**DRUG SALES DATA ANALYSIS FOR OUTBREAK DETECTION OF** **Influenza-like Illness IN CHINA**

by

Yufei Zhang

BA, Applied Technology College of Soochow University, China, 2014

Submitted to the Faculty of

Health Policy and Management

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Public Health

University of Pittsburgh

2018

**ABSTRACT**

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This essay is submitted

by

Yufei Zhang

on

February 20, 2018

and approved by

Essay Advisor:

Gerald Mark Barron, MPH \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Associate Professor and MPH Director

Department of Health Policy and Management

Deputy Director of the Center for Public Health Practice

Graduate School of Public Health

University of Pittsburgh

Essay Reader:

Inmaculada Hernandez, PharmD, PhD \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Assistant Professor

Department of Pharmacy and Therapeutics

School of Pharmacy

University of Pittsburgh

Copyright © by Yufei Zhang

2018

Gerald M. Barron, MPH

**DRUG SALES DATA ANALYSIS FOR OUTBREAK DETECTION OF Influenza-like Illness IN CHINA**

Yufei Zhang, MPH

University of Pittsburgh, 2018

Data analysis is increasingly being used in syndromic surveillance system. This paper studies the drug sales analysis because of its public health relevance to outbreak ILI disease detection. Influenza-like illness (ILI) currently has a particularly high public health significance in China because the 2009 swine flu and the 2013 H7N9 outbreaks caused hundreds of deaths. Drug sales analysis can use real-time drug sales data to provide early detection of potential flu outbreaks. One example of syndromic surveillance system is the use of OTC medication sales in estimating annual influenza epidemics.  Moreover, there is no established syndromic surveillance system that leverages drug sales data in China.

This essay aims to investigate the correlation between influenza surveillance data and drug sales data to help public health department make a timely outbreak detection in China. Specifically, this study collected and compared OTC drug sales data from a drug retailing pharmacy chain in Danyang, China with ILI trends surveillance system data from local CDC department. The collected ILI drugs categories includes medications for the treatment or control of cough, cold, and other respiratory ailments both in children and adults. The incidence rate of ILI or influenza was provided by the local CDC. To compare weekly lagged sales of selected drugs with ILI% data from 2015 to 2017, we computed Pearson’s correlation coefficients. The correlation coefficient between the ILI % and drug sales data from 2016 Oct to 2017 Mar was 0.848 (p<0.05) which showed that the drug sales data can be used as a source to predict outbreaks. Moreover, we observed that peaks for drug sales appear 1-2 weeks before the local ILI% peak.

In conclusion, we observed a strong correlation between sales of influenza drugs and the number of influenza-like illness visits.

**Public health significance/relevance**. These results suggest that the analysis of OTC medication sales can potentially predict disease outbreaks, which will allow for advanced preparation by the local health authorities. With increasing concern over influenza-like disease outbreaks, public health is being required to have early outbreak detection systems in China.

TABLE OF CONTENTS

[PREFACE ix](#_Toc509420867)

[1.0 Introduction 1](#_Toc509420868)

[2.0 Methods and mAterials 3](#_Toc509420869)

[2.1 STUDY POPULATION 4](#_Toc509420870)

[2.2 OTC DRUG SALES 4](#_Toc509420871)

[2.3 ILI incidence 7](#_Toc509420872)

[2.4 statistical analyses 8](#_Toc509420873)

[3.0 results 10](#_Toc509420874)

[3.1 Data analysis of January 2015 - August 2017 10](#_Toc509420875)

[3.1 Data analysis from October 2016 to March 2017. 15](#_Toc509420876)

[3.2 Data analysis from 2015 to March 2016. 21](#_Toc509420877)

[4.0 discussion 26](#_Toc509420878)

[bibliographY 30](#_Toc509420879)

List of tables

[Table 1. Influenza drug sales data among different drug groups in 2015, 2016 and 2017 6](#_Toc509422180)

[Table 2.Influenza cases distribution among different age groups from 2015 to 2017 8](#_Toc509422181)

[Table 3. The maximum positive CCF value of cross correlation analysis between drugs and ILI%, 2015-2017 11](#_Toc509422182)

[Table 4.Pearson correlation test of OTC drugs with ILI% in 2015-2017 14](#_Toc509422183)

[Table 5.The maximum positive CCF value of cross correlation analysis between drugs and ILI%, 2016 Oct-2017 Mar 17](#_Toc509422184)

[Table 6.Pearson correlation test of OTC drugs with ILI% in 2016 Oct-2017 Mar 18](#_Toc509422185)

[Table 7.The maximum positive CCF value of cross correlation analysis between drugs and ILI%, Apr-Sept, 2016 19](#_Toc509422186)

[Table 8.Pearson correlation test of OTC drugs with ILI%, Apr-Sept, 2016 20](#_Toc509422187)

[Table 9. The maximum positive CCF value of cross correlation analysis between drugs and ILI%, 2015 Oct-2016 Mar 23](#_Toc509422188)

[Table 10. Pearson correlation test of OTC drugs with ILI% in 2015 Oct-2016 Mar 24](#_Toc509422189)

[Table 11. The maximum positive CCF value of cross correlation analysis between drugs and ILI%, Apr-Sept, 2015 25](#_Toc509422190)

List of figures

[Figure 1.Comparison between ILI% and sum of OTC drug sales in Danyang, China, 2015-2017 10](#_Toc509422191)

[Figure 2.Comparison between ILI% and sum of adult drug sales in Danyang, China, 2015-2017 10](#_Toc509422192)

[Figure 3.Comparison between ILI% and sum of children drug sales in Danyang, China, 2015-2017 11](#_Toc509422193)

[Figure 4.Cross correlation analysis between sum, adult, children and ILI%, 2015-2017 13](#_Toc509422194)

[Figure 5. Comparison between ILI% and sum of OTC drug sales in Danyang, China, 2016/2017 influenza season. 15](#_Toc509422195)

[Figure 6. Comparison between ILI% and sum of adult drug sales in Danyang, China, 2016/2017 influenza season. 16](#_Toc509422196)

[Figure 7. Comparison between ILI% and sum of children drug sales in Danyang, China, 2016/2017 influenza season. 16](#_Toc509422197)

[Figure 8. Comparison between ILI% and sum of children drug sales in Danyang, China, 2016 non-influenza season. 20](#_Toc509422198)

[Figure 9. Comparison between ILI% and sum of OTC drug sales in Danyang, China, 2015/2016 influenza season. 21](#_Toc509422199)

[Figure 10. Comparison between ILI% and sum of adult drug sales in Danyang, China, 2015/2016 influenza season. 22](#_Toc509422200)

[Figure 11. Comparison between ILI% and sum of children drug sales in Danyang, China, 2015/2016 influenza season. 22](#_Toc509422201)

# PREFACE

I would like to thank the local CDC department and local pharmacies for their support. Manager of pharmacies, Mingfang Kuang, provided the drug sales data. Her secretory, and Li Ni, answered numerous questions and clarified a lot of data questions. I appreciate the data and information support from local CDC surveillance department in Danyang, China. Thank you to my advisor, Professor Barron, and my second reader, Inmaculada Hernandez for their love and support. Also, I would like to thank my friend, Li, for giving me help and support during data analysis. Finally, thanks to the Department of Health Policy and Management in Graduate School of Public Health in University of Pittsburgh. I hope to make you all proud.

# Introduction

Influenza-like illness (ILI) leads to a set of flu-like symptoms including fever, chills, cough, body aches, nausea. [1] {, 2005 #2}ILI has a great public health significance because the influenza virus has high potential to cause a national-wide influenza outbreak and even global pandemic. The Chinese government has paid more attention to influenza-like illness control since the 2009 swine flu and 2013 H7N9 outbreaks. During the 2009 flu season, thousands of cases of swine flu were reported in the media in China. [2] Moreover, in the past years, China also has had deadly annual outbreaks of H7N9. This H7N9 virus is a subtype of influenza viruses which have been detected in birds and was first found in March 2013. In 2017, the Chinese CDC reported 759 cases of H7N9 and 281 deaths. [2]

Due to the high infectious potential of the influenza virus, it is important for the public health authorities to take measures to detect, control and prevent these outbreaks. From the public health perspective, we need to think how can Chinese government, at national or local level, to take measures to control and prevent flu outbreaks and Influenza-like illness. In doing so, it is particularly important to evaluate the use of analytic methods and technology that can leverage big data in predicting these outbreaks. Also, we need to considerate how can we predict this outbreak by big data analytics or technology. However, in China, there is no comprehensive and advanced syndromic surveillance system using network data,. In predicting influenza outbreaks, syndromic surveillance can be particularly useful. Syndromic surveillance is defined as a dynamic analysis of real-time medical data to detect or anticipate disease outbreaks.[3,4] A real-time surveillance system can predict emerging infectious diseases, bioterrorism, and pandemics.[3] Though syndromic surveillance has been used to investigate potential cases, its utility is increasingly being explored by public health officials.[4] For example, once an influenza outbreak begins to affect the population, some people may call in sick for work or school firstly. Then, some others may visit drug stores, to purchase over the counter (OTC) medicine. Finally, others who cannot control their symptom by drugs may go to hospital.

# Methods and mAterials

This paper utilizes qualitative and quantitative analysis to examine data from a city-wide area, Danyang, in China. The original data in this essay was obtained from local CDC and pharmacies.

