ISSN 2277 – 5528
Impact Factor- 5.085

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES &MANAGEMENT

SPATIAL CLUSTERING OF UPPER RESPIRATORY TRACT INFECTIONS IN BANDUNG CITY BY MEANS LOCAL MORAN'S I AND SPATIAL SCAN STATISTICS

I Gede Nyoman Mindra Jaya[*1], Bertho Tantular[2] & Zulhanif[3]
[1,2&3]Statistics Department, Padjadjaran University, Bandung, Indonesia

ABSTRACT

Upper respiratory tract infection (URTI) is a serious problem for public health in Bandung city. This disease is caused by bocavirus and the number incidence increase in the rainy season every year. The number of incidences is found in all sub-districts and some of the districts have a high relative risk with a value higher than 1.The early warning system for controlling the disease spread is needed. In this study, we introduce the Local Moran's I and Spatial Scan statistics to detect the high-risk spatial cluster of URTI. We found the western and eastern regions are riskier than the northern and southern district.

Keywords: *Local Moran's I, Relative Risk, Spatial Scan Statistics, URTI*

I. INTRODUCTION

Upper respiratory tract infection (URTI) is a serious problem for public health caused by boca virus. Both of Children and adult can be infected by boca virus. URTI is not only health but also an economic problem. It has a cost to society for treatment and covers for absenteeism from school and works (Cotton, et al. 2004).

URTI is one of the infectious diseases in Bandung, West Java-Indonesia that has become a serious concern of the government due to the rapid transmission and high numbers of incidence of the disease (Ratnadewi, 2015). In 2016 is found 12,579 much higher than the other city in West Java, Indonesia. The high incidence of URTIs in Bandung is caused by the high population density and high mobility of the population.

Develop an early warning system is needed to reduce the negative impact of URTI and controlling the spread of this disease. Early warning system must be developed based on knowledge about the spatial distribution of the URTI disease. The advantage of knowing the spatial distribution of diseases such as provides information about the high risk-low risk regions. This information can be used to design the effective treatment for controlling the spread of the disease. Several methods were developed to analyze the spatial distribution of a particular disease such as disease mapping and clustering studies (Jaya, et al. 2017). In this study we focus on disease clustering technique to find the high risk cluster of URTI disease. Disease cluster may help to know that the high risk areas geographically clustered or randomly spread. Spatial cluster detection is an important tool in URTI surveillance to identify areas of elevated risk.

Several method have been introduced for disease clustering such as Local Moran's I (Anselin, 1995) and Spatial Scan Statistics (Kulldorff, 1997). These method should be applied together for better result.

The structure of the remainder of this paper is as follows; Section 2 presents the material and methods. Section 3 present application for URTI disease data in Bandung City and section 4 presents the conclusion.

II. MATERIAL AND METHOD

2.1. Study Area

Bandung city is the capital city of West Java province in Indonesia. It has an area around 167.73 km. Total population in 2016 is 2,575,477 distributed in 30 sub-districts. Bandung is the most populous city in West Java with a variety of social and economic activities(Bandung, 2018). Economic activities affect the movement of people in Bandung city. The high population movement, density and high annual rainfall is very influential in the easy transmission of various diseases where the people may be easily infected by various infectious diseases. There are several serious infectious diseases are found especially in rainy season-dengue fever, dengue fever, diarrhea and upper respiratory tract infection(Sari, 2018). In this study we focus on the analysis spatial clustering of (URTI) using spatial data in 2016.

2.2. Relative Risk

The common statistics in disease study to measure the high-low risk area is relative risk. This statistics is defined as ratio of number incidence on expected rate(Jaya, Folmer, Ruchjana, Kristiani, & Yudhie, 2017):

$$r_i = \frac{y_i}{e_i}; i = 1, \dots, m \tag{1}$$

where r_i the value of relative risk r at location i; y_i denote the number of cases and e_i is expected rate which defined as:

$$e_i = N_i \times \frac{\sum_{i=1}^{m} y_i}{\sum_{i=1}^{m} N_i} \tag{2}$$

with N_i is the number population at risk at location i

2.3. Disease Clustering

Several method have been introduced for disease clustering such as Local Moran's I (Anselin, 1995) and Spatial Scan Statistics(Kulldorff, 1997).

