EVALUATION OF HEPATITIS A SURVEILLANCE DATA AND OUTBREAK DETECTION IN YUNNAN PROVINCE, CHINA, FROM 2004 THROUGH 2009

Liu Xiaoqiang1,2,3, Virasakdi Chongsuvivatwong2, Chuleeporn Jiraphongsa3, Lu Lin1, Yang Jun1, Li Qiongfen1 and Vorasith Sornsrivichai2

1Yunnan Provincial Centers for Disease Control and Prevention, Kunming, China; 2Epidemiology Unit, Faculty of Medicine, Prince of Songkla University, Hat Yai, Thailand; 3International Field Epidemiology Training Program (FETP), Bureau of Epidemiology, Department of Disease Control, Ministry of Public Health, Nonthaburi, Thailand

Abstract. Hepatitis A outbreaks may be averted if detected early. The current study objectives were to evaluate the quality and timeliness of hepatitis A surveillance data from Yunnan Province, China, and to evaluate the sensitivity of the system for reporting outbreaks. The study period was from January 1, 2004 through December 31, 2009. Records from the National Infectious Diseases Surveillance System (NIDSS) were compared with those from local hospitals. The timeliness of case detecting, reporting and updating was also analyzed. The numbers of cases in a specific location during a moving time interval were computed to identify past outbreaks which were then validated against reported cases. The NIDSS received 38,095 reports during the study period; of which 6% were duplicates, and 26% had serological confirmation. The sensitivity and positive predictive value of cases were 96% and 98%, respectively. Time from onset to diagnosis remained constant over the 6-year period. Delays in reporting became shorter, and quality control of data improved over the period, but the timeliness of identifying duplicate records did not. Based on data from NIDSS, 9 outbreaks should have been reported, but only 3 were reported. The 3 reported outbreaks were shorter in duration than the 6 unreported ones, but the numbers of cases involved were not significantly different. Surveillance data needs improvement in updating timeliness. The system for outbreak detection and reporting needs to be improved.

Keywords: hepatitis A, surveillance evaluation, sensitivity, positive predictive value, timeliness, outbreak detection

INTRODUCTION

Hepatitis A is a widespread acute illness caused by infection with the hepatitis A virus (HAV), a nonenveloped ribonucleic acid (RNA) picornavirus that infects only primates. The disease varies in clinical severity from a mild illness that lasts 1-2 weeks, to a severely disabling disease that lasts several months. Clinical manifestations of hepatitis A include fever, malaise, anorexia, nausea, and abdominal discomfort, followed within a few days by jaundice. Prolonged, relapsing hepatitis for up to one year occurs
in 15% of cases; no chronic infection is known to occur. The disease is spread primarily through ingestion of food or water that has been contaminated with the fecal matter of someone who is infected or through close, personal contact (WHO, 2000; Cuthbert, 2001; Klevens et al, 2008; CDC, 2010).

Even with improvement in sanitation and the use of vaccination (since 1995) have lowered the incidence of hepatitis A in the last 20 years it still causes a high disease burden and is considered a public health problem in both developed and developing countries. Over the first decade of the 21st century, outbreaks of hepatitis A occurred in many countries worldwide, including the United States of America (Wheeler et al, 2005), England (Edelstein et al, 2010), Australia (Lee et al, 2004; Rowe et al, 2009), and Spain (Syed et al, 2003; Arce et al, 2004; Tortajada et al, 2009).

In China, hepatitis A is a serious public health problem. In 1988, 47 deaths (case-fatality rate of 0.015%) were reported in a large epidemic in Shanghai, China, which involved more than 300,000 adolescents and young adults (Cooksley, 2000). Serum IgG positive prevalence among the general population has been reported to be as high as 81% in China (Xiao et al, 2007). From 1990 to 1999, more than 200,000 cases (incidence rate >17/105 population) have been reported annually. From 2000 to 2007, the reported incidence rate declined from 10.8 to 5.9/105 population. The China Centers for Diseases Control and Prevention (CDC) categorizes all provinces of mainland China into three strata: high (>8/105 population/year), intermediate (4-8/105 population/year), or low incidence (<4/105 population/year). Yunnan, where this study was conducted, is one of the 10 high-incidence provinces (Cui et al, 2009).

