

201134023A

厚生労働科学研究費補助金

健康安全・危機管理対策総合研究事業

健康危機事象への迅速な対応に有効なサーベイランス解析

手法・ツールの開発・普及に関する研究

平成23年度 総括研究報告書

研究代表者 高橋邦彦

平成24（2012）年 3月

## 目 次

### I. 総括研究報告

健康危機事象への迅速な対応に有効なサーベイランス解析手法・ツールの開発・普及に関する研究	-----	1
(資料)	-----	7
・健康危機事象への迅速な対応に有効なサーベイランス解析手法・ツールの開発・普及に関する研究 (成果発表会資料)	-----	9
・ On a flexible scan statistic with a restricted likelihood ratio (2011 年度日本計量生物学会年会発表予稿)	-----	15
・ 心肺機能停止傷病者救急搬送件数の時間集積性の検出 (シンポジウム「生物統計における統計的推測とその応用」発表資料)	-----	17
・ 疾病集積性の検定の適用例 (公開研究集会「イベントの空間及び時間集積性に対する統計的手法の研究」)	-----	33

# 総括研究報告

厚生労働科学研究費補助金（健康安全・危機管理対策総合研究事業）  
総括研究報告書

健康危機事象への迅速な対応に有効なサーベイランス解析手法・ツールの  
開発・普及に関する研究

研究代表者 高橋邦彦 国立保健医療科学院政策技術評価研究部主任研究官

研究要旨: 欧米の症候サーベイランス解析としても利用することができる疾病集積性の検定の一つとして、我々が提案・開発している集積性検出の手法とソフトウェア FleXScan を中心に、その改良、普及、実際への適用について検討を行った。まずソフトウェアについてユーザーからの要望をもとに改良を試みた。また、実証的な研究として、健康危機管理上重要な情報となりうる救急車による搬送件数について、その発生に関する検討を行い、年末年始時期の発生増加を検出することができた。

研究分担者

・丹後俊郎(昭和女子大学客員教授・医学統計学  
研究センター長)

・灘岡陽子(東京都健康安全研究センター主任  
研究員)

A. 研究目的

近年のバイオテロリズムやSARS、新型インフルエンザの発生などのように、その発生をいち早く検出するためのサーベイランスシステムの必要性が世界的に高まってきている。ある症候が突発的・集中的に発生するなどの事象をいち早く検出するためには、日々それらの症候の発生状況を継続的に観察・監視し、その中で事象の発生が突発的・集中的に発生したと疑われる

場合に、それが偶然なのか、または意味のあるものなのかを、客観的に判定することが必要となる。その方法として集積性の検定という統計手法が利用できる。中でも米国ハーバード大学のKulldorff博士の提案する手法とそのソフトウェア「SaTScan」は世界的に広く利用されており実際米国でのいくつかの監視システムにも組み込まれ日々解析が行われている。この方法は簡便ではあるが、一方でいくつかの問題があることが最近指摘されてきている。その問題の一つを解決するため我々が開発したflexible scan法とそのソフトウェア「FleXScan」も注目され利用され、海外の専門書 (Pfeiffer et al. “Spatial Analysis in Epidemiology” ; Shekhar and Xiong(eds.) “Encyclopedia of GIS “など)にも取り上げられている。実際、一般的な研究者やユーザーが容易に、無料で解析を行えるソフトウェアは世界においてKulldorffによる「SaTScan」と我々の「FleXScan」の2つだけ

であると認識されている状況であり、近年もこの2つのソフトウェアを用いた疫学研究などが世界的に増えてきている。さらにニューヨーク市の保健局 (New York City Department of Health and Mental Hygiene (DOHMH)) においても保健局のGIS Center of Excellenceを中心にFlexScanの利用に関する検討が始められている。本研究では、我々の提案するFlexScan法を、国内において実際のサーベイランス業務として各地で中心となると考えられる地方衛生研究所での利用の可能性、また有用性について検討を行う。さらに救急搬送件数として総務省消防庁で公開されている心肺機能停止傷病者の全国データ (ウツタインデータ) を用い、その発生の特徴について明らかにする。さらにこれらの検討を通しFlexScanのさらなる統計解析手法の開発およびツールとなるソフトウェアの改良を行う。

## B. 研究方法

### (1) 開発研究：

よりサーベイランスに適し、実用性のある統計解析手法の開発として、Tango and Takahashi (2005)で提案するflexible scan statistic (FlexScan法)を改良した新たな手法を提案した。さらにそれを組み込んだソフトウェア「FlexScan」とその改良を行った。

### (2) 普及のための研究：

手法の普及に向けて実データにあてはめた適用例として実例での実証研究について検討を行った。特に健康危機管理上の重要な情報となりうる救急搬送データについての実証研究を行った。