Local Moran's I

URTI hotspots can be clustered by means local autocorrection function. In this study, spatial clusters of URTI would be locations with a high relative risk of URTI surrounded by other location with a high relative risk. It can be identified using Local Moran's I (Anselin, 1995), (Getis & Ord, 1996):

$$I_i = \frac{r_i - \bar{r}}{s^2} \sum_{j \sim i}^{m} \left[w_{ij} \left(r_j - \bar{r} \right) \right] \tag{3}$$

where$\bar{r}$ – is the average of the relative risk over m sub districts, r_j is the value of the relative risk at all other locations (where $j{\sim}i$ means location i share border with location j); $s^2$is the variance estimate of relative risk; and w_{ij} weight which can be defined as distance or contiguity function (Zhang, Luo, Xu, & Ledwith, 2008). A high positive local Moran's I value indicates that the location interest has similarly high or low values as its neighbours, thus the locations are spatial clusters. A high negative local Moran's I value indicates that the location interest is a spatial outlier. Spatial cluster by means local Moran's index include high-high clusters (high values in high values neighbourhood) and low-low (low values in low values neighbourhood). Spatial outlier include high-low (high values in a low value neighbourhood), low-high (low values in high values neighbourhood). To obtain the spatial cluster and outlier, p-value generally used. P-values are obtained based on hypothesis testing which can be calculated by mean normal assumption or permutation test (Anselin, 1995).

Spatial Scan Statistics

Spatial scan statistics was introduced by (Kulldorff & Nagarwalla, 1995) and (Kulldorff, 1997)for the Bernoulli and Poisson Model. This is a spatial version of the scan statistics with a variable windows size and is a generalization of cluster evaluation permutation procedure(Tango, 2010). The objective of spatial scan statistics is to detect hot-spot cluster within the study interval divided into several non overlapping interval and do the significant testing.

Null hypothesis and alternative hypothesis for spatial clustering. In order to test the significant of spatial cluster let we assumes that:

$$(y_1, \ldots, y_m) \sim Multinomial(n, \boldsymbol{p})\,;\, \boldsymbol{p} = (p_1, \ldots, p_m) \tag{4}$$

where m is number of regions with n is number of trial and p_i is probability to draw y_i which defines as:

$$p_i = \frac{N_i}{\sum_{i=1}^{n} N_i}; i = 1, . . , m \tag{5}$$

where N_i denotes the population at risk in subinterval ith and the expected rate of cases in subinterval ith is $e_i = np_i$. The null hypothesis and the alternative hypothesis is presented below:

$$H_0 : E[y(J)] = e(J), \text{for all } J$$
$$H_1 : E[y(J)] > e(J), \text{for some } J$$

with $J = [i, i + k - 1]$, an interval of length k, and $y(J)$ and $e(J) = \sum_{i \in J} np_i$ denote the random number of incidence and the null expected number of cases within the specified interval J respectively. Rejected H_0indicates there is an elevated risk within some interval J compare with outside (Tango, 2010).

Test statistics

The likelihood ratio maximized for J similar to Nagarwalla's scan statistics for continuous data is used as the test statistics.

$$\lambda = \sup_J \left(\frac{n(J)}{e(J)}\right)^{n(J)} \left(\frac{n - n(J)}{n - e(J)}\right)^{n-n(J)} I\left(\frac{n(J)}{e(J)} > \frac{n - n(J)}{n - e(J)}\right) \tag{6}$$

Where $n(J)$ and $e(J)$are the observed number and the null expected number of cases within the interval J and $I(\quad)$ is the indicator function. The interval J^* that attains the maximum likelihood is defined as the most likely cluster(Tango, 2010).

Null distribution

The null distribution of λ is obtained by mean Monte Carlo hypothesis and package SatScan from Kulldorf can be used.