Access to ILI% data and to pharmacy sales was granted by the local CDC and by pharmacies in Danyang upon request, respectively. Because these data are not public data, it is hard to get. Therefore, we tried to communicate with external people. When we collect the data from external organizations, drug stores and local CDC in China, we meet some difficulties to access these data. But we overcome these difficulties by persistent efforts and communication skills. By repeatedly asking local CDC department that whether they are willing to provide the ILI% data to us to do the research, we have been authorized to access to the local CDC data system. Then, the data were collected and sorted by author. Also, by connecting with managers in pharmacies in Danyang again and again, we are authorized to access the sales data of different drugs and put them in to the computer to classify them.

Specifically, local CDC provided the ILI cases and percentage of ILI. OTC sales data were provided by local pharmacy chain HTH. The percentage of ILI was estimated on the basis of the number of ILI cases detected by the hospitals or clinics and sorted and provided by local CDC. The drug sales database consists of most weekly over-the-counter (OTC) drug sales from 2015 to 2017 and were provided by the manager of Heart to Heart (HTH) drugstores in Danyang. Further details on the description of the data sets and the categorization of drugs may be found below.

## STUDY POPULATION

The study site is located at Danyang. It is close to Shanghai and approximately 200 km from Shanghai. Danyang is a [c](https://en.wikipedia.org/wiki/County-level_city)ounty-level city under the administration of Jiangsu province located in eastern coastal of China. It has a population of about 90 million as estimated in the year 2016. [5] Danyang has a total area of 1,059 km2. [5] As a city within the Shanghai economic sphere of influence, its economy was 16th in a 2010 ranking of Chinese top county-level cities and it is considered as the top 100 rich city in China.[6] Therefore, Danyang has attracted economic power, enough population and great public health influence in eastern China.

Additionally, the climate in Danyang is mild and warm and the annual temperature is about 18 ℃. The temperature in summer is much higher than winter. Its coldest month on average is January with a temperature about 3 ℃. Therefore, Danyang has a typical temperate climate with an influenza season from October to March. Because of the influenza season is from October to March, this essay will focus on annual winter influenza season from 2015 to 2017, in the following analysis.

## OTC DRUG SALES

This study aims to compare drug sales data from a pharmacy chain, HTH, in Danyang with ILI% trends surveillance system data from local CDC department.

The ILI drug sales data is collected from HTH pharmacy originally named as Heart to Heart pharmacy was founded in Danyang, in 2002. HTH Pharmacy is currently the largest pharmacy [chain](https://en.wikipedia.org/wiki/Chain_store) in Danyang by number of locations which is over 30. It sells prescription drugs and a wide kind of general goods, including over-the-counter drugs, medical device, nutritious products and beauty products. HTH pharmacy took the largest market share in Danyang and its market share is more than 50% in Danyang. Therefore, this study mainly collected the OTC drug sales data from more than 30 HTH pharmacies in Danyang. The drug sale data from HTH are reliable and meaningful for the following syndromic surveillance analysis because of the big database and market share.

This study monitored the purchase of OTC medications by computer database system of HTH drugstores. We collected weekly sales data on the number ILI OTC drugs by pharmacies' database records, including adult and pediatric fever/cold, cough medications and so on in liquid or tablet form. This study not only collect the pharmacy sales data but also classify them to different groups, such us children drug group and adult drug group.

Specifically, we selected 19 different kinds of drugs, including 3 kinds of pediatric ILI drugs and 16 kinds of adult drugs in total. (See Table 1). As is shown in the following table, the drugs are marked as number 1-16 are adult drugs and number 17- 19 are pediatric drugs. Compared with pediatric influenza or cold drugs, sales volume of adult cold drugs is apparently higher. The sales volume of adult cold drugs is roughly 10 times than the sales volume of children cold drug in total. In order to make it more easier to understand, we translated all the names of the 19 kinds of drug were translated from Chinese to English. ( See Table 1)

Table . Influenza drug sales data among different drug groups in 2015, 2016 and 2017

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| No. | Drug Name | | English Name | total sales in 2015 | total sales in 2016 | total sales in 2017 (AUGUST 12) |
| 1 | | 维C银翘片 | Vc Yinqiao Tablets | 3145 | 3583 | 1161 |
| 2 | | 氨咖黄敏胶囊 | Paracetamol,Caffein,Atificial Cow-bezoar and chlorphenamine Maleate capsules | 55502 | 57425 | 30362 |
| 3 | | 风寒感冒颗粒 | Fenghan Ganmao Granules | 1172 | 1570 | 936 |
| 4 | | 感冒清热颗粒 | Ganmao Qingre Granules | 848 | 1068 | 807 |
| 5 | | 复方氨酚烷胺胶囊 | Compound Paracetamol and Amantadine Hydrochloride Capsules | 24709 | 26326 | 15265 |
| 6 | | 复方氨酚烷胺片 | Compound Paracetamol and Amantadine Hydrochloride Tablets | 7389 | 10321 | 6584 |
| 7 | | 四季感冒片 | Four Seasons cold tablets | 2086 | 2233 | 1664 |
| 8 | | 小柴胡颗粒 | Xiaochaihu Granule | 1835 | 2002 | 1463 |
| 9 | | 桑姜感冒片 | Sangjiang cold tablets | 1026 | 458 | 1285 |
| 10 | | 氨酚伪麻美芬片 | Compound Pseudoephedrin Hydrochlorid Tablets | 1338 | 1365 | 731 |
| 11 | | 蒲地蓝消炎口服液 | Po blue to oral anti-inflammatory | 7710 | 9684 | 6692 |
| 12 | | 蓝芩口服液 | Lanqin oral | 1695 | 5431 | 4156 |
| 13 | | （复方）板蓝根颗粒 | Compound Indigowoad Root | 13349 | 14406 | 7567 |
| 14 | | 复方金银花颗粒 | Compound lotion of honeysuckle | 748 | 3307 | 1812 |
| 15 | | 连花清瘟胶囊 | Lianhuaqingwen Capsule | 249 | 318 | 243 |
| 16 | | 盐酸吗啉胍片 | Moroxydine Hydrochloride Tablets | 294 | 365 | 197 |
|  | | **adult drug sum** | **adult drug sum** | **123095** | **139862** | **80925** |
| 17 | | 小儿伪麻美芬滴剂 | Antuss (Pediatric Pseudoephedrine Hydrochloride and Dextromethorphan Hydrobromide Drops) | 195 | 317 | 142 |
| 18 | | 小儿氨酚黄那敏颗粒 | Pediatric Paracetamol, Atificial Cow-bezoar and Chlorphenamine Maleate Granules. | 8678 | 11468 | 6854 |
| 19 | | 小儿氨酚烷胺颗粒 | Pediatric Paracetamol and Amantadine Hydrochloride | 3511 | 2814 | 1309 |
|  | | **children drug sum** | **children drug sum** | **12384** | **14796** | **8305** |

**Table 1 Continued**

## ILI incidence

This essay collected the influenza surveillance from January, 2015 to August, 2017(136 weeks in total). Local CDC collects the ILI cases data by the reports from local hospitals and clinics. Physicians in primary hospitals and clinics diagnosed ILI cases and reported them by age group 0–4, 5–14, 15–24, 25–59, ≥60 years[7]. Designated hospital staff reported the number of ILI cases by age group and total number of consultations daily on line.[8] 2015-2017 ILI cases population who are 0-4, 5-14,15-24, 25-59, and older than 60 year old accounted for 54.86%, 26.44%, 3.12%, 11.72% and 3.87% respectively. (See Table 2) As shown in the table 2 below, from 2015 to 2017, age distribution of ILI cases can be interpreted as 0-4> 5-14> 25-24> 60 ~> 15-24, indicating that children and the elderly are the most susceptible population and should be put more attention. Also, children and the elderly is the key part in infectious disease prevention and control chain.

According to the local CDC surveillance data, ILI% numbers in Danyang from January, 2015 to August, 2017 were about 2.42% in 2015, 3.51% in 2016, 3.52% as of August 13, 2017. As can be seen from the Table 2 below, the proportion of ILI cases in 2016 (53-104 week) and 2017 (105-136 week) was significantly higher than in 2015 (1-52 week).

Table .Influenza cases distribution among different age groups from 2015 to 2017

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Age 0-4 | 5-14 | 15-24 | 25-59 | >60 |
| Year 2015 | 51.16% | 24.04% | 3.24% | 15.47% | 6.10% |
| 2016 | 57.39% | 25.68% | 2.72% | 11.05% | 3.15% |
| 2017 | 56.88% | 25.93% | 3.85% | 10.85% | 2.49% |
| avg. | 54.86% | 26.44% | 3.12% | 11.72% | 3.87% |

ILI was defined as the presence of fever (≥38°C) and cough and/or sore throat in the absence of any other confirmed diagnosis [9]. ILI% was defined as the number of influenza-like illness visits for cases among the total number of outpatient and emergency department visits. The incidence rate of ILI considered the ILI cases proportion of the census population who seek care in the sites when got ILI symptom. [7] Incidence rate of ILI/influenza (week) =  Number of ILI/influenza cases (weak )/total number total number of consultations (week) . [10]

## statistical analyses

This paper utilizes statistical analysis which completed by time series function in SPSS 15.0. In order to test the theory that sales of drugs can detect a seasonal influenza epidemic early, the data from an influenza surveillance system were compared to medication sales and a predictive model was developed. [10] Weekly ILI incidence rates from the local CDC’s surveillance system were compared to sales of nineteen kinds of medications including nasal decongestants, medicines for sore throat (MST), antitussives and anti-inflammatory drugs and so on to determine the correlation of drug sales with the surveillance system and their predictive power in this study. [10]

Cross-correlation analysis by SPSS 15.0 is a standard method for analyzing the correlation between two independent variables and their time sequence. The amount of influenza drug sales and ILI% with changes over time are two different sources of surveillance information. There is a time sequence between these two variables therefore they can be applied into cross-correlation analysis.

