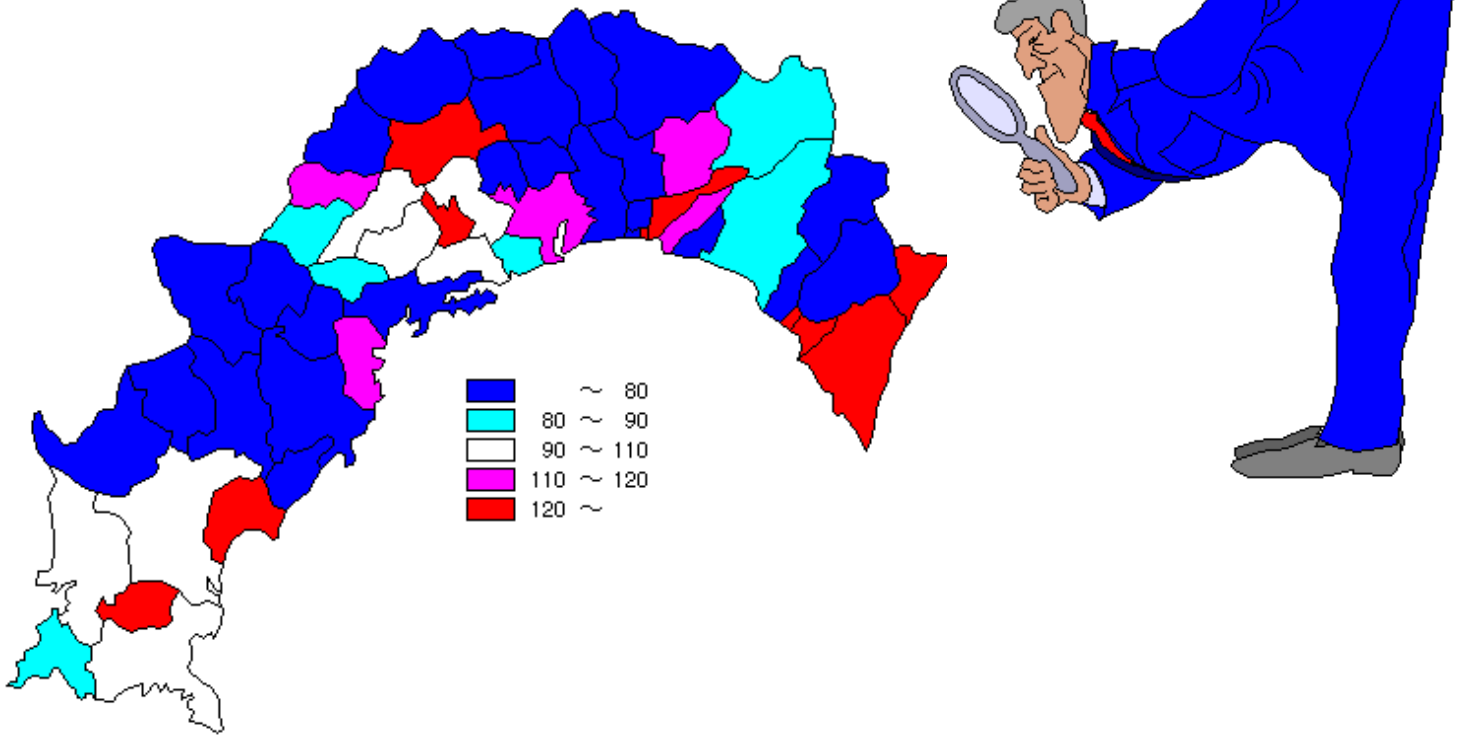
統計的発想を刺激するVisual tool: S-PLUS

研究事例：疾病地図と疾病集積性について

Toshiro Tango

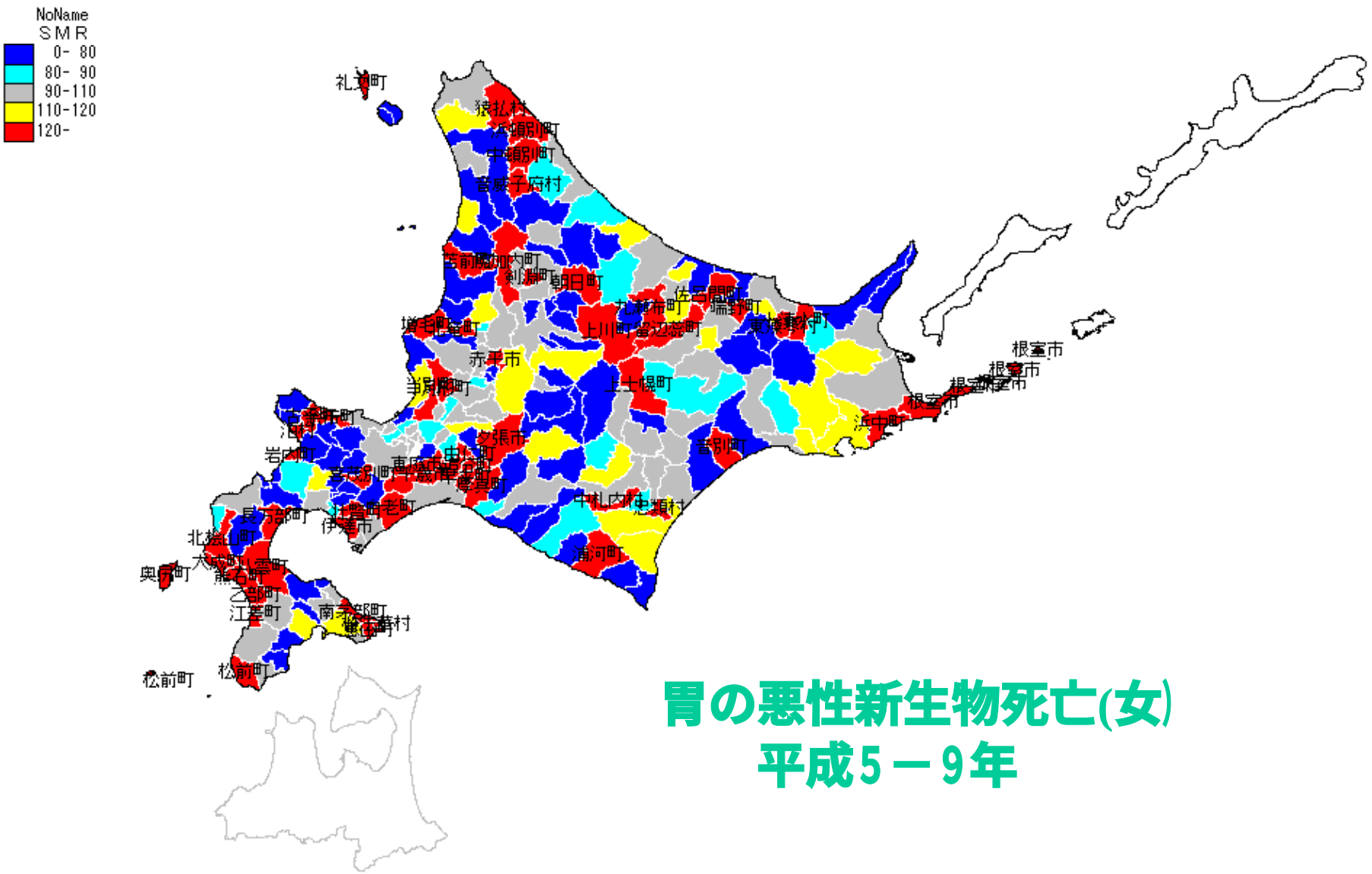
Departement of Technology Assessment and Biostatistics

The National Institute of Public Health

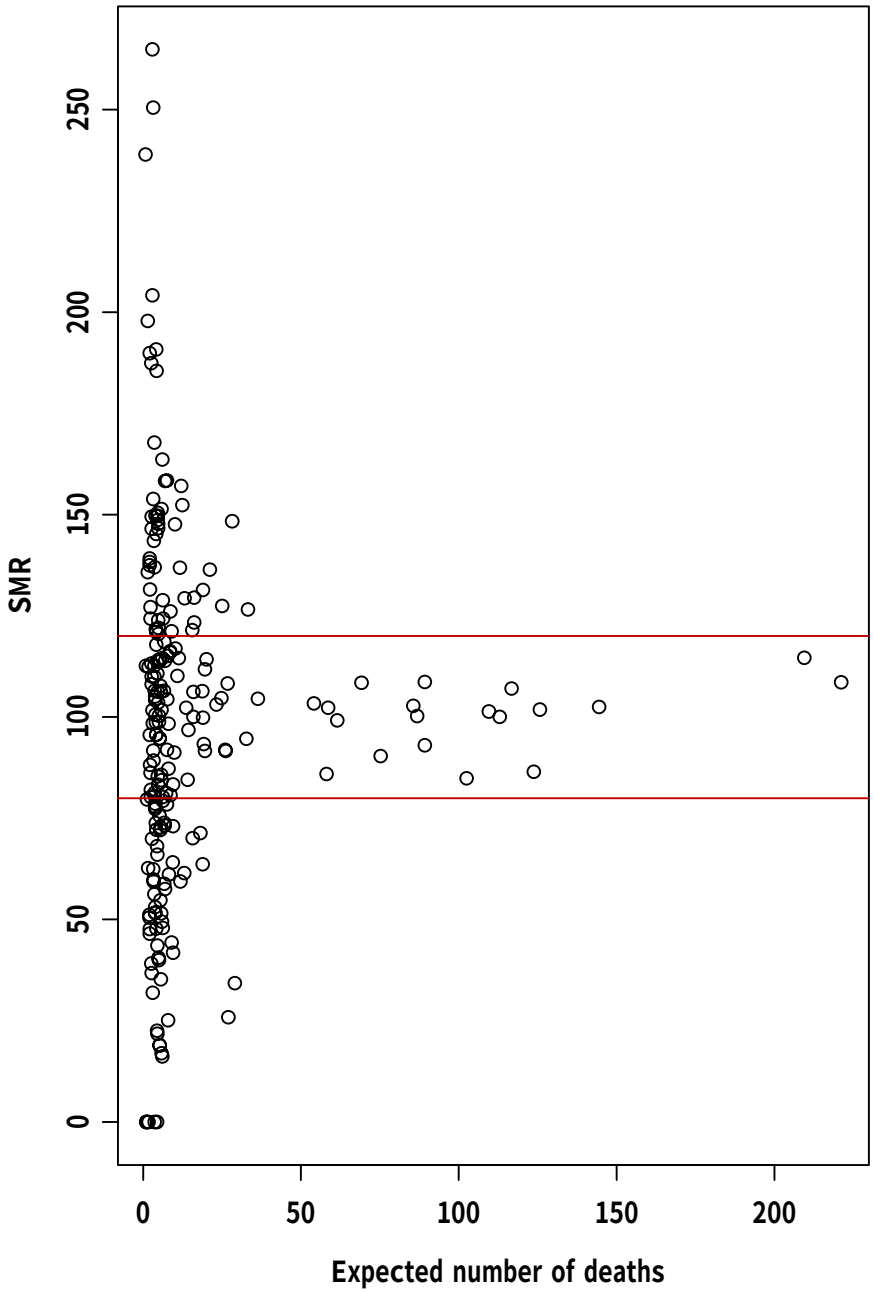


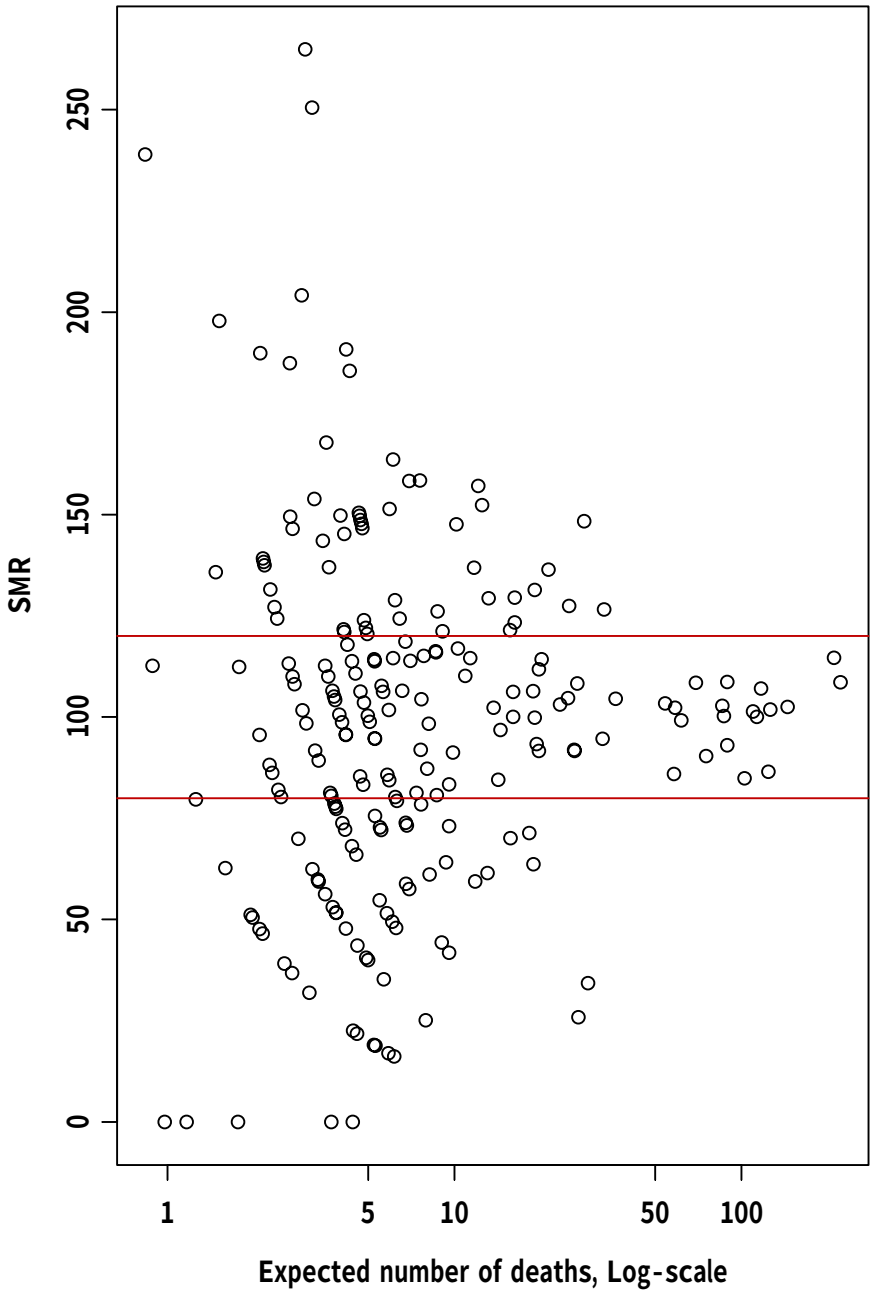
WHO専門者会議(1997, Rome)
Disease Mapping and Risk Assessment for Public Health Decision Making

- 信頼性のある統計指標
- 国際比較が可能な統計指標
- 疾病地図の推定法
- 疾病の集積性を表現する指標



胃の悪性新生物死亡(女)
平成5－9年


SMR
0
50
100
150
200
250
0
50
100
150
200
Expected number of deaths


SMR
0
50
100
150
200
250
1
5
10
50
100
Expected number of deaths, Log-scale

健康指標の「計算」

$$\text{標準化死亡比} \overset{\text{def}}{=} \frac{\text{観測死亡数}}{\text{期待死亡数}}$$

$$\text{SMR}(\lambda) = \frac{d}{E} \times 100$$

健康指標の「推測」

$$\hat{\lambda} \quad \text{Est.} = \frac{\text{観測死亡数}}{\text{期待死亡数}} = \frac{d}{E}$$

＜仮定＞

λ：未知のSMR

観測死亡数 d ----- ポアッソン分布 (λE)

期待死亡数 $E = \sum_i$ (人口)$_i$ (標準死亡率)$_i$

λ を母数と考えた最尤推定値

統計的推測に関する二つの立場

Frequentist (頻度論学派)

λ を**母数(定数)** と考える

伝統的な統計学

Bayesian (ベイジアン ,ベイズ学派)