III. RESULTS

3.1. Statistics of URTI

Table 1 provides information on URTI incidence in Bandung city. The median URTI cases was 301 with maximum incidence is 1435. The maximum incidence of URTI is found in sub district Andir.

Table 1. Statistics of URTI Incidence

m	Minimum	Median	Maximum
30	51	301	1435

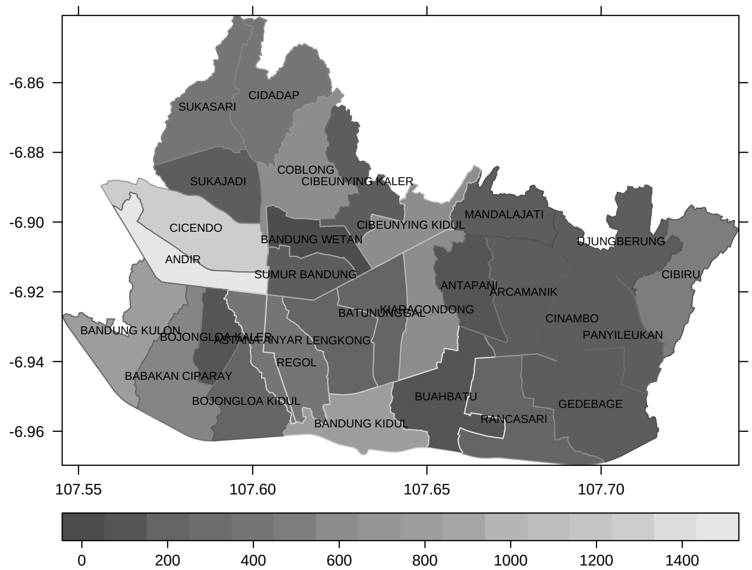

Figure 1. Spatial distribution of URTI Incidence

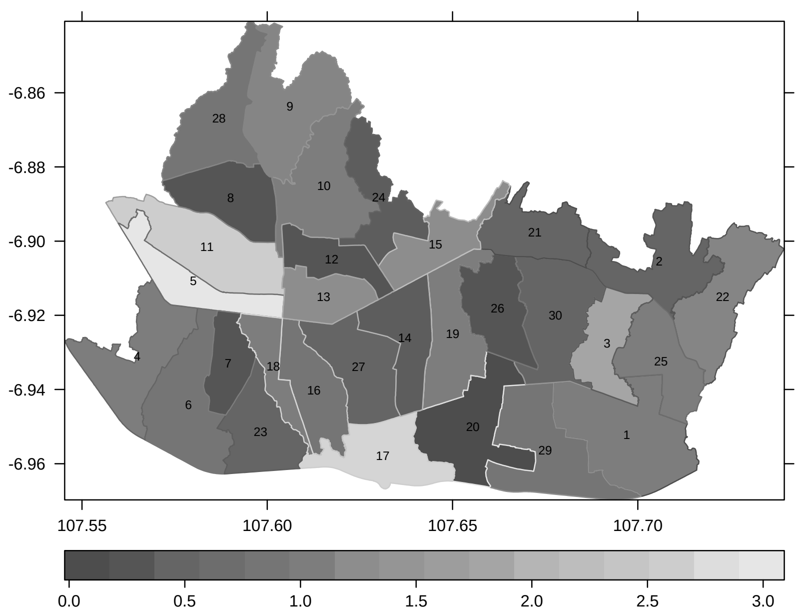

Figure 2. Spatial distribution of the relative risk of URTI

Figure 1 and 2 show the spatial distribution of URTI incidence and the relative risk respectively. Figure 2 clearly present that there are three sub districts which have high relative risk (11=Cinambo, and 5=Andir, 17=Bandung Kidul). Spatial cluster analysis is needed to present the significant spatial cluster which have high risk of URTI.