総務省消防庁では心肺機能停止傷病者の救急搬送の全報告が救急蘇生統計として公表されている (ウツタイン様式)。これまで目撃・バイスタンダー・心肺蘇生の有無など、救命活動による蘇生や1ヶ月後の生存率、回復程度など、救命効果に関する検討などが行われている。一般

に心肺機能停止は冬季に多いことが知られており、これまでの検討会の集計などでも明らかになっている。しかし発生に関する詳細な分析はされていない。そこで本研究では、性、年齢、都道府県、発生週、曜日、心原性・非心原性の別などで集計を行い、さらに発生日に注目し、サーベイランス解析手法を適用して、発生の時間集積性の検出を試みた。

## C. 結果

### (1) 開発研究：

Tango and Takahashi (2005)の提案するflexible scan statisticにTango (2008)の制限付き尤度比統計量を組み込んだ手法を提案したことで従来のFlexScan法で問題点であった最大地域数の制限が不要になり、さらに計算時間も大幅に短縮することができた。またシミュレーションにより集積地の同定精度もSaTScan法などの従来法よりも高いことが確かめられた。

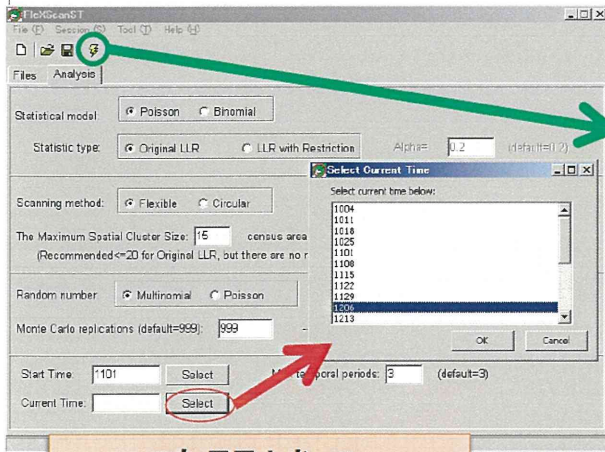
さらにソフトウェアFlexScanの改良としては時間域の範囲指定や時空間での集積地域毎に色分け表示も取り入れるよう開発をすすめている (図1)。

### (2) 普及のための研究：

心肺機能停止傷病者の救急搬送件数について詳細に日単位で検討を行ったところ、年末年始時期に特異的な増加が認められた。時間集積性の検定によって、いずれの年においても年末年始時期に最も有意な集積が検出された。地域 (都道府県)、男女別、心原性・非心原性を問わず、いずれも年末年始時期に高くなっていた。一方で検出された集積期間内とそれ以外 (12月～1月)で搬送された傷病者の年齢構成を比較しても大きな違いが見られなかった。サーベイランスの手法を用いることで、このような顕在化されていない健康危機管理上の問題を浮き彫りにすることもできた。

# 結果1

アプリケーション FlexScan の改良,  
解析方法の追加など



時間域の  
範囲指定

時空間での  
集積地域毎に  
色分け表示も

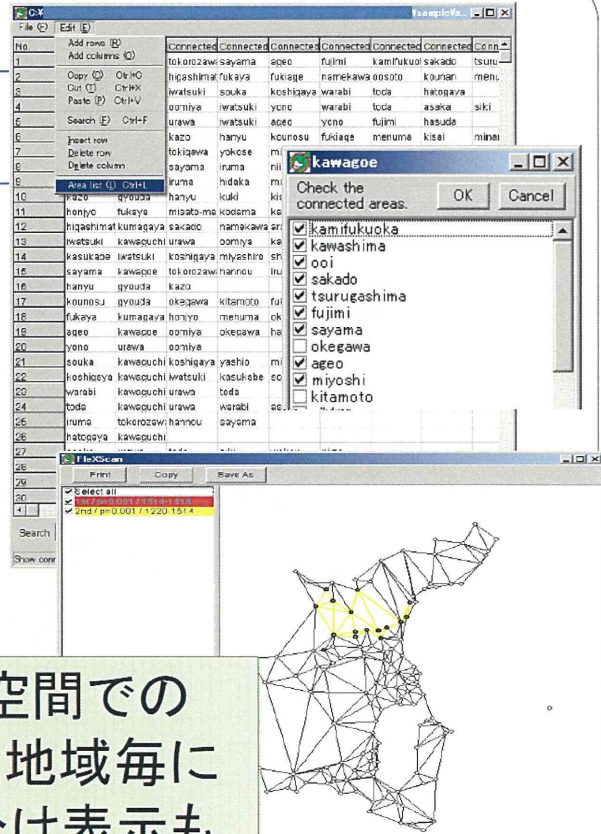


図1. ソフトウェアFlexScanの新たなバージョンの政策

## E. 結論

本年度の研究では、サーベイランスに向けたより適切な解析ができるような統計量の改良の検討について、理論的側面および実際の側面から様々な検討を行い、より適切な手法・ツールを提案したが今後も引き続き検討を行っていきさらに解析ツールも改良を行っていきまた今年度の結果を踏まえ、救急搬送データの解析を更にすすめ、サーベイランスへの利用可能性について検討を行う。また東京都をはじめ自治体で活用できるデータと実際のツールの適用について検討を続け、効果的な普及方法についても検討し試行していきたい。

