2.2 A New Space–Time Scan Statistic

2.2.1 *Temporal overdispersion.* We will assume that, under the null hypothesis of no outbreaks, the number of cases N_{it} in region i $(i = 1, \ldots, m)$ at some surveillance time t follows an independent negative binomial distribution $NB(\mu_{it}, \phi_{it})$ by taking the possibility of nonnegligible time-to-time variation of Poisson mean or *temporal overdispersion*, into account

$$H_0 : N_{it} \sim NB(\mu_{it}, \phi_{it}), \tag{9}$$

where the negative binomial distribution used here is given by

$$\Pr\{N_{it} = n_{it} \mid \mu_{it}, \phi_{it}\} = \frac{\Gamma(\phi_{it} + n_{it})}{\Gamma(\phi_{it})\, n_{it}!} \left(\frac{\phi_{it}}{\phi_{it} + \mu_{it}}\right)^{\phi_{it}} \left(\frac{\mu_{it}}{\phi_{it} + \mu_{it}}\right)^{n_{it}} \tag{10}$$

and

$$E(N_{it}) = \mu_{it}, \quad \mathrm{Var}(N_{it}) = \mu_{it} + \mu_{it}^2/\phi_{it} = \mu_{it} w_{it}. \tag{11}$$

The temporal overdispersion is given by

$$w_{it} = 1 + \mu_{it}/\phi_{it} \tag{12}$$

and ϕ_{it} denote the parameter regulating overdispersion. Needless to say, if we do not observe any overdispersion in region i, then we have only to set $\phi_{it} = \infty$ or $w_{it} = 1$.

2.2.2 *How to estimate* (μ_{it}, ϕ_{it}). The expected number of cases μ_{it} and the parameter ϕ_{it} should be estimated using data from a predefined *baseline period.* If we can use the surveillance data reported for several years in the past, then we can apply an appropriate *negative binomial regression model* within the framework of generalized linear mixed models (for example, see Hardin and Hilbe, 2007; Hilbe, 2007)

$$\log E(Y_{it}) = \sum_{j=1} x_{itj}\beta_j + b_i, \quad Y_{it} \sim NB\left(\mu_{it}^{(Y)}, \phi_i^{(Y)}\right), \tag{13}$$

where Y_{it} denotes a random variable for the observed counts data y_{it} in the predefined baseline period, x_{itj} denotes the value of covariate j such as a month, a day-of-week, and so on; b_i denotes a random regional effects independently normally distributed with mean 0; and ϕ_{it} is usually set constant over time, i.e., $\phi_{it} = \phi_i$. This model is an extension of generalized linear mixed models used by Kleinman et al. (2004) and Kleinman (2005). The independence assumption among the b_i is made for ease of application, the same reason used by Kleinman et al. (2004). Using this regression model, we can set $\mu_{it} = \hat{\mu}_{it}^{(Y)}$ and $\phi_{it} = \hat{\phi}_i^{(Y)}$.

However, if the surveillance system starts in the recent past as in the case of our example illustrated later, we cannot adopt the above-stated regression approach. In this case, we can use a so-called *moving average* method. This says that the null expected number of cases μ_{it} is assumed to be constant during the *baseline period* defined such as $\{t-1, t-2, \ldots t-B\}$ with baseline length B. Then these two parameters can be simply estimated using baseline mean $\bar{y}_{i\cdot}$ and baseline variance s_i^2 by the moment method, i.e.,

$$\mu_{it} = \hat{\mu}_{it}^{(Y)} = \bar{y}_{i\cdot} \tag{14}$$

$$\phi_{it} = \hat{\phi}_i^{(Y)} = \begin{cases} \dfrac{\bar{y}_{i\cdot}^2}{s_i^2 - \bar{y}_{i\cdot}}, & \text{if } s_i^2 > \bar{y}_i \\ \infty, & \text{otherwise.} \end{cases} \tag{15}$$

2.2.3 *An outbreak model.* Under the above-stated situation, the alternative hypothesis is given by

$$H_1 : N_{it} \sim NB(\theta_{it}\mu_{it}, \phi_{it}), \tag{16}$$

where (μ_{it}, ϕ_{it}) are known and θ_{it} denote the unknown relative risk in region i at surveillance time t. We propose the following *outbreak model* to capture localized emerging disease outbreaks:

$$\theta_{it} = \begin{cases} h(\tau + \beta_W (t - t_p + u)), & \text{if } (i,t) \in W = Z \times I_u \\ 1, & \text{otherwise,} \end{cases} \tag{17}$$

where $h(\tau)$ denote the relative risk at $t = t_p - u$ and the *initial slope* of emerging disease outbreak which starts just after the time point $t_p - u$ within the domain W is

$$\left[\frac{\partial \theta_{it}}{\partial t}\right]_{t=t_p-u} = \beta_W h'(\tau) \tag{18}$$

and $h(.)$ denotes any monotonically increasing function with $h(\tau) = 1$ and the first and second differentials $h'(.)$ and $h''(.)$ are assumed to be finite. Then, the above-stated hypothesis testing is simply reduced to the following hypothesis testing over all possible sets of domains $W = Z \times I_u$:

$$H_0 : \beta_W = 0, \quad H_1 : \beta_W > 0. \tag{19}$$

Under the outbreak model (17), the likelihood is

$$L(\beta \mid \mu_{it}, \phi_{it}, t_p, u, t) = \prod_{i \in Z} \prod_{t \in I_u} \frac{\Gamma(\phi_{it} + n_{it})}{\Gamma(\phi_{it})\, n_{it}!} \left(\frac{\phi_{it}}{\phi_{it} + \theta_{it}\mu_{it}}\right)^{\phi_{it}} \left(\frac{\theta_{it}\mu_{it}}{\phi_{it} + \theta_{it}\mu_{it}}\right)^{n_{it}} \times \prod_{(i,t) \in W^c} \frac{\Gamma(\phi_{it} + n_{it})}{\Gamma(\phi_{it})\, n_{it}!} \left(\frac{\phi_{it}}{\phi_{it} + \mu_{it}}\right)^{\phi_{it}} \left(\frac{\mu_{it}}{\phi_{it} + \mu_{it}}\right)^{n_{it}}. \tag{20}$$

Then, as with Kulldorff's space–time scan statistics, we can construct a likelihood ratio test for testing the null hypothesis $H_0 : \beta_W = 0$ against $H_1 : \beta_W > 0$. However, the likelihood ratio test statistic requires the functional form $h(.)$ and the maximum likelihood estimator for β_W. Therefore, we will derive here an efficient score test that does not depend on the functional form $h(.)$ and does not require the maximum likelihood estimator for β_W, which is asymptotically equivalent to the likelihood ratio test. The efficient score test statistic for the null hypothesis $H_0 : \beta_W = 0$ is

$$S = \sup_{Z \in Z,\ 1 \le u \le T} \frac{\sum_{i \in Z} \sum_{t \in I_u} (n_{it} - \mu_{it})(t - t_p + u)/w_{it}}{\sqrt{\sum_{i \in Z} \sum_{t \in I_u} \mu_{it}(t - t_p + u)^2/w_{it}}} \sim N(0,1). \tag{21}$$

The derivation of the efficient score test is given in the Appendix. Needless to say, if the Poisson distribution can be

used instead of the negative binomial distribution for all regions, then we have only to set $w_{it} = 1$ for all regions. The domain $W^* = Z^* \times I_{u^*}$ for which the efficient score is maximized identifies the *Most Likely Outbreak* (MLO). In a similar manner to Kulldorff's scan statistic, secondary outbreaks and Monte Carlo simulated p-value can be defined.