Mass post-exposure vaccination to contain the spread of HAV in established outbreaks has been proven to be efficacious in curbing emerging epidemics (Delpech et al, 2000; Arce et al, 2004, 2007; Shen et al, 2008; Diaz et al, 2010; Edelstein et al, 2010). The World Health Organization (WHO), the Advisory Committee on Immunization Practices (ACIP) of the USA (Fiore et al, 2006), and the Ministry of Health of China (Wang, 2006) have recommended mass hepatitis A vaccination to interrupt outbreaks at an early stage.

In China, hepatitis A has been included in the Notifiable Infectious Disease Surveillance System (NIDSS) since 1990. In 2004, an online, real-time reporting system was established (Wang et al, 2008). At 93% of hospitals and 66% of township health centers (in 2006), 18 terms of specific information (including age, sex, address, occupation, onset date, diagnosis date) about suspected or confirmed cases are collected with a standard structured card and entered into the NIDSS within 24 hours of diagnosis (Ma et al, 2006). County-level CDCs then check the information, delete duplicates, and confirm the report through the NIDSS within another 24 hours. Duplicate records are identified by comparing name, age, onset date, and diagnosis. If the NIDSS is unavailable, health centers or hospitals will fax the paper card to county-level CDCs to perform the entering process. Upper-level CDCs can view the database and results of the format-structured analysis online and retrieve the database for further analysis. If five or more cases occur in one village within one week, the county government will announce an outbreak. Immediate responses include active case finding, case administration and treatment, mass vaccination, improving sanitation, and endorsement of health education. The event
information is reported through another online real-time outbreak report system that is linked to NIDSS.

The objectives of this study were to evaluate the validity, quality, and timeliness of hepatitis A surveillance data, and to evaluate the system’s sensitivity in reporting outbreaks in Yunnan Province based on the analysis of NIDSS data. These are necessary components to trigger reliable outbreak announcements and subsequent control and prevention measures.

MATERIALS AND METHODS

Study setting

Yunnan is a province in southwestern China with a population of almost 45 million. It has 16 prefectures containing 129 counties.

This study was divided into three parts. The first part involved the use of a field investigation to compare the validity and quality of data from the online surveillance system with that of the medical records. The second part was based on analysis of the online surveillance system data to evaluate the timeliness. In the third part, the outbreaks reported through the outbreak surveillance system were compared with the outbreaks detected through the NIDSS.

Validity and quality of surveillance data

An evaluation of the validity of a surveillance system generally includes two measurements: sensitivity and positive predictive value (PVP) (German et al., 2001; Buehler et al., 2004). Sensitivity refers to the probability that an actual case has been correctly reported. PVP refers to the probability that a person reported as a case actually is a case. Data quality refers to the proportion of consistent data values in the surveillance system compared to the “true” values from relevant medical records (German et al., 2001; Buehler et al., 2004; Porta, 2008). Since hepatitis A is a seasonal disease in China (Cui et al, 2009), we studied the cases that occurred in June and December during each year from 2004 to 2009, to represent high and low incidence months.

Six counties were randomly selected as study sites. In each selected county, the county hospital and two randomly chosen township health centers that had previously reported hepatitis A were used. Information from medical records (inpatient records and outpatient log books) was treated as the gold standard for evaluating surveillance information.

According to Chinese national diagnostic criteria, hepatitis A is classified into probable or confirmed (Ministry of Health of China, 2008). A probable case is defined as any acute illness with discrete onset of symptoms compatible with hepatitis A (e.g., fever, weakness, and digestive tract symptoms, such as anorexia, nausea, jaundice, bilirubin in the urine), elevated serum aminotransferase levels, with evidence that links this probable case with a confirmed case or outbreak. A confirmed case was defined as any probable case testing positive for anti-HAV IgM (in the absence of recent vaccination) or a 4-fold rise in specific IgG antibody. Acute hepatitis without laboratory confirmation but associated with an outbreak was defined as unclassified viral hepatitis. All probable, confirmed, and unclassified cases must be reported through the NIDSS.

In our study, cases of hepatitis A (confirmed and probable) in the data base were evaluated regardless of whether they were laboratory confirmed, or diagnosed as probable, unclassified, or not a real
Case. Data from these cases were evaluated for sensitivity and PVP (Table 1). We also checked the onset date, diagnosis date, and report date from the NIDSS database compared to information from the medical record.

**Timeliness of hepatitis A surveillance**

Specific case information for 2004-2009 was retrieved from the NIDSS computer system. Nine variables for specific cases, including index number, gender, age, address, onset date, diagnosis date, report date, checking date, and deletion date of duplicate cases were obtained for use in evaluation of timeliness. There were five time points during the process of case reporting: onset, diagnosis, report, check, and duplicate case deletion. Six intervals were introduced as illustrated in Fig 1 where the time intervals are assigned abbreviations that were defined and used for the results.

