Methods and Applications

Comparison of Statistical Signal Detection Methods in Adverse Events Following Immunization — China, 2011–2015

Lanfang Xia[1,2]; Keli Li[1,2]; Yan Li[1,2]; Zhijie An[1,2]; Quanwei Song[1,2]; Lei Wang[1,2]; Zundong Yin[1,2]; Huaqing Wang[1,2,#]

ABSTRACT

Introduction: The current study aims to assess the performance of data mining techniques in detecting safety signals for adverse events following immunization (AEFI) using routinely obtained data in China. Four different methods for detecting vaccine safety signals were evaluated.

Methods: The AEFI data from 2011 to 2015 was collected for our study. We analyzed the data using four different methods to detect signals: the proportional reporting ratio (PRR), reporting odds ratio (ROR), Bayesian confidence propagation neural network (BCPNN), and multi-item gamma Poisson shrinker (MGPS). Each method was evaluated at 1–3 thresholds for positivity. To assess the performance of these methods, we used the published signal rates as gold standards to determine the sensitivity and specificity.

Results: The number of identified signals varied from 602 for PRR1 (with a threshold of 1) to 127 for MGPS1. When considering the common reactions as the reference standard, the sensitivity ranged from 0.9% for MGPS1/2 to 38.2% for PRR1/2, and the specificity ranged from 85.2% for PRR1 and ROR1 to 96.7% for MGPS1. When considering the rare reactions as the reference standard, PRR1, PRR2, ROR1, ROR2, and BCPNN exhibited the highest sensitivity (73.3%), while MGPS1 exhibited the highest specificity (96.9%).

Discussion: For common reactions, the sensitivities were modest and the specificities were high. For rare reactions, both the sensitivities and specificities were high. Our study provides valuable insights into the selection of signal detection methods and thresholds for AEFI data in China.

Data mining techniques have been widely employed since the late 1990s for identifying safety signals in databases containing spontaneously reported adverse reactions of drugs and vaccines (*1*–*5*). The primary objective is to generate hypotheses for further evaluation of potential safety concerns. Disproportionality analysis, a case/non-case method that compares observed rates with expected rates, is the most commonly used technique for signal detection (*1*–*5*). However, the performance of these methods and the impact of different thresholds on their performance in detecting safety signals in adverse events following immunization (AEFI) reports in China remain unknown. It is crucial to assess the performance of each signal detection method to establish a reference for routine vaccine safety signal detection.

We evaluated the performance of safety signal detection algorithms in detecting AEFI using data collected in China from 2011 to 2015. The number of signals detected and the operating characteristics of these algorithms were analyzed. Sensitivity and specificity were estimated using published data as gold standards, with different threshold values for each algorithm (*2*–*7*). The findings of this study can guide the selection of suitable detection methods and threshold values for vaccine safety surveillance in China.

METHODS

Data Sources

The study utilized spontaneous AEFI reports from the national AEFI information system (*8*) from 2011 to 2015. Data preparation involved several steps, including the removal of confidential information (*9*) and duplicate reports, as well as reports without valid AEFI clinical diagnosis that would not contribute to the vaccine-AEFI pair. Additionally, reports with multiple AEFI clinical diagnoses or suspected vaccines were separated into multiple individual reports, each with a unique AEFI clinical diagnosis and suspected vaccine (*10*). The AEFI clinical diagnoses were coded

using the Medical Dictionary for Regulatory Activities (MedDRA) dictionary, version 24.0 (*11*). All vaccines and AEFI clinical diagnoses were included in the analysis.

Signal Detection Methods and Thresholds

Statistical signal detection was performed by analyzing the reporting rates of specific adverse events associated with specific vaccines. Four commonly used methods for analyzing disproportionality were applied: proportional reporting ratio (PRR) (*4*), reporting odds ratio (ROR) (*2*–*3*), Bayesian confidence propagation neural network (BCPNN) (*1*,*3*), and multi-item gamma Poisson shrinker (MGPS) (*5*). PRR and ROR are frequentist methods, while BCPNN and MGPS are Bayesian methods. The computation techniques for each method can be found in the supplementary table S1 and relevant publications (*1*–*5*).

