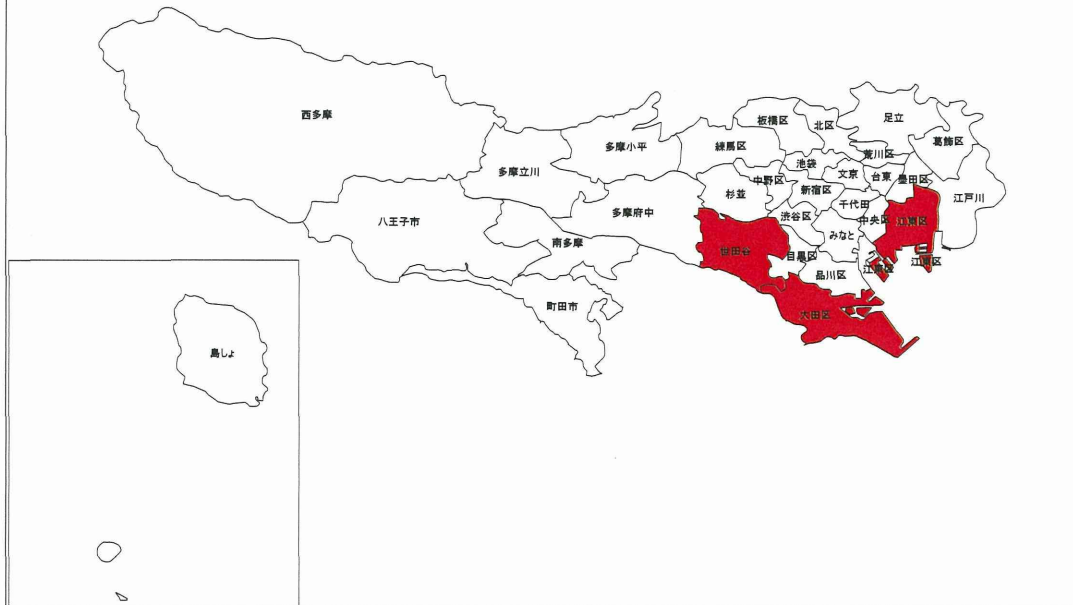


2011_w52_3

□ <その他の値すべて>

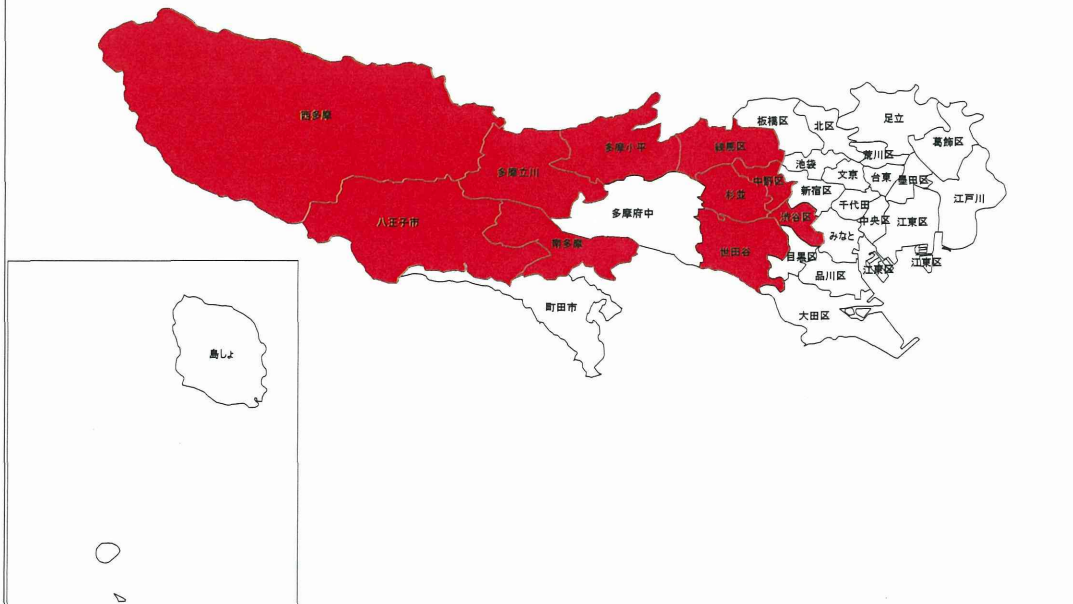
w52

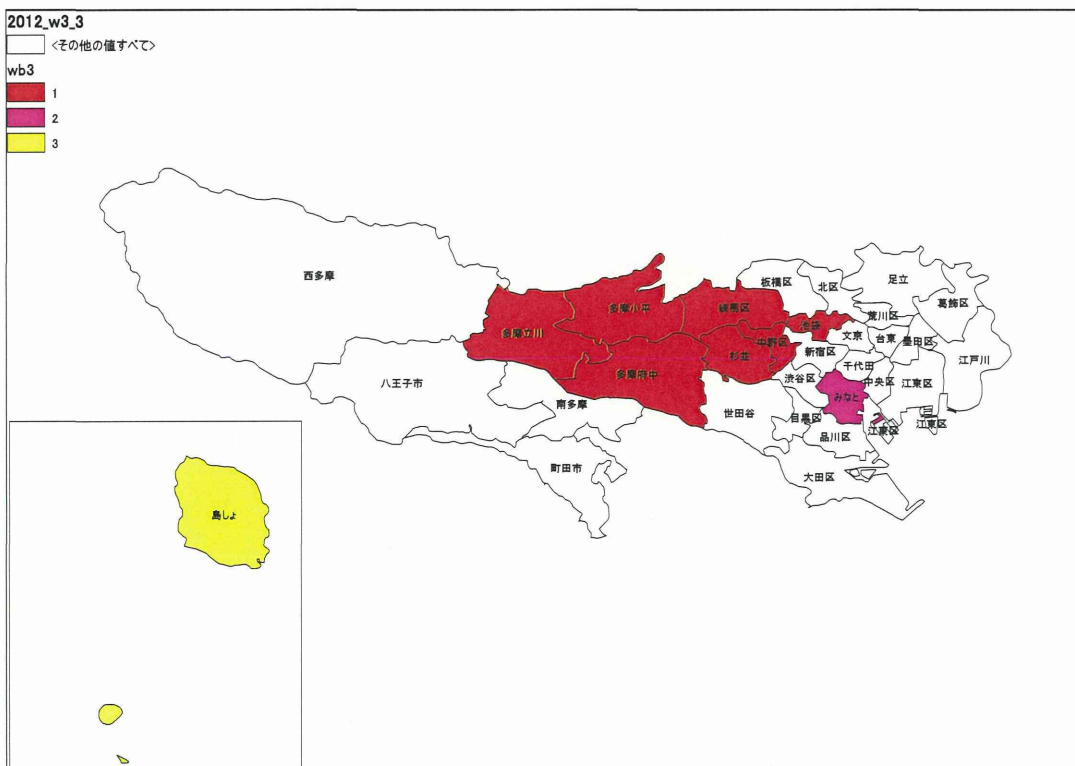
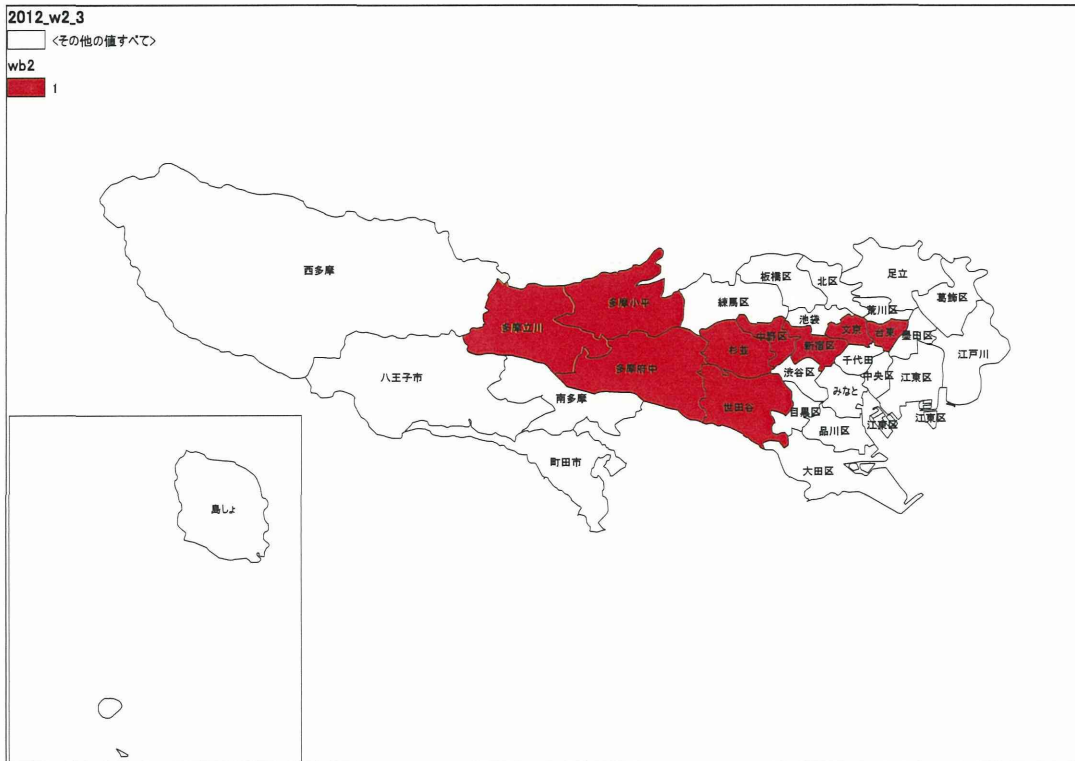


