APPLICATION OF THE PROSPECTIVE SPACE-TIME SCAN STATISTIC FOR DETECTING MALARIA CASES HOTSPOTS IN BANGKA DISTRICT, INDONESIA

Asep Saefuddin[1] and Etih Sudarnika[2]
[1]Department of Statistics, Faculty of Mathematics and Science, IPB, 16680, Darmaga, Bogor, Indonesia
[2]Laboratory of Epidemiology, Faculty of Veterinary Medicine, IPB, 16680, Darmaga, Bogor, Indonesia
E-mail: [1]asaefuddin@gmail.com, [2]etih23@yahoo.com

ABSTRACT

Malaria is one of the complex health problems in Indonesia. The disease is still listed high priority due to its high mortality rate, especially among children under five years old, and its fatal impact on pregnant woman. The aim of this study was to apply cluster detection method for detecting disesase outbreaks in the malaria surveillance system in Bangka District, Bangka Belitung Province, Indonesia. Bangka is one of the malaria endemic areas in Indonesia. The cluster detection method used is the Prospective Kulldorff's space-time scan statistic. The research was conducted in a one year period, starting from September 2007 to September 2008. The malaria cases data were recorded from eleven public health center facilities in Bangka Districts, based on confirmed lab examinations. The results showed that the primary emerging cluster was Sinar Baru with a cluster period from January 2008 to July 2008. The secondary emerging cluster was Kenanga with a cluster period from January 2008 to July 2008. The conclusions were Sinar Baru and Kenanga were the areas which need further investigation and priority in the disease control and surveillance.

Keywords: malaria, space-time clustering, spatial statistics

1. INTRODUCTION

Background

Malaria is an infectious disease caused by a parasite from the *plasmodium* genus and is spread through the bite of the *Anopheles* mosquito. Its general symptoms include periodic fevers, anemia, bubonic, and other symptoms related to its effect on other organs such as the brain, liver, and kidneys (WHO 2007, Wikipedia 2007).

Indonesia is a country endemic of malaria. Indonesia is a tropical region with high rainfall and swampy topography which then enhance the life cycle of the Anopheles mosquito as the vector of malaria (Jamal, 2009). According to the Indonesian Malaria Endemic Map, it is predicted *that* in 2007 45% of Indonesia's population live in areas endemic with malaria. One of the areas endemic with malaria is the Bangka district, Bangka Belitung province. This area is categorized

as medium endemic with malaria with an AMI of 29.3 per 1000 person in the year 2007 (Indonesian Health Department, 2008).

Generally, malaria is endemic in small villages with poor sanitation, bad transportation and communication facilities, hard to access healthcare services, low social economy levels, and unhealthy community lifestyles. Locations and social conditions like these are not uncommon in Indonesia. Aside to activities such as early diagnosis, quick and proper medication, *fogging,* and larva and vector control, surveillance is also an essential activity in the effort to reduce the number of cases and deaths caused by malaria.

Through surveillance, quick preventive actions can be conducted in certain malaria hotspots to better control the disease. Thus, it is necessary to conduct a geospatial statistical analysis to help analyze the surveillance data in order to find out when and where serious actions must be conducted, and so that the area gets first priority in handling malaria.

The objective of this research was to conduct a *space-time scan statistic* to detect malaria hotspots in Bangka district, Indonesia.

2. RESEARCH METHODS

2.1. Study Area

The Bangka district is located on Bangka Island covering an area of approximately 295.068 Ha populated by 237.053 people or 80 people per sq km. It has a tropical climate with rainfalls between 18.5 to 394.7 mm each month, reaching its lowest point in August. The temperature ranges between 26.2^0C to $28,3^0$C and the humidity varies from 71% to 88%. The light intensity is around 18% to 66.1% and the air pressure varies from 1009.1 to 1011.1 mb.
The mining sector is one of Bangka's main sectors. The district is rich with tin and other mining goods and has a relatively high reserve. Thus, Bangka district has many abandoned mines containing ponds and puddles resulting from former mining activities. These ponds and puddles are potential breeding grounds for mosquitoes. The Bangka district has different topography, consisting of 4% hills, 25% swamps, and lowlands for the rest of the areas (Central Statistics Agency (BPS) and the Bangka District Regional Development Agency (BAPEDA) 2007).

