



Article

# Subset Analysis for Screening Drug–Drug Interaction Signal Using Pharmacovigilance Database

Yoshihiro Noguchi <sup>1,\*</sup> , Tomoya Tachi <sup>1</sup> and Hitomi Teramachi <sup>1,2,\*</sup>

<sup>1</sup> Laboratory of Clinical Pharmacy, Gifu Pharmaceutical University, 1-25-4, Daigakunishi, Gifu-shi, Gifu 501-1196, Japan; tachi@gifu-pu.ac.jp

<sup>2</sup> Laboratory of Community Healthcare Pharmacy, Gifu Pharmaceutical University, Daigakunishi, Gifu-shi, Gifu 501-1196, Japan

\* Correspondence: noguchiy@gifu-pu.ac.jp (Y.N.); teramachih@gifu-pu.ac.jp (H.T.); Tel.: +81-58-230-8100 (Y.N. & H.T.)

Received: 13 July 2020; Accepted: 10 August 2020; Published: 12 August 2020



**Abstract:** Many patients require multi-drug combinations, and adverse event profiles reflect not only the effects of individual drugs but also drug–drug interactions. Although there are several algorithms for detecting drug–drug interaction signals, a simple analysis model is required for early detection of adverse events. Recently, there have been reports of detecting signals of drug–drug interactions using subset analysis, but appropriate detection criterion may not have been used. In this study, we presented and verified an appropriate criterion. The data source used was the Japanese Adverse Drug Event Report (JADER) database; “hypothetical” true data were generated through a combination of signals detected by three detection algorithms. The accuracy of the signal detection of the analytic model under investigation was verified using indicators used in machine learning. The newly proposed subset analysis confirmed that the signal detection was improved, compared with signal detection in the previous subset analysis, on the basis of the indicators of *Accuracy* (0.584 to 0.809), *Precision* (= *Positive predictive value*; *PPV*) (0.302 to 0.596), *Specificity* (0.583 to 0.878), *Youden’s index* (0.170 to 0.465), *F-measure* (0.399 to 0.592), and *Negative predictive value* (*NPV*) (0.821 to 0.874). The previous subset analysis detected many false drug–drug interaction signals. Although the newly proposed subset analysis provides slightly lower detection accuracy for drug–drug interaction signals compared to signals compared to the  $\Omega$  shrinkage measure model, the criteria used in the newly subset analysis significantly reduced the amount of falsely detected signals found in the previous subset analysis.

**Keywords:** subset analysis; signal detection algorithms; drug–drug interaction; spontaneous reporting systems

## 1. Introduction

Drug-induced adverse events (AEs) caused by individual drugs and drug combinations not only hinder treatment but also cause new health hazards. To alleviate this problem, AEs caused by individual drug candidates are closely monitored and investigated during the drug development and approval process [1]. Pre-marketing randomized clinical trials are performed under certain conditions associated with age, gender, and co-morbidities, and some AEs may not be detected. In particular, in pre-marketing randomized clinical trials, patients on combination therapy are usually excluded because the focus is to establish the safety and efficacy of single drugs and not to investigate drug–drug interactions [2]. However, in the real world, many patients suffer from a variety of co-morbidities and use a number of drugs to treat them. The concomitant use of two or more drugs increases the

risk of AEs due to drug–drug interactions; the proportion of such AEs is estimated to be up to 30% of unexpected AEs [3].

Therefore, in order to use drugs appropriately in the real world, it is important to understand, in advance, the AEs caused by drug–drug interactions. Post-marketing analysis of AE reports could significantly contribute to the discovery of AEs caused by single drugs or drug–drug interactions that could not be detected before marketing.

For the safety surveillance of drugs, AE reports collected post-marketing are maintained by regulatory agencies as a spontaneous reporting system. There are several algorithms for detecting adverse event signals using the spontaneous reporting system [4]. Of these, the algorithms commonly used for quantitative signal detection include the proportional reporting ratio (PRR) [5], the reporting odds ratio (ROR) [6], the Bayesian confidence propagation neural network (BCPNN) [7], and the empirical Bayesian geometric mean (EBGM) [8].

Additionally, multiple statistical algorithms have been proposed for detecting drug–drug interaction signals [9,10]. However, calculation of the PRR, similar to the risk ratio, and the ROR, similar to the odds ratio, is simple, but that of other algorithms (particularly the algorithm for detecting drug–drug interaction signals) is very complicated.

