

特集：地域診断・症候サーベイランスに向けた空間疫学の新展開

疾病集積性の検定を用いた症候サーベイランス解析

高橋邦彦，丹後俊郎

国立保健医療科学院技術評価部

Syndromic Surveillance Using a Cluster Detection Test

Kunihiko TAKAHASHI, Toshiro TANGO

Department of Technology Assessment and Biostatistics

National Institute of Public Health

要旨

突発的な健康危機事象を早期発見することは保健医療政策上において重要な課題である。実際，米国などでは定期的な観測データを用いた症候サーベイランスが運用されており，その解析を行う統計手法として cylindrical scan statistic とそのソフトウェア SaTScan は広く利用されている。本論では，そのようなスキャン統計量に基づくサーベイランスの手法について論じる。またいくつかのサーベイランスデータをもちいた適用例についても検討する。

キーワード： 症候サーベイランス，スキャン統計量，疾病集積性，集積性の検定

Abstract

Early detection of disease outbreaks enables public health officials to implement disease control and prevention measures at the earliest possible time. A time periodic geographical disease surveillance system based on a cylindrical space-time scan statistic has been used extensively for disease surveillance along with the SaTScan software. In this paper, we discuss a syndromic surveillance method based on the prospective space-time scan statistic. Some surveillance data are used to illustrate the method with real data.

Keywords: syndromic surveillance, scan statistic, disease clustering, clustering test

1 はじめに

2001年アメリカにおける炭疽菌によるバイオテロリズムの発生，2002年に中国で始まり，その後他国へも広がっていった SARS (Severe Acute Respiratory Syndromes; 重症急性呼吸器症候群) の発生，さらに最近の新型インフルエンザの世界的な大流行 (パンデミック) の懸念など，近年我々の健康を脅かす様々な問題が出現してきている。そのため，このような脅威に対して対策を講じることが保健医療・公衆衛生上の重要な課題のひとつとなってきた¹⁾。特にここ数年，欧米を中心に症候サーベイランス (Syndromic Surveillance) やバイオサーベイランス (Biosurveillance) とよばれるサーベイランスを目的とし

た取り組みが活発になっている。わが国においても健康危機管理の一環として，その情報や予防，対応などが重要な課題とされてきている。なかでも健康危機事象の発生をいち早く発見することを目的としたサーベイランスの重要性は国際的に高まってきている。特に欧米ではバイオテロリズムを対象とした議論が活発であり，実際に米国では2001年9月11日のテロの発生以降いくつかのサーベイランスシステムが稼動し日々監視が行われている²⁾。例えば，Washington, DC における ESSENCE (the Early Notification of Community-Based Epidemics system) や New York における NYC-DOHMH (the New York City Department of Health and Mental Hygiene) system などがある。近年，国際疾病サーベイランス学会 (the International Society for Disease

〒351-0197 埼玉県和光市南 2-3-6

2-3-6 Minami Wako-shi, Saitama-ken 351-0197, Japan.

FAX: 048-469-3875 email: kunihiko@niph.go.jp

Surveillance (ISDS)) の主催する会議 (Syndromic Surveillance Conference) などでも、サーベイランスに関する様々な発表・討論が行われている³⁾。

公衆衛生におけるサーベイランスとは米国の CDC (The Centers for Disease Control and Prevention; 疾病予防管理センター) では以下のように定義されている⁴⁾。

the ongoing, systematic collection, analysis, and interpretation of health data essential to the planning, implementation, and evaluation of public health practice, closely integrated with the timely dissemination of these data to those who need to know. The final link of the surveillance chain is the application of these data to prevention and control. A surveillance system includes a functional capacity for data collection, analysis, and dissemination linked to public health programs. (Thacker, 1994)

この定義のように、一般にサーベイランスの議論はそのデータの収集方法、システムから議論され、最終的には予防やそのコントロールまで包括して検討されるものである。しかし解析部分においては統計学が重要な役割を果たすことになる⁵⁾。バイオテロリズムのように突発的な症候の発生を発見するためには、日頃から関連の症状の発生状況を監視しておき、患者数が通常の状態に比べ突発的に集中した場合、それが重要なシグナルになっていると考えることができる。もちろん疾病によっては通常ではなかなか起きないもので、1件でも患者が発見されれば直ちに対応が必要なものもあるが、一般的には、日常的に似たような症状・疾患が少数ながらも起きてもおかしくないものも多い。このようなデータからシグナルを統計的に検出するため、いくつかの手法が用いられる。ある特定の地点における発生状況を追い続け、その変化を検出するための Process Control Charts などは代表的なものである。また、地域的な変化も観察するための方法としては、疾病地図を推定する手法なども適用される。

