

環境と発がんの関係 地理情報解析システムを使って

Relationship between the environment and cancer development

Trial analysis using geographical information system

三上 春夫*

1. はじめに

地域がん登録はがん罹患の地域的年代的特性を明らかにするものであるが、これまで罹患率や生存率といった指標は市町村などの行政区分を単位としてしか集計不能であり解析には自ずと限界があった。また近年がん罹患と環境要因の関連が明らかになるにつれて、環境モニタリングデータや人工衛星観測データなど行政区画とは異なった区分のデータを利用する解析の必要性も出てきた。そこでPC上での利用が急速に進歩してきた地理情報システム (GIS; Geographical Information System) を採り入れた地理疫学情報解析ソフトウェアの開発を行ったので、その概要と成果を紹介する。

2. 開発環境

開発にはIBM-PCAT互換機とMicrosoft社のWindows 98/NT上で動作するMapInfo社製地理情報システムMapInfo(R) Ver5.0およびMapInfo上の開発環境MapBasicを使用した。地図等のデータでは、国勢調査人口データ、メ

GIS: MapInfo Ver5.0, MapBasic (MapInfo社) 等高線ライタ for MapInfo(リンク情報システム) OS: Windows 98/NT/2000/XP CPU: MMX Pentium 300MHz + RAM: 128MB + HDD: 500MB + Video: SVGA (800*600) + Data: 住宅地図、道路地図、メッシュ国勢調査地図 Developing Partner: (株)プラムシックス(千葉市)

図1. 開発環境 System requirement

ッシュ地図データ、道路地図データ、町丁目地図データ、住宅地図データなどがアルプス社および三井造船システム技研よりMapInfo専用データとして提供されているが、最近国土地理院より住宅レベルでマッピング可能な地図データの提供が開始された。

3. 解析機能

基本的な処理の流れに沿って解析機能を説明する。

解析対象のデータファイルはCSVファイルで用意する。ID、住所、診断コード、性別、年齢区分、発症年代などのデータ構成は初期設定ファイルに記述する。ファイルが読み込まれた後、住宅地図上に患者住所をマッピングする「住所ジオコード」を実行する。通常都市部では80%以上のマッピング成功率が得られる。マッピングされたデータを、性別や診断区分(ICDコードなど)や年齢階級などにより選択して地図上に表示する。この機能により各がんの分布特性などを目視確認できる。

続いて関心領域(バッファ)を生成して領域内の患者を抽出する。バッファには円、扇形、折れ線(道路や送電線など)、不定形、メッシュ(JIS 1kmメッシュ)を基準とした領域が設定可能で、各バッファ間の和・積・補集合などの演算結果を新たなバッファに指定できる。不定形バッファでは大気汚染など観測点の住所と観測データによりスプライン補間を用いてコンターマップ(等高線図)を描いた後、等高

*千葉県がんセンター疫学研究部

〒260-0801 千葉市中央区仁戸名町 666-2

線区間の患者抽出が可能である。上記のバッファで抽出した患者情報を出力する際、円・扇形・折れ線バッファでは基点(基線)からの距離を計測し出力する。これにより「道路から50m圏内」などのデータを集計できる。また患者情報出力の際、患者の属するメッシュ区画の人口を付加して出力することにより簡易の人口密度補正が可能である。

MapInfo での基本的な印刷範囲は表示画面単位に限られるので(スケールは任意) 現在画面を中央とし周囲8面を加えた9面連続印刷機能を追加した。これにより広範囲の一括地図印刷が可能である。

さらに、自動化されていないが、国勢調査に基づいたメッシュ内年齢階級別人口を面積比で案分してバッファの年齢階級別人口を推計することにより、環境要因によるがん罹患のリスク評価を試みている。

4. 解析の成果と課題

これまでに 1) 胃がんや大腸がん、肺がんなどがんの種類によって道路からの距離により分布のパターンが異なること(国際がん登録学会 2001)、2) 送電線近傍に比較的若年のリンパ腫が集積すること(表 1. 国際がん登録学会 2002)、3) 肺がんが道路 50m 圏に集積する傾向があり(図 1. 図 2. 日本癌学会 2002) 幹線道路では道路 500m 圏に比べ 1.8~2.0 倍のリスクを有すること(表 2. 厚生労働省地域がん登録研究班 2003)などを報告した。

実際には環境要因への曝露からがん罹患までには潜伏期間があり、過去のデータに基づいた曝露状況の確認や関心領域の設定が必要である。また類似の解析事例が少なく、今後解析の精度を上げるとともに、他のがんや他地域で同様の試算を行い解析方法の確立に努める必要がある。

表 1. リンパ腫の特性 Properties of Lymphoma Cases

Gender			
Distance	Female	Male	Total
<100m	5	8	13
100-500 m	26	27	53
Total	31	35	66

Age Class										
Distance	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-	Total
<100m	2	2		3	4		2			13
100-500m	1	2	4	5	13	11	13	3	1	53
Total	3	4	4	8	17	11	15	3	1	66

ICD-10						
Distance	C81	C82	C83	C84	C85	Total
<100m	1		6		6	13
100-500m	4	1	11	1	36	53
Total	5	1	17	1	42	66

