Analysis of Individual Differences in Vaccine Pharmacovigilance Using VAERS Data and MedDRA System Organ Classes: A Use Case Study With Trivalent Influenza Vaccine

Jingcheng Du[1*], Yi Cai[2*], Yong Chen[3], Yongqun He[4] and Cui Tao[1]

[1]School of Biomedical Informatics, The University of Texas, Houston, TX, USA. [2]Pieces Technology, Dallas, TX, USA. [3]Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA, USA. [4]Medical School, University of Michigan, Ann Arbor, MI, USA.

ABSTRACT: Personalized and precision vaccination requires consideration of an individual's sex and age. This article proposed systematic methods to study individual differences in adverse reactions following vaccination and chose trivalent influenza vaccine as a use case. Data were extracted from the Vaccine Adverse Event Reporting System from years 1990 to 2014. We first grouped symptoms into the Medical Dictionary for Regulatory Activities System Organ Classes (SOCs). We then applied zero-truncated Poisson regression and logistic regression to identify reporting differences among different individual groups over the SOCs. After that, we further studied detailed symptoms of 4 selected SOCs. In all, 19 of the 26 SOCs and 17 of the 434 symptoms under the 4 selected SOCs show significant reporting differences based on sex and/or age. In addition to detecting previously reported associations among sex, age group, and symptoms, our approach also enabled the detection of new associations.

KEYWORDS: Vaccine pharmacovigilance, trivalent influenza vaccine, VAERS, MedDRA, system organ classes

Introduction

Vaccination is one of the most effective methods to protect the public against various infectious diseases.[1] Although their benefits far overweigh risks and costs, vaccines are accompanied by specific adverse events (AEs). Although most of the adverse effects are very mild, rare but severe problems do happen, including serious allergic reactions, long-term seizures, coma or lowered consciousness, deafness or permanent brain damage, and even death.[2]

As required by the US Food and Drug Administration (FDA), vaccines are tested for safety before they are allowed to enter the market, and their performance is continuously evaluated over time to identify any risks.[3] An assessment of vaccine safety usually starts at the preapproval stage when the information about AEs is collected during clinical trials. Due to relatively small sample sizes, short follow-up periods, and homogeneous testing populations, rare AEs and AEs more prevalent in subpopulations (eg, different sex and age groups) are likely missed during clinical trials.[4]

Vaccine Adverse Event Reporting System (VAERS) was created in 1990 as a postmarketing surveillance system to accept reports of possible AEs following vaccinations in the United States.[5] Although there may not be a cause-and-effect relationship between vaccines and AEs reported in VAERS, it can be used to detect signals of potential safety problems with more rigorous methods.[6] Although VAERS can serve as a postlicensure surveillance system to assess vaccine safety, it has several limitations, including underreporting, lack of verification of reported signals, wide range of data quality, and absence of an unvaccinated control group.[7]

To address some of the above limitations, various statistical and data mining approaches have been developed to identify potential safety signals using VAERS data. These studies use methods such as empirical Bayesian,[8] proportional reporting ratio,[9,10] and Delphic approach[11] to detect the signal between AEs and vaccines such as those for influenza, rotavirus, typhoid fever, and anthrax. Most existing analyses focus on identifying safety signals and causal relationships of vaccine AEs. To improve vaccine safety, however, individual differences (especially age and sex) also need to be considered in the context of precision vaccination.

As the most common vaccine type reported in VAERS, trivalent influenza virus vaccine (FLU3) was chosen in this study as a use case due to its prevalence. Trivalent influenza virus vaccine is the traditional flu vaccine to protect people against 3 different flu viruses.[12] Our previous work proposed a novel method that combines terminology-based and statistical analyses to study the