なかでも疾病集積性の検定 (disease clustering test) は、そのままサーベイランスの問題に適用可能であり、これらの解析の重要な手法のひとつとなっている。たとえば、ある地域において突発的な事象が発生したかどうかを検出するためには、日頃から定期的に観測される経時データを用いた伝統的な時間集積性の解析手法が利用できる。観測された疾病の頻度が通常期待される頻度に比べて大きい現象が、ある時期に集中していれば、そこで何らかの事象が発生したと考えられ重要なシグナルとなる。この場合、疾病集積性の検定を行うことで集積 (cluster) の有無が判定できる。一方、ある時点における空間 (平面) 的な疾病の集積性の検討には空間集積性 (地域集積性) の解析手法が用いられる。この解析では、ある時点において、注目してい

る事象 (死亡など) が対象地域内のどこかに集積しているかどうかを判定することができる。なかでも集積の有無の判定と同時に集積地域の同定も行う Cluter Detection Test (CDT) は有用であり、とくに Kulldorff による spatial scan statistic⁶⁻⁷⁾ を用いる方法と、そのアプリケーションソフトウェア “SaTScan” は空間疫学の研究において広く利用されている。しかし最近、Kulldorff の方法の問題点として非円状の地域をうまく同定できないことが指摘され、それを改良した方法として Duczmal and Assunção⁸⁾, Patil and Taillie⁹⁾, Tango and Takahashi¹⁰⁾, Assunção *et al.*¹¹⁾, Kulldorff *et al.*¹²⁾ などいくつかの方法が提案されている。

さて、バイオテロリズムの発生など突発的な事象を早期発見するサーベイランスの目的としては、それが「いつから」発生したか、「どこで」発生したか、空間的・時間的の両面の検出が重要となる。この目的のために、疾病の空間・時間集積性 (space-time clustering) の検定を利用した方法がいくつか提案されているが、現時点で実際の解析に即利用できる形で提供されている手法は多くない。そこで Kulldorff¹³⁾ は空間集積性の検定のための spatial scan statistic に時間集積性の検出も含めた空間・時間スキャン統計量 (space-time scan statistic) を提案した。この方法により突発的な事象が「いつ」「どこで」発生したのかを検出し、その3次元的な cluster を同定することが可能となる。実際アメリカのいくつかのサーベイランスシステムにはこの方法が取り入れられ、日々解析が行われている。さらにこの改良として Takahashi *et al.*¹⁴⁾ による flexible space-time scan statistic も提案されている。本論では、これらの space-time scan statistic による検定に焦点をあて、サーベイランスのための解析について論じる。

2 後ろ向き研究と前向き研究

ある地点において継続して観測されたデータから、その発生の集積性を検出するためには、時間集積性の検定法が利用できる。また地域的な集積性も同時に検出するためには空間・時間集積性の検定法が利用できる。一般的にこれらの方法は過去に得られたデータから集積のあった (過去の) 集積性を検出していることになる。つまり後ろ向き (retrospective) の方法である。しかし、バイオテロリズムの発生の監視のようなサーベイランスを目的とした解析においては、すでに終結した集積性の発見よりも、むしろ現時点でも続いている集積を、発生時点から時間をあけずそれをいち早く発見・同定することが重要である。そのため Kulldorff¹³⁾ では、解析時点を含んだ「**生きているクラスター (alive cluster)**」を同定する前向き (prospective) の方法を提案している。この解析によって、解析時点において、まさに起こっている突発的な事象の発生が「いつから起きていたのか」を検出することができる。そのためサーベイランスの解析では、(目的に応じて) 短い間隔で定期的 (例えば毎日1回や毎週1回など) に解析するこ

とが求められるのである。

ところで、一般的な統計的検定の有意性の判定基準としては、 p 値が0.05や0.01などの値を用いることが多い。しかし、日々のサーベイランスにおいて5%の確率で起こるということは、 $1/0.05 = 20$ でほぼ20日に1回の頻度で起こっても不思議がないということになる。このように「〇〇日に1回の頻度よりも稀である」という期間を Recurrence Interval (RI) という¹⁵⁾。その考えから、毎日行われるサーベイランスにおいては、その有意性の判定基準として、ほぼ1年に1回の頻度よりも稀な現象 (RI=365日) に対応する $p = 1/365 = 0.0027$ がひとつの基準として用いられることもある。週単位のデータであれば、1年間の52週前後を基準として $p = 1/50 = 0.02$ などが適当であろう。