The disproportionality analysis is based on a 2×2 table (Table 1). In this table, cell a represents the number of reports containing both the vaccine of interest and the AEFI of interest. Cell b represents the number of reports containing the vaccine of interest with AEFIs other than the AEFI of interest. Cell c represents the number of reports containing the AEFI of interest with vaccines other than the vaccine of interest. Cell d represents the number of reports containing AEFIs and vaccines other than the ones of interest.

TABLE 1. Two-by-two contingency table for signal detection.

	AEFI of interest	Other AEFIs	Total
Vaccine of interest	a	b	a+b
Other vaccines	c	d	c+d
Total	a+c	b+d	N

Note: "a" means number of reports containing both the vaccine of interest and the AEFI of interest; "b" means number of reports containing the vaccine of interest with AEFIs other than the AEFI of interest; "c" means number of reports containing the AEFI of interest with vaccines other than the vaccine of interest; "d" means number of reports containing AEFIs and vaccines other than the ones of interest.
Abbreviation: AEFI=adverse events following immunization.

Table 2 presents the signal detection methods and the threshold values to be assessed. Each signal detection method was evaluated using up to three signal threshold values. Vaccine-AEFI combinations with statistical values exceeding the threshold values were deemed as positive signals.

Performance Evaluation

We calculated the operating characteristics (sensitivity and specificity) of each signal detection algorithm to classify each vaccine-AEFI combination as either a signal or a non-signal. We used published reference standards as our gold standards (*14*).

Two sets of reference standards based on the global manual on the surveillance of adverse events following immunization by the World Health Organization (WHO) (*15*) and safety signals from previous studies (*16*–*20*) were created (Table 3). Sensitivity and specificity were determined and presented in Table 4.

$$\textbf{Sensitivity} = \frac{A}{A+C} \times 100\%$$

$$\textbf{Specificity} = \frac{D}{B+D} \times 100\%$$

TABLE 2. Signal detection methods and thresholds to be evaluated.

Signal detection method	Signal detection algorithm*	Threshold†
PRR	PRR1	Lower limit of 95% *CI* of PRR > 1(7) and a ≥ 3
	PRR2	Lower limit of 95% *CI* of PRR > 1(7) and a ≥ 5
	PRR3	PRR > 2 and $\chi^2 \geq 4$ and a ≥ 3 (7)
ROR	ROR1	Lower limit of 95% *CI* of ROR > 1 and a ≥ 3 (7,12)
	ROR2	Lower limit of 95% *CI* of ROR > 1 and a ≥ 5 (12)
BCPNN	BCPNN	Lower limit of 95% *CI* of IC > 0 (12)
MGPS	MGPS1	5th percentile of EBGM (EB05) > 2 (7)
	MGPS2	5th percentile of EBGM (EB05) ≥ 1.8 and EBGM ≥ 2.5 (12)
	MGPS3	EBGM ≥ 2 (7,13)

Abbreviation: PRR=proportional reporting ratio; ROR=reporting odds ratio; BCPNN=Bayesian confidence propagation neural network; MGPS=multi-item gamma Poisson shrinker; *CI*=confidence interval; IC=information component; EBGM=empirical Bayesian geometric mean.
* The number refers to various thresholds.
† The variable "a" represents the number of reports that include both the specific vaccine being studied and the AEFI being investigated.

TABLE 3. Reference standard for performance evaluation.

