation and the onset of non-bleeding adverse events (eg, renal, hepatic, skin, and gastrointestinal disease) by comparing symmetry.

Table 2
Statistical methods for ADR detection in SRS data.

Systems Author	Category of method	Method	Source	Purpose
Candore et al ^[11]	Disproportionate method	Almost all	Multiple SRS data	To compare the performance of commonly used algorithms detecting ADRs
Monaco et al ^[28]	Disproportionate method	PRR	EudraVigilance	to find out suspected ADRs
Raschi et al ^[29]	Disproportionate method	ROR	FAERS	To assess the hepatic safety of novel oral anticoagulants
Fukazawa et al ^[30]	Disproportionate method	ROR	FAERS	To conduct a disproportionality analysis and categorized these signals into groups which are signals with statistical significance and those without signals
Rahman et al and Alatawi et al ^[31,32]	Disproportionate method	ROR	FAERS	To compare whether adverse event reporting patterns are similar between brand and generic drugs
Hoffman et al ^[33]	Disproportionate method	ROR	FAERS	To construct a list of signal ADRs
Takada et al ^[34]	Disproportionate method	ROR	FAERS	To test that the use of sodium channel-blocking antiepileptic drugs are inversely associated with cancer
Yu et al ^[35]	Disproportionate method	ROR	FAERS	To assess the extent of sex differences in ADRs
Yue et al ^[36]	Disproportionate method	ROR	FAERS	To investigate acute kidney injury events associated with the concomitant use of oral acyclovir or valacyclovir with a Nonsteroidal anti-inflammatory drugs
Cai et al ^[37]	LRT	Likelihood ratio test	VAERS	To propose a powerful testing procedure for signal detection of temporal variation in ADR
Tong et al ^[38]	LRT	Likelihood ratio test based on zero-inflated poisson (ZIP-LRT)	VAERS	To identify four adverse events that are rare and have significantly different reporting rates for FLU4 vaccine
Zhao et al ^[39]	LRT	The extended likelihood ratio with Poisson model and zero-inflated Poisson model (Ext-ZIP-LRT)	FAERS	To identify ADR signals that have disproportionately high reporting rates
Wang et al ^[40]	LRT	Count-dependent probability mixture drug-count response model (MDRM)	FAERS and OMOP CDM	To introduce two novel mixture drug-count response models for detecting drug combinations of high dimension that induce myopathy
handler et al ^[42]	Disproportionate method	PRR	VigiBase database	To explore reporting patterns for HPV vaccine
Sugawara et al ^[43]	Disproportionate method	ROR	National specific SRS data	To evaluate the incidence of respiratory depression by use of opioids
Tan et al ^[61]	Disproportionate method	PRR	SRS data	To explore risks injection-related ADRs
Trinh et al ^[62]	Disproportionate method	PRR	National specific SRS data	To optimize signal detection investigating the interest of time-series analysis
Chan et al ^[63]	LRT	Sequential Probability Ratio Test	National specific SRS data	To detect signals of disproportionate reporting with the hRRs
Marbac et al ^[64]	Regression	Logistic regression with Metropolis–Hastings algorithm	SRS data	To identify a logistic regression with metropolis–hastings algorithm
Xu et al ^[65]	Regression	Logistic	FAERS	To identify secondary medications for mitigating the adverse effects of a primary drug
Pettit et al ^[66]	Regression	Logistic	FAERS	To figure out between posaconazole serum concentrations and toxicity
Lerch et al ^[67]	Disproportionate method	Signals of disproportionate reporting	SRS data	To detect unknown causal associations between drugs and unexpected events

ADR = adverse drug reaction.

3.2. Machine learning methods for ADR detection

Various computational methods, ranging from statistical methods to machine learning methods, have been used to detect

and predict new links between drugs and ADR. Machine learning, rooted in artificial intelligence framework, can be used to train computers with specific data patterns.^[2] It is considered more useful than statistical methods for analyzing complex

Table 3
Statistical methods for ADR detection in EMR data.

Author	Category of method	Method	Source	Purpose
Uozumi et al ^[44]	Regression	Survival	EMR	To investigate skin toxicity which is a common adverse event during cetuximab treatment
Jeong et al ^[50]	Regression	Comparison of Extreme Laboratory Test results, among others	EMR and SIDER	To propose a model that enables ADR signal detection from existing algorithms based on the EHR laboratory results for inpatient
Nishihara et al ^[68]	Regression	Survival	EMR	To investigate the relationships between increased blood pressure and bevacizumab administration
Otake et al ^[69]	Regression	Survival	EMR	To assess whether chemotherapy-induced neutropenia could be a prognostic factor and clarify other prognostic factors with metastatic pancreatic cancer patients
Kucharz et al ^[70]	Regression	Survival	EMR	To investigate cabozantinib-induced adverse events which are predictive factors of survival in case of sunitinib or axitinib
Dona et al ^[71]	Regression	Survival	EMR	To confirm that nonsteroidal anti-inflammatory drugs induced urticaria/angioedema
Gadelha et al ^[72]	Regression	Survival	EMR	To identify risk factors for death in patients who have suffered noninfectious ADR
Andrade et al ^[73]	Regression	Survival	EMR	to identify the risk factors for ADRs in pediatric inpatients
Westberg et al ^[74]	Regression	Survival	EMR	To assess the association of DTP likelihood of harm severity score, as measured by comprehensive medication management pharmacist after hospital discharge
Sobhonslidsuk et al ^[75]	Regression	Survival	EMR	To confirm that toxic liver diseases are mainly caused by drug-induced liver injury
Cordiner et al ^[76]	Regression	Survival	EMR	To test for Antipsychotic polypharmacy runs the risk of additional ADR and drug interactions
Merid et al ^[77]	Regression	Survival	EMR	To assess incidence and predictors of major adverse drug events among drug resistant tuberculosis patients
Oshikoya et al ^[78]	Regression	Survival	EMR	To determine the risk of serious ADR when oral azithromycin or intravenous/intramuscular fentanyl are used off-label compared to on-label in pediatric ICU
Okamoto et al ^[79]	Regression	Survival	EMR	To examine adverse event occurrence rates by grade, deaths and the appearance of severe ADR
Dedefo et al ^[80]	Regression	Logistic	EMR	To assess the incidence and determinants of medication errors and adverse drug events among hospitalized children
Blumenthal et al ^[81]	Regression	Logistic	EMR	To address inpatient penicillin allergies results in more broad-spectrum antibiotic use, treatment failures, and adverse drug events
Sellick et al ^[82]	Regression	Logistic	EMR	To measure the incidence and risk factors for fluoroquinolone-associated psychosis or delirium
Degu et al ^[83]	Regression	Logistic	EMR	To figure out hospital admissions which are due to drug related problems
Mill et al ^[84]	Regression	Logistic	EMR	To assess the accuracy and the negative predictive value of the graded provocation challenge in a cohort of children referred with suspected allergy to amoxicillin
llich et al ^[85]	Regression	Logistic	EMR	To determine whether female colorectal cancer patients experienced a higher incidence of dose-limiting toxicity than men when treated with adjuvant capecitabine
Khong et al ^[86]	Regression	Negative binomial	EMR	To affect the interleukin-2 therapy for metastatic melanoma and renal cell carcinoma
Daley et al ^[87]	Regression	Conditional poisson	EMR	To evaluate the safety for influenza vaccine in children
Vock et al ^[88]	other	Inverse Probability of Censoring Weighting	EMR	To propose a technique for mining right-censored time-to-event data