# results

## Data analysis of January 2015 - August 2017

The following Figure 1, 2 and 3 carried out by SPSS 15.0 is the result of data analysis from 2015 to 2017.

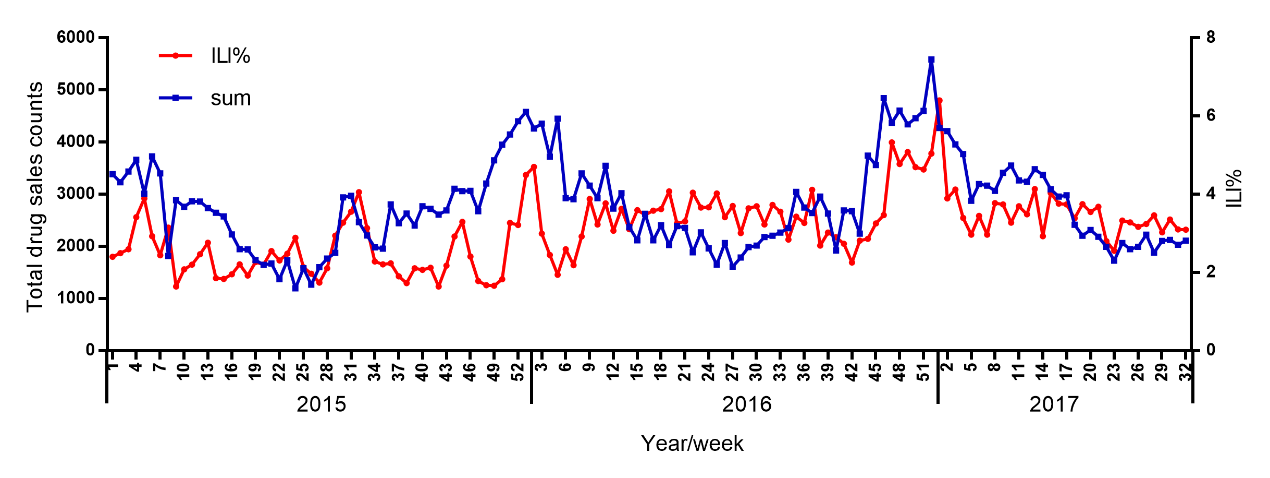
****

Figure .Comparison between ILI% and sum of OTC drug sales in Danyang, China, 2015-2017

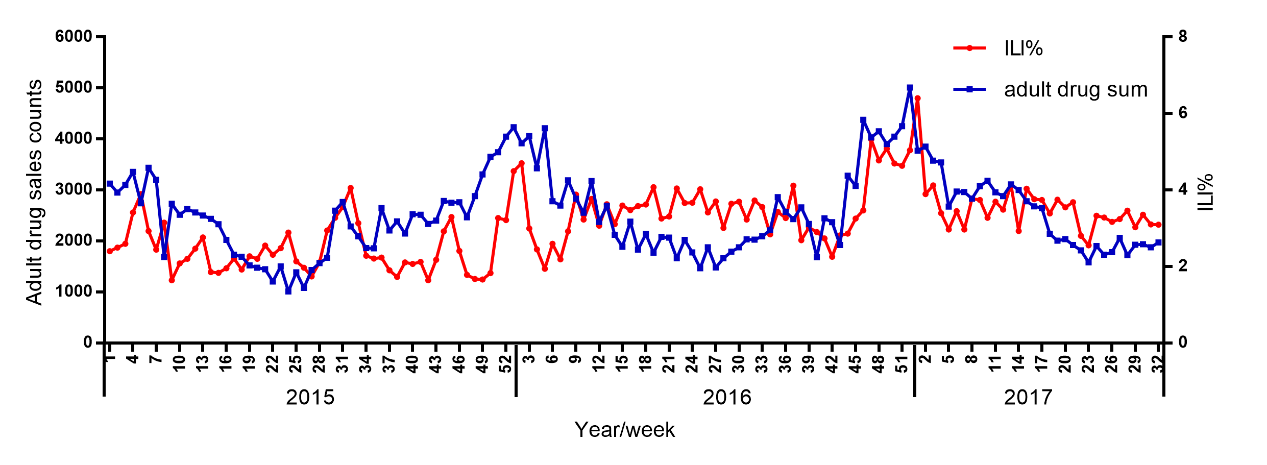
****

Figure .Comparison between ILI% and sum of adult drug sales in Danyang, China, 2015-2017

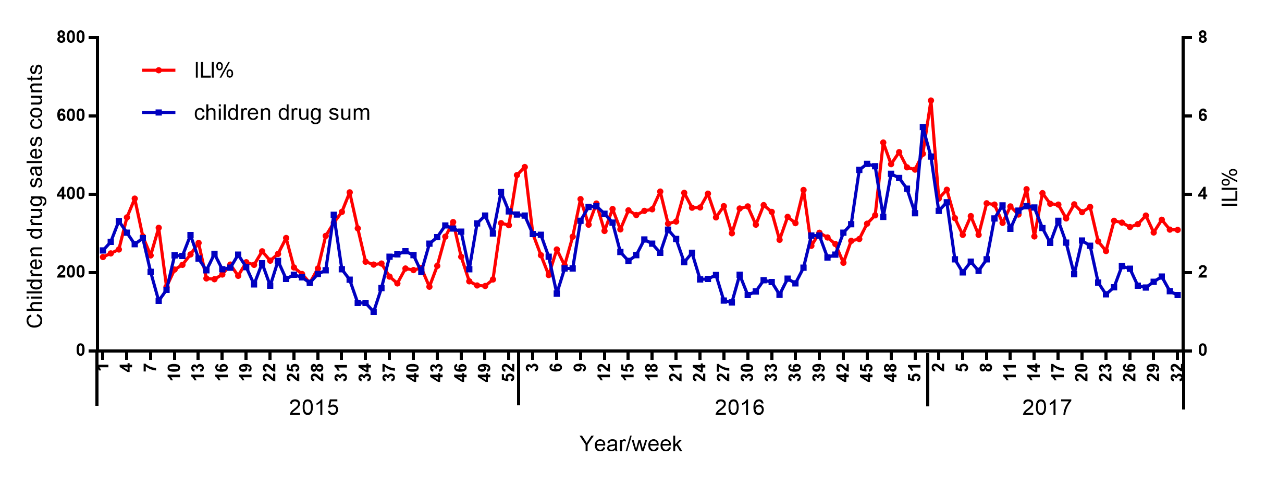
****

Figure .Comparison between ILI% and sum of children drug sales in Danyang, China, 2015-2017

First, cross-correlation is a [measure of similarity](https://en.wikipedia.org/wiki/Similarity_measure) of two series and is useful for describe the relationship between two time series (y and x), the series y may be related to past lags of the x-series. [11]. Therefore, CCF is used to determine the time delay between two signals or identify lags of the x-variable that might be useful predictors of y [11]. The cross correlation analysis is carried out by SPSS, and the results for coefficient of Cross Correlation Function (CCF) are as follows (See Table 3):

Table . The maximum positive CCF value of cross correlation analysis between drugs and ILI%, 2015-2017

|  |  |  |
| --- | --- | --- |
| **Category** | **CCF (SE)** | **Lag (week)** |
| adult drug sum | 0.443 (0.086) | 1 |
| children drug sum | 0.474 (0.086) | 1 |
| sum | 0.458 (0.086) | 1 |