## F. 健康危機情報

なし

## G. 研究発表

### 1. 論文発表

なし

### 2. 学会発表

- Tango T and Takahashi K. FlexScan with a restricted likelihood ratio: a spatial scan statistic for irregularly shaped clusters. 32nd Annual Conference of the International Society for Clinical Biostatistics, August 21-25, 2011, Ottawa, Canada.

- 丹後俊郎, 高橋邦彦. On a flexible scan statistic with a restricted likelihood ratio. 2011年度日本計量生物学会年会, 大阪大学, 2011年6月2日-3日.

・高橋邦彦, 武村真治, 金谷泰宏, 齋藤大蔵.  
わが国における年間の心肺停止傷病者救急搬送  
件数の時間集積性の検出. 第70回日本公衆衛  
生学会総会, 秋田市, 2011年10月20日.

・高橋邦彦. 心肺機能停止傷病者救急搬送件数  
の時間集積性の検出. シンポジウム「生物統計に  
おける統計的推測とその応用」, 宮崎市, 2011  
年12月11日.

・高橋邦彦. 疾病集積性の検定の適用例. 統計  
数理研究所・共同利用研究集会「イベントの空  
間及び時間集積性に対する統計的手法の研究」  
統計数理研究所, 2012年3月8日.

### 3. 著書

なし

・

### 4. その他

・FleXScan v3.1: Software for the Flexible  
Scan Statistic. 国立保健医療科学院,  
[http://www.niph.go.jp/soshiki/gijutsu/index\\_j.html](http://www.niph.go.jp/soshiki/gijutsu/index_j.html)

## 資料



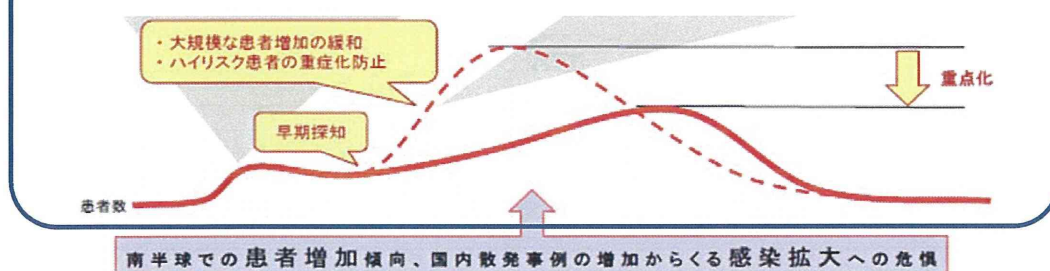
## 健康危機事象への迅速な対応に有効なサーベイランス 解析手法・ツールの開発・普及に関する研究 (H23-健危-一般-003)

研究代表者 高橋 邦彦 (国立保健医療科学院)  
研究分担者 丹後 俊郎 (医学統計学センター)  
灘岡 陽子 (東京都健康安全研究センター)

### 症候サーベイランス

- 日々の監視の中で、突発的な事象(症候の発生)が「いつ」「どこで」発生したかをいち早く発見することが重要。

厚生労働省資料より



## Surveillance System

- NY市: NYC-DOHMH (the New York City Department of Health and Mental Hygiene) system
  - Washington, DC: ESSENCE (the Early Notification of Community-Based Epidemics system)
  - Boston, Denver, Minneapolis, etc...
- そのシステムの中での統計手法として“集積性の検定”が利用され有用なツールとなっている。
- 上記の米国システムではSaTScan法が組み込まれて実際に運用されている。

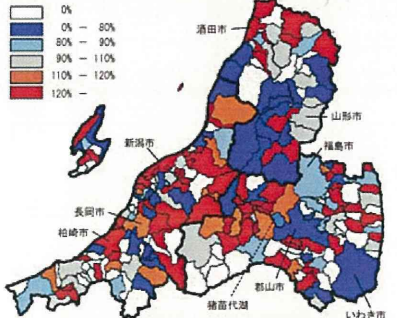
# 疾病集積性の検定

発生がどこかに集中(集積)していないか?