ADR = adverse drug reaction.

datasets. The machine learning methods were classified according to the criteria described in Section 2.2.

DrugBank was the most used database in machine learning, followed by the EMR, SIDER, and FAERS databases (Fig. 3). For method classification, other data sources accounted for the

largest proportion (21/40 cases, 52.5%). Among the other methods, the k-nearest neighbor method comprised the highest proportion, followed by matrix factorization. The supervised method had the second-largest proportion (13/40 cases, 32.5%). Among the supervised methods, the detection method using RF

Table 4
Statistical methods for ADR detection in other data sources.

Author	Category of method	Method	Source	Purpose
Wang et al ^[56]	LRT	Maximum log likelihood ratio	Common Data Model	to propose tree- based scan statistics to detect ADR signal
Maura et al ^[57]	Other	Sequence Symmetry Analysis	Health Insurance system	to assess the association between DOAC initiation and the onset of nonbleeding adverse events

ADR = adverse drug reaction, DOAC = Direct Oral Anti-Coagulants.

was the most common, followed by SVM and gradient-boosted trees. The Bayesian and semisupervised methods followed.

3.2.1. Machine learning methods for ADR detection in SRS.

For the machine learning approach, the Bayesian method has been used as a flexible and practical method that incorporates prior information (Table 5).^[12–16] It has also been used to identify important signals in ADR detection. Xiao et al^[12] used the MCEM and signal combination to determine drug safety signals. The authors extracted drug and ADR datasets to find a significant edge pair between the drug and ADR multiple Gamma Poisson Shrinkers. The authors calculated the selected ADR pair final MGPS score, which is an algorithm that derives the posterior probability. Signal combinations and MCEM were used to extract useful pairs, while other drugs were considered confounders and filtered out.

In addition, matrix-preprocessing method and a semi-supervised method were classified as “other methods” in the present study. Li et al used the inductive matrix completion (IMC) algorithm to predict potential drug–ADR associations using multiple data sources.^[14] The IMC method created drug and ADR matrices using the drug and ADR low-rank matrix based on chemical structure, cosine, or Jaccard similarity. Ren

et al used VAERS data with block matrices composed of correlation information.^[15] The block matrices that were merged by these vectors were calculated using neighboring information to calculate the distance between the vaccine and the ADR. Liu et al proposed a machine learning framework and identified potential high-priority DDIs.^[16] The authors used an auto-encoder-based semisupervised learning algorithm and a weighted SVM. They created reliable samples by combining labeled (FAERS) and unlabeled samples (ONC high-priority and DDI list), stacked the samples with an auto-encoder, and classified them using a weighted SVM to detect ADR. Assessments of the risk of bias for Table 5 were performed (see Supplemental Digital Content 3, <http://links.lww.com/MD2/B62>, <http://links.lww.com/MD2/B63>, <http://links.lww.com/MD2/B64>, <http://links.lww.com/MD2/B65>, <http://links.lww.com/MD2/B66>, <http://links.lww.com/MD2/B67>).

3.2.2. Machine learning methods for ADR detection in EMR data.

When using EMR data, all machine learning methods were classified as supervised methods (Table 6).^[17–21,45] Wang et al^[17] developed a data mining method for the systematic and automated detection of ADR. The authors used the RF method based on a set of positive or negative signals for known drug and