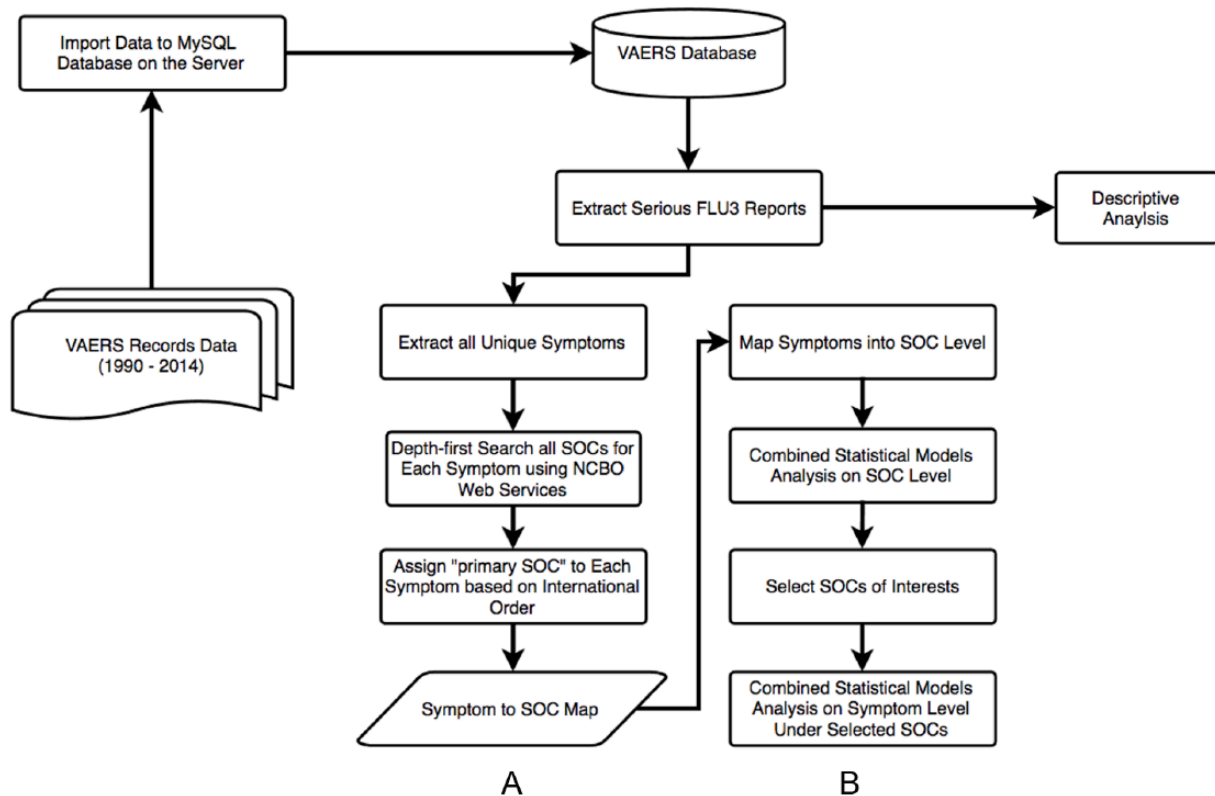* J.D. and Y.C. contributed equally to this work.

Figure 1. General project workflow. (A) Import VAERS records data to MySQL server and map symptoms to MedDRA SOC level. (B) Statistical analysis of VAERS data using the symptoms to MedDRA SOC map. FLU3 indicates trivalent influenza virus vaccine; MedDRA, Medical Dictionary for Regulatory Activities; NCBO, National Center for Biomedical Ontology; SOCs, System Organ Classes; VAERS, Vaccine Adverse Event Reporting System.

differences in postvaccination reactions among age groups and between sexes.[13,14] Our preliminary study using 1-year VAERS data (year 2011) indicated that 15 of the 26 System Organ Classes (SOCs) showed significant association with the age groups or between sexes.[14] In this study, we extended the preliminary work to conduct a more comprehensive analysis on the whole VAERS data set (1990-2014) and extended the methods to further analyze the individual difference at the symptoms level. We also did a detailed literature search to validate our results and highlighted new findings detected by our methods.

Materials and Methods

Method overview

In this study, instead of considering several single vaccine-AE associations, we took advantage of all the adverse symptom information. Figure 1 shows our workflow. In step A, we first imported all the VAERS reports (from years 1990 to 2014) to a MySQL relational database on our server and then extracted all the unique symptoms from FLU3 reports with serious reactions. As VAERS contains thousands of different types of symptoms, we sought to improve computational efficiency and avoid using sparse data by grouping symptoms to the Medical Dictionary for Regulatory Activities (MedDRA) SOC level using the mapping methods described in our previous studies.[13,14]

In step B, after grouping the symptoms into 26 SOCs, we identified differences among different age groups and between

sexes over the SOCs using zero-truncated Poisson regression and logistic regression. As the SOCs only cover high-level classification of the symptoms, we went one step further to study detailed symptoms for selected SOCs of interests.

Data source and descriptive analysis

The VAERS database currently contains more than 400 000 vaccine AE reports. Each VAERS report has been manually annotated with MedDRA Preferred Terms (PT) by domain experts.[15] We searched the VAERS for US reports submitted after FLU3 vaccination from 1990 to 2014 and extracted reports with serious reactions (ie, death, life-threatening illness, hospitalization, prolonged hospitalization, or permanent disability).[5]

We calculated certain descriptive statistics, including the total number of cases, the number of cases for each sex, and the number of cases for different age groups. We grouped the case reports into 5 age groups based on cut points (below 0.5 years, 0.5-17 years, 18-49 years, 50-64 years, and 65+ years) suggested by the Centers for Disease Control and Prevention (CDC).[16] As the number of reports from age group 0 to 0.5 years is very small (around 4% of all reports), we excluded these reports.

