INTRODUCTION

Highly pathogenic avian influenza (HPAI) H5N1 is an infectious disease that can be transmitted from animals to humans. The disease is widely spread throughout the world, including Thailand. Many researchers have studied the risk factors associated with the presence of H5N1 and spatial techniques are commonly used for evaluation. The present study investigated outbreaks of HPAI H5N1 in Thailand between January 2004 and November 2005 using kernel smoothing and Kulldorff’s scan statistics. A total of 1,493 HPAI H5N1 outbreak points from 288 districts in 60 provinces around Thailand were recorded using national outbreak data. For the kernel smoothing, the provinces with the highest risk were Phitsanulok, Phichit, Suphan Buri, Ang Thong, Samut Prakan, Bangkok, Chon Buri, Kamphaeng Phet and Saraburi. Kulldorff’s spatial scan statistics showed that the high-risk districts were in the central and lower northern parts of Thailand. The findings confirmed that central Thailand had the highest risk for HPAI H5N1 outbreaks. The appropriate authorities should focus on this area for disease control and prevention, and should pay special attention to this area when outbreaks occur in neighboring countries. This may help authorities to prevent outbreaks or decrease the magnitude of outbreaks when they occur.

Keywords: avian influenza, spatial analysis, Thailand

ABSTRACT

Highly pathogenic avian influenza (HPAI) H5N1 is an infectious disease that can be transmitted from animals to humans. The disease is widely spread throughout the world, including Thailand. Many researchers have studied the risk factors associated with the presence of H5N1 and spatial techniques are commonly used for evaluation. The present study investigated outbreaks of HPAI H5N1 in Thailand between January 2004 and November 2005 using kernel smoothing and Kulldorff’s scan statistics. A total of 1,493 HPAI H5N1 outbreak points from 288 districts in 60 provinces around Thailand were recorded using national outbreak data. For the kernel smoothing, the provinces with the highest risk were Phitsanulok, Phichit, Suphan Buri, Ang Thong, Samut Prakan, Bangkok, Chon Buri, Kamphaeng Phet and Saraburi. Kulldorff’s spatial scan statistics showed that the high-risk districts were in the central and lower northern parts of Thailand. The findings confirmed that central Thailand had the highest risk for HPAI H5N1 outbreaks. The appropriate authorities should focus on this area for disease control and prevention, and should pay special attention to this area when outbreaks occur in neighboring countries. This may help authorities to prevent outbreaks or decrease the magnitude of outbreaks when they occur.

Keywords: avian influenza, spatial analysis, Thailand
Livestock Development, 2008). Many studies on the risk factors associated with HPAI H5N1 use spatial analysis, which has indicated that the central part of Thailand and the connecting areas, such as northern Thailand, have the highest risks of infection. They revealed that the relationship between free-grazing duck husbandry and rice paddy fields is a major risk factor (Gilbert, 2005; Gilbert et al., 2006; Songserm et al., 2006; Gilbert et al., 2007; Gilbert et al., 2008). During the 2004 outbreak, the relative risk of outbreak in the central region was 3.7 (95% confidence interval = 3.4–4.1), which was higher than for other regions (Tiensin et al., 2005). Many free-grazing ducks in the central regions of Thailand graze in the rice paddy fields, and the movement of the ducks from field to field is a major risk factor for the spread of HPAI H5N1. Based on Thailand’s national outbreak data between 2004 and 2005, Tiensin et al. (2009) used Bayesian smoothing and a local spatial cluster statistical test to measure the incidence of HPAI H5N1 and to detect the significant spatial clusters at the subdistrict level. They found that the central region contained multi-clusters as well as the highest incidence of the disease. The Q statistic was used to evaluate the risk of the neighborhood infection of H7N1 (Mulatti et al., 2010). In addition, Paul et al. (2010) used Bayesian spatial analysis to define the HPAI H5N1 risk area between July 2004 and May 2005 and found the highest level of relative risk was in the central part of Thailand. Moreover, they found that areas with free-grazing ducks and rice production were at a higher risk for HPAI H5N1.

