A Spatial Scan Statistic Scanning Only the Regions with Elevated Risk

Toshiro Tango

Department of Technology Assessment and Biostatistics
National Institute of Public Health, Japan

OBJECTIVE

To propose a new spatial scan statistic that has higher ability of pinpointing the true cluster.

BACKGROUND

The spatial scan statistic proposed by Kulldorff[1] has been applied to a wide variety of epidemiological studies and also to disease surveillance for the detection of disease clusters along with SaTScan software. However, it does not seem to be well recognized that maximizing the likelihood ratio tends to detect the most likely cluster much larger than the true cluster by swallowing neighbouring regions with non-elevated risk, Tango and Takahashi[2] have shown such typcial examples which casted a doubt on the validity of the model selection based on maximizing the likelihood ratio.

METHODS

To avoid such undesirable phenomena, we propose the following modified likelihood ratio test statistic by taking individual region's risk into account:

$$\lambda_T(\mathbf{Z}) = \sup_{\mathbf{Z}\in\mathcal{Z}} \left(\frac{n(\mathbf{Z})}{\xi(\mathbf{Z})}\right)^{n(\mathbf{Z})} \left(\frac{n-n(\mathbf{Z})}{n-\xi(\mathbf{Z})}\right)^{n-n(\mathbf{Z})} \cdot I\left(\frac{n(\mathbf{Z})}{\xi(\mathbf{Z})} > \frac{n-n(\mathbf{Z})}{n-\xi(\mathbf{Z})}\right) \prod_{i\in\mathbf{Z}} I\left(p_i < \alpha_1\right)$$

where where $n()$ and $\xi()$ denote the random number of cases and the null expected number of cases within the specified window $\mathbf{Z}$, respectively, p_i is the p-value for the i-th region's risk and α_1 is the pre-specified individual region's significance level. Individual region's significance level α_1 is set equal to the significance level α_0 for detecting MLC.

RESULTS

The proposed circular spatial scan statistic was, via Monte Carlo simulations, shown to have quite high ability of pinpointing the true circular cluster assumed in the simulation (ex., see Table 1).

CONCLUSIONS

Our result suggests that, if other spatial scan statistics for detecting arbitrary shaped clusters adopt the proposed likelihood ratio test statistic, their performance is expected to be improved.

REFERENCES

[1] Kulldorff M. A spatial scan statistic. *Communications in Statistics: Theory and Methods* 1997; **26**, 1481–1496.

[2] Tango T, Takahashi T, A flexibly shaped spatial scan statistic for detecting clusters. *International Journal of Health Geographics* 2005; **4**:11.

Table 1. Simulated bivariate power distributions $P(l, s) \times 1000$ of Kulldorff's circular spatial scan statistic and the proposed circular spatial scan statistic for the hot-spot circular cluster $\mathbf{A} = \{14, 15, 20\}$ with length $l = 3$ and relative risk $\theta = 3.0$ that was assumed in the areas of Tokyo Metropolis and Kanagawa prefecture in Japan (113 regions). Total number of cases was set as 235. Two kinds of α-levels were set as $\alpha_0 = \alpha_1 = 0.05$ and 1000 trials were carried out. The power of accurate detection ($l = s = 3$) was indicated by "$*$".

Kulldorff's circular scan

Length l	Include s hot-spot regions: 0	1	2	3
1	1	0		
2	0	0	0	
3	0	0	0	∗672
4	0	0	0	147
5	0	0	0	38
6	0	0	0	26
7-10	0	0	0	15
11-20	0	0	0	50
21-47	0	0	0	14
Total	1	0	0	962

usual power=0.963

Proposed circular scan

Length l	Include s hot-spot regions: 0	1	2	3
1	1	0		
2	1	0	0	
3	0	0	0	∗845
4	0	0	0	38
Total	2	0	0	883

usual power=0.885