Abbreviations: CCF= Coefficient of Cross Correlation Function

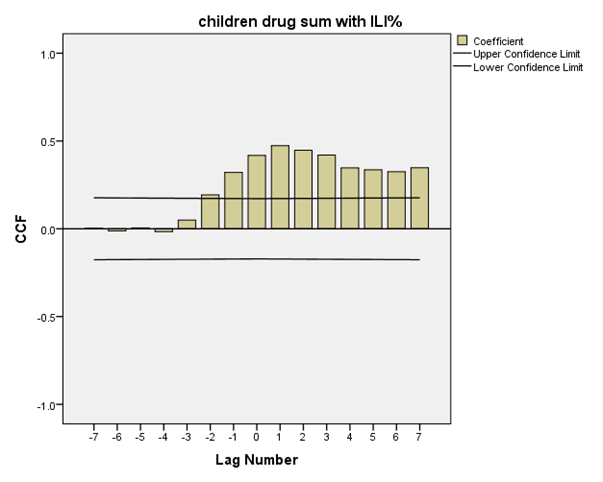
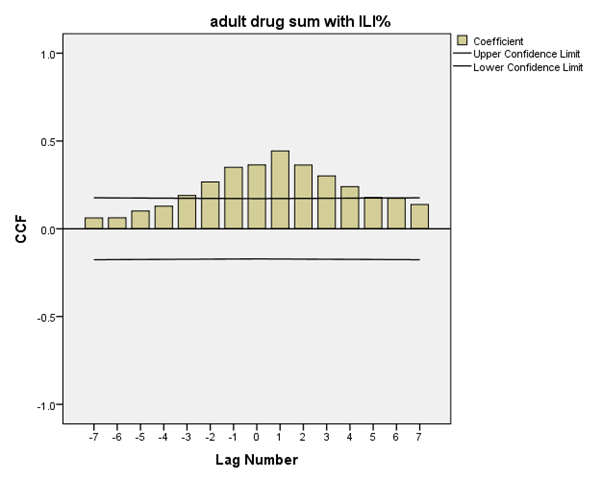
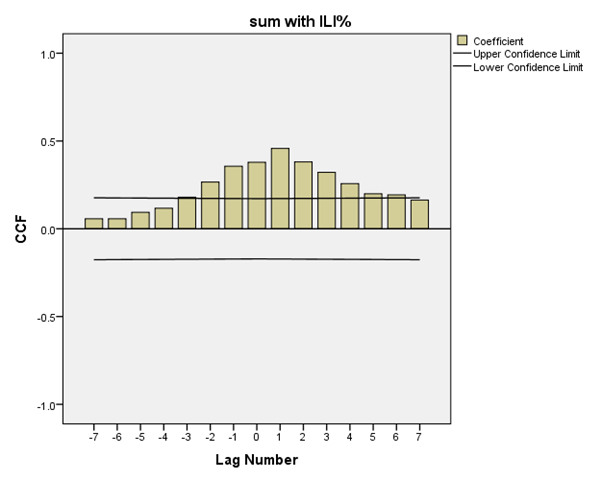


Figure .Cross correlation analysis between sum, adult, children and ILI%, 2015-2017

During the reporting period, January 2015 - August 2017, the total cases of reported influenza-like illness was 19,924. Citywide trends in the ratio of ILI to OTC drug sales from 2015 to 2017 followed a consistent cyclical pattern with relative highs in the winter and relative lows in the summer (See Figure 1, 2 and 3). These significant increases in winter were related to the epidemic influenza winter and spring seasons. This local ILI% trend was strongly correlated with OTC drug sales trends in the previous 1 week during this period (Cross Correlation Function (CCF ) = 0.443) (See Table 3 and Figure 4)

According to the data from the national and provincial surveillance system, 2016 and 2017 are in the influenza epidemic period among the whole country. (See Figure 1, 2 and 3) Since Danyang People's Hospital built a system based on HIS respiratory symptoms monitoring system to monitor the influenza case until 2016, quality of surveillance data since 2016 is higher than before. This essay will pay more attention to analyze the data collected since 2016.

The above table shows that from 3 years' overall data, whether the adult drugs or the children drugs both predict the influenza 1 week ahead of time (lag=1), and if lagging the sales volume 1 week later, the adult drugs (CCF=0.443, SE=0.086) or the children drugs (CCF=0.474, SE=0.086) will be the most relevant to ILI%. ( See Table 3 and Figure 4) On the basis, the data of adult and children drugs and total OTC data were lagged 1 week later and the Pearson correlation test was calculated. The results were as follows in the Table 4:

Table .Pearson correlation test of OTC drugs with ILI% in 2015-2017

|  |  |  |
| --- | --- | --- |
| **Category** | **ILI%** | |
| **r** | **p** |
| **adult drug sum** | **0.445** | **<0.01** |
| 1. 维C银翘片 | 0.222 | 0.01 |
| 2. 氨咖黄敏胶囊 | 0.274 | 0.001 |
| 3. 风寒感冒颗粒 | 0.453 | <0.01 |
| 4. 感冒清热颗粒 | 0.476 | <0.01 |
| 5. 复方氨酚烷胺胶囊 | 0.309 | <0.01 |
| 6. 复方氨酚烷胺片 | 0.545 | <0.01 |
| 7. 四季感冒片 | 0.285 | 0.001 |
| 8. 小柴胡颗粒 | 0.409 | <0.01 |
| 9. 桑姜感冒片 | 0.107 | 0.216 |
| 10. 氨酚伪麻美芬片 | 0.278 | 0.001 |
| 11. 蒲地蓝消炎口服液 | 0.627 | <0.01 |
| 12. 蓝芩口服液 | 0.498 | <0.01 |
| 13.（复方）板蓝根颗粒 | 0.355 | <0.01 |
| 14. 复方金银花颗粒 | 0.409 | <0.01 |
| 15. 连花清瘟胶囊 | 0.366 | <0.01 |
| 16. 盐酸吗啉胍片 | 0.323 | <0.01 |
| **children drug sum** | **0.478** | **<0.01** |
| 17. 小儿伪麻美芬滴剂 | 0.359 | <0.01 |
| 18. 小儿氨酚黄那敏颗粒 | 0.526 | <0.01 |
| 19. 小儿氨酚烷胺颗粒 | 0.107 | 0.219 |
| **sum** | **0.460** | **<0.01** |

The above table 4 shows that the correlation between children drug sum (r=0.478, p<0.01) and ILI% is higher than that of adult drug sum (r=0.445, p<0.01) at the year-round level, which indicates that the accuracy and reliability of children cold medicine is higher than that of adults in the whole year. Among the children's drugs, the relationship between No.18 (Pediatric Paracetamol, Atificial Cow-bezoar and Chlorphenamine Maleate Granules) and ILI% was the greatest (r=0.526, p<0.01). Among the adult drugs, No.11 (Po blue to oral anti-inflammatory) and ILI% shows the maximum correlation (r=0.627, p<0.01). There was no correlation between the No.9 (Sangjiang cold tablets) (r=0.107, p=216), No.19 (Pediatric Paracetamol and Amantadine Hydrochloride) (r=0.107, p=219) and the ILI%.

## Data analysis from October 2016 to March 2017

The following Figure 5, 6 and 7 is the result of data analysis of the flu season from October 2016 to March 2017.

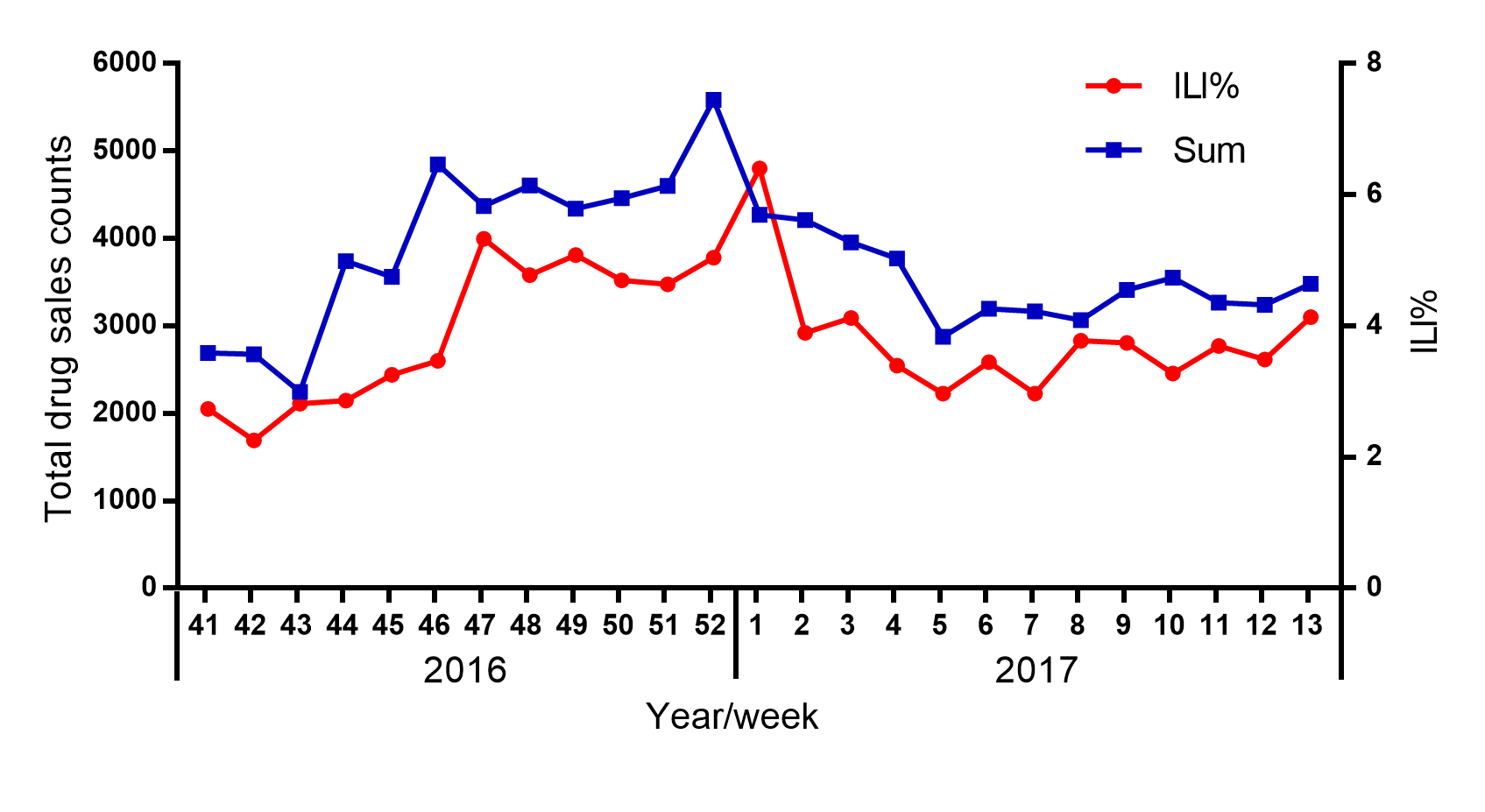


Figure . Comparison between ILI% and sum of OTC drug sales in Danyang, China, 2016/2017 influenza season.

