



# A New Search Method Using Association Rule Mining for Drug-Drug Interaction Based on Spontaneous Report System

Yoshihiro Noguchi\*, Anri Ueno, Manami Otsubo, Hayato Katsuno, Ikuto Sugita, Yuta Kanematsu, Aki Yoshida, Hiroki Esaki, Tomoya Tachi and Hitomi Teramachi\*

## OPEN ACCESS

### Edited by:

Vassilis Koutkias,  
Institute of Applied Biosciences  
(INAB), Centre for Research and  
Technology Hellas, Greece

### Reviewed by:

Domenico Criscuolo,  
Genovax S.r.l., Italy  
Maria Margarita Salazar-Bookaman,  
Central University of Venezuela,  
Venezuela

### \*Correspondence:

Yoshihiro Noguchi  
noguchiy@gifu-pu.ac.jp  
Hitomi Teramachi  
teramachih@gifu-pu.ac.jp

### Specialty section:

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

Received: 26 May 2017

Accepted: 21 February 2018

Published: 09 March 2018

### Citation:

Noguchi Y, Ueno A, Otsubo M,  
Katsuno H, Sugita I, Kanematsu Y,  
Yoshida A, Esaki H, Tachi T and  
Teramachi H (2018) A New Search  
Method Using Association Rule  
Mining for Drug-Drug Interaction  
Based on Spontaneous Report  
System. *Front. Pharmacol.* 9:197.  
doi: 10.3389/fphar.2018.00197

Laboratory of Clinical Pharmacy, Gifu Pharmaceutical University, Gifu, Japan

**Background:** Adverse events (AEs) can be caused not only by one drug but also by the interaction between two or more drugs. Therefore, clarifying whether an AE is due to a specific suspect drug or drug-drug interaction (DDI) is useful information for proper use of drugs. Whereas previous reports on the search for drug-induced AEs with signal detection using spontaneous reporting systems (SRSs) are numerous, reports on drug interactions are limited. This is because in methods that use “a safety signal indicator” (signal), which is frequently used in pharmacovigilance, a huge number of combinations must be prepared when signal detection is performed, and each risk index must be calculated, which makes interaction search appear unrealistic.

**Objective:** In this paper, we propose association rule mining (AR) using large dataset analysis as an alternative to the conventional methods (additive interaction model (AI) and multiplicative interaction model (MI)).

**Methods:** The data source used was the Japanese Adverse Drug Event Report database. The combination of drugs for which the risk index is detected by the “combination risk ratio (CR)” as the target was assumed to be true data, and the accuracy of signal detection using the AR methods was evaluated in terms of sensitivity, specificity, Youden’s index, *F*-score.

**Results:** Our experimental results targeting Stevens-Johnson syndrome indicate that AR has a sensitivity of 99.05%, specificity of 92.60%, Youden’s index of 0.917, *F*-score of 0.876, AI has a sensitivity of 95.62%, specificity of 96.92%, Youden’s index of 0.925, and *F*-score of 0.924, and MI has a sensitivity of 65.46%, specificity of 98.78%, Youden’s index of 0.642, and *F*-score of 0.771. This result was about the same level as or higher than the conventional method.

**Conclusions:** If you use similar calculation methods to create combinations from the database, not only for SJS, but for all AEs, the number of combinations would be so

enormous that it would be difficult to perform the calculations. However, in the AR method, the “Apriori algorithm” is used to reduce the number of calculations. Thus, the proposed method has the same detection power as the conventional methods, with the significant advantage that its calculation process is simple.

**Keywords:** drug-drug interaction, spontaneous reporting systems, signal detection, association rule mining, apriori algorithm

## INTRODUCTION

Pharmacovigilance is defined as “the science and activities relating to the detection, assessment, understanding and prevention of drug-related problems” by the World Health Organization (WHO) (World Health Organization, 2002). For adverse drug events (ADEs) that cannot be found in clinical trials, numerous risk assessments of drugs using spontaneous reporting systems (SRSs) based on spontaneous reports of drug-adverse event (AE) pairs including post-marketing survey data have been reported (Poluzzi et al., 2013; Zorych et al., 2013; Fujimoto et al., 2014; Ali et al., 2015; Noguchi et al., 2015; Mizuno et al., 2016; Gahr et al., 2017).

For risk assessment, the data mining algorithms of the quantitative signal detection from such a large database, include the proportional reporting ratio (PRR) (Evans et al., 2001) is used in the Medicines and Healthcare Products Regulatory Agency (MHRA), the reporting odds ratio (ROR) (van Puijenbroek et al., 2002) which is used by the Netherlands Pharmacovigilance Centre Lareb, the information component (IC) as Bayesian Confidence Propagation Neural Network (BCPNN) (Bate et al., 1998) is used by the Uppsala Monitoring Centre, Sweden, and the Gamma-Poisson Shrinker (GPS) (Szarfman et al., 2002) as empirical Bayes geometric mean (EBGM). These methods are useful for early detection of unknown ADEs. However, in pre-marketing randomized clinical trials, patients with multiple drug use are usually excluded because focus is on establishing the safety and efficacy of single drugs and not on the investigation of drug-drug interactions (DDIs).

However, the reality is that in post-marketing pharmaceutical products, unlike pre-marketing trials, not only ADEs caused by a single drug but also ADEs caused by DDI by two or more drugs exist. The proportion of ADEs due to DDI is estimated to be up to 30% of unexpected ADEs (Pirmohamed and Orme, 1998). Therefore, clarifying whether the ADEs expressed are due to a specific suspected drug or DDI would be useful when a patient uses the drugs.