Mapping symptoms to MedDRA SOC level

According to the MedDRA user guide,[17] each PT must be linked to at least 1 SOC, and a PT can be linked to as many

SOCs as appropriate. This indicates that each PT could be grouped under more than 1 SOC. It results from the lack of a well-defined hierarchical structure in MedDRA.[18] Therefore, double counting might happen when we group 1 PT to multiple SOCs. To avoid this problem, we applied a mapping method that was described in our previous work[14] by leveraging National Center for Biomedical Ontology (NCBO) web services[19] and the Depth-First Search (DFS) algorithm,[20] we first used a recursive tree-traversing method to retrieve all the SOCs that are linked by a symptom (PT term). Then, we assigned a "primary SOC" to each term using an internationally agreed order SOC list (see Table 1). If 1 PT term belongs to several SOCs, we chose the SOC that is ranked the highest in this internationally agreed order SOC list to be its primary SOC. The order of the SOCs was based on the relative importance of each SOC in adverse drug reaction/AE reports.[21]

Statistical methods

To further investigate the grouped results, we conducted statistical analyses to explore the associations of SOCs with demographic factors, such as age and sex, and the correlations among individual SOCs.

Zero-truncated Poisson regression model. As all the subjects included in the data set have at least 1 SOC (ie, the number of reported SOCs are nonzero), we used zero-truncated Poisson regression instead of regular Poisson regression[22] to fit the data. First, an intercept-only regression model was fitted to estimate the average number of symptoms for all subjects. After that, 2 covariates of age category and sex were added to the model to estimate and compare the number of symptoms across different age and sex strata. The goodness of fit for the zero-truncated regression model was then evaluated by residual plots.

Logistic regression model. To explore the association of age and sex with the occurrence of individual SOCs, we dichotomized the original count number of SOCs to binary outcomes (0: SOC = 0; 1: SOC $\geqslant$ 0). For each SOC, we then conducted a logistic regression analysis,[23] with covariates being sex and age category in the model. The women in age group 1 (0.5-17 years) were considered the reference group. To avoid multiple testing problems, we used Bonferroni correction for hypothesis testing.[24] As we had 26 SOCs in total, we set ; if the *P* value of the Wald test for regression coefficient was less than .0019 (.05/26), we concluded that the covariate is statistically significantly associated with the occurrence of SOC. The Hosmer-Lemeshow test statistics were calculated to evaluate the goodness of fit for each logistic regression model.

For the association of age and sex with the occurrence of AEs at symptom level, we followed the steps of the SOC level and used Bonferroni correction for hypothesis testing for each selected SOC.

Spearman rank correlation coefficient. To measure the correlations among all SOCs, we calculated the correlation coefficients

Table 1. MedDRA terminology SOC list based on internationally agreed order.

SOC	ORDER
Infections and infestations	1
Neoplasms benign, malignant, and unspecified (including cysts and polyps)	2
Blood and lymphatic system disorders	3
Immune system disorders	4
Endocrine disorders	5
Metabolism and nutrition disorders	6
Psychiatric disorders	7
Nervous system disorders	8
Eye disorders	9
Ear and labyrinth disorders	10
Cardiac disorders	11
Vascular disorders	12
Respiratory, thoracic, and mediastinal disorders	13
Gastrointestinal disorders	14
Hepatobiliary disorders	15
Skin and subcutaneous tissue disorders	16
Musculoskeletal and connective tissue disorders	17
Renal and urinary disorders	18
Pregnancy, puerperium, and perinatal conditions	19
Reproductive system and breast disorders	20
Congenital, familial, and genetic disorders	21
General disorders and administration site conditions	22
Investigations	23
Injury, poisoning, and procedural complications	24
Surgical and medical procedures	25
Social circumstances	26

Abbreviations: MedDRA, Medical Dictionary for Regulatory Activities; SOC, System Organ Class.

of reported number of SOCs. As the number of SOCs is highly right-skewed and non-normally distributed, we used the rank-based Spearman ρ statistic to better estimate the correlation coefficient.[25] A correlation coefficient matrix was constructed to present the pairwise correlation among SOCs.