An estimated 30-day incubation period for hepatitis A was used for this study as a surrogate measure for the period of communicability (Cuthbert, 2001; Klevens et al., 2008). This period is critical for considering when to implement effective prevention and control measures (Jajosky and Groseclose, 2004). To assess whether data from the NIDSS support the timely identification of an outbreak, the percentage of NIDSS cases reported during one incubation period was calculated.

**Outbreak report and detection**

Reported outbreaks were obtained from the Yunnan CDC, which received its reports from the county CDC office. For analyzing of the NIDSS data set, a probable outbreak was defined as an occurrence of hepatitis A in one township during any continuous 30-day period in 2009 greater than the baseline calculated as the mean plus two times the standard deviation of

---

**Table 1**

Symbols for number of cases classified by type of report and diagnosis, Yunnan Province, China, 2004-2009.

<table>
<thead>
<tr>
<th>Confirmed hepatitis A</th>
<th>Probable hepatitis A</th>
<th>Unclassified hepatitis</th>
<th>Non-hepatitis A (other type of hepatitis)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reported as confirmed hepatitis A</td>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>Reported as probable hepatitis A</td>
<td>e</td>
<td>f</td>
<td>g</td>
<td>h</td>
</tr>
<tr>
<td>Reported as hepatitis unclassified</td>
<td>i</td>
<td>j</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Did not report as hepatitis A</td>
<td>k</td>
<td>l</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>M</td>
<td>N</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

a The third column denotes negative result for reported case checking. The total number was not used for calculations.

To evaluate the surveillance system, sensitivity is the ability of the surveillance system to identify true cases, and PVP is the proportion of cases reported by a surveillance system that are true cases. Thus, we used the following formulas:

- Total sensitivity = \((a+b+e+f)/(M+N)\) x 100%
- Sensitivity for confirmed cases = \(a/M\) x 100%
- Total PVP = \((a+b+e+f)/(P+Q)\) x 100%
- PVP for confirmed cases = \(a/P\) x 100%
the number of cases that occurred in the same area and the same time period in each of the years 2004 - 2008 (Yang et al, 2004). Probable outbreaks were then scrutinized by a panel of experts, which included epidemiologists and medical specialists (Buehler et al, 2004). Finally, six attributes were given to each identified outbreak: specified place of outbreak, date outbreak started, date outbreak peaked, date outbreak stopped, number of cases involved in the outbreak, and occupation of most victims.

RESULTS

Incidence and diagnosis of hepatitis A

From January 1, 2004 to December 31, 2009, a total of 38,095 cases were reported to the NIDSS in Yunnan Province (Table 2). The numbers of reported cases rose from 2004 to 2007 and declined during subsequent years. The proportion of duplicate records ranged from 6% to 9%, mainly due to the fact that some patients were seen at more than one hospital. The proportion of laboratory-confirmed cases varied from 16% to 33%. These levels are much lower than the national average of 52%.

Sensitivity and positive predictive value for case detection, and validity of reporting value

Table 3 displays cross-classification between data from the report (row) and from the NIDSS (column) at the 6 selected county hospitals and 12 selected township health centers. The inconsistency between the two sources was minor. NIDSS had a sensitivity of 96% and a positive predictive value of 98%.

Of 136 confirmed and probable cases, 130 (96%) had consistent information about onset date and diagnosis date with medical record and 118 (87%) had consistent date of report.

Timeliness of surveillance

Table 4 summarizes variation in the six types of time intervals defined in Fig 1. The time interval from onset to diagnosis (Tod) remained quite constant over the 6-year period, whereas the time interval from diagnosis to report (Tdr), time interval from onset to report (Tor), and time interval from report to checking (Trc) continued to improve, and the time interval from checking to duplicate deletion (Tcd) and time interval from report...
### Table 2
Diagnosed and reported cases of hepatitis A, Yunnan Province, China, 2004-2009.