To take into consideration the interaction between Drug A and Drug B, as in the case of a single drug, a “Drug A and Drug B-AE 2 × 2 contingency table” should first be prepared and the respective signal values calculated. Furthermore, to detect the signal at the time of concomitant use as a DDI, it must be stronger than the signal of each single drug.

Methods such as the additive interaction model (AI) (= risk difference) and the multiplicative interaction model (MI) (= risk ratio) are used to evaluate interactions with ADEs expression when two drugs are used in combination (Thakrar

et al., 2007). However, in searching for drug interactions, Susuta et al. proposed “combination risk ratio (CR),” in which the ratio of signal detection using a plurality of drugs in combination to the signal detection calculated by focusing on one drug is utilized (Susuta and Takahashi, 2014). If the PRR value of ADE when two concomitant drugs are used is more than twice the PRR value of each drug, then it is regarded as a DDI. This method proposed by Susuta et al. is used searching for drug interactions in Japan (Noguchi et al., 2015; Mizuno et al., 2016).

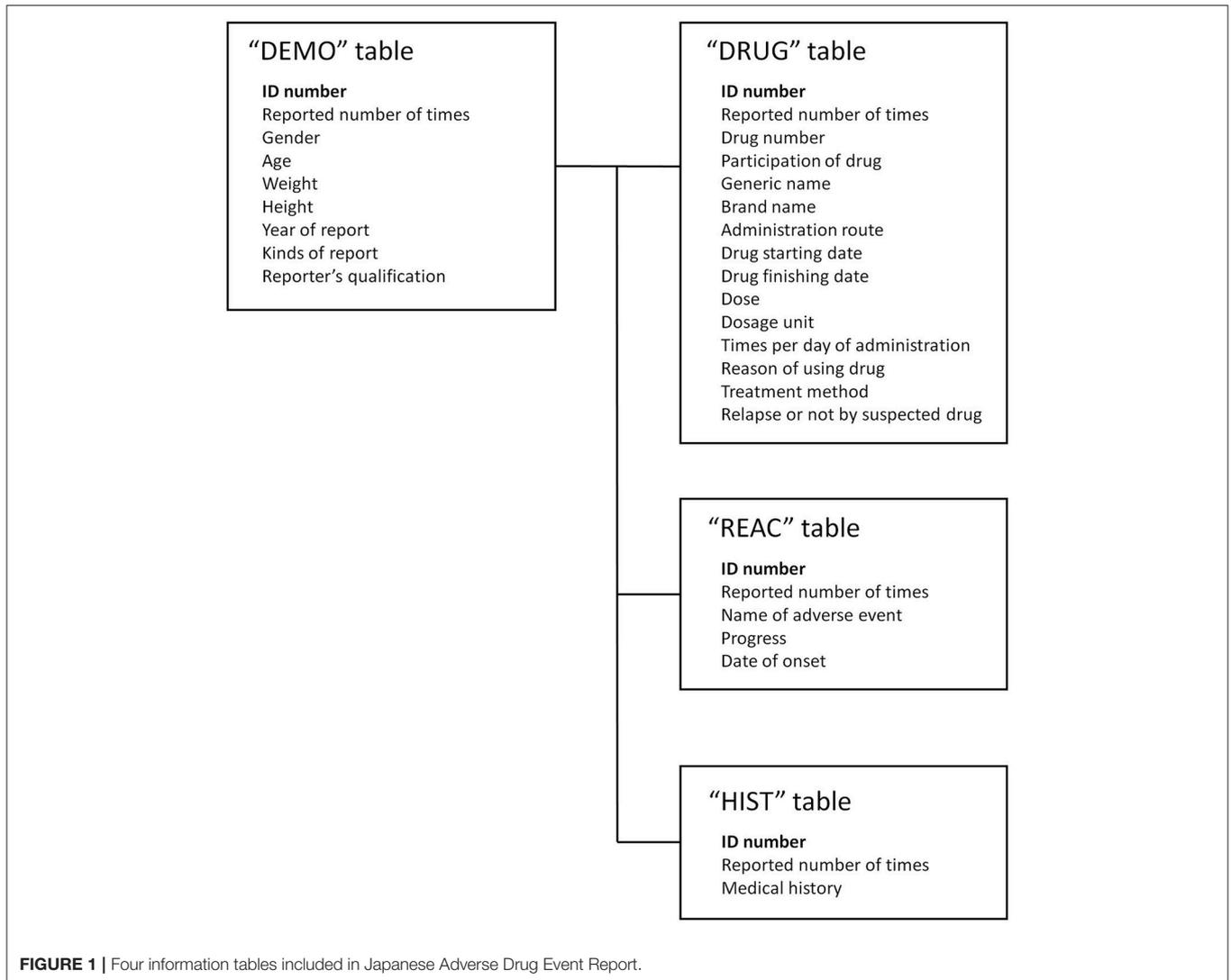
However, in these methods as AI, MI, and CR, it is necessary to compose a huge combination of medicines from SRS and compare the signal of these drugs with the signal of a single drug in order to detect drug interaction. Consequently, it is difficult to efficiently conduct a cross-sectional search aimed at early detection of DDI.

Association rule mining (AR) is used to find “interesting patterns hidden in a database.” Algorithms that simplify operations such as the “Apriori algorithm” (Agrawal and Srikant, 1994) are used for AR and, in recent years, utilization of ADEs as signal detection has also been reported (Shirakuni et al., 2009; Harpaz et al., 2010; Wang et al., 2012; Fujiwara et al., 2016). Therefore, to construct a simple ADE search method for DDI, we modified existing AR such that risk variation due to a combination of drugs can be evaluated and investigated combinations of drugs that may cause Stevens-Johnson syndrome (SJS) as DDI.