Therefore, in order to detect the drug–drug interaction signals between *drug D<sub>1</sub>* and *drug D<sub>2</sub>*, the subset analysis that detects the signal of *drug D<sub>1</sub>* using the ROR, which is easy to calculate in a subset of patient groups, is often reported [11–13].

Of previous studies, several [11,12] have used animal experiments and/or pharmacological data to ensure signal reliability, but the signals obtained with this analysis model are not strictly drug–drug interaction signals; they only showed the effect of drug combinations for the following two reasons [14]:

1. The subset analysis used in this study detects signals from the target AE when the patient group using *drug D<sub>1</sub>* takes *drug D<sub>2</sub>*. In all patient groups, when the signal value of the target AE is large for drug *D<sub>2</sub>*, the signal is detected regardless of whether the patient group is using *drug D<sub>1</sub>*.
2. Target AE signal intensities when a patient group using *drug D<sub>1</sub>* takes *drug D<sub>2</sub>* vs. that when a patient group using *drug D<sub>2</sub>* takes *drug D<sub>1</sub>* do not necessarily match. In other words, the value to be adopted as the target AE signal value when *drug D<sub>1</sub>* and *drug D<sub>2</sub>* are used concomitantly has not been fixed (i.e., no clear detection criteria have been defined for detecting drug–drug interaction signals).

On the other hand, because the ROR, often used in subset analysis, is easy to calculate, if these shortcomings are improved and the appropriate detection criterion can be set, it might lead to early detection of AEs caused by drug–drug interactions.

In this study, we proposed a new detection criterion for the subset analysis (the newly proposed subset analysis) and verified the detection power using the spontaneous reporting system.

## 2. Materials and Methods

The design of this study is based on a previous paper that discussed trends in methods to detect the signals of AEs caused by individual drugs [15] or drug–drug interactions [14].

### 2.1. Data Sources

The validation dataset was created from the Japanese Adverse Drug Event Report database (JADER), using data from the first quarter of 2004 to the fourth quarter of 2015. The JADER consists of four comma-separated values (csv) files as data tables: DEMO.csv (patient information), DRUG.csv (medicine information), HIST.csv (patient past history), and REAC.csv (AE event information). This study used 374,327 cases registered in the verification dataset.

However, the Japanese authority, the Pharmaceuticals and Medical Devices Agency (PMDA), which owns these data, does not permit sharing the data directly. Therefore, we do not own the

JADER. It can be accessed directly here: [<http://www.info.pmda.go.jp/fukusayoudb/CsvDownload.jsp>] (in Japanese only).

## 2.2. Definitions of Adverse Drug Events

The drugs targeted for the survey are all registered and classified as “suspect drugs” in the validation dataset. The AEs in JADER are based on the preferred terms (PTs) in the Medical Dictionary for Regulatory Activities Japanese version (MedDRA/J). The AE targeted for this study was Stevens–Johnson syndrome (SJS), which was extracted from the dataset using the PT in MedDRA/J. However, the choice of target adverse events is the same as in previous similar studies [14,15], and there was no medical or pharmacological reason for this choice.

## 2.3. “Hypothetical” True Data of Adverse Events for Comparative Verification

The signals obtained from the spontaneous reporting system including the JADER used in this study included unknown AEs that were also detected, which needs to be verified in order to confirm they were true AEs. Moreover, the information provided by the regulatory authorities, of course, does not include unknown AEs. That is, there are no “real” true data for every AE. Therefore, we set “hypothetical” true data because we cannot use “real” true data for validation in this study.

To verify the power of the subset analysis, we prepared “hypothetical” true data of AEs. To generate “hypothetical” true data, we excluded the  $\Omega$  shrinkage measure model [16] that detected the most conservative signal and the combination risk ratio model [17], which would not detect a signal with a small number of reports, from the five detection algorithms. That is, this study used the combination of signals detected by three algorithms (the additive model [18], the multiplicative model [18], and the chi-square statistical model [19]) as “hypothetical” true data.

## 2.4. Statistical Models and Criteria

### 2.4.1. Subset Analysis

To detect the signal for the interaction between *drug D*<sub>1</sub> and *drug D*<sub>2</sub>, we created subsets of the patient group using *drug D*<sub>1</sub> (or the patient group using *drug D*<sub>2</sub>) (Table 1).