Results

Descriptive results

For the study period (1990-2014), we imported 440 663 VAERS reports to our local database. Among all 83 857 reports for FLU3 AEs, 7331 were serious. Of these serious reports, 3966 were for
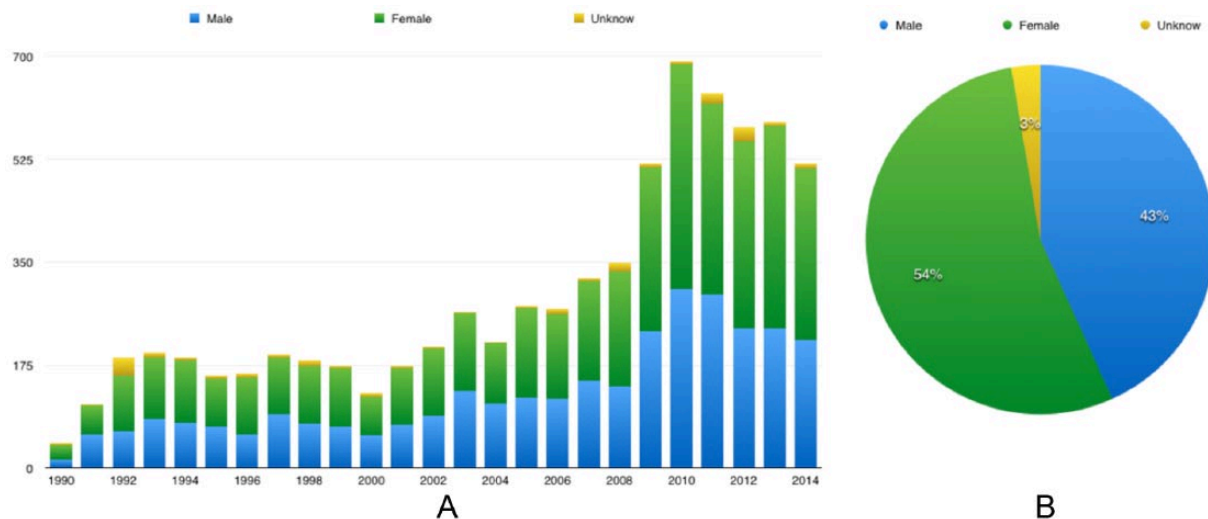
Figure 2. Sex distribution of serious trivalent influenza virus vaccine reports, from 1990 to 2014: (A) sex distribution for each year and (B) cumulative distribution per sex group between 1990 and 2014.
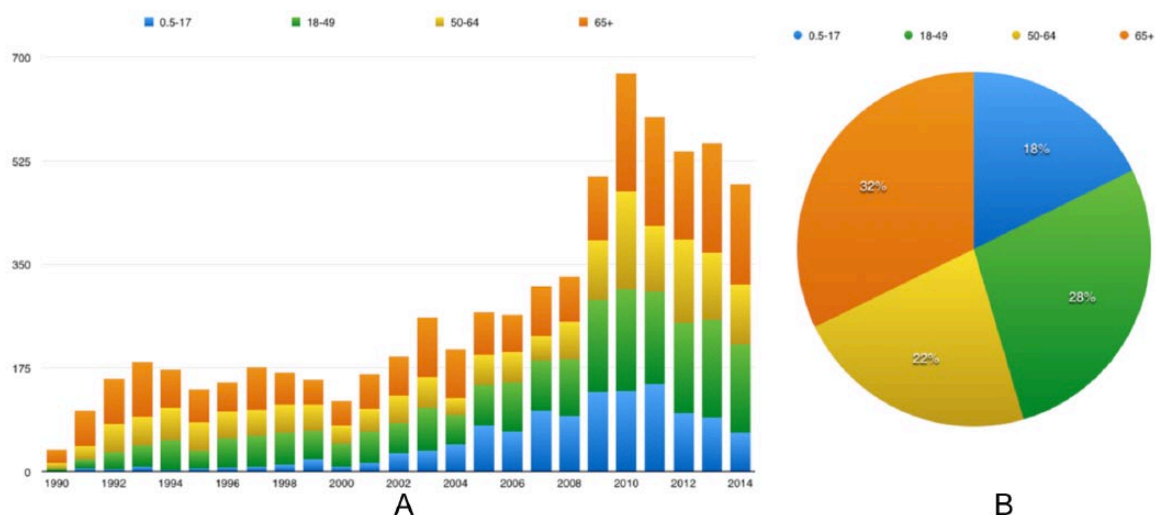

Figure 3. Age distribution of serious trivalent influenza virus vaccine reports, from 1990 to 2014: (A) age distribution for each year and (B) cumulative distribution per age group between 1990 and 2014.

female patients, 3169 were for male patients, and 196 were with unknown sex (Figure 2). There were 2230 patients older than 65 years (rounded to 32%), followed by age group 2 (18-49 years) which included 1922 patients (rounded to 28%) (Figure 3). These serious reports contained 72 076 symptom terms (not unique) and 3904 symptoms (unique). Among those unique symptoms, 189 cannot be mapped to any SOC, 2721 can be mapped to one of the SOCs, and 994 can be mapped to multiple SOCs. Three of the most frequent SOCs in these serious reports are Investigation, Nervous system disorders, and Infections and infestations. As MedDRA versions continue to change, some symptom terms cannot be assigned to the current MedDRA SOCs.

General results on SOC level

By applying the zero-truncated Poisson model with intercept only, the result showed that the average number of symptoms per subject from 1990 to 2014 was 9.818 (SE = 1.01). We found statistically significant differences in the number of symptoms among age groups and between sexes. Specifically, age group 4 (65+ years) showed the least number of symptoms, 10.1% lower than the reference age group (0.899 compared with intercept). Age group 3 (50-64 years) showed the highest number of symptoms, which was 5% higher than the reference age group (1.050 compared with intercept). For difference in sex, men experienced 6.2% higher number of symptoms than women, after adjusting for age (1.062 compared with intercept).

Logistic regression on individual SOC level showed differences in SOC among age groups and between sexes. There were a total of 13 SOCs that showed statistically significant differences among age groups or sexes. Six SOCs showed sex difference and 12 SOCs showed age group difference. Detailed logistic regression results can be seen in Table 2.

Table 2. Descriptive and logistic regression results on SOC level (1990-2014).