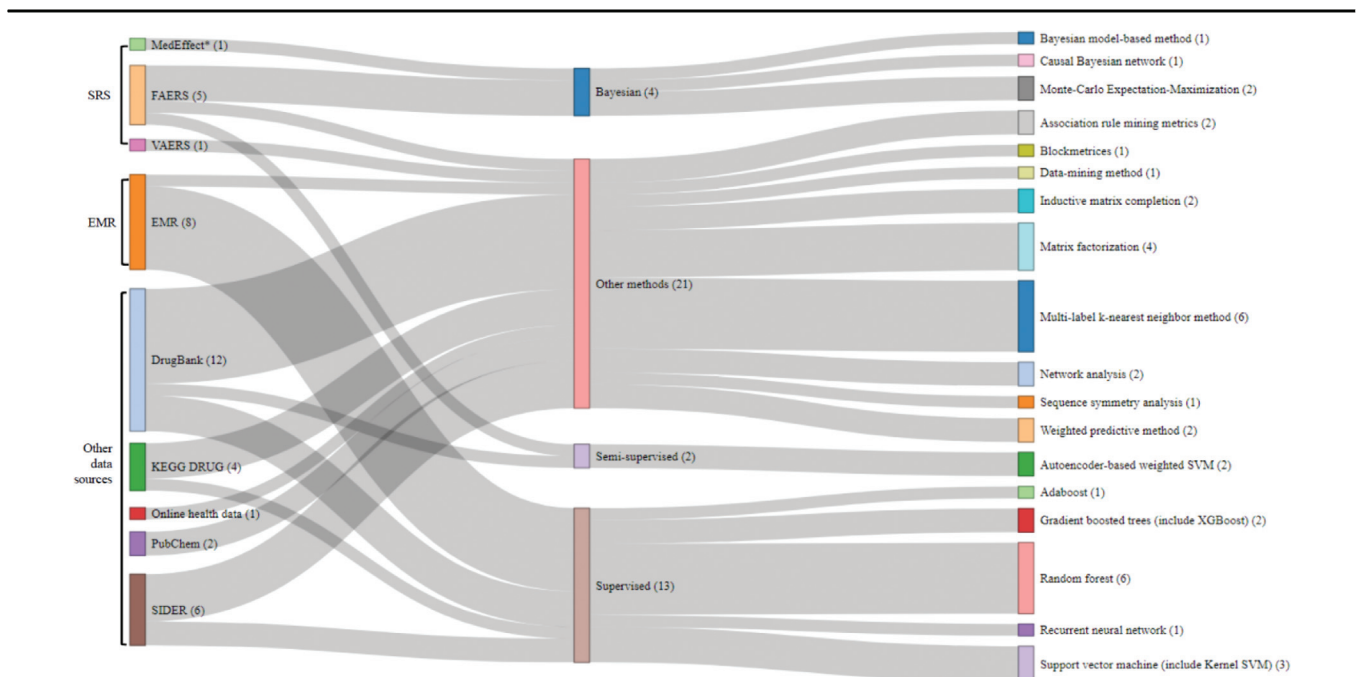


Figure 3. MedEffect* = National SRS data. The use of multiple algorithms within one study may result in duplicate inclusions. Figure 3. Sankey diagram for machine learning methods.

Table 5
Machine learning methods for ADR in SRS data.

Author	Category of method	Method	Source	Purpose
Xiao et al ^[12]	Bayesian	Monte-Carlo Expectation-Maximization procedure	FAERS, MedEffect, among others	To detects exact drug safety signals from multiple data sources via Monte Carlo Expectation Maximization and signal combination step
Cai et al ^[13]	Bayesian	Causal Bayesian Network	FAERS	To discover DDIs
Li et al ^[14]	Other methods	Inductive matrix completion	FAERS, DrugBank, among others	To find a random matrix value which minimized the distance between the drug and the ADR by the loss function and regularization
Ren et al ^[15]	Other methods	Blockmetrics with correlation	VAERS	To use correlation matrices to detect the adverse events or symptoms after vaccination
Liu et al ^[16]	Other methods	Autoencoder-Based Semi-Supervised Learning Algorithm and weighted SVM	FAERS and ONC High-Priority	To propose a machine learning framework to extract useful features and identify potential highpriority DDIs

ADR = adverse drug reaction, DDI = Drug-Drug Interaction.

ADR pairs from databases, such as clinical notes and DrugBank. Zhao et al^[18] explored data using diagnostic information, drug administration, clinical measurements, laboratory tests, and clinical notes. They used ADR-related diagnosis codes as class labels with a RF. In another study, the RF method was used to calculate weights that indicated the importance of clinical events for ADR detection using drugs, diagnoses, measurements from laboratory tests, and clinical notes.^[19] The authors extracted the contribution of each variable as a weight (weight aggregation and sampling) and then applied the RF algorithm. Wang et al calculated the feasibility of multiclass classification for identifying ADR using regularized logistic regression and SVM.^[45] Boyce examined the value of text mining by identifying suspected bleeding ADR from admission notes.^[20] They used RF analysis and other classification methods. Wunnava et al developed rule-based tokenization techniques to minimize noise in EMR notes using an embedding method with a recurrent neural network.^[22] These notes were annotated with medication information (eg, medication name, dose, route, frequency, and duration), ADR, indications, and other signs and symptoms. We assessed the risk of bias in Table 6 (see Supplemental Digital Content 3, <http://links.lww.com/MD2/B62>, <http://links.lww.com/MD2/B63>, <http://links.lww.com/MD2/B64>, <http://links.lww.com/MD2/B65>, <http://links.lww.com/MD2/B66>, <http://links.lww.com/MD2/B67>).

3.2.3. Machine learning methods for ADR detection in other data sources.

The list included in the other data sources is presented in Table 7.^[41,46–49,51–53,55,58–60,90] Bean et al considered ADR caused by lead compounds and predicted new ADR from available information on marketed drugs.^[46] Their analysis used the weighted predictive method. The authors constructed a knowledge graph that consisted of four types of nodes and edges, which were composed of drugs, protein targets, indications, and adverse reactions, each indicating the weighted feature score for the ADR pair. Zhang et al determined unobserved drug side effects based on known associations between drugs, ADR, and available drug features.^[47] They used feature-derived graph-regularized matrix factorization. The drug and ADR association matrix was decomposed into 2 low-rank matrices, which uncovered the latent features of the drugs and ADR.

Several articles were classified under supervised methods for other data sources. Zhao et al used a RF algorithm to identify ADR based on drug similarity.^[48] Muñoz et al used a machine learning method with a feature selection-based multilabel k-nearest neighbor method to explore the effects of knowledge graph-machine-readable interlinked representations of biomedical knowledge as a convenient uniform representation of heterogeneous data, and ADR detection as a multi-label ranking problem.^[49] Song et al used a pairwise kernel SVM classifier.^[51] The established similarity measures included molecular structure similarity, interaction profile fingerprint similarity (between two drugs that codify the known interaction), target similarity (integration of drug targets, enzymes, transporters, and carrier data to calculate the drug target fingerprints using the Jaccard score), and adverse drug effect similarity as a vector that codifies the presence (1) or absence (0) of the adverse effects in different bit positions from DrugBank and SIDER. After calculating the similarity, the authors generated a similarity matrix and predicted unknown drug and ADR pairs using a pairwise kernel SVM classifier. Davazdahemami et al used gradient-boosted trees. The authors calculated the similarity index of drugs from DrugBank and predicted drug and ADR associations, especially high-risk ADR from MEDLINE.^[53] Liu et al used the XGBoost algorithm to predict the risk of analgesic side effects and provided information on the interpretability of the model.^[58] The authors developed a model based on the Osteoarthritis Initiative dataset, which includes the demographic features, medical history, and physical examination data of patients with osteoarthritis. The assessments of the risk of bias in Table 7 are presented in Supplemental Digital Content 3, <http://links.lww.com/MD2/B62>, <http://links.lww.com/MD2/B63>, <http://links.lww.com/MD2/B64>, <http://links.lww.com/MD2/B65>, <http://links.lww.com/MD2/B66>, <http://links.lww.com/MD2/B67>.