疾病集積性の検定

~統計的に客観的に判定~

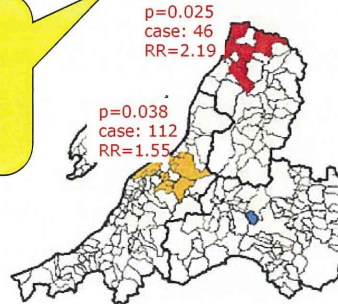
## ◎ 空間(地域)集積性



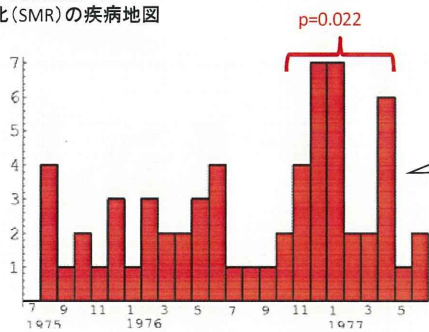
1996年~2000年男性胆のうがんの標準化死亡比(SMR)の疾病地図

SaTScanで同定できない非円状など任意の形状の地域をうまく同定可能

## Tango & Takahashi's Flexible scan statistic



## ◎ 時間集積性



Wallenstein, 1980. NYCの3病院で報告された自然流産での染色体異常の月別頻度

# ニューヨーク市: 新型インフルエンザの発生

救急病院(55病院中50病院)へのvisitを毎日報告。患者の居住地(zip code)で解析

## Flexible and Circular ILI Clusters by NYC ZIP Code Area, 4/26 - 4/28

Date	H1N1
4/26	3
4/27	2
4/28	2

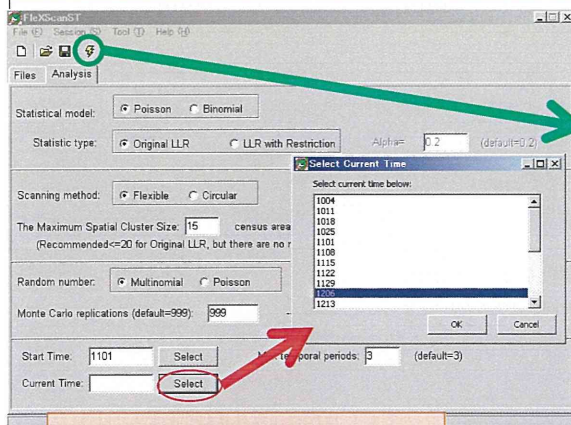
The screenshot also shows the website's navigation menu, search bar, and a list of online geographic data resources. A red box highlights 'FlexScan' in the list of tools used by the GIS Center.

# 本研究の目的・方法

- 開発研究：
  - よりサーベイランスに適し、実用性の**統計解析手法の開発**
  - それらの解析ができる**ツール(ソフトウェア)の開発**
    - ⇒ ソフトウェア FlexScanの改良。時空間解析や高速化など、より適した手法を組み込んだソフトの開発
- 普及のための研究：
  - 実データにあてはめた適用例、実例での実証研究
  - 現場のニーズに合わせて、解析を行う仕組みづくりの検討
    - ⇒ 東京都での救急搬送サーベイランスや感染症発生動向調査など実際のデータを想定し、その実現性の検討
    - ⇒ 特に健康危機管理上の重要な情報となりうる救急搬送データについての実証研究