## MATERIALS AND METHODS

### Data Sources

We used the SRS dataset from the first quarter of 2004 to the fourth quarter of 2015 from the Japanese Adverse Drug Event Report (JADER) database. The JADER database was downloaded from the website of the Pharmaceuticals and Medical Devices Agency in Japan. The structure of JADER complies with the international safety reporting guidelines (ICH E2B). The database consists of four tables: “DEMO,” “DRUG,” “REAC,” and “HIST.” The “DEMO” table contains information on gender, age, weight, etc., The “DRUG” table contains information on suspect drug, concomitant usage, etc. The “REAC” table contains information on AE and outcome. The “HIST” table contains information on primary disease, secondary disease, etc. In this study, we combined these four tables (Figure 1) and analyzed 374,327 cases, with cases having gender and age deficiency and cases that have ambiguous reports of age such as “youth” and “elderly” being excluded.



### Definitions of Adverse Drug Events

The drugs targeted for the survey are all registered and classified as “suspect drugs” in JADER. The AE investigated was SJS. To extract the AEs from JADER, the preferred term (PT) in the Medical Dictionary for Regulatory Activities (Japanese version (MedDRA /J) version 18.1) was used. Further, the definition of SJS was the same as that used by Susuta et al. in a previous study (Susuta and Takahashi, 2014).

### Association Rule Mining Model

In this paper, we propose AR as an alternative to the conventional ADE search methods for DDI. However, in this study, because the degree of the influence on Drug A-induced AE C by Drug B is used as an indicator of the DDI, the association rule “ $B \rightarrow A \cap C$ ” (where B is Drug B and  $A \cap C$  is the Drug A-induced AE C) was used instead of the typically used association rule “ $A \cap B \rightarrow C$ ” (where A is Drug A, B is Drug B, and C is the AE C). In association rule “ $A \cap B \rightarrow C$ ”, the signal by the combination of Drug A and Drug B is detected, but the

influence due to the addition of Drug B cannot be investigated. In addition, we used “lift” and “conviction” as the detection criteria for searching the association rule. The calculation methods are shown in **Figures 2, 3** and the following formulas.

$$\begin{aligned}
 \text{lift } (B \rightarrow A \cap C) &= \text{confidence } (B \rightarrow A \cap C) / \text{support } (A \cap C) \\
 &= \{n_{AB1} / (n_{AB+} + n_{B+})\} / \{(n_{AB1} + n_{A1}) / n_{++}\} \\
 \text{conviction } (B \rightarrow A \cap C) &= (1 - \text{Support } (A \cap C)) / (1 - \text{Confidence } (B \rightarrow A \cap C)) \\
 &= \{1 - (n_{AB1} + n_{A1}) / n_{++}\} / \{1 - n_{AB1} / (n_{AB+} + n_{B+})\}
 \end{aligned}$$

“Lift” is an index that indicates the relative magnitude of the probability of observing  $A \cap C$  under the condition of B, compared to the overall probability of observing  $A \cap C$ . When lift = 1, the two events, B and  $A \cap C$ , are independent of each other. When the lift value is greater than one, the two events, B and

$A \cap C$ , are not independent, and the higher the value, the greater the relevance of the interaction (Hahsler et al., 2005).

On the other hand, “conviction” is an indicator that evaluates whether the rule makes a wrong prediction, paying particular attention to the exclusion event of the conclusion part of the obtained rule. If the conviction value is large, it is less likely that

the conclusion  $A \cap C$  is not true for the premise B (Brin et al., 1997).

In this study,  $lift > 1$  and  $conviction > 1$  were used as the detection standard for the AR method. However, in order to investigate the power of detection, because the combination of drugs for which the risk index is detected by the CR as the target was assumed to be true data, it was found that  $n_{AB1} < 3$  was not detected in the “combined risk ratio.” That combination was therefore excluded from the signal and  $n_{AB1} \geq 3$  was added to the detection condition.

AR Criteria of DDI for this study:

$$n_{AB1} \geq 3, Lift(B \rightarrow A \cap C) > 1 \text{ and Conviction}(B \rightarrow A \cap C) > 1$$

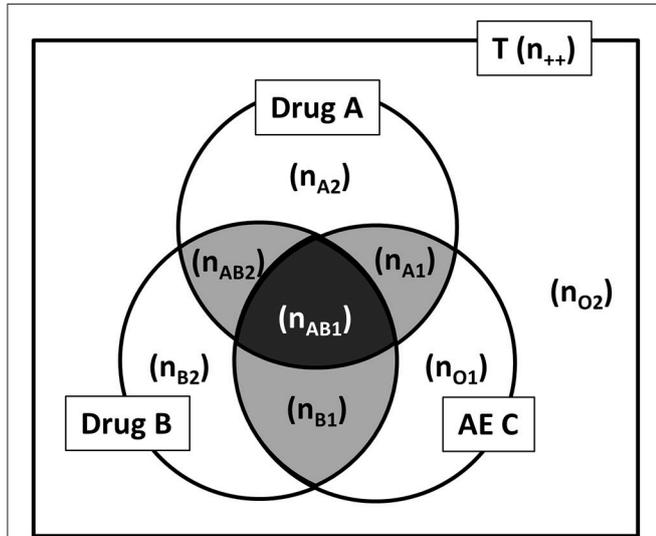
### Combination Risk Ratio Model

Susuta et al. proposed a CR model to generate DDI signals. In their method, AE C as DDI risk of drug A and drug B is estimated by the following procedure (Susuta and Takahashi, 2014):

