

機関番号：82602  
 研究種目：若手研究（B）  
 研究期間：2010～2011  
 課題番号：22700301  
 研究課題名（和文） 地域差を考慮した空間疫学・地域診断のためのスキャン統計量の開発とその利用  
 研究課題名（英文） Development and use of scan statistics for community diagnosis with considering regional differences in spatial epidemiology  
 研究代表者  
 高橋 邦彦（TAKAHASHI KUNIHICO）  
 国立保健医療科学院・政策技術評価研究部・主任研究官  
 研究者番号：50323259

## 研究成果の概要（和文）：

空間疫学・地域診断のための統計手法において、各地域の人口規模の違いに起因する地域差を考慮した新たな疾病集積性の検定のためのスキャン統計量ならびにわが国における国民健康保険の地域比較のための指標のベイズ推定法を提案した。シミュレーションからいずれの方法も従来の方法よりも精度よい解析を行うことができることが確かめられた。

## 研究成果の概要（英文）：

We proposed several scan statistics for detecting disease clusters and Bayesian approach to indices of regional healthcare expenditure in Japan for spatial epidemiology and community diagnosis, with taking account of regional differences which are caused by population size of each area. Our simulation studies could show that the proposed methods can accurately detect the clusters and stably estimate the level of regional medical expenses in compared with conventional methods.

## 交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	1,600,000	0	1,600,000
2011年度	1,400,000	0	1,400,000
年度			
年度			
年度			
総計	3,000,000	0	3,000,000

研究分野：総合領域

科研費の分科・細目：情報学・統計科学

キーワード：空間・環境統計，空間疫学，地域診断，スキャン統計量

## 1. 研究開始当初の背景

2009年に世界的流行となった新型インフルエンザ A(H1N1)をはじめ、近年、保健医療分野におけるさまざまな健康危機事象の発生が問題となってきており、特にその突発的発生を早期発見することを目的としたサーベイランスが重要となっている。そのような問題に対して、疾病や症状発生の空間的な分布を観察し解析を行う空間疫学の統計解

析手法が有用であり、実際ニューヨーク市をはじめいくつかのサーベイランスシステムでは日単位の解析も組み込まれ運用されている。一方で疾病の発生だけでなく、医療費などの保健福祉における地域格差の問題が論じられており、それらの空間的な分布に注目した適切な地域診断の方法の必要性も高まっている。このような保健医療分野における空間データの解析・検討においては、日本

での市区町村や二次医療圏のように、一般に小地域ごとに集計された観測値に基づく議論が中心となる。その場合、各地域の人口サイズの違いなどによる地域差の影響が大きく、統計学の基本となる同一分布からの標本に基づく理論が適用できない。一般的にはそれぞれの地域で個々に推定を行う手法がとられるが、この地域差の影響は地域ごとに求められる推定値の推定誤差に大きな影響を与えてしまい、そのため単純な地域比較ができないことになる。たとえば疫学研究や実際のサーベイランス解析では Kulldorff(1997)の提案する疾病集積性の検定法(SaTScan法)が広く利用されている。最近、我々はSaTScan法の集積地の同定精度に関する問題を指摘し、それを改良した新たなスキャン統計量を提案しており(Tango and Takahashi, 2005)、最近ではFlexScan法として海外においても利用されるようになってきている。実際、ニューヨーク市の保健局(Department of Health and Mental Hygiene)が行うサーベイランス解析の担当者からも注目され、従来用いられているSaTScan法とともに我々の提案するFlexScan法での検出について、現実の2009年の新型インフルエンザ発生データの解析と検討も行われている(Goranson et al, 2009)。しかしこれらの方法はいずれも同じ最大尤度比に基づく統計量であり、統計量を最大化するという点においては、人口の少ない地域での影響が相対的に小さくなってしまいうため、本来検出すべきでない地域も多く含んで集積地として同定してしまうという問題点が指摘されている。その解決としていくつかの方法が提案されており、我々も各地域のばらつきを考慮した統計量などの検討を行っている(Takahashi and Tango, 2007)。一方、本研究で検討している集積性の検定は症候サーベイランス以外の研究にも適用でき、スキャン統計量を疾病発生と関連ある遺伝子の検出に適用した先行研究なども行われている。これらの場面においても上述のような同一分布を仮定できないデータによって生じる問題点があり、これらの影響の詳細な検討や、それに対処する方法の開発は重要となる。