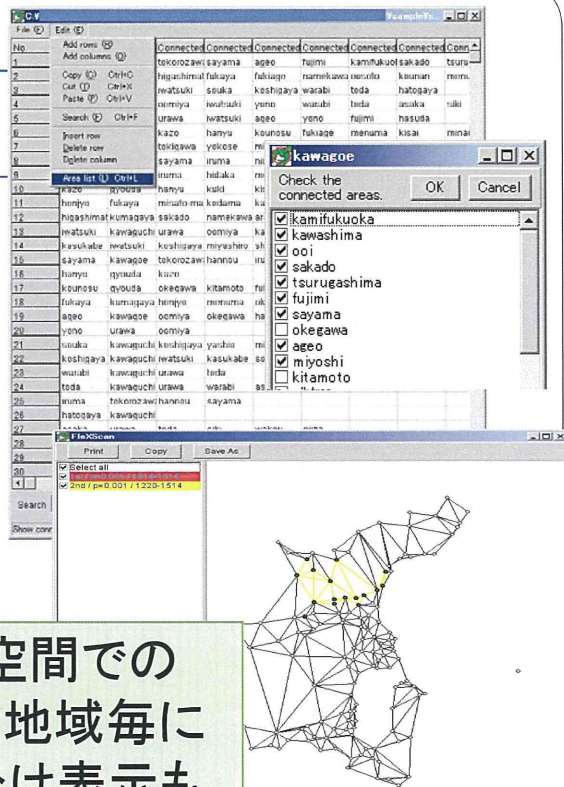
## 結果 1

アプリケーション FlexScan の改良、  
解析方法の追加など



時間域の  
範囲指定

時空間での  
集積地域毎に  
色分け表示も



## 結果2 心肺機能停止傷病者救急搬送の時間集積性の検出

- 総務省消防庁：心肺機能停止傷病者の救急搬送の全報告が救急蘇生統計として公表されている。（ウツタイン様式）
- 目撃・バイスタンダー・心肺蘇生の有無など，救命活動による蘇生や1ヶ月後の生存率，回復程度など，救命効果に関する検討。
- 一般に心肺機能停止は冬季に多いことが知られており，これまでの検討会の集計などでも明らかになっている。
- しかし発生に関する詳細な分析はされていない。
- 本研究では，性，年齢，都道府県，発生週，曜日，心原性・非心原性の別などで集計を行い，さらに発生日に注目し，サーベイランス解析手法を適用して，発生の時間集積性の検出を試みた。  
（現在，結果を論文投稿中）

図1 2005年1月1日～2007年12月31日における週ごとの心肺機能停止傷病者救急搬送件数(男女別，心原性・非心原性別)。横軸：週(日曜～土曜)，縦軸：件数。  
実線—■：総数，破線—●：心原性，▲：非心原性。1月1日が日曜でない場合は，第1週と前年最終週は同数。

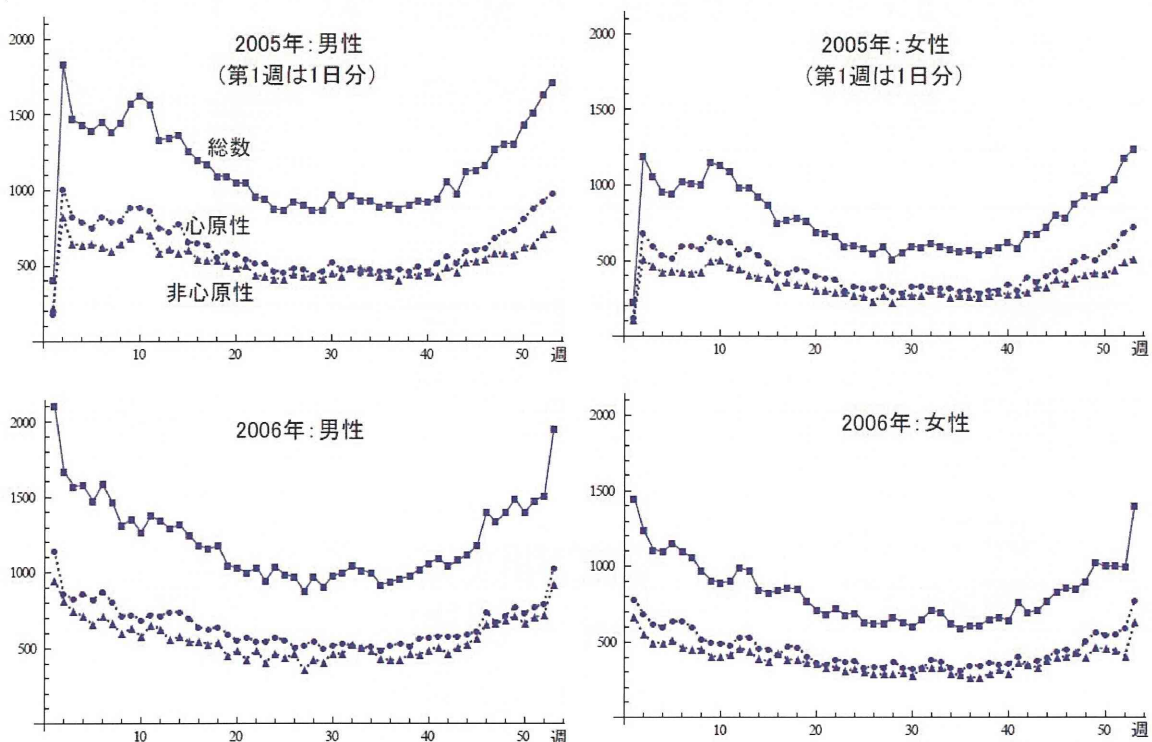
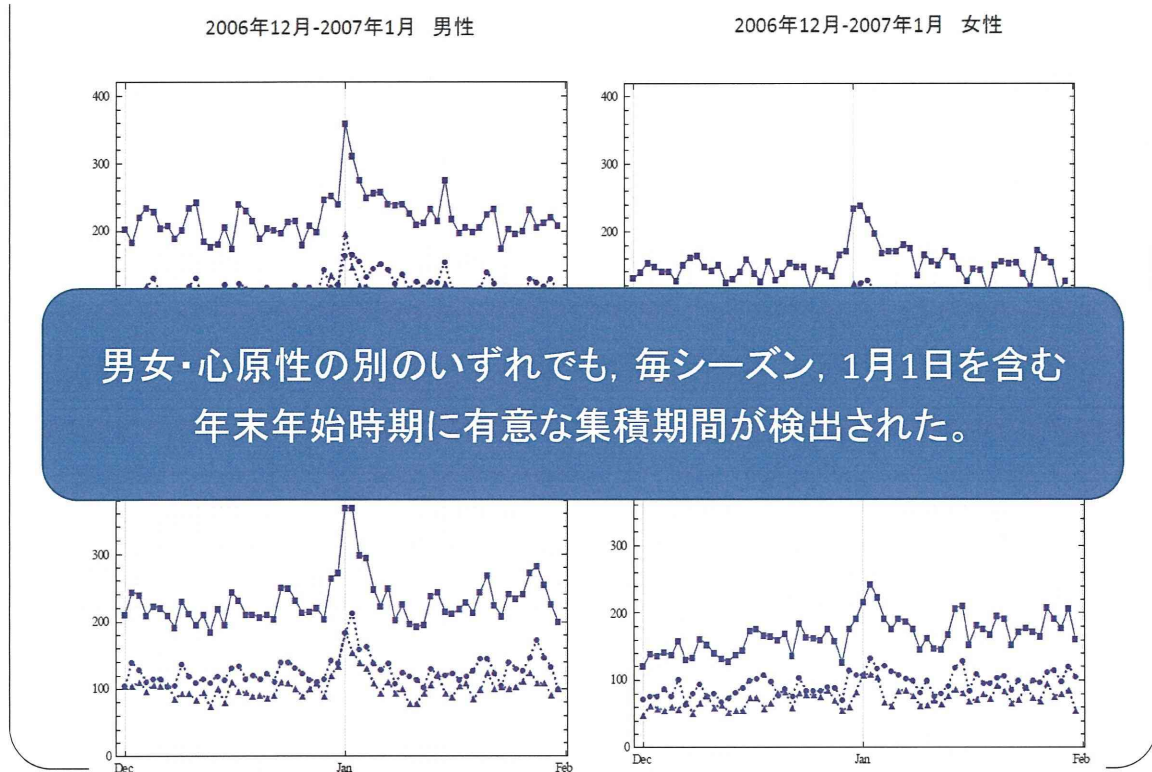