- 1) PRR, which is a statistical signal detection indicator, is used to evaluate the increase in AE C as DDI risk of Drug A and Drug B.
- 2) PRR about one drug (= Drug A) of two target suspicious drugs is calculated individually.
- 3) The case where the ratio of the PRR of the two drug concomitant usage to the PRR obtained individually exceeds two is judged as a DDI risk.

$$DDI \text{ risk ratio} = PRR_{Drug A \text{ and } B} / PRR_{Drug A}$$

$$DDI \text{ risk criteria : } n_{AB1} \geq 3, PRR_{Drug A \text{ and } B} > 2 \text{ and } \chi^2_{Drug A \text{ and } B} > 4 \text{ and DDI risk ratio} > 2$$



**FIGURE 2** | Drug A, Drug B, and AE C by Venn diagram from the spontaneous report system. AE, adverse events; T, total of report (all reports); n, the number of reports (e.g.,  $n_{++}$ , the number of all reports;  $n_{AB1}$ , the number of Drug A and Drug B induced AE C report).

	Suspected AE C	All other AEs	Total
Drug A and Drug B	$n_{AB1}$	$n_{AB2}$	$n_{AB+}$
Drug A, Not Drug B	$n_{A1}$	$n_{A2}$	$n_{A+}$
Drug B, Not Drug A	$n_{B1}$	$n_{B2}$	$n_{B+}$
All other drugs	$n_{O1}$	$n_{O2}$	$n_{O+}$
Total	$n_{+1}$	$n_{+2}$	$n_{++}$

**FIGURE 3** | Four-by-two contingency table for AR from the Venn diagram in Figure 2. AR, association rule mining; AE, adverse events; n, the number of reports (e.g.,  $n_{++}$ , the number of all reports;  $n_{AB1}$ , the number of Drug A and Drug B induced AE C reports).

AE C	Drug B (B = 1)	Not drug B (B = 0)
Drug A (A = 1)	$P_{11}$ (= $n_{AB1} / n_{AB1} + n_{AB2}$ )	$P_{10}$ (= $n_{A1} / n_{A1} + n_{A2}$ )
Not drug A (A = 0)	$P_{01}$ (= $n_{B1} / n_{B1} + n_{B2}$ )	$P_{00}$ (= $n_{O1} / n_{O1} + n_{O2}$ )

**FIGURE 4 |** Two-by-two contingency table for calculation of AI and MI. AI, additive interaction; MI, multiplicative interaction; AE, adverse events; n, the number of reports (e.g.,  $n_{AB1}$ , the number of Drug A and Drug B induced AE C reports), P, the proportion of reports (e.g.,  $P_{11}$ , the proportion of Drug A and Drug B induced AE C report to Drug A and Drug B induced all AEs).

### Additive Interaction Model

Thakrar et al. proposed an AI model to generate DDI signals (Thakrar et al., 2007). In their proposed method, under the additive assumption, no interaction is established when the excess risk associated with Drug A in the absence of Drug B is the same as the excess risk associated with Drug A in the presence of Drug B:

$$\text{risk}_{(\text{Drug A, not B})} - \text{risk}_{(\text{not Drug A, not B})} = \text{risk}_{(\text{Drug A, B})} - \text{risk}_{(\text{not Drug A, B})}$$

This equality implies:

$$\text{Risk Difference (RD)}_{11} = (\text{RD}_{10} + \text{RD}_{01})$$

i.e. the excess risk associated with the combination is the same as the sum of the excess risks associated with each exposure in the absence of the other, where  $\text{RD}_{11} = \text{risk}_{(\text{DrugA,B})} - \text{risk}_{(\text{notDrugA,notB})}$  is the excess risk associated with the two drugs in combination and similarly for  $\text{RD}_{10}$  and  $\text{RD}_{01}$ .

The AI is calculated from the two-by-two contingency table shown in **Figure 4** and the following formulas.

$$\text{AI} = \text{RD}_{11} - (\text{RD}_{10} + \text{RD}_{01}) = (P_{11} - P_{00}) - \{(P_{10} - P_{00}) + (P_{01} - P_{00})\} = P_{11} - P_{10} - P_{01} + P_{00}$$

P: the proportion of reports

(e.g.  $P_{11}$ : the proportion of Drug A and Drug B induced reports to Drug A and Drug B induced all AEs)

When  $\text{RD}_{11} > \text{RD}_{10} + \text{RD}_{01}$  (i.e.,  $\text{RD}_{11} - (\text{RD}_{10} + \text{RD}_{01}) > 0$ ) there is a potential interaction with an increased risk for the combination compared with that expected based on the individual drugs. On the other hand, in the absence of an interaction under the additive assumption, the excess risk associated with the combination is the same as the sum of the excess risk associated with each drug separately (Thakrar et al., 2007). However, in this study, because it was found that  $n_{AB1} < 3$  is not detected in the “combined risk ratio,” that combination was excluded from the signal, and  $n_{AB1} \geq 3$  was added to the detection condition.