## 2. 研究の目的

疾病集積性の検定を目的として、従来用いられている手法について人口規模を中心とした地域差の影響を詳細に検討することにより、より正確な同定をできるようなスキャン統計量の開発を行い、従来の方法との比較を行いシミュレーション等によってその精度比較を行う。さらにその方法の実際への適用、ならびに他の問題への適用について検討を行う。また地域診断を目的として、それぞ

れの地域の人口規模の影響を考慮するリスク推定方法の開発とその実例について検討を行う。

## 3. 研究の方法

疾病集積性の検定を目的としたスキャン統計量の開発として、以下の視点から改良を試みた。(1) 症候サーベイランスを目的とし、事象の発生時点からの立ち上がりを数理的にモデル化する。また時間的な変動について従来のポアソンモデルではなく負の二項分布で記述することでより安定化させる。

(2) 一般的なポアソンモデルに基づき、分散安定化変換を用いたスキャン統計量を用いることで、いったん各地域の変動をある程度同じにした上で、その違いを検討する。(3) 人口規模の大きさの影響を検定のp値ではかることを目的に、各地域の人口規模に基づくp値がある基準よりも小さくなる地域だけに制限したクラスで最も尤度比が大きくなる集積地域を見つけ出す。

これらの方法について理論的な導出とともに、シミュレーション等によって従来の方法と検出精度の比較を行った。さらにスキャン統計量を用いた集積性検定の応用について、実際に難病の集積性の検出を行った。さらに遺伝子のゲノムワイドCNVデータへスキャン統計量の応用についても検討を行った。

一方、地域のリスク推定として、わが国における医療費地域差指数について検討を行い、各地域の人口規模を考慮すべくベイズ推定を行い、特に隣接地域の影響を考慮した空間相関を取り込めるような階層モデルをいくつか提案し、従来の方法との比較を行いながら、それぞれの方法の特性を比較した。

## 4. 研究成果

(1) 症候サーベイランスを目的とし、事象の発生時点からの立ち上がりを数理的にモデル化した方法として、新たなspace-time scan統計量を提案した。北九州市の学校欠席数のデータ解析やそれに基づくシミュレーションによって、従来のcylindrical scan統計量ならびにflexible space-time scan統計量よりも、数時点早くアウトブレイクを検出することが確かめられた。

(2) 一般的なポアソンモデルに基づき、分散安定化変換を用いたスキャン統計量を提案した。この方法では従来のモデルで想定されているホットスポットモデルを仮定せず、最も有意なクラスターを検出することを目的としている。シミュレーションによって提案法はよりクラスターの中心部分をfalse positiveが少なくなるよう同定する様子が確認できた。

(3) p値を用いて制限された尤度比統計量に

基づくスキャン統計量では、従来の方法では決して同定できないようなクラスターも高精度で同定することができるようになった。さらに計算時間も大幅に改善することができた。

(4) 実際の適用として難病の地域差について、医療受給者証所持者数を保健所管轄で集計したデータについて集積性の検定をおこなった。多発性硬化症、クローン病、サルコイドーシスなどについて男女別に検討を行った結果、それぞれ地域差があることが確認でき、さらに疾患ごとに地域分布が異なる様子が観察された。今回用いた受給者証所持者数は厳密な意味での患者数ではないが、これらのデータを用いた地域比較の結果を活用することは、今後の難病研究推進に有用であると考えられた。

(5) 遺伝子のゲノムワイド CNV データへのスキャン統計量の応用について、CNV データの解析に合わせたスキャン統計量を提案した。シミュレーション結果から、従来の方法では検出することができない遺伝子座についても、提案法によって検出できることが確かめられた。この方法を更に拡張し、より精度よく検出する方法の検討が必要であると考えられた。