ORDER	SOC NAME	COUNTS	SYMPTOMS	SEX	AGE2	AGE3	AGE4
1	Infections and infestations	4477	244	1.373 (1.245–1.515)	N/A	1.697 (1.447–1.991)	1.348 (1.16–1.566)
2	Neoplasms benign, malignant and unspecified (including cysts and polyps)	362	82	1.464 (1.165–1.84)	3.374 (1.895–6.007)	5.83 (3.317–10.251)	4.577 (2.618–8.001)
3	Blood and lymphatic system disorders	969	87	N/A	N/A	N/A	N/A
4	Immune system disorders	2059	106	0.648 (0.572–0.733)	1.367 (1.14–1.638)	N/A	0.682 (0.564–0.824)
5	Endocrine disorders	287	40	N/A	N/A	2.234 (1.395–3.874)	2.777 (1.713–4.504)
6	Metabolism and nutrition disorders	1438	90	N/A	N/A	N/A	N/A
7	Psychiatric disorders	2588	167	N/A	0.626 (0.526–0.746)	0.588 (0.489–0.707)	0.642 (0.543–0.76)
8	Nervous system disorders	11442	350	N/A	N/A	N/A	0.694 (0.6–0.803)
9	Eye disorders	370	49	N/A	N/A	N/A	N/A
10	Ear and labyrinth disorders	117	24	N/A	N/A	N/A	N/A
11	Cardiac disorders	3287	122	N/A	1.605 (1.338–1.926)	1.663 (1.353–1.971)	1.89 (1.585–2.254)
12	Vascular disorders	1471	119	N/A	N/A	1.48 (1.176–1.861)	1.681 (1.585–2.254)
13	Respiratory, thoracic, and mediastinal disorders	2430	113	N/A	N/A	N/A	N/A
14	Gastrointestinal disorders	2890	138	N/A	N/A	N/A	N/A
15	Hepatobiliary disorders	102	24	N/A	N/A	N/A	N/A
16	Skin and subcutaneous tissue disorders	1903	77	0.801 (0.706–0.909)	N/A	N/A	0.607 (0.502–0.734)
17	Musculoskeletal and connective tissue disorders	3299	118	N/A	1.774 (1.484–2.122)	1.824 (1.517–2.193)	1.426 (1.195–.702)
18	Renal and urinary disorders	575	54	1.397 (1.143–1.708)	2.05 (1.368–3.072)	2.781 (1.863–4.15)	2.626 (1.783–3.868)
19	Pregnancy, puerperium, and perinatal conditions	96	32	N/A	N/A	N/A	N/A
20	Reproductive system and breast disorders	46	24	N/A	N/A	N/A	N/A
21	Congenital, familial, and genetic disorders	2	1	N/A	N/A	N/A	N/A
22	General disorders and administration site conditions	6876	99	N/A	N/A	N/A	N/A
23	Investigations	21020	1361	N/A	0.638 (0.548–0.743)	0.586 (0.5–0.686)	0.452 (0.389–0.524)
24	Injury, poisoning, and procedural complications	401	43	N/A	N/A	N/A	N/A
25	Surgical and medical procedures	1835	152	1.291 (1.129–1.475)	N/A	N/A	N/A
26	Social circumstances	383	28	N/A	N/A	N/A	N/A

Abbreviation: SOC, System Organ Class.
"Counts" is the total number of occurrences of SOC. "Symptom" is the number of unique symptoms belonging to SOC. "Sex" and "AGE2, 3, 4" are the odds ratios (OR) of the risks of having adverse events of male patients and age groups 2, 3, and 4 compared with the reference group; "N/A" means the P value of the cell is larger than or equal to .0019 (.05/26).

To study the correlation between SOCs, we calculated the pairwise correlation matrix of the SOCs by applying the Spearman method (see Figure 4). The color and size of spots represent the strength of correlation among SOCs. The detailed correlation coefficient matrix is given in Supplementary Table E. As shown in the correlation coefficient matrix, there was a moderate correlation (correlation coefficient $\geqslant 0.3$, according to Cohen's (1988) convention) between SOC 25 (Surgical and medical procedures) and SOC 23 (Investigations). Other pairs showing moderate correlation include SOC 25 with SOC 8 (Nervous system disorders), SOC 23 with SOC 8, and SOC 13 (Respiratory, thoracic and mediastinal disorders) with SOC 11 (Cardiac disorders).

Analyses at the symptom level for 4 selected SOCs

To further explore individual SOCs, we focused on 4 SOCs of interests and applied logistic regression model on the symptom level within these 4 SOCs: Blood and lymphatic system disorders (SOC 3), Immune system disorders (SOC 4), Cardiac disorders (SOC 11), and Vascular diseases (SOC 12). Blood-system–related SOC 3, SOC 11, and SOC 12 were selected due to the fact that cardiac and vascular diseases are reported as leading causes of death globally,[26] which together resulted in 17.3 million deaths in 2013.[27] SOC 4 was chosen due to its relatively larger occurrence and significant difference in sexes detected by previous step (0.648 in men compared with intercept).