Ordinary statistical methods, such as descriptive statistics and inferential statistics, may not be adequate to analyze disease patterns because such patterns always involve spatial factors. Spatial analysis or spatial statistics use many different techniques to analyze the different properties of spatial data. Previous studies have only used some techniques of spatial analysis to visualize and quantify the risk of HPAI H5N1. This study investigated the outbreak density (using kernel smoothing to evaluate an outbreak point pattern) and examined the local spatial-point clustering using Kulldorff’s scan statistic for the HPAI H5N1 outbreak at the district level in Thailand between January 2004 and November 2005. This study determined whether the central and lower northern parts of Thailand are high-risk areas for HPAI H5N1.

MATERIALS AND METHODS

Data collection

The national outbreak data were sourced from the Thai Department of Livestock Development (DLD), which also served as the database for the present study. The data were collected between January 23, 2004 and November 9, 2005 during which time there were three outbreak waves of HPAI H5N1: January 2004–June 2004 (wave 1), July 2004–June 2005 (wave 2) and July 2005–November 2005 (wave 3). The outbreak data contained the date of disease detection, the location and coordinate points. The poultry population census was also obtained from the DLD website (http://www.dld.go.th). To analyze the data for the whole study period and the second wave, the average from the poultry census for 2004 and 2005 was used. For the first and third waves, the number of poultry in 2004 and 2005, respectively, was used. A map of Thailand was obtained from the DLD.

Statistical analysis

Detection of spatial variation

Kernel smoothing is a nonparametric estimation of the probability density function that can be used for spatial point processing (Bailey and Gatrell, 1995; Gatrell and Bailey, 1996; Dohoo et al., 2010). The present study analyzed all outbreak points in each wave as well as the total outbreaks in Thailand. Default grid cells, bandwidth and the
nearest neighbor pyramid re-sampling method in ArcGIS (version 10; (Esri; Redlands, CA, USA)) were used for the technique calculation and mapping at the provincial level. Bailey and Gatrell (1995), Gatrell and Bailey (1996) and Barreto et al. (2008) previously described the calculation formula (Equation 1):

$$\hat{\lambda}_x(s) = \frac{1}{n} \sum_{i=1}^{n} k \left( \frac{s - s_i}{\tau} \right)$$

(1)

where $\hat{\lambda}_x()$ is the kernel density estimator that estimates the outbreak areas, $\tau$ is the bandwidth or smoothing parameter, $k()$ is the kernel weighting function, $s$ is the center of the outbreak area and $s_i$ is the location of each outbreak point.

**Detection of local point clustering**

A spatial cluster analysis for infectious diseases can be used to investigate questions related to causal risk factors. A disease cluster means that there is a localized risk factor. Normally, spatial processes can be characterized in terms of first-order and second-order properties (Bailey and Gatrell, 1995). The first-order properties are associated with studies that examine the entire region of interest. However, both first-order and second-order properties are involved in studies that take place within a specific study period. The second-order properties imply a relationship between the number of events and a subregion. The present study used Kulldorff’s spatial scan statistics to detect the second-order properties of the HPAI H5N1 outbreak pattern (Kulldorff and Nargarwalla, 1995). Similar to the kernel smoothing analysis, an overall outbreak status and each individual outbreak wave were calculated. The unit of interest was the district level, and the test statistic was based on the previously described discrete Poisson models (Waller and Gotway, 2004; Dohoo et al., 2010). SaTScan, version 9.1 (Kulldorff, 2009) was used to calculate the spatial scan test. The software detected all the outbreak points by simulated circles based on geography.

The circular scans were used on all the areas that had the outbreak points. The circular size ran from zero to the maximal default using the space-time permutation in the exponential model in Equation 2 (Bivand et al., 2008):

$$K = \max \left( \frac{O_Z}{E_Z} \left( \frac{O_z - O_Z}{E_z - E_Z} \right) \right)$$

(2)

where $K = \text{Kulldorf's statistic}$, $z$ is the outbreak point of $Z_i$ in each cluster, $O$ is the observed outbreak points in each cluster, $E$ is the expected outbreak points in each cluster and max indicates that this model uses the highest value of the likelihood ratio.