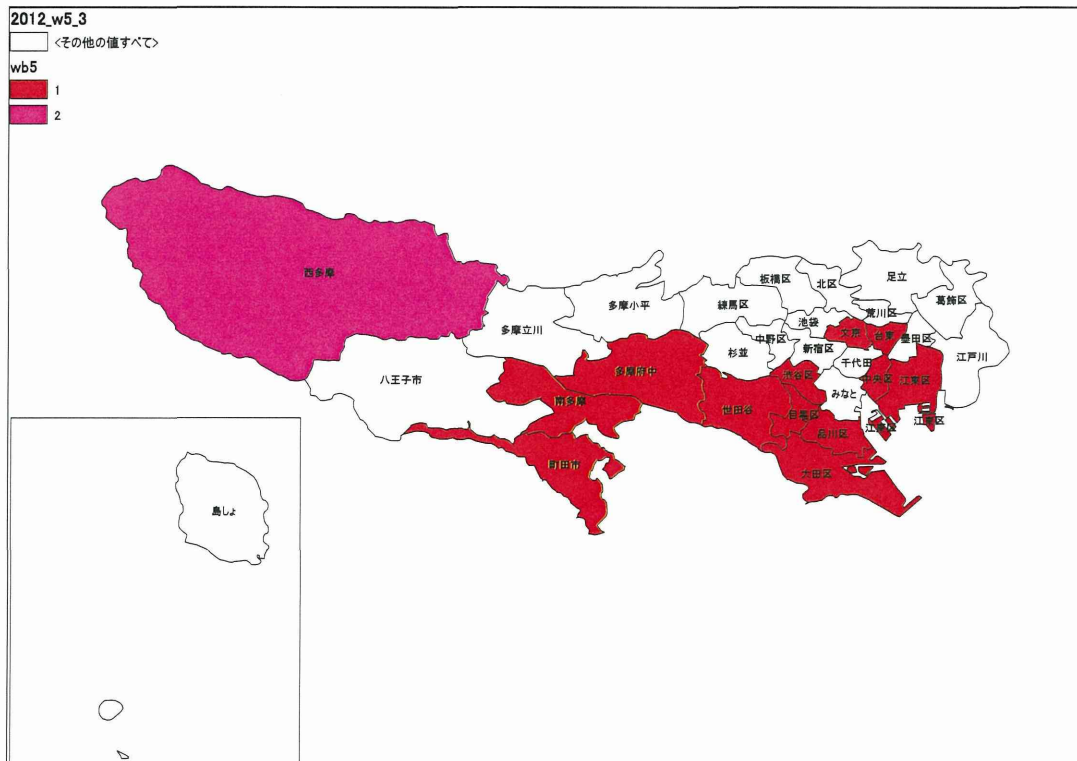
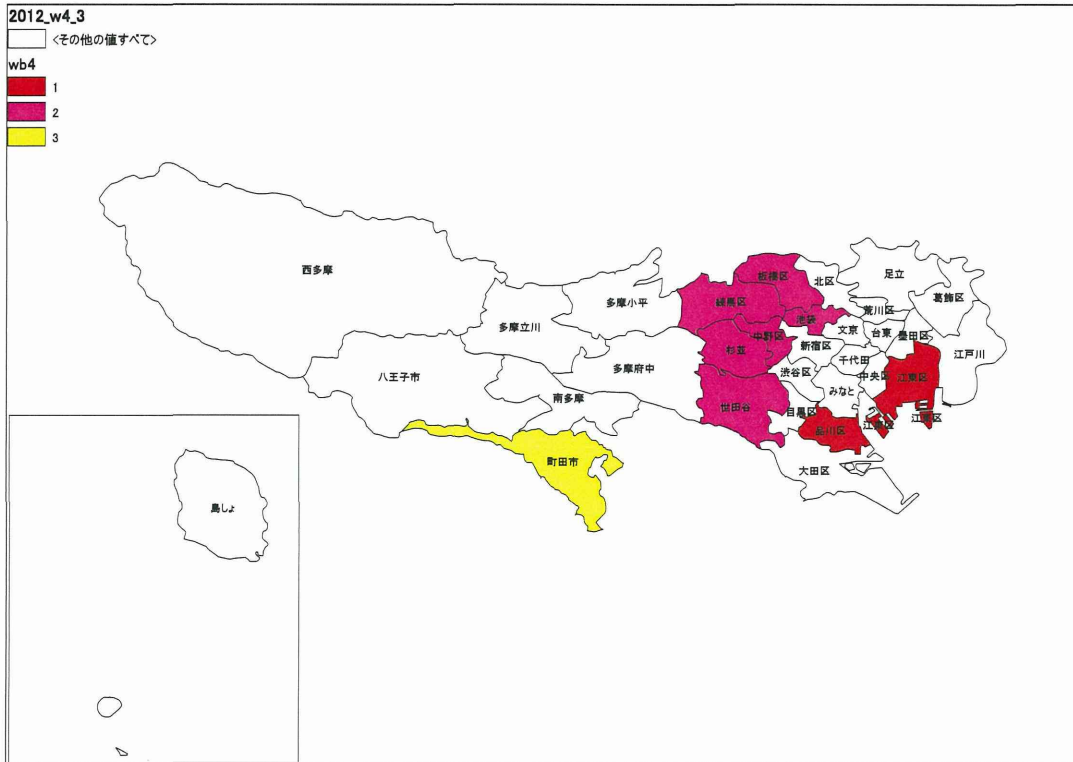
2012_w1_3