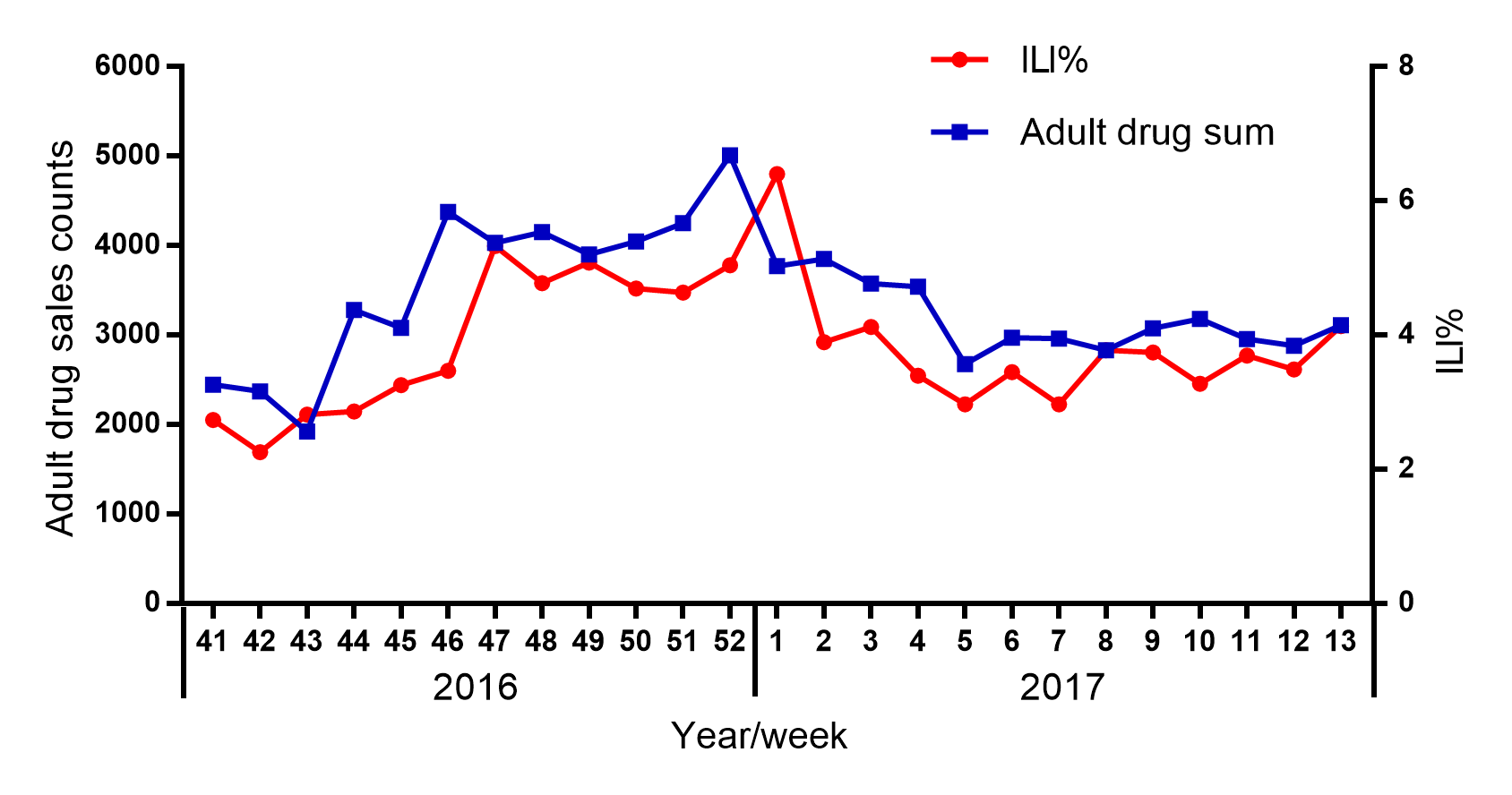
****

Figure . Comparison between ILI% and sum of adult drug sales in Danyang, China, 2016/2017 influenza season.

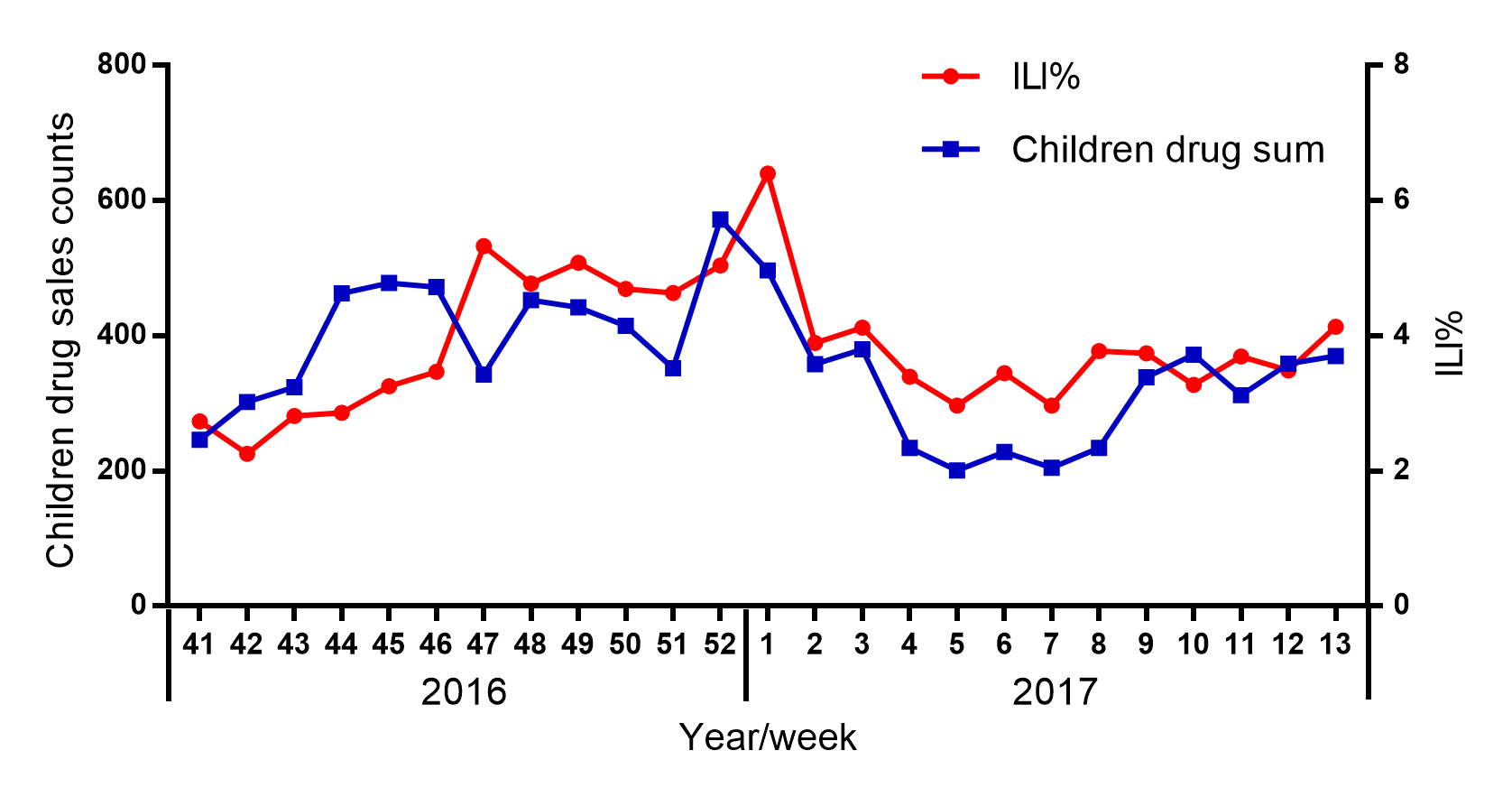
****

Figure . Comparison between ILI% and sum of children drug sales in Danyang, China, 2016/2017 influenza season.

Cross correlation analysis is carried out, and the results are as follows in Table 5:

Table .The maximum positive CCF value of cross correlation analysis between drugs and ILI%, 2016 Oct-2017 Mar

|  |  |  |
| --- | --- | --- |
| **Category** | **CCF (SE)** | **Lag (week)** |
| adult drug sum | 0.844 (0.204) | 1 |
| children drug sum | 0.720 (0.200) | 0 |
| sum | 0.848 (0.204) | 1 |

The CCF value (CCF=0.848, SE=0.204) in the above Table 5 is significantly higher than the total CCF value (CCF=0.458, SE=0.086) of 2015-2017, indicating that the sales volume of drugs can better predict the occurrence of influenza during the high stage of influenza. Among them, the adult drugs sum (CCF=0.844, SE=0.204) and the total sales amount (CCF=0.848, SE=0.204) still had the maximum correlation with ILI% after 1 week (lag=1), while the children drug sum (CCF=0.720, SE=0.200) showed the greatest correlation with ILI% in the case of not moving back (lag=0). First, this may be related to the selection of the minimum unit value. Also, it may be caused by the ILI population since the children accounted for more than 50% of the whole population. Finally, if the day is a lagging unit, there should be a lag number of days. Therefore, the result showed that the CCF for children is lower and without weekly lag.