なお、発生直後の集積は一般的に (i) 空間的に集積地域は狭い地域、(ii) 時間的に発生時点は解析時点に近い最近、であることが想定されるであろう。また感染性の症候などは、いったん発生した後、徐々に近隣の地域へ広がっていくことなども考えられよう。このような空間的な広がりを把握するためのモニタリングもサーベイランスにとって重要な問題であるが、この目的のためにも空間集積性、もしくは空間・時間集積性の検定法などが適用可能となる。

3 サーベイランスのための Cluster Detection Test

時間的なデータを観察しながら集積性を検出する方法としては時間集積性の検定法が利用できる。一方で、サーベイランスを目的とした場合、その発生時点とともに発生地域の同定が重要となる。そこで Kulldorff¹³⁾ は平面上での疾病集積性の検定である spatial circular scan statistic に時間のデータを入れた cylindrical space-time scan statistic (SaTScan) を提案した。

いま、解析を行う対象地域 G が m 個の region (市区町村, counties, zip codes など) に分割されているものとし、各 region に対して、時点 $Y_1, Y_1 + 1, Y_1 + 2, \dots, Y_2$ のデータが存在するとする。この対象地域において、 $[Y_1, Y_2]$ の時点では集積が存在しない、つまり定常的に症候が観測されているということが帰無仮説となる。この帰無仮説のもとでは、 i 地区の時点 d における case の数 (観測数) N_{id} が互いに独立に期待値 μ_{id} の Poisson 分布

$$N_{id} \sim \text{Poisson} (\mu_{id})$$

に従うとし、その観測値を n_{id} とする。ただし、 μ_{id} は i 地区の人口に比例する値、または性・年齢などの共変量を調整した期待観測数とする。ここで、日々のサーベイランスを目的とした解析のため、解析時点を含んだ cluster のみを考える前向き研究として prospective space-time 解析を考える。つまり、時点 Y_2 のデータが得られた段階で解析

を行い、クラスターの時間 $[s, t]$ として常に $t = Y_2$ となる集積のみを考える。これによって同定されるクラスターの時間は $[s, Y_2]$ となる alive cluster になる。さらに、発生時点 s も解析時点 Y_2 からそれほど遠くない時点であると想定される。そこで、実際にデータは $[Y_1, Y_2]$ の全てを用いるが、発生時点はある時点 Y_T (Y_2 から T 時点前) 以降である、と考え、 $[s, t]$ として

$$[Y_2 - T + 1, Y_2], [Y_2 - T + 2, Y_2], \dots, [Y_2 - 1, Y_2], [Y_2, Y_2]$$

のいずれかであるような cluster を探しだすことになる。この方法が prospective scan 法 (図1) であり、このときの T を maximum temporal length という。

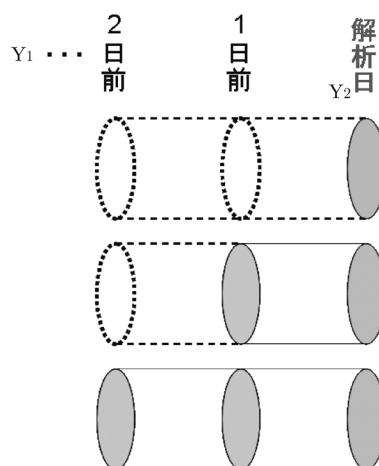


図1 prospective space-time 解析の模式図

さらに $Z_{ik} (k=1, 2, \dots, K)$ を region i から近い順に、 i 自身を含む k 個の region からなる集合とする。ただし各 i の座標はその region の代表点1点 (市区町村役場の所在地など) であらわすものとする。ここで K は予め定めておく maximum geographical length とする。このとき cylindrical scan statistic では、底面を

$$\mathcal{Z}_1 = \{Z_{ik} \mid 1 \leq i \leq m, 1 \leq k \leq K\}$$

および高さを

$$\mathcal{Y} = \{[Y_2 - s + 1, Y_2] \mid 1 \leq s \leq T\}$$

とした全体集合 $W = \mathcal{Z}_1 \times \mathcal{Y}$ からの要素となる円柱状の window W を考える。

しかし、この方法は底面が円状の正円柱状の window だけを考えているため、平面上での circular scan 同様、非円状の地域を同定することは出来ない。そこで Takahashi *et al.*¹⁴⁾ は spatial flexible scan における window $Z \in \mathcal{Z}_2$ を底面にもつ多角柱の window を scan することで非円状の地域の space-time の解析が可能となる flexible space-time

scan 法を提案した. まず region i を中心として i 自身を含み i から近い順に K 個の地域からなる集合 Z_{ik} を定める. この Z_{ik} から, i を含み, 連結している部分集合を考え, その全体 Z_2 を考える. つまり Z_{ik} の中で i を含んで k 個の region からなる連結したものが j_{ik} 個あるとすると, 底面を

$$Z_2 = \{Z_{ik(j)} \mid 1 \leq i \leq m, 1 \leq k \leq K, 1 \leq j \leq j_{ik}\}$$

とする $W = Z_2 \times \mathcal{Y}$ から要素となる window W を考える.