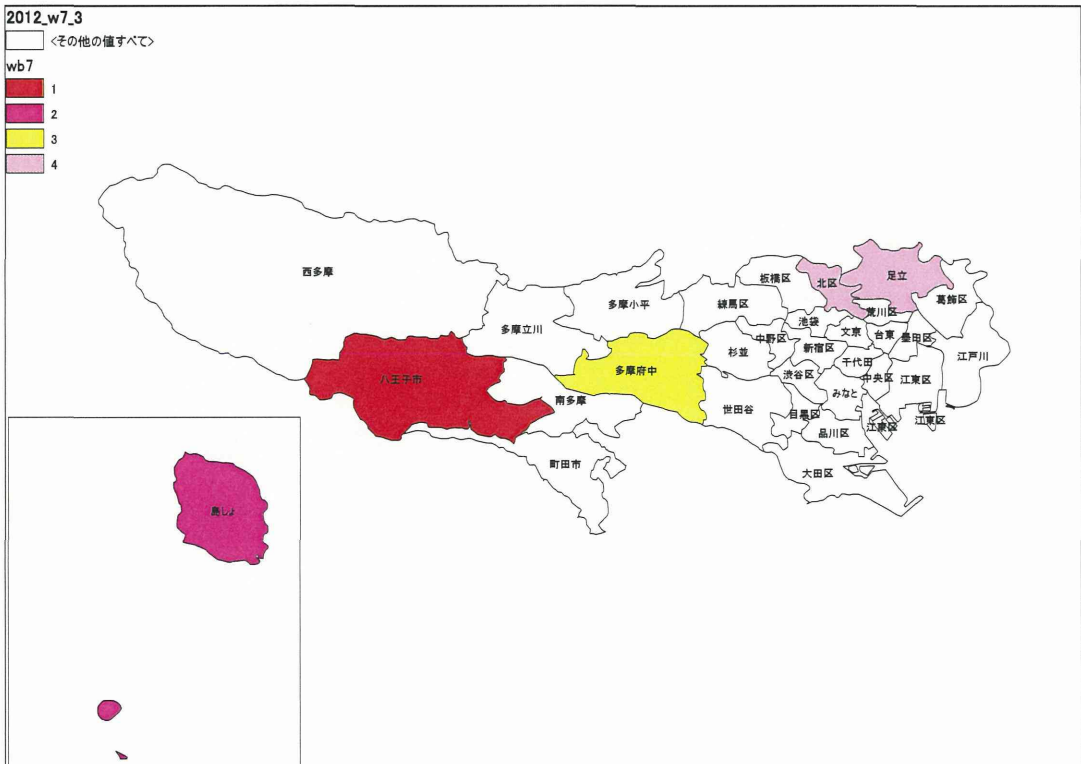
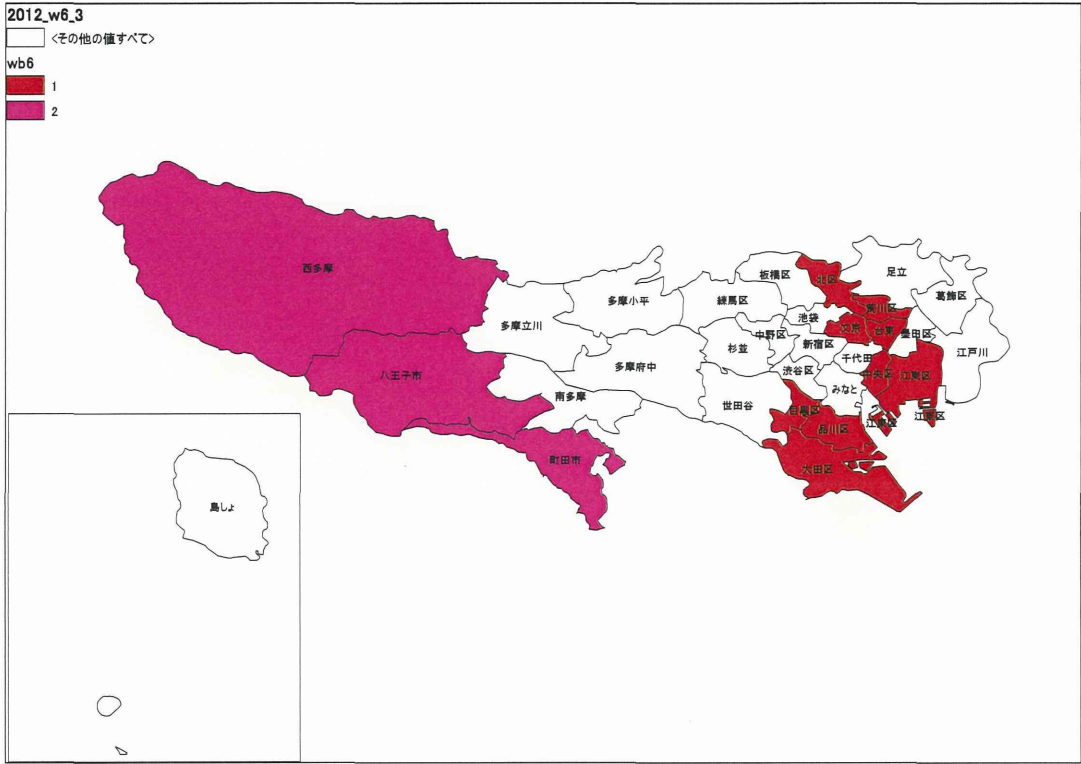
□ <その他の値すべて>

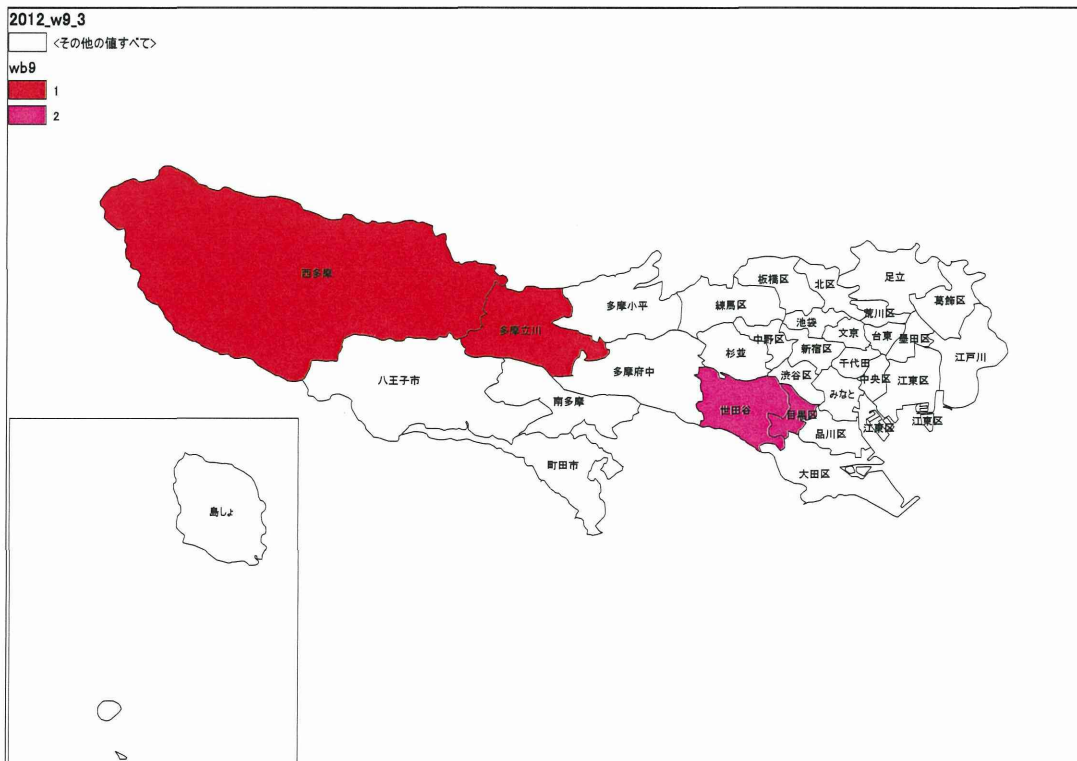
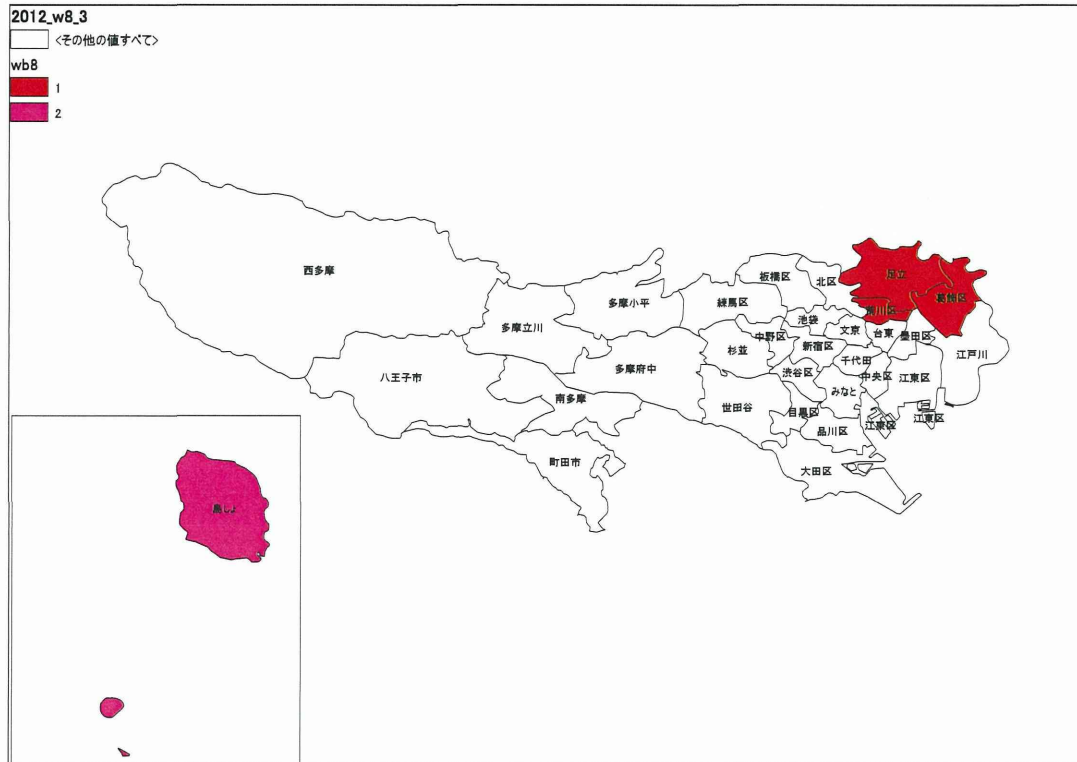
wb1