Kulldorff’s statistic calculates the radius of clusters as an outcome. The relative risk is used to compare the outbreak points inside the circle which is then divided by the outbreak points outside the circle. The null hypothesis is that no clustering will illustrate the relative risk as one. In a case where the relative risk is higher than one, this indicates that the outbreak points in the circle are clustering (Bivand et al., 2008). ArcGIS, version 10 was used for the map display.

**RESULTS AND DISCUSSION**

A total of 1,493 HPAI H5N1 outbreak points in 288 districts from 60 provinces around Thailand were recorded. Nine records from three districts were removed from the analysis because of incomplete data.

**Mapping and spatial variation**

Kernel smoothing was used to calculate the first-order spatial property for HPAI H5N1 outbreaks in Thailand between January 2004 and November 2005. An overall distribution of outbreak points was projected on the central and lower northern parts of Thailand. The four provinces with the highest risk were Phitsanulok, Phichit, Suphan Buri and Ang Thong. When each epidemic wave was considered, it was found that
Phitsanulok had the highest density area in the first and second waves, while Suphan Buri had a high density of H5N1 cases in the second and third waves. Other provinces, including Samut Prakan, Bangkok, Chon Buri, Ang Thong, Kamphaeng Phet and Saraburi only experienced one wave of HPAI H5N1 outbreak. The present study showed that most of the outbreaks occurred in the central and lower northern parts of Thailand (Figure 1), which confirmed previous studies by showing that these provinces were high-risk areas (Tiensin et al., 2007; Paul et al., 2010; Souris et al., 2010). Moreover, these results correlated with Gilbert et al. (2007) who indicated that high-risk areas have more free-grazing ducks due to the paddy fields in the regions. Based on the current analysis, the governmental authority should provide an early warning system and institute special disease control measures to prevent outbreaks in these areas. To this end, an active surveillance system must be implemented for disease detection.

Previous studies have used the empirical Bayes smoothing method to measure the incidence of HPAI H5N1 in Thailand (Tiensin et al., 2009). This technique is suitable for aggregated data; however, this study analyzed point data. Therefore,

**Figure 1**  Spatial distribution of HPAI H5N1 outbreaks in Thailand between January 2004 and November 2005 (1a–4a) and a comparison of kernel smoothing (1b–4b) in the: (1) whole period; (2) first wave; (3) second wave; and (4) third wave. The outbreak points are represented by different shades for each analysis.
Kernel smoothing was a more suitable method (kernel smoothing was not used for the adjusted disease rates because of the lack of high-quality data on the at-risk populations). Savill et al. (2006) used the kernel density based on Euclidean distance to detect the spatial pattern of transport links during the outbreak of foot and mouth disease in UK. They found that the multiplicity of transmission routes was a significant factor for the spreading of diseases. In terms of human epidemic diseases, Barreto et al. (2008) used these techniques to evaluate the pattern of the spreading dengue epidemic in Brazil. The results showed that the disease occurred in common areas. Furthermore, Chaikaew et al. (2009) predicted the areas where diarrhea would be prevalent in Chiang Mai, Thailand.

The advantage of the kernel smoothing method is that it allows analyses to be performed with optimal bandwidth. Hjort and Walker (2001) found that if a result lies outside the confidence interval around the empirical distribution function, then the probability tends toward one as the sample size increases. To avoid this problem, the researcher should not provide an optimal bandwidth when analyzing the data.