However, the drug sales peak for both children and adult are appear around the 52th week which are one week before the peak of ILI% (See Figure 6 and 7). Therefore, it also shows that in the high incidence of influenza season, both adult and children drugs are good at predicting the occurrence of influenza but the adult drug sales data is more statistically predictable.( 1 week >0 week)

On the basis, the data of adult drugs and total OTC data were lagged 1 week later, the children drugs were not lagged, and the Pearson correlation test was calculated in the following Table 6. The results are listed in the table below (See Table 6):

Table .Pearson correlation test of OTC drugs with ILI% in 2016 Oct-2017 Mar

|  |  |  |
| --- | --- | --- |
| **Category** | **ILI%** | |
| **r** | **p** |
| **adult drug sum** | **0.867** | **<0.01** |
| 1. 维C银翘片 | 0.682 | <0.01 |
| 2. 氨咖黄敏胶囊 | 0.736 | <0.01 |
| 3. 风寒感冒颗粒 | 0.723 | <0.01 |
| 4. 感冒清热颗粒 | 0.725 | <0.01 |
| 5. 复方氨酚烷胺胶囊 | 0.787 | <0.01 |
| 6. 复方氨酚烷胺片 | 0.818 | <0.01 |
| 7.四季感冒片 | 0.682 | <0.01 |
| 8. 小柴胡颗粒 | 0.788 | <0.01 |
| 9. 桑姜感冒片 | 0.303 | 0.141 |
| 10. 氨酚伪麻美芬片 | 0.623 | 0.001 |
| 11. 蒲地蓝消炎口服液 | 0.830 | <0.01 |
| 12. 蓝芩口服液 | 0.567 | 0.003 |
| 13.（复方）板蓝根颗粒 | 0.828 | <0.01 |
| 14. 复方金银花颗粒 | 0.112 | 0.595 |
| 15. 连花清瘟胶囊 | 0.464 | 0.020 |
| 16. 盐酸吗啉胍片 | 0.722 | <0.01 |
| **children drug sum** | **0.538** | **0.006** |
| 17. 小儿伪麻美芬滴剂 | 0.593 | 0.002 |
| 18. 小儿氨酚黄那敏颗粒 | 0.493 | 0.012 |
| 19. 小儿氨酚烷胺颗粒 | 0.540 | 0.005 |
| **sum** | **0.871** | **<0.01** |

During the flu season from October 2016 to March 2017, children drug sum (r=0.538, p=0.006) was less correlated with ILI% than adult drug sum (r=0.867, p<0.01), indicating that children cold medicine is less accurate and less reliable in predicting ILI than adult cold medicine. Among the children drugs, all three drugs were associated with ILI% (p=0.002; p=0.012; p=0.005) and the relationship between NO.17小儿伪麻美芬滴剂 (Antuss-Pediatric Pseudoephedrine Hydrochloride and Dextromethorphan Hydrobromide Drops) and ILI% was the greatest (r=0.593, p=0.002). Among the adult drugs, there was no correlation between the No. 9 桑姜感冒片 (Sangjiang cold tablets) (r=0.303, p=0.141), No.14复方金银花颗粒(Compound lotion of honeysuckle) (r=0.112, p=0.595) and ILI%, and the rest were all correlated with ILI%. Among them, the maximum correlation is between the No.11蒲地蓝消炎口服液(Po blue to oral anti-inflammatory) (r=0.830, p<0.01) and ILI%.

The following is data analysis of the non-influenza season, Apr-Sept, 2016.

Cross correlation analysis is carried out, and the results are as follows (See Table 7):

Table .The maximum positive CCF value of cross correlation analysis between drugs and ILI%, Apr-Sept, 2016

|  |  |  |
| --- | --- | --- |
| **Category** | **CCF (SE)** | **Lag (week)** |
| adult drug sum | 0.103 (0.218) | 6 |
| children drug sum | 0.412 (0.209) | 4 |
| sum | 0.144 (0.218) | 6 |

It is found that adult drugs (CCF=0.103, SE=0.218) can predict influenza 6 weeks early, but the CCF value is too low and less reliable. Children drugs (CCF=0.412, SE=0.209) were predicted 4 weeks in advance and the CCF value was reliable. This shows that the sales data are much statistically reliable in influenza season since CCFs are lower in non-influenza season than in influenza season. From the statistic perspective, the season data from October to March is much more meaningful to predict the ILI outbreaks.

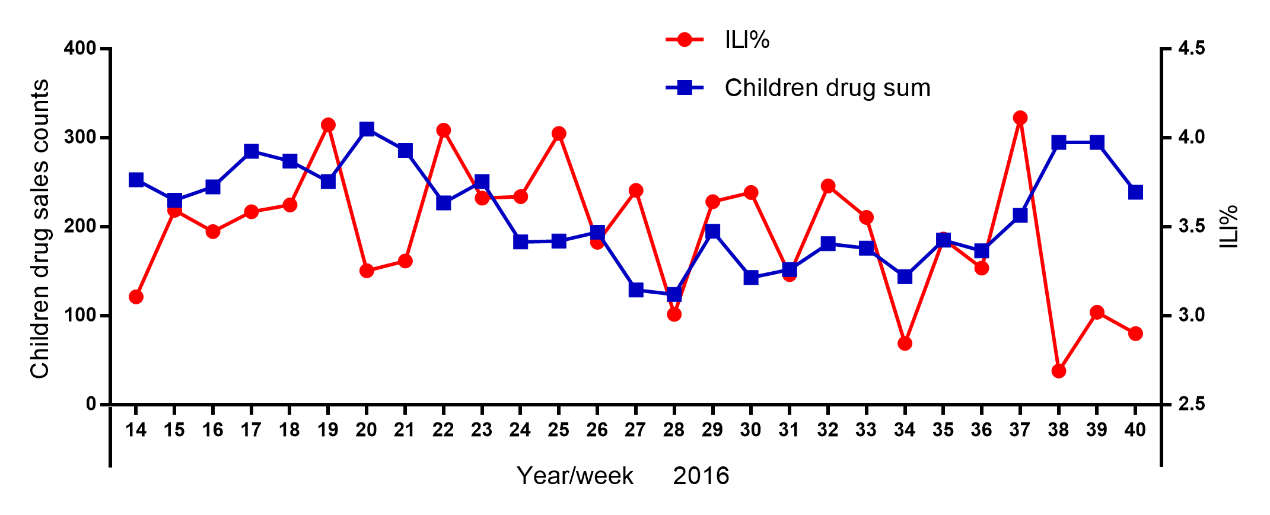


Figure . Comparison between ILI% and sum of children drug sales in Danyang, China, 2016 non-influenza season.

On the basis, the data of children drugs were lagged 4 weeks later (See Table 7), and the Pearson correlation test was carried out. The results were as follows (See Table 8):

Table .Pearson correlation test of OTC drugs with ILI%, Apr-Sept, 2016

|  |  |  |
| --- | --- | --- |
| **Category** | **ILI%** | |
| **r** | **p** |
| **children drug sum** | **0.143** | **0.514** |
| 17. 小儿伪麻美芬滴剂 | 0.285 | 0.187 |
| 18. 小儿氨酚黄那敏颗粒 | 0.469 | 0.024 |
| 19. 小儿氨酚烷胺颗粒 | 0.315 | 0.143 |

Overall, the correlation between children drugs and ILI% was low (r=0.143, p>0.05), indicating that children drugs could not predict the outbreak of influenza during the non-influenza period. However, we found a correlation between sales of No.18小儿氨酚黄那敏颗粒 (Pediatric Paracetamol, Atificial Cow-bezoar and Chlorphenamine Maleate Granules) and ILI% (r=0.469, p=0.024). There was no significant correlation between the other two drugs and ILI% (p=0.187, p=0.143).

## Data analysis from 2015 to March 2016.

The following Figure 9, 10 and 11 is the results of data analysis of the flu season from October 2015 to March 2016.