Reference standard	Vaccine	AEFI
Reference standard 1	All vaccines (*15*)	Fever (temperature ≥38.6 ℃)
(common events)	All injectable vaccines (*15*)	Vaccination site erythema (diameter >2.5 cm), Vaccination site induration (diameter >2.5 cm)
Reference standard 2 (rare events)	Live-attenuated Hepatitis A vaccine (*16*–*20*)	Anaphylactic shock
	Varicella vaccine (*16*–*20*)	Anaphylactic shock
	BCG (*15*)	Vaccination site abscess, lymphadenitis, disseminated BCG infection
	Live-attenuated oral Polio vaccine (*15*)	Vaccine-associated paralytic poliomyelitis
	Measles containing vaccines (*15*)	Thrombocytopenic purpura
	Measles containing vaccines (*15*)	Rash morbilliform

Abbreviation: AEFI=adverse events following immunization; BCG=Bacillus Calmette-Guérin.

TABLE 4. Two-by-two contingency table for performance evaluation.

Test	Reference standard		Total
	Positive	Negative	
Positive	True positive (A)	False positive (B)	A+B
Negative	False negative (C)	True negative (D)	C+D
Total	A+C	B+D	N

Note: "A" means number of vaccine-AEFI combinations listed in reference standard and detected in this study; "B" means number of vaccine-AEFI combinations not listed in reference standard but detected in this study; "C" means number of vaccine-AEFI combinations listed in reference standard but not detected in this study; "D" means number of vaccine-AEFI combinations not listed in reference standard and not detected in this study.

Analyses

The baseline characteristics of AEFI data were analyzed to assess disproportionality. We examined the number of signals detected by each signal detection algorithm and calculated the cumulative distribution of signals for each algorithm. The distribution of signals was determined by dividing the number of signals with a specific number of reports by the total number of signals detected by each algorithm. Sensitivity and specificity for each detection method were also determined as described previously. Analyses were conducted using R software (version 4.3.1, The R Foundation for Statistical Computing, Lucent Technologies, Auckland, New Zealand) and the PhViD package (version 1.0.8, The R Foundation for Statistical Computing, Lucent Technologies, Auckland, New Zealand).

RESULTS

The original dataset consisted of 587,149 reports documenting AEFI. After removing 762 duplicate records and 15,844 records without a valid AEFI clinical diagnosis, these unique records were further separated into individual reports, resulting in 871,647 records that contained both an AEFI and a suspected vaccine. After removing 485 records without a valid vaccine name and 132 records with duplicate AEFI-vaccine pairings, the final analyzable data set consisted of 871,030 records. This data set included 41 different vaccines, 771 specific AEFI events, and 3,893 unique combinations of vaccines and AEFI events.

Table 5 presents the number of signals detected by each signal detection algorithm. PRR1 and ROR1 identified the highest number of signals, while MPGS1 identified the lowest number of signals. MPGS methods detected fewer signals compared to PRR, ROR, and BCPNN.

PRR3, PRR1, ROR1, and MGPS3, exhibited the highest number of signals detected when the number of reports was five or fewer. Conversely, MGPS1 did not detect any signals when the number of reports was 5 or fewer. Among the signals detected by PRR3, 38.9% had a number of reports equal to or less than 5, while for MGPS3 this percentage was 33.7%. On the other hand, PRR1, PRR2, ROR1, ROR2, and BCPNN identified the greatest number of signals when the number of reports exceeded five.

Algorithm Performance

Table 6 presents the sensitivity and specificity values for each signal detection algorithm using two reference

TABLE 5. Cumulative distribution of number of reports for signals detected by each signal detection algorithm using AEFI in China from 2011–2015.