ここで window W における観測数を $N(W)$ とし, その観測値を $n(W)$, さらに W が cluster でないときの期待観測数を $\mu(W)$ とする. cluster の有無の検定として

$$\text{帰無仮説: } E(N(W)) = \mu(W) \text{ for } \forall W \in \mathcal{W}$$

$$\text{対立仮説: } E(N(W)) > \mu(W) \text{ for } \exists W \in \mathcal{W}$$

を考え, window W が cluster となる対立仮説のもとでの尤度を $L(W)$ とし, 帰無仮説のもとでの尤度を L_0 とする. 尤度比統計量

$$\lambda(W) = \left(\frac{L(W)}{L_0} \right) = \begin{cases} \left(\frac{n(W)}{\mu(W)} \right)^{n(W)} \left(\frac{\mu(W)}{n(W)} \right)^{n(W)} & (n(W) > \mu(W)) \\ 1 & (\text{その他}) \end{cases}$$

を考え

$$\lambda^* = \lambda(W^*) = \max_{W \in \mathcal{W}} \lambda(W)$$

のように最大尤度比 λ^* をとる window W^* を most likely cluster (MLC) とし, これを cluster の候補と考える. ここで, この MLC が統計的に有意な集積性をもつかどうかの評価が必要となる. そのため帰無仮説のもとでの $\max_{W \in \mathcal{W}} \lambda(W)$ の分布を使って有意性を見るが, 一般的には Monte Carlo 法を利用して数値的に求めた p 値によってその有意性が検討される. その結果有意となれば, W^* が有意な cluster となり, その地域, その期間に疾病が集積していたと判定できる.

4 米国 Massachusetts 州東部のサーベイランスデータ

このデータは Harvard Vanguard Medical Associate によって利用されているデータで, zip code ごとにまとめられた電子的な医療記録として毎日集計され¹⁶⁻¹⁷⁾ SaTScan を利用して毎日解析が行われている. ここでは2005年8月の呼吸器疾患 (respiratory) の発生について, Kulldorff の cylindrical scan 法 (SaTScan) と Takahashi *et al.* の flexible scan 法 (FlexScan) によって解析を行った. 解析に利用したデータ, パラメータは以下のとおりである.

• データ:

2005年8月における respiratory の発生

Massachusetts 州東部にある $m = 385$ の zip codes からなる地区

• 解析:

解析モデル: Poisson モデル

2005年8月7日から30日まで毎日解析

scan する window の平面における最大長 (Maximum geographical length): $K = 20$ area

scan する window の最大時間 (Maximum temporal length): $T = 7$ 日

Monte Carlo シミュレーションの回数: 999

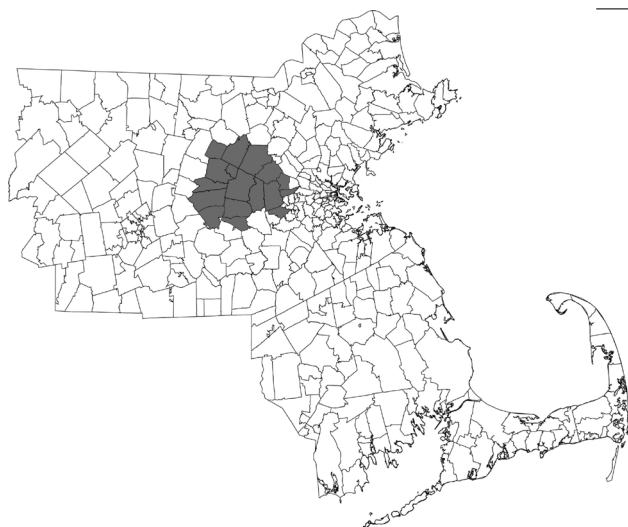
有意と判断する recurrence interval: 6 ヶ月以上 (すなわち $p < 0.0054$)

なお, 期待観測数は同地域における過去1年分以上のデータを用い, 性別, 年齢以外にも季節, 月, 曜日などの影響も調整されたものを利用している.