(6) 隣接地域の影響を考慮した医療費地域差指数のベイズ推定としては、ベイズ推定をすることで各市町村の人口規模に起因するばらつきを同程度にできることが確かめられた。特に空間相関の程度を変動させる mixture model では、データへの適合度も高くなり、実際、ばらつきも人口規模によらずほぼ一定になっていた。

本研究の結果から、地域差を考慮した空間疫学・地域診断のための方法として、従来法を改善する方法を提案できた。しかし、シミュレーション結果からも、いずれの方法にも更なる改善が求められる課題が残されていることが明らかになった。これらの点についてさらなる改善の検討が必要であると考えられる。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 4 件)

・Tango T, Takahashi K (2012). A flexible spatial scan statistic with a restricted likelihood ratio for detecting disease clusters. *Statistics in Medicine* (印刷中).

・Nishiyama T, Takahashi K, Tango T, Pinto D, Scherer SW, Takami S, Kishino H (2011).

A scan statistic to extract causal gene clusters from case-control genome-wide rare CNV data. *BMC Bioinformatics* 12:205.

・Tango T, Takahashi K, Kohriyama K (2011). A space-time scan statistic for detecting emerging outbreaks. *Biometrics* 67:106-115.

・Doi Y, Yokoyama T, Tango T, Takahashi K, Fujimoto K, Nakano I (2010). Temporal trends and geographic clusters of mortality from amyotrophic lateral sclerosis in Japan, 1995-2004. *Journal of the Neurological Sciences* 298:78-84.

[学会発表] (計 6 件)

・高橋邦彦, 横山徹爾, 金谷泰宏, 土井由利子 特定疾患 (難病) 医療受給者証所持者数の地域比較. 第 22 回日本疫学会学術総会, 東京都, 2012 年 1 月 28 日.

・高橋邦彦, 土井由利子, 横山徹爾, 金谷泰宏, FleXScan による疾病集積性の検出～特定疾患 (難病) の地域比較への適用～. 2011 年度統計関連学会連合大会企画セッション 「医学データの分類と判別」. 九州大学, 2011 年 9 月 6 日.

・Tango T, 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.

・西山毅, 高橋邦彦, 丹後俊郎, 高見聖, 岸野洋久. スキャン統計量を用いた遺伝子パスイェイに基づくゲノムワイド Copy Number Variation (CNV) データ解析法の開発. 第 21 回日本疫学会学術総会, 北海道札幌市, 2011 年 1 月 21 日-22 日.

・Takahashi K, Tango T. A maximum scan score-type statistic based on Anscombe's variance stabilization transformation for disease clustering. XXVth International Biometric Conference, December 5-10, 2010, Florianopolis, Brazil.

・高橋邦彦, 飛田英祐, 山岡和枝, 丹後俊郎. 隣接地域の影響を考慮した医療費地域差指数のベイズ推定. 第 69 回日本公衆衛生学会総会, 東京国際フォーラム, 2010 年 10 月 27 日-29 日.

[図書] (計 3 件)

・高橋邦彦. 空間スキャン統計量. In 丹後俊郎, 小西貞則 (編). 医学統計学の事典.

108-109. 朝倉書店, 2010.

・高橋邦彦. 症候サーベイランス. In 丹後俊郎, 小西貞則 (編). 医学統計学の事典.

109-110. 朝倉書店, 2010.

・高橋邦彦. 疾病地図. In 松原望, 美添泰人, 岩崎学, 林文, 金明哲, 山岡和枝, 竹村和久 (編) 統計応用の百科事典. 506-507. 丸善, 2011.

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

なし

## 6. 研究組織

### (1) 研究代表者

高橋 邦彦 (TAKAHASHI KUNIHICO)

国立保健医療科学院・政策技術評価研究部・主任研究官

研究者番号 : 50323259

### (2) 研究分担者

なし

### (3) 研究連携者

なし