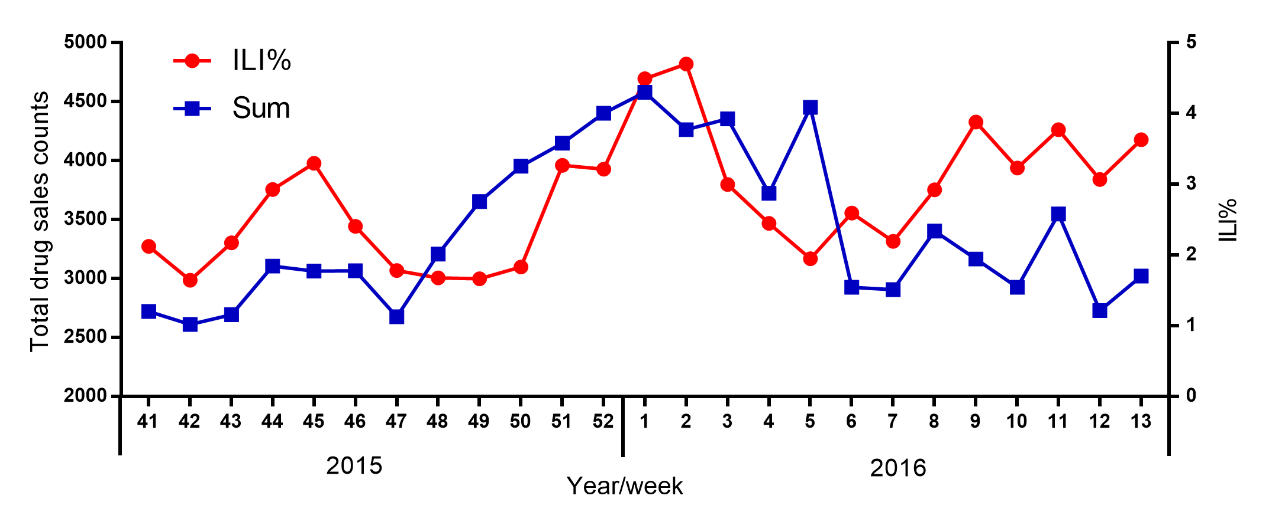


Figure . Comparison between ILI% and sum of OTC drug sales in Danyang, China, 2015/2016 influenza season.

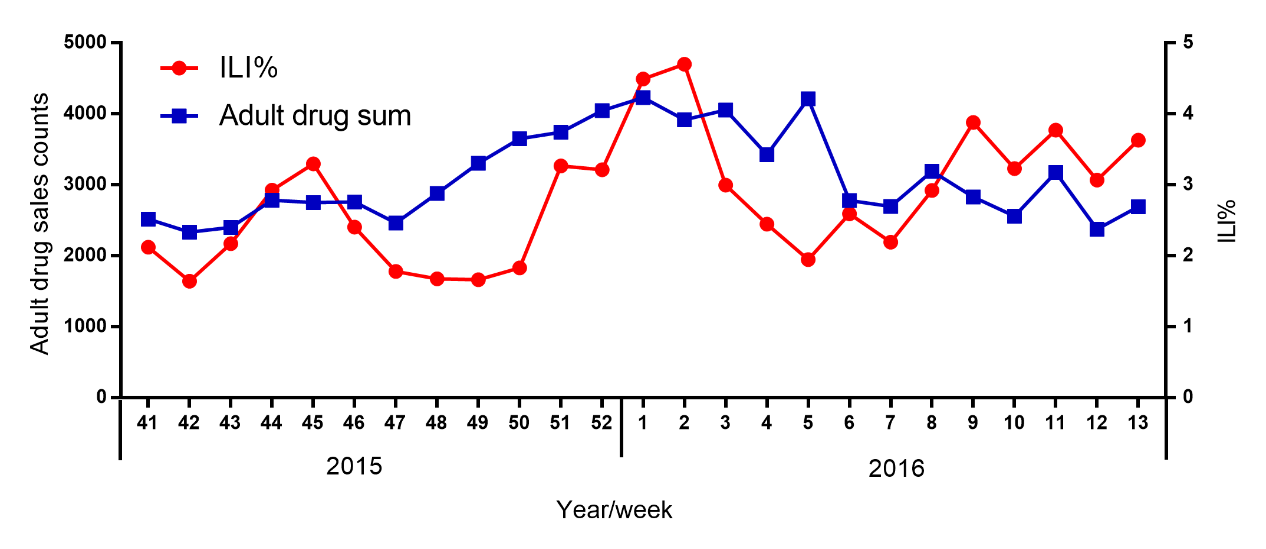
****

Figure . Comparison between ILI% and sum of adult drug sales in Danyang, China, 2015/2016 influenza season.

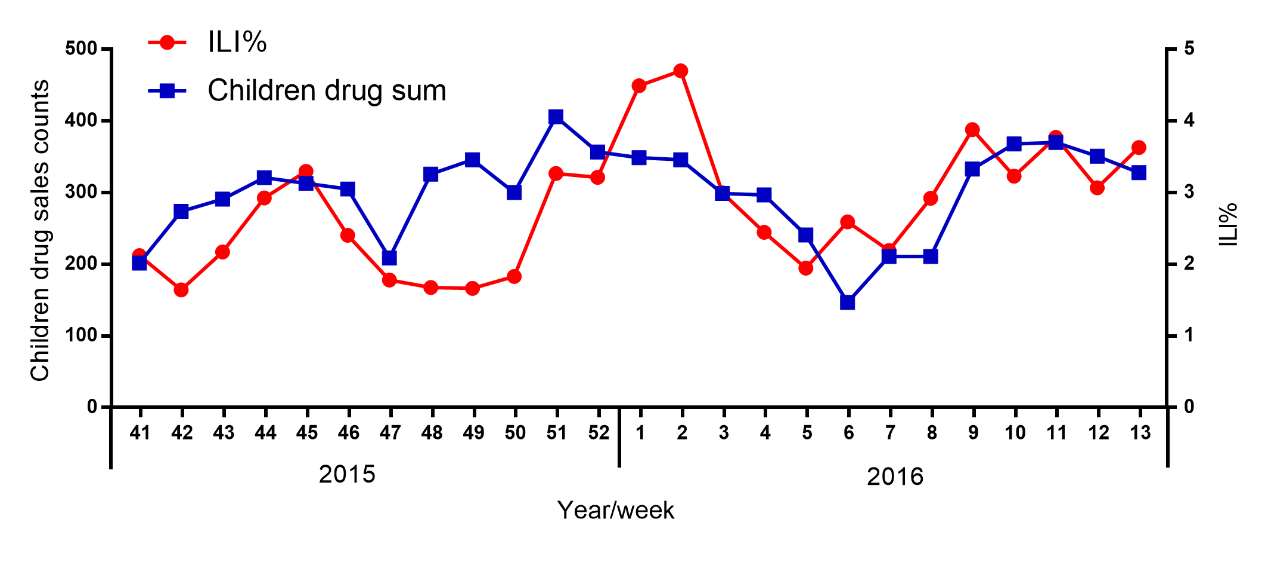


Figure . Comparison between ILI% and sum of children drug sales in Danyang, China, 2015/2016 influenza season.

Cross correlation analysis is carried out, and the results are as follows:

Table . The maximum positive CCF value of cross correlation analysis between drugs and ILI%, 2015 Oct-2016 Mar

|  |  |  |
| --- | --- | --- |
| **Category** | **CCF (SE)** | **Lag (week)** |
| adult drug sum | 0.363 (0.204) | 1 |
| children drug sum | 0.492 (0.200) | 0 |
| sum | 0.391 (0.204) | 1 |

The CCF value (CCF=0.391, SE=0.204) in the above Table 9 is significantly lower than the total CCF value (CCF=0.848, SE=0.204) of 2016 Oct-2017 Mar, the possible reason is that after 2016, the local hospital has built a respiratory symptom monitoring system based on HIS system, which has improved the quality of influenza surveillance. Among them, the adult drugs sum (CCF=0.363, SE=0.204) and the total sales amount (CCF=0.391, SE=0.204) still had the maximum correlation with ILI% after 1 week (lag=1), while the children drug sum (CCF=0.492, SE=0.200) showed the greatest correlation with ILI% in the case of not moving back (lag=0). This may be related to the selection of the minimum unit value. If the day is a lagging unit, there should be a lag number of days.

On the basis, the data of adult drugs and total OTC data were lagged 1 week later, the children drugs were not moved, and the Pearson correlation test was carried out. The results were as follows (See Table 10):

Table . Pearson correlation test of OTC drugs with ILI% in 2015 Oct-2016 Mar

|  |  |  |
| --- | --- | --- |
| **Category** | **ILI%** | |
| **r** | **p** |
| **adult drug sum** | **0.389** | **0.054** |
| 1. 维C银翘片 | 0.443 | 0.027 |
| 2. 氨咖黄敏胶囊 | 0.282 | 0.172 |
| 3. 风寒感冒颗粒 | 0.378 | 0.063 |
| 4. 感冒清热颗粒 | 0.229 | 0.272 |
| 5. 复方氨酚烷胺胶囊 | 0.399 | 0.048 |
| 6. 复方氨酚烷胺片 | 0.112 | 0.593 |
| 7. 四季感冒片 | -0.506 | 0.010 |
| 8. 小柴胡颗粒 | 0.249 | 0.230 |
| 9. 桑姜感冒片 | -0.355 | 0.081 |
| 10. 氨酚伪麻美芬片 | 0.292 | 0.156 |
| 11. 蒲地蓝消炎口服液 | 0.486 | 0.014 |
| 12. 蓝芩口服液 | 0.050 | 0.811 |
| 13.（复方）板蓝根颗粒 | 0.398 | 0.049 |
| 14. 复方金银花颗粒 | 0.576 | 0.003 |
| 15. 连花清瘟胶囊 | ***0.668*** | ***<0.01*** |
| 16. 盐酸吗啉胍片 | 0.131 | 0.534 |
| **children drug sum** | **0.492** | **0.012** |
| 17. 小儿伪麻美芬滴剂 | 0.017 | 0.936 |
| 18. 小儿氨酚黄那敏颗粒 | ***0.703*** | ***<0.01*** |
| 19. 小儿氨酚烷胺颗粒 | 0.056 | 0.789 |
| **sum** | **0.417** | **0.038** |

During the flu season from October 2015 to March 2016, children drug sum (r=0.492, p=0.012) was more correlated with ILI% than adult drug sum (r=0.389, p>0.05), indicating that children cold medicine is more accurate and less reliable than adult cold medicine. Among the children drugs, the relationship between No.18小儿氨酚黄那敏颗粒 (Pediatric Paracetamol, Atificial Cow-bezoar and Chlorphenamine Maleate Granules) and ILI% was the greatest (r=0.593, p<0.01). Among the adult drugs, the maximum correlation is between the No.15连花清瘟胶囊 (Lianhuaqingwen Capsule) (r=0.668, p<0.01) and ILI%. There was no correlation between most drugs and ILI%, and the rest were all correlated with ILI%, the possible reason is that the local hospital has not built a respiratory symptom monitoring system based on HIS system until 2016.