**Table 1.** The 4 × 2 contingency table for signal detection: AE: adverse event; n: the number of reports (e.g.,  $n_{+++}$ : the number of all reports,  $n_{111}$ : the number of *drug D*<sub>1</sub>- and *drug D*<sub>2</sub>-induced target AE reports).

	Target AE	Other AEs	Total
Concomitant use of <i>drug D</i> <sub>1</sub> and <i>drug D</i> <sub>2</sub>	$n_{111}$	$n_{110}$	$n_{11+}$
only <i>drug D</i> <sub>1</sub>	$n_{101}$	$n_{100}$	$n_{10+}$
only <i>drug D</i> <sub>2</sub>	$n_{011}$	$n_{010}$	$n_{01+}$
Neither <i>drug D</i> <sub>1</sub> or <i>drug D</i> <sub>2</sub>	$n_{001}$	$n_{000}$	$n_{00+}$
Total	$n_{++1}$	$n_{++0}$	$n_{+++}$

The following equations (Equations (1) and (2)) were used to calculate the ROR and 95% confidence interval (95% CI) of the target AE caused by *drug D*<sub>1</sub> (or *drug D*<sub>2</sub>) from the generated subset, respectively. For the signal of a patient group on *drug D*<sub>1</sub> that takes *drug D*<sub>2</sub>, the number of each report can be expressed as follows:  $N_{11} = n_{111}$ ,  $N_{10} = n_{110}$ ,  $N_{01} = n_{101}$ ,  $N_{00} = n_{100}$ . On the other hand, for the signal of a patient group on *drug D*<sub>2</sub> that takes *drug D*<sub>1</sub>, the number of each report can be expressed as follows:  $N_{11} = n_{111}$ ,  $N_{10} = n_{110}$ ,  $N_{01} = n_{011}$ ,  $N_{00} = n_{010}$ .

$$\text{ROR} = \frac{N_{11}/N_{00}}{N_{01}/N_{10}} \quad (1)$$

$$\text{ROR (95\% CI)} = e^{\ln(\text{ROR}) \pm 1.96 \sqrt{\frac{1}{N_{11}} + \frac{1}{N_{10}} + \frac{1}{N_{01}} + \frac{1}{N_{00}}}} \quad (2)$$

In previous studies [11–13], if the signal for drug  $D_2$  was detected in the subset of a patient group using drug  $D_1$  or if the signal for drug  $D_1$  was detected in the subset of a patient group using drug  $D_2$ , this signal was considered the drug–drug interaction signal. The criterion that a signal only needs to be detected from a subset of either patient group is ambiguous, highlighting the two shortcomings mentioned earlier. Therefore, for the newly proposed subset analysis, a case was redefined as the drug–drug interaction signal if a signal was detected in both subsets of a patient group using drug  $D_1$  and a patient group using drug  $D_2$ .

#### 2.4.2. $\Omega$ Shrinkage Measure Model

The  $\Omega$  shrinkage measure model [16] is based on a measure calculated as the ratio of the observed reporting ratio of the AE associated with the combination of two drugs and its expected value; this model is used by the Uppsala Monitoring Center (UMC) and the World Health Organization (WHO) Collaborating Centre for International Drug Monitoring for signal analysis of drug–drug interactions (Table 1, Equations (3)–(7)).

$$\Omega = \log_2 \frac{n_{111} + 0.5}{E_{111} + 0.5} \quad (3)$$

where  $n_{111}$  is the reported number of AEs caused by the combination of two drugs, and  $E_{111}$  is the expected value of AEs caused by the combination of two drugs.

$\phi(0.975)$  is 97.5% of the standard normal distribution and  $\Omega_{025} > 0$  is used as a threshold to screen for signals under the combination of two drugs (Equation (4)).

$$\Omega_{025} = \Omega - \frac{\phi(0.975)}{\log(2) \sqrt{n_{111}}} \quad (4)$$

To calculate  $E_{111}$ , we used the following Equations (5)–(7).

$$f_{00} = \frac{n_{001}}{n_{00+}}, f_{10} = \frac{n_{101}}{n_{10+}}, f_{01} = \frac{n_{011}}{n_{01+}}, f_{11} = \frac{n_{111}}{n_{11+}} \quad (5)$$

$$g_{11} = 1 - \frac{1}{\max\left(\frac{f_{00}}{1-f_{00}}, \frac{f_{10}}{1-f_{10}}\right) + \max\left(\frac{f_{00}}{1-f_{00}}, \frac{f_{01}}{1-f_{01}}\right) - \frac{f_{00}}{1-f_{00}} + 1} \quad (6)$$

When  $f_{10} < f_{00}$  (which denotes no risk of AE caused by drug  $D_1$ ), the most sensible estimator  $g_{11} = \max(f_{00}, f_{01})$  is yielded and vice versa when  $f_{01} < f_{00}$ .