3.2. Statistics of Local Moran's I

Table 2. Statistics of Local Moran's I

ID	Sub District	Ii	Pr(z > 0)
1	Gedebage	0.0070	0.4630
2	Ujungberung	-0.0324	0.4979
3	Cinambo	-0.2664	0.6979
4	Bandung Kulon	0.2134	0.3523
5	Andir	0.5259	0.0401
6	BabakanCiparay	-0.0964	0.5550
7	BojongloaKaler	-0.6298	0.9084
8	Sukajadi	-0.6524	0.9164
9	Cidadap	-0.1072	0.5551
10	Coblong	-0.0048	0.4699
11	Cicendo	0.5423	0.0712
12	Bandung Wetan	-0.2999	0.7742
13	Sumur Bandung	0.1784	0.2531
14	Batununggal	-0.0976	0.5781
15	CibeunyingKidul	-0.2976	0.7483
16	Regol	-0.1571	0.6225
17	Bandung Kidul	-1.2390	0.9997
18	Astana Anyar	0.0475	0.4081
19	Kiaracondong	-0.0741	0.5401
20	Buahbatu	0.0573	0.3973
21	Mandalajati	0.2244	0.2551
22	Cibiru	-0.1146	0.5488
23	BojongloaKidul	-0.1096	0.5666
24	CibeunyingKaler	0.1072	0.3937
25	Panyileukan	0.0082	0.4620
26	Antapani	0.6911	0.0524
27	Lengkong	-0.2146	0.6564
28	Sukasari	0.0732	0.4346
29	Rancasari	0.1535	0.3602
30	Arcamanik	0.2274	0.2068

Table 2 presents the statistics of Local Moran's I. Some sub-districts have positive local Moran's I which indicates the sample location is surrounded by high-risk sub-district. There are two sub-district which indicate have significant high risk of URTI and p-value lower than 0.1 (Andir and Cicendo) and one district indicates has a significantly low risk is Antapani. The spatial cluster clearly sees in Figure 3 below.

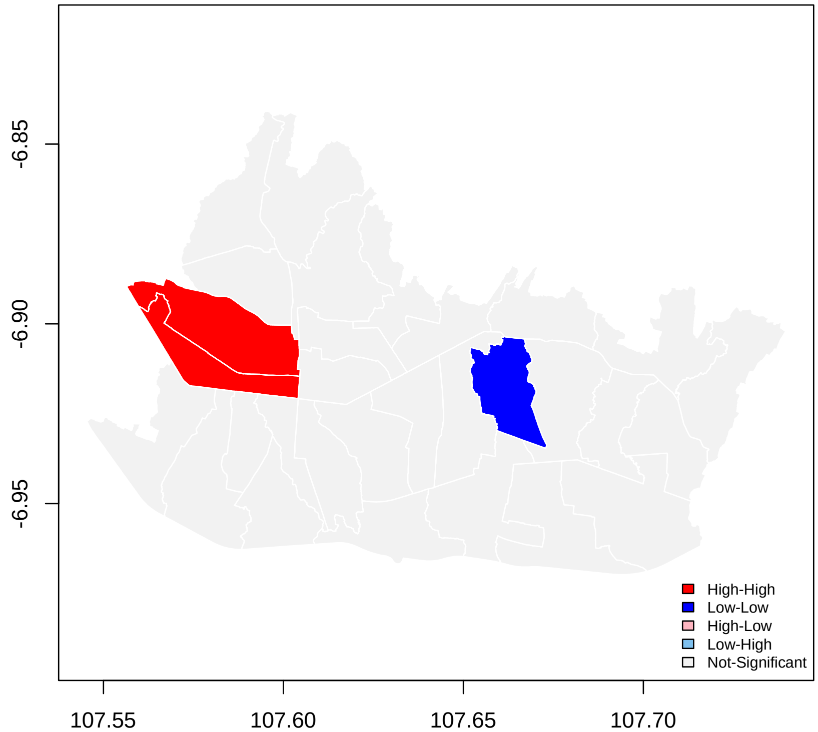

Figure 3. Spatial Cluster by Means Local Moran's I

Figure 3 present the spatial cluster by means Local Moran's I. Local Moran's I informing two different spatial clusters. The red colour is the high risk spatial cluster and the blue one is low risk spatial cluster. Two sub-districts in high risk spatial cluster are Cicendo and Andir. While the low spatial cluster consists one sub-district (Antapani). The clusters is created based on the significant value of the Local Moran's Index.