λ を**確率変数**と考える

２１世紀の統計学の方向

Frequentist

推定したいパラメータ λ は母数(定数) と考える

繰り返し可能な実験、標本調査を想定

頻度分布(ヒストグラム) => 確率分布

f (データ | パラメータ)

(例)　死亡数　d　　ポアッソン分布　$f(d|\lambda)$

　　計測値　x　　正規分布　$f(x|\mu, \sigma 2)$

Bayesian

推定するパラメータ λ を定数と考えるのは不自然

(why?) ヒトの臨床検査値, 死亡率, などほとんどのパラメータは時間差、地域差、個人差などの不確実性 (variability)があるもの

λ も確率分布にしたがう確率変数と考える

繰り返し可能でない事象にも適用可能

(例:) あいつが彼を殺した確率は 0.9 以上だ！

Bayesian computation

- パラメータ λ の不確実性を確率分布
  p(λ)　(事前分布, prior distribution)
- データの確率分布 f (d | λ)
- 同時分布 p(λ, d)　を考えることができる
  p(λ , d) = p(λ) f (d | λ)

- データをとる前の不確実性 p(λ) をデータの情報で更新する．更新された不確実性を
  p(λ | d)　(事後分布、posterior distribution)
- 点推定としては「事後分布の期待値(平均値)」

事後分布の計算
--Bayesの定理--

$$p(\lambda \mid x) = \frac{p(\lambda, x)}{p(x)} = \frac{p(\lambda, x)}{\int p(\lambda, x)\, d\lambda}$$

事後分布

$$= \frac{p(\lambda)\, f(x \mid \lambda)}{\int p(\lambda)\, f(x \mid \lambda)\, d\lambda}$$

$$\propto p(\lambda)\ f(x \mid \lambda) = \text{事前分布} \times \text{尤度}$$

Bayesianの問題

- 事前分布 p(λ) の選び方
  - 従来：自然共役(natural conjugate)
  - 最近：なんでもO.K.
- 事前分布p(λ)に含まれるパラメータの推定法
  - Full Bayes：パラメータも確率変数
  - Empirical Bayes：データから推定
- Bayesの定理に含まれる積分の計算法
  - 従来：不可能であきらめることが多かった
  - 最近：Markov Chain Monte Carlo

Empirical Bayes estimate for SMR

観測死亡数 d_k は期待死亡数 $\lambda_k E_k$ をもつポアッソン分布：

$$f(d_k|\lambda_k, E_k) = \frac{(\lambda_k E_k)^{d_k} \exp(-\lambda_k E_k)}{d_k!}$$

パラメータ λ_k の事前分布にはガンマ分布：

$$g(\lambda_k \mid \alpha, \beta) = \frac{\alpha(\alpha\lambda_k)^{\beta-1} exp(-\alpha\lambda_k)}{\Gamma(\beta)}$$

$$E(\lambda_k) = \frac{\beta}{\alpha}$$

$$\mathrm{Var}(\lambda_k) = \frac{\beta}{\alpha^2}$$

ポアッソン分布とガンマ分布は自然共役

λ_k の事後分布はBayesの定理より，$\boldsymbol{\eta} = (\alpha, \beta)$ とおいて、

$$h(\lambda_k \mid E_k, d_k, \boldsymbol{\eta}) = \frac{p(\lambda_k \mid \boldsymbol{\eta}) f(d_k \mid \lambda_k, E_k)}{\int_0^\infty p(\lambda_k \mid \boldsymbol{\eta}) f(d_k \mid \lambda_k, E_k) d\lambda_k}$$

したがって、SMR$(= \lambda_k)$ の推測は、事後分布からの期待値

$$\hat{\lambda}_k \Leftarrow E(\lambda_k \mid E_k, d_k, \boldsymbol{\eta}) = \int_0^\infty \lambda_k h(\lambda_k \mid E_k, d_k, \boldsymbol{\eta}) d\lambda_k$$

$$= \frac{\int_0^\infty \lambda_k p(\lambda_k \mid \boldsymbol{\eta}) f(d_k \mid \lambda_k, E_k) d\lambda_k}{\int_0^\infty p(\lambda_k \mid \boldsymbol{\eta}) f(d_k \mid \lambda_k, E_k) d\lambda_k}$$

事後分布 h もガンマ分布

Bayes 推定値は

$$\hat{\lambda}_{EB,k} = \frac{\hat{\beta} + d_k}{\hat{\alpha} + E_k}$$
$$= \frac{E_k}{\hat{\alpha} + E_k} \frac{d_k}{E_k} + \frac{\hat{\alpha}}{\hat{\alpha} + E_k} \frac{\hat{\beta}}{\hat{\alpha}}$$

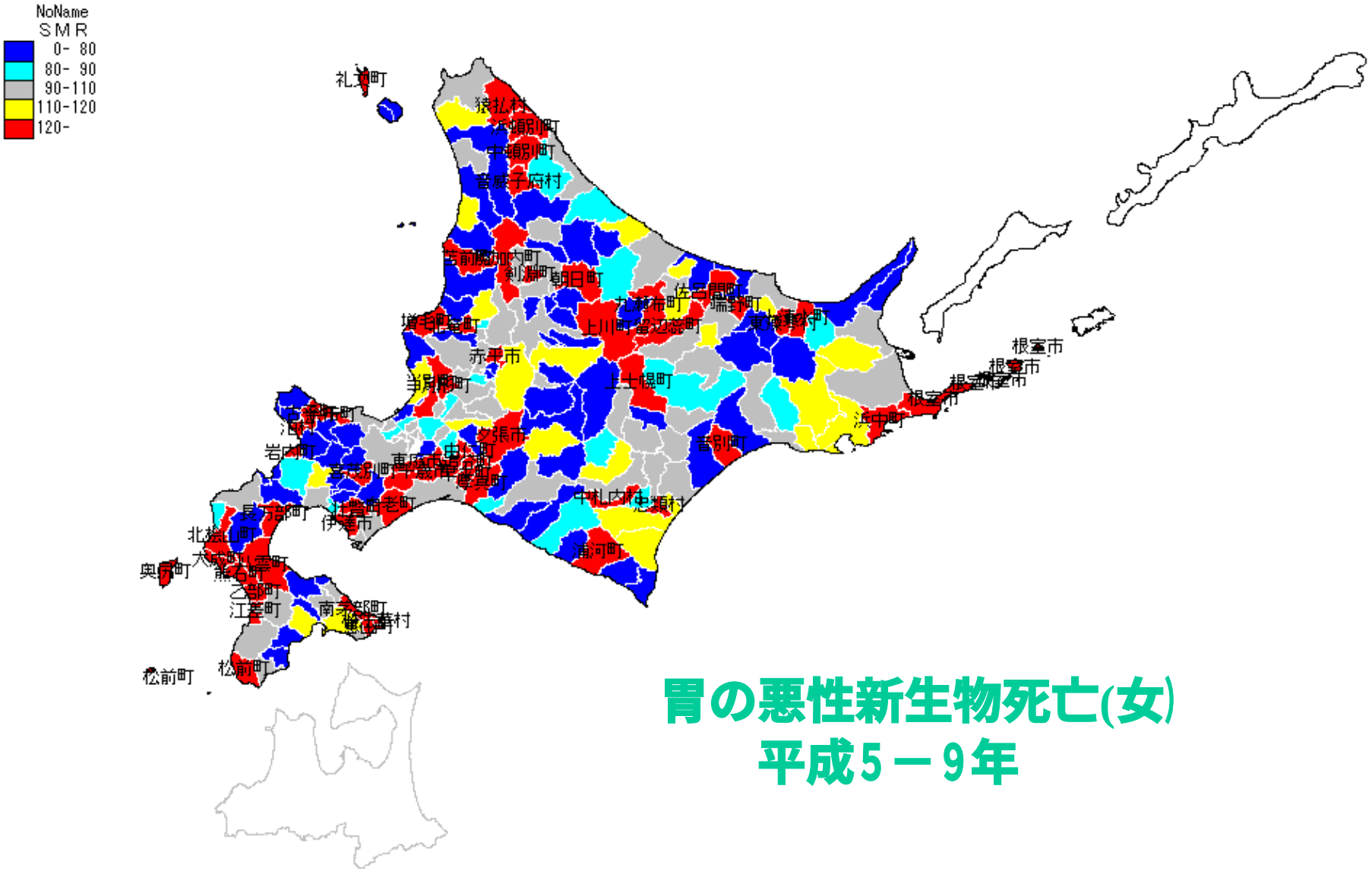
となる。この式の形から $\hat{\lambda}_{EB,k}$ は

1) 人口が大きい場合には（$E_k \to$ 大)、通常の標準化死亡比 $\hat{\lambda}_k = d_k/E_k$ に近づき、
2) 人口が少ない場合には（$E_k \to$ 小)、地域全体の平均値 $\hat{\beta}/\hat{\alpha}$ に近づく。

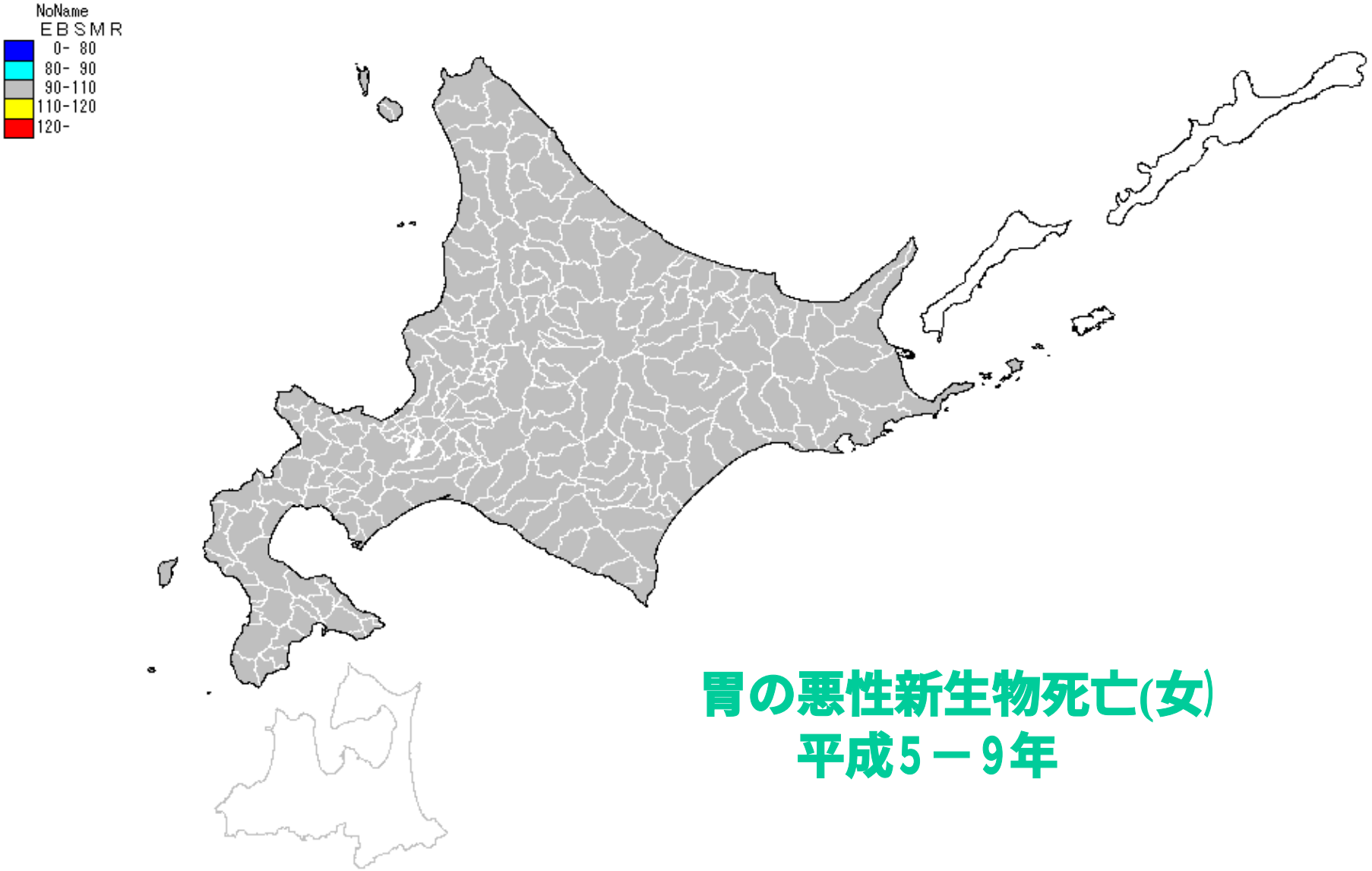
Bayesian inference

- 情報の無いときは地域全体の平均値
- 情報が増加してくればその地域のデータ

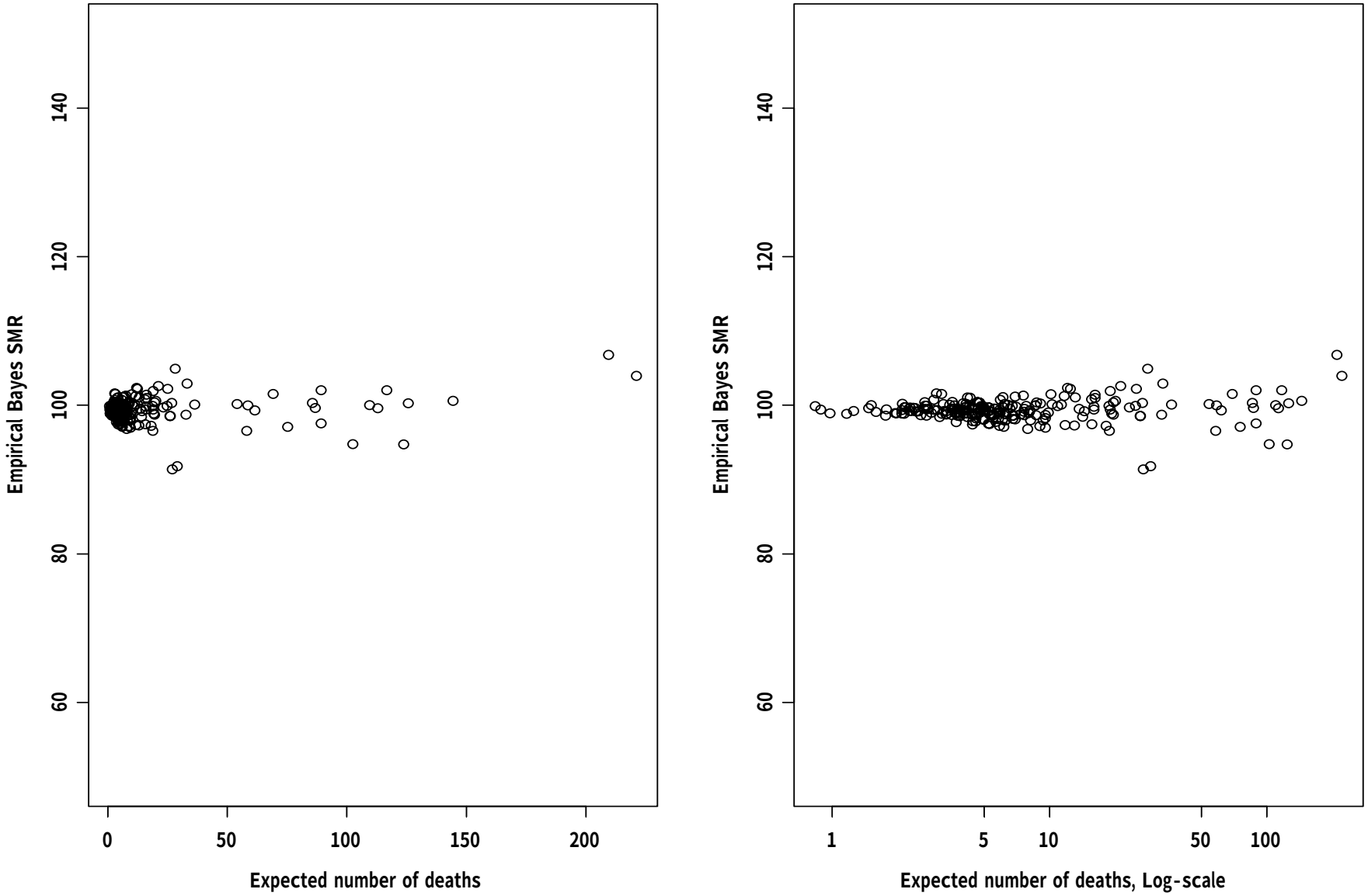
- 診断･治療における臨床検査値の読み方
  1） 初診時に正常範囲（他のヒトのデータ）
  2） 入院患者の状態は患者の過去のデータ