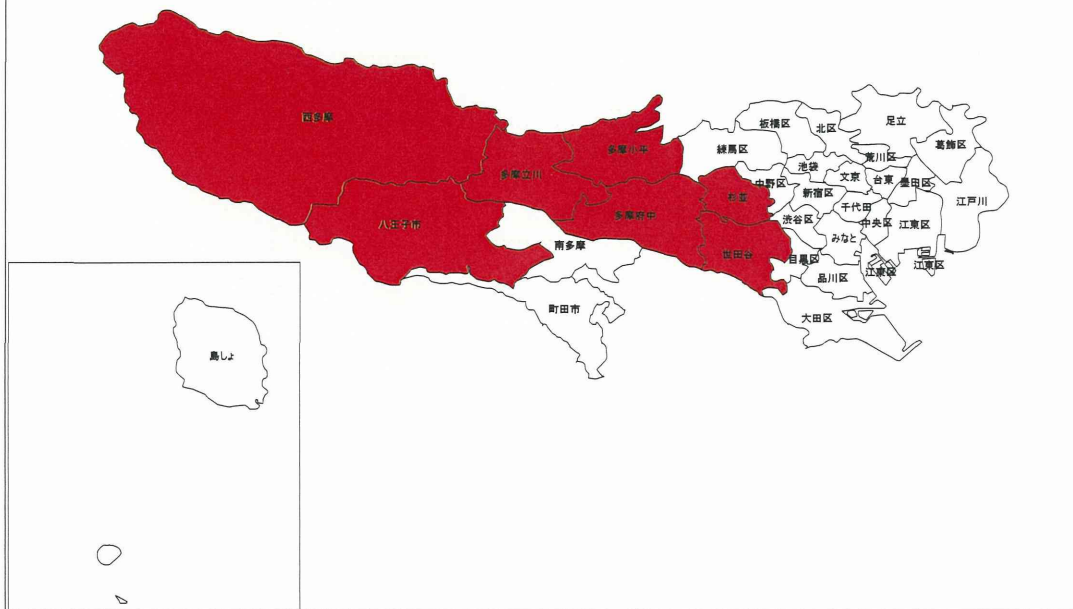


2012_w10_3

□ <その他の値すべて>

wb10

■ 1



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）
分担研究報告書

東京都におけるインフルエンザ学級閉鎖状況の地理情報システムを用いた
可視化に関する研究（2011-2012年シーズン）

研究協力者 東京都健康安全研究センター 高橋琢理
研究協力者 東京都健康安全研究センター 杉下由行
研究分担者 東京都健康安全研究センター 灘岡陽子

研究要旨：東京都におけるインフルエンザによる学級閉鎖実施状況の地理的な分布とその規模を地理情報システム（GIS）によって地図上に描画することで可視化し、動画・静止画によってその流行状況を把握することを試みた。また、従来から実施されているインフルエンザ定点医療機関からの報告と比較を行った。その結果、学級閉鎖の時系列変化を示した動画の分析から、学級閉鎖の実施状況が定点報告と関連することが示唆された。学級閉鎖状況からインフルエンザの流行状況を推測する方法を開発し、公衆衛生行政の疫学情報として還元することが今後の課題である。

A. 研究目的

地理情報システム（GIS）による情報活用と公開が様々な分野で進められている。自治体においても、行政データの分析や住民への情報提供などへの地理情報の活用が求められている¹⁾。公衆衛生行政では、地理情報に基づいた疫学情報の分析とその提供が課題であり、例えば、季節性インフルエンザについては、発生状況の把握と分析、流行拡大への対策、住民への情報提供などに地理情報を活用することが求められる。そこで、本研究ではインフルエンザによる学級閉鎖実施状況の地理的な分布とその規模をGISによって地図上に描画することで可視化し、動画・静止画によってその流行状況を把握することを試みた。

B. 研究方法

東京都内の高校、中学校、小学校、幼稚園、保育園、各種学校（以下、学校）で2011-2012年シーズン（2011年9月1日～2012年8月31日）に

報告された学級閉鎖・学年閉鎖の状況を対象とした。

学級閉鎖・学年閉鎖の実施状況に関しては学校から管轄保健所に対し以下の内容が報告された。市区町村、学校種別、学校名、臨時休業の種別（学級閉鎖、あるいは学年閉鎖）、閉鎖する学年・年齢、閉鎖学年数、閉鎖学級数、在籍者数、インフルエンザ様疾患による欠席者数、インフルエンザ様疾患罹患登校者数、閉鎖期間開始日、閉鎖期間終了日。

学級閉鎖・学年閉鎖の実施された学校の緯度経度情報を、市区町村、学校種別および学校名をもとに、国土交通省国土政策局の国土数値情報公共施設データ（東京都、平成18年度）から得た²⁾。新設や統廃合、あるいは名称変更などがなされた学校では、上記データで緯度経度が得られないため、インターネットから検索した住所をもとにGeocodingサービス³⁾から緯度経度情報を得た。これらの情報によりGISソフト（

ArcGIS 10 ESRI社製) によって学校の位置を地図上に表示した。また、在籍者の数を色で分類し、学級閉鎖・学年閉鎖の期間と規模を表示する動画として可視化した。また、従来から実施されているインフルエンザ定点医療機関からの報告と比較を行った。

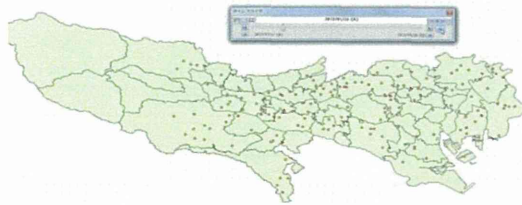


図1 2011-2012年シーズンの都内学級閉鎖状況例

C. 研究結果

2011-2012年シーズンに東京都内で報告されたインフルエンザによる学級閉鎖・学年閉鎖の分布状況を静止画として可視化した。また、その実施状況の時系列変化を動画として示した。これにより、学校における学級閉鎖・学年閉鎖の状況が地理情報として把握可能となった。図1に学級閉鎖状況の例を示す。点は学級閉鎖等の実施された学校の所在地であり、2012年第3週に都内全域で学級閉鎖が実施されたことが示された。

また、時系列変化の動画から、2011-2012年シーズンの学級閉鎖は、2012年1月の第3週に都内全域で一斉に増加し、3月20日ごろに終息したことが明らかとなった。これはインフルエンザ定点医療機関からの報告の増加および終息の傾向と同じであった。資料に時系列動画表示を分解して示した。

D. 考察

学級閉鎖の時系列変化を示した動画の分析からは、学級閉鎖の実施状況が定点報告と関連することが示唆された。公立の小学校、中学校、幼稚園は各地域で均等に配置されるように学区が

設けられている。これらの学校に通う子供たちは、学区住民の子供たちである。そのため、学区内のインフルエンザ流行状況が子供たちの学級閉鎖にあらわれていると仮定すると、地域におけるインフルエンザの流行状況がある程度反映していると考えられる。また、成人における職場を中心とした流行は、定点医療機関からの報告や、薬局サーベイランスなどが実施されている。しかしながら、基礎的なデータは不足しており、地理情報についても十分なデータが得られていない。一方、学校は、子供の集団生活の場であり、学校における学級閉鎖の状況を分析することで、職場を介した成人におけるインフルエンザ流行状況を推測する方法も得られる可能性が考えられる。