The total occurrence of Immune system disorders was 2059, and there were 106 unique symptoms under this group. The most frequent symptoms under this SOC were "Urticaria," "Hypersensitivity," and "Asthma." Due to the nonreported occurrence of one or more subpopulation groups for some symptoms, 33 models were fit for 33 unique symptoms. As we set $\alpha = 0.05$, if the P value of the Wald test for regression coefficient is less than .0015 (.05/33), we concluded that the covariate is statistically significantly associated with the occurrence of this symptom. Four symptoms were significantly different among age groups or between sexes. Specifically, 1 symptom showed differences in different sex and 3 showed differences among different age groups. Detailed results of the symptoms showing difference at sex or age can be seen in Supplementary Table A.

For Blood and lymphatic system disorders, 87 unique symptoms were found. The total occurrence for those symptoms was 969. "Leukocytosis," "Anemia," and "Thrombocytopenia" were the most frequent symptoms for this SOC. In all, 18 logistic regression models were fit for 18 unique symptoms. If the P value regression coefficient is less than .0028 (.05/18), we concluded that the covariate is statistically significantly associated with the occurrence of this symptom. Only 2 symptoms were found with significantly different prevalence among different age groups. Detailed results can be seen in Supplementary Table B.

For Cardiac disorders, there were 122 unique symptoms found, and the total occurrence of these symptoms was 3287. "Dyspnea," "Chest pain," and "Tachycardia" were the most frequent symptoms for Cardiac disorders. In all, 31 logistic regression models identified 8 symptoms with significantly different prevalence among different sex or different age groups ($P <$ .0016 [.05/31]). Detailed results can be seen in Supplementary Table C.

The total occurrence of Vascular disorders was 1471, and 119 unique symptoms were found. "Hypertension," "Hypotension," and "Pallor" were the most frequent symptoms for Vascular Disorders. In all, 26 logistic regression models identified 3 symptoms with significantly different prevalence among different age groups (P < .0019 [.05/26]). Detailed results can be seen in Supplementary Table D.

Evaluation

Findings which are consistent with previous studies

Many of our results are consistent or partially consistent with previous studies. This can be an indication that our methods can be used to identify personal differences among age groups and between sexes. In terms of sex differences, our method detected that women have a higher probability of reporting immune system disorders after vaccination, which is consistent with previous studies that women have higher responses to various types of vaccination.[28] In addition, our finding indicates that men have a higher probability of reporting renal and urinary disorders after vaccination. Previous studies have shown that there are sex differences in urinary disorders, such as urinary tract infection,[29] interstitial cystitis,[30] and kidney diseases.[31] These sex differences may also cause different reactions of the urinary system toward FLU3. In terms of age differences, our results indicate that age group 4 (65+ years) reported the least number of symptoms among all age groups. Previous studies have shown that aging may decrease the immune response to vaccination,[32] which may also be relevant to adverse reactions. Age group 3 (50-64 years) showed a higher number of reported symptoms than age group 2 (18-49 years), which is consistent with the CDC National Health Interview Survey showing that persons above 45 years experienced a larger number of diseases than persons below 44 eyars.[33]

On the symptom level, our study indicated that men had lower rates of reporting "Dyspnea" than women after vaccination. This is consistent with previous study indicating that women were experiencing more dyspnea in the symptoms of acute coronary syndromes.[34] In terms of age difference, our study also indicated that the adult population (aged $\geqslant 18$ years) showed more responses on some symptoms under Cardiac and Vascular disorders, such as "Dyspnea," "Chest pain," and "Chest discomfort," than the younger population. This is consistent with previous findings reporting that aging can cause changes in heart and blood vessels, which could increase the probability of acquiring cardiovascular diseases.[32] In addition, a previous
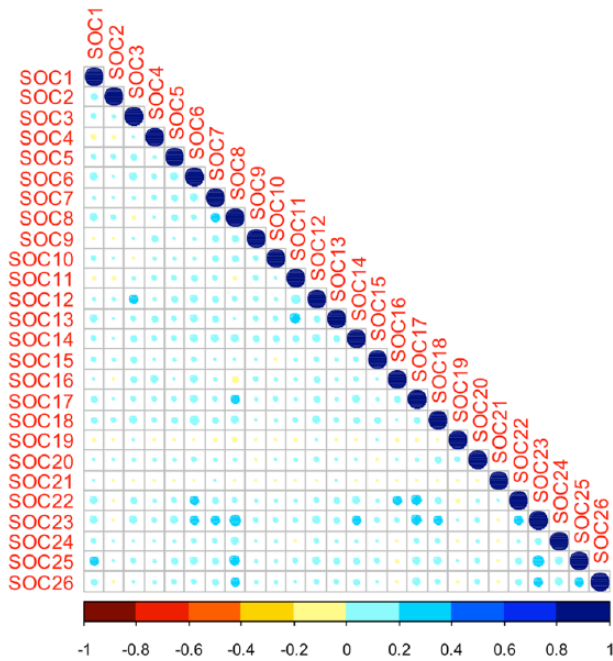
Figure 4. The pairwise correlation matrix of System Organ Classes (SOCs) determined by the Spearman method.

study showing that aging causes decreased immune responses to vaccination[35] supports our findings that many symptoms under immune system disorders show lower probability in the age groups older than 17 years.