If we assume a *hotspot cluster model* instead of the proposed *outbreak model*

$$\theta_{it} = \begin{cases} \tau_W\,(>1), & \text{if } (i,t) \in W = Z \times I_u \\ 1, & \text{otherwise} \end{cases} \tag{22}$$

then, the efficient score test statistic for the null hypothesis $H_0 : \tau_W = 1$ is

$$S_2 = \sup_{Z \in \mathcal{Z},\ 1 \le u \le T} \frac{\sum_{i \in Z} \sum_{t \in I_u} (n_{it} - \mu_{it})/w_{it}}{\sqrt{\sum_{i \in Z} \sum_{t \in I_u} \mu_{it}/w_{it}}} \sim N(0,1). \tag{23}$$

Furthermore, if the Poisson distribution can be used, then the efficient score statistic is reduced to

$$S_3 = \sup_{W \in \mathcal{W}} \frac{n(W) - \mu(W)}{\sqrt{\mu(W)}} \sim N(0,1) \tag{24}$$

which is asymptotically equivalent to the *unconditional* likelihood ratio test in the sense that the expected number of cases $\mu(W)$ in the domain W are calculated unconditionally from the baseline data.

As it is well known that a cylindrical scan statistic performs better than a prismatic scan statistic in detecting a circular area and the latter performs better than the former in detecting a noncircular area. Therefore, in what follows, we will consider two cylindrical space time scan statistics for comparison: Kulldorff's cylindrical scan statistic and the proposed cylindrical scan statistic. We will sometimes refer to these statistics as KC and PC, respectively.

3. Application

We will illustrate here these two space time scan statistics with data from weekly surveillance of the absentees in 131 primary school districts in Kitakyushu-shi (city), Japan, during April 12 (1st week) to December 20 (30th week) in 2006, where the total number of school children was 52,177. Figure 1 shows the weekly number of infectious enteritis patients in Kitakyushu-shi from April 2006 to March 2007, taken from Infectious Disease Weekly Report in this city. This clearly shows the occurrence of an overall outbreak during November 15 to December 20. Needless to say, there are several other reasons for students' absence from school. However, we assumed here that the proportion of absentees unrelated to infectious enteritis would be approximately constant for each primary school since the detailed data on the reason for absence was not available.

3.1 *Design*

In this application, we used an 8-week moving average method (14) and (15) to estimate the baseline information (μ_{it}, ϕ_{it}) for the analysis week t. Figure 2 shows the weekly number of absentees in four selected school districts, 8-week moving average line (from the 10th week, June 21st) and two kinds of 95% UCL (upper control limit) lines based on a Poisson

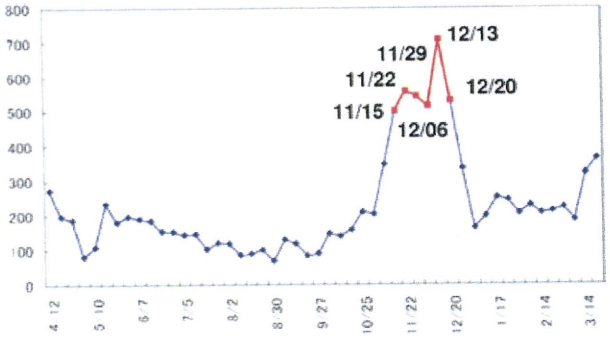

Figure 1. Weekly number of infectious enteritis patients in Kitakyushu-shi, Japan, from April 2006 to March 2007 (from Infectious Disease Weekly Report). This figure appears in color in the electronic version of this article.

distribution (dotted line) and the negative binomial distribution (dash dotted line). We can observe that three school districts No. 69, No. 70, No. 76, and No. 77 have nonnegligible temporal overdispersion but school No. 70 does not have any overdispersion. We started the weekly analysis from October 11 (20th week) and continued the analysis until December 20 (30th week), wherein the maximum spatial length was set as $K = 15$, the maximum temporal length was set as $T = 2$, the significance level was set as $\alpha = 0.02$ corresponding to one expected false alarm every 50 weeks or approximately 1 year (the recurrence interval, RI) and the p-value was based on the Monte Carlo hypothesis testing with $M = 999$ of replicates. As *detected areas*, we considered all significant clusters (outbreaks) including significant secondary clusters (outbreaks) that did not overlap with MLC (MLO).

3.2 *Results*

Table 1 summarizes the results of two cylindrical scan statistics in which we show the baseline expected number of absentees $\mu_{+t} = \sum_{i \in Z^*} \mu_{it}$ and the observed number of absentees $n_{+,t} = \sum_{i \in Z^*} n_{it}$ within the detected areas Z^* at week t. Using these statistics, we will indicate the overall temporal trend in the number of absentees of detected areas such as

$$(\mu_{+,t_p} : n_{+,t_p-1} \to n_{+,t_p}).$$

Figure 3 highlighted the detected areas by analysis week. For each of the analyses during October 11 to November 8, two cylindrical scan statistics did not detect any areas. On November 15, when the infectious enteritis outbreak seems to start (Figure 1), PC detected a significant outbreak O_1 that covered nine schools and had a temporal trend $(50 : 61 \to 87)$ with $p = 0.013$ and temporal length $u^* = 2$ (from November 8 to 15). But KC did not detect any clusters.

On November 22, both cylindrical scan statistics detected the same significant area $C_1 = O_1$ with 13 schools. On November 29, when we observed a decreasing trend in the number of infectious enteritis patients, PC did not detect any outbreaks. On the other hand, KC detected a significant cluster C_1 ($p = 0.001$) with a decreasing temporal trend $(97 : 168 \to 116)$. On December 6, both PC and KC detected the same significant two areas $C_1 = O_1$ and $C_2 = O_2$, respectively.