Norén et al. re-expressed the observed and expected RRR  $f_{11}$  and  $g_{11}$  in terms of the observed number of reports  $n_{111}$  and expected numbers of reports  $E_{111} = g_{11} \times n_{11+}$ , respectively:

$$\frac{f_{11}}{g_{11}} = \frac{n_{111}/n_{11+}}{E_{111}/n_{11+}} = \frac{n_{111}}{E_{111}} \quad (7)$$

### 2.5. Evaluation of Detection Models

#### 2.5.1. Using Evaluations of Classification in Machine Learning

The evaluation indicators that we set were *Accuracy* (Table 2, Equation (8)), *Precision (Positive predictive value; PPV)* (Table 2, Equation (9)), *Recall (Sensitivity)* (Table 2, Equation (10)), *Specificity* (Table 2, Equation (11)), *Youden's index* (Table 2, Equation (12)), *F-measure* (Table 2, Equation (13)), and *Negative predictive value (NPV)* (Table 2, Equation (14)).

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (8)$$

$$\text{Precision (Positive predictive value; PPV)} = \frac{TP}{TP + FP} \quad (9)$$

$$\text{Recall (Sensitivity)} = \frac{TP}{TP + FN} \quad (10)$$

$$\text{Specificity} = \frac{TN}{FP + TN} \quad (11)$$

$$\text{Youden's index} = \text{Sensitivity} + \text{Specificity} - 1 \quad (12)$$

$$F\text{-measure} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (13)$$

$$\text{Negative predictive value (NPV)} = \frac{TN}{TN + FN} \quad (14)$$

**Table 2.** Agreement between the criterion A and the “hypothetical” true data. AE: adverse event, TP: True Positive, FP: False Positive, FN: False Negative, TN: True Negative.

		“Hypothetical” True Data	
		AE	non-AEs
analysis model	signal	TP	FP
	Non-signal	FN	TN

### 2.5.2. Cohen’s Kappa Coefficient

The commonality of the signals detected by each statistical model was evaluated using *Cohen’s kappa coefficient* ( $\kappa$ ), proportionate agreement for positive rating ( $P_{\text{positive}}$ ), and proportionate agreement for negative rating ( $P_{\text{negative}}$ ), as reported in a previous study [14,15]. In this study, we investigated the similarities with  $\Omega$  shrinkage measure model for the previous/newly proposed subset analysis.

### 2.6. Analysis Software

The analysis software in this study used Visual Mining Studio (NTT DATA Mathematical Systems Inc., Shinjuku-ku, Tokyo, Japan) version 8.4 and Microsoft Excel 2019 (Microsoft Corp., Redmond, WA, USA).

## 3. Results

### 3.1. Evaluations of Classification in Machine Learning

Among all 374,327 cases analyzed, there were 3924 drug  $D_1$ –drug  $D_2$ –SJS combinations. Of these, 923 combinations were detected by all three algorithms—the additive model [18], the multiplicative model [18], and the chi-square statistics model [19]. In this study, these combinations were treated as “hypothetical” true data.

The evaluation of the analysis model is shown in Tables 3 and 4.

**Table 3.** The number of True positive, False positive, True negative, and False negative.

Analysis Model	TP	FP	TN	FN
Previous subset analysis	542	1251	1750	381
Newly proposed subset analysis	542	367	2634	381
$\Omega$ shrinkage measure model	538	174	2827	385

TP: True positive, FP: False positive, TN: True negative, FN: False negative.

**Table 4.** Evaluation of detected drug–drug interaction signals.

Analysis Model	Accuracy	Precision (PPV)	Recall (Sensitivity)	Specificity	Youden's Index	F-Measure	NPV
Previous subset analysis	0.584	0.302	0.587	0.583	0.170	0.399	0.821
Newly proposed subset analysis	0.809	0.596	0.587	0.878	0.465	0.592	0.874
$\Omega$ shrinkage measure model	0.858	0.756	0.583	0.942	0.525	0.658	0.880

PPV: Positive predictive value, NPV: Negative predictive value.

Table 3 shows the number of True positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN).