E. 結論

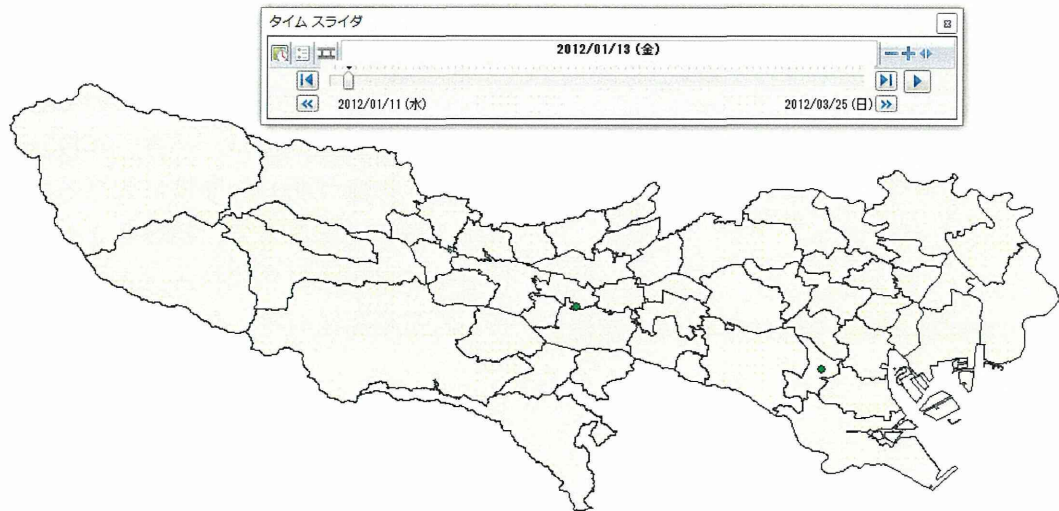
インフルエンザによる学級閉鎖・学年閉鎖の状況を、GISによって地理的な分布と規模、さらに時間変化を静止画と動画により可視化した。これらの結果から、学校における学級閉鎖状況が、インフルエンザ定点報告の傾向を反映する可能性が示された。本研究では、学級閉鎖と定点報告との関連性の数理的な検討が不十分である。また、地域の流行状況の反映程度についての分析も限定されており、これらは制限として挙げられる。今後、数理的な検討によって、学級閉鎖と定点報告との関連性を明らかにするとともに、学級閉鎖状況から流行状況を推測する方法を開発し、疫学情報として還元することが、今後の課題である。

参考

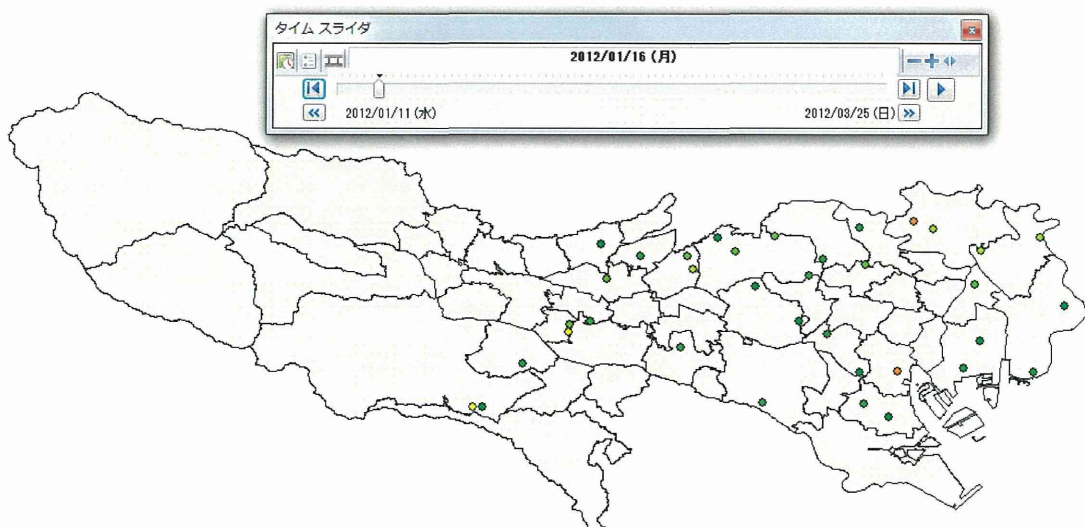
- 1) 地理空間情報活用推進基本法
(平成十九年五月三十日法律第六十三号)
- 2) 国土交通省国土政策局 国土数値情報ダウンロードサービスhttp://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-P02-v4_0.html
(2013年4月10日閲覧確認)
- 3) Geocoding <http://www.geocoding.jp/>
(2013年4月10日閲覧確認)

資料

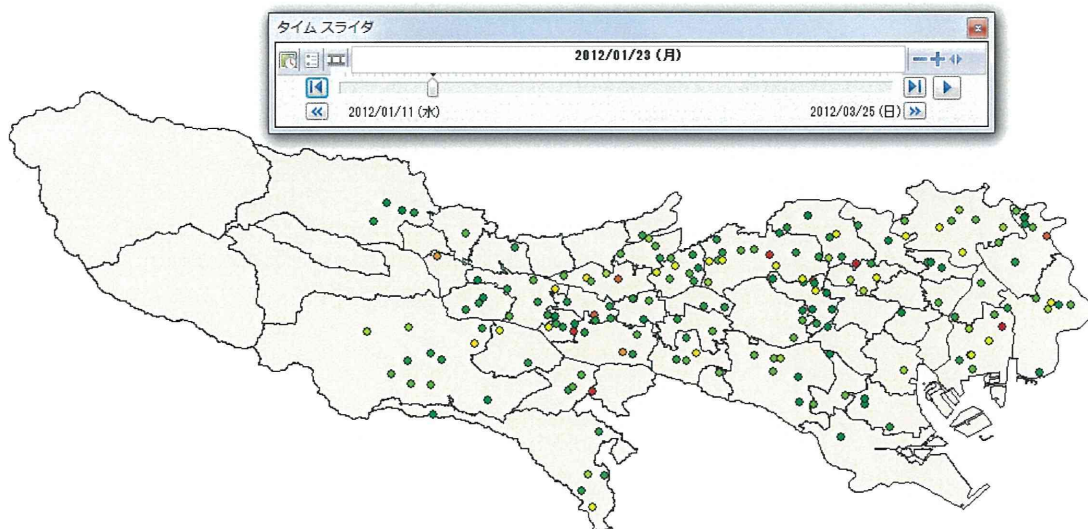
動画によるインフルエンザ流行状況 1 2012年1月13日



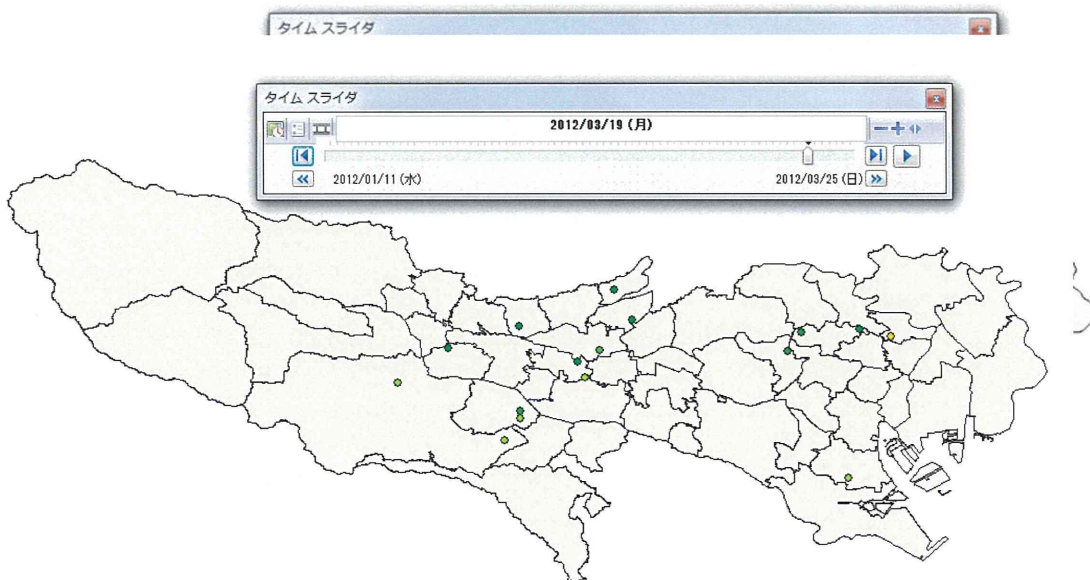
動画によるインフルエンザ流行状況 2 2012年1月16日



動画によるインフルエンザ流行状況 3 2012年1月23日



動画によるインフルエンザ流行状況 4 2012年3月19日



F. 研究発表

1. 論文発表

なし

2. 学会発表

なし

3. 著書

なし

G. 知的所有権の取得状況

なし

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）
分担研究報告書

わが国における救急蘇生統計を用いた冬季の心肺機能停止傷病者救急搬送件数の
時間的集積性の検出

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
研究協力者	武村真治	国立保健医療科学院健康危機管理研究部
	長谷川学	総務省消防庁救急企画室
	金谷泰宏	国立保健医療科学院健康危機管理研究部
	齋藤大蔵	防衛医科大学校防衛医学研究センター
	阪本敏久	防衛医科大学校病院救急部