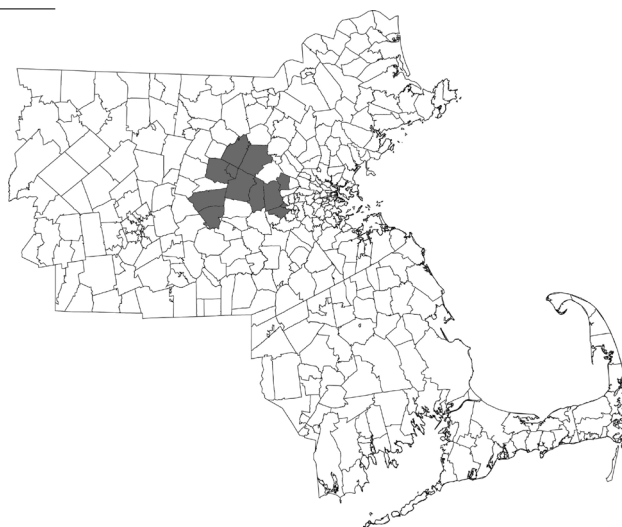
この解析によって, 8月12日~15日の解析において有意な集積が検出された (図2). まず8月12日において, SaTScan, FlexScan とも $RI=2.7$ 年 ($p=0.001$) の集積が検出された. SaTScan においては18地区からなる地域で8月11日~12日の2日間に集積が同定された. 一方, FlexScan では12地区からなる地域が同定され, 微妙に同定された地域が異なっていた (11地区は共通). 8月13日, 14日においては SaTScan は12日と同じ地域を同定したが, それらの RI は短く有意とはならなかった. 一方, FlexScan では13地区からなる有意な集積地域が同定された. 8月15日ではそれぞれ有意な集積地域が同定されたがその集積の時間が異なった (SaTScan では5日間, FlexScan では6日間). このとき FlexScan では集積地域がまた少し変化した. SaTScan では集積地域の変化の様子は観察されなかった.

5 北九州市内小学校欠席数によるサーベイランス

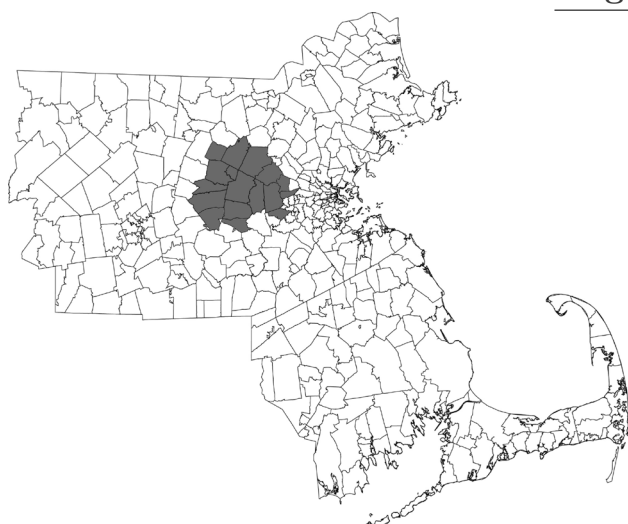
次に, 小学校の欠席児童数によるサーベイランス解析について検討を行った. このデータは2006年度福岡県北九州市内の132小学校 (全児童数52,189名) について, 毎週水曜日 (小学校の休校日は除く40週) にその日の欠席者数が報告されたものである. ここでは特に目立つ欠席増減がなかった4~9月 (18週) の小学校別平均欠席児童数を各小学校のベースライン (期待欠席数) として用い, 10月11日から12月冬休み前まで毎週解析を行った. 解析には FlexScan (maximum length: $K=15$, $T=2$) とし, その有意性の判定として $p < 0.02$ ($RI \approx 1$ 年) を用いた. その結果, 11月15日までの解析 (10月11日, 18日, 25日, 11月1日, 8日, 15日) では有意な集積は検出されなかった. しかし, 11月22日に有意な欠席数の集積が検出され, その後, 冬休み前の12月20日までの5週間にわたり集積が認められた (表1, 図3). 特に12月20日においてはその欠席の集積地域が小さい範囲になっている様子が観察された.

Aug. 12**SaTScan**

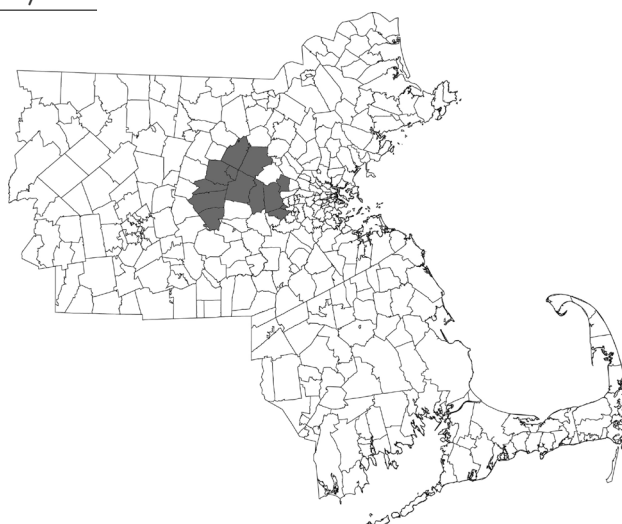
Cluster period: Aug.11–12
 RI=2.7 years ($p = 0.001$)
 No. of cases: 51, RR: 2.25
 No. of zip codes: 18