A total of 1793 combinations were detected by the previous subset analysis (*True positive*: 542, *False positive*: 1251). On the other hand, the newly proposed subset analysis detected 909 combinations of signals (*True positive*: 542, *False positive*: 367) (Table 3).

The detection accuracy shown in Table 4 was calculated from the values shown in Table 3.

In addition, the newly proposed subset analysis confirmed that the signal detection was improved with respect to the indicators of *Accuracy* (0.584 to 0.809), *Precision (PPV)* (0.302 to 0.596), *Specificity* (0.583 to 0.878), *Youden's index* (0.170 to 0.465), *F-measure* (0.399 to 0.592), and *NPV* (0.821 to 0.874) as compared with the signal detection in the previous subset analysis (Table 3).

The values of each indicator of the  $\Omega$  shrinkage measure model were *Accuracy* (0.858), *Precision (PPV)* (0.756), *Recall (Sensitivity)* (0.583), *Specificity* (0.942), *Youden's index* (0.525), *F-measure* (0.658), and *NPV* (0.880) (Table 4).

### 3.2. Cohen's Kappa Coefficient

The similarity between the detection results of the  $\Omega$  shrinkage measure model and that of the newly proposed subset analysis was  $\kappa$  (95% CI): 0.375 (0.355–0.395),  $P_{\text{positive}}$ : 0.502, and  $P_{\text{negative}}$ : 0.870. The similarity was  $\kappa$  (95% CI): 0.355 (0.327–0.384),  $P_{\text{positive}}$ : 0.678, and  $P_{\text{negative}}$ : 0.674 when targeting three or more reports (Table 5).

**Table 5.** The similarity between the  $\Omega$  shrinkage measure model and subset analysis.

Analysis Model	All Case			$n_{111} \geq 3$		
	$\kappa$ (95% CI)	$P_{\text{positive}}$	$P_{\text{negative}}$	$\kappa$ (95% CI)	$P_{\text{positive}}$	$P_{\text{negative}}$
Previous subset analysis	0.088 (0.071–0.105)	0.325	0.684	−0.120 (−0.151–0.088)	0.556	0.296
Newly proposed subset analysis	0.375 (0.355–0.395)	0.502	0.870	0.355 (0.327–0.384)	0.678	0.674

$n_{111}$ : targeting three or more reports,  $\kappa$ : Cohen's kappa coefficient,  $P_{\text{positive}}$ : proportionate agreement for positive rating,  $P_{\text{negative}}$ : proportionate agreement for negative rating.

## 4. Discussions

In this study, we evaluated the accuracy of drug–drug interaction signals for the newly proposed subset analysis that modified two shortcomings of the previous subset analysis on the basis of data from the spontaneous reporting system.

There were 3924 pairs of drug  $D_1$ –drug  $D_2$ –SJS in the spontaneous reporting system, JADER. There are several known combinations of drugs that onset SJS by drug–drug interactions [20]. On the other hand, there are some combinations that have not yet been reported. Recently, we used the  $\Omega$  shrinkage measure model to report potential drug combinations for the onset of SJS in concomitant use with antiepileptic drugs [21]. Not all AEs have been identified and there are still many unknown AEs. Unfortunately, unknown AE data do not exist anywhere in the world; there were no “real” true data for AEs. Therefore, to verify the accuracy of the subset analysis, we needed to prepare “hypothetical” true data of AEs. A previous comparative study [14] of five algorithms for detecting drug–drug interaction signals revealed that the  $\Omega$  shrinkage measure model [16] detected the most conservative signal, while

the combination risk ratio model [17] did not detect any interaction signal in less than three reports due to the detection criterion. Therefore, of the five algorithms, we used the combination of signals detected by the three algorithms (the additive model, the multiplicative model, and the chi-square statistical model) as “hypothetical” true data.

Among the previous subset analysis, the newly proposed subset analysis, and the  $\Omega$  shrinkage measure model, most signals were detected by the previous subset analysis with 1793 pairs (45.7% of the total combinations, *Accuracy*: 0.584, *Precision (PPV)*: 0.302, *Recall (Sensitivity)*: 0.587, *Specificity*: 0.583, *Youden’s index*: 0.170, *F-measure*: 0.399, and *NPV*: 0.821), followed by the newly proposed subset analysis with 909 pairs (23.2% of the total combinations, *Accuracy*: 0.809, *Precision (PPV)*: 0.596, *Recall (Sensitivity)*: 0.587, *Specificity*: 0.878, *Youden’s index*: 0.465, *F-measure*: 0.592, and *NPV*: 0.874). In contrast, the  $\Omega$  shrinkage measure model detected the fewest signals with 712 pairs (18.1% of the total combinations, *Accuracy*: 0.858, *Precision (PPV)*: 0.756, *Recall (Sensitivity)*: 0.583, *Specificity*: 0.942, *Youden’s index*: 0.525, *F-measure*: 0.658, and *NPV*: 0.880) (Table 2, Table 4).