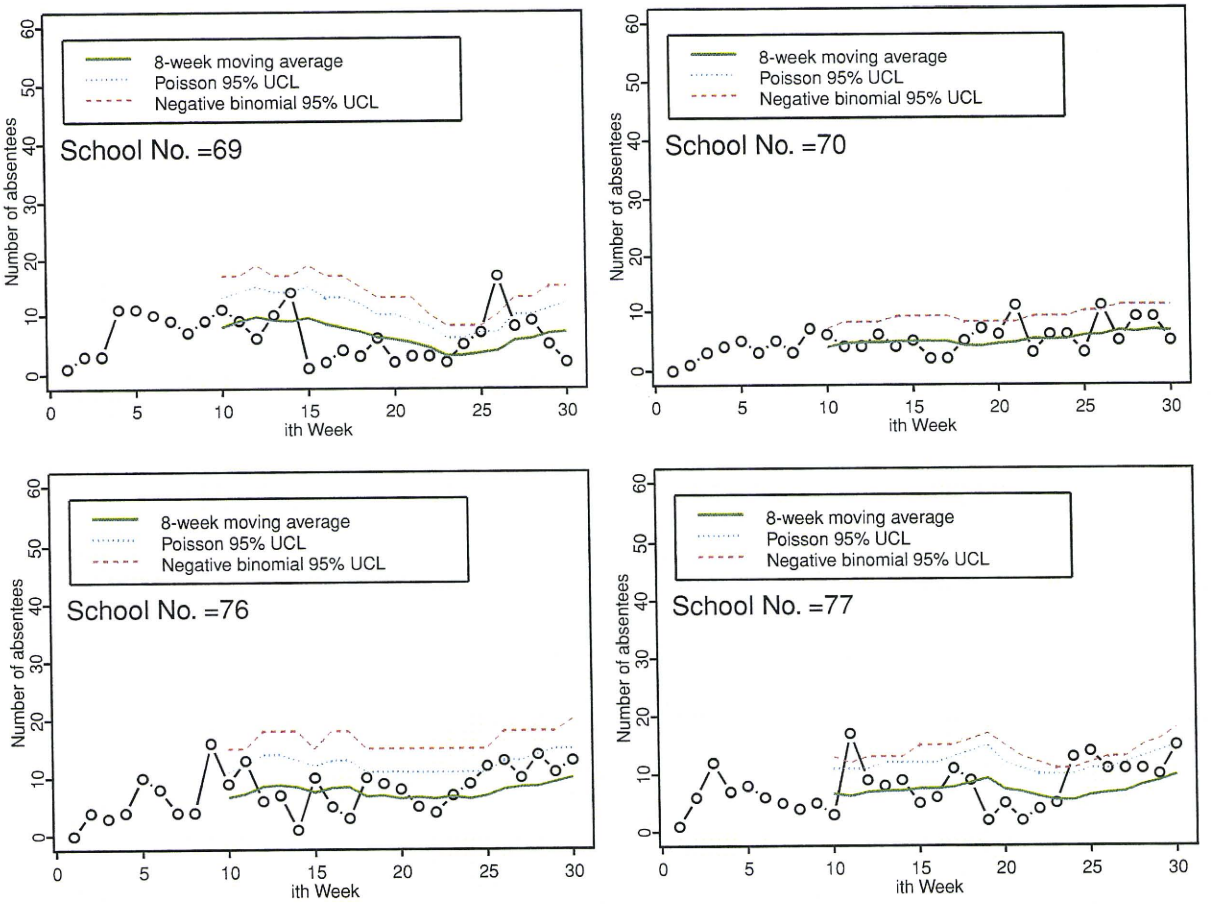

Figure 2. Weekly number of absentees in four selected school districts with 8-week moving average (line) and two kinds of 95% UCL (upper control limit) based on Poisson distribution (dotted line) and the negative binomial distribution (dash dotted line). Calculation of 8-week moving average was started from the 10th week (June 21st). This figure appears in color in the electronic version of this article.

On December 13, the difference in performance of these two scan statistics was also conspicuous. This week was the peak in the number of infectious enteritis patients. While KC detected one significant cluster C_1 ($p = 0.001$) with 10 schools with a trend ($69 : 130 \rightarrow 141$), PC detected five significant outbreaks in which the area O_1 includes the area C_1. All of those detected areas had an increasing trend in the number of absentees. On December 20, KC detected only one small area C_1 ($p = 0.001$) with two schools, which is a part of the area C_1 detected on the previous week while PC detected two new areas that were not detected before, indicating a possibility of spreading to this area.

4. Simulation Study

To compare the performance of two space time scan statistics, an extensive Monte Carlo simulation study was conducted. We report here the results for the situation where localized outbreaks occurred in the circular area.

4.1 *Data Simulated*

To mimic the situation of weekly surveillance of the absentees in 131 primary school districts in Kitakyushu-shi, we assumed that the true school-district specific (μ_i, ϕ_i) are constant during the baseline period and were estimated from data during 9 weeks from September 13($t = -8$) to November 8($t = 0$) in 2006. Then, for the ith school district, a set of baseline weekly number of absentees, $\{y_{it} : t = -8, -7, \ldots, 0\}$, are generated independently according to $NB(\mu_i, \phi_i)$. And then we calculated (μ_{i0}, ϕ_{i0}) using the 8-week moving average method (equations (14) and (15)) as the parameter values under null hypothesis of no outbreaks. Variability of calculated expected number of cases μ_{i0} and temporal overdispersion $w_{i0} = 1 + \mu_{i0}/\phi_{i0}$ were as follows:

$$\mu_{i0} : \min = 0.28,\ 25\% = 4.69,\ 50\% = 6.61,$$
$$75\% = 9.63,\ \max = 19.8$$
$$w_{i0} : \min = 1.0,\ 25\% = 1.27,\ 50\% = 1.63,$$
$$75\% = 1.94,\ \max = 11.88.$$

Then, to get the critical values of the test statistics, 9999 independent random data setss were generated using the negative binomial distribution $NB(\mu_{i0}, \phi_{i0})$ for PC and using the Poisson distribution Poisson(μ_{i0}) for KC. Throughout the

Table 1

Significant clusters detected by Kuldorff's cylindrical *scan statistic and significant outbreaks detected by the proposed* cylindrical *scan statistic applied to data from weekly surveillance of the absentees in* 131 *primary school districts in Kitakyushu-shi, Japan, during October* 11 *to December* 20, 2006, *where the maximum temporal length is set as* $T = 2$. *Bold number indicates the temporal period detected.*

	Kulldorff's *cylindrical* scan					Proposed *cylindrical* scan				
Current week t_p	Detected areas (p-value)	No. schools Z^*	Baseline μ^a_{+,t_p}	n^a_{+,t_p-1}	n_{+,t_p}	Detected areas (p-value)	No. schools Z^*	Baseline μ_{+,t_p}	n_{+,t_p-1}	n_{+,t_p}
10/11										
10/18										
10/25										
11/01										
11/08										
11/15						O_1 (0.013)	9	50	**61**	**87**
11/22	C_1 (0.001)	11	64	**101**	**130**	O_1 (0.003)	11	64	**101**	**130**
11/29	C_1 (0.001)	15	97	**168**	**116**					
12/06	C_1 (0.001)	5	32	42	**74**	O_1 (0.002)	5	32	42	**74**
	C_2 (0.004)	4	19	**30**	**44**	O_2 (0.008)	4	19	**30**	**44**
12/13	C_1 (0.001)	10	69	**130**	**141**	O_1 (0.002)	14	100	**171**	**191**
						O_2 (0.003)	14	99	**131**	**163**
						O_3 (0.008)	14	127	**159**	**200**
						O_4 (0.009)	9	75	**87**	**131**
						O_5 (0.013)	6	38	**56**	**75**
12/20	C_1 (0.001)	2	18	**44**	**36**	O_1 (0.004)	15	183	**243**	**261**
						O_2 (0.013)	14	108	**191**	**135**