Overall, Bangka's climate, topography, and environmental conditions are suitable for sustaining the *Anopheles* mosquito life cycle.

2.2. Data Collection and Management

Data collection was done monthly by collecting lab examination notes from every public health center (*puskesmas*) in the Bangka district. There were 11 *puskesmas* working areas spread throughout 8 sub districts in the Bangka district. They were *Puskesmas* Sungai Liat, Sinar Baru and Kenanga in sub district Sungai Liat, *Puskesmas* Bakam in sub district Bakam, *Puskesmas* Petaling in sub district Mendo Barat, *Puskesmas* Gunung Muda and Belinyu in sub district Belinyu, *Puskesmas* Pemali in sub district Pemali, *Puskesmas* Riau Silip in sub district Riau Silip, *Puskesmas* Puding Besar in sub district Puding Besar, and *Puskesmas* Batu Rusa in sub district Merawang. The definition of a malaria case in the study was a person positively diagnosed with malaria after a *Plasmodium* parasite lab check. Data resulting from clinical examinations and *Rapid Diagnostic Tests* were not considered as cases. If a *puskesmas* had a

positive case but the person did not live in the area, then the case was moved to the *puskesmas* where the person lives. Each *puskesmas* location was mapped using Geography Positioning System (GPS) technology. The data used was the malaria cases that occured from June 2007 to July 2008. The demographic data used was obtained from the Central Statistics Agency (BPS) and the Bangka district Regional Development Agency (Bapeda). Every person in the area is assumed possible of being infected by malaria.

2.3. Statistical Analysis

The geospatial analysis used in this research is the *prospective space-time scans statistics* developed by Kulldorff in 1997. The data was assumed to spread in a Poisson spread and was analyzed using *SatScan* 7.0 software. Kulldorff 2001 mentioned that the prospective space-time scan statistics uses a cylindrical window in three dimensions. The base of cylinder represents space, whereas height represents time. The cylinder is flexible its circular geographical base as well as in its starting date, independently of each other. In mathematical notation, let $[Y_1, Y_2]$ be the time interval for which data exist, and let s and t be the start and end dates of the cylinder respectively. We then consider all cylinders for which $Y_1 \leq s \leq t \leq Y_2$ and $t \geq Y_m$, where Y_m is the time periodic surveillance began. Conditioning on the observed total number of cases, *N*, the definition of the spatial scan statistic S is the maximum likelihood ratio over all possible circles Z,

$$S = \frac{\max_{z}\{L(Z)\}}{L_0} = \max_{z}\left\{\frac{L(Z)}{L_0}\right\} \qquad (1)$$

where $L(Z)$ is the maximum likelihood for circle Z, expressing how likely the observed data are given a differential rate of events within and outside the zone, and where L_0 is the likelihood function under the null hypothesis.

Let n_z be the number of cases in circle Z. For the Poisson model, let $\mu(Z)$ be the expected number under the null hypothesis, so that $\mu(A) = N$ for A, the total region under study. It can then be shown that

$$\frac{L(Z)}{L_0} = \left\{\frac{n_z}{\mu(Z)}\right\}^{n_z}\left\{\frac{N - n_Z}{N - \mu(Z)}\right\}^{N - n_Z} \qquad (2)$$

if $n_Z > \mu(Z)$ and $L(Z)/L_0 = 1$ otherwise.