4. Discussion

4.1. Summary of findings

Here, we reviewed the landscape of technological methods for ADR signal detection using multiple data sources. There have been several systematic review articles on ADR signal detection. There are existing reference papers that employed different

Table 6
Machine learning methods for ADR in EMR data.

Author	Category of method	Method	Source	Purpose
Wang et al ^[17]	Supervised	Random Forest	EMR, DrugBank, and etc.	To develop data-mining method for detection of ADRs
Zhao et al ^[18]	Supervised	Random Forest	EMR	To detect the drug-induced diagnosis ADRs
Zhao et al ^[19]	Supervised	Random Forest	EMR	To learn weights for ADRs detection
Boyce et al ^[20]	Supervised	Random Forest and so on	Admission notes	To show the value of text mining for identifying suspected bleeding ADRs
Desautels et al ^[21]	Supervised	AdaBoost	EMR	To identify patients who suffer from ICU readmission
Wunnava et al ^[22]	Supervised	Bi-directional long short-term memory, among others	EMR notes	To develop rule-based tokenization techniques for ADRs detection
Wang et al ^[45]	Supervised	Regularized logistic regression, linear support vector machine	EMR notes	To evaluate the feasibility of multiclass classification for ADRs

ADR = adverse drug reaction.

methods according to SRS data, social media, and others.^[2] However, it is difficult to intuitively check the relationship between the database and the method. Another paper summa-

rized the latest machine learning methods based on multiomics and SRS data. However, these failed to provide clarification on which statistical methods are mainly used for SRS data.^[2,3]

Table 7
Machine learning methods for ADR in other data sources.

Author	Category of method	Method	Source	Purpose
Kastrin et al ^[41]	Other methods	Unsupervised and supervised method	DrugBank, KEGG, NDF-RT, and Twosides	To represent the process of discovering potential DDIs and to evaluate performance of unsupervised and supervised machine learning methods
Bean et al ^[46]	Other methods	Weighted predictive method	DrugBank, SIDER, and EMR	To use knowledge about drugs known to cause an ADR to predict new causes
Zhang et al ^[47]	Other methods	Feature-derived graph regularized matrix factorization	SIDER, DrugBank, KEGG DRUG and PubChem	To predict ADRs based on known drug-side effect associations
Zhao et al ^[48]	Supervised	Random forest	STITCH, KEGG, DrugBank, RDKit and SIDER	To detect the ADRs
Muñoz et al ^[49]	Supervised	Feature selection-based multi-label k-nearest neighbour	PubChem and Bio2RDF dataset (DrugBank, SIDER, KEGG, etc.)	To explore effects of: using knowledge graphs as a representation of heterogeneous data; and casting ADRs prediction as a multilabel ranking problem
Song et al ^[51]	Supervised	Pairwise kernel SVM classifier	DrugBank and SIDER	To predict drug pairs and check if they truly interact with each other
Zhang et al ^[52]	Supervised	Feature selection-based multi-label k-nearest neighbor method	SIDER, PubChem, DrugBank, KEGG DRUG, and etc.	To build the association between feature and ADR vector for multilabel learning
Davazdahemami et al ^[53]	Supervised	Gradient boosted trees	MEDLINE and DrugBank	To predict the drug and ADR associations
Hoang et al ^[55]	Supervised	Sequence symmetry method	DrugBank	To assess the utility of supervised machine learning as a signal detection tool for ADRs
Liu et al ^[58]	Supervised	XGBoost	Osteoarthritis Initiative dataset	To identify high-risk features of cardiovascular diseases caused by analgesics OA patients
Ross et al ^[59]	Bayesian	Bayesian method	Cardiology's National Cardiovascular Data Registry	To provide insights into whether multiple methods used as an ensemble to detect all safety signals
Cotterill et al ^[60]	Bayesian	Bayesian method	Simulated data	To account for a subgroup effect to ADRs by including covariates
Yang et al ^[90]	Supervised	Association rule mining metrics	MedHelp	To propose a framework for drug safety signal detection by harnessing online health community data which associated ADRs and DDIs

ADR = adverse drug reaction, DDI = Drug-Drug Interaction.

Therefore, our study aimed to observe the latest trends and intuitively suggest which analysis method can be selected based on the relationship between data and methods. These methods consist of statistical and machine learning approaches, and the summary was achieved using a visualization tool and a Sankey diagram.

The type of data used to detect ADR signals differed according to the category of the method. More than 90% of the statistical methods were used to analyze SRS and EMR data, whereas 43% of machine learning methods were used for SRS and EMR data. More than 50% of the studies used other data sources, such as DrugBank or KEGG, which include drug interactions, pathways, and structures. Statistical methods generally used a single database for each study. Conversely, machine learning methods made use of multiple databases for detection of ADR signals.

For the statistical approach, >80% of the total results were related to statistical analysis with SRS and EMR data. The disproportionality, LRT, and regression methods were utilized for the SRS data. The EMR data were analyzed exclusively using the regression method.

A different set of characteristics was observed for the machine learning approach. Machine learning has a trend-complex dataset for ADR signal detection compared to statistical methods. DrugBank was the most used database in machine learning, followed by EMR, SIDER, and FAERS databases. The remaining databases accounted for the largest proportion (52.5%). Among the other methods, the k-nearest neighbor method comprised the highest proportion. The supervised method had the second-largest proportion (32.5%). In our study, RF was most frequently used in ADR signal detection, followed by SVM and gradient-boosted trees. Furthermore, the risk of bias evaluation showed that while the statistical component was generally high, the machine learning component was generally low.

There was an additional difference with respect to the time of the study between the statistical and machine learning methods (see Supplemental Digital Content 1, <http://links.lww.com/MD2/B55>, <http://links.lww.com/MD2/B56>). Until 2016 and 2017, statistical methods accounted for a higher proportion of studies than machine learning methods. In 2018, however, 38.1% of the studies involved machine learning methods, whereas 33.4% involved statistical methods.