<table>
<thead>
<tr>
<th>Year</th>
<th>No. of cases reported</th>
<th>Percent with duplicate records</th>
<th>Percent confirmed by laboratory</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>5,780</td>
<td>6.5</td>
<td>26.3</td>
</tr>
<tr>
<td>2005</td>
<td>6,228</td>
<td>6.7</td>
<td>32.9</td>
</tr>
<tr>
<td>2006</td>
<td>7,565</td>
<td>6.7</td>
<td>30.0</td>
</tr>
<tr>
<td>2007</td>
<td>8,104</td>
<td>7.4</td>
<td>24.4</td>
</tr>
<tr>
<td>2008</td>
<td>5,969</td>
<td>8.6</td>
<td>20.5</td>
</tr>
<tr>
<td>2009</td>
<td>4,448</td>
<td>8.0</td>
<td>15.9</td>
</tr>
<tr>
<td>Total</td>
<td>37,095</td>
<td>6.1</td>
<td>25.6</td>
</tr>
</tbody>
</table>

### Table 3
Number of cases reported or diagnosed as hepatitis A in six counties of Yunnan Province, China, during June and December, 2004-2009.

<table>
<thead>
<tr>
<th>Confirmed hepatitis A</th>
<th>Probable hepatitis A</th>
<th>Unclassified hepatitis</th>
<th>Non-hepatitis A (other type of hepatitis)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reported as confirmed hepatitis A</td>
<td>38</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Reported as probable hepatitis A</td>
<td>5</td>
<td>88</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Reported as hepatitis unclassified</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Did not report as hepatitis A</td>
<td>0</td>
<td>6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>43</td>
<td>96</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

\[
\text{Sensitivity} = \frac{38+5+2+88}{38+5+2+88+6} = 95.7\%
\]

\[
\text{Positive predictive value} = \frac{38+5+2+88}{38+5+2+88+1+2} = 97.8\%
\]

The three reported outbreaks were announced 3-10 days after the number of cases peaked. In response, the local government performed control measures, such as case management, mass vaccination, and health education. The outbreaks stopped within two incubation periods.

All three reported outbreaks were confirmed in, and confined to, schools, whereas four of the six unreported outbreaks also involved persons in the community. However, the proportion of to duplicate deletion (Trd) did not show any specific trend.

**Outbreak report and detection**

Table 5 lists the hepatitis A outbreaks in Yunnan Province in 2009. Nine outbreaks were identified from NIDSS historical data; however, only three of these were actually reported. The numbers of reported probable or confirmed cases in the three reported outbreaks were 31, 41 and 12. Thirty more cases were detected in the NIDSS for the second outbreak, making this the largest outbreak (71 cases) in this series.
Type of time interval | Year | Total
--- | --- | ---
No. of cases reported | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | Total
--- | --- | --- | --- | --- | --- | --- | ---
5,780 | 6,228 | 7,565 | 8,104 | 5,969 | 4,448 | 37,095 |
Percent confirmed by laboratory | 26.3 | 32.9 | 30.0 | 24.4 | 20.5 | 15.9 | 25.6%

**Tod** (from onset to diagnosis)
- Median (days): 6, 6, 6, 6, 5, 6
- IQR (days): 3-9, 3-9, 3-9, 3-9, 3-9, 2-9

**Tdr** (from diagnosis to report)
- Median (days): 2, 1, 1, 1, 0, 0, 1
- IQR (days): 1-7, 0-2, 0-1, 0-1, 0-1, 0-1
- % within 1 day: 42.7, 66.3, 76.8, 86.7, 95.8, 99.9
- % within 1 incubation period: 89.0, 93.4, 95.8, 96.7, 96.7, 97.6

**Tor** (from onset to report)
- Median (days): 10, 8, 7, 7, 6, 6, 7
- IQR (days): 6-18, 4-12, 4-11, 4-11, 3-9, 3-9, 4-11
- % within 1 incubation period: 89.0, 93.4, 95.8, 96.7, 96.7, 97.6

**Trc** (from report to checking)
- Median (days): 0, 0, 0, 0, 0, 0, 0
- IQR (days): 0-1, 0-1, 0-1, 0-0, 0-0, 0-0, 0-1
- % within 1 day: 88.0, 86.5, 89.7, 95.1, 98.9, 100
- % within 1 day: 92.7

No. of duplicate cases reported | 377 | 417 | 504 | 603 | 515 | 356 | 2,272 |

**Tcd** (from checking to delete)
- Median (days): 2, 4, 4, 5, 2, 1.5, 3
- IQR (days): 0-27, 1-16, 1-20, 1-19, 0-10, 0-7, 0-14

**Trd** (from report to delete)
- Median (days): 3, 5, 5, 5, 3, 2, 4
- IQR (days): 0-28, 1-16, 1-20, 1-19, 1-10, 1-7, 1-15

---

Table 4
Timeliness of hepatitis A surveillance, Yunnan Province, China, 2004-2009.