### Multiplicative Interaction Model

Thakrar et al. proposed an MI model to generate DDI signals (Thakrar et al., 2007). In their proposed model, when there is no interaction on the multiplicative scale, the relative risk associated with Drug A is the same in both the absence and presence of exposure to Drug B.

$$\frac{\text{risk}_{(\text{Drug A, not B})}}{\text{risk}_{(\text{not Drug A, not B})}} = \frac{\text{risk}_{(\text{Drug A, B})}}{\text{risk}_{(\text{not Drug A, B})}}$$

This equality implies:

$$\text{Risk Ratio (RR)}_{11} = (\text{RR}_{10} \times \text{RR}_{01})$$

where  $\text{RR}_{11} = \text{risk}_{(\text{DrugA,B})} / \text{risk}_{(\text{notDrugA,notB})}$  is the relative risk associated with the two drugs in combination and similarly for  $\text{RR}_{10}$  and  $\text{RR}_{01}$ .

The MI is calculated from the two-by-two contingency table shown in **Figure 4** and the following formulas.

$$\text{MI} = \text{RR}_{11} / (\text{RR}_{10} \times \text{RR}_{01}) = (P_{11} / P_{00}) / \{(P_{10} / P_{00}) \times (P_{01} / P_{00})\} = (P_{11} \times P_{00}) / (P_{10} \times P_{01})$$

If  $\text{RR}_{11} / (\text{RR}_{10} \times \text{RR}_{01})$  is statistically different from one, it is considered evidence of an interaction. In particular, whenever this measure is greater than one, there is a positive interaction that is of interest from a safety perspective. In such a situation, the relative risk associated with two drugs administered in combination is greater than the product of the relative risks associated with each drug separately (Thakrar et al., 2007). However, in this current study, because it was found that  $n < 3$  is not detected in the “combined risk ratio,” that combination was excluded from the signal, and  $n \geq 3$  was added to the detection condition.

### Application of the Models

Since true data of unknown DDIs cannot be prepared, in this study, we treated the drug-drug combination detected CR model as true data. This model is used for searching DDI signals in Japan (Noguchi et al., 2015; Mizuno et al., 2016), but, if a calculation model that simply creates combinations from the database was used, the number of combinations that need to be calculated would be enormous. Therefore, there is a need for a simple method to obtain a search result similar to this method.

The accuracy of signal detection using the AR, AI, and MI methods was examined using sensitivity, specificity, Youden’s index, positive predictive value (PPV), negative predictive value (NPV), F-score and area under the receiver operating characteristic (ROC) curve (AUC).

### Analysis Software

Data management and analyses were performed using Visual Mining Studio software (version 8.1; Mathematical Systems, Inc. Tokyo, Japan). Drawing of the ROC curve and AUC calculation were performed using JMP 11.2.0 (SAS Institute Inc.,).

### RESULTS

Among all the cases analyzed (374,327 cases), there were 7,848 drug–AE (SJS) combinations. The number of signals detected using the “combined risk ratio” was 1,672 pairs.

**Table 1** shows the signal detection power of AR, AI, and MI. For AR, the sensitivity was 99.05%, specificity was 92.60%, Youden’s

index was 0.917, positive predictive value (PPV) was 78.57%, negative predictive value (NPV) was 99.72% and *F*-score was 0.876. For AI, the sensitivity was 95.62%, specificity was 96.92%, Youden's index was 0.925, PPV was 89.47%, NPV was 98.78% and *F*-score was 0.924. For MI, the sensitivity was 65.46%, specificity was 98.78%, Youden's index was 0.642, PPV was 93.64%, NPV was 91.26% and *F*-score was 0.771. Among the three models, AR had the highest sensitivity and the highest NPV, whereas MI had the highest Specificity and the highest PPV. On the other hand, for Youden's index and *F*-score, AI had slightly higher values than those obtained by the AR model proposed in this study.

**Figure 5** shows the ROC curve. For AR, the AUC was 0.959; for AI, the AUC was 0.956; and for MI, the AUC was 0.959. For all three models, the AUC was greater than 0.95, which shows that AR has detection power that is on par with conventional methods such as AI and MI.

**Figure 6** shows the correlation between the lift value and DDI risk ratio signal intensity. These show a positive correlation, the decision coefficient ( $R^2$ ) = 0.170.

Since there are many combinations of SJS-expressing drugs with 7848 pairs, in this paper, we will explain the case where Drug B is a combination that is loxoprofen. There were 195 pairs where Drug B was loxoprofen. Among them, 67 pairs of signals were detected. **Table 2** shows some of them.

## DISCUSSION

Whereas previous reports on the search for drug-induced AEs using SRS are numerous, reports on drug interactions are limited.

This is because in methods that use a safety signal indicator, which is frequently used in pharmacovigilance, a huge number of combinations must be prepared when signal detection is performed, and each risk index must be calculated, which makes interaction search appear unrealistic.

A potential solution to this issue would be to utilize AR. However, although AR is often used to efficiently analyze large datasets, there are only a few examples of it being used in the medical field, especially in SRS analysis.

Wang et al. recently conducted a simulation study that is considered the gold standard of AR signal detection criteria for searching for ADEs due to single drugs (Wang et al., 2012). However, there are few examples of utilization of AR for early detection of drug interaction.