3.3. Statistics of Spatial Scan Statistics

Table 2. Statistics of spatial cluster by mean spatial scan statistics.

Cluster	Region	Number of Regions	RR	Radius (km)	Log Likelihood Ratio	p-value
1	11, 5	2	3.20	1.09	1142.932	0.000
2	17	1	2.82	0	307.308	0.000
3	25, 3, 1, 22	4	1.30	2.66	32.241	0.000
4	15	1	1.37	0	31.961	0.000
5	13	1	1.36	0	7.487	0.011
6	4	1	1.15	0	7.487	0.011
7	9	1	1.22	0	6.330	0.019
8	19	1	1.10	0	3.304	0.297
9	10	1	1.09	0	2.320	0.616
10	18	1	1.08	0	0.947	0.991

Table 2 shows the statistics of spatial cluster by mean spatial scan statistic (SatScan). There are 14 sub district which have relative risk higher than one and grouping become 10 spatial clusters. Cluster 1 is consisted 2 districts (Cicendo and Andir) and have highest average relative risk i.e. 3.20. This cluster is located at western region of Bandung (small blue circle in Figure 4). The biggest cluster is cluster 3 which consisted 4 sub-districts (Panyileukan, Cinambo, GedeBage, and Ujung Berung). This cluster is located at eastern regions of Bandung (Big blue circle in Figure 4). From 10 different clusters, there are seven cluster are significant with average relative risk differ from 1.

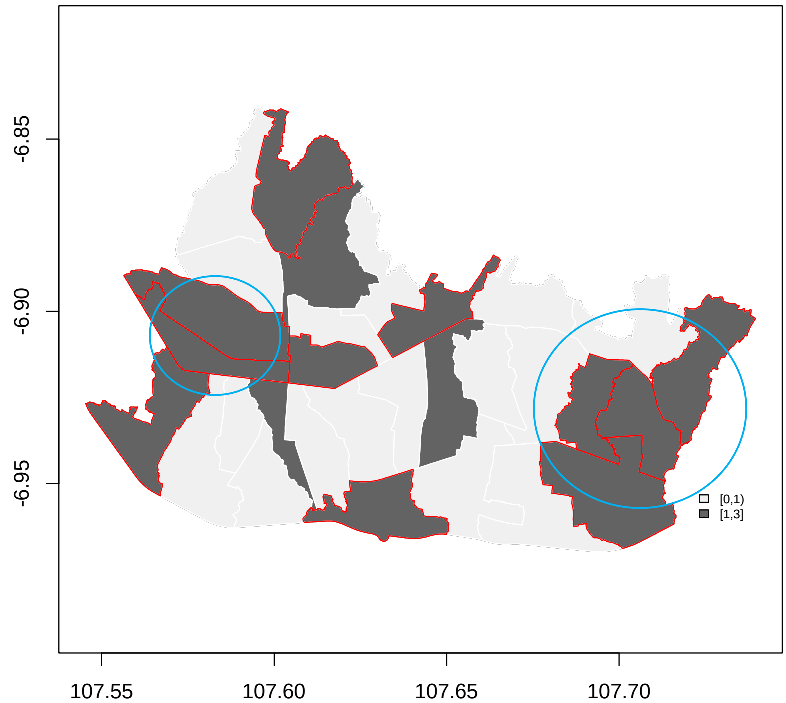

Figure 4. Spatial Cluster by Means Spatial Scan Statistics

Figure 4 shows the spatial cluster of URTI relative risk by means spatial scan statistics. There are 10 clusters which consisted 14 sub-districts which have relative risk higher than 1. The blue circle indicates the clusters consist more than one sub districts. While the sub-districts with red border line indicate the region with significant high risk of URTI. This figure informs that the sub-district where located in western and eastern regions of Bandung, indicated as high-risk sub-district. The western regions are the regions which have high population density and the eastern regions known as the polluted regions in Bandung. Two sub-districts which have high risk of URT are Cicendo and Andir. The spatial cluster information was obtained consistent with Local Moran's I.