New findings detected by our methods

In addition to the findings that can be confirmed by previous studies, our methods have also identified some novel information for trivalent influenza AEs. Here, we report the results that are statistically significant but cannot be found, to the best of our knowledge, from literature review through PubMed.

On the SOC level, for sex differences, we have found that men have lower probabilities (19.9% lower) of acquiring skin and subcutaneous tissue disorders after influenza vaccination than women. To the best of our knowledge, there are no similar studies on the relationship between influenza vaccinations and skin disorders. For age differences, we also found that aging people have more chances of acquiring renal and urinary disorders after influenza vaccinations.

On the symptom level, we have identified some new associations between symptoms with influenza vaccination. Our results showed that men have higher chances of reporting "Demyelinating polyneuropathy" and "Myocarditis." We also found that the younger population (younger than 17 years) shows higher probabilities of having "Pallor" and "Contusion" than the elder population.

These new findings can potentially be used to provide directions for future studies to uncover factors contributing to sex and age differences in susceptibility to influenza infection, which may be helpful in adjusting the dose and ingredients of the influenza vaccines for people of different age and sex groups.

As there are few studies on the analysis of SOC level, we believe most of our results on the correlations of SOCs have not been reported by other studies. For example, there are correlations between SOC 25 (Surgical and medical procedures) and SOC 23 (Investigations), SOC 13 (Respiratory, thoracic and mediastinal disorders) and SOC 11 (Cardiac disorders), and so on. These newly found correlations might be helpful to study the complications of FLU3 AEs.

Discussion and Conclusions

Although vaccines are held to the highest standard of safety, they can cause adverse reactions that can sometimes be serious. Vaccine Adverse Event Reporting System is one of the biggest and most comprehensive postmarketing surveillance systems in the United States to monitor and accept vaccination AE reports. Using trivalent influenza vaccine as the use case, this article proposed systematic methods that combined multiple statistical models to detect potential sex and age AE differences at SOCs and symptom level for specific vaccines. The following evaluation on the results showed that our methods not only provide support for the findings from previous studies but also identify some novel associations that have not been reported by other studies.

Our results can provide direction for future studies to uncover factors contributing to sex and age differences in susceptibility to influenza infection, which may be helpful in adjusting the dose and ingredients of the influenza vaccines for people of different age and sex groups. Our methods could also be used to provide similar direction for other types of vaccines, such as human papillomavirus, measles-mumps-rubella, and typhoid fever vaccines. We can also apply these methods to other data sets, such as the FDA Adverse Event Reporting System (FAERS), or data from social networks, such as Twitter or Facebook, to study the AE individual differences.

A significant limitation of our study is that the differences in rates of reporting among different subpopulations could also lead to identified differences in AEs for different subpopulations. Further rigorous clinical studies need to be done to verify our findings. Another limitation comes from the mapping process when we mapped the symptoms to MedDRA SOCs. A symptom could belong to several SOCs; however, we only assigned 1 "Primary SOC" to each symptom term based on an international agreed order. That international agreed order could lead to biased mapping and thus lead to biased findings. Due to the deficiencies of MedDRA, such as domain completeness and lack of well-defined hierarchy, we may consider using Ontology of Adverse Events (OAE)[18] in the future to group the adverse symptoms more accurately.

Author Contributions

JD collected that data, wrote the initial draft, and revised the subsequent draft. JD and YC developed the method, preformed the evaluation, and conducted analysis of the results. YC guided the statistical analysis design and modeling. YH provided

expertise in vaccine and vaccine terminologies. CT provided institutional support, contributed to research design, and guided the data analysis. All authors read and approved the final manuscript.