胃の悪性新生物死亡(女)
平成5－9年



胃の悪性新生物死亡(女)
平成5－9年


Empirical Bayes SMR
60
80
100
120
140
0
50
100
150
200
Expected number of deaths
Empirical Bayes SMR
60
80
100
120
140
1
5
10
50
100
Expected number of deaths, Log-scale

c:¥tan¥asak¥aaa.dvi - dviout

File Jump Search Display View Option Help

An ordinary ecological regression

このような図を見ると、論文等でよくみかける次のような回帰分析が如何に馬鹿げているか理解できるはずである。

$$\mathrm{SMR} = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m + \text{誤差}$$

このように、地域の比較を行うためには、「人口の大きさを調整」しなければならない。

Page: 1/7, number 1/7 | dpi: x=300/2, y=300/2 | Gamma = 800/1000 | Size: x = 21.00cm, y = 29.70cm

スタート | Microso... | MS-DO... | S-PLUS... | エクスプロ... | aaa.tex ... | Jniph99... | c:¥tan... | 22:21

疾病の集積性

- Bayesian approach は人口のサイズを調整する方法として重要であることを解説した
- しかし、どんな指標であっても小さい順に並べれば必ず、最小値、最大値が存在する
- 本当に健康の良い、思わしくない地域の同定は

- **疾病の集積性の検定**

(Test for disease clustering)

疾病の集積性検定のタイプ

Global test:

- **Study area の特定の地域（未知）に疾病が集積している？**
- **感染症のように地域のいたるところで集積性が点在している？**

Focused test:

- **原発施設、ごみ焼却施設、危険物廃棄施設等の周辺に疾病が集積している？**

二つの集積性検定

Kulldorff ’s Spatial Scan Test:

- 集積している地域の推定
- 過疎地域の集積性の検出力が高い

(1995, Stat in Med; 1997, Com in Stat)

Tango’s Maximized Excess Events Test:

- Tango’s index (1984, Biometrics) の拡張
- 集積している地域の中心を推定
- 都市部の集積性の検出力が高い

(1995, 2000, Stat in Med)

Focused test : **Tango(1995, 2000)**

$$
\begin{aligned}
d_k &\sim \text{Poisson}(E(d_k)),\ k = 1, ..., K \\
H_0 : E(d_k) &= \tau n_k \quad \text{, no clustering} \\
H_1 : E(d_k) &= \tau n_k (1 + a_{k,k_0}\theta), \text{cluster around } k_0 \\
a_{kh}(\lambda) &= \exp\{-4(\frac{d_{kh}}{\lambda})^2\} \\
d_{kh} &= \text{distance between } (k, h)
\end{aligned}
$$

where λ denotes the cluster size.

Efficient score for this test is

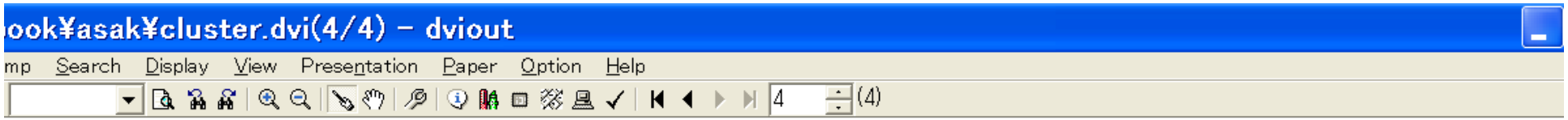
$$U_{k_0}(\lambda) = \sum_{k=1}^{K} a_{k_0,k}(\lambda)(d_k - E(d_k))$$

$$= \sum_{k=1}^{K} w_{k_0} a_{k_0,k}(\lambda)(d_k - E(d_k))$$

where $w_k = 1(k = k_0); = 0$ (otherwise). When the cluster region is unknown, by putting $w_k = d_k - E(d_k)$, the above score can lead to a **Global test**:

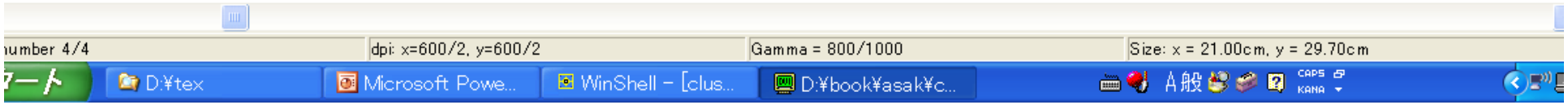
$$
\begin{aligned}
C_\lambda &= \sum_{k=1}^{K} \left\{ \sum_{h=1}^{K} a_{k,h}(\lambda)(d_k - E(d_k))(d_h - E(d_h)) \right\} \\
&= \sum_{k=1}^{K} U_k(\lambda) \sim \chi^2 \text{ approximation (quite good)} \\
P_{min} &= \min_{\lambda} \Pr\{C_\lambda > c_\lambda \mid H_0, \lambda\} \\
&= \Pr\{C_\lambda > c_\lambda \mid H_0, \lambda = \lambda^*\}
\end{aligned}
$$

where λ should move in the range $0 < \lambda \leq d_{max}/4$.



If test result is significant, the most likely cluster center could be the region which has the largest contribution in terms of

$$\frac{U_k(\lambda^*)}{C_{\lambda^*}} \times 100(\%)$$



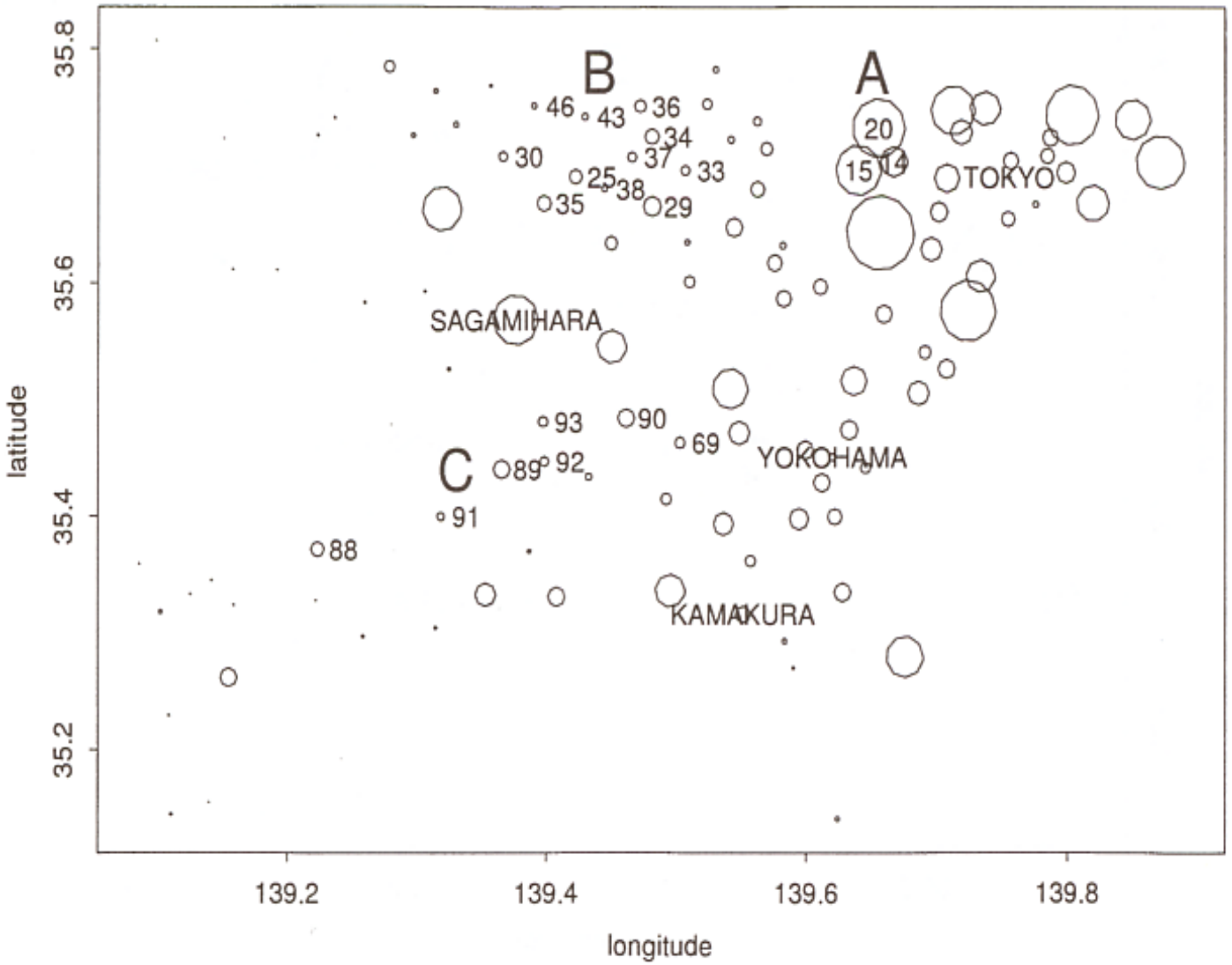


Figure 1. The 113 regions comprising wards, cities and villages in the area of Tokyo Metropolis and Kanagawa prefecture in Japan. The centre of a circle is the location of population centroid of the corresponding region and the radius is set proportional to the population size. Three types of clusters are indicated by region numbers and symbols 'A', 'B' and 'C'. For more details, see text

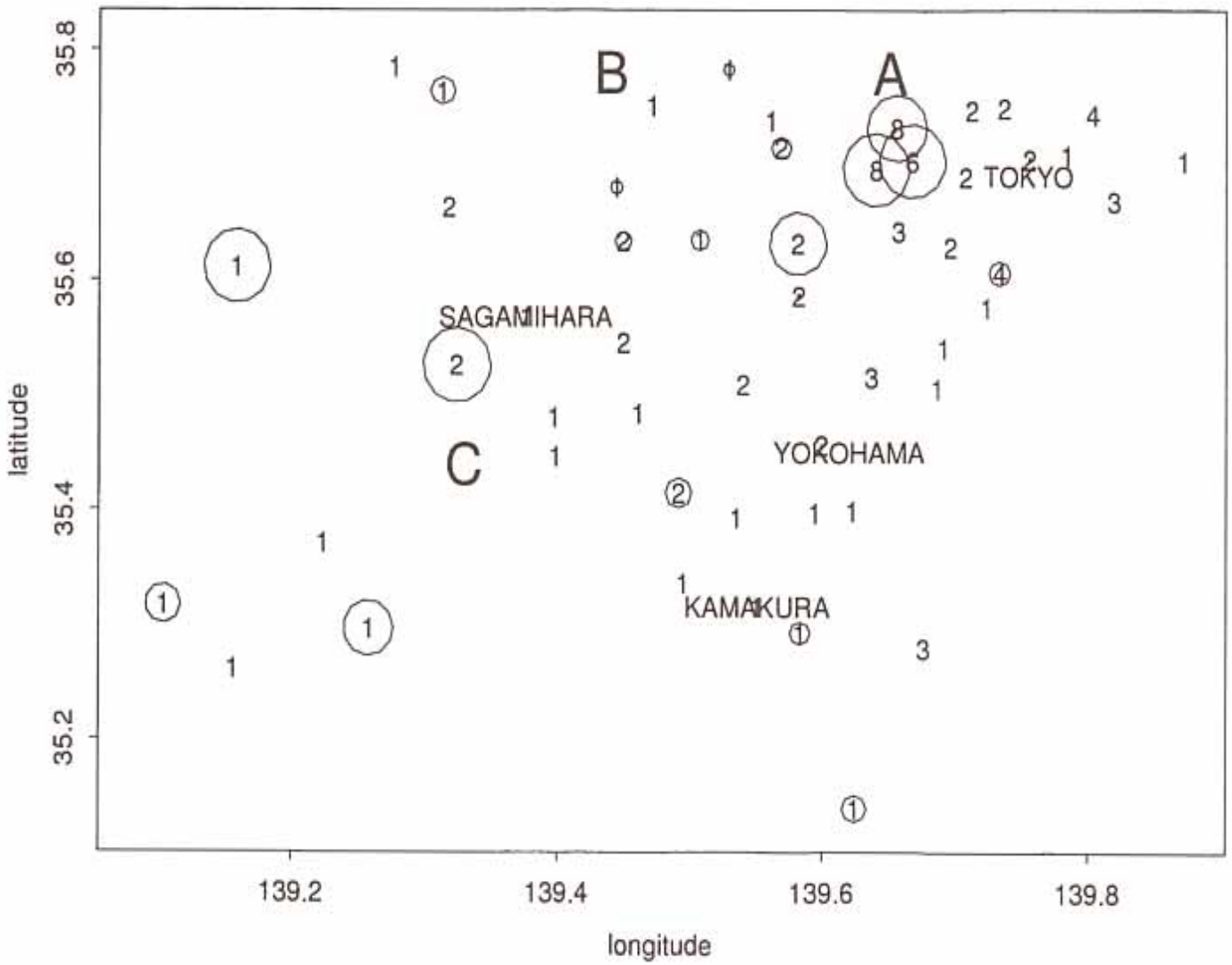
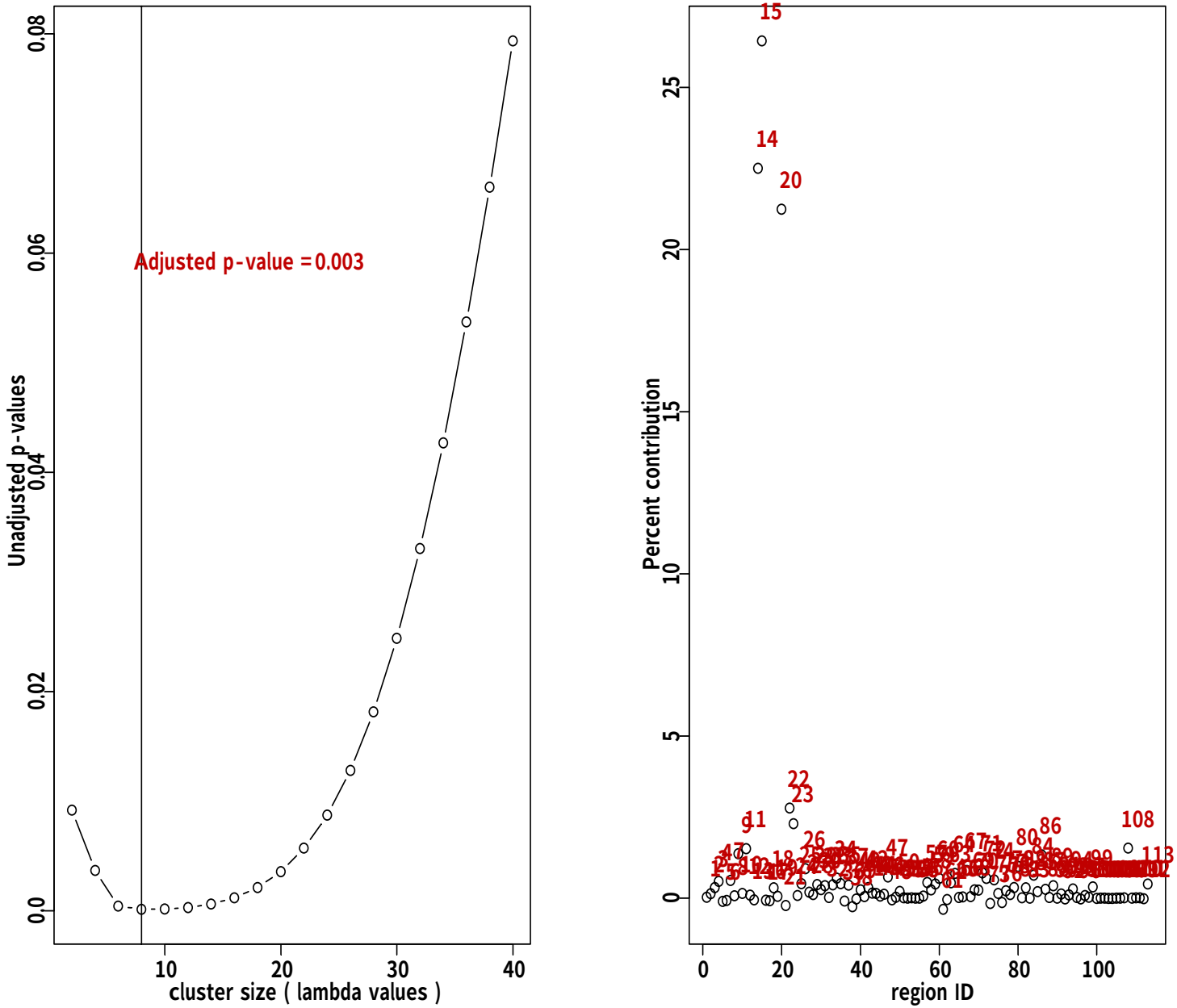
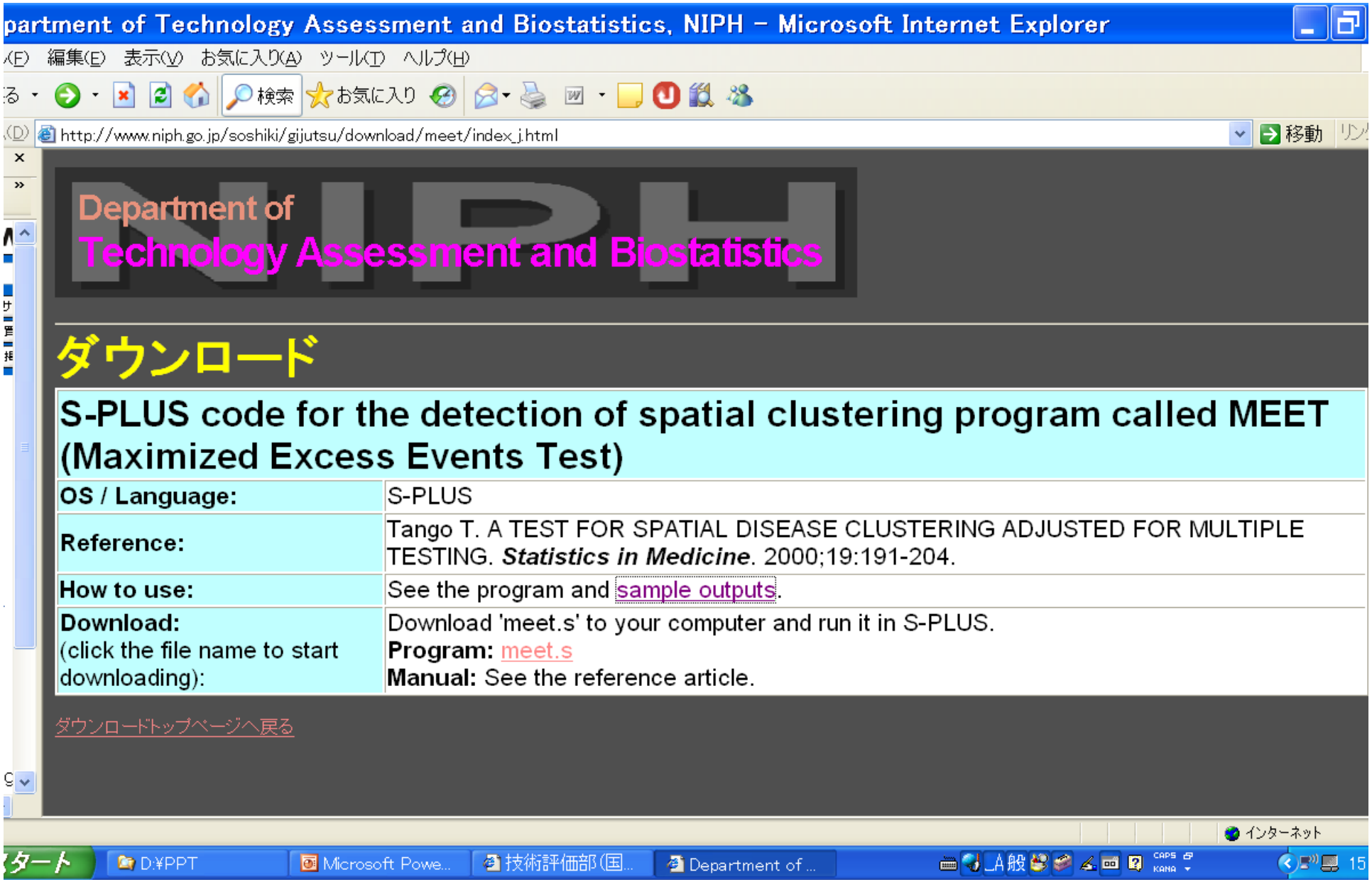