[a] $\mu_{+,t} = \sum_{i \in Z^*} \mu_{it}$ and $n_{+,t} = \sum_{i \in Z^*} n_{it}$.

simulation, we used a maximum length of the geographic window $K = 15$, a maximum temporal length $T = 2$ weeks, and significance level $\alpha = 0.02$ corresponding to one expected false alarm every 50 weeks. Then, to estimate the size or the probability of false alarm for KC in our overdispersed condition, we applied KC to 1000 random data sets generated from $NB(\mu_{i0}, \phi_{i0})$. The resultant size of KC was inflated to 0.198 or one expected false alarm every 5 weeks, a result that does not take temporal overdispersion present into account. On the other hand, the estimated size of PC using these 1000 random data sets was 0.020, which is close to the nominal significance level as expected.

The primary purpose of our proposed space time scan statistic is an early detection of localized emerging disease outbreaks with geographical accuracy, not a detection of disease spreading over space and time. Thus, a total of eight alternative outbreak scenarios $(S_1, \ldots, S_8)$ were evaluated, based on three different outbreak areas shown in Figure 4:

1. A circular area **A**: consisting of five connected regions with large μ_{i0} $(= 11.1 \sim 19.6)$.
2. A circular area **B**: consisting of five connected regions with small μ_{i0} $(= 2.4 \sim 6.8)$.
3. Two Areas **A** and **B**.

Eight scenarios considered are as follows:

S1: the first outbreak at week $t = 1$: $\theta_{i1} = 1.5$
S2: the first outbreak at week $t = 1$: $\theta_{i1} = 2.0$
S3: hot-spot outbreak model (analysis week $t = 2$): $\theta_{i1} = 1.5, \theta_{i2} = 1.5$
S4: hot-spot outbreak model (analysis week $t = 2$): $\theta_{i1} = 2.0, \theta_{i2} = 2.0$
S5: increasing outbreak model (analysis week $t = 2$): $\theta_{i1} = 1.5, \theta_{i2} = 2.0$
S6: decreasing outbreak model (analysis week $t = 2$): $\theta_{i1} = 1.5, \theta_{i2} = 1.2$
S7: monotonically increasing outbreak model (analysis week $t = 3$): $\theta_{i1} = 1.5$, $\theta_{i2} = 2.0$, $\theta_{i3} = 2.5$
S8: Nonmonotonic outbreak model (analysis week $t = 3$): $\theta_{i1} = 1.5$, $\theta_{i2} = 1.2$, $\theta_{i3} = 2.0$.

Values of relative risk (1.2, 1.5, 2.0, 2.5) were chosen by considering the values of $n_{+,t}/\mu_{+,t}$ for the detected areas shown in Table 1. For each of eight scenarios, a set of weekly number of absentees, $\{n_{is} : s = 1, \ldots, t\}$, were generated independently according to $NB(\theta_{it}\mu_i, \phi_i)$. And then we calculated (μ_{is}, ϕ_{is}), $s = 1, \ldots, t$ by using the 8-week moving average method from data $\{y_{is}, s = -8, \ldots, 0, n_{is}, s = 1, \ldots, t\}$. At analysis week t, 1,000 independent random data sets were generated from $NB(\theta_{it}\mu_{it}, \phi_{it})$. In this simulation study, the *detected cluster or outbreak* is defined as the domain W whose corresponding test statistic is larger than the critical value, which includes significant secondary clusters (outbreaks).

4.2 *Bivariate Power Distribution*

In order to compare the performance of the cluster detection tests, the usual power has been treated in the same manner as in the usual hypothesis tests. However, it should be noted that the usual power estimates reflected the "power to reject the null hypothesis for whatever reasons," while the