The following is data analysis of the non-influenza season, Apr-Sept, 2015.

Cross correlation analysis is carried out, and the results are as follows (See Table 11):

Table . The maximum positive CCF value of cross correlation analysis between drugs and ILI%, Apr-Sept, 2015

|  |  |  |
| --- | --- | --- |
| **Category** | **CCF (SE)** | **Lag (week)** |
| adult drug sum | 0.507 (0.234) | -7 |
| children drug sum | 0.431 (0.224) | -7 |
| sum | 0.531 (0.204) | -7 |

In both adults (CCF=0.507, SE=0.234) and children (CCF=0.431, SE=0.224), the data had the greatest positive cross-correlation only 7 weeks earlier (lag=-7), meaning that the two were not predictive of influenza. Therefore, no further analysis is conducted. This also suggests that during the non-influenza season, drug sales are less accurate than during the influenza season in predicting influenza.

# discussion

According to three years' overall data, the local trends in the ratio of ILI to OTC drug sales followed a cyclical pattern with relative highs in the influenza winter and spring season. The local ILI% trend was strongly correlated with OTC drug sales trends in the previous week during this period. Even though the sales of children drugs are lower than the sales of adult drugs, children drugs sales data were more effective at predicting ILI among the long-term unseasonal period.

In addition to analyzing data from every year, we also compared the seasonal data from October to March and non-seasonal data from April to September. We observed that drug sales surveillance is more suitable for monitoring seasonal flu in winter or spring because the correlation coefficients are affected by seasons, being higher in winter.

Since the local hospital has not built a respiratory symptom monitoring system based on HIS system until 2016, the data from 2016 to 2017 are more reliable. Therefore, this paper will pay more attention to discuss the statistic result from 2016 to 2017. In 2016, the weekly reported ILI% fluctuated but did not increase over the summer months. The ILI% peak appears located nearly from week 97 to week 108 which is from October 2016 to January 2017. It should be noted that the number of infections with H7N9 virus in the fifth epidemic wave whose onset is near October 2016 has been greater than earlier waves. [12] The onset of the outbreak in our graph is best matched by the onset of H7N9 outbreak in real world. [12] The earlier peak of drug sales, about one or two week before the ILI% peak, perfectly predicted the H7N9 outbreak in real world.

We also manually classified each OTC medication product to different ILI categories. According to the data from both 3 years analysis and 2016-2017 analysis, the ILI curve is the best matched by the drug No. 11 (Po blue to oral anti-inflammatory) sales curve. This drug is a kind of traditional Chinese herb medicine which can be used to prevent early symptom of cough, cold and anti-inflammatory swelling. Therefore, we propose that this drug can be used as the primary indicator in the drug sales surveillance.

The result carried out by statistic tool proved that there is a statistical correlation between drug sales volume and ILI%. The above r value indicates that there may be a strong, constant and stable correlation between the two kinds of variables. [13]This close correlation has strongly proved that the sales volume of OTC influenza drugs can play an effective role in detecting the high incidence or outbreaks of influenza-like illness.[13]

By analyzing the changes in pharmacy drug sales, we can predict the disease outbreak and control the disease outbreak in advance because we can find that the peak for Drug sales appear one or two weeks before the ILI% peak. This study also found that the children’s drug has higher correlation with ILI% than adult drug. By the pharmacy sales data, we can also provide information, such as drug sales, to make health policy to improve public health because the children’s drug can better predict the ILI% peak.

We can conclude syndromic surveillance systems in conjunction with traditional surveillance systems can improve influenza surveillance.[14]

Among detection of dynamic process of influenza-like illness, the traditional hospital surveillance is better in confirmation of cases. But the sensitivity and timeliness of the OTC drug sales surveillance is better than the traditional hospital surveillance because of the time lag.[14]Both of them have their own advantages. [13]The effective and scientific surveillance system should integrate surveillance data sources from traditional epidemiological surveillance data and laboratory data, and clinical data, human behavior data and syndromic real-time data.[13]Therefore, using drug sales data as a complementary approach for syndromic surveillance can improve detection of ILI outbreaks. [15]Drug sales surveillance with laboratories, hospitals, and death certificates surveillance may make the outbreaks prediction more timely.[15]

Extending the use of this drug sales surveillance can be feasible because drug sales are available to public health agencies, such local CDC department, in China. However, real-time surveillance of ILI is still seldom set up by Chinese public health agencies. In other words, surveillance in China is based on traditional hospital and laboratory surveillance system currently which does not include the electronic real-time drug sales surveillance. [15]

This real-time syndromic surveillance conducted by sales records in pharmacies can detect influenza activity and outbreaks in the 2015-2017 influenza season and enabled estimation of the incidence of ILI cases. It also offers a scientific data for public health measures. After data analysis, we need to synthesize the data and information to provide evidence to make policy to reduce outbreak risk.

It is critical that a national government to make a set of effective policy to guide the local public health agency to strengthen surveillance system to ensure timely and effective detection or response to outbreaks. In order to caution local agency to prevent and control outbreak in a timely and effective manner, government should make a set of comprehensive policies to guide these local agencies to do early outbreak prediction and surveillance. The quantitative sales data is considered as a useful supplementary system. Therefore, an advanced policy related to drug sales surveillance in China can ensure that outbreak cases can be identified and managed timely. [15]

# 

# bibliographY

1. *Influenza-like illness*.Wikipedia [cited 2018 Feb 10]; Available from: https://en.wikipedia.org/wiki/Influenza-like\_illness

2. Doucleff, M. *Deadly Bird Flu In China Evolves, Spreads To New Regions*. 2017 [cited 2018 Jan 1]; Available from: <https://www.npr.org/sections/goatsandsoda/2017/09/07/549069924/deadly-bird-flu-in-china-evolves-spreads-to-new-regions>

3. Cochrane, J., et al., *Alcohol use in China.* Alcohol and Alcoholism, 2003. **38**(6): p. 537-542.

4. Syndromic Surveillance: an Applied Approach to Outbreak Detection, What is Syndromic Surveillance. *webcitation.org.* 2006; Available from:https://www.webcitation.org/5LKEYsrRT.

5. *Bulletin on Reforming Jiangsu's GDP Accounting and Data Release System*. 2017; Available from: [www.people.com.cn](file:///C:\Users\jpegher\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\9LDUD2BR\www.people.com.cn).

6. *From Danyang to Shanghai*. 2012 [cited 2017; Available from: Travelchinaguide.com.

7. Zhang, Y., et al., *Influenza surveillance and incidence in a rural area in China during the 2009/2010 influenza pandemic.* PloS one, 2014. **9**(12): p. e115347.

8. Wang, X., et al., *Influenza surveillance in Shenzhen, the largest migratory metropolitan city of China, 2006-2009.* Epidemiol Infect, 2011. **139**(10): p. 1551-9.

9. Metzger, K.B., et al., *How many illnesses does one emergency department visit represent? Using a population-based telephone survey to estimate the syndromic multiplier.* Morbidity and Mortality Weekly Report, 2004: p. 106-111.

10. Socan, M., V. Erculj, and J. Lajovic, *Early detection of influenza like illness through medication sales.* Central European journal of public health, 2012. **20**(2): p. 156.

11. *Cross Correlation Functions and Lagged Regressions*. Applied Time Series Analysis [cited 2018 march 20]; Available from: [https://onlinecourses.science.psu.edu/stat510/node/74.](https://onlinecourses.science.psu.edu/stat510/node/74)

12. World Health Organization, Emergency Preparedness, Response. *Human infection with avian influenza A(H7N9) virus – China*. 2017 [cited 2018 Feb 10]; Available from: [http://www.who.int/csr/don/13-september-2017-ah7n9-china/en/.](http://www.who.int/csr/don/13-september-2017-ah7n9-china/en/)

13. Guo,J.,et al, *Analysis on Correlation Between Sales of the Over-counter Drugs and the Incidence of Influenza-like-illness.* [cited 2018 Feb 10]; Available from: http://www.doc88.com/p-9995785726366.html.

14. Dong, X., et al., *Syndromic surveillance for influenza in Tianjin, China: 2013–14.* Journal of Public Health, 2016. **39**(2): p. 274-281.

15. Vergu, E., et al., *Medication Sales and Syndromic Surveillance, France.* Emerging Infectious Diseases, 2006. **12**(3): p. 416-421.