**FleXScan**

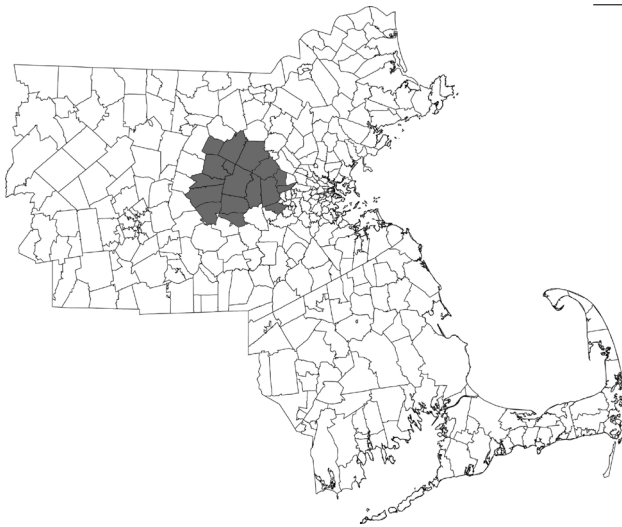
Cluster period: Aug.11–12
 RI=2.7 years ($p = 0.001$)
 No. of cases: 42, RR: 3.37
 No. of zip codes: 12

Aug. 13 / 14**SaTScan**

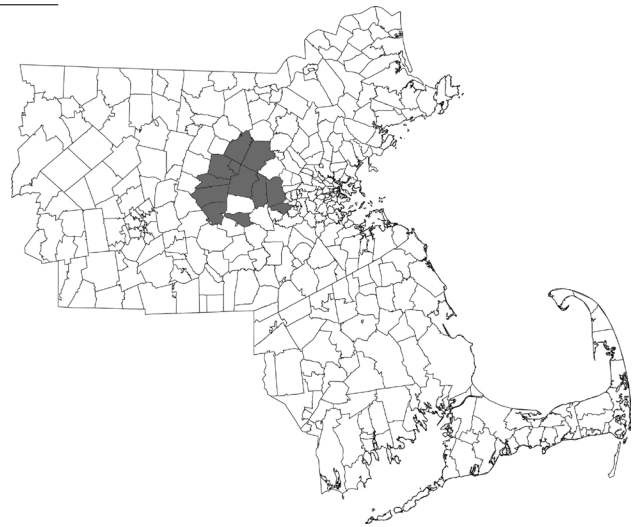
Cluster period: Aug.11–13 / 11–14
 RI=91 days ($p = 0.011$) / 30 days ($p = 0.033$)
 No. of cases: 55 / 59, RR: 2.31 / 2.19
 No. of zip codes: 18

**FleXScan**

Cluster period: Aug.11–13 / 11–14
 RI=333 days ($p = 0.003$) / 250 days ($p = 0.004$)
 No. of cases: 46 / 49, RR: 3.08 / 2.89
 No. of zip codes: 13

Aug. 15**SaTScan**

Cluster period: Aug.11-15
 RI=200 days ($p = 0.005$)
 No. of cases: 82, RR: 2.00
 No. of zip codes: 18

**FleXScan**

Cluster period: Aug.10-15
 RI=1.4 years ($p = 0.002$)
 No. of cases: 72, RR: 2.40
 No. of zip codes: 12