Number of reports*	PRR1 *n* (%)	PRR2 *n* (%)	PRR3 *n* (%)	ROR1 *n* (%)	ROR2 *n* (%)	BCPNN *n* (%)	MGPS1 *n* (%)	MGPS2 *n* (%)	MGPS3 *n* (%)
a≤3	84 (14.0)	0	86 (18.3)	84 (14.0)	0	25 (5.0)	0	0	42 (12.0)
a≤4	139 (23.1)	0	141 (30.0)	139 (23.1)	0	74 (14.8)	0	0	88 (25.1)
a≤5	181 (30.1)	42 (9.1)	183 (38.9)	181 (30.1)	42 (9.1)	106 (21.2)	0	0	118 (33.7)
a≤6	217 (36.0)	78 (16.8)	218 (46.4)	216 (35.9)	77 (16.7)	136 (27.1)	2 (1.6)	4 (2.8)	146 (41.7)
a≤7	241 (40.0)	102 (22.0)	242 (51.5)	240 (39.9)	101 (21.9)	158 (31.5)	3 (2.4)	6 (4.2)	163 (46.6)
a≤8	271 (45.0)	132 (28.5)	271 (57.7)	270 (44.9)	131 (28.4)	182 (36.3)	11 (8.7)	19 (13.2)	184 (52.6)
a≤9	284 (47.2)	145 (31.3)	284 (60.4)	283 (47.1)	144 (31.2)	195 (38.9)	16 (12.6)	25 (17.4)	194 (55.4)
a≤10	300 (49.8)	161 (34.8)	297 (63.2)	299 (49.8)	160 (34.6)	208 (41.5)	21 (16.5)	32 (22.2)	201 (57.4)
a>5	421 (69.9)	421 (90.9)	287 (61.1)	420 (69.9)	420 (90.9)	395 (78.8)	127 (100)	144 (100)	232 (66.3)
a>10	302 (50.2)	302 (65.2)	173 (36.8)	302 (50.2)	302 (65.4)	293 (58.5)	106 (83.5)	112 (77.8)	149 (42.6)

Note: %=Number of signals in each category divided by the total number of signals detected by each method multiplied by 100. The number after each method refers to various thresholds.
Abbreviation: PRR=proportional reporting ratio; ROR=reporting odds ratio; BCPNN=Bayesian confidence propagation neural network; MGPS=multi-item gamma Poisson shrinker.
* A represents the number of reports containing both the vaccine of interest and the AEFI of interest.

standards: reference standard one for common adverse events and reference standard two for rare adverse events. Based on reference standard one, the algorithms PRR1 and PRR2 demonstrated the highest sensitivity at 38.2%, closely followed by ROR1 and ROR2 at 37.3%. MGPS1 exhibited the lowest sensitivity at 0.9%. On the other hand, MGPS1 exhibited the highest specificity at 96.7%, followed by MGPS2 at 96.2%. MGPS sensitivity was significantly lower than that of PRR, ROR, and BCPNN, while its specificity was higher than that of PRR, ROR, and BCPNN.

Based on reference standard 2, the diagnostic tests PRR1, PRR2, ROR1, ROR2, and BCPNN exhibited the highest sensitivity (73.3%), while PRR3, MGPS1, MGPS2, and MGPS3 showed a lower sensitivity (53.3%). Among the tests, MGPS1 demonstrated the highest specificity (96.9%). Although MGPS had lower sensitivity compared to PRR, ROR, and BCPNN, its specificity was higher than those three tests.

DISCUSSION

Our study aimed to assess the main features of commonly employed algorithms for detecting signals in spontaneous reporting datasets. Specifically, we examined the performance of four signal detection methods in identifying vaccine safety signals within AEFI data collected in China from 2011 to 2015. To do this, we analyzed the data using different thresholds of signal positivity. In order to evaluate the accuracy of the algorithms, we compared their results to reference standards from published scientific analyses, which were considered as the gold standard. From these comparisons, we calculated the sensitivities and specificities of each algorithm.

The number of signals detected varied significantly among the algorithms, which aligns with the findings of Kubota and colleagues (*21*). The PRR and ROR methods identified the highest number of safety signals, while MGPS method identified the fewest signals. Specifically, PRR1 found 475 more signals than MGPS1. The distribution of signals differed significantly among algorithms when the number of reports was five or fewer, but not when the number exceeded five. PRR1 and ROR1 demonstrated similar performance in signal identification, as did PRR2 and ROR2. The variation in the number of signals identified by PRR1, ROR1, and PRR3 was related to the variability in signals for more commonly reported events (i.e., those with more than five reports). On the other hand, the variability in PRR1 (ROR1) compared to PRR2, ROR2, BCPNN, and MGPS was due to differences in signal identification when the number of reports was fewer than five.