Although SRS data have been the cornerstone of signal detection for drug safety, the data have some issues, such as underreporting, selective reporting, or the absence of information about actual exposed patients. This can potentially hinder the identification of safety signals.^[91] Monitoring of SRS is voluntary, and SRS does not include all ADRs. Thus, studies have found that as many as 90% of serious ADR are unreported.^[92] The underreporting of ADR can trigger a delay in the signal detection of ADR and marketing of a drug and raise risk estimates, resulting in false positives.^[92] We examined the use of statistical methods, considering the limitations of these data. Two statistical methods have been used to control sparsity: the metropolis-hasting algorithm^[64] which deals with the sparsity of logistic regression, and the zero-inflated Poisson model^[38,39] for a true zero distribution. Another analysis method is the LRT. The LRT comprised the third largest proportion, followed by the disproportionate and regression methods. The LRT is primarily used because it can address false positives, such as those typically seen in FAERS data, and display satisfactory power and sensitivity, controlling type-1 errors.^[39] Several studies have attempted to

modify the LRT method. Since the LRT method itself is difficult to analyze patient-level data using the LRT method itself, Zhao et al stratified LRT statistics according to age and sex among demographic factors.^[39] Wang et al used the LRT method to represent the probability of risk-associated DDIs for multiple drugs as a log-likelihood function.^[40]

The Bayesian method comprised the lowest proportion even though it had flexible and practical characteristics in the machine learning method.^[60] Therefore, it was confirmed to be useful in research on ADR signal detection. The Bayesian method reduces the computational cost and robustly preserves the predictive performance of high-dimensional data. However, complexity is a critical problem when using this method. RF is the most used supervised machine learning method. The RF method, which is a strong classifier that exhibits favorable performance, can be applied to signal detection models.^[48] XGBoost is the most used RF method. The XGBoost is a sparsity-aware algorithm. It can predict side effects in large cases of disease.^[93]

In addition, the multilabel k-nearest neighbor method has been used as a flexible and robust approach; it is similar to the Bayesian method.^[49] This method can control the dimensionality by embedding it into low-rank feature spaces. It can also determine the optimal feature dimensions and derive high accuracy.^[94] Other machine learning methods are matrix based. Among matrix-based methods, matrix factorization can detect unobserved associations based on a known association matrix. The method suggests that centrality-based matrices predict the network edges of drug and ADR pairs better than existing similarity-based matrices.^[53]

4.2. Limitations

4.2.1. Limitations of the methods used in ADR assessment.

There are some limitations to the statistical methods. Except for the regression method, patient-level analysis using SRS or other databases is difficult to achieve. It is mainly an analysis method for group-level analysis and cannot use the demographic information of an individual. A stratified statistic adjusting the baseline covariate can be suggested as an alternative to overcome this limitation.^[39] This method is appropriate for analyzing baseline covariates. However, there are some limitations to this. Stratified statistics become complicated when the number of adjusted baseline covariates increases. When considering multiple prescriptions, the concept of DDI was introduced to consider multiple counts and determine the count-dependent probability.^[40] However, it is difficult to determine whether these DDIs consider actual drug interactions. Additionally, they are limited with respect to weight adjustment according to the number of drugs administered. In a statistical method, more intensive research is needed to properly reflect these covariates. Further studies are needed to determine how to use the interaction information for multiple drugs.

There are also some limitations to machine learning methods. Owing to the limitations of data-driven methods, the results of the predictive model may not be comprehensive, and it is difficult to distinguish clinically relevant signals. The weighted method was used in the RF, but the interpretation for results is also an additional consideration when using other methods (eg, information gain) rather than weight.^[19]

4.2.2. Limitations of the review. The limitation of XGBoost is the complexity and highly relevant relationships between risk factors.^[93] There are limitations not only for the research

analysis methods but also for the present study. Both statistical and machine learning methods capture an appropriate risk window for the calculation of ADR.^[56] Because there are no “gold standard” safety signal detection methods in databases, it is unknown which method is the most effective and reliable.^[59] Although both statistical and machine learning methods have been reviewed, it is difficult to conclude which method has the best performance because of the diverse characteristics of the data and methods used in the studies. In this review, publication bias may have existed depending on the selected article.

5. Conclusions

Seventy-two articles using statistical and machine learning methods for predicting ADRs were identified. This systematic review followed the PRISMA guidelines to analyze the methods thoroughly. Our main finding was that, for the statistical analysis, >90% of the cases were analyzed by disproportionate or regression analysis with SRS or EMR data. In contrast, in machine learning research, there was a strong tendency to analyze various data combinations. Only 50% of the DrugBank database containing drug and drug target information was available. For the machine learning research, we expected the more conventional supervised analysis to dominate. However, detection of ADRs using various other methods was more common. Out of these, the k-nearest neighbor method accounted for the largest proportion. Our trend analysis is expected to serve as a guideline for researchers in the future. Our future work will reveal which method is optimal by comparing performance indicators to measure the utility of the methodology.

Author contributions

HRK developed the study concept and design, data analysis, data interpretation, drafting, and revision of the manuscript; MDS, JAP, and HHK developed the study concept and design, and reviewed this manuscript; SHL developed the review of this manuscript; YRP developed the study concept and design, data interpretation, drafting, and revision of the manuscript. All authors contributed to and approved the final manuscript.

Conceptualization: Yu Rang Park

Data curation: Hae Kim, Yu Rang Park

Funding acquisition: Yu Rang Park

Methodology: Hae Kim

Project administration: Yu Rang Park

Resources: Yu Rang Park

Software: Hae Kim

Validation: MinDong Sung, Yu Rang Park

Visualization: Hae Kim

Writing – original draft: Hae Kim

Writing – review & editing: Ho Kim, Ji Park, Kyeongseob Jeong, MinDong Sung, Suehyun Lee, Yu Rang Park