REFERENCES

1. Immunization Safety Review Committee. *Immunization Safety Review: Vaccines and Autism*. Washington, DC: National Academies Press; 2004.
2. Centers for Disease Control and Prevention. Possible side effects from vaccines. http://www.cdc.gov/vaccines/vac-gen/side-effects.htm. Published 2012.
3. Clayton EW, Rusch E, Ford A, Stratton K. *Adverse Effects of Vaccines: Evidence and Causality*. Washington, DC: National Academies Press; 2012.
4. Zhang Y, Wu P, Luo Y, Tao C. Identification of sex-associated network patterns in Vaccine-Adverse Event Association Network in VAERS. *J Biomed Semantics*. 2015;6:33.
5. Centers for Disease Control and Prevention, US Food and Drug Administration. *VAERS Data User Guide*. 2015. https://vaers.hhs.gov/data/READMEJanuary2015.pdf
6. Vellozzi C, Burwen DR, Dobardzic A, Ball R, Walton K, Haber P. Safety of trivalent inactivated influenza vaccines in adults: background for pandemic influenza vaccine safety monitoring. *Vaccine*. 2009;27:2114–2120.
7. Banks D, Woo EJ, Burwen DR, Perucci P, Braun MM, Ball R. Comparing data mining methods on the VAERS database. *Pharmacoepidemiol Drug Saf*. 2005;14:601–609.
8. 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;19:4627–4634.
9. Zhou W, Pool V, DeStefano F, Iskander JK, Haber P, Chen RT. A potential signal of Bell's palsy after parenteral inactivated influenza vaccines: reports to the Vaccine Adverse Event Reporting System (VAERS)—United States, 1991–2001. *Pharmacoepidemiol Drug Saf*. 2004;13:505–510.
10. Begier EM, Burwen DR, Haber P, Ball R; Vaccine Adverse Event Reporting System Working Group. Postmarketing safety surveillance for typhoid fever vaccines from the Vaccine Adverse Event Reporting System, July 1990 through June 2002. *Clin Infect Dis*. 2004;38:771–779.
11. Sever JL, Brenner AI, Gale AD, et al. Safety of anthrax vaccine: an expanded review and evaluation of adverse events reported to the Vaccine Adverse Event Reporting System (VAERS). *Pharmacoepidemiol Drug Saf*. 2004;13:825–840.
12. Centers for Disease Control and Prevention. Seasonal influenza: flu basics. http://www.cdc.gov/flu/about/disease/index.htm. Published 2014.
13. Tao C, Du J, Cai Y, Chen Y. Trivalent influenza vaccine adverse event analysis based on MedDRA system organ classes using VAERS data. *Stud Health Technol Inform*. 2014;216:1076.
14. Du J, Cai Y, Chen Y, Tao C. Trivalent influenza vaccine adverse symptoms analysis based on MedDRA terminology using VAERS data in 2011. *J Biomed Semantics*. 2016;7:13.
15. US Food and Drug Administration. What is a serious adverse event? http://www.fda.gov/safety/medwatch/howtoreport/ucm053087.htm. Published 2014.
16. Centers for Disease Control and Prevention. Estimates from the Behavioral Risk Factor Surveillance System (BRFSS), National Immunization Survey (NIS), and the National 2009 H1N1 Flu Survey (NHFS). http://www.cdc.gov/flu/fluvaxview/trends/age-groups.htm. Published 2012.
17. Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Introductory guide. http://www.who.int/medical_devices/innovation/MedDRAintroguide_version14_0_March2011.pdf. Published 2015.
18. He Y, Sarntivijai S, Lin Y, et al. OAE: the ontology of adverse events. *J Biomed Semantics*. 2014;5:29.
19. The National Center for Biomedical Ontology. NCBO Web services and the development of semantic applications. http://www.bioontology.org/wiki/index.php/NCBO_Web_Services_and_the_Development_of_Semantic_Applications. Published 2012.
20. Wikipedia. Depth-first search. https://en.wikipedia.org/wiki/Depth-first_search.
21. Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Introductory guide MedDRA Version 14.0. http://www.who.int/medical_devices/innovation/MedDRAintroguide_version14_0_March2011.pdf
22. Stata Data Analysis Examples. Zero-truncated Poisson regression. http://www.ats.ucla.edu/stat/stata/dae/ztp.htm.
23. Hosmer DW Jr, Lemeshow S. *Applied Logistic Regression*. Hoboken, NJ: John Wiley & Sons; 2004.
24. Bonferroni Correction. From mathworld–A wolfram web resource. http://mathworld.wolfram.com/BonferroniCorrection.html.
25. Zar JH. Significance testing of the spearman rank correlation coefficient. http://www.jstor.org/stable/2284441?seq=1#page_scan_tab_contents. Accessed April 6, 2015.
26. Mendis S, Puska P, Norrving B. *Global Atlas on Cardiovascular Disease Prevention and Control*. Geneva, Switzerland: World Health Organization; 2011.
27. Naghavi M, Wang H, Lozano R, et al. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2015;385:117–171.
28. Giefing-Kröll C, Berger P, Lepperdinger G, Grubeck-Loebenstein B. How sex and age affect immune responses, susceptibility to infections, and response to vaccination. *Aging Cell*. 2015;14:309–321.
29. Sammon JD, Sharma P, Rahbar H, et al. Predictors of admission in patients presenting to the emergency department with urinary tract infection. *World J Urol*. 2014;32:813–819.
30. Roberts RO, Bergstralh EJ, Bass SE, Lightner DJ, Lieber MM, Jacobsen SJ. Incidence of physician-diagnosed interstitial cystitis in Olmsted County: a community-based study. *BJU Int*. 2003;91:181–185.
31. Iseki K. Gender differences in chronic kidney disease. *Kidney Int*. 2008;74:415–417.
32. Bill PM. Aging changes in the heart and blood vessels. https://medlineplus.gov/ency/article/004006.htm
33. Blackwell DL, Lucas JW, Clarke TC. Summary health statistics for U.S. adults: national health interview survey, 2012. *Vital Health Stat* 10 2014;260:1–161.
34. Patel H, Rosengren A, Ekman I. Symptoms in acute coronary syndromes: does sex make a difference? *Am Heart J*. 2004;148:27–33.
35. Chit A, Roiz J, Briquet B, Greenberg DP. Expected cost effectiveness of high-dose trivalent influenza vaccine in US seniors. *Vaccine*. 2015;33:734–741.