---

**DISCUSSION**

Early detection of hepatitis A outbreaks may assist in the control of further spread of the disease because they can be stopped by effective vaccination programs (Arce et al., 2007; Xiao et al., 2007; Shen et al., 2008; Cui et al., 2009). For further studies of the performance of prospective outbreak detection, it is necessary to evaluate data from surveillance both for individual cases and outbreaks (German et al., 2001; Buehler et al., 2004).

In our study, duplicate reports of hepatitis A cases were not uncommon.
Table 5
List of outbreaks of hepatitis A reported and detected from Yunnan Province, China, 2009.

<table>
<thead>
<tr>
<th>ID</th>
<th>Extent of outbreak</th>
<th>Date outbreak starts</th>
<th>Date outbreak peaks</th>
<th>Date outbreak ends</th>
<th>Date outbreak announced</th>
<th>Duration of outbreak (days)</th>
<th>Number of cases involved</th>
<th>Median (IQR) of cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reported outbreak</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No.1</td>
<td>S</td>
<td>Mar 23</td>
<td>Mar 28</td>
<td>Apr 5</td>
<td>Apr 1</td>
<td>18</td>
<td>31</td>
<td>5 (4, 8)</td>
</tr>
<tr>
<td>No.2</td>
<td>S</td>
<td>Apr 16</td>
<td>May 1</td>
<td>Jun 18</td>
<td>May 8</td>
<td>63</td>
<td>71</td>
<td>5 (4, 7.5)</td>
</tr>
<tr>
<td>No.3</td>
<td>S</td>
<td>Oct 12</td>
<td>Oct 20</td>
<td>Nov 2</td>
<td>Oct 30</td>
<td>27</td>
<td>12</td>
<td>7 (5, 8)</td>
</tr>
<tr>
<td>Unreported outbreak</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No.4</td>
<td>S+C</td>
<td>Apr 17</td>
<td>Apr 30</td>
<td>Jul 17</td>
<td>-</td>
<td>91</td>
<td>58</td>
<td>5 (1.2, 7)</td>
</tr>
<tr>
<td>No.5</td>
<td>S+C</td>
<td>Apr 19</td>
<td>Jun 2</td>
<td>Jul 28</td>
<td>-</td>
<td>100</td>
<td>40</td>
<td>5 (4.7)</td>
</tr>
<tr>
<td>No.6</td>
<td>S</td>
<td>Apr 20</td>
<td>Jul 1</td>
<td>Jul 10</td>
<td>-</td>
<td>81</td>
<td>18</td>
<td>7.5 (5.2, 11)</td>
</tr>
<tr>
<td>No.7</td>
<td>S+C</td>
<td>May 1</td>
<td>May 27</td>
<td>Jul 13</td>
<td>-</td>
<td>74</td>
<td>40</td>
<td>7 (4.8, 12)</td>
</tr>
<tr>
<td>No.8</td>
<td>S+C</td>
<td>May 5</td>
<td>Jul 27</td>
<td>Sep 5</td>
<td>-</td>
<td>123</td>
<td>43</td>
<td>4 (3.6)</td>
</tr>
<tr>
<td>No.9</td>
<td>S</td>
<td>Jun 4</td>
<td>Jun 16</td>
<td>Aug 5</td>
<td>-</td>
<td>63</td>
<td>26</td>
<td>6 (5.8)</td>
</tr>
</tbody>
</table>

S, school; C, community

*Time interval in days from date of onset to date of report calculated from timeliness of outbreak-associated cases

Not statistically significant using the rank-sum test, $p=0.824$
Laboratory confirmation was obtained in fewer than one-third of reported cases. Sensitivity and positive predictive value for case notification were both >95%. The timeliness of notification improved over time, but the timeliness for deleting duplicate records did not. Nine outbreaks were detected from analysis of NIDSS data, but only three of them were reported to the system. All nine outbreaks involved schools. Unreported outbreaks lasted longer, but did not involve more cases than reported outbreaks. Although not all the outbreaks involving the community were reported, the number of outbreaks was too small to test whether a report not in the system was associated with involvement of the community.

Undeleted duplicate records included in a recent data set may increase the probability of an outbreak being falsely detected. Duplicate records are difficult to avoid because patients are free to seek medical service anywhere they wish, and each time they visit a hospital a report must be made to the surveillance system. However, the complicated process of deleting duplicate reports involves a significant amount of work and subjective judgment (Hu et al, 2008; Huang et al, 2010). Surveillance systems in some countries assign a unique number for each person to avoid duplicate reports (German et al, 2001). We suggest a unique identification number, such as a health insurance number should be used to distinguish patients with the same name, date of birth, and other attributes.