References

- Lee CY, Chen YPP. Machine learning on adverse drug reactions for pharmacovigilance. *Drug Discov Today* 2019;24:1332–43.
- Ho TB, Le L, Thai DT, Taewijit S. Data-driven approach to detect and predict adverse drug reactions. *Curr Pharm Des* 2016;22:3498–526.
- Alomar M, Tawfiq AM, Hassan N, Palaian S. Post marketing surveillance of suspected adverse drug reactions through spontaneous reporting: current status, challenges and the future. *Ther Adv Drug Saf* 2020;11:2042098620938595.
- Ibrahim H, Abdo A, El Kerdawy AM, Eldin AS. Signal detection in pharmacovigilance: a review of informatics-driven approaches for the discovery of drug-drug interaction signals in different data sources. *Artificial Intelligence in the Life Sciences* 2021;1:100005.
- Faillie J-L, Montastruc F, Montastruc J-L, Pariente A. Pharmacoepidemiology and its input to pharmacovigilance. *Therapies* 2016;71:211–6.
- Vallano A, Cereza G, Pedrós C, et al. Obstacles and solutions for spontaneous reporting of adverse drug reactions in the hospital. *Br J Clin Pharmacol* 2005;60:653–8.
- Hasford J, Goettler M, Munter KH, Müller-Oerlinghausen B. Physicians' knowledge and attitudes regarding the spontaneous reporting system for adverse drug reactions. *J Clin Epidemiol* 2002;55:945–50.
- Zhang X et al., Comparing Pharmacovigilance Outcomes Between FAERS and EMR Data for Acute Mania Patients, IEEE International Conference on Healthcare Informatics Workshop (ICHI-W), 2018.
- Wang X, Li L, Wang L, Feng W, Zhang P. Propensity score-adjusted three-component mixture model for drug-drug interaction data mining in FDA Adverse Event Reporting System. *Stat Med* 2020;39:996–1010.
- Noguchi Y, Tachi T, Teramachi H. Detection algorithms and attentive points of safety signal using spontaneous reporting systems as a clinical data source. *Brief Bioinform* 2021;22:bbab347.
- Candore G, Juhlin K, Manlik K, et al. Comparison of statistical signal detection methods within and across spontaneous reporting databases. *Drug Saf* 2015;38:577–87.
- Xiao C, Li Y, Baytas IM, Zhou J, Wang F. An MCEM framework for drug safety signal detection and combination from heterogeneous real world evidence. *Sci Rep* 2018;8:1806.
- Cai R, Liu M, Hu Y, et al. Identification of adverse drug-drug interactions through causal association rule discovery from spontaneous adverse event reports. *Artif Intell Med* 2017;76:7–15.
- Li R, Dong Y, Kuang Q, et al. Inductive matrix completion for predicting adverse drug reactions (ADRs) integrating drug-target interactions. *Chemosometrics and Intelligent Laboratory Systems* 2015;144:71–9.
- Ren JJ, Sun T, He Y, Zhang Y. A statistical analysis of vaccine-adverse event data. *BMC Med Inform Decis Mak* 2019;19:101.
- Liu N, Chen CB, Kumara S. Semi-supervised learning algorithm for identifying high-priority drug-drug interactions through adverse event reports. *IEEE J Biomed Health Inform* 2020;24:57–68.
- Wang G, Jung K, Winnenburg R, Shah NH. A method for systematic discovery of adverse drug events from clinical notes. *J Am Med Inform Assoc* 2015;22:1196–204.
- Zhao J, Henriksson A, Asker L, Bostrom H. Predictive modeling of structured electronic health records for adverse drug event detection. *BMC Med Inform Decis Mak* 2015;15(suppl 4):S1.
- Zhao J, Henriksson A. Learning temporal weights of clinical events using variable importance. *BMC Med Inform Decis Mak* 2016;16(suppl 2):71.
- Boyce RD, Jao J, Miller T, Kane-Gill SL. Automated screening of emergency department notes for drug-associated bleeding adverse events occurring in older adults. *Appl Clin Inform* 2017;8:1022–30.
- Desautels T, Das R, Calvert J, et al. Prediction of early unplanned intensive care unit readmission in a UK tertiary care hospital: a cross-sectional machine learning approach. *BMJ Open* 2017;7:e017199.
- Wunnavu S, Qin X, Kakar T, Sen C, Rundensteiner EA, Kong X. Adverse drug event detection from electronic health records using hierarchical recurrent neural networks with dual-level embedding. *Drug Saf* 2019;42:113–22.
- Nguyen DA, Nguyen CH, Mamitsuka H. A survey on adverse drug reaction studies: data, tasks and machine learning methods. *Brief Bioinform* 2019.
- Park RW. The distributed research network, observational health data sciences and informatics, and the South Korean research network. *Korean J Med* 2019;94:309–14.
- Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLOS Med* 2009;6:e1000100.
- McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): an R package and Shiny web app for visualizing risk-of-bias assessments. *Research Synthesis Methods* 2020;n/a.
- Jakic M, Jager M, Kosnik M. Predictive value of a negative oral provocation test in patients with hypersensitivity to analgesics. *Acta Dermatovenerol Alp Pannonica Adriat* 2016;25:27–30.