3. RESULTS AND DISCUSSION

With its swampy and water surrounded ecology and also the presence of palm plantations and abandoned tin mines, Bangka district is a potential breeding ground for the *Anopheles* mosquito. Incidences of malaria vary according to *puskesmas* working area and observation period. Image 1 shows malaria incidence per 10000 people per year in the 11 *puskesmas* monitored from June 2007 until July 2008. The darker the color gradation, the higher the malaria incidence occur. During the observation period, the Bangka district had a total 1.778% malaria incidence. The area with the highest incidence is *puskesmas* Sinar Baru with 3.449% followed by *puskesmas* Kenanga with 3.138%. These two *puskesmas* are surrounded by beaches, an area potential for *Anopheles* breeding.
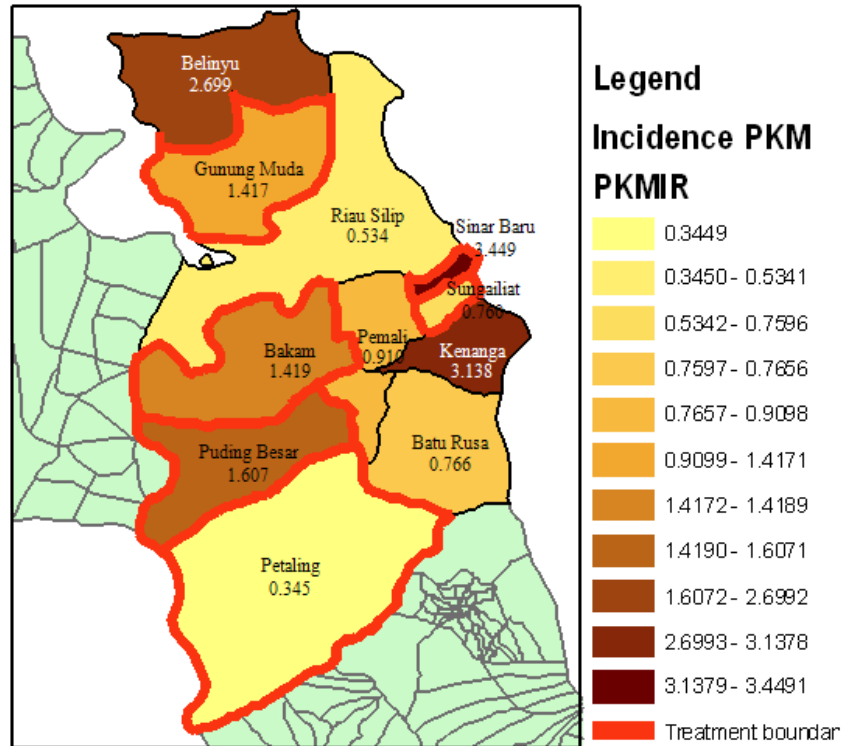
Image 1. Malaria Incidence per 10000 people in Bangka district *Puskesmas*,
June 2007 – July   2008

The fluctuation of malaria incidence in every *puskesmas* working area in the Bangka district through June 2007 to July 2008 is shown in image 2. In the beginning of the observation, the malaria incidence in *puskesmas* Belinyu shows a significantly high value compared to the other regions, but the value declines during the rest of the observations. Other areas with high malaria incidence are Sinar Baru and Kenanga.

The result of the surveillance on malaria incidence per month during the observation period is spatially shown in image 3. The darker gradation shows areas with higher malaria incidence. Image 3 shows that spatial and temporal analysis can quickly be done by observing the color changes in each district every month. For example, in the beginning of the observation, Belinyu has the highest malaria incidence value, but shows a decline until the end of the observation period. The opposite occurs in Bakam, having a low value in the beginning but increasing until the end of the observation period.