研究要旨：一般に心肺機能停止傷病者の発生は冬季に多いことが知られており、消防庁救急統計活用検討会でのデータの集計結果でも明らかになっている。本研究では心肺機能停止傷病者の救急搬送件数が増える冬季に注目し、その中で特異的に件数が高く集積している期間があるかどうかを客観的に判断するために解析、検討を行った。2005～2008年の日本全国において救急搬送された心肺機能停止傷病者の救急蘇生統計データを用い、性別、心原性・非心原性の別などについて集計を行い、12～1月の期間に注目し、発生の有意な集積期間があるか集積性の検定を行った。いずれの年も、男女ともに年末年始時期に最も有意な時間集積が認められた。また、心原性・非心原性とも同様の集積があり、原因による差は認められなかった。わが国における心肺機能停止傷病者の救急搬送件数が年末年始時期に統計的に有意に集積していることが明らかになった。

A. 研究目的

近年、少子高齢化、核家族化の進展や住民意識の変化等に伴い、救急需要が増加し、平成20年の救急出場件数は約510万件で、平成10年からの10年間で約38%増加していた。その中で毎年約10万件が心肺機能停止傷病者の搬送となっている。心肺機能停止症例について地域間・国際間での蘇生率等の統計比較を可能とするために、1990年の国際蘇生会議において「ウツタイン様式」とよばれる記録に関する統一的なガイドラインが提唱された。わが国においては、平成15年「救急業務高度化推進検討会」によるウツタイン様式導入の提言をもとに、平成17年1月

より全国の消防本部における全ての心肺機能停止傷病者のウツタイン様式に基づくデータ収集が開始され、救急蘇生統計として公表されている。ウツタイン様式の調査項目では、病院搬送前の心肺機能停止傷病者に対する救命活動による蘇生、1ヶ月後の生存率や回復程度など、救命処置の成績に関する項目に重点が置かれている。その一方で、心肺機能停止傷病者が発生するリスク要因に関しては十分に検討されていないのが実情である。

心肺機能停止のリスク要因として、患者個人の要因とそれを取り巻く環境の要因に大別できるが、予防方策を検討する上では後者の分析が

不可欠である。このうち最も重要なリスク要因の一つとして「季節性」が挙げられる。一般に心肺機能停止は冬季に多いことが知られており平成21年度第2回救急統計活用検討会の集計や新潟市における検討などでも明らかになっている。しかし具体的にどの時期にどの程度発生しているかなど、その発生時期についての詳細な分析や厳密な統計学的検討は十分に行われていない。一般に疾病などの発生を経時的に観測している場合、その発生が特定の期間に集積しているとき「時間集積性」があるという。時間集積性が認められた場合、その期間の発生増加をもたらす何らかの要因が背後にあることが示唆される。この集積性の検出のためには疾病集積性の検定手法が利用できる。そこで本研究ではウツタイン様式に基づき集計され消防庁救急企画室から公開されている救急蘇生統計データを用い、特に心肺機能停止傷病者発生が多い冬季12～1月の2ヶ月間に注目し、その期間内において発生が集中している時期があるかどうか時間集積性の検出を検討する。

B. 研究方法

2005～2008年の日本全国において救急搬送された心肺機能停止傷病者を対象として、匿名化・電子化されたうえで消防庁救急企画室から提供されているウツタイン様式に基づく救急蘇生統計データを用いた。本研究では特にその発生の特徴をつかむことを目的とし、発生日、都道府県、性別、年齢、心原性・非心原性の別、および覚知の時刻の項目を用いた。データは年毎のファイルに分けられ提供されているが、入力されている発生日の年が収録されている年と異なっていたものについては、収録されているファイルの年に合わせた。また発生日が未入力のものについては解析から除外した。以上により対象解析例数は2005年が102,738件（うち発生日不明2件）、2006年が105,943件（うち発生日不明4件）、2007年が109,461件（うち発生日不明0件）、2008年が113,827（うち発生日不明6件）件であった。

12～1月における搬送件数の時間集積性の検

出を行うため、各年12～1月の2ヶ月について日単位の発生件数を用いて、特定の期間に統計的に有意な時間集積性が確かめられるか、男女別の総数、ならびに心原性・非心原性別に集積性の検定を行った。検定には疾病集積性の検定法であるKulldorff's scan statisticを用いた。ここではこの2ヶ月内に集積期間がないことを帰無仮説とし、1日あたり期待件数は、当該2ヶ月間の搬送件数の総数を日数（62日）で割った数を用いた。また検定に必要なパラメータとして、検出される集積期間の最大期間(maximum temporal length)を15日間とし、モンテカルロ検定のための繰り返し数を9,999回、有意水準5%として検定を行った。最もp値が小さい期間をmost likely clusterとし、それ以外にも $p < 0.05$ となる有意な集積期間があったものは、secondary clusterとしてp値の小さい順に挙げた。この解析には集積性の検定を行うソフトとしてMS-Windows用に開発提供されているFlexScan Version 3.19)を用いた。

C. 結果

毎年12～1月の2ヶ月間における男女別、総数と心原性・非心原性別での時間集積性の検定を行った結果は表1のようになった。まず総数での結果をみると、2005年12月～2006年1月においては、男性で2005年12月30日～2006年1月10日の12日間が有意な集積期間($p=0.0001$)として検出された。12～1月の2ヶ月間を基準とした集積期間内の発生の相対リスク(RR)は1.23であった。同様に女性については、2005年12月30日～2006年1月11日の13日間が有意な集積期間($p=0.0001$, $RR=1.21$)として検出された。2006年12月～2007年1月、2007年12月～2008年1月においても、男女とも有意な集積期間が検出された。なお2006年12月～2007年1月の男性では最も有意な集積が①12月29日～1月9日の12日間($p=0.0001$)であり、次いで②1月15日の1日間が2番目の集積期間($p=0.0298$)であったことを示している。いずれに年においても最も有意な集積が検出されたのは、男女とも1月1日を含む年末年始時期であり、相対リスクも1.2倍を

超えていた。

心原性・非心原性別の検定においても、いずれの場合でも有意な時間的集積が検出された。男女問わず、心原性、非心原性とも、それぞれ最も有意な集積期間として検出されたのは1月1日を含む年末年始時期であった。特に非心原性で検出された集積期間では、相対リスクが1.63（2005年12月～2006年1月，男性），1.72（2006年12月～2007年1月，男性），1.51（同，女性）など、心原性のみの場合の年末年始時期の集積期間の相対リスク（1.22～1.40）よりもさらに高いリスクをもつ期間が検出されていた。

D. 考察

近年、救急車による搬送件数の増加が指摘され、その要因についていくつかの検討が行われており、救急搬送率には医学的要因以外のものも関与していることが示唆されている。また救急医療システムと地域の健康指標の関連も認められている。本研究では心肺機能停止傷病者の救急搬送が網羅されているウツタイン様式に基づいた救急蘇生統計のデータを用いて、その搬送件数の発生時期について詳しく検討を行った。わが国においては心肺機能停止傷病者の発生は冬季に高くなる傾向にあることが知られており、本研究での集計からも夏季よりも冬季に件数が多い傾向が確認された。特に発生件数の多い冬季12～1月に注目し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。この増加が誤差範囲と考えられるかどうか、時間集積性の検定による解析を行ったところ、いずれの年でも1月1日を含む年末年始時期に最も有意な集積が検出された。この結果は人口の多い都道府県だけの影響ではなく、全国的に心肺機能停止傷病者の救急搬送件数がこの時期に増加していると認められた。また男女別、心原性・非心原性の別のいずれの場合でも1月1日を含む年末年始時期に最も有意な集積期間が検出された。このことから年末年始の件数増加に関しては、例えば正月に多く発生する餅による窒息などの外的要因での事故が顕著に増えているのではなく、性別や心原性・非心原