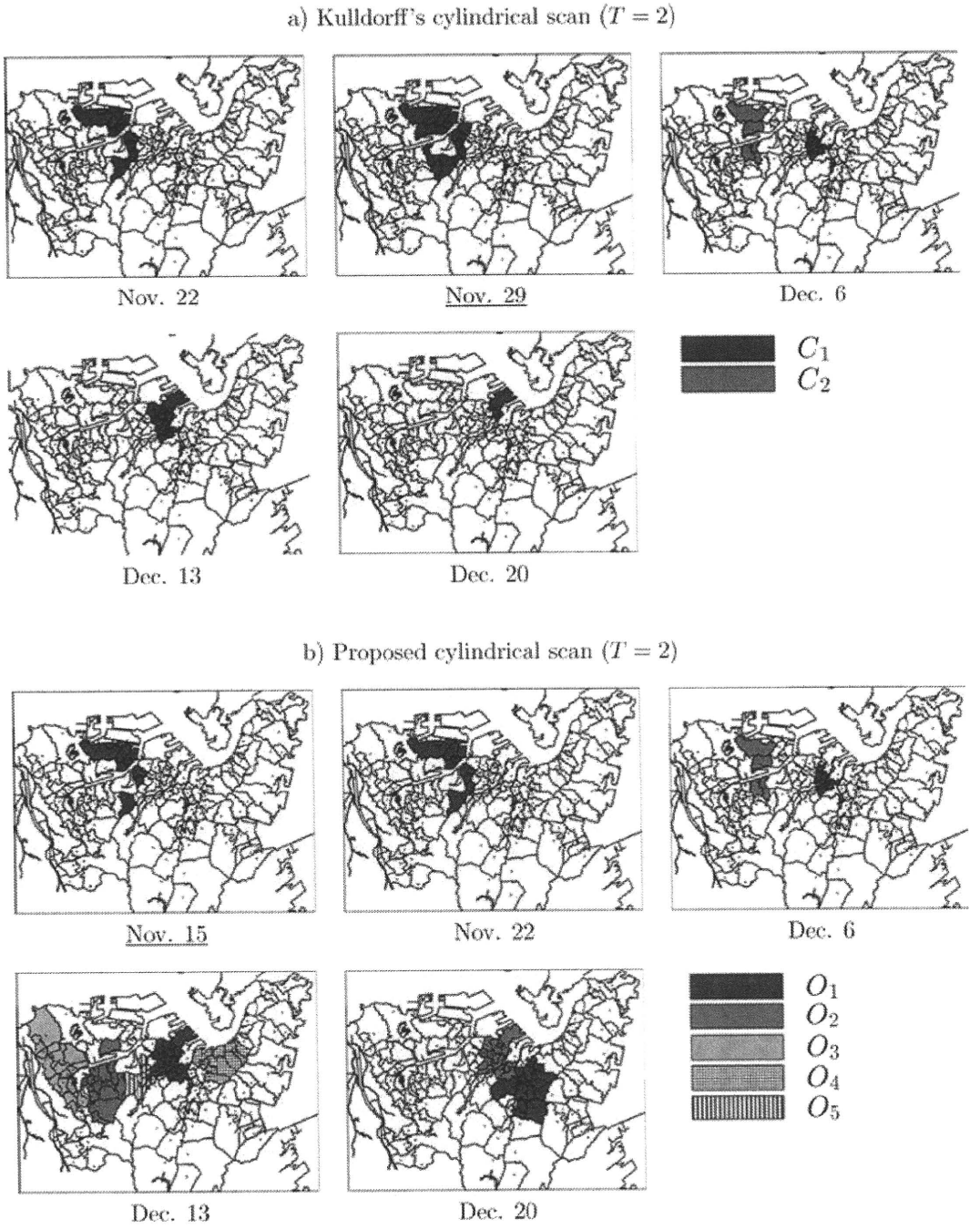

Figure 3. (a) Significant clusters detected by the Kulldorff's *cylindrical* scan statistic and (b) significant outbreaks detected by the proposed *cylindrical* scan statistic, applied to data from weekly surveillance of the absentees in 131 primary schools in Kitakyushu-shi, Japan, during October 2006 to December 2006, where the maximum temporal length is set as $T = 2$.

probability of both rejecting the null hypothesis and accurately identifying the true cluster is a different matter altogether. To compare the performance of the purely spatial cluster detection tests, Tango and Takahashi (2005) proposed a spatial bivariate power distribution $P(l, s \mid s^*)$ based on Monte Carlo simulation where l is the length of significant MLC, s is the number of regions identified out of the assumed true cluster with s^* regions.

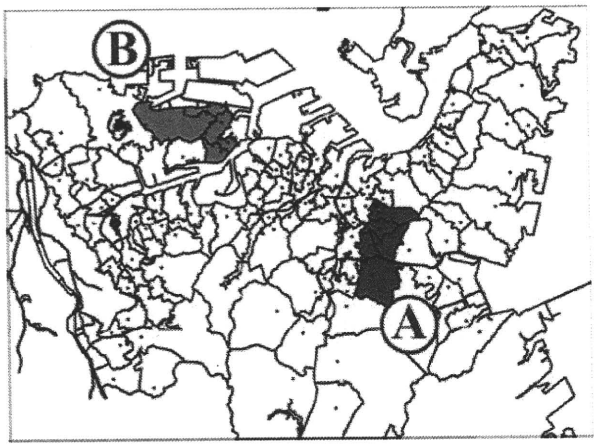

Figure 4. Outbreak areas $\boldsymbol{A}$ and $\boldsymbol{B}$ assumed in the simulation. Both areas are circular.

$$P(l, s \mid s^*) = \Pr\{L = l, S = s \mid s^*\} = \frac{\#\{\text{significant MLC has length } l \text{ and includes } s \text{ true regions}\}}{\#\{\text{trials for each simulation}\}}, \tag{25}$$

where L and S denote the random variable of l and s under the specified model, respectively, and $l \geq 1$ and $0 \leq s \leq s^*$. To evaluate and compare the statistical method to detect changes or aberrations in public health surveillance time-series data, the probability of false alarm or equivalently the average run length under the null hypothesis of no outbreaks (ARL^0) and the average run length under the alternative hypothesis (ARL^1) are often used (Sonesson and Bock, 2003). However, in surveillance to detect localized changes in the spatial pattern of a disease, we have to evaluate another type of false alarm under the alternative hypothesis such as detection of nonoutbreak areas. Therefore, based on $P(l, s \mid s^*)$, we considered the following four powers for each analysis week:

Q_1: the probability of exact detection, $P(s^*, s^* \mid s^*)$.

Q_2: the probability that regions detected are all outbreak regions, $\sum_{s=1}^{s^*} P(s, s \mid s^*)$.

Q_3: the probability that regions detected include at most two nonoutbreak regions,

$$\sum_{s=1}^{s^*} \sum_{j=0}^{2} P(s + j, s \mid s^*).$$

Q_4: the usual power, $\sum_{l \geq 1} \sum_{s=0}^{s^*} P(l, s \mid s^*)$.

Each of the three powers Q_1, Q_2, and Q_3 can be considered as a measure of the geographical accuracy of detection. It should be noted that the usual power of KC is expected to be larger than PC due to much inflated actual size (KC = 0.198, PC = 0.020).

4.3 *Results*

The results are shown in Table 2. For most scenarios applied to outbreak areas $\boldsymbol{A}$ and $\boldsymbol{A} + \boldsymbol{B}$, we can observe that (i) the usual powers Q_4 of KC were sometimes surprisingly much higher than those of PC and well beyond our expectation and (ii) the other powers Q_1, Q_2, and Q_3 of KC, on the other hand, were often much lower than those of PC and the usual powers of KC. Furthermore, when we observe the change from the usual powers of scenario S_6 (a decreasing outbreak model) at week $t = 1$, i.e., scenario S_1, to those of scenario S_6 at week $t = 2$, the usual powers of KC are strangely increasing while those of PC are decreasing irrespective of outbreak areas. These results clearly indicate that KC has quite low power to detect emerging disease outbreaks correctly and tends to detect many nonoutbreak regions, which makes KC's apparent usual power surprisingly high depending on the situation. When applied to the outbreak area $\boldsymbol{B}$ that is a set of regions with small expected number of cases, the usual powers of KC often falls lower than (a) its actual size (indicated by a in Table 2) and (b) those of PC contrary to our expectation, suggesting that KC does not satisfy the requirement for *sound* statistical tests. The conclusion from these results did not change even when we took the 95% confidence interval of the estimated powers into account.

5. Discussion

In this article, we have proposed a new space time scan statistic to cope with the following three questions associated with both Kulldorff's and Takahashi et al.'s space–time scan statistic where (i) why is the conditional expected number of cases e_{it} defined in equation (3) or (4) used as the basis for detecting emerging disease outbreaks?, (ii) a time-to-time variation of Poisson mean could not be taken into account, and iii) the alternative *hotspot* hypothesis was assumed for emerging disease outbreaks. The proposed space time scan statistic is not a usual likelihood ratio statistic but an efficient score test statistic based on the *initial slope of localized emerging disease outbreaks*, which has the important property that it does not depend on the functional form, or the pattern of emerging disease outbreaks.