The signal-finding algorithms showed considerable variation in sensitivity and specificity. PRR1 and PRR2 demonstrated the highest sensitivity, followed by ROR1, ROR2, and BCPNN, which were also sensitive but to a lesser extent. However, MGPS1 exhibited the highest specificity, but had the lowest sensitivity. Further research is needed to investigate the reasons

TABLE 6. Performance of each signal detection algorithm.

Signal detection method	No. of signals	True positive (A)	False positive (B)	False negative (C)	True negative (D)	Sensitivity (%)	Specificity (%)
Based on reference standard 1							
PRR1	602	42	560	68	3,223	38.2	85.2
PRR2	463	42	421	68	3,362	38.2	88.9
PRR3	470	5	465	105	3,318	4.5	87.7
ROR1	601	41	560	69	3,223	37.3	85.2
ROR2	462	41	421	69	3,362	37.3	88.9
BCPNN	501	40	461	70	3,322	36.4	87.8
MGPS1	127	1	126	109	3,657	0.9	96.7
MGPS2	144	1	143	109	3,640	0.9	96.2
MGPS3	350	2	348	108	3,435	1.8	90.8
Based on reference standard 2							
PRR1	602	11	591	4	3,287	73.3	84.8
PRR2	463	11	452	4	3,426	73.3	88.3
PRR3	470	8	462	7	3,416	53.3	88.1
ROR1	601	11	590	4	3,288	73.3	84.8
ROR2	462	11	451	4	3,427	73.3	88.4
BCPNN	501	11	490	4	3,388	73.3	87.4
MGPS1	127	8	119	7	3,759	53.3	96.9
MGPS2	144	8	136	7	3,742	53.3	96.5
MGPS3	350	8	342	7	3,536	53.3	91.2

Note: "A" means number of vaccine-AEFI combinations listed in reference standard and detected in this study; "B" means number of vaccine-AEFI combinations not listed in reference standard but detected in this study; "C" means number of vaccine-AEFI combinations listed in reference standard but not detected in this study; "D" means number of vaccine-AEFI combinations not listed in reference standard and not detected in this study.
Abbreviation: PRR=proportional reporting ratio; ROR=reporting odds ratio; BCPNN=Bayesian confidence propagation neural network; MGPS=multi-item gamma Poisson shrinker. The number after each method refers to various thresholds.

behind this finding. When using the reference standard for rare side effects, PRR1, PRR2, ROR1, ROR2, and BCPNN were more sensitive than PRR3, MGPS1, MGPS2, and MGPS3. MGPS1 was found to be the most specific. In summary, our study indicates that PRR, ROR, and BCPNN are more sensitive than MGPS for detecting safety signals, while MGPS is more specific. These findings align with previous studies (*12*,*21*–*22*).

The initial analysis of our data highlighted the significance of data preparation. The standardized processing of data is crucial for ensuring consistent signal detection analyses (*9*). In order to perform signal detection analyses, it is necessary to preprocess spontaneous AEFI reports. This involves eliminating duplicate and invalid records, as well as separating AEFI-vaccine pairs in reports that contain multiple pairs.

Variations in the number of reports, as well as sensitivity and specificity, can be attributed to several factors. First, computation methods differ when dealing with a small number of reports (*21*–*22*). Bayesian shrinkage calculations used by BCPNN and MGPS result in more stable but conservative results compared to PRR and ROR. Second, variations arise from the different thresholds selected (*21*). Future research should systematically evaluate the impact of threshold values on sensitivity and specificity. Therefore, the variations observed in the number of reports and sensitivity and specificity highlight the importance of selecting appropriate signal detection methods and threshold values based on specific use case scenarios.