> > 100m 圏で若年者の比率が有意に高い

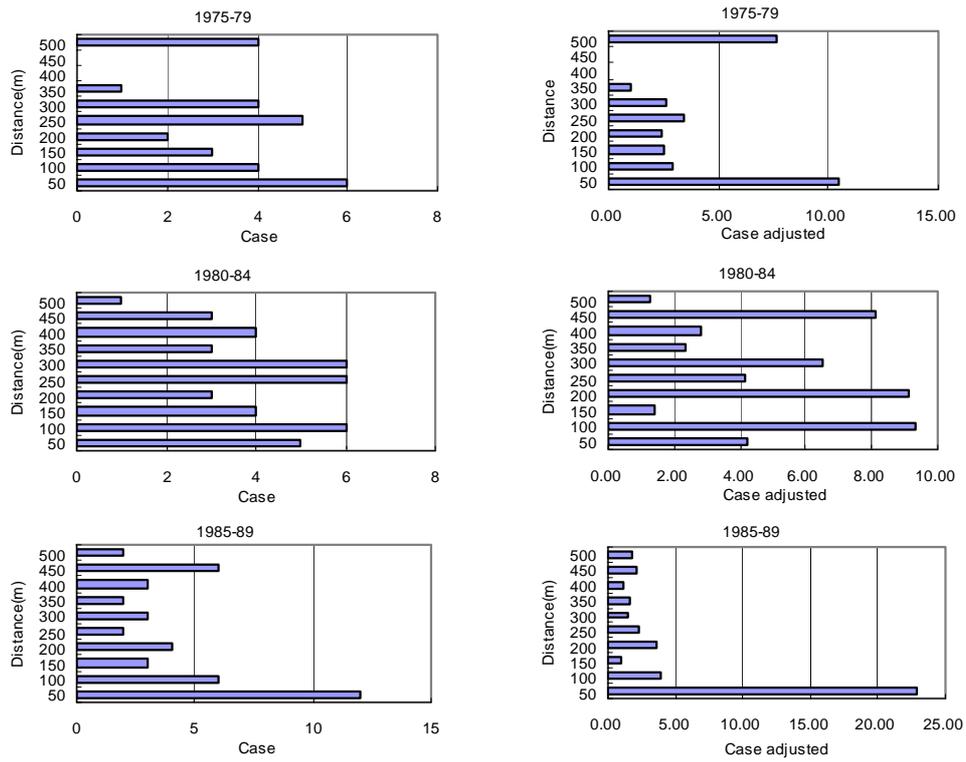


図 1. 肺がんの罹患年代別距離区分別度数分布<1>

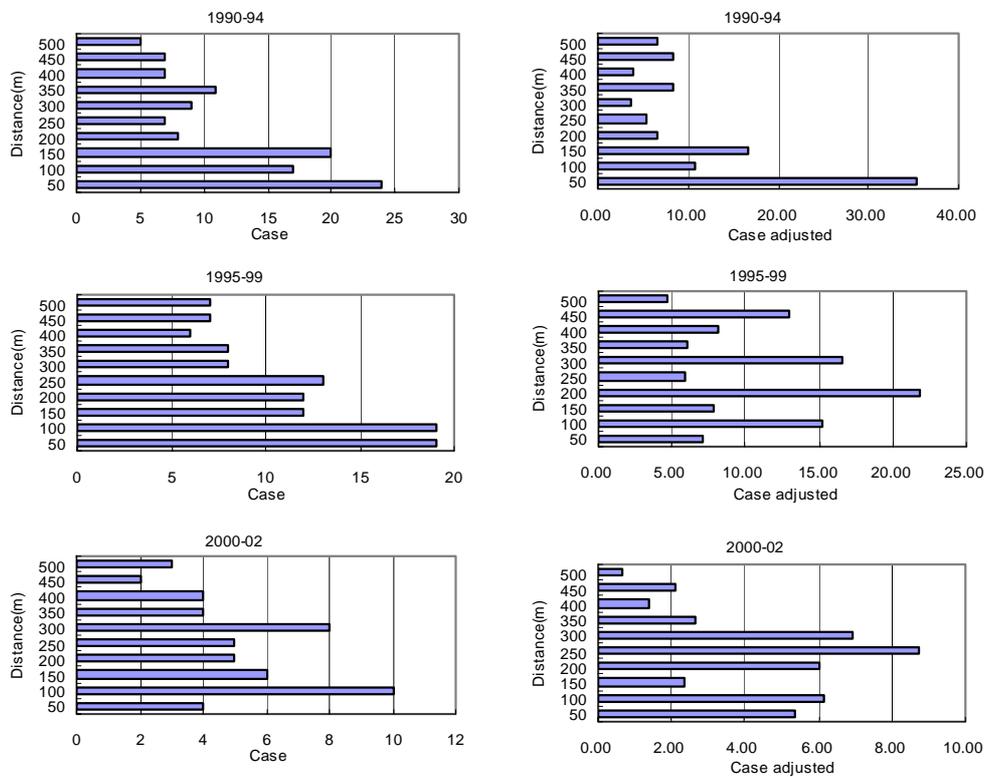


図 2. 肺がんの罹患年代別距離区分別度数分布<2>

表 2. 肺がん和大腸がんの幹線道路 50m 圏罹患リスク推計

1990-94	肺			大腸		
	男	女	計	男	女	計
観察罹患数 O	17	7	24	11	0	11
期待罹患数 E	9.64	3.50	13.14	8.27	7.98	16.25
O/E 比	1.76	2.00	1.83	1.33	0.00	0.68
カイ 2 乗検定	4.877	2.571	8.163	0.600	7.008	1.388
ポアソン検定	0.037	0.065	0.005	0.212		0.069

>> 肺がんの男と男女計で幹線道路 50m 圏に有意のリスクを認めるが、同地域の大腸がんではリスクを認めない。