図2 年末年始における日単位での心肺機能停止傷病者救急搬送件数。横軸：日，縦軸：件数  
 実線■は総数，破線●：心原性，▲：非心原性



- 詳細に日単位で検討を行ったところ，年末年始時期に特異的な増加が認められた。
- 時間集積性の検定によって，いずれの年においても年末年始時期に最も有意な集積が検出された。
- 地域(都道府県)，男女別，心原性・非心原性を問わず，いずれも年末年始時期に高くなっている(集積が検出される)
- 検出された集積期間内とそれ以外(12月～1月)で搬送された傷病者の年齢構成を比較しても大きな違いが見られなかった。
- サーベイランスの手法を用いることで，このような顕在化されていない健康危機管理上の問題を浮き彫りにすることもできた。  
 ⇒公衆衛生上の観点からも更なる検討を行うための足がかりになる。

## 今後の研究計画

- 開発研究：
  - 更にサーベイランスに適して実用性のある統計解析手法の開発とそれを組み込んだ解析ツール(ソフトウェア)の開発  
⇒ 国内外の研究者・実務者にコメントをもらいながら進める。
- 普及のための研究：
  - 救急搬送データの解析⇒サーベイランスへの利用可能性
  - 東京都をはじめ自治体で活用できるデータと実際のツール適用についての検討を続ける。  
⇒ 効果的な普及方法について検討し試行する。

**On a flexible scan statistic with a restricted likelihood ratio**T. Tango<sup>1</sup>, K. Takahashi<sup>2</sup><sup>1</sup>Center for Medical Statistics, Tokyo, Japan, Graduate School of Human Ecology, Showa Women's University, Tokyo, Japan,<sup>2</sup>Dept. Technology Assessment and Biostatistics, National Institute of Public Health, Saitama, Japan  
[tango@medstat.jp](mailto:tango@medstat.jp)