The proposed space time scan statistic was then illustrated with data from weekly surveillance of the absentees in 131 primary schools in Kitakyushu-shi. Two cylindrical scan statistics were illustrated and compared in this article. We also applied two prismatic scan statistics, Takahashi et al. (2008)'s space time scan statistic and our proposed space time scan statistic with prismatic domain, and the results are shown in Web Table 1. The difference in performance of the two prismatic scan statistics was more or less similar to that of the two cylindrical scan statistics. For example, on November 15, when the infectious enteritis outbreak seems to start, neither Kulldorff's cylindrical scan statistic nor Takahashi et al.'s prismatic scan statistic detected any clusters but our proposed cylindrical and prismatic scan statistic detected a significant outbreak and three outbreaks, respectively. These results might give us an example of timely detection of emerging outbreaks for the proposed scan statistic. On November 29, Kulldorff's cylindrical scan statistic and Takahashi et al.'s prismatic scan statistic detected a significant but slightly different cluster, respectively. Our proposed two scan statistics,

Table 2

Four kinds of estimated powers of Kulldorff's space–time scan and the proposed cylindrical space–time scan by simulation scenario and outbreak area, where the significance level $\alpha = 0.02$ *and* 1000 *trials for each scenario were carried out. Estimated size or the probability of false alarm:* $Kulldorff = 0.198(KC)$ *and Proposed with cylindrical scan* $= 0.020(PC)$.

		Circular area $\boldsymbol{A}$ with large μ_{it}				Circular area $\boldsymbol{B}$ with small μ_{it}				Area $\boldsymbol{A}+\boldsymbol{B}$			
Scenario	Method	Q_1	Q_2	Q_3	Q_4	Q_1	Q_2	Q_3	Q_4	Q_1	Q_2	Q_3	Q_4
S_1: 1.5	PC	0.7	3.0	3.3	5.6	0.4	1.7	2.0	4.3[a]	0.0	5.8	6.4	8.7
	KC	0.4	0.7	0.8	54.0	0.0	0.0	0.0	16.8[a]	0.0	0.2	0.6	48.6
S_2: 2.0	PC	28.5	52.4	59.4	62.5	6.1	13.4	17.1	18.8	2.5	57.0	65.5	72.0
	KC	7.8	9.0	11.2	88.2	0.2	0.4	0.4	16.7[a]	0.1	9.7	12.7	85.8
S_3: 1.5-1.5	PC	4.6	10.3	12.8	15.0	0.1	1.7	2.6	5.0[a]	0.0	12.6	16.2	19.1
	KC	1.5	1.6	2.6	87.5	0.0	0.1	0.1	18.9[a]	0.0	1.1	2.1	84.8
S_4: 2.0-2.0	PC	61.5	85.1	91.4	93.4	20.0	32.3	38.4	39.9	11.2	82.8	93.4	95.6
	KC	25.8	26.5	30.8	99.9	5.7	6.5	8.1	34.4	1.2	30.4	36.6	99.8
S_5: 1.5-2.0	PC	41.4	56.1	63.7	76.4	8.0	16.0	20.0	21.9	3.3	66.8	76.8	80.3
	KC	9.5	10.0	13.4	98.6	1.5	1.8	2.4	22.2	0.1	11.3	12.4	98.2
S_6: 1.5-1.2	PC	0.0	0.2	0.2	2.8	0.1	0.9	0.9	2.9[a]	0.0	0.8	1.3	3.6
	KC	0.1	0.2	0.6	66.5	0.0	0.1	0.1	18.0[a]	0.0	0.2	0.4	62.7
S_7: 1.5-2.0-2.5	PC	81.5	96.7	99.5	99.8	38.9	61.8	67.8	68.9	33.8	92.2	99.0	99.8
	KC	49.4	50.6	55.3	100	21.4	25.5	29.9	63.4	11.6	52.7	61.2	100
S_8: 1.5-1.2-2.0	PC	5.1	22.1	24.5	26.8	1.3	6.2	6.9	8.6	0.2	25.9	28.6	31.0
	KC	5.9	11.6	14.9	90.9	1.1	2.1	2.4	38.5	0.2	11.5	15.4	89.3

Q_1: The probability of exact detection $P(s^*, s^*)$.

Q_2: The probability that regions detected are all outbreak regions, $\sum_{s=1}^{s^*} P(s, s)$.

Q_3: The probability that regions detected include at most two nonoutbreak regions, $\sum_{s=1}^{s^*} \sum_{j=0}^{2} P(s+j, s)$.

Q_4: The usual power.

[a]The usual power is less than its size.

on the other hand, did not detect any outbreaks. Because the temporal trend in the number of absentees of clusters detected by two ordinary scan statistics was decreasing from the previous week, these two clusters will not be considered outbreaks. On December 13, when the difference in performance between two cylindrical scan statistics was conspicuous, the difference in performance between two prismatic scan statistics was also similarly noticeable. These results seemed to suggest that Kulldorff's and Takahashi et al.'s space–time scan statistics have lower power in detecting localized outbreaks correctly compared with our proposed space time scan statistics.

We have not included a comparison of our proposed scan test with Kleinman et al. (2004)'s model-based approach which computes the p-value of having the observed number of cases or more case for a given surveillance time in a given region. Though their approach performs well in some contexts, for the problem under study we do not consider it a reasonable option. Specifically, it faces a considerable problem of multiple testing due to many regions, it has difficulty in detecting elevated counts over several contiguous regions or localized clustered outbreaks, and our negative binomial regression model (13) is an extension of Kleinman et al.'s Poisson regression model taking temporal overdispersion into account.

Furthermore, we carried out an extensive Monte Carlo simulation study to compare the performance of two cylindrical and two prismatic space time scan statistics. In Section 4, we report only the results of two cylindrical scan statistics, but the results of the two prismatic scan statistics are shown in Web Table 2. The simulation study reveals a good performance of our proposed space time scan statistics and an undesirable performance of the ordinary space time scan statistics in the sense that (i) Kulldorff's and Takahashi et al.'s space time scan statistics tend to detect many nonoutbreak regions, which makes their apparent usual power surprisingly high, (ii) Kulldorff's and Takahashi et al.'s powers relating to the geographical accuracy are much lower than those of the proposed space time scan statistics, and (iii) the usual powers of Kulldorff's and Takahashi et al.'s often falls lower than even their actual size when applied to the situation where the null expected number of cases is small in the outbreak area.

However, our proposed space time scan statistic still adopts the same cylindrical and/or prasmatic domain W in the same manner as Kulldorff's and Takahashi et al.'s space–time scan statistics, i.e., the proposed domain cannot adjust itself as the localized disease outbreak grows or shrinks spatially over time. Implementing such flexible spatial change over time will be an important and challenging issue. While this manuscript has been submitted, Neill (2009) proposed an *expectation-based Poisson statistic*, which is a likelihood ratio statistic asymptotically equivalent to the efficient score statistics S_3 derived under the Poisson model.

6. Supplementary Materials

Web Tables referenced in Section 5 are available under the Paper Information link at the *Biometrics* website `http://www.biometrics.tibs.org`.