Local point clustering

Table 1 presents the results for Kulldorff’s spatial scan statistics at the district level. Throughout the study period, 10 disease clusters were found in various regions of Thailand. The highest clustering or the most likely cluster was cluster 1, which was located in the Ko Sichang district in Chon Buri province. This cluster had a high strength of association (relative risk >100, $P < 0.001$), which means that a member district in the cluster has a significantly greater risk of being infected (that is, more than 100-times greater risk). Furthermore, this was an isolated district with 81 outbreaks during the study period. It is important to note that the disease is easier to control when the outbreaks occur in isolated districts. However, when the avian population of the area was considered, only 626 birds in cluster 1 were found. Therefore, the number of outbreaks in Ko Sichang seemed very high given the low avian population. Alternatively, there could have been a problem in the data collection for the avian census. The local officials may have misunderstood the case definition. For example, they may have counted a chicken as one outbreak equally, or one geographical point may have had multiple outbreaks during the same period. These factors could explain the relative high values of risk observed. This point should be considered in the disease reporting system. Most of the other significant clusters were located in districts in the central part of Thailand; however, some were located in the lower northern part of the country and elsewhere. The radii of the clusters ranged from 11.64 km to more than 100 km, except for clusters 6 and 10. The disease was more widely spread in the clusters with larger areas. The number of poultry in the different areas varied from 626 to 6,956,465 birds. Some areas had lower numbers of infected districts but higher numbers of outbreaks with smaller populations. These areas may have had more backyard poultry. Therefore, control measures for outbreaks should be different in areas that have more commercial farms.

For the first wave, four clusters were located in the central part of Thailand connected to the lower northern part of Thailand. The most likely cluster was cluster 1, which was located in the Bangkoknoi district in Bangkok province and the Bangphi district in the Samut Prakan province with a 31.47 km radius. This cluster had a relative risk that was greater than 100 ($P < 0.001$) and the interpretation was similar to that previously described. In addition, districts in one cluster could share risk factors with districts in another cluster. Furthermore, the number of poultry in the first wave varied from 9,196 to 371,236 birds.

In the second wave, six clusters were found throughout Thailand. The most likely cluster was cluster 1, which was located in the
Ko Sichang district in Chon Buri province. This cluster had a relative risk that was greater than 100 ($P < 0.001$) and one outbreak. In addition, the number of poultry in the second wave varied from 626 to 15,206,646 birds. In the third wave, three clusters were found, most of which were located in the central part of Thailand. The most likely cluster was cluster 1, which included eight districts in three provinces with a radius greater than 100 km. This cluster had a relative risk value that was greater than 100 ($P < 0.001$) and 29 outbreaks. Furthermore, the number of poultry in the third wave varied from 312,999 to 889,509 birds (Table 1).

Previous studies have shown that most clusters at the subdistrict level were located in the central part of Thailand (Tiensin et al., 2009). The current study confirmed these results. The outbreak points in the same cluster indicate that each point could share the same risk factors of disease occurrence. In terms of disease control and prevention, the relevant authorities should implement control measures for all outbreak points in the cluster. For example, the DLD

<table>
<thead>
<tr>
<th>Study period</th>
<th>Cluster number</th>
<th>Number of districts in the cluster</th>
<th>Number of districts in the outbreak</th>
<th>Number of poultry</th>
<th>Cluster radius (km)</th>
<th>Relative risk</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 04 – Nov 05</td>
<td>1</td>
<td>1</td>
<td>81</td>
<td>626</td>
<td>0</td>
<td>&gt;100</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>(whole period)</td>
<td>2</td>
<td>47</td>
<td>260</td>
<td>6,264,557</td>
<td>&gt;100</td>
<td>6.02</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>8</td>
<td>42</td>
<td>31,731</td>
<td>19.61</td>
<td>&gt;100</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>7</td>
<td>108</td>
<td>2,578,012</td>
<td>27.29</td>
<td>5.37</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>42</td>
<td>165</td>
<td>6,956,465</td>
<td>&gt;100</td>
<td>3.11</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1</td>
<td>32</td>
<td>626,535</td>
<td>0</td>
<td>6.21</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>5</td>
<td>11</td>
<td>151,986</td>
<td>22.44</td>
<td>8.67</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>2</td>
<td>12</td>
<td>244,637</td>
<td>11.64</td>
<td>5.88</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>12</td>
<td>31</td>
<td>1,423,489</td>
<td>&gt;100</td>
<td>2.63</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>1</td>
<td>9</td>
<td>182,500</td>
<td>0</td>
<td>5.90</td>
<td>0.024</td>
</tr>
<tr>
<td>Jan 04 – Jun 04</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>13,868</td>
<td>31.47</td>
<td>&gt;100</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>(first wave)</td>
<td>2</td>
<td>6</td>
<td>6</td>
<td>371,236</td>
<td>68.48</td>
<td>12.54</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>68,138</td>
<td>68.20</td>
<td>14.62</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>9,196</td>
<td>40.95</td>
<td>53.27</td>
<td>0.003</td>
</tr>
<tr>
<td>Jul 04 – Jun 05</td>
<td>1</td>
<td>1</td>
<td>81</td>
<td>626</td>
<td>0</td>
<td>&gt;100</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>(second wave)</td>
<td>2</td>
<td>71</td>
<td>375</td>
<td>15,206,646</td>
<td>&gt;100</td>
<td>3.96</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>45</td>
<td>164</td>
<td>5,839,376</td>
<td>&gt;100</td>
<td>3.62</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>7</td>
<td>21</td>
<td>17,862</td>
<td>19.61</td>
<td>&gt;100</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>2</td>
<td>12</td>
<td>244,637</td>
<td>11.64</td>
<td>5.57</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1</td>
<td>9</td>
<td>182,500</td>
<td>0</td>
<td>5.59</td>
<td>0.040</td>
</tr>
<tr>
<td>Jul 05 – Nov 05</td>
<td>1</td>
<td>8</td>
<td>29</td>
<td>662,859</td>
<td>&gt;100</td>
<td>8.25</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>(third wave)</td>
<td>2</td>
<td>7</td>
<td>12</td>
<td>312,999</td>
<td>&gt;100</td>
<td>5.60</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>16</td>
<td>889,509</td>
<td>22.62</td>
<td>2.63</td>
<td>0.045</td>
</tr>
</tbody>
</table>