Serologic diagnosis can be clinically useful because different types of hepatitis have different prognoses and complication patterns. From a descriptive epidemiological point of view, hepatitis A and E are water- and food-borne, whereas types B and C are transmitted by blood and other body fluids (Cuthbert, 2001; Sharapov and Hu, 2010). Vaccines are currently available for types A and B but not for types C, D, and E (CDC, 2010). However, during routine surveillance, hepatitis is difficult to divide into specific types of virus based only on clinical presentation (Sharapov and Hu, 2010). Lack of diagnostic facilities and limited patient finances were the main reasons for a low proportion of laboratory confirmed diagnoses for other diseases in China (Xu et al, 2010). For hepatitis A, if serologic confirmation is required, the patient must pay 10 Yuan (USD 1.5) more for their medical examination (Yunnan Provincial Health Bureau, 2005). Health financial reform is needed if these data are to be obtained.

Sensitivity and PVP are the most important validity measurements for a surveillance system and for outbreak detection. False-positive reports can lead to unnecessary interventions and subsequent false outbreak reports (German et al, 2001). Specificity and negative predictive value cannot be defined because the number of true negatives has not yet been clearly defined. Compared with hepatitis A surveillance in other countries or regions, such as in Taiwan (32%) and Italy (79.4%), mainland China’s surveillance system has a higher sensitivity (Lopalco et al, 2002; Hsieh et al, 2009).

With the Chinese reporting system, Tdr, but not Tod, improved over the six years of the study. This difference was probably due to the fact the Tod was based partly on the patient’s promptness in seeking health care and partly on the ability of the health system to diagnose the case in a timely manner (Jajosky and Groseclose, 2004; Yoo et al, 2009). The Tod for hepatitis A has not improved, whereas the Tdr improved continuously. This discrepancy indicates improvement in the
surveillance data was more on the health service side than on the community side. The challenge is how to shorten Tcd and Trd, which reflect the health service side (Huang et al, 2010).

Our study shows reported outbreaks were of shorter duration but do not involve fewer cases than unreported outbreaks. A shorter outbreak is easier to confirm and is more likely to be reported. The absence of a report leads to less intense intervention and may allow the outbreak to last longer. All the reported outbreaks were followed by mass vaccination for hepatitis A in schools and communities. It is difficult to judge whether the vaccine or other measures stopped the outbreaks. The cost-benefit of the mass vaccination also needs further study.

Our study found students are a high-risk group for hepatitis A since all the outbreaks we evaluated were closely related to schools. An earlier country-wide study reported the same pattern. About a half (49%) of outbreaks occurred in schools during 2004 - 2007 (Cui et al, 2009). This may be explained by the high population density, close contact, and other risk-generating behaviors found among the school pupils (WHO, 2000; Cuthbert, 2001; Roberts and Palmer, 2006; Pelletier et al, 2010). School health education has been found useful for reducing the occurrence of hepatitis in several outbreak investigations (Leoni et al, 1998; Marks et al, 2001). Further studies to predict outbreaks among school pupils are needed.

Our study was based on NIDSS data, which is more complete than surveillance in the past which only reported cases monthly (Wang et al, 2008). The data demonstrate the size of the outbreak and the way to report them. One study limitation was assuming the logbook was the gold standard. However, with such a low percentage of laboratory-confirmed cases of hepatitis A and a lack of laboratory confirmation in the medical records, there may be unreported cases.

In conclusion, using the medical logbook as a gold standard, the sensitivity and specificity of hepatitis A cases in NIDSS were high. However, duplicate records were not uncommon, and the percent of cases with laboratory confirmation was less than 30%. Two-thirds of the outbreaks were not reported. This needs further improvement.

ACKNOWLEDGEMENTS

This research was a part of the first author’s thesis to fulfill the requirement of PhD in epidemiology at Prince of Songkla University. The authors would like to thank the CDC for Yuxi, and the CDC for Honghe in Yunnan Province for supporting the field investigations, and Dr Xu Wen and Dr Fu Xiaoqing for collecting the data. Financial support was provided by the Yunnan Provincial CDC, China.

REFERENCES


Huang WL, Chen CH, Hong RT. Analysis on duplicate cards in infectious disease surveillance system in Fujian Province, China. Pract Prev Med 2010; 17: 3.


Pelletier AR, Mehta PJ, Burgess DR, et al. An outbreak of hepatitis A among primary