Spatial scan statistics are important and widely used tools for cluster detection. Especially, SaTScan (circular version) proposed by Kulldorff (Com Stat: Theory and Methods, 1997) has been utilized to a wide variety of epidemiological studies and disease surveillance. However, as it cannot detect non-circular irregularly shaped clusters, many authors have proposed different spatial scan statistics, including an elliptic version of SaTScan. Above all, FleXScan proposed by Tango and Takahashi (Int J Health Geo, 2005) has also been used for detecting irregularly shaped clusters. However, this method allows a maximum of 30 regions in cluster size due to heavy computational load. Many authors who conducted comparison of tests for disease clusters pointed out this limitation. Tango (Jap J Biometrics, 2008), on the other hand, proposed a new restricted likelihood ratio (rLR) for spatial scan statistics, motivated by a different point of view, i.e., to avoid the property that ordinary spatial scan statistics tend to detect the cluster much larger than the true cluster by swallowing neighboring regions with non-elevated risk. Recently, we have confirmed that FleXScan implemented with the rLR

- (1) has surprisingly less computational time than the ordinary FleXScan, leading to no restrictions on cluster size and
- (2) tends to be able to detect *any clusters* reasonably well as the relative risk of the cluster gets large.

Other existing spatial scan statistics seem to have no such nice property. However, FleXScan with the rLR still suffers from the above-stated property due to likelihood ratio to a certain extent. FleXScan with the rLR is illustrated with three kinds of artificial cluster models created in the areas of Tokyo Metropolis and Kanagawa prefecture.

Center for Medical Statistics  
SAN Bldg. 4F 2-9-6 Higashi Shimbashi  
Minato-ku Tokyo 105-0021 Japan  
Tel: +81-3-6450-1680  
Email: [tango@medstat.jp](mailto:tango@medstat.jp)

# 心肺機能停止傷病者救急搬送件数の 時間集積性の検出

国立保健医療科学院  
政策技術評価研究部・健康危機管理研究部  
高橋邦彦

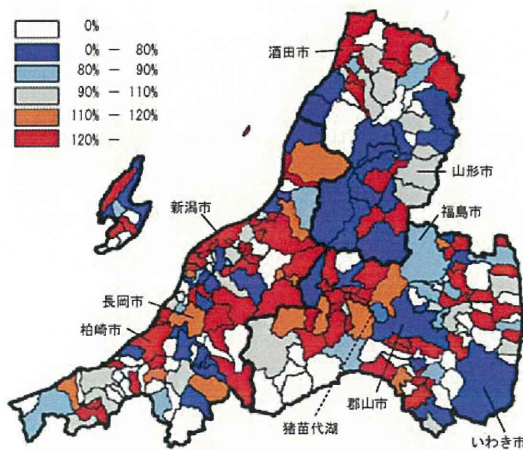
2011/12/11 シンポジウム(宮崎市)  
「生物統計における統計的推測とその応用」

## 疾病集積性の検定

～ 空間集積性と時間集積性 ～



## 疾病集積性の検定: FleXScan法

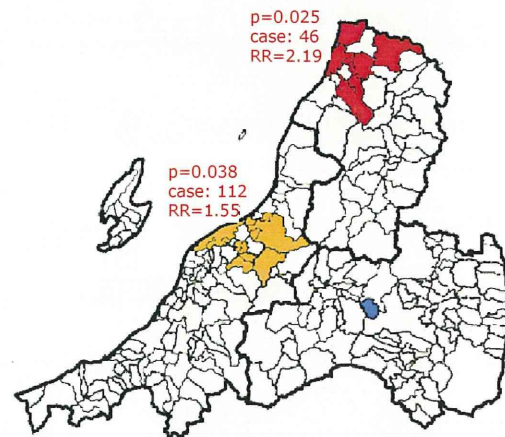


1996年～2000年男性胆のうがんの標準化死亡比(SMR)の疾病地図

発生がどこかに集中(集積)していないか?



Tango & Takahashi's Flexible scan statistic



疾病集積性の検定

～統計的に客観的に判定～

## 集積性の検定手法の実用化

- 現時点で、実務家が容易に利用できる地域集積性の検定アプリケーション:
  - **SaTScan**: Kulldorff
  - **FleXScan**: Takahashi *et al*の2つのみ(研究者レベルでの方法やプログラムは他にもあるが...)
- 実際、海外などでも、この2つの手法を用いた研究が増えている。

# FlexScan

FlexScan ver.1.1

Scanning method: Flexible spatial scan by data length.  
 Program run on: Fri Jan 14 20:58:31 2005

Purely Spatial analysis  
 scanning for clusters with high rates using the Poisson model.