This result indicates that the accuracy of signal detection has been greatly improved in the newly proposed subset analysis with a simple modification of the previous subset analysis. However, the newly proposed subset analysis exhibited slightly lower power and accuracy for detecting the drug–drug interaction signals compared to the  $\Omega$  shrinkage measure model.

Verification by the number of reports showed that when the number of reports ( $N_{11}; n_{111}$ ) < 2, the accuracy (*Youden’s index*, *F-measure*) of signal detection was higher in the newly proposed subset analysis than in the  $\Omega$  shrinkage measure model (*Youden’s index*: the newly proposed subset analysis (0.337) vs. the  $\Omega$  shrinkage measure model (0.174), *F-measure*: the newly proposed subset analysis (0.448) vs. the  $\Omega$  shrinkage measure model (0.298)).

However, as the number of reports increased, the  $\Omega$  shrinkage measure model became more accurate (*Youden’s index*: the newly proposed subset analysis (0.465) vs. the  $\Omega$  shrinkage measure model (0.525), *F-measure*: the newly proposed subset analysis (0.592) vs. the  $\Omega$  shrinkage measure model (0.658)) (Table 4).

Additionally, the *True positive* values for the previous subset analysis and the newly proposed subset analysis were the same (Table 3). Since all signals obtained by the newly proposed subset analysis were included in the previous subset analysis, this result indicates that the detection criterion of the previous subset analysis was loose and that the data contained false positives.

The similarity between the newly proposed subset analysis and the  $\Omega$  shrinkage measure model was  $\kappa$  (95% CI): 0.375 (0.355–0.395),  $P_{\text{positive}}$ : 0.502, and  $P_{\text{negative}}$ : 0.870. On the other hand, the similarity between the previously subset analysis and the  $\Omega$  shrinkage measure model was  $\kappa$  (95% CI): 0.088 (0.071–0.105),  $P_{\text{positive}}$ : 0.325, and  $P_{\text{negative}}$ : 0.684. Thus, the newly proposed subset analysis was more similar to the  $\Omega$  shrinkage measure model than the previously subset analysis. However, the similarity of the newly proposed subset analysis and the  $\Omega$  shrinkage measure model is not very high. Additionally, when the number of reports ( $N_{11}; n_{111}$ ) was  $\geq 3$ , no significant change was observed in the similarity between the  $\Omega$  shrinkage measure model and the newly proposed subset analysis. Despite not being similar to the  $\Omega$  shrinkage model, the newly subset analysis showed a high degree of accuracy. This result suggests that the newly subset analysis may be detecting signals that the  $\Omega$  shrinkage model has failed to detect.

This study has the following three limitations. First, unfortunately, unknown AE data do not exist anywhere in the world [14]. Therefore, there were no “real” true data for AEs. Thus, for the purpose of verification, it was necessary to set “hypothetical” true data for AEs instead of “real” true data. Therefore, of the five algorithms for detecting drug–drug interaction signals, we used the combination of signals detected by the three algorithms (the additive model, the multiplicative model, and the chi-square statistical model) as “hypothetical” true data in this study. In other words, the hypothetical true data consisted of statistically based drug  $D_1$ –drug  $D_2$ –AE combinations, not pharmacologically based combinations.

Second, usually it is important to compare detection trends using all AEs recorded in the validation dataset created on the basis of a spontaneous reporting system; however, it takes an extremely long time to calculate signal values for all combinations of drug–drug interactions. Such a study design is not realistic. Therefore, this study targeted SJS, the same AE used in previous comparative studies [14,15]; if different reference sets were used, the possibility of obtaining different performance characteristics might not be ruled out. There are fewer enrolled cases than in the global dataset because JADER is limited to cases in Japan. However, the signal detection is based on a comparison between the ratio of reported cases ( $N$ ) to expected values ( $E$ ). Therefore, differences in the number of cases enrolled in the spontaneous reporting system had only a very small statistical impact in this study. Recently, validation of the number of cases enrolled in the spontaneous reporting system has also been reported by Caster et al. [22]. Moreover, differences in the way regulatory authorities think may result in a different tendency to register AEs to the spontaneous reporting system. For example, the Food and Drug Administration Adverse Events Reporting System (FAERS) in the United States has also registered reports from non-medical professionals, but JADER has not registered reports from patients until recently. It is unknown how the differences in registration tendencies affect the results of this study.