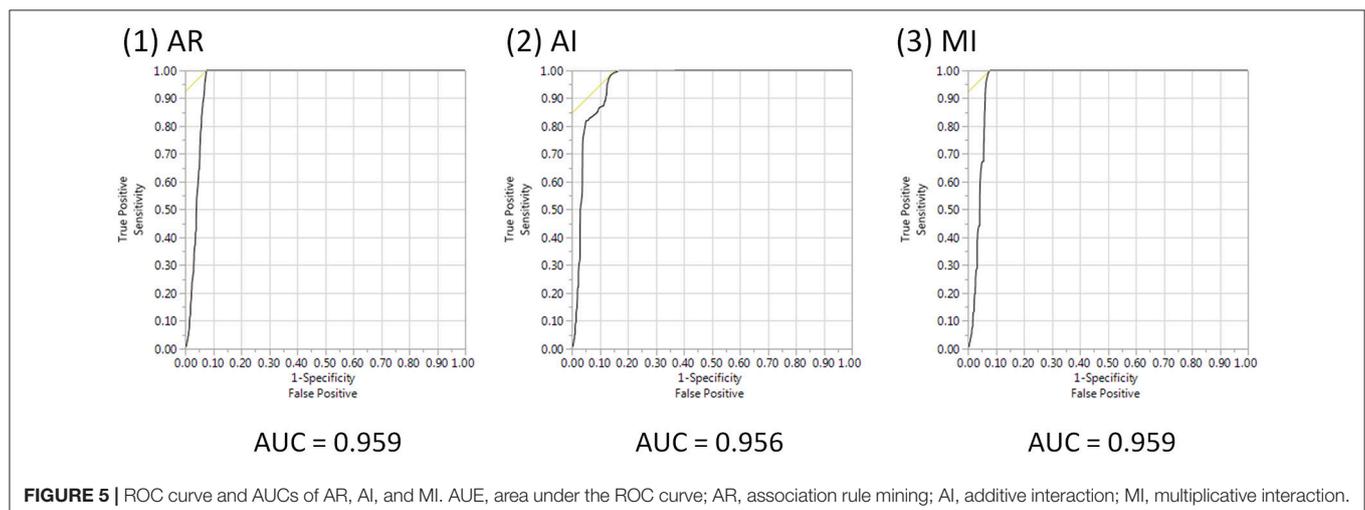
The method proposed by Susuta et al. can be considered as a prior research method in the search for DDI induced ADEs (Susuta and Takahashi, 2014). In this study, we treated the drug-drug combination detected by Susuta et al.'s method (= CR) as true data. Then, we evaluated the detection power of our proposed method by calculating the sensitivity, specificity, PPV, NPV, and *F*-score.

In general, when searching for AE C with drugs A and B, the AR model uses " $A \cap B \rightarrow C$ ". However, this model only shows the association between the combination used and the AE C, it does not show the influence of DDI. To show the influence of DDI, the addition of Drug B requires that the indicator be larger than the original Drug A -induced AE C (" $B \rightarrow A \cap C$ "). Therefore, in this study, the degree of influence on Drug A-induced AE C by Drug B is used as an indicator of the DDI. In this way, the lift value calculated in this study is a variation value of the probability of occurrence of AE C caused by Drug A by using Drug B in combination.

**TABLE 1** | Ability of AR, AI, and MI to detect signals.

Models	Sensitivity (%)	Specificity (%)	Youden's index	PPV (%)	NPV (%)	F-score
AR	99.05	92.60	0.917	78.57	99.72	0.876
AI	95.62	96.92	0.925	89.47	98.78	0.924
MI	65.46	98.78	0.642	93.64	91.26	0.771

AR, Association rule mining; AI, additive interaction model (= risk difference), MI, multiplicative interaction model (= risk ratio), PPV, positive predictive value; NPV, negative predictive value.



For AR Criteria, the lift value is common. However, an indicator that evaluates whether the rule has made a wrong prediction needs to be included among the criteria. Therefore, the conviction value was also added as a criterion.

The sensitivity for AR was 99.05%. That is, using AR, it was revealed that almost all of the combination of drug interactions expressing SJS indicated as signal detection in the combination risk index was detected. In addition, the sensitivity of AR was higher than both the 95.62% of AI and the 65.46% of MI, which are conventional methods. For Youden's index, which considers sensitivity to sensitivity, AR was 0.917, which is comparable to the 0.925 of AI, and exceeds the 0.642 of MI. Also, for *F*-score, AR was 0.876, which is near to the 0.924 of AI, and exceeds the 0.771 of MI. From these results, it is clear that the detection accuracy of AR is as high as that of AI and higher than that of MI.

The NPV of AR was 99.72%, that of AI was 98.78%, and that of MI was 91.26%. None of the detection methods had a negative misdetection signal. Of the three methods, AR was the least

erroneous. On the other hand, with PPV, AR: 78.57% was lower than AI: 89.47%, and MI: 93.64%.

This result is because data using positive combination risk index, which is one of the signal detection criteria of existing drug interaction, were used as true data. AR can detect a combination of DDI induced SJS that cannot be detected with a combination risk index; hence, there is a possibility that PPV was underestimated.

However, as true data of "true risk" containing "unknown AEs" does not exist in the world, these results for AR, AI, and MI cannot be verified. It merely showed combinations of drugs with possible risks due to combination of sensitivity index and high sensitivity for the CR proposed by Susuta et al.

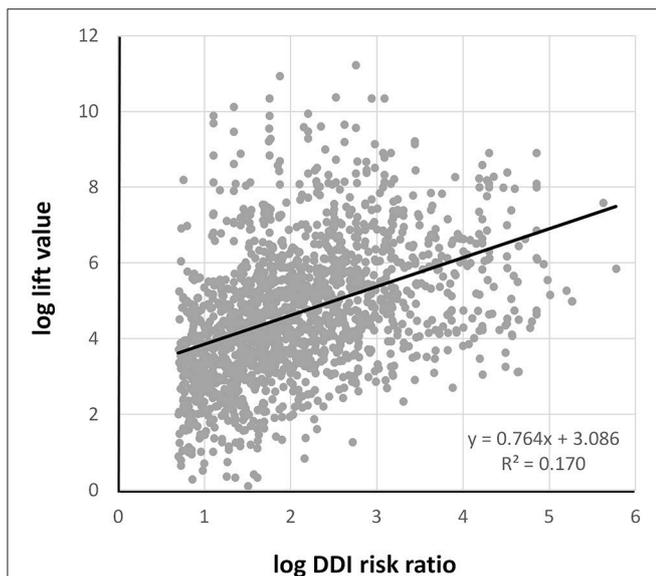
Since AR model proposed by us in this paper is represented by this formula: " $B \rightarrow A \cap C$ ", if " $\text{Drug B} \rightarrow \text{AE C}$ " is a known effect, " $\text{Drug B} \rightarrow \text{Drug A} \cap \text{AE C}$ " might be considered to be true for a random Drug A. However, when we performed investigations in accordance with this suggestion, we encountered a case in which the result was not necessarily the same as this suggestion.