To the best of our knowledge, this is the first study to investigate the reference standard for performance evaluation. We systematically evaluated the variation in the number of reports, as well as the sensitivity and specificity of the signal detection method, using the AEFI database in China. Our findings can offer valuable insights for the selection of signal detection

methods and corresponding threshold values for the routine signal detection system in China's AEFI data. It is important to strike a balance between sensitivity and specificity when choosing signal detection methods and threshold values, while considering factors such as the ability to investigate detected signals (*12*), the severity of the AEFI under investigation, and the potential impact on public health if a true safety signal is missed.

Based on our study and an extensive review of relevant scientific literature, we propose different approaches for the detection of AEFI, depending on the severity and prevalence of the events, as well as the type of vaccine. For common or mild AEFIs, we recommend utilizing more specific signal detection methods such as the BCPNN or the MGPS, along with more stringent thresholds such as PRR2 or ROR2. These methods and thresholds can effectively reduce the number of false positives. In contrast, for rare or severe AEFIs, or for new licensed vaccines, we advise using more sensitive signal detection methods like the PRR or the ROR, along with less stringent thresholds. These approaches are designed to minimize the risk of missing true signals.

This study has some limitations. First, there is no universally accepted gold standard for evaluating the performance of signal detection (*12*). In this study, we used reference standards based on the World Health Organization's global manual on surveillance of adverse events following immunization (*15*) and safety signals identified in previous studies (*16*–*20*) as the gold standards. Second, AEFI data are collected through a passive surveillance system, and the quality of the reports may affect the detection of signals. Additionally, AEFI data is subject to known limitations, such as under-reporting, selective reporting, or over-reporting (*23*). Therefore, safety signals identified solely based on AEFI data in this study cannot determine causality and should be interpreted cautiously.

In our study, we conducted a comprehensive analysis of the number of signals detected and the performance of various methods for vaccine safety signal detection. The analysis was based on data from a passive, spontaneously reported database of AEFI. We recommend further research to evaluate the specific characteristics of the identified signals and assess the impact of different thresholds on signal detection accuracy. This additional research will provide valuable insights for enhancing the accuracy of vaccine safety signal detection in the context of vaccines used in China.

Conflicts of interest: All authors declare no competing interests.

Acknowledgments: Dr. Lance Rodewald for his review and valuable feedback.

Funding: Supported by the Beijing Natural Science Foundation (L212058).

doi: 10.46234/ccdcw2024.066

[#] Corresponding author: Huaqing Wang, wanghq@chinacdc.cn.

[1] National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases (NITFID), Beijing, China; [2] National Immunization Program, Chinese Center for Disease Control and Prevention, Beijing, China.

Submitted: November 28, 2023; Accepted: January 24, 2024

REFERENCES

1. Bate A, Lindquist M, Edwards IR, Olsson S, Orre R, Lansner A, et al. A Bayesian neural network method for adverse drug reaction signal generation. Eur J Clin Pharmacol 1998;54(4):315 – 21. https://doi.org/10.1007/s002280050466.
2. Van Puijenbroek EP, Diemont WL, Van Grootheest K. Application of quantitative signal detection in the Dutch spontaneous reporting system for adverse drug reactions. Drug Saf 2003;26(5):293 – 301. https://doi.org/10.2165/00002018-200326050-00001.
3. Van Puijenbroek EP, Bate A, Leufkens HGM, Lindquist M, Orre R, Egberts ACG. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. Pharmacoepidemiol Drug Saf 2002;11(1):3 – 10. https://doi.org/10.1002/pds.668.
4. Evans SJW, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. Pharmacoepidemiol Drug Saf 2001;10(6):483 – 6. https://doi.org/10.1002/pds.677.
5. Dumouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. Am Stat 1999;53(3):177 – 90. https://doi.org/10.1080/00031305.1999.10474456.
6. Orre R, Lansner A, Bate A, Lindquist M. Bayesian neural networks with confidence estimations applied to data mining. Comput Stat Data Anal 2000;34(4):473 – 93. https://doi.org/10.1016/S0167-9473(99)00114-0.
7. Deshpande G, Gogolak V, Smith SW. Data mining in drug safety: review of published threshold criteria for defining signals of disproportionate reporting. Pharm Med 2010;24(1):37 – 43. https://doi.org/10.1007/BF03256796.
8. Cao LS, Yuan P. Thinking of China national immunization program information management system construction. Chin J Vaccines Immun 2010;16(6):553 – 7. https://doi.org/10.19914/j.cjvi.2010.06.026.
9. Gu CM, Li Y, Wang BH. Real world data and evidence. Scientific and Technical Documentation Press. Beijing, 2022. (In Chinese).
10. Li CJ. Researches on theory and application of adverse drug reaction signal detection [dissertation]. 2009. https://kns.cnki.net/kcms2/article/abstract?v=wcPNn8Zia7Ohp-UUl8oj7LKtjFIfAVpXcdzz9hIelBhHz7eLSkqvM37V14qbdr_zZW4zBKjpJLXt9iWOvmJCbSCZ26XThEsYhwGFVpqngeCT4LgnO12Aw8Npt6kIYa3j8RvfewILU8qL3XZB8jKlGQ==&uniplatorm=NZKPT&languag=CHS. [2023-03-18]. (In Chinese)
11. Brown EG, Wood L, Wood S. The medical dictionary for regulatory activities (MedDRA). Drug Saf 1999;20(2):109 – 17. https://doi.org/10.2165/00002018-199920020-00002.
12. Candore G, Juhlin K, Manlik K, Thakrar B, Quarcoo N, Seabroke S, et