IV. CONCLUSION

Spatial clustering is an important tool in epidemiology to detect the high-risk cluster of particular disease. Two methods are commonly used-local Moran's I and Spatial Scan Statistics. These methods are important to apply together because they have different information in case of disease clustering. Local Moran's I identify the spatial cluster based on the similarity of relative risk around sample locations. The high-risk cluster is defined as the location which surrounded by other high-risk location. Usually, it methods present the small clusters size. While spatial scan statistics may present the wide spatial cluster size of high risk location without considering the high risk neighbour location(Laohasiriwong, Puttanapong, & Luenam, 2018).These methods are applied to identify the spatial cluster of upper respiratory tract infection (URTI) which is a serious problem for public health in Bandung city. We found that some of the districts have a high relative risk with a value higher than 1.We found the western and eastern regions are riskier than the northern and southern district. The highest one are Cicendo and Andir. This information can be used to develop an early warning system in controlling the impact of URTI disease.

V. ACKNOWLEDGEMENT

This paper is funded by RFU Unpad contract: 1732 d/UN6.RKT/LT/2018. The authors thank Rector Universit as Padjadjaran and to the anonymous referee whose valuable checking has improved this paper

REFERENCES

1. *Anselin, L. (1995). Local indicators of spatial association – LISA. Geographical Analysis, 27, 93-115.*
2. *Bandung. (2018). Data Basis Pembangunan Daerah Kota Bandung 2018. Bandung: UniversitasPadjadjaran.*
3. *Cotton, Innes, Jaspan, Madide, & Rabie. (2004). Management of upper respiratory tract infections in children. S Afr Fam Pract, 50(2), 6-12.*
4. *Getis, A., & Ord, J. (1996). Local spatial statistics: an overview. In P. Longley, & M. Batty, Spatial Analysis: Modelling in a GIS Environment (pp. 261-277). England: Cambridge.*
5. *Jaya, I. M., Folmer, H., Ruchjana, B. N., Kristiani, F., &Yudhie, A. (2017). Modeling of Infectious Diseases: A Core Research Topic for the Next Hundred Years. In R. Jackson, & P. Schaeffer, Regional Research Frontiers - Vol. 2 Methodological Advances, Regional Systems Modeling and Open Sciences (pp. 239-254). USA: Springer International Publishing.*
6. *Kulldorff, M. (1997). A Spatial Scan Statistics. Commun. Statist.-Theory Meth., 26(6), 1481-1496.*
7. *Kulldorff, M., &Nagarwalla, N. (1995). Spatial Disesae Clusters: Detection and Inference. Statistics in Medicine, 14, 799-810.*
8. *Laohasiriwong, W., Puttanapong, N., &Luenam, A. (2018). A comparison of spatial heterogeneity with local cluster detection methods for chronic respiratory diseases in Thailand. F1000Research, 1819, 1-16.*
9. *Ratnadewi, Y. (2015, Agustus 2015). Waspadai ISPA, MeningkatKasusnya di Kota Bandung. Retrieved from Pikiran Rakyat: http://www.pikiran-rakyat.com/bandung-raya/2015/08/04/337026/waspadai-ispa-meningkat-kasusnya-di-kota-bandung*
10. *Sari, W. C. (2018, April 26). Selain DBD, WaspadaDiaredan ISPA di MusimIni. Retrieved from Pikiran Rakyat: http://www.pikiran-rakyat.com/bandung-raya/2018/04/26/selain-dbd-waspada-diare-dan-ispa-di-musim-ini-423428*
11. *Tango, T. (2010). Statistica Method for Disease Clustering. London: Springer.*
12. *Zhang, C., Luo, L., Xu, W., &Ledwith, V. (2008). Use of local Moran's I and GIS to identify pollution hotspots of Pb in urban soils of Galway, Ireland. Science of The Total Environment, 398, 212-221.*