Acknowledgements

The authors thank the coeditor, the associate editor, and anonymous referees for constructive comments that substantially improved the presentation of the article. This research was supported in part by Grant-in-Aid for Scientific Research

(Grant No. 20300101) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Assunção, R., Costa, M., Tavares, A., and Ferreira, S. (2006). Fast detection of arbitrarily shaped disease clusters. *Statistics in Medicine* **25,** 723–742.

Duczmal, L. and Assunção, R. (2004). A simulated annealing strategy for the detection of arbitrarily shaped spatial clusters. *Computational Statistics and Data Analysis* **45,** 269–286.

Dwass, M. (1957). Modified randomization tests for nonparametric hypotheses. *Annals of Mathematical Statistics* **28,** 181–187.

Hardin, J. W. and Hilbe, J. M. (2007). *Generalized Linear Models and Extensions*, 2nd edition. College Station, Texas: A Stata Press Publication.

Heffernan, R., Mostashari, F., Das, D., Karpati, A., Kulldorff, M., and Weiss, D. (2004). Syndromic surveillance in public health practice, New York City. *Emerging Infectious Diseases* **10,** 858–864.

Hilbe, J. M. (2007). *Negative Binomial Regression*. Cambridge: Cambridge University Press.

Kleinman, K. (2005). Generalized linear models and generalized linear mixed models for small-area surveillance. In *Spatial & Syndromic Surveillance for Public Health*, A. Lawson and K. Kleinman (eds), 77–94. Chichester, U.K.: John Wiley.

Kleinman, K., Lazarus, R., and Platt, R. (2004). A generalized linear mixed models approach for detecting incident clusters of disease in small areas, with an application to biological terrorism. *American Journal of Epidemiology* **159,** 217–224.

Kleinman, K., Abrams, A., Kulldorff, M., and Platt, R. (2005). A model-based space-time scan statistic with an application to syndromic surveillance. *Epidemiology and Infection* **133,** 409–419.

Kulldorff, M. (2001). Prospective time periodic geographical disease surveillance using a scan statistic. *Journal of the Royal Statistical Society, Series A* **164,** 61–72.

Kulldorff, M. and Information Management Services, Inc. (2009). *SaTScan version 7.0: software for the spatial and space-time scan statistics*. `http://www.satscan.org`, accessed January 4, 2010.

Kulldorff, M., Heffernan, R., Hartman, J., Assunção, R., and Mostashari, F. (2005). A space-time permutation scan statistic for disease outbreak detection. *PLoS Medicine* **2,** e59.

Kulldorff, M., Huang, L., Pickle, L., and Duczmal, L. (2006). An elliptic spatial scan statistic. *Statistics in Medicine* **25,** 3929–3943.

Lawson, A. B. and Kleinman, K. eds. (2005). *Spatial & Syndromic Surveillance for Public Health*. Chichester, U.K.: Wiley.

Lazarus, R., Kleinman, K., Dashevsky, I., Adams, C., Kludt, P., DeMaria, A., and Platt, R. (2002). Use of automated ambulatory-care encounter records for detection of acute illness clusters, including potential bioterrorism events. *Emerging Infectious Diseases* **8,** 753–760.

Lombardo, J., Burkom, H., Elbert, E., Magruder, S., Lewis, S. H., Loschen, W., Sari, J., Sniegoski, C., Wojcik, R., and Pavlin, J. (2003). A systems overview of the electronic surveillance system for the early notification of community-based epidemics (ESSENCE II). *Journal of Urban Health* **80** (Suppl. 1), i32–i42.

Montgomery, D. C. (1991). *Introduction to Statistical Quality Control*. New York: John Wiley & Sons.

Mostashari, F., Fine, A., Das, D., Adams, J., and Lyton, M. (2003). Use of ambulance dispatch data as an early warning illness, New York City. *Journal of Urban Health* **80** (Suppl. 1), i43–i49.

Neill, D. B. (2009). An empirical comparison of spatial scan statistics for outbreak detection. *International Journal of Health Geographics* **8,** 20.

Patil, G. P. and Taillie, C. (2004). Upper level set scan statistic for detecting arbitrarily shaped hotspots. *Environmental and Ecological Statistics* **11,** 183–197.

Platt, R., Bocchino, C., Caldwell, B., Harmon, R., Kleinman, K., Lazarus, R., Nelson, A. F., Nordin, J. D., and Ritzwoller, P. (2003). Syndromic surveillance using minimum transfer of identifiable data: The example of the National Bioterrorism Syndromic Surveilance Demonstration Program. *Journal of Urban Health* **80** (Suppl. 1), i25–i31.

Sonesson, C. and Bock, D. (2003). A review and discussion of prospective statistical surveillance in public health. *Journal of the Royal Statistical Society, Series A* **166,** 5–21.

Tango, T. and Takahashi, K. (2005). A flexibly shaped spatial scan statistic for detecting clusters. *International Journal of Health Geographics* **4,** 11.

Takahashi, K., Kulldorff, M., Tango, T., and Yih, K. (2008). A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring. *International Journal of Health Geographics* **7,** 14.

Takahashi, K., Yokoyama, T., and Tango, T. (2009). *FleXScan: Software for the Flexible Spatial Scan Statistic*. National Institute of Public Health, Japan. `http://www.niph.go.jp/soshiki/gijutsu/index_e.html`, accessed January 4, 2010.

Zhou, H. and Lawson, A. B. (2008). EWMA smoothing and Bayesian spatial modeling for health surveillance. *Statistics in Medicine* **27,** 5907–5923.

Received February 2009. *Revised January* 2010. *Accepted January* 2010.

Appendix

Derivation of Efficient Score Test

From the likelihood (20), we have the following partial derivatives:

$$\frac{\partial \log L(\beta_W)}{\partial \beta_W} = \sum_{i\in Z}\sum_{t\in I_u}\left(\frac{n_{it}}{\theta_{it}} - \frac{\mu_{it}(\phi_{it}+n_{it})}{\phi_{it}+\theta_{it}\mu_{it}}\right) \times (t-t_p+u)h'(\tau+\beta_W(t-t_p+u))$$

$$\frac{\partial^2 \log L(\beta_W)}{\partial \beta_W^2} = \sum_{i\in Z}\sum_{t\in I_u}\left(-\frac{n_{it}}{\theta_{it}^2} + \frac{\mu_{it}^2(\phi_{it}+n_{it})}{(\phi_{it}+\theta_{it}\mu_{it})^2}\right) \times \{(t-t_p+u)h'(\tau+\beta_W(t-t_p+u))\}^2 + \sum_{i\in Z}\sum_{t\in I_u}\left(-\frac{n_{it}}{\theta_{it}} + \frac{\mu_{it}(\phi_{it}+n_{it})}{\phi_{it}+\theta_{it}\mu_{it}}\right) \times (t-t_p+u)^2 h''(\tau+\beta_W(t-t_p+u)).$$

Then, the efficient score $U(0)$ and the Fisher information $J(0)$ evaluated under the null hypothesis $H_0 : \beta_W = 0$ are given by

$$U(0) = \left[\frac{\partial \log L(\beta_W)}{\partial \beta_W}\right]_{\beta_W=0} = \sum_{i\in Z}\sum_{t\in I_u}\left(\frac{\phi_{it}(n_{it}-\mu_{it})}{\phi_{it}+\mu_{it}}\right)(t-t_p+u)h'(\tau)$$

$$J(0) = -E\left[\frac{\partial^2 \log L(\beta_W)}{\partial \beta_W^2}\right]_{\beta_W=0} = \sum_{i\in Z}\sum_{t\in I_u}\left(\frac{\phi_{it}\mu_{it}}{\phi_{it}+\mu_{it}}\right)((t-t_p+u)h'(\tau))^2.$$

Therefore, the efficient score test statistic is given by $U(0)/\sqrt{J(0)}$ as shown in (21).