al. Comparison of statistical signal detection methods within and across spontaneous reporting databases. Drug Saf 2015;38(6):577 – 87. https://doi.org/10.1007/s40264-015-0289-5.
13. Niu MT, Erwin DE, Braun MM. Data mining in the US Vaccine Adverse Event Reporting System (VAERS): early detection of intussusception and other events after rotavirus vaccination. Vaccine 2001;19(32):4627 – 34. https://doi.org/10.1016/s0264-410x(01)00237-7.
14. Grunau G, Linn S. Commentary: sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice. Front Public Health 2018;6:256. https://doi.org/10.3389/fpubh.2018.00256.
15. WHO. Global manual on surveillance of adverse events following immunization. Geneva: WHO; 2016. https://www.who.int/publications/i/item/9789241507769.
16. Wu WD, Li KL, Zheng JS, Liu DW, Xu DS, Yang H, et al. Analysis on surveillance data of adverse events following immunization in China, 2011. Chin J Vaccines Immun 2013;19(2):97 – 109. https://doi.org/10.19914/j.cjvi.2013.02.001.
17. Wu WD, Liu DW, Li KL, Zheng JS, Xu DS, Wang YM, et al. Analysis on surveillance data of adverse events following immunization in China, 2012. Chin J Vaccines Immun 2014;20(1):1 – 12,66. https://doi.org/10.19914/j.cjvi.2014.01.001.
18. Ye JK, Li KL, Xu DS, Wu WD, Liu DW, Zheng JS, et al. Analysis of surveillance for adverse events following immunization in China, 2014. Chin J Vaccines Immun 2016;22(2):125 – 37. https://doi.org/10.19914/j.cjvi.2016.02.002.
19. Ye JK, Li KL, Xu DS, Wu WD, Liu DW, Zheng JS, et al. Evaluation of the adverse events following immunization information management system in China, 2013. Chin J Vaccines Immun 2015;21(2):121 – 31, 200. https://doi.org/10.19914/j.cjvi.2015.02.001.
20. Ye JK, Li KL, Xu DS, Wu WD, Zheng JS, Cao L, et al. Surveillance of adverse events following immunization in China, 2015. Chin J Vaccines Immun 2017;23(5):481 – 92,511. https://doi.org/10.19914/j.cjvi.2017.05.001.
21. Almenoff JS, LaCroix KK, Yuen NA, Fram D, DuMouchel W. Comparative performance of two quantitative safety signalling methods: implications for use in a pharmacovigilance department. Drug Saf 2006;29(10):875 – 87. https://doi.org/10.2165/00002018-200629100-00005.
22. Kubota K, Koide D, Hirai T. Comparison of data mining methodologies using Japanese spontaneous reports. Pharmacoepidemiol Drug Saf 2004;13(6):387 – 94. https://doi.org/10.1002/pds.964.
23. Chen RT, Rastogi SC, Mullen JR, Hayes SW, Cochi SL, Donlon JA, et al. The vaccine adverse event reporting system (VAERS). Vaccine 1994;12(6):542 – 50. https://doi.org/10.1016/0264-410x(94)90315-8.

