

疫学統計ソフト Epi Info の使い方 応用編

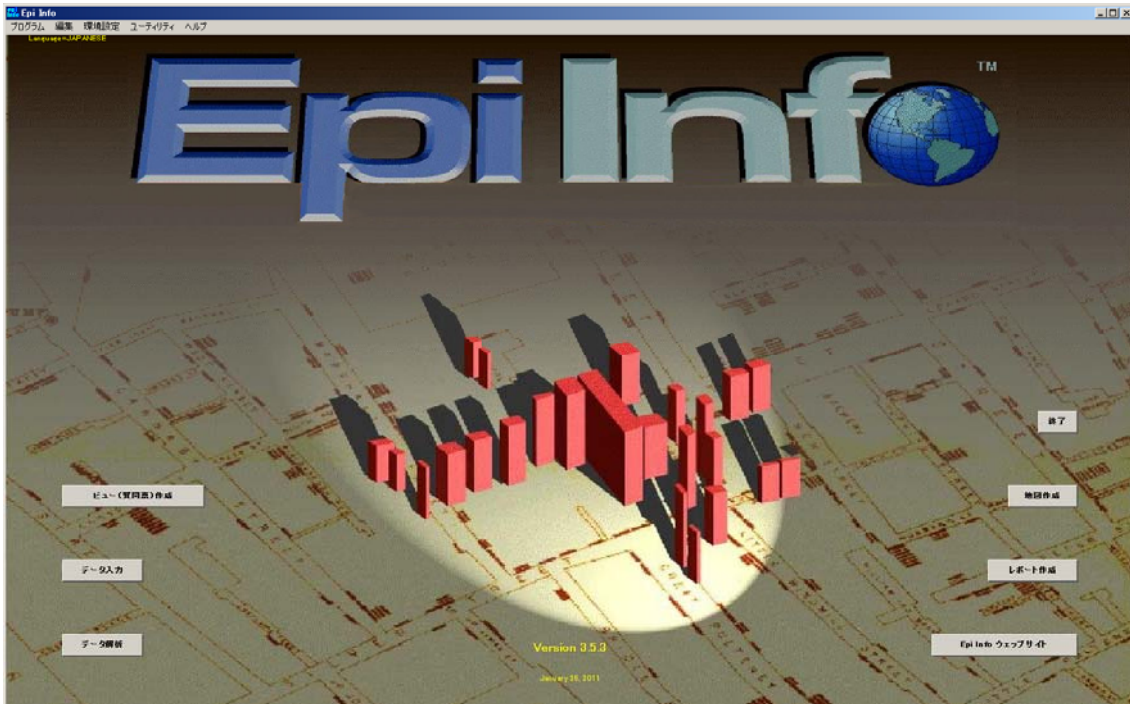


図1 Epi Info 日本語版のトップページ。ジョン・スノーによる1954年夏ロンドン・ブロードストリート周辺のコレラ症例分布を示したスポット・マップが背景に使われている。Epi Info™ はCDCのトレードマークです。

学んで欲しいポイント

1. 症例の定義の変更をおこなう。
2. 層別分析をおこなう。
3. 質問票作成、データ入力、地図作成の機能を知る。

I 症例の定義の変更及び層別分析

疫学統計ソフト Epi Info 3.5.3j の基礎編では、納豆オクラ事例でのデータ解析を扱ったが、応用編では、症例の定義の変更と層別分析を引き続き行う。後半では調査票作成、データ入力、疫学地図作成の紹介を行う。

では基礎編を参照して、Epi Info を立ち上げ、データ解析をクリック。「データ/読込」で、C:\¥Epi_Info¥EXAMPLES ¥納豆オクラ.xls を読み込み、データ解析の準備を行う。症例の定義を変更する準備として、症例中の症状別発症状況を調べた。これを Epi Info で行うには、全被調査者161名から症例のみの集団を抽出し、この症例集団の症状別発症状況を調べればよい。

1 データ選択

データを抽出するには解析コマンド:「選択/IF」メニューの「選択」を使う。「選択」をクリックすると次の図2 選択画面が表示されるので、「選択基準」に「症状="有"」と入力する。入力方法は

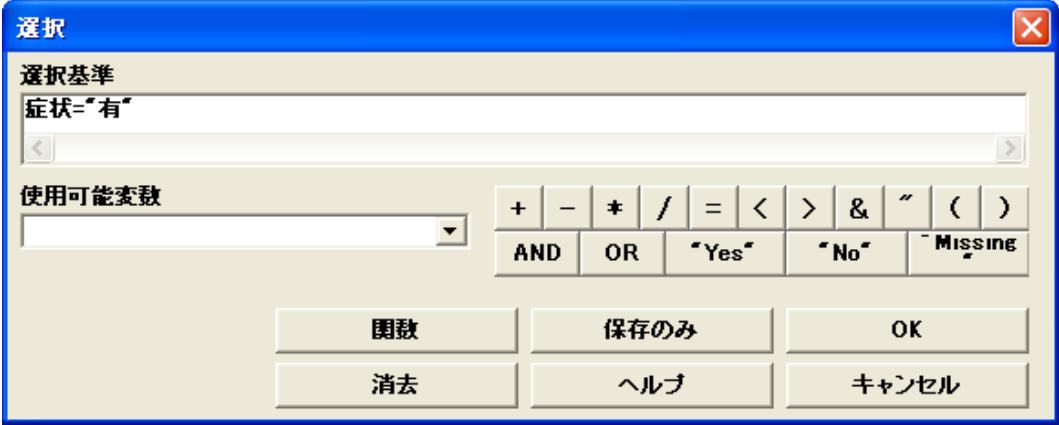


図2 選択画面

「使用可能変数」の逆三角印ボタンをクリックすると表示される変数リストから「症状」を選び、「記号メニュー」から、イコール「=」及びダブルコーテーション「"」がマウスクリックで得られる。「有」はキーボードから入力する。OK ボタンをクリックすると、症例72名のデータセットが作成される。出力画面上では「レコード数: 72 nbsp:」でレコード数は確認出来るが、「基本統計/リスト」で変数を「*」、表示形式をグリッドにすれば72名分のデータの確認が出来る。なお、「選択/IF」メニューの「選択解除」を行うと、元の161名のデータに戻る事ができる。「選択解除」を行わずに、「異なる選択」を繰り返すと選択基準が重複してしまい、症例72名のデータセットからの「異なる選択」となり、元の161名のデータからの「異なる選択」とならないことに注意が必要である。なお、選択基準は「使用可能変数」、「記号メニュー」から可能な限り文字・記号を選び、キーボードからの入力を最小限にすれば、記入ミスを防ぐ事ができる。

2 頻度・発現割合

では、腹痛から眼症状までのこの症例72名集団における頻度を調べる。「基本統計/頻度」を選ぶと頻度画面になる。ここでは、図3 頻度画面の様に、全16症状を一度に指定すると手間が省けて良いので、その方法を示す。「頻度」の▼をクリックし、16の症状を選ぶと