- [28] Monaco L, Melis M, Biagi C, et al. Signal detection activity on EudraVigilance data: analysis of the procedure and findings from an Italian Regional Centre for Pharmacovigilance. *Expert Opin Drug Saf* 2017;16:271–5.
- [29] Raschi E, Poluzzi E, Koci A, et al. Liver injury with novel oral anticoagulants: assessing post-marketing reports in the US Food and Drug Administration adverse event reporting system. *Br J Clin Pharmacol* 2015;80:285–93.
- [30] Fukazawa C, Hinomura Y, Kaneko M, Narukawa M. Significance of data mining in routine signal detection: analysis based on the safety signals identified by the FDA. *Pharmacoepidemiol Drug Saf* 2018;27:1402–8.
- [31] Rahman MM, Alatawi Y, Cheng N, et al. Methodological considerations for comparison of brand versus generic versus authorized generic adverse event reports in the US Food and Drug Administration Adverse Event Reporting System (FAERS). *Clin Drug Investig* 2017;37:1143–52.
- [32] Alatawi Y, Rahman MM, Cheng N, et al. Brand vs generic adverse event reporting patterns: an authorized generic-controlled evaluation of cardiovascular medications. *J Clin Pharm Ther* 2018;43:327–35.
- [33] Hoffman KB, Dimbil M, Tatonetti NP, Kyle RF. A Pharmacovigilance signaling system based on FDA regulatory action and post-marketing adverse event reports. *Drug Saf* 2016;39:561–75.
- [34] Takada M, Fujimoto M, Motomura H, Hosomi K. Inverse association between sodium channel-blocking antiepileptic drug use and cancer: data mining of spontaneous reporting and claims databases. *Int J Med Sci* 2016;13:48–59.
- [35] Yu Y, Chen J, Li D, Wang L, Wang W, Liu H. Systematic analysis of adverse event reports for sex differences in adverse drug events. *Sci Rep* 2016;6:24955.
- [36] Yue Z, Shi J, Li H, Li H. Association between Concomitant Use of Acyclovir or Valacyclovir with NSAIDs and an Increased Risk of Acute Kidney Injury: Data Mining of FDA Adverse Event Reporting System. *Biol Pharm Bull* 2018;41:158–62.
- [37] Cai Y, Du J, Huang J, et al. A signal detection method for temporal variation of adverse effect with vaccine adverse event reporting system data. *BMC Med Inform Decis Mak* 2017;17:76.
- [38] Tong J, Huang J, Du J, Cai Y, Tao C, Chen Y. Identification of rare adverse events with year-varying reporting rates for FLU4 vaccine in VAERS. *AMIA Annu Symp Proc* 2018;1544–51.
- [39] Zhao Y, Yi M, Tiwari RC. Extended likelihood ratio test-based methods for signal detection in a drug class with application to FDA's adverse event reporting system database. *Stat Methods Med Res* 2018;27:876–90.
- [40] Wang X, Zhang P, Chiang CW, et al. Mixture drug-count response model for the high-dimensional drug combinatorial effect on myopathy. *Stat Med* 2018;37:673–86.
- [41] Kastrin A, Ferk P, Leskosek B. Predicting potential drug-drug interactions on topological and semantic similarity features using statistical learning. *PLoS One* 2018;13:e0196865.
- [42] Chandler RE, Juhlin K, Fransson J, Caster O, Edwards IR, Noren GN. Current safety concerns with human papillomavirus vaccine: a cluster analysis of reports in VigiBase(R). *Drug Saf* 2017;40:81–90.
- [43] Sugawara H, Uchida M, Suzuki S, et al. Analyses of respiratory depression associated with opioids in cancer patients based on the Japanese Adverse Drug Event Report Database. *Biol Pharm Bull* 2019;42:1185–91.
- [44] Uozumi S, Enokida T, Suzuki S, et al. Predictive value of cetuximab-induced skin toxicity in recurrent or metastatic squamous cell carcinoma of the head and neck. *Front Oncol* 2018;8.
- [45] Wang Y, Coiera E, Runciman W, Magrabi F. Using multiclass classification to automate the identification of patient safety incident reports by type and severity. *BMC Med Inform Decis Mak* 2017;17:84.
- [46] Bean DM, Wu H, Iqbal E, et al. Knowledge graph prediction of unknown adverse drug reactions and validation in electronic health records. *Sci Rep* 2017;7:16416.
- [47] Zhang W, Liu X, Chen Y, Wu W, Wang W, Li X. Feature-derived graph regularized matrix factorization for predicting drug side effects. *Neurocomputing* 2018;287:154–62.
- [48] Zhao X, Chen L, Lu J. A similarity-based method for prediction of drug side effects with heterogeneous information. *Math Biosci* 2018;306:136–44.
- [49] Muñoz E, Nováček V, Vandenbusche PY. Facilitating prediction of adverse drug reactions by using knowledge graphs and multi-label learning models. *Brief Bioinform* 2019;20:190–202.
- [50] Jeong E, Park N, Choi Y, Park RW, Yoon D. Machine learning model combining features from algorithms with different analytical methodologies to detect laboratory-event-related adverse drug reaction signals. *PLoS One* 2018;13:e0207749.
- [51] Song D, Chen Y, Min Q, et al. Similarity-based machine learning support vector machine predictor of drug-drug interactions with improved accuracies. *J Clin Pharm Ther* 2019;44:268–75.
- [52] Zhang W, Liu F, Luo L, Zhang J. Predicting drug side effects by multi-label learning and ensemble learning. *BMC Bioinform* 2015;16:365.
- [53] Davazdahemami B, Delen D. A chronological pharmacovigilance network analytics approach for predicting adverse drug events. *J Am Med Inform Assoc* 2018;25:1311–21.
- [54] Raja K, Patrick M, Elder JT, Tsoi LC. Machine learning workflow to enhance predictions of Adverse Drug Reactions (ADRs) through drug-gene interactions: application to drugs for cutaneous diseases. *Sci Rep* 2017;7:3690.
- [55] Hoang T, Liu J, Roughead E, Pratt N, Li J. Supervised signal detection for adverse drug reactions in medication dispensing data. *Comput Methods Programs Biomed* 2018;161:25–38.
- [56] Wang SV, Maro JC, Baro E, et al. Data mining for adverse drug events with a propensity score-matched tree-based scan statistic. *Epidemiology* 2018;29:895–903.
- [57] Maura G, Billionnet C, Coste J, Weill A, Neumann A, Pariente A. Non-bleeding adverse events with the use of direct oral anticoagulants: a sequence symmetry analysis. *Drug Saf* 2018;41:881–97.
- [58] Liu L, Yu Y, Fei Z, et al. An interpretable boosting model to predict side effects of analgesics for osteoarthritis. *BMC Systems Biol* 2018;12:105.
- [59] Ross JS, Bates J, Parzynski CS, et al. Can machine learning complement traditional medical device surveillance? A case study of dual-chamber implantable cardioverter-defibrillators. *Med Devices (Auckl)* 2017;10:165–88.
- [60] Cotterill A, Jaki T. Dose-escalation strategies which use subgroup information. *Pharm Stat* 2018;17:414–36.
- [61] Tan L, Li M, Lin Y. Safety concerns of traditional Chinese medicine injections used in chinese children. *Evid Based Complement Altern Med* 2019;2019.
- [62] Trinh NTH, Sole E, Benkebil M. Benefits of combining change-point analysis with disproportionality analysis in pharmacovigilance signal detection. *Pharmacoepidemiol Drug Saf* 2019;28:370–6.
- [63] Chan CL, Rudrappa S, Ang PS, Li SC, Evans SJW. Detecting signals of disproportionate reporting from singapore's spontaneous adverse event reporting system: an application of the sequential probability ratio test. *Drug Saf* 2017;40:703–13.
- [64] Marbac M, Tubert-Bitter P, Sedki M. Bayesian model selection in logistic regression for the detection of adverse drug reactions. *Biom J* 2016;58:1376–89.
- [65] Xu D, Ham AG, Tivis RD, et al. MSBIS: a multi-step biomedical informatics screening approach for identifying medications that mitigate the risks of metoclopramide-induced Tardive Dyskinesia. *EBioMedicine* 2017;26:132–7.
- [66] Pettit NN, Miceli MH, Rivera CG, et al. Multicentre study of posaconazole delayed-release tablet serum level and association with hepatotoxicity and QTc prolongation. *J Antimicrob Chemother* 2017;72:2355–8.
- [67] Lerch M, Nowicki P, Manlik K, Wirsching G. Statistical signal detection as a routine pharmacovigilance practice: effects of periodicity and resigalling criteria on quality and workload. *Drug Saf* 2015;38:1219–31.
- [68] Nishihara M, Morikawa N, Yokoyama S, Nishikura K, Yasuhara M, Matsuo H. Risk factors increasing blood pressure in Japanese colorectal cancer patients treated with bevacizumab. *Pharmazie* 2018;73:671–5.
- [69] Otake A, Tsuji D, Taku K, et al. Chemotherapy-induced neutropenia as a prognostic factor in patients with metastatic pancreatic cancer treated with gemcitabine. *Eur J Clin Pharmacol* 2017;73:1033–9.
- [70] Kucharz J, Dumnicka P, Kusnier-Cabala B, Demkow T, Wiechno P. The correlation between the incidence of adverse events and progression-free survival in patients treated with cabozantinib for metastatic renal cell carcinoma (mRCC). *Med Oncol* 2019;36:19.
- [71] Dona I, Barrionuevo E, Salas M, et al. Natural evolution in patients with nonsteroidal anti-inflammatory drug-induced urticaria/angioedema. *Allergy* 2017;72:1346–55.
- [72] Gadelha GO, Paixao H, Prado PRD, Viana R, Amaral TLM. Risk factors for death in patients with non-infectious adverse events. *Rev Lat Am Enfermagem* 2018;26:e3001.
- [73] Andrade PHS, Lobo IMF, da Silva WB. Risk factors for adverse drug reactions in pediatric inpatients: a cohort study. *PLoS One* 2017;12:e0182327.