A Space-Time Scan Statistic for Detecting Emerging Outbreaks

Toshiro Tango, Kunihiko Takahashi, and Kazuaki Kohriyama

Web-based Supplementary Materials

Table 1 (*Web Table 1*) shows the results of two prismatic scan statistics, Takahashi *et al.* (2008)'s space-time scan statistic and our proposed space-time scan statistic with prismatic domain, applied to data from weekly surveillance of the absentees in 131 primary school districts in Kita kyushu shi, during October 11 to December 20, 2006, Japan.

Table 2 (*Web Table 2*) shows the results of an extensive Monte Carlo simulation study to compare the performance of two prismatic space-time scan statistics.

Table 1: Significant clusters detected by Takahashi et al's *prismatic* scan statistic and significant outbreaks detected by the proposed *prismatic* scan statistic applied to data from weekly surveillance of the absentees in 131 primary school districts in Kita kyushu shi, during October 11 to December 20, 2006, Japan, where the maximum temporal length is set as $T = 2$. Bold number indicates the temporal period detected.

Current Week t_p	Takahashi *et al.*'s *prismatic* scan					Proposed *prismatic* scan				
	detected areas (p-value)	No. schools Z^*	baseline $\mu^{a)}_{+,t_p}$	$n^{a)}_{+,t_p-1}$	n_{+,t_p}	detected areas (p-value)	No. schools Z^*	baseline μ_{+,t_p}	n_{+,t_p-1}	n_{+,t_p}
10/11										
10/18										
10/25										
11/01										
11/08										
11/15						O_1 (0.012)	8	65	**99**	**111**
						O_2 (0.016)	11	68	**81**	**115**
						O_2 (0.018)	9	78	**114**	**117**
11/22	C_1 (0.001)	12	75	**120**	**144**	O_1 (0.003)	13	81	**123**	**155**
11/29	C_1 (0.001)	11	76	**141**	**98**					
12/06	C_1 (0.001)	5	33	42	**80**	O_1 (0.002)	7	49	**70**	**106**
12/13	C_1 (0.001)	13	93	**166**	**183**	O_1 (0.002)	12	90	**162**	**178**
						O_2 (0.003)	12	88	**121**	**159**
						O_3 (0.003)	10	53	**89**	**105**
						O_4 (0.003)	9	78	**90**	**148**
						O_5 (0.012)	11	90	**123**	**142**
12/20	C_1 (0.015)	2	18	**44**	**36**	O_1 (0.002)	10	118	**166**	**193**
						O_2 (0.013)	11	57	**97**	**90**

[a)] $\mu_{+,t} = \sum_{i \in Z^*} \mu_{it}$ and $n_{+,t} = \sum_{i \in Z^*} n_{it}$

Table 2: Four kinds of estimated powers of Takahashi *et al.*'s space-time scan (TP) and the proposed prismatic space-time scan (PP) by simulation scenario and outbreak area, where the significance level $\alpha = .02$ and 1000 trials for each scenario were carried out. Estimated size or the probability of false alarm: Takahashi *et al.*=0.280 and Proposed with prismatic scan= 0.022.

Scenarcio	Me th od	circular area $\boldsymbol{A}$ with large μ_{it}				circular area $\boldsymbol{B}$ with small μ_{it}				Area $\boldsymbol{A}+\boldsymbol{B}$			
		Q_1	Q_2	Q_3	Q_4	Q_1	Q_2	Q_3	Q_4	Q_1	Q_2	Q_3	Q_4
S_1: 1.5	PP	0.1	2.2	3.8	8.6	0.0	1.6	1.9	4.3	0.0	4.4	6.3	11.1
	TP	0.1	0.1	0.4	64.6	0.0	0.0	0.0	23.6#	0.0	0.0	0.4	58.3
S_2: 2.0	PP	3.3	36.8	63.8	69.9	2.7	15.7	18.5	20.2	0.1	23.0	54.9	77.4
	TP	0.5	0.8	5.0	92.8	0.1	0.4	0.4	22.5#	0.0	0.5	3.7	91.3
S_3: 1.5-1.5	PP	0.4	4.4	11.6	22.2	0.0	2.0	2.6	4.5	0.0	6.7	14.2	24.0
	TP	0.0	0.0	0.8	91.3	0.0	0.1	0.1	25.2#	0.0	0.0	0.6	91.3
S_4: 2.0-2.0	PP	19.7	51.2	83.3	95.3	9.3	32.9	37.7	39.6	1.6	51.7	84.0	97.0
	TP	1.5	2.3	12.9	99.9	3.4	5.8	9.0	35.9	0.1	1.8	15.4	100
S_5: 1.5-2.0	PP	8.9	34.6	65.5	82.7	3.0	17.8	20.2	22.2	0.5	35.5	66.2	85.2
	TP	0.3	0.4	3.9	99.1	1.2	1.9	2.4	25.8#	0.0	0.3	4.7	99.3
S_6: 1.5-1.2	PP	0.0	0.2	0.5	4.4	0.0	0.7	0.7	2.4	0.0	0.7	1.1	4.9
	TP	0.0	0.0	0.1	75.4	0.0	0.1	0.1	25.5#	0.0	0.0	0.2	71.1
S_7: 1.5-2.0-2.5	PP	37.5	77.5	96.4	100	17.1	61.3	67.6	68.8	7.6	73.5	96.2	100
	TP	6.8	8.4	26.3	100	13.6	25.6	33.9	62.1	1.3	10.7	36.0	100
S_8: 1.5-1.2-2.0	PP	0.7	15.8	24.5	31.9	0.3	6.4	7.5	9.1	0.0	19.0	27.5	36.0
	TP	0.1	2.5	6.5	96.4	0.2	1.7	2.0	55.8	0.0	2.3	6.6	95.2

Q_1: the probability of exact detection $P(s^*, s^*)$.

Q_2: the probability that regions detected are all outbreaked regions, $\sum_{s=1}^{s^*} P(s, s)$.

Q_3: the probability that regions detected include at most two non-outbreaked regions, $\sum_{s=1}^{s^*} \sum_{j=0}^{2} P(s+j, s)$.

Q_4: the usual power

#: The usual power is less than its size.