Figure 2. A random sample from clustering model A. Circles are drawn only for the regions whose standardized risk ratios are statistically significantly larger than 1 at $\alpha = 0{\cdot}05$. The radius is set inversely proportional to the tail probability. The number shown in the map indicates the number of observed cases $n_i (> 0)$


Adjusted p-value = 0.003
Unadjusted p-values
cluster size (lambda values)
Percent contribution
region ID


partment of Technology Assessment and Biostatistics, NIPH - Microsoft Internet Explorer
(F) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)
検索 お気に入り
http://www.niph.go.jp/soshiki/gijutsu/download/meet/index_j.html
移動
Department of
Technology Assessment and Biostatistics
ダウンロード
S-PLUS code for the detection of spatial clustering program called MEET (Maximized Excess Events Test)
OS / Language:
S-PLUS
Reference:
Tango T. A TEST FOR SPATIAL DISEASE CLUSTERING ADJUSTED FOR MULTIPLE TESTING. Statistics in Medicine. 2000;19:191-204.
How to use:
See the program and sample outputs.
Download:
(click the file name to start downloading):
Download 'meet.s' to your computer and run it in S-PLUS.
Program: meet.s
Manual: See the reference article.
ダウンロードトップページへ戻る
インターネット
スタート
D:¥PPT
Microsoft Powe...
技術評価部(国...
Department of ...

Spatial scan statistic (Poisson model) --- Kulldorff (1997)

Assume that the study area consisting of ***m*** regions and the number of cases ***Xi*** in the region ***i*** is a Poisson random variable,

$$X_i \overset{\text{i.i.d.}}{\sim} Po(\eta_i p_i) \quad (i = 1, 2, \cdots, m)$$
$$\eta_i > 0, \ 0 < p_i < 1$$

Define the **windows** Z (the set of regions) and the **collection** Ω of all the windows. Then, for some Z $(\in \Omega)$, we consider the **hot spot model**:

$$p_i = p \ \ (i \in Z); \quad p_i = q \ \ (i \notin Z)$$

Spatial scan statistic (Poisson model)

$$H_0 : p = q \qquad H_1 : p > q$$

$$n(Z) := \sum_{i \in Z} x_i \ , \quad \mu(Z) := \sum_{i \in Z} \eta_i \ , \quad G = Z \cup Z^c$$

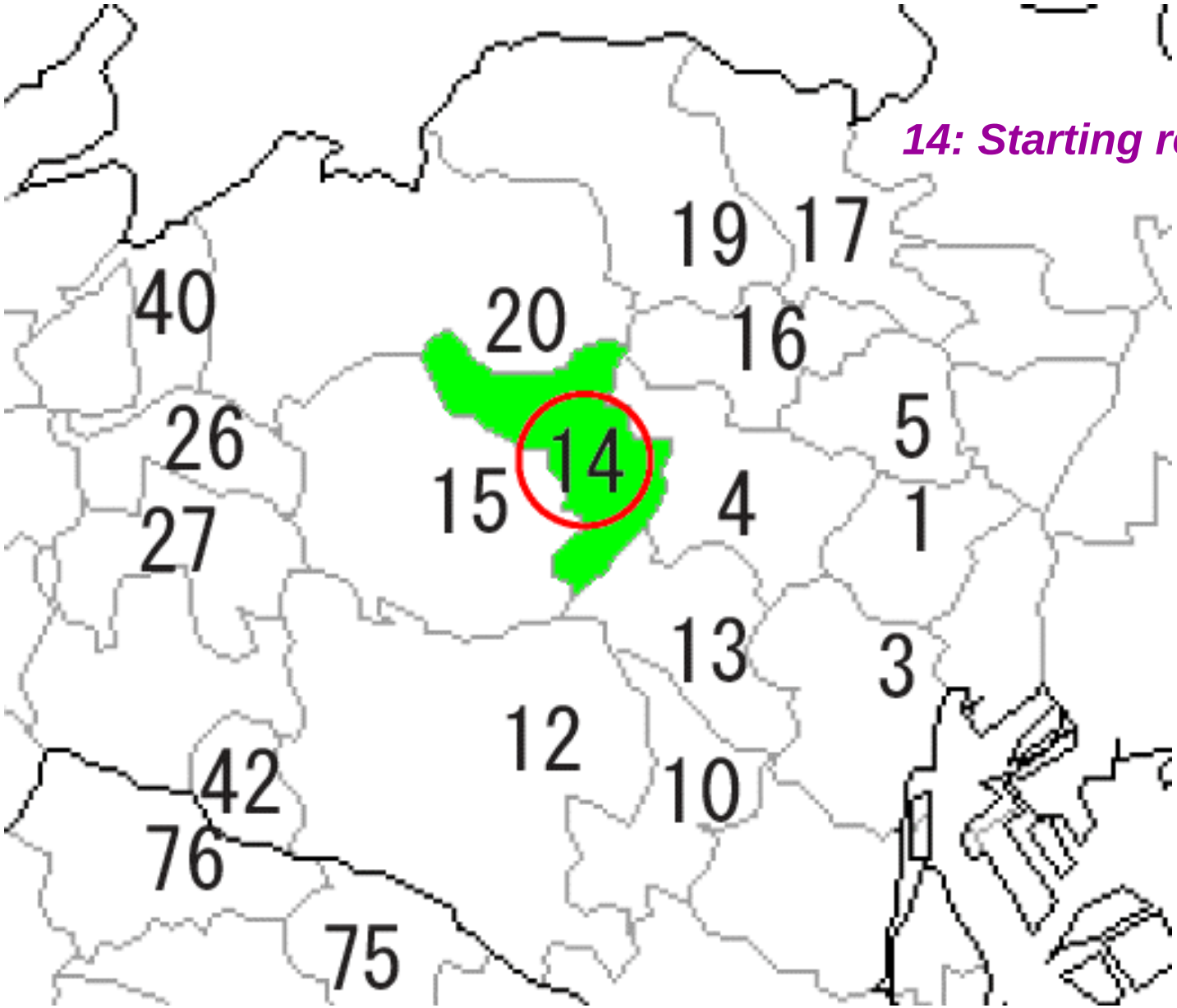
$$\lambda := \sup_{Z \in \Omega} \frac{\left(\frac{n(Z)}{\mu(Z)}\right)^{n(Z)} \left(\frac{n(Z^c)}{\mu(Z^c)}\right)^{n(Z^c)}}{\left(\frac{n(G)}{\mu(G)}\right)^{n(G)}} I\left(\frac{n(Z)}{\mu(Z)} > \frac{n(Z^c)}{\mu(Z^c)}\right)$$

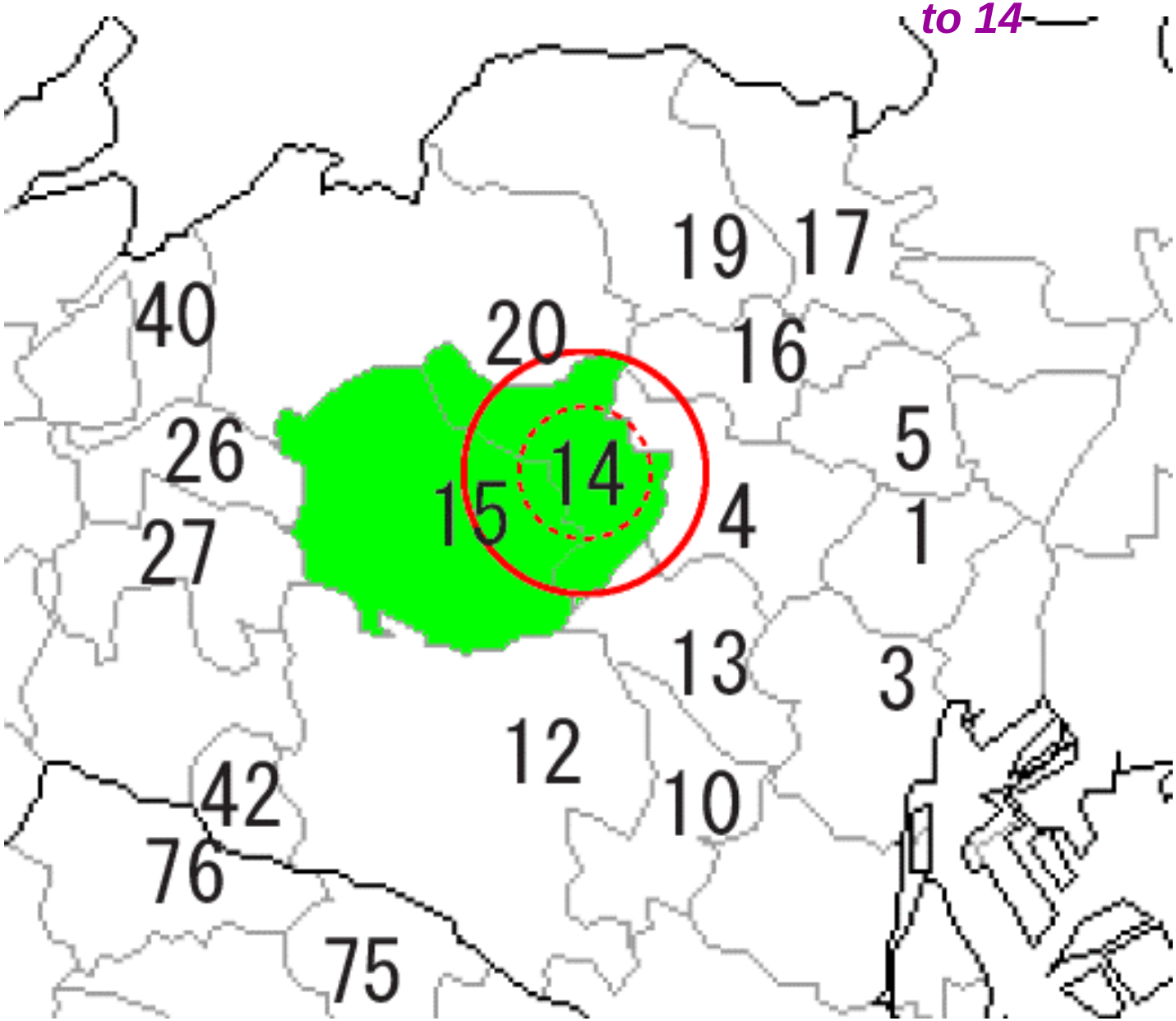
The window $Z \ (\in \Omega)$ which attains the maximum likelihood is defined as the most likely cluster (MLC).