性の別なく全体に発生の集積が起きていると考えられた。さらに検出された集積期間内とそれ以外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ことから、特定の年代だけが年末年始に増加しているのではなく、どの年代でも増加していると考えられる。なお時間集積性の検定では、発生が集中している期間を検出するため、最も有意な集積期間の相対リスクが必ずしも最大になるわけではない。実際、2006年12月～2007年1月の男性総数において最も有意な集積期間は2006年12月29日からの12日間で、相対リスクは1.20 ($p=0.0001$)であったが、2番目の集積として2007年1月15日の1日だけが検出され、その相対リスクは1.26 ($p=0.0298$)であり、最も有意な集積期間のリスクよりも高かった。また時間集積性の検定法はいくつか提案されているが、本研究のように事前に集積のある期間を想定せず、検定と同時に集積期間の同定も行う解析では、特に検定の多重性を考慮しているという点でKulldorff's scan statisticの利用が推奨されている⁸⁾。実際の解析には本研究で用いたFlexScanの他にSaTScanなどを利用することも可能である。

一方、2005年12月～2006年1月における最も件数の多い東京都での元日の発生総数は71件であった。これは日単位でみると、年末年始時期以外における多い日より東京都全体で20件程度の増加であり、それを消防署単位や病院単位として単純に考えれば1～2件増えた程度になる。さらに人口の少ない県などでは県全体でも数件増える程度になっている。また救命救急センターなどの救急外来では、心肺機能停止傷病者に限らず総じて受診患者が年末年始時期に増加している。このことから各地域や医療機関では年末年始時期に急激に心肺機能停止傷病者の搬送件数が増えているとは認識されていないのではないかと思われる。しかし本研究によって、この年末年始時期における心肺機能停止傷病者搬送件数の特異的に集中した増加は全国的に起きていることが明らかになった。実際、12～1月の2ヶ月間において、年末年始時期の集積期間では1.2倍を超えるリスクで発生が増加してい

た。このことは、心肺機能停止や救急搬送の問題を、個々の患者ではなく、地域の課題として取り組む必要があることを示唆している。今回用いた時間集積性の検定による検討では、集積の原因を明らかにすることは困難であるが、いくつかの仮説と考察が挙げられる。一つには、心肺機能停止傷病者の発生要因として心筋梗塞のリスクファクターである低温環境が考えられる。しかし年間を通して最も寒冷になるのは、年末年始時期ではなくむしろ1月下旬から2月である。本研究の結果からは、年末年始時期の搬送件数は男女、心原性・非心原性とも1月下旬～2月の件数よりも多いことが観察された。冬季にその発生が増加していることから、低温環境による人体に対するストレスが大きいこともひとつの要因になる可能性はあるが、それのみが年末年始に心肺機能停止傷病者が集積する要因とは考えられない。もう一つは、長期休日による医療アクセシビリティの低下が考えられ、それが高血圧、糖尿病等の基礎疾患を有する者の健康リスクをさらに高めている可能性は否定できない。急性心筋梗塞による死亡と地域の救急医療資源の関連も示唆されており、この時期多くの医療機関が休診になることや、それが患者の医療受診行動に少なからず影響を与えている可能性も考えられる。実際、週単位の集計で若干一時的な増加がみられた2007年の33週はお盆の時期になっていた。この年は記録的な猛暑に見舞われた年でもあり、この増加も医療機関の休診が重なったことが影響した可能性も考えられる。また、年末年始においては大都市圏への野宿者の流入が生じることが指摘されており、越年者については特に健康状態が相対的に重篤化していることが指摘されており、年末年始における社会・文化的な特徴が影響していることも考えられる。

しかし本研究で用いたウツタイン様式に基づく救急蘇生統計で集計されるのは、病院外心停止患者の搬送・治療成績に焦点が当てられた記録集計であり、その発生の原因についての項目は限られている。そのため心原性の傷病者も、非心原性の傷病者も、ともに年末年始に搬送件

数が集中して増加していることや、その期間内外で年齢構成や覚知時間に大きな違いが見られないことは明らかにできたものの、本研究で用いたデータからは発生集積の原因を明らかにしえなかった。これらの発生増加の原因を探るためには、より詳細な救急搬送記録や医療データなどを用いて、さらなる解析、検討が必要となる。また海外においてもこの年末年始時期における心肺停止傷病者の発生集積などの研究は、我々が知る限り行われていないが、海外での状況についての検討も行い、比較を行うことも必要であろう。なお平成22年度救急業務高度化推進検討会報告書にあるとおり、平成25年から新たな調査項目が追加された救急業務統計の報告が行われる予定である。そこで追加された「疾病」の分類や「発生場所」の区分などを利用した解析を行うことができれば、予防対策をする上でも有用な要因の分析が可能になると考えられる。今後、今回明らかになったわが国における年末年始時期における心肺機能停止傷病者発生の集積の要因について明らかにしていくことは、社会的な対策を探る上での重要な情報になると考えられ、公衆衛生・救急医学における重要な研究課題として引き続き検討を行う必要がある。

E. 結論

わが国における心肺機能停止傷病者の救急搬送件数について、その発生が多くなると知られている冬季において、その中でも特に年末年始時期に統計的に有意な搬送数の集積が検出された。今後の対策や原因を検討する上で有益な情報になると考えられる。

F. 研究発表

1. 論文発表

- ・高橋邦彦, 武村真治, 長谷川学, 金谷泰宏, 齋藤大蔵, 阪本敏久. わが国における救急蘇生統計を用いた冬季の心肺機能停止傷病者救急搬送件数の時間集積性の検出. 日本臨床救急医学会雑誌 2012;15(5):652-661.

2. 学会発表

- ・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
件数の時間的集積性の検出. 第70回日本公衆衛
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, 20
11年12月11日.

3. 著書

なし

表1 検出された有意な集積期間

(男女別, 総数)

		全国総件数	1日あたり期待件数	検出された集積期間	日数	期間内件数	p値	RR
2005年12月 ~2006年1月	男性	14,442	232.94	① 12/30-1/10	12	3,439	0.0001	1.23
	女性	10,216	164.77	① 12/30-1/11	13	2,587	0.0001	1.21
2006年12月 ~2007年1月	男性	13,629	219.82	① 12/29-1/9	12	3,171	0.0001	1.20
				② 1/15	1	276	0.0298	1.26
	女性	9,437	152.21	① 1/1-1/4	4	889	0.0001	1.46
				② 1/5-1/9	5	869	0.0102	1.14
2007年12月 ~2008年1月	男性	14,356	231.55	① 1/1-1/4	4	1,331	0.0001	1.44
				② 1/27-1/29	3	809	0.0022	1.16
	女性	10,388	167.55	① 12/31-1/7	8	1,620	0.0001	1.21
				② 1/14-1/28	15	2,725	0.0005	1.08

(男女別, 心原性のみ)

		全国総件数	1日あたり期待件数	検出された集積期間	日数	期間内件数	p値	RR
2005年12月 ~2006年1月	男性	8,177	131.89	① 12/30-1/10	12	1,930	0.0001	1.22
	女性	6,020	97.10	① 12/29-1/11	14	1,652	0.0001	1.22
2006年12月 ~2007年1月	男性	7,422	119.71	① 1/1-1/7	7	1,053	0.0001	1.26
	女性	5,209	84.02	① 12/30-1/9	11	1,177	0.0001	1.27
2007年12月 ~2008年1月	男性	7,931	127.92	① 1/1-1/4	4	718	0.0001	1.40
				② 1/27-1/29	3	467	0.0040	1.22
	女性	5,880	94.84	① 12/30-1/6	8	927	0.0001	1.22
				② 1/14-1/15	2	248	0.0073	1.31
				③ 1/27-31	5	552	0.0324	1.16

(男女別, 非心原性のみ)