Aside to descriptive data analysis by visualizing incidence maps, statistical analysis is also needed to determine malaria hotspot areas for immediate preventive actions. The statistical analysis was done by using a prospective space time scan statistic. The scan statistic is a statistical method designed to detect a local excess of events and to test if such an excess can reasonably have occurred by chance. Three basic properties of the scan statistic are the geometry of the area being scanned, the probability distribution generating events under the null hypothesis, and the shapes and sizes of the scanning window (Kulldorff 1997). The spatial scan statistic has the following features, which make it particularly suitable as a tool for surveillance purposes:
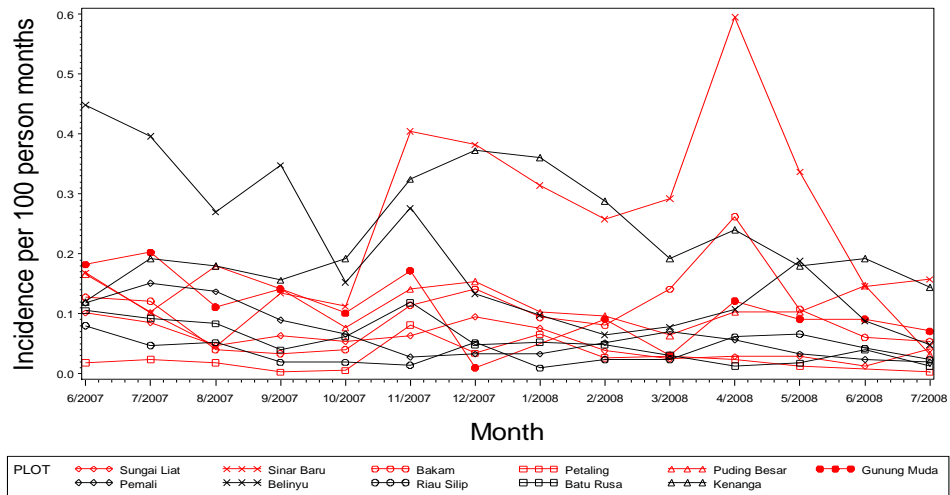
251

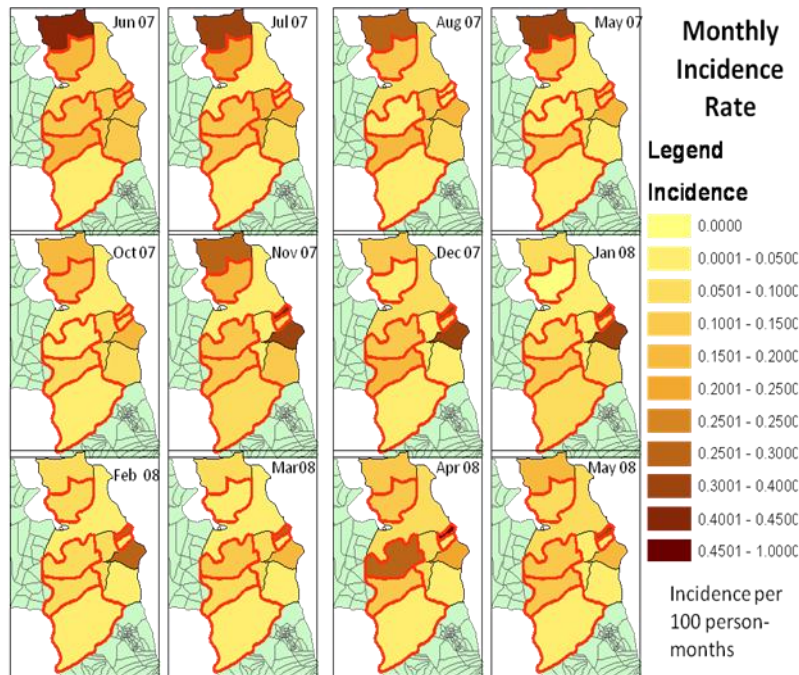**Image 2**. Malaria Incidence per 100 people in Bangka District, June 2007 – July 2008

Image 3. Malaria Incidence per Month in every *Puskesmas* in Bangka district

a) it adjusts both for the inhomogeneous population density and for any number of confounding variables;
b) by searching for clusters without specifying their size or location the method ameliorates the problem of pre selection bias;
c) the likelihood-ratio-based test statistic takes multiple testing into account and delivers a single p-value for the test of the null hypothesis;
d) if the null hypothesis is rejected, we can specify the approximate location of the cluster that caused the rejection. (Kulldorff 2001).