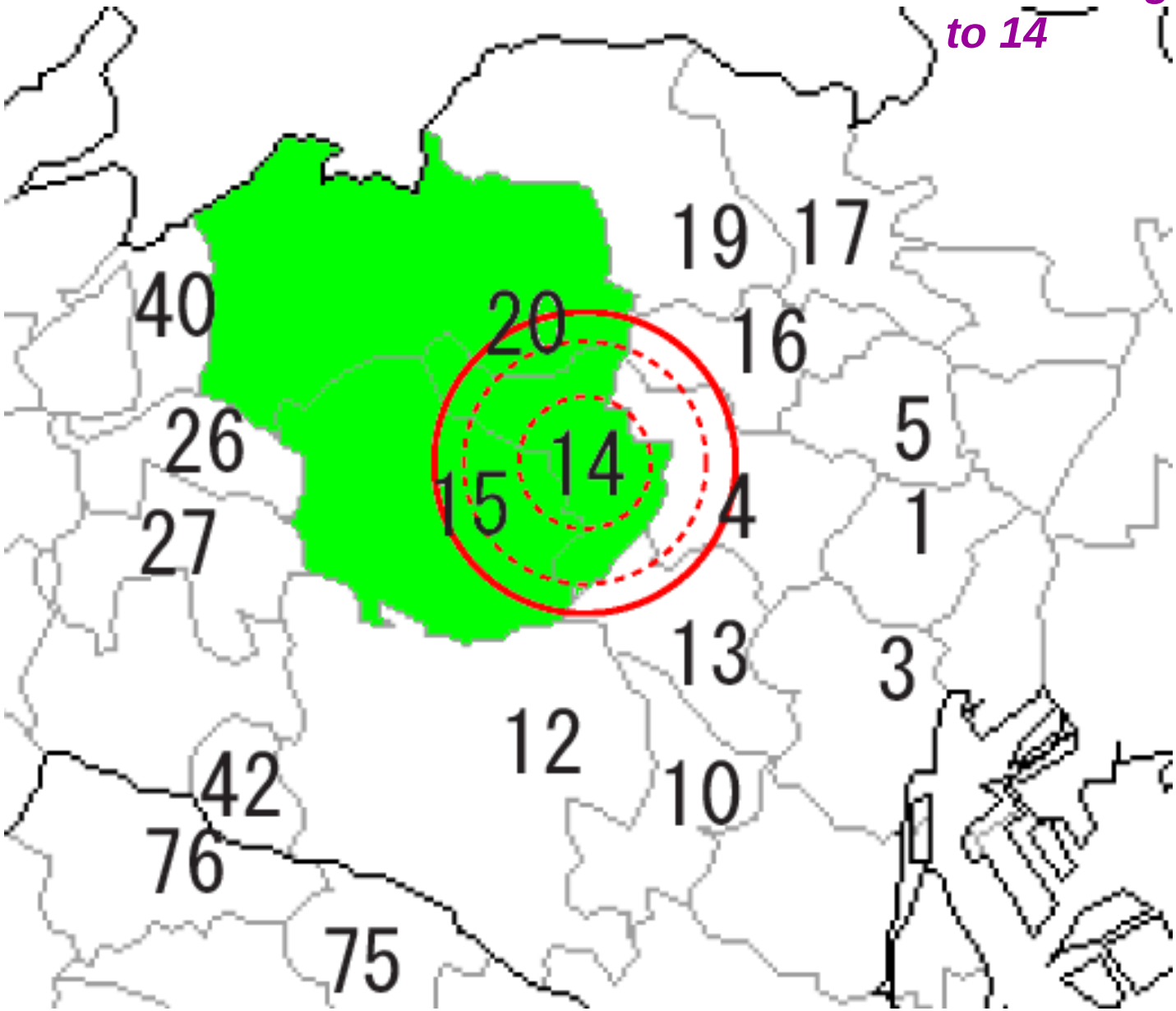
Note: The null distribution of λ is calculated by Monte Carlo samplings

Kulldorff's spatial scan

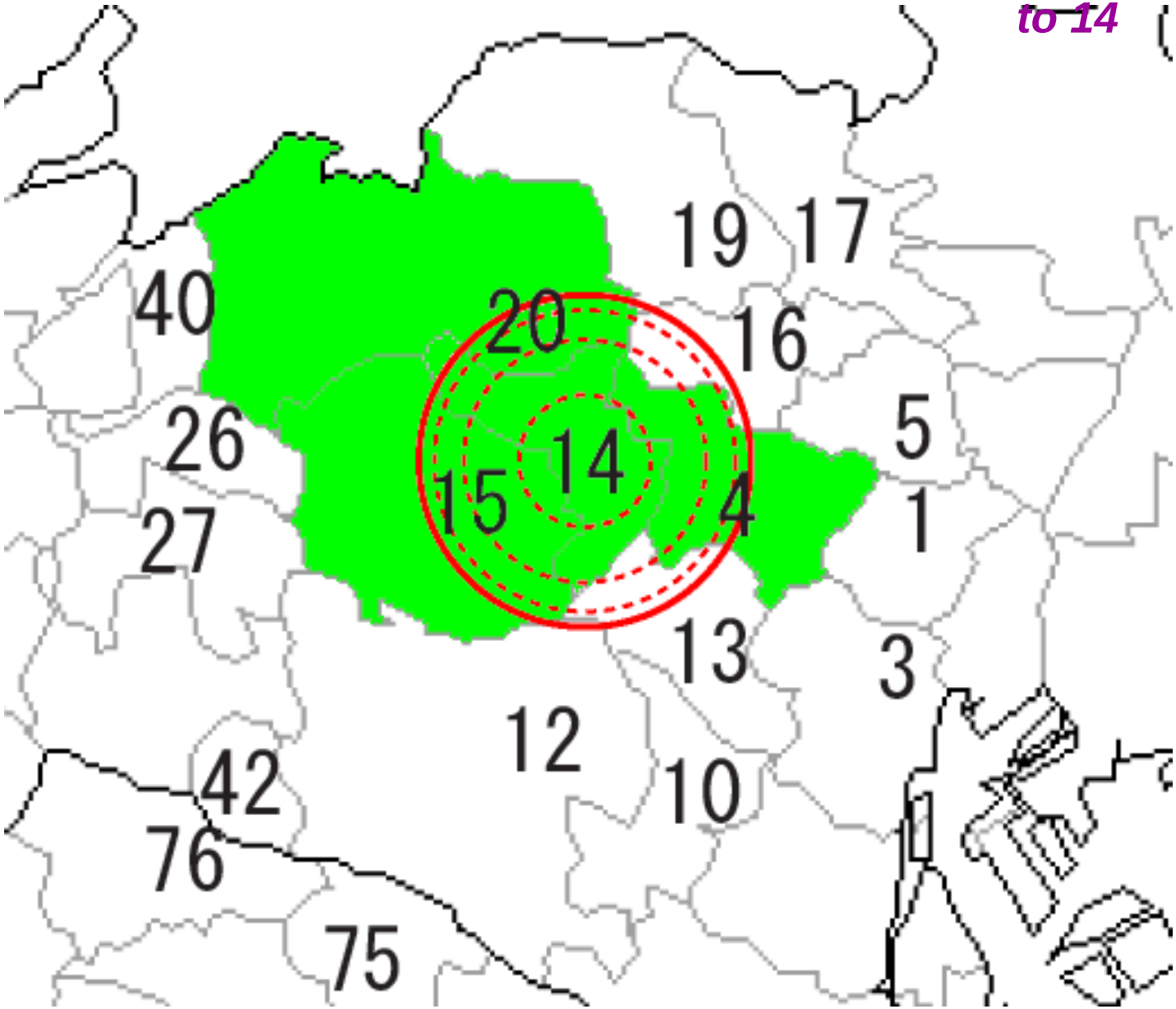
Let $K = 15$.

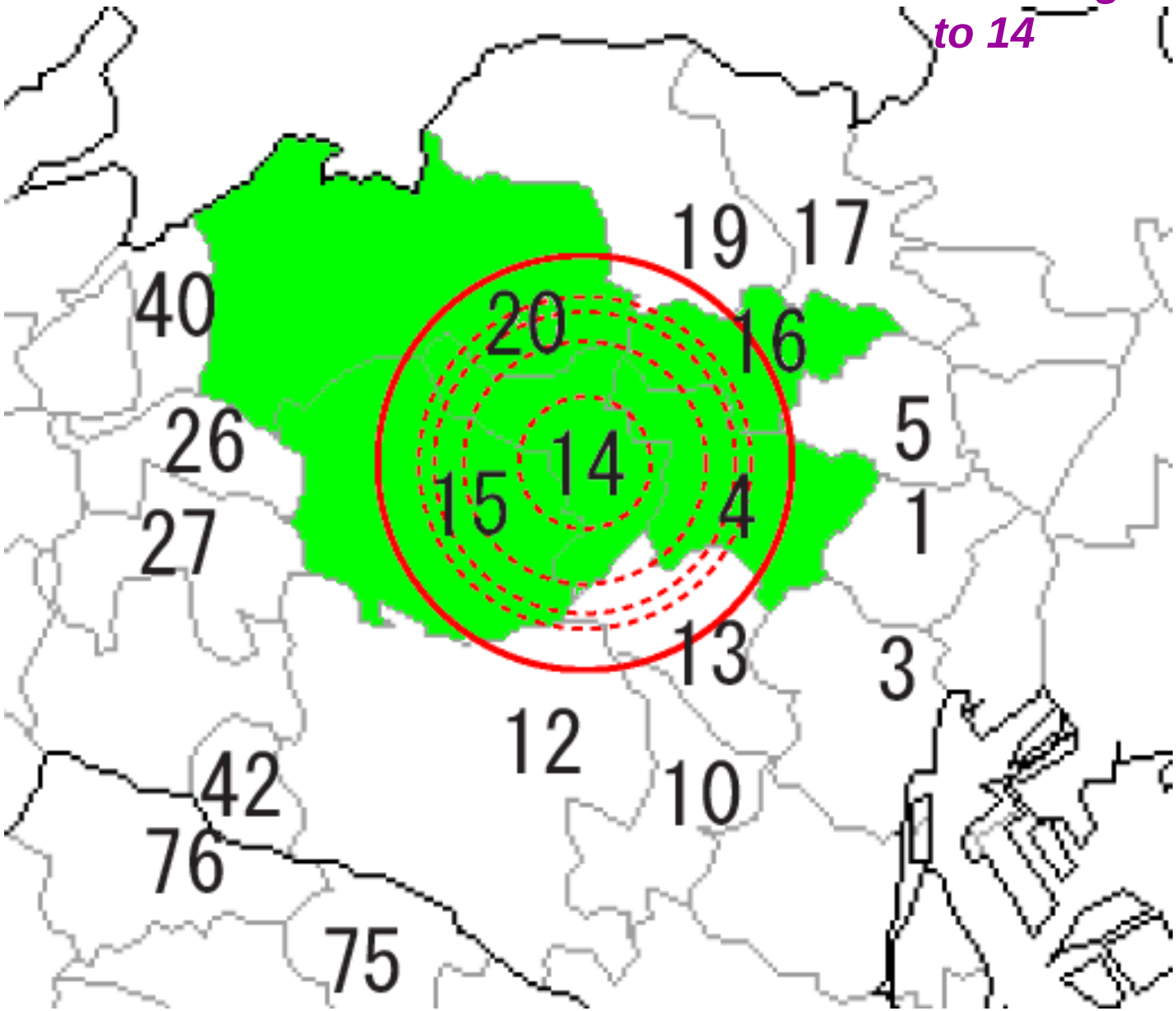


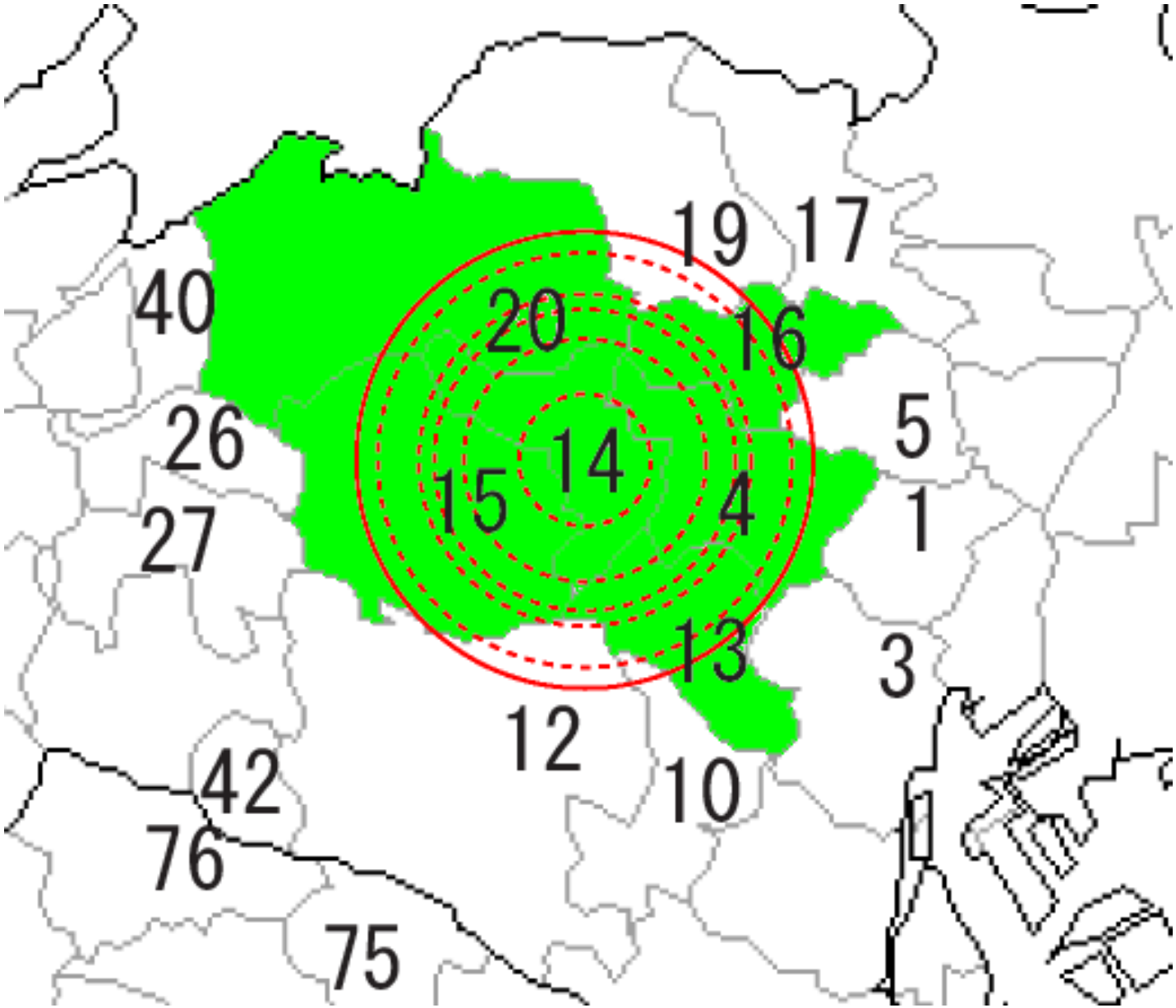

15: 1st-Nearest neigbour to 14
19
17
40
20
16
26
14
5
15
4
1
27
13
3
12
42
10
76
75

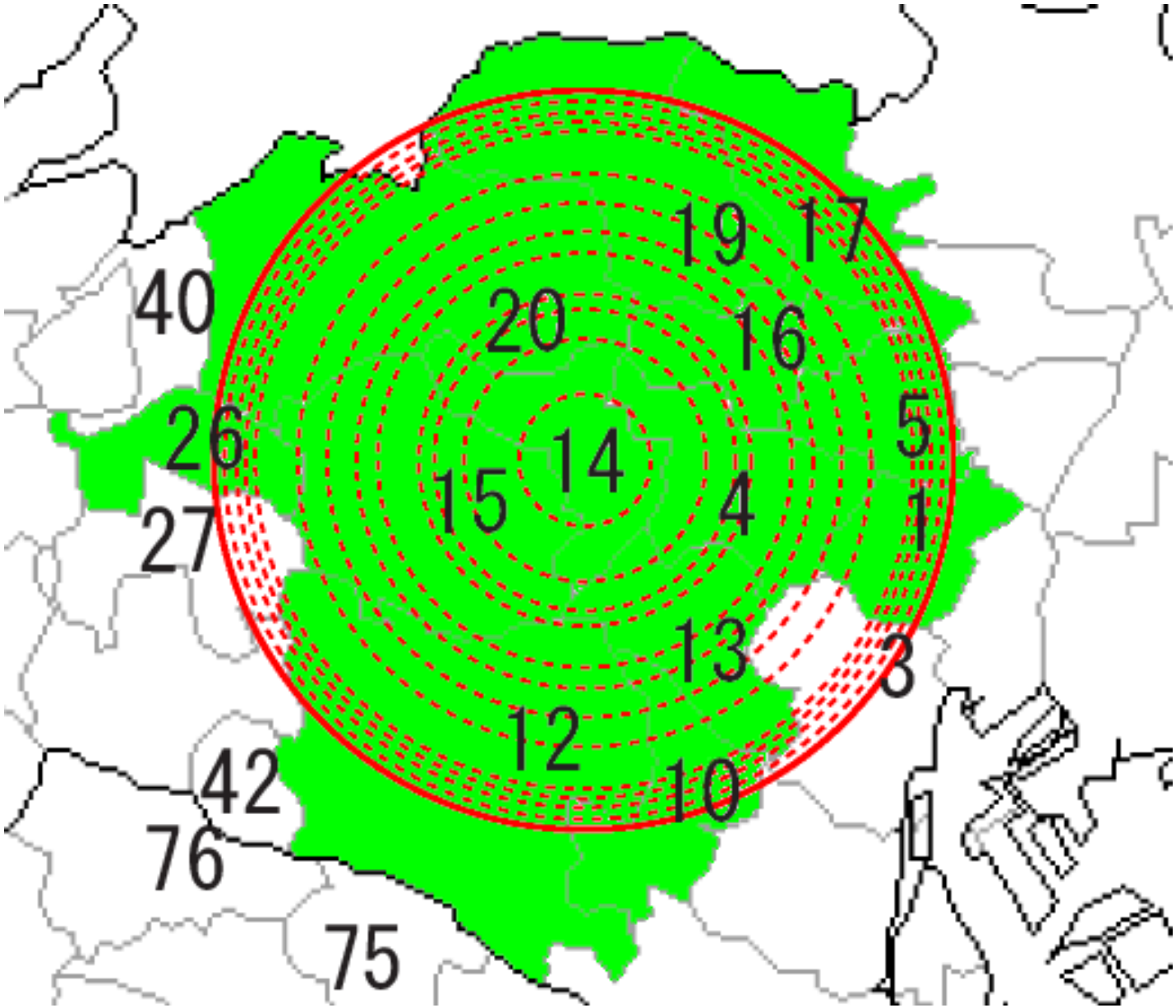

20: 2nd-Nearest neigbour to 14
19
17
40
20
16
26
14
5
15
4
1
27
13
3
12
42
10
76
75