- [74] Westberg SM, Yarbrough A, Weinhandl ED, et al. Drug therapy problem severity following hospitalization and association with 30-day clinical outcomes. *Ann Pharmacother* 2018;52:1195–203.
- [75] Sobhonslidsuk A, Poovorawan K, Soonthornworasiri N, Pan-Ngum W, Phaosawasdi K. The incidence, presentation, outcomes, risk of mortality and economic data of drug-induced liver injury from a national database in Thailand: a population-base study. *BMC Gastroenterol* 2016;16:135.
- [76] Cordiner M, Shajahan P, McAvoy S, Bashir M, Taylor M. Effectiveness of long-acting antipsychotics in clinical practice: 2. Effects of antipsychotic polypharmacy on risperidone long-acting injection and zuclopenthixol decanoate. *Ther Adv Psychopharmacol* 2015;6:66–76.
- [77] Merid MW, Gezie LD, Kassa GM, Muluneh AG, Akalu TY, Yenit MK. Incidence and predictors of major adverse drug events among drug-resistant tuberculosis patients on second-line anti-tuberculosis treatment in Amhara regional state public hospitals; Ethiopia: a retrospective cohort study. *BMC Infect Dis* 2019;19:286.
- [78] Oshikoya KA, Wharton GT, Avant D, et al. Serious adverse events associated with off-label use of azithromycin or fentanyl in children in intensive care units: a retrospective chart review. *Paediatr Drugs* 2019;21:47–58.
- [79] Okamoto I, Sato H, Kondo T, et al. Efficacy and safety of nivolumab in 100 patients with recurrent or metastatic head and neck cancer—a retrospective multicentre study. *Acta Otolaryngol* 2019;139:918–25.
- [80] Dedefo MG, Mitike AH, Angamo MT. Incidence and determinants of medication errors and adverse drug events among hospitalized children in West Ethiopia. *BMC Pediatr* 2016;16:81.
- [81] Blumenthal KG, Wickner PG, Hurwitz S, et al. Tackling inpatient penicillin allergies: assessing tools for antimicrobial stewardship. *J Allergy Clin Immunol* 2017;140: 154-161.e156.
- [82] Sellick J, Mergenhagen K, Morris L, et al. Fluoroquinolone-related neuropsychiatric events in hospitalized veterans. *Psychosomatics* 2018;59:259–66.
- [83] Degu A, Njogu P, Weru I, Karimi P. Assessment of drug therapy problems among patients with cervical cancer at Kenyatta National Hospital, Kenya. *Gynecol Oncol Res Pract* 2017;4.
- [84] Mill C, Primeau MN, Medoff E, et al. Assessing the diagnostic properties of a graded oral provocation challenge for the diagnosis of immediate and nonimmediate reactions to amoxicillin in children. *JAMA Pediatr* 2016;170:e160033.
- [85] Ilich AI, Danilak M, Kim CA, et al. Effects of gender on capecitabine toxicity in colorectal cancer. *J Oncol Pharm Pract* 2016;22:454–60.
- [86] Khong B, Lawson BO, Ma J, et al. Rigor prophylaxis in stage IV melanoma and renal cell carcinoma patients treated with high dose IL-2. *BMC Cancer* 2018;18:1007.
- [87] Daley MF, Clarke CL, Glanz JM, et al. The safety of live attenuated influenza vaccine in children and adolescents 2 through 17 years of age: A Vaccine Safety Datalink study. *Pharmacoepidemiol Drug Saf* 2018;27:59–68.
- [88] Vock DM, Wolfson J, Bandyopadhyay S, et al. Adapting machine learning techniques to censored time-to-event health record data: a general-purpose approach using inverse probability of censoring weighting. *J Biomed Inform* 2016;61:119–31.
- [89] Choi YH, Han CY, Kim KS, Kim SG. Future directions of pharmacovigilance studies using electronic medical recording and human genetic databases. *Toxicol Res* 2019;35:319–30.
- [90] Yang CC, Yang H. Mining heterogeneous networks with topological features constructed from patient-contributed content for pharmacovigilance. *Artif Intell Med* 2018;90:42–52.
- [91] Arnaud M, Bégau B, Thurin N, Moore N, Pariente A, Salvo F. Methods for safety signal detection in healthcare databases: a literature review. *Expert Opin on Drug Safety* 2017;16:721–32.
- [92] Ventola CL. Big data and pharmacovigilance: data mining for adverse drug events and interactions. *P & T: a peer-reviewed journal for formulary management* 2018;43:340–51.
- [93] Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. 2016: arXiv:1603.02754
- [94] Bajpai P, Kumar M. Genetic algorithm—an approach to solve global optimization problems. *Indian Journal of Computer Science and Engineering* 2010;1:199–206.