図2: Massachusetts 州における respiratory の発生の集積

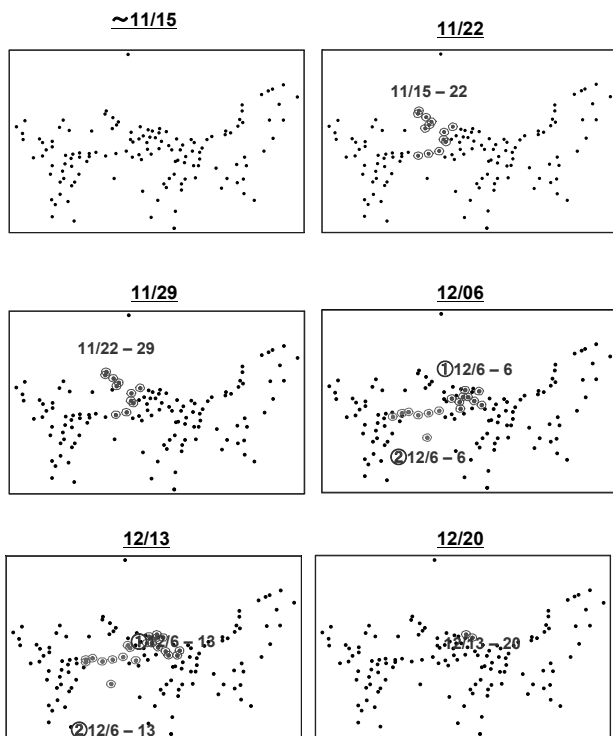


図3 北九州市小学校欠席サーベイランスで検出されたクラスター (10月～12月). 点は小学校の座標を表わす。

表1 北九州市内小学校欠席数によるサーベイランス解析の結果 (10～12月の各週)

解析日	校数	cluster period	cases	expectd	p-value
11月22日	13校	11月15～22日	283	191.4	0.001
11月29日	11校	11月22～29日	238	164.9	0.001
12月6日	9校	12月6～6日	141	84.5	0.002
	7校	12月6～6日	84	45.8	0.006
12月13日	11校	12月6～13日	305	225.4	0.001
	10校	12月6～13日	203	142.3	0.006
12月20日	2校	12月13～20日	80	44.3	0.012

もちろん、この解析は欠席数のみからの検討であり、その欠席理由がなんであるかは特定できない。しかし、事後的に得られたこの期間の北九州市内の定点観測での感染性腸炎の発生 (図4) と見比べると、この結果がうまく符合していることがうかがえ、この欠席は感染性腸炎のものであったのではないかと推測される。

なお、1月以降の解析においては3月初旬まで有意な集積が検出された。これはインフルエンザの流行によるものと推測される。

6 サーベイランス解析における解釈

統計学的に有意な集積性が検出された場合でも、サーベイランスにおいてはいくつかの注意が必要となる。たとえば先に述べたように、 $p < 0.05$ となったものが直ちに問題として考えることはせず、そのデータに応じた RI などを考えることになる。データが日単位のものなのか、週単位

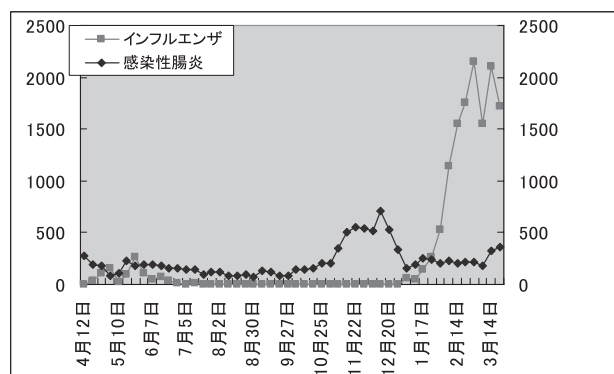


図4 北九州市内定点観測によるインフルエンザと感染性腸炎の発生報告件数

のものなのかでそのシグナルの有意性の判断が違ってくることになる。

また、単に有意性だけみればよいわけではない。表2は、先ほどの例と同じく Massachusetts において2005年8月に発疹 (Rash) の発生を解析した結果 $p < 0.0054$ となったものである。この解析では SaTScan, FleXScan と同じ結果が得られ、8月7日～10日の解析で、1箇所の zip code だけが同定され、その有意性はいずれも強いものであった ($p = 0.001$)。ここから「この突発的な発生は8月2日であり、そこから10日までおきていた」と解釈されそうである。しかし、この結果をよく眺めると、8月7日の解析では、8月2日～7日の6日間が同定され、その観測数は7件であった。次に8月7日の解析でも集積性が検出されたが、そこで同定されたクラスターは8月2日～8日の7日間で観測数は7件のままである。つまり、8月8日にはこの地区で疾病発生は起きていないことが読み取れる。さらに、9日、10日には、期待観測数が同じであるにも係わらず、観測数は減少している。つまりこの場合、8

月7日の解析で検出されたクラスターは検討しなくてはいけませんが、8日～10日の有意なクラスターはそれほど問題にならないと考えられよう。

また、同定されたクラスターの情報 (症候、場所、期間など) を検討し、このクラスターが本当に問題となるかどうかを検討する必要がある。例えば発生した曜日の問題 (週明けは患者が多い) や特定のイベントがあった日など、様々な検討がされる。これは統計学者だけではなく、疫学や医学、保健医療、行政、データ解析など、それぞれの専門家がチームとなって検討されるべきものである。