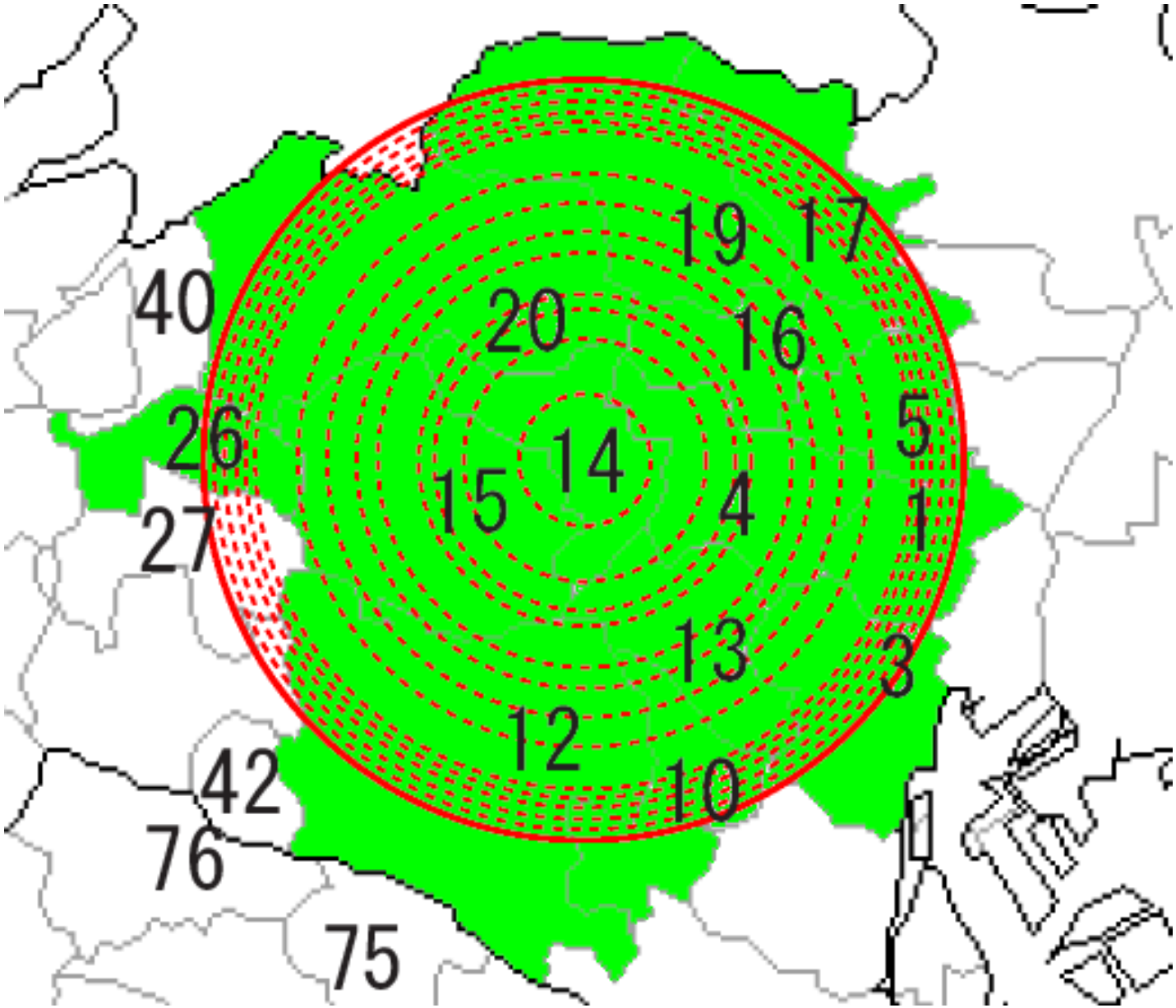
4: 3rd-Nearest neigbour to 14

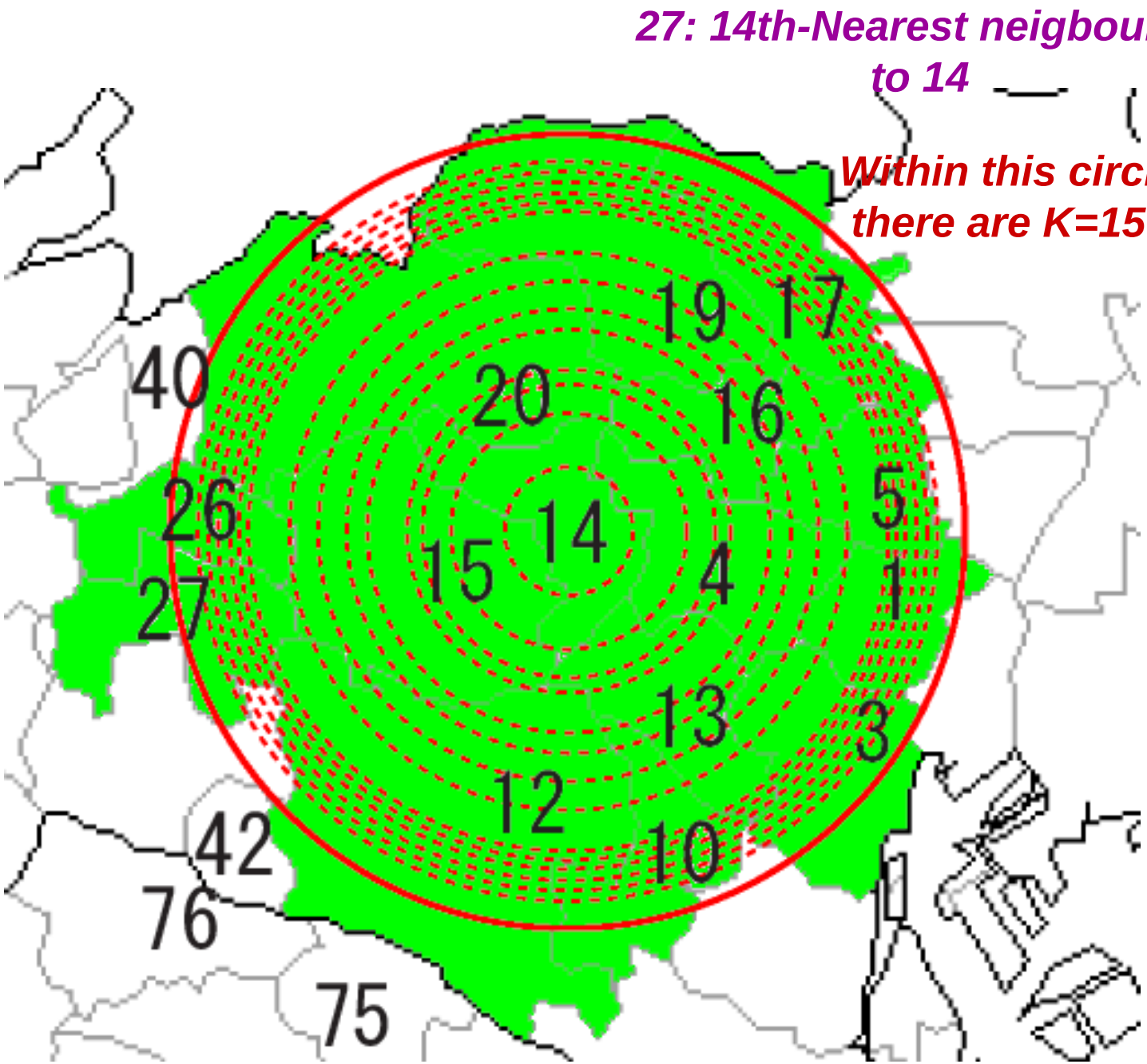



16: 4th-Nearest neigbour to 14
19
17
40
20
16
26
5
15
14
4
1
27
13
3
12
42
10
76
75


19
17
40
20
16
26
5
14
15
4
27
1
13
3
12
42
10
76
75


19
17
40
20
16
26
5
14
15
4
27
1
13
3
12
42
10
76
75


19
17
40
20
16
26
5
14
15
4
27
1
13
3
12
42
10
76
75


27: 14th-Nearest neigbour
to 14
Within this circle,
there are K=15 regions
19
17
40
20
16
26
5
14
15
4
27
1
13
3
12
10
42
76
75