Table 1 Results of Kulldorff’s spatial scan statistics showing significant clusters of HPAI H5N1 outbreaks in Thailand between January 2004 and November 2005.
must destroy all poultry in the cluster within the radius indicated in Table 1. These results may indicate the benefits of using spatial tools for disease control. However, the authorities must also consider control measures for the entire country. The results of both studies are valuable for improving knowledge and understanding of the spatial patterns of HPAI H5N1 by indicating that the disease is likely to occur in specific areas. In addition, Tiensin et al. (2009) suggested that the disease may spread across Thailand during the first wave, but is likely to spread locally thereafter. In a similar way, Viel et al. (2000) use Kulldorff’s statistic to evaluate the cases of soft-tissue sarcoma and non-Hodgkin’s lymphoma in France. They found that the most likely cluster was located around the municipal solid waste incinerator. In veterinary science, researchers have conducted similar studies on infectious diseases. For example, Schwermer et al. (2007) detected clusters of bovine spongiform encephalopathy surrounding feed producers. However, a purely spatial analysis should be applied to future analyses, because the method is not suitable for calculating many years of data due to the low power of statistical analysis. This is especially true for detecting diseases in an emerging cluster. Space-time scan statistics are suitable for analyzing many years of data (Kulldorff, 2001).

Future studies should focus on a disease pattern that makes use of advanced spatial methods. Furthermore, it is necessary to construct a model for disease prediction that can be used to describe spatial variability. An example is the Bayesian spatial regression model, which introduces prior knowledge to a data-driven property. Indeed, information on the risk of an outbreak can be defined using existing data.

CONCLUSION

Two techniques were used to confirm that the central and connected lower northern areas of Thailand were high-risk areas. The focus was on two provinces—Suphanburi and Phitsanulok—that require an early warning system. The human network for the rapid reporting and monitoring of diseases is an easy and effective potential method. The cluster detection demonstrated that disease control measures must be improved. In particular, the radius around the infected area for preemptive culling and restricted animal movement should be varied in different areas because the disease pattern is usually spatially different. The results from this study will be useful for effectively controlling HPAI H5N1 in Thailand.

ACKNOWLEDGEMENTS

The authors would like to thank Assistant Professor Suwicha Kasemsuwan for her helpful advice and also the Department of Livestock Development of Thailand for access to the database and their cooperation. This work was supported by the “Strategic Scholarships for Frontier Research Network” and the Commission of Higher Education, Thai Ministry of Education.

LITERATURE CITED


Tiensin, T., P. Chaitaweesub, T. Songserm, A. Chaisingh, W. Hoonsuwan, C. Buranathai,