さらに、北九州市内の小学校欠席児童数による解析では、特定の疾患などの情報はなく、理由を問わず学校欠席のデータから集積が検出されるだけであり、その欠席理由までは断定できない。むしろこの結果で検出された集積地域で「何か異常がおきているのではないか」、「感染性の疾患が発生しているのではないか」という疑いが出たのであって、実際のサーベイランスとしては、同定された小学校の詳細調査を即時に行うことで、その原因究明が必要となるであろう。しかし、定常的に観測し続けるなかで何らかの健康危機事象の発生をいち早く見つけ出すという観点では、この小学校欠席数によるサーベイランスも有用であると考えられる。

本論のはじめに述べたように、近年、サーベイランスのための研究は重要となってきた。その解析手法として空間疫学における手法は有用なツールとなっている。サーベイランスを目的とした手法の開発もさらに進んでいくことになるであろう。しかし、同じくはじめに述べたように、サーベイランスの検討は解析手法だけではなく、データの収集から解析、それを行うシステム、さらに対策までを総合的に考える必要もある。わが国においても健康危機管理のひとつとしてこれらのサーベイランスがますます重要な課題になってくると考えられる。

表2: Detected outbreaks of Rash based on daily syndromic surveillance data in eastern Massachusetts during August 1-30, 2005.

Day	zip codes	cluster period	cases	expectd	R.I. (p-value)
Aug.07	01951	Aug.02-07	7	0.0427	2.7years (0.001)
Aug.08	01951	Aug.02-08	7	0.0545	2.7years (0.001)
Aug.09	01951	Aug.03-09	6	0.0545	2.7years (0.001)
Aug.10	01951	Aug.04-10	5	0.0545	2.7years (0.001)

参考文献

- 1) 丹後俊郎, 横山徹爾, 高橋邦彦. 空間疫学への招待. 東京: 朝倉書店; 2007.
- 2) Brookmeyer R, Stroup DF, editors. Monitoring the health of populations. New York: Oxford University Press, Inc.; 2004.
- 3) International Society for Disease Surveillance. <http://www.syndromic.org/>
- 4) Lawson AB, Kleinman K, editors. Spatial & syndromic surveillance for public health. Chichester: John Wiley & Sons; 2005.
- 5) Wilson AG, Wilson GD, Olwell DH, editors. Statistical methods in counterterrorism. New York: Springer; 2006.
- 6) Kulldorff M, Nagarwalla N. Spatial disease clusters: detection and inference. *Statistics in Medicine* 1995; 14: 799-810.
- 7) Kulldorff M. A spatial scan statistic. *Communications in Statistics: Theory and Methods* 1997; 26: 1481-1496.
- 8) Duczmal L, Assunção R. A simulated annealing strategy for the detection of arbitrarily shaped spatial clusters. *Computational Statistics & Data Analysis* 2004; 45: 269-286.
- 9) Patil GP, Taillie C. Upper level set scan statistic for detecting arbitrarily shaped hotspots. *Environmental and Ecological Statistics* 2004; 11: 183-197.
- 10) Tango T, Takahashi K. A flexibly shaped spatial scan statistic for detecting clusters. *International Journal of Health Geographics* 2005; 4: 11.
- 11) Assunção R, Costa M, Tavares A, Ferreira S. Fast detection of arbitrarily shaped disease clusters. *Statistics in Medicine* 2006; 25: 723-742.
- 12) Kulldorff M, Huang L, Pickle L, Duczmal L. An elliptic spatial scan statistic *Statistics in Medicine* 2006; 25: 3929-3943.
- 13) Kulldorff M. Prospective time periodic geographical disease surveillance using a scan statistic. *Journal of the Royal Statistical Society, Series A* 2001; 164: 61-72.
- 14) Takahashi K, Kulldorff M, Tango T, Yih K. A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring. *International Journal of Health Geographics* 2008; 7: 14.
- 15) Kleinman K, Lazarus R, Platt R. 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* 2004; 159:217-224.
- 16) Lazarus R, Kleinman K, Dashevsky I, DeMaria A, Platt R. Using automated medical records for rapid identification of illness syndromes (syndromic surveillance): The example of lower respiratory infection. *BMC Public Health* 2001; 1: 9.
- 17) Lazarus R, Kleinman K, Dashevsky I, Adams C, Kludt P, DeMaria A, Platt R. Use of automated ambulatory-care encounter records for detection of acute illness clusters, including potential bioterrorism events. *Emerging Infectious Diseases* 2002; 8(8): 753-760.