Visual Disease Mapping System の開発

日本全国疾病地図

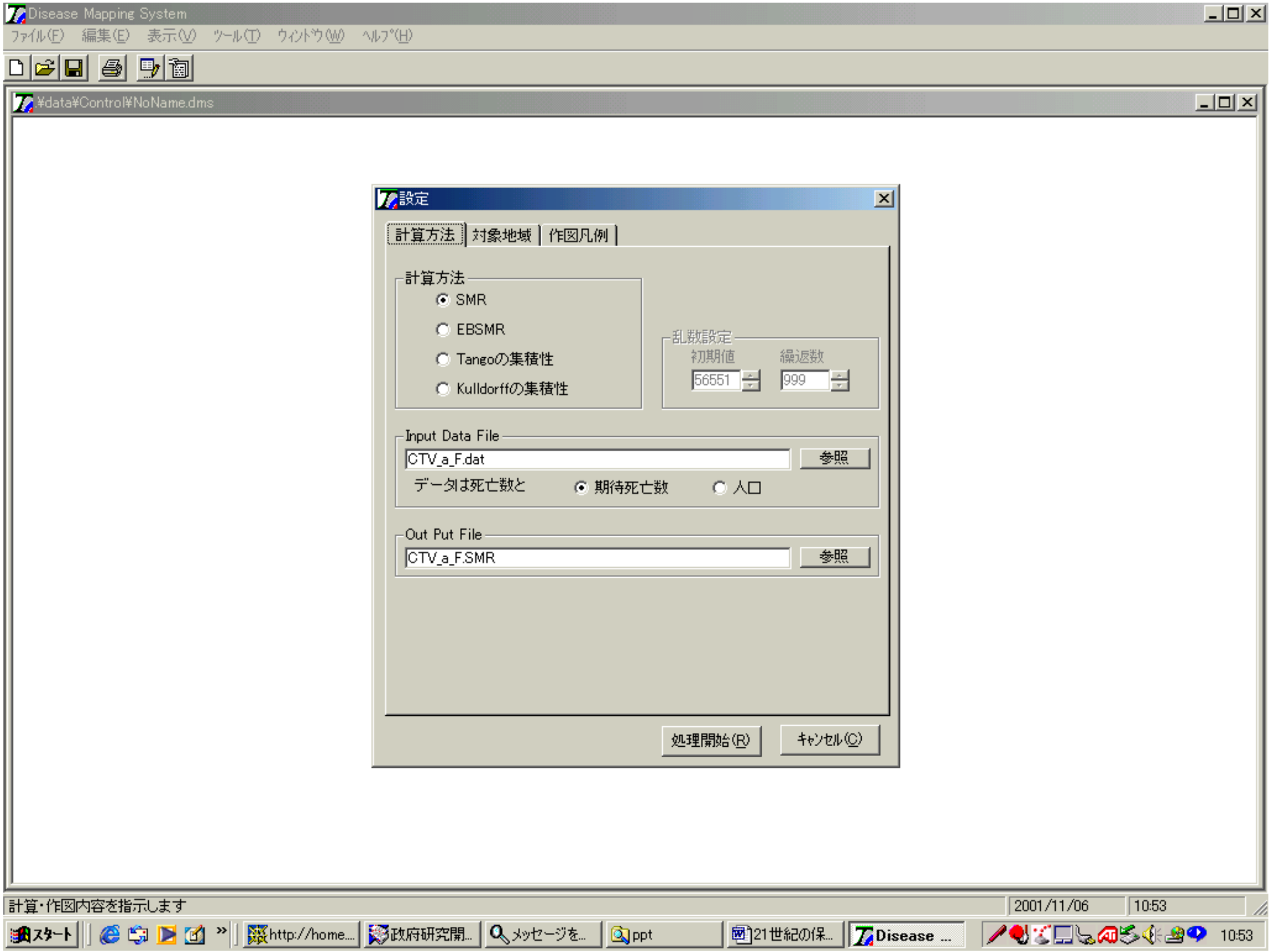
- **都道府県別**
- **二次医療圏別**
- **市区町村別**

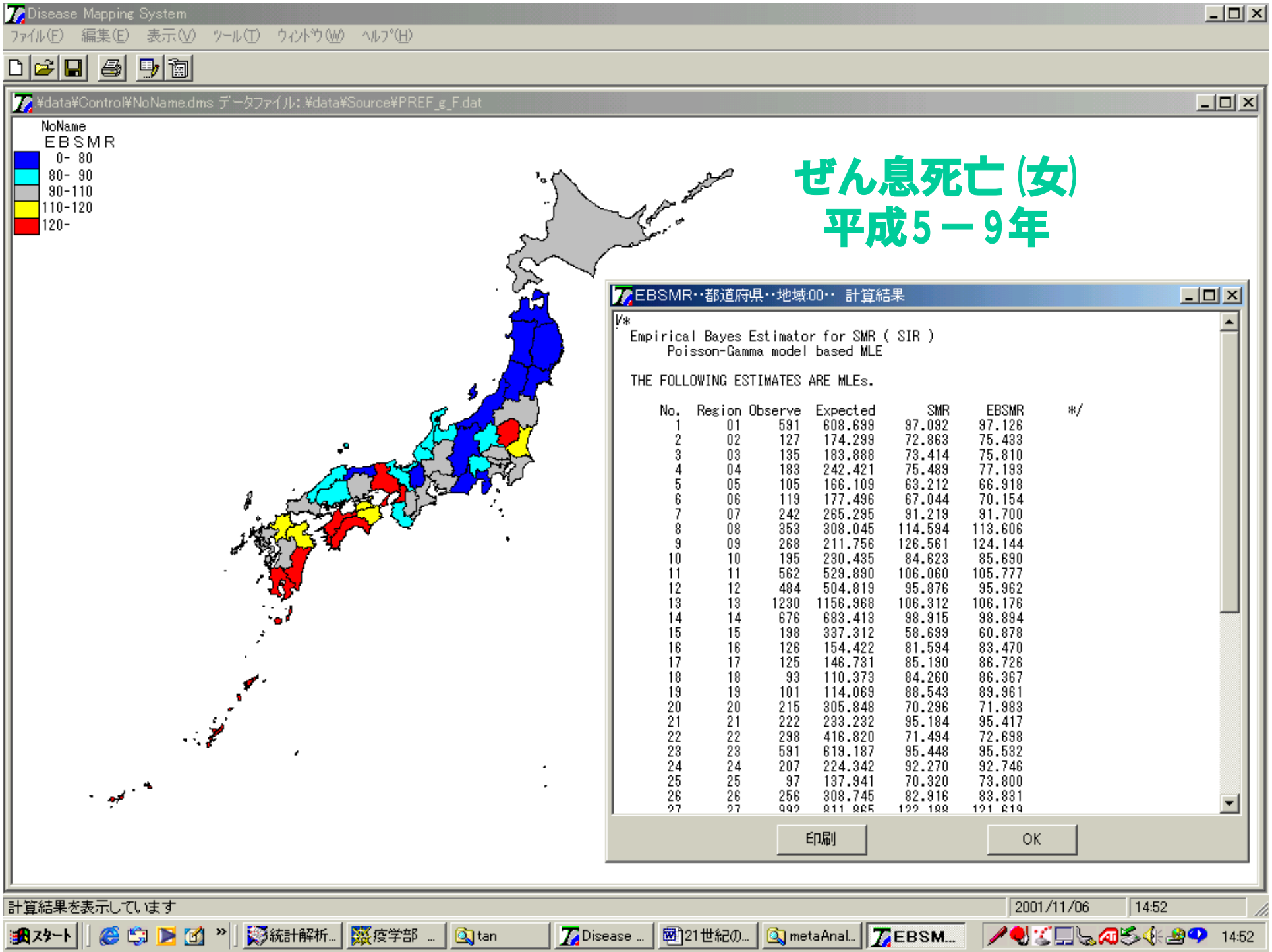
a. SMR

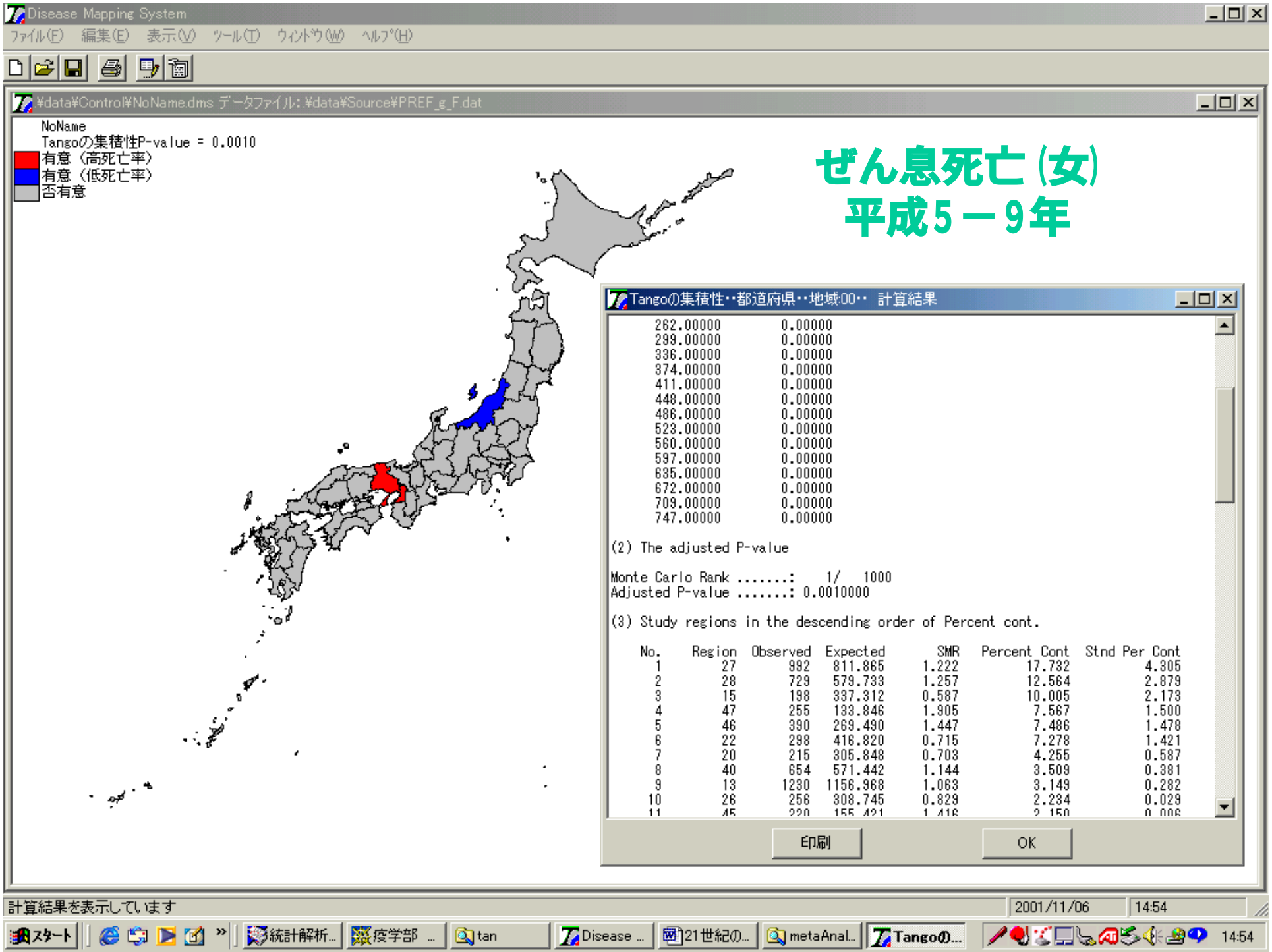
b. Empirical Bayes SMR

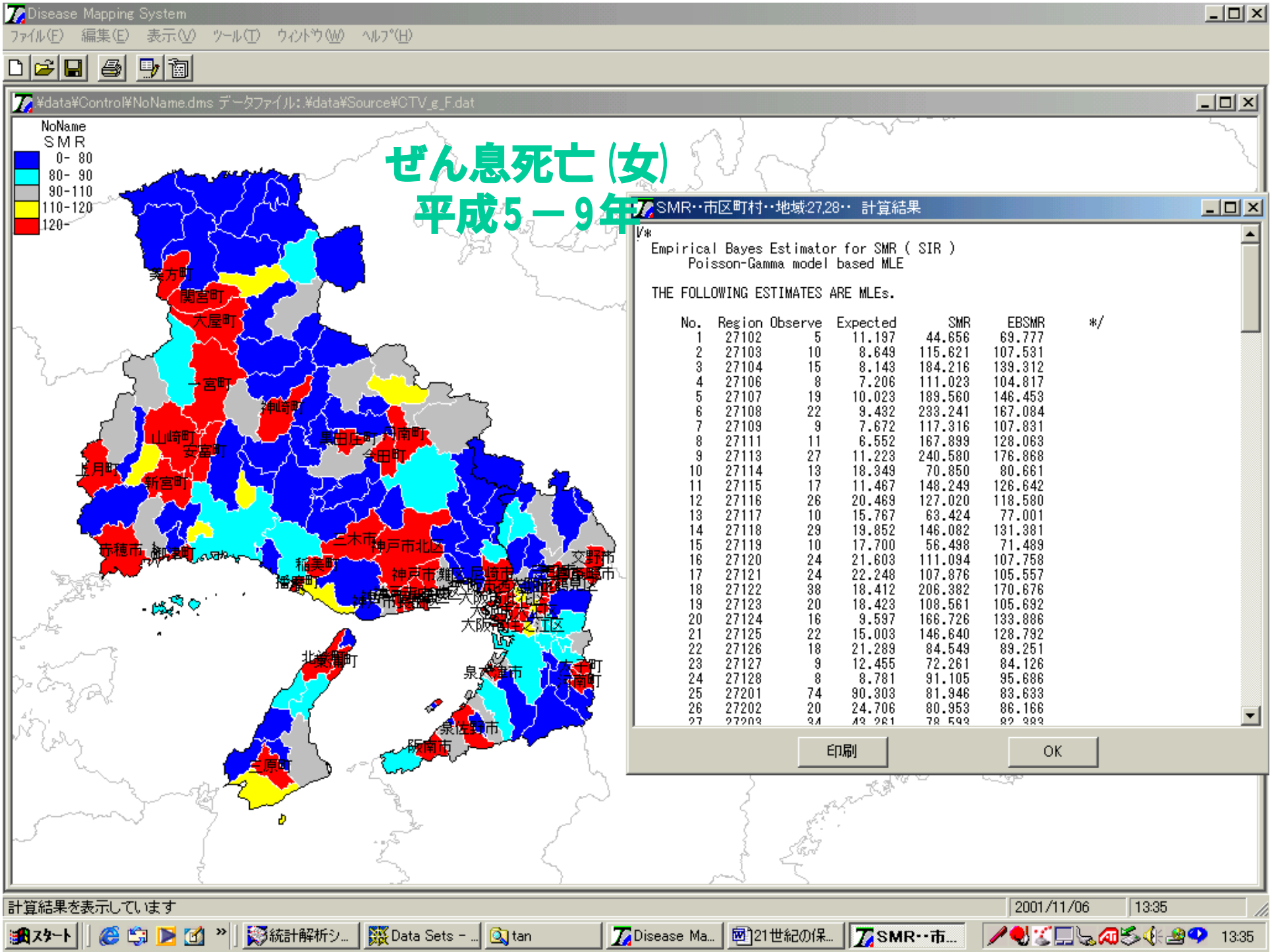
疾病集積性の検定

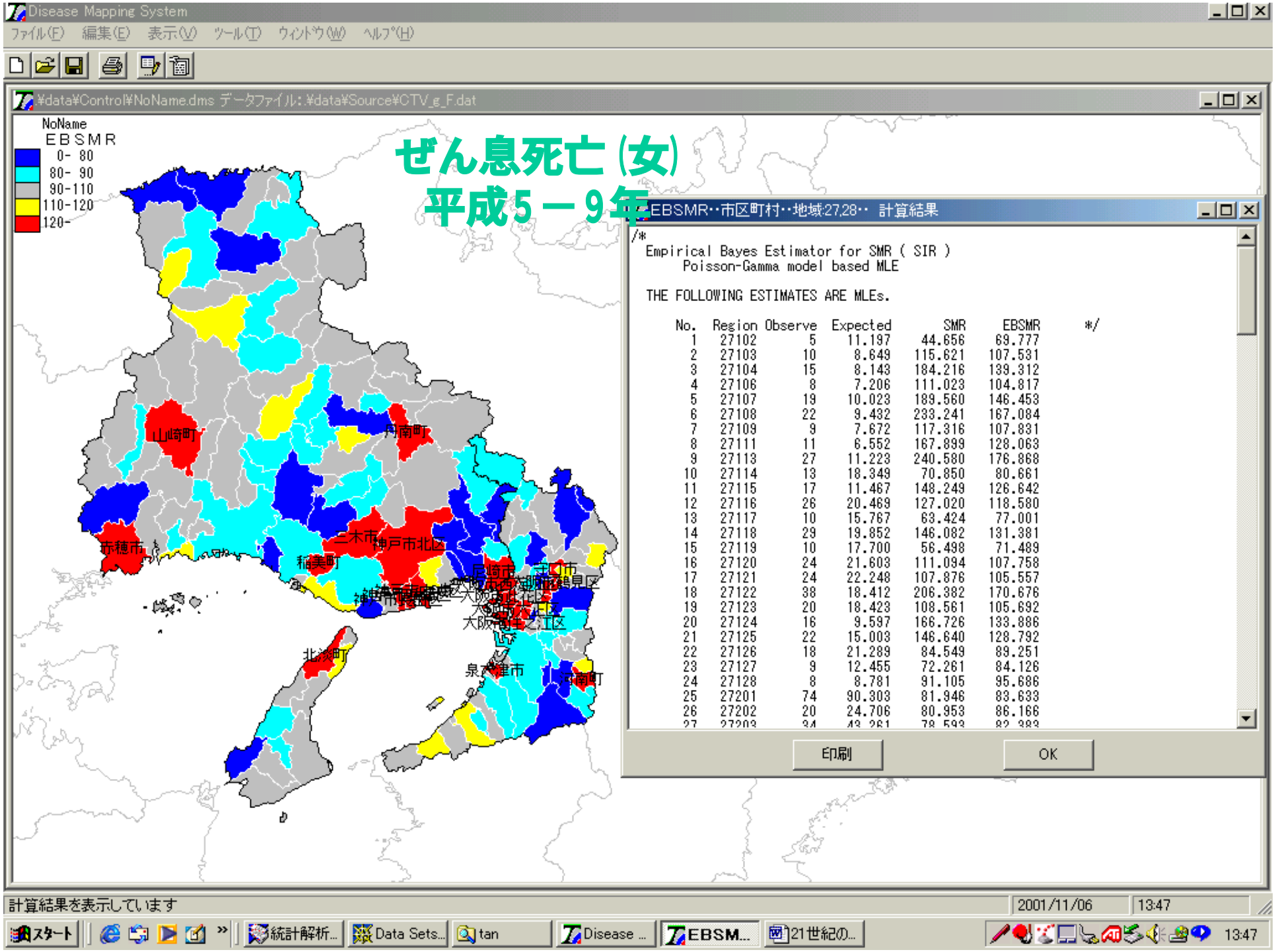
- **Kulldorff ’s Spatial Scan Test**
- **Tango’s Maximized Excess Events Test**

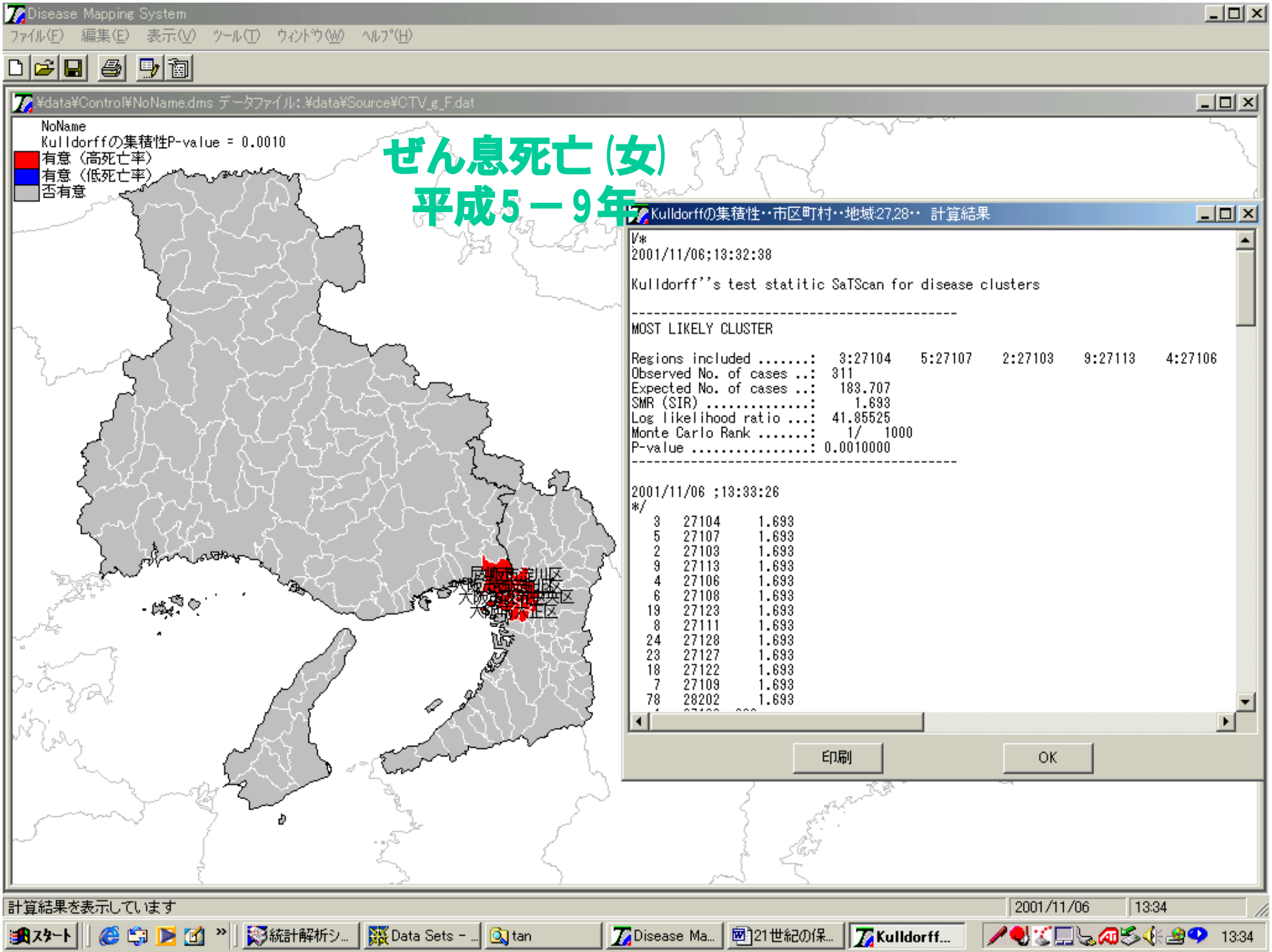

Disease Mapping System
ファイル(F) 編集(E) 表示(V) ツール(T) ウィンドウ(W) ヘルプ(H)
¥data¥Control¥NoName.dms
設定
計算方法
対象地域
作図凡例
計算方法
SMR
EBSMR
Tangoの集積性
Kulldorffの集積性
乱数設定
初期値
56551
繰返数
999
Input Data File
CTV_a_F.dat
参照
データは死亡数と
期待死亡数
人口
Out Put File
CTV_a_F.SMR
参照
処理開始(R)
キャンセル(C)
計算・作図内容を指示します
2001/11/06
10:53
スタート
http://home...
政府研究開...
メッセージを...
ppt
21世紀の保...
Disease ...
10:53