Finally, neither the general algorithms for detecting drug–drug interaction signals nor the proposed subset analysis in this study were antagonistic; only signals of synergistic interactions were detected [10].

## 5. Conclusions

In recent years, the need for safety signal screening has been demanded, not only for single drugs but also for drug–drug interactions. Although several methods for detecting signals of drug interaction have been reported, it is difficult to say that these methods are used because many of them are complicated in calculation. Therefore, there were several cases [11–13] where subset analysis using the algorithm for detecting signals of single drugs (e.g., ROR [6]) was used for signal detection of drug–drug interactions before its validity was verified.

This study showed that there were many false positives in the existing subset analysis, albeit under limited conditions. Additionally, very simple modifications of the detection criteria were made to solve two problems associated with the previous subset analysis for exploring recently reported drug–drug interaction signals. This modification helped to reduce falsely detected signals found in the previous subset analysis.

Moreover, the newly proposed subset analysis is more similar to the  $\Omega$  shrinkage measure model than the previous subset analysis, but the similarity with the  $\Omega$  shrinkage measure model is not as high. However, the newly proposed subset analysis showed that although the detection accuracy of the drug–drug interaction signal was slightly lower than that of the  $\Omega$  shrinkage measure model, the detection accuracy was sufficient. This result may also indicate the possibility of detecting signals that cannot be detected by the  $\Omega$  shrinkage measure model.

**Author Contributions:** Conceptualization, Y.N.; funding acquisition, Y.N.; investigation, Y.N.; methodology, Y.N. and T.T.; project administration, Y.N. and H.T.; supervision, Y.N.; validation, Y.N. and T.T.; visualization, Y.N.; writing—original draft, Y.N.; writing—review and editing, Y.N., T.T., and H.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by JSPS KAKENHI grant number 19K20731.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Berlin, J.A.; Glasser, S.C.; Ellenberg, S.S. Adverse event detection in drug development: Recommendations and obligations beyond phase 3. *Am. J. Public Health* **2008**, *98*, 1366–1371. [[CrossRef](#)] [[PubMed](#)]
2. Noguchi, Y.; Ueno, A.; Otsubo, M.; Katsuno, H.; Sugita, I.; Kanematsu, Y.; Yoshida, A.; Esaki, H.; Tachi, T.; Teramachi, H. A New Search Method Using Association Rule Mining for Drug-Drug Interaction Based on Spontaneous Report System. *Front. Pharmacol.* **2018**, *9*, 197. [[CrossRef](#)] [[PubMed](#)]