SUPPLEMENTARY TABLE S1. Computation and application of each signal detection method.

Method	Computation	Advantages (*1*)	Limitations (*1*)
PRR (*2*)	$PRR = \frac{a/(a+b)}{c/(c+d)}$ $95\%CI = e^{ln(PRR)\pm1.96\times SE(lnPRR)}$ $= e^{ln(PRR)\pm1.96\times\sqrt{\frac{1}{a}-\frac{1}{a+b}+\frac{1}{c}-\frac{1}{c+d}}}$	1. Easily applicable; 2. Easily interpretable; 3. More sensitive as compared to Bayesian method*.	1. Can not be calculated for all vaccine-AEFI pairs, e.g. PRR can not be calculated if cell c is 0, 95% CI can not be calculated if cell a or c is 0; 2. Low specificity*.
ROR (*3*–*4*)	$ROR = \frac{a/b}{c/d} = \frac{ad}{bc}$ $95\%CI = e^{ln(ROR)\pm1.96\times SE(lnROR)}$ $= e^{ln(ROR)\pm1.96\times\sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}}$	1. Easily applicable; 2. Easily interpretable; 3. More sensitive as compared to Bayesian method*; 4. Different adjustment for covariates in logistic regression.	1. Can not be calculated for all vaccine-AEFI pairs, e.g. ROR can not be calculated if cell b and c are 0, 95% CI can not be calculated if cell a or b or c or d is 0; 2. Low specificity*.
BCPNN (*4*–*5*)**	$\log_2 a \times (a+b+c+d)/(a+b)\times(a+c)$	1. Always applicable; 2. More specific as compared to the frequentist method*.	1. Computation is complex; 2. Low sensitivity*.
MGPS (*6*)***	$\frac{a}{\frac{(a+b)\times(a+c)}{a+b+c+d}}$	1. Always applicable; 2. More specific as compared to frequentist method*.	1. Computation is complex; 2. Low sensitivity*.

* when commonly cited thresholds are used.

** Basic computation is listed here. More details regarding Bayesian shrinkage can be found in the paper (*4*–*5*).

*** Basic computation is listed here. More details regarding Bayesian shrinkage can be found in the paper (*6*).

REFERENCES

1. Karahoca A. Data mining applications in engineering and medicine. IntechOpen. London, United Kingdom. 2012.
2. Evans SJW, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. Pharmacoepidemiol Drug Saf 2001;10(6):483 – 6. https://doi.org/10.1002/pds.677.
3. Van Puijenbroek EP, Diemont WL, Van Grootheest K. Application of quantitative signal detection in the Dutch spontaneous reporting system for adverse drug reactions. Drug Saf 2003;26(5):293 – 301. https://doi.org/10.2165/00002018-200326050-00001.
4. Van Puijenbroek EP, Bate A, Leufkens HGM, Lindquist M, Orre R, Egberts ACG. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. Pharmacoepidemiol Drug Saf 2002;11(1):3 – 10. https://doi.org/10.1002/pds.668.
5. Bate A, Lindquist M, Edwards IR, Olsson S, Orre R, Lansner A, et al. A Bayesian neural network method for adverse drug reaction signal generation. Eur J Clin Pharmacol 1998;54(4):315 – 21. https://doi.org/10.1007/s002280050466.
6. Dumouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. Am Stat 1999;53(3):177 – 90. https://doi.org/10.1080/00031305.1999.10474456.