For example, When Drug B is loxoprofen, AE C is SJS, the lift value of " $\text{loxoprofen} \rightarrow \text{SJS}$ " is 5.78. As shown in Table 2, among some of the combinations where drug B is loxoprofen, DDI risk is detected when Drug A is carbamazepine or clarithromycin. On the other hand, when Drug A is lamotrigine, the lift value is 0.66. Also, when Drug A is phenytoin, the lift value is 0.89. Even if " $\text{Drug B} \rightarrow \text{AE C}$ " is well-known, " $\text{Drug B} \rightarrow \text{random Drug A} \cap \text{AE C}$ " is not always true.

The magnitude of the lift value and that of the DDI risk ratio signal intensity are positively correlated. However, the hypothesis obtained from the signal should be verified separately for ADEs with a large lift value.

The CR is used for searching DDI signals in Japan (Noguchi et al., 2015; Mizuno et al., 2016), and there is a need for a simple method to obtain a search result similar to this method. If a similar calculation method that simply creates combinations from the database was used instead of AR, the number of combinations would be enormous and it would be difficult to perform the calculations within a realistic time. However, in the AR method, the "Apriori algorithm" can be used to reduce the number of calculations. The Apriori algorithm is based on the principle that "support of a certain item set is always less than or equal to support of its partial item set" (Agrawal and Srikant, 1994). Therefore, it is not necessary to calculate indices for all combinations, as in the conventional methods including the CR.

In this study, to compare the signal detection powers, all combinations were calculated in the AR analysis without using the apriori algorithm. Therefore, it is unknown how much time can be



**FIGURE 6 |** The correlation between log lift value and log DDI risk ratio. DDI, drug-drug interaction.

**TABLE 2 |** Signal of some combination of drug B: loxoprofen as examples of DDI.

Drug A	$n_{AB}$	Combination risk ratio			Association rule mining	
		PRR Drug A and B	$\chi^2$ Drug A and B	DDI risk ratio	lift	conviction
Carbamazepine	20	18.70	320.29	2.75	4.77	1.00 (1.00343)
Clarithromycin	34	16.52	483.04	2.81	14.28	1.00 (1.00069)
lamotrigine	3	23.76	45.17	1.79	0.66	1.00 (0.99966)
phenytoin	1	9.90	1.59	1.35	0.89	1.00 (0.99997)

DDI, drug-drug interaction;  $n_{AB}$ , the number of Drug A and Drug B induced AE C reports; PRR, proportional reporting ratio.

shortened by the apriori algorithm, compared with the conventional method. However, the pairs that can be a combination of drug interactions are predicted to be enormous. In our proposed method, since the calculation can be simplified by apriori algorithm in practice, it is expected that the signal search time will definitely be shortened in actual search.

So, it is possible to efficiently perform a cross-sectional search aiming at early detection of drug interactions. Furthermore, our proposed method can be applied not only to JADER but also to the SRS of other countries, such as the US FDA Adverse Event Reporting System (FAERS).

In this study, the true data for verification was created by CR, not the “real” true data. Therefore, we cannot deny the possibility that false positives are detected. We could not verify all combinations. Therefore, our proposed method requires verification for ADEs other than SJS. However, although only a part of the large data is used, this result obtained in this study would suggest that calculation is easy with the same detection power as the conventional method.

Because SRS is the result of spontaneous reporting and is influenced by reporting bias including underreporting, the value of the signal easily varies depending on the time of investigation; therefore, the analysis result is not a true risk but a risk hypothesis. Consequently, in order to clarify the true risk, further fundamental pharmacological research and clinical research are needed based

on the obtained risk hypothesis. However, in order to discover an unknown ADE at an early stage, the issue of how to establish this hypothesis early can be considered an important factor. In pharmacovigilance, building a correct hypothesis is also important, but we believe that simple methods such as our proposed method with the same detection power as the conventional methods are required.

## AVAILABILITY OF DATA AND MATERIAL

The Japanese Adverse Drug Event Report database (JADER). [<http://www.info.pmda.go.jp/fukusayouddb/CsvDownload.jsp>] (in Japanese only).

## AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: YN and HT. Performed the experiments: YN, AU, MO, and HK. Analyzed the data: YN, IS, YK, AY, and HE. Contributed reagents, materials, analysis tools: YN, HT, and TT. Wrote the paper: YN and HT.

## FUNDING

This study was carried out with the aid of JSPS scientific research fund 16K19175.