252

By using a prospective space-time scan statistic analysis, observation period from June 2007 to July 2008, time aggregation length 1 month, maximum spatial cluster size 50% of population at risk, and maximum temporal cluster size 50% of study period, the results are shown in table 1.

Table 1. Detection of recently emerging clusters of malaria in Bangka

Most likely cluster	Cluster period	Cases	Expected	Relative risk	p
Primary cluster:					
Sinar Baru	2008/1/1 - 2008/7/31	187	56.76	3.455	0.001
Secondary cluster:					
Kenanga	2008/1/1 - 2008/7/31	133	53.05	2.581	0.001
Bakam	2008/3/1 - 2008/7/31	93	68.12	1.378	0.145

Table 1 show that the first emerging cluster is *puskesmas* Sinar Baru with a cluster period from January 1st 2008 to July 31st 2008 and a relative risk value 3.455. The second emerging cluster is *Puskesmas* Kenanga with 133 cases and a relative risk value 2.581. Another emerging cluster is *Puskesmas* Bakam, but this cluster does not have a significant p-value.

The working areas of *puskesmas* Sinar Baru and Kenanga are located in lowlands mainly consisting of coasts and swamps. These areas also have a lot of abandoned and unmaintained unauthorized tin mines that contain many ponds and puddles, the potential breeding ground for mosquitoes. These topographic conditions along with the residential areas neighboring with mosquito breeding grounds are the main cause of the high malaria incidence in these areas. The Bakam *puskesmas* working area shows low malaria incidence rates in the beginning of the observation period. The rates then increase in the following observations reaching its highest in April 2008, but there is insufficient data to categorize the area as an emerging cluster.

Quantitative analysis using prospective space time scan statistics allows us to discover the emerging clusters that need attention based on the surveillance results so that the authority can necessary priority actions in controlling the disease. The methods used to prevent and control malaria Includes installing mosquito nets, covering the body while sleeping, and using natural mosquito repellants. Nets with insecticides are available for children and pregnant woman in the Bangka district. The nets are distributed by UNICEF and the Health Department through *puskesmas, posyandu* (Integrated Community Health Service Center)*,* and birth Clinics. Other methods include exterminating mosquito nests by cleaning up puddles and conducting *Indoor Residual Spraying* (RIS).

Kulldorf 2001 states bahwa the space-time scan statistic can serve as an important tool for systematic time periodic geographical disease surveillance. It is possible to detect emerging clusters, and we can adjust for multiple tests performed over time. No a priori hypothesis about cluster location, size or length need to be made. The method can be used at different levels of geographical and temporal aggregation, and for different types of disease. Although it is computer intensive, the method is not overly complex.

4. CONCLUSION

The primary emerging cluster is Sinar Baru with a time frame starting from January 2008 until July 2008. The secondary emerging cluster is Kenanga with a time frame starting from January

2008 until July 2008 and Bakam with a time frame starting from March 2008 until July 2008, but is not significant for Bakam. Sinar Baru and Kenanga were the areas which need further investigation and priority in the control.

ACKNOWLEDGMENTS

REFERENCES

Badan Pusat Statistik Kabupaten Bangka (2007). *Bangka Dalam Angka 2006*. Bangka: BPS & BAPEDA Kab. Bangka.

Jamal S. (2009). Apa yang perlu anda ketahui tentang malaria?

http://www.isfinational.or.id/pt-isfi-penerbitan/124/443-apakah-malaria-itu- [3/16/2009].

Kulldorff, M. (1997). *A Spatial Scan Statistic*. Communs Statist. Med., **26**. 1481-1496.

Kulldorff, M. (2001). *J. R. Statist. Soc. A*, **164**. Part 1. 61-72

Ministry of Health of Republik Indonesia (2008). *Indonesian Health Profile*. Jakarta: MOH RI.

World Health Organization (2007). *Malaria*. http://www.who.int/ topics/malaria/en/ [6/23/22007]

Wikipedia (2007). Malaria. http://en.wikipedia.org/wiki/Malaria [6/23/22007]