		全国総件数	1日あたり期待件数	検出された集積期間	日数	期間内件数	p値	RR
2005年12月 ~2006年1月	男性	6,265	101.05	① 1/1-1/3	3	495	0.0001	1.63
	女性	4,196	67.68	① 12/30-1/5	7	614	0.0001	1.30
2006年12月 ~2007年1月	男性	6,207	100.11	① 1/1-1/2	2	344	0.0001	1.72
	女性	4,228	68.19	① 1/1-1/4	4	413	0.0001	1.51
2007年12月 ~2008年1月	男性	6,425	103.63	① 12/31-1/4	5	747	0.0001	1.44
	女性	4,508	72.71	① 1/1-1/3	3	323	0.0001	1.48

研究成果の刊行に関する一覧表

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Tango T. Takahashi K.	A flexible spatial scan statistic with a restricted likelihood ratio for detecting disease clusters.	Statistics in Medicine	31	4207-4218	2012
高橋邦彦, 武村真治, 長谷川学, 金谷泰宏, 齋藤大蔵, 阪本敏久	わが国における救急蘇生統計を用いた冬季の心肺機能停止傷病者救急搬送件数の時間的集積性の検出	日本臨床救急医学会雑誌	15	652-661	2012

A flexible spatial scan statistic with a restricted likelihood ratio for detecting disease clusters

Toshiro Tango^{a,*†} and Kunihiro Takahashi^b

Spatial scan statistics are widely used tools for detection of disease clusters. Especially, the circular spatial scan statistic proposed by Kulldorff (1997) has been utilized in a wide variety of epidemiological studies and disease surveillance. However, as it cannot detect noncircular, irregularly shaped clusters, many authors have proposed different spatial scan statistics, including the elliptic version of Kulldorff's scan statistic. The flexible spatial scan statistic proposed by Tango and Takahashi (2005) has also been used for detecting irregularly shaped clusters. However, this method sets a feasible limitation of a maximum of 30 nearest neighbors for searching candidate clusters because of heavy computational load. In this paper, we show a flexible spatial scan statistic implemented with a restricted likelihood ratio proposed by Tango (2008) to (1) eliminate the limitation of 30 nearest neighbors and (2) to have surprisingly much less computational time than the original flexible spatial scan statistic. As a side effect, it is shown to be able to detect clusters with any shape reasonably well as the relative risk of the cluster becomes large via Monte Carlo simulation. We illustrate the proposed spatial scan statistic with data on mortality from cerebrovascular disease in the Tokyo Metropolitan area, Japan. Copyright © 2012 John Wiley & Sons, Ltd.

Keywords: cluster detection; hotspot cluster; likelihood ratio statistic; Monte Carlo testing; spatial epidemiology

1. Introduction

Many authors have proposed and introduced in several textbooks many different statistical methods for detecting disease clustering and disease clusters [1–4]. Especially, the spatial scan statistic proposed by Kulldorff and Nagarwalla [5] and Kulldorff [6] along with SaTScan software [7] has been applied in a wide variety of epidemiological studies and also in disease surveillance for detection of disease clusters. The spatial scan statistic tries to identify the *most likely cluster* (MLC), which is defined as the set of connected regions that attains the maximum likelihood ratio. However, because it uses a circular window to scan the potential cluster areas, it has difficulty in correctly detecting actual noncircular clusters. To detect arbitrarily shaped clusters that cannot be detected by the circular spatial scan statistic, many authors [8–11] have proposed different spatial scan statistics. Kulldorff *et al.* [12] also proposed an elliptic version of spatial scan statistic. Above all, the flexible spatial scan statistic proposed by Tango and Takahashi [10] along with FleXScan software [13] has also been used for detecting irregularly shaped clusters. However, this method allows only a maximum of 30 nearest neighbors for searching candidate clusters because of heavy computational load. Several authors who conducted comparison of tests for disease clusters pointed out this limitation (e.g., [14]).

On the other hand, Tango [15] and Tango and Takahashi [10] have shown that these spatial scan statistics based on maximizing the ordinary likelihood ratio tend to detect an MLC that is much larger than the true cluster by absorbing neighboring regions with nonelevated risk of disease occurrence. To

^aCenter for Medical Statistics, 2-9-6 Higashi Shimbashi, Minato-ku, Tokyo, 105-0021, Japan

^bDepartment of Technology Assessment and Biostatistics, National Institute of Public Health, 3-6 Minami 2 chome, Wako, Saitama, 351-0197, Japan

*Correspondence to: Toshiro Tango, Center for Medical Statistics, 2-9-6 Higashi Shimbashi, Minato-ku, Tokyo, 105-0021, Japan.

†E-mail: tango@medstat.jp

avoid such an undesirable property, Tango [16] proposed a new spatial scan statistic by restricting the likelihood ratio so that it scans only the regions with elevated risk. Tango [16] compared the performance of detecting *circular clusters* by the two methods, that is, the original Kulldorff's circular spatial scan statistic and the circular spatial scan statistic implemented with the proposed restricted likelihood ratio, via a Monte Carlo simulation study. The results indicated that the latter procedure had the ability to identify the whole, or a part, of the true cluster assumed in the simulation more often than the former.

In this paper, we shall propose a flexible spatial scan statistic implemented with the restricted likelihood ratio. The performance of the proposed spatial scan statistic with the restricted likelihood ratio is compared with those of Tango and Takahashi's original flexible spatial scan statistic and Kulldorff's circular spatial scan statistic via a Monte Carlo simulation study. We illustrate the proposed procedure with data on mortality from cerebrovascular disease in the areas of Tokyo Metropolis and Kanagawa prefecture in Japan.

2. Spatial scan statistics

Consider the situation where the entire study area is divided into m regions (e.g., counties and enumeration districts). The number of cases in region i is denoted by the random variable N_i with observed value n_i ($i = 1, \dots, m$) and $n = n_1 + \dots + n_m$. Under the null hypothesis H_0 of no clustering, the N_i are independent Poisson variables such that

$$H_0 : E(N_i) = \xi_i, N_i \sim \text{Poisson}(\xi_i), \quad i = 1, \dots, m, \quad (1)$$

where $\text{Poisson}(\xi)$ denotes the Poisson distribution with mean ξ and the ξ_i are the expected number of cases in region i under the null hypothesis. For calculation of the expected number of cases adjusted for potential confounders such as age, we can use indirect standardization or a Poisson mixed-effects regression model [7]. If we ignore the confounders, we can calculate ξ_i as

$$\xi_i = n \frac{w_i}{\sum_{k=1}^m w_k}, \quad i = 1, \dots, m, \quad (2)$$

where w_i denotes the population size in region i . To specify the geographical position of each region, we will use the coordinates of the administrative population centroid.

2.1. Windows to be scanned

In the aforementioned situation, Kulldorff's circular spatial scan statistic imposes a circular window \mathbf{Z} on each centroid of regions. For any of those centroids, the radius of the circle varies continuously from zero up until 50% of the population at risk is covered, which is the standard option in SaTScan. If the window contains the centroid of a region, then that whole region is included in the window. In total, a very large number of different but overlapping circular windows are created, each with a different location and size, and each being a potential cluster. Let \mathbf{Z}_{ik} ($k = 1, \dots, K_i$) denote the window composed of the $(k-1)$ -nearest neighbors to region i . Then, all of the windows to be scanned by the circular spatial scan statistic are included in the set

$$\mathcal{Z} = \mathcal{Z}_1 = \{\mathbf{Z}_{ik} \mid 1 \leq i \leq m, 1 \leq k \leq K_i\}.$$

Tango and Takahashi's flexible spatial scan statistic, on the other hand, imposes a flexibly shaped window \mathbf{Z} on each centroid of region by connecting its adjacent regions. For any given region i , we create the set of flexibly shaped windows with *length* k consisting of k connected regions including i and let k move from 1 to the prespecified maximum length K of nearest neighbors. To avoid detecting a cluster of *unlikely peculiar shape*, the connected regions are restricted to the subsets of the set of regions i and K -nearest neighbors to the region i . In total, as in the circular spatial scan statistic, a very large number of different but overlapping arbitrarily shaped windows are created. Let $\mathbf{Z}_{ik(j)}$, $j = 1, \dots, j_{ik}$ denote the j th window, which is a set of k regions connected starting from the region i , where j_{ik} is the number of j satisfying $\mathbf{Z}_{ik(j)} \subseteq \mathbf{Z}_{ik}$ for $k = 1, \dots, K_i = K$. Then, all of the windows to be scanned are included in the set

$$\mathcal{Z} = \mathcal{Z}_2 = \{\mathbf{Z}_{ik(j)} \mid 1 \leq i \leq m, 1 \leq k \leq K, 1 \leq j \leq j_{ik}\}.$$