## REFERENCES

- Agrawal, R., and Srikant, R. (1994). “Fast algorithms for mining association rules,” in *Proceedings of the 20th International Conference on Very Large Databases* (San Francisco, CA), 487–499.
- Ali, T. B., Schleret, T. R., Reilly, B. M., Chen, W. Y., and Abagyan, R. (2015). Adverse effects of cholinesterase inhibitors in dementia, according to the pharmacovigilance databases of the United-States and Canada. *PLoS ONE* 10:e0144337. doi: 10.1371/journal.pone.0144337
- Bate, A., Lindquist, M., Edwards, I. R., Olsson, S., Orre, R., Lansner, A., et al. R.M. (1998). A Bayesian neural network method for adverse drug reaction signal generation. *Eur. J. Clin. Pharmacol.* 54, 315–321. doi: 10.1007/s002280050466
- Brin, S., Motwani, R., Ullman, J. D., and Tsur, S. (1997). Dynamic itemset counting and implication rules for market basket data. *ACM SIGMOD Rec.* 26, 255–264. doi: 10.1145/253262.253325
- 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
- Fujimoto, M., Hosomi, K., and Takada, M. (2014). Statin-associated lower urinary tract symptoms: data mining of the public version of the FDA adverse event reporting system, FAERS. *Int. J. Clin. Pharmacol. Ther.* 52, 259–266. doi: 10.5414/CP202033
- Fujiwara, M., Kawasaki, Y., and Yamada, H. A. (2016). Pharmacovigilance approach for post-marketing in Japan using the Japanese Adverse Drug Event Report (JADER) database and association analysis. *PLoS ONE* 11:e0154425. doi: 10.1371/journal.pone.0154425
- Gahr, M., Connemann, B. J., Schönfeldt-Lecuona, C., and Zeiss, R. (2017). Sensitivity of quantitative signal detection in regards to pharmacological neuroenhancement. *Int. J. Mol. Sci.* 18:E101. doi: 10.3390/ijms18010101
- Hahsler, M., Grün, B., and Hornik, K. (2005). arules - A computational environment for mining association rules and frequent item sets. *J. Stat. Softw.* 14, 1–25. doi: 10.18637/jss.v014.i15
- Harpaz, R., Chase, H. S., and Friedman, C. (2010). Mining multi-item drug adverse effect associations in spontaneous reporting systems. *BMC Bioinformatics* 11:S7. doi: 10.1186/1471-2105-11-S9-S7
- Mizuno, T., Umemura, T., Sakai, T., Fukatsu, M., Yamada, T., Kajiguchi, T., et al. (2016). Signal detection on the concomitant use of deferasirox with other drugs and acute renal failure using data mining of the Japanese adverse drug event report database and evaluation by a case-control study. *Jpn. J. Pharm. Health Care Sci.* 42, 717–726. doi: 10.5649/jphcs.42.717
- Noguchi, Y., Esaki, H., Asano, S., Yokoi, T., Usui, K., Kato, M., et al. (2015). Analysis of effects of the diuretics on levels of blood potassium and blood sodium with angiotensin receptor blockers and thiazide diuretics combination therapy: data mining of the Japanese adverse drug event report database, JADER. *Jpn. J. Pharm. Health Care Sci.* 41, 488–496. doi: 10.5649/jphcs.41.488
- Pirmohamed, M., and Orme, M., (1998). “Drug interactions of clinical importance,” in *Davies’s Textbook of Adverse Drug Reactions. 5th Edn.*, eds D. Davies, R. Ferner and H. de Glanville (London: Chapman & Hall Medical), 888–912.
- Poluzzi, E., Raschi, E., Koci, A., Moretti, U., Spina, E., Behr, E. R., et al. (2013). Antipsychotics and torsadogenic risk: signals emerging from the US FDA adverse event reporting system database. *Drug Saf.* 36, 467–479. doi: 10.1007/s40264-013-0032-z
- Shirakuni, Y., Okamoto, K., Kawashita, N., Yasunaga, T., and Takagi, T. (2009). Signal detection of drug complications applying association rule learning for Stevens-Johnson Syndrome. *J. Com. Aid. Chem.* 10, 118–127. doi: 10.2751/jcac.10.118
- Susuta, Y., and Takahashi, Y. (2014). Safety risk evaluation methodology in detecting the medicine concomitant use risk which might cause critical drug rash. *Jpn. J. Pharmacoepidemiol.* 19, 39–49. doi: 10.3820/jjpe.19.39
- Szarfman, A., Machado, S. G., and O’Neill, R. T. (2002). Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the US FDA’s spontaneous reports database. *Drug Saf.* 25, 381–392. doi: 10.2165/00002018-200225060-00001
- Thakrar, B. T., Grundschober, S. B., and Doesseger, L. (2007). Detecting signals of drug-drug interactions in a spontaneous reports database. *Br. J. Clin. Pharmacol.* 64, 489–495. doi: 10.1111/j.1365-2125.2007.02900.x
- van Puijnenbroek, 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, C., Guo, X. J., Xu, J. F., Wu, C., Sun, Y. L., Ye, X. F., et al. (2012). Exploration of the association rules mining technique for the signal detection of adverse drug events in spontaneous reporting systems. *PLoS ONE* 7:e40561. doi: 10.1371/journal.pone.0040561
- World Health Organization (2002). “The importance of pharmacovigilance,” in *WHO Collaborating Centre for International Drug Monitoring* (Geneva), 44.
- Zorych, I., Madigan, D., Ryan, P., and Bate, A. (2013). Disproportionality methods for pharmacovigilance in longitudinal observational databases. *Stat. Methods. Med. Res.* 22, 39–56. doi: 10.1177/0962280211403602

**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 Noguchi, Ueno, Otsubo, Katsuno, Sugita, Kanematsu, Yoshida, Esaki, Tachi and Teramachi. 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 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.