3. Iyer, S.V.; Harpaz, R.; LePendu, P.; Bauer-Mehren, A.; Shah, N.H. Mining clinical text for signals of adverse drug–drug interactions. *J. Am. Med. Inf. Assoc.* **2013**, *21*, 353–362. [[CrossRef](#)] [[PubMed](#)]
4. Suling, M.; Pigeot, I. Signal Detection and Monitoring Based on Longitudinal Healthcare Data. *Pharmaceutics* **2012**, *4*, 607–640. [[CrossRef](#)] [[PubMed](#)]
5. Evans, S.J.; Waller, P.C.; Davis, S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiol. Drug Saf.* **2001**, *10*, 483–486. [[CrossRef](#)] [[PubMed](#)]
6. Rothman, K.J.; Lanes, S.; Sacks, S.T. The reporting odds ratio and its advantages over the proportional reporting ratio. *Pharmacoepidemiol. Drug Saf.* **2004**, *13*, 519–523. [[CrossRef](#)]
7. Bate, A.; Lindquist, M.; Edwards, I.R.; Olsson, S.; Orre, R.; Lansner, A.; De Freitas, R.M. A Bayesian neural network method for adverse drug reaction signal generation. *Eur. J. Clin. Pharmacol.* **1998**, *54*, 315–321. [[CrossRef](#)]
8. DuMouchel, W. Bayesian Data Mining in Large Frequency Tables, with an Application to the FDA Spontaneous Reporting System. *Am. Stat.* **1999**, *53*, 177–190. [[CrossRef](#)]
9. Vilar, S.; Friedman, C.; Hripcsak, G. Detection of drug–drug interactions through data mining studies using clinical sources, scientific literature and social media. *Brief. Bioinform.* **2018**, *19*, 863–877. [[CrossRef](#)]
10. Noguchi, Y.; Tachi, T.; Teramachi, H. Review of Statistical Methodologies for Detecting Drug–Drug Interactions Using Spontaneous Reporting Systems. *Front. Pharmacol.* **2019**, *10*, 1319. [[CrossRef](#)]
11. Nagashima, T.; Shirakawa, H.; Nakagawa, T.; Kaneko, S. Prevention of antipsychotic-induced hyperglycaemia by vitamin D: A data mining prediction followed by experimental exploration of the molecular mechanism. *Sci. Rep.* **2016**, *6*, 26375. [[CrossRef](#)] [[PubMed](#)]
12. Uno, T.; Wada, K.; Hosomi, K.; Matsuda, S.; Ikura, M.M.; Takenaka, H.; Terakawa, N.; Oita, A.; Yokoyama, S.; Kawase, A.; et al. Drug interactions between tacrolimus and clotrimazole troche: A data mining approach followed by a pharmacokinetic study. *Eur. J. Clin. Pharmacol.* **2020**, *76*, 117–125. [[CrossRef](#)] [[PubMed](#)]
13. Sanagawa, A.; Hotta, Y.; Kondo, M.; Nishikawa, R.; Tohkin, M.; Kimura, K. Tumor lysis syndrome associated with bortezomib: A post-hoc analysis after signal detection using the US Food and Drug Administration Adverse Event Reporting System. *Anti-Cancer Drugs* **2020**, *31*, 183–189. [[CrossRef](#)] [[PubMed](#)]
14. Noguchi, Y.; Tachi, T.; Teramachi, H. Comparison of signal detection algorithms based on frequency statistical model for drug–drug interaction using spontaneous reporting systems. *Pharm. Res.* **2020**, *37*, 86. [[CrossRef](#)]
15. Kubota, K.; Koide, D.; Hirai, T. Comparison of data mining methodologies using Japanese spontaneous reports. *Pharmacoepidemiol. Drug Saf.* **2004**, *13*, 387–394. [[CrossRef](#)]
16. Norén, G.N.; Sundberg, R.; Bate, A.; Edwards, I.R. A statistical methodology for drug–drug interaction surveillance. *Stat. Med.* **2008**, *27*, 3057–3070. [[CrossRef](#)]
17. Susuta, Y.; Takahashi, Y. Safety risk evaluation methodology in detecting the medicine concomitant use risk which might cause critical drug rash. *Jpn. J. Pharmacoepidemiol.* **2014**, *19*, 39–49. [[CrossRef](#)]
18. Thakrar, B.T.; Grundschober, S.B.; Doessegger, L. Detecting signals of drug–drug interactions in a spontaneous reports database. *Br. J. Clin. Pharmacol.* **2007**, *64*, 489–495. [[CrossRef](#)]
19. Gosho, M.; Maruo, K.; Tada, K.; Hirakawa, A. Utilization of chi-square statistics for screening adverse drug–drug interactions in spontaneous reporting systems. *Eur. J. Clin. Pharmacol.* **2017**, 73779–73786. [[CrossRef](#)]
20. Cheng, F.J.; Syu, F.K.; Lee, K.H.; Chen, F.C.; Wu, C.H.; Chen, C.C. Correlation between drug–drug interaction-induced Stevens–Johnson syndrome and related deaths in Taiwan. *J. Food Drug Anal.* **2016**, *24*, 427–432. [[CrossRef](#)]
21. Noguchi, Y.; Takaoka, M.; Hayashi, T.; Tachi, T.; Teramachi, H. Antiepileptic combination therapy with Stevens–Johnson syndrome and toxic epidermal necrolysis: Analysis of a Japanese pharmacovigilance database. *Epilepsia* **2020**. [[CrossRef](#)] [[PubMed](#)]
22. Caster, O.; Aoki, Y.; Gattepaille, L.M.; Grundmark, B. Disproportionality Analysis for Pharmacovigilance Signal Detection in Small Databases or Subsets: Recommendations for Limiting False-Positive Associations. *Drug Saf.* **2020**, *43*, 479–487. [[CrossRef](#)] [[PubMed](#)]