SUMMARY OF DATA

Limit length of cluster: 15  
 Number of census areas: 92  
 Total cases .....: 16462

MOST LIKELY CLUSTER

1. Census areas included .: kawaguchi, souka, koshigaya.  
 Maximum distance.....: 24.2393 km (areas: kawaguchi  
 Number of cases .....: 3097 (2812.38 expected)  
 Overall relative risk ..: 1.1012  
 Log likelihood ratio ...: 16.3276  
 Monte Carlo rank .....: 1/1000  
 P-value .....: 0.001

No.	Area name	Connected	Connected	Connected	Connected	Connected	Connected	Connected	Connected	Connected
8	tokorozawa	kawagoe	sayama	iruma	niiza	miyoshi				
9	tokorozawa				moroyama	osose	naguri	tokigawa	yok	
10	tokorozawa				kisai	kitakawabe	ootone	kurihashi	wasi	
11	tokorozawa				kamisato	okabe				
12	tokorozawa				arashiyama	kawashima	yoshimi	hatoyama	oosc	
					ukabe	koshigaya	hasuda	shiraoka		
					aoka	sugito	matsubushi	syouwa		

FlexScan

File (F) Session (S) Tool (T) Help (H)

Files Analysis

Statistical model:  Poisson  Binomial

Statistic type:  Original LLR  LLR with Restriction Alpha= 0.2

Scanning method:  Flexible  Circular

The Maximum Spatial Cluster Size: 15 census areas (default=15, recommended)

Random number:  Multinomial  Poisson

Monte Carlo replications (default=999): 999 -- Seed of Random Number

Preview

Top: 50 Width: 170  
 Left: 20 Height: 240

Just Fit default Print

Print Setup Print

FlexScan ver.3.0 -- purely spatial

Scanning method: Flexible spatial scan.

STATISTICS

Log likelihood ratio with restriction.  
 $(Pr(X \geq x) + 0.5 * Pr(X = x)) < 0.200000$

Program run on: Mon Feb 02 17:09:00 2009

Purely Spatial analysis  
 scanning for clusters with high rates using the Poisson model.

SUMMARY OF DATA

Limit length of cluster: 15  
 Number of census areas: 118  
 Total cases .....: 235  
 (expected number is adjusted by Total cases as \*Total expected\* + \*Total cases\*)

MOST LIKELY CLUSTER

1. Census areas included .: Area-14, Area-15, Area-26, Area-27, Area-33  
 Maximum distance.....: 14.5143 km (areas: Area-14 to Area-33)

## Poisson model

- 対象空間  $G$  が  $m$  個の region (市区町村) に分割されているとする。帰無仮説(集積なし)のもとで

$N_i$  : region  $i$  の case 数 ( $n_i$  : 観測値)       $\mu_i$  : 期待観測数

$$N_i \sim \text{Poisson}(\mu_i) \quad (i = 1, 2, \dots, m)$$

とする。

- window  $Z$  に対し hot-spot モデルの検定仮説を考える。

$$N_i \sim \text{Poisson}(\theta_a \mu_i) \quad (i \in Z)$$

$$N_i \sim \text{Poisson}(\theta_b \mu_i) \quad (i \notin Z)$$

$$H_0 : \theta_a = \theta_b \quad \text{vs.} \quad H_1 : \theta_a > \theta_b.$$

## Scan statisticsによる集積性の検定

$Z \in \mathcal{Z}$  : window (clusterの候補)

$L(Z)$  : 対立仮説のもとでの  $Z$  が cluster であるという尤度

$L_0$  : 帰無仮説のもとでの尤度

としたときの最大尤度比

$$\begin{aligned} \lambda &= \sup_{Z \in \mathcal{Z}} \frac{L(Z)}{L_0} \\ &= \sup_{Z \in \mathcal{Z}} \left( \frac{n(Z)}{\mu(Z)} \right)^{n(Z)} \left( \frac{n(Z^c)}{\mu(Z^c)} \right)^{n(Z^c)} I \left( \frac{n(Z)}{\mu(Z)} > \frac{n(Z^c)}{\mu(Z^c)} \right) \end{aligned}$$

(Poisson modelの場合)

を検定統計量とする検定.

帰無仮説(「集積はない」)のもとでの $\lambda$ の分布をモンテカルロシミュレーションによって考え, 集積性の有無を検定.

この検定で有意となった場合、尤度比を最大にした **most likely cluster (MLC)**  $Z = Z^*$  を集積地域と同定する。

$$\lambda = \sup_{Z \in \mathcal{Z}} \left( \frac{n(Z)}{\mu(Z)} \right)^{n(Z)} \left( \frac{n(Z^c)}{\mu(Z^c)} \right)^{n(Z^c)} I \left( \frac{n(Z)}{\mu(Z)} > \frac{n(Z^c)}{\mu(Z^c)} \right)$$

➡  $Z$ の全体集合  $\mathcal{Z}$  の定め方が問題  
(Scan Statistic の違い)

- **Circular scan statistic**はある制限された範囲内の全ての **circular window**をscanして探す.
- **Flexible scan statistic**はある制限された範囲内の可能な全ての **(arbitrarily shaped) window**をscanして探す.