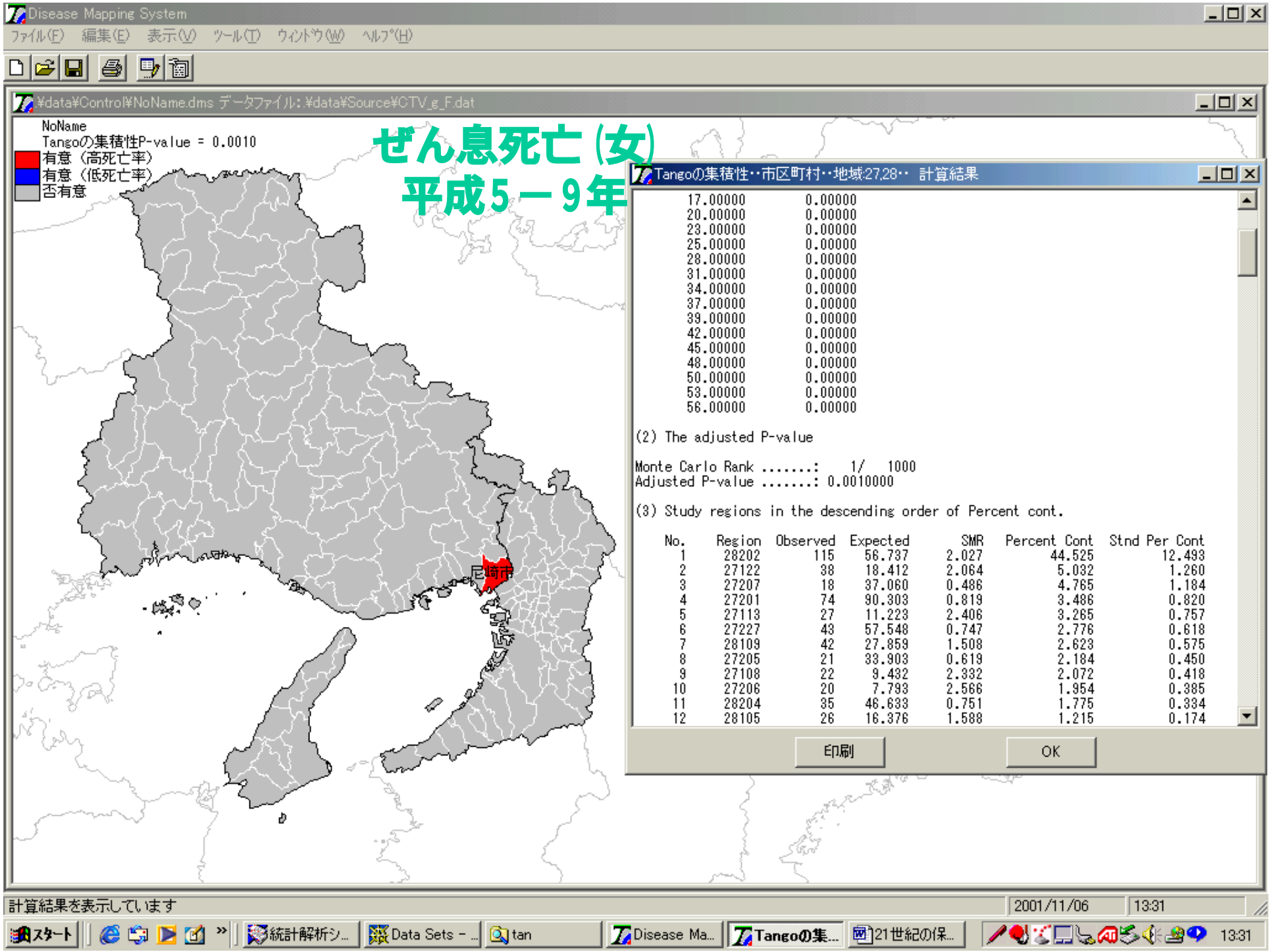

Disease Mapping System
ファイル(F) 編集(E) 表示(V) ツール(T) ウィンドウ(W) ヘルプ(H)
¥data¥Control¥NoName.dms データファイル:¥data¥Source¥PREF_g_F.dat
NoName
EBSMR
0- 80
80- 90
90-110
110-120
120-
ぜん息死亡(女)
平成5－9年
EBSMR--都道府県--地域:00-- 計算結果
/*
Empirical Bayes Estimator for SMR (SIR)
Poisson-Gamma model based MLE
THE FOLLOWING ESTIMATES ARE MLEs.
No. Region Observe Expected SMR EBSMR */
1 01 591 608.699 97.092 97.126
2 02 127 174.299 72.863 75.433
3 03 135 183.888 73.414 75.810
4 04 183 242.421 75.489 77.193
5 05 105 166.109 63.212 66.918
6 06 119 177.496 67.044 70.154
7 07 242 265.295 91.219 91.700
8 08 353 308.045 114.594 113.606
9 09 268 211.756 126.561 124.144
10 10 195 230.435 84.623 85.690
11 11 562 529.890 106.060 105.777
12 12 484 504.819 95.876 95.962
13 13 1230 1156.968 106.312 106.176
14 14 676 683.413 98.915 98.894
15 15 198 337.312 58.699 60.878
16 16 126 154.422 81.594 83.470
17 17 125 146.731 85.190 86.726
18 18 93 110.373 84.260 86.367
19 19 101 114.069 88.543 89.961
20 20 215 305.848 70.296 71.983
21 21 222 233.232 95.184 95.417
22 22 298 416.820 71.494 72.698
23 23 591 619.187 95.448 95.532
24 24 207 224.342 92.270 92.746
25 25 97 137.941 70.320 73.800
26 26 256 308.745 82.916 83.831
印刷
OK
計算結果を表示しています
2001/11/06
14:52
スタート
統計解析...
疫学部 ...
tan
Disease ...
21世紀の...
metaAnal...
EBSM...
14:52


Disease Mapping System
ファイル(F) 編集(E) 表示(V) ツール(T) ウィンドウ(W) ヘルプ(H)
¥data¥Control¥NoName.dms データファイル:¥data¥Source¥PREF_g_F.dat
NoName
Tangoの集積性P-value = 0.0010
有意（高死亡率）
有意（低死亡率）
否有意
ぜん息死亡(女)
平成5－9年
Tangoの集積性‥都道府県‥地域:00‥ 計算結果
262.00000 0.00000
299.00000 0.00000
336.00000 0.00000
374.00000 0.00000
411.00000 0.00000
448.00000 0.00000
486.00000 0.00000
523.00000 0.00000
560.00000 0.00000
597.00000 0.00000
635.00000 0.00000
672.00000 0.00000
709.00000 0.00000
747.00000 0.00000
(2) The adjusted P-value
Monte Carlo Rank .....: 1/ 1000
Adjusted P-value .....: 0.0010000
(3) Study regions in the descending order of Percent cont.
No. Region Observed Expected SMR Percent Cont Stnd Per Cont
1 27 992 811.865 1.222 17.732 4.305
2 28 729 579.733 1.257 12.564 2.879
3 15 198 337.312 0.587 10.005 2.173
4 47 255 133.846 1.905 7.567 1.500
5 46 390 269.490 1.447 7.486 1.478
6 22 298 416.820 0.715 7.278 1.421
7 20 215 305.848 0.703 4.255 0.587
8 40 654 571.442 1.144 3.509 0.381
9 13 1230 1156.968 1.063 3.149 0.282
10 26 256 308.745 0.829 2.234 0.029
印刷
OK
計算結果を表示しています
2001/11/06
14:54
スタート
統計解析...
疫学部 ...
tan
Disease ...
21世紀の...
metaAnal...
Tangoの...
14:54


Disease Mapping System
ファイル(F) 編集(E) 表示(V) ツール(T) ウィンドウ(W) ヘルプ(H)
¥data¥Control¥NoName.dms データファイル:¥data¥Source¥CTV_g_F.dat
NoName
SMR
0- 80
80- 90
90-110
110-120
120-
ぜん息死亡(女)
平成5－9年
SMR‥市区町村‥地域27,28‥ 計算結果
/*
Empirical Bayes Estimator for SMR (SIR)
Poisson-Gamma model based MLE
THE FOLLOWING ESTIMATES ARE MLEs.
No. Region Observe Expected SMR EBSMR */
1 27102 5 11.197 44.656 69.777
2 27103 10 8.649 115.621 107.531
3 27104 15 8.143 184.216 139.312
4 27106 8 7.206 111.023 104.817
5 27107 19 10.023 189.560 146.453
6 27108 22 9.432 233.241 167.084
7 27109 9 7.672 117.316 107.831
8 27111 11 6.552 167.899 128.063
9 27113 27 11.223 240.580 176.868
10 27114 13 18.349 70.850 80.661
11 27115 17 11.467 148.249 126.642
12 27116 26 20.469 127.020 118.580
13 27117 10 15.767 63.424 77.001
14 27118 29 19.852 146.082 131.381
15 27119 10 17.700 56.498 71.489
16 27120 24 21.603 111.094 107.758
17 27121 24 22.248 107.876 105.557
18 27122 38 18.412 206.382 170.676
19 27123 20 18.423 108.561 105.692
20 27124 16 9.597 166.726 133.886
21 27125 22 15.003 146.640 128.792
22 27126 18 21.289 84.549 89.251
23 27127 9 12.455 72.261 84.126
24 27128 8 8.781 91.105 95.686
25 27201 74 90.303 81.946 83.633
26 27202 20 24.706 80.953 86.166
印刷
OK
計算結果を表示しています
2001/11/06
13:35
スタート
統計解析シ...
Data Sets - ...
tan
Disease Ma...
21世紀の保...
SMR‥市...
13:35


Disease Mapping System
ファイル(F) 編集(E) 表示(V) ツール(T) ウィンドウ(W) ヘルプ(H)
¥data¥Control¥NoName.dms データファイル: ¥data¥Source¥CTV_g_F.dat
NoName
EBSMR
0- 80
80- 90
90-110
110-120
120-
ぜん息死亡(女)
平成5－9年
山崎町
河南町
赤穂市
二木市
神戸市北区
稲美町
尼崎市
北淡町
泉大津市
河南町
EBSMR··市区町村··地域:27,28·· 計算結果
/*
Empirical Bayes Estimator for SMR (SIR)
Poisson-Gamma model based MLE
THE FOLLOWING ESTIMATES ARE MLEs.
No. Region Observe Expected SMR EBSMR */
1 27102 5 11.197 44.656 69.777
2 27103 10 8.649 115.621 107.531
3 27104 15 8.143 184.216 139.312
4 27106 8 7.206 111.023 104.817
5 27107 19 10.023 189.560 146.453
6 27108 22 9.432 233.241 167.084
7 27109 9 7.672 117.316 107.831
8 27111 11 6.552 167.899 128.063
9 27113 27 11.223 240.580 176.868
10 27114 13 18.349 70.850 80.661
11 27115 17 11.467 148.249 126.642
12 27116 26 20.469 127.020 118.580
13 27117 10 15.767 63.424 77.001
14 27118 29 19.852 146.082 131.381
15 27119 10 17.700 56.498 71.489
16 27120 24 21.603 111.094 107.758
17 27121 24 22.248 107.876 105.557
18 27122 38 18.412 206.382 170.676
19 27123 20 18.423 108.561 105.692
20 27124 16 9.597 166.726 133.886
21 27125 22 15.003 146.640 128.792
22 27126 18 21.289 84.549 89.251
23 27127 9 12.455 72.261 84.126
24 27128 8 8.781 91.105 95.686
25 27201 74 90.303 81.946 83.633
26 27202 20 24.706 80.953 86.166
印刷
OK
計算結果を表示しています
2001/11/06
13:47
スタート
統計解析...
Data Sets...
tan
Disease ...
EBSM...
21世紀の...
13:47


Disease Mapping System
ファイル(F) 編集(E) 表示(V) ツール(T) ウィンドウ(W) ヘルプ(H)
¥data¥Control¥NoName.dms データファイル:¥data¥Source¥CTV_g_F.dat
NoName
Kulldorffの集積性P-value = 0.0010
有意（高死亡率）
有意（低死亡率）
否有意
ぜん息死亡(女)
平成5－9年
Kulldorffの集積性‥市区町村‥地域:27,28‥ 計算結果
/*
2001/11/06;13:32:38
Kulldorff''s test statitic SaTScan for disease clusters
MOST LIKELY CLUSTER
Regions included .....: 3:27104 5:27107 2:27103 9:27113 4:27106
Observed No. of cases ..: 311
Expected No. of cases ..: 183.707
SMR (SIR) .....: 1.693
Log likelihood ratio ...: 41.85525
Monte Carlo Rank .....: 1/ 1000
P-value .....: 0.0010000
2001/11/06 ;13:33:26
*/
3 27104 1.693
5 27107 1.693
2 27103 1.693
9 27113 1.693
4 27106 1.693
6 27108 1.693
19 27123 1.693
8 27111 1.693
24 27128 1.693
23 27127 1.693
18 27122 1.693
7 27109 1.693
78 28202 1.693
印刷
OK
計算結果を表示しています
2001/11/06
13:34
スタート
統計解析シ...
Data Sets - ...
tan
Disease Ma...
21世紀の保...
Kulldorff...
13:34


Disease Mapping System
ファイル(F) 編集(E) 表示(V) ツール(T) ウィンドウ(W) ヘルプ(H)
¥data¥Control¥NoName.dms データファイル:¥data¥Source¥CTV_g_F.dat
NoName
Tangoの集積性P-value = 0.0010
有意（高死亡率）
有意（低死亡率）
否有意
ぜん息死亡(女)
平成5－9年
Tangoの集積性‥市区町村‥地域:27,28‥ 計算結果
17.00000 0.00000
20.00000 0.00000
23.00000 0.00000
25.00000 0.00000
28.00000 0.00000
31.00000 0.00000
34.00000 0.00000
37.00000 0.00000
39.00000 0.00000
42.00000 0.00000
45.00000 0.00000
48.00000 0.00000
50.00000 0.00000
53.00000 0.00000
56.00000 0.00000
(2) The adjusted P-value
Monte Carlo Rank .....: 1/ 1000
Adjusted P-value .....: 0.0010000
(3) Study regions in the descending order of Percent cont.
No. Region Observed Expected SMR Percent Cont Stnd Per Cont
1 28202 115 56.737 2.027 44.525 12.493
2 27122 38 18.412 2.064 5.032 1.260
3 27207 18 37.060 0.486 4.765 1.184
4 27201 74 90.303 0.819 3.486 0.820
5 27113 27 11.223 2.406 3.265 0.757
6 27227 43 57.548 0.747 2.776 0.618
7 28109 42 27.859 1.508 2.623 0.575
8 27205 21 33.903 0.619 2.184 0.450
9 27108 22 9.432 2.332 2.072 0.418
10 27206 20 7.793 2.566 1.954 0.385
11 28204 35 46.633 0.751 1.775 0.334
12 28105 26 16.376 1.588 1.215 0.174
印刷
OK
尼崎市
計算結果を表示しています
2001/11/06
13:31
スタート
統計解析シ...
Data Sets - ...
tan
Disease Ma...
Tangoの集...
21世紀の保...
13:31

S-PLUSの新しいGIS機能として

- **標準の関数**として市区町村別の地図
- 距離（緯度、経度）が計算できる
- 隣接しているか否かの情報がわかる
  （問題）市区町村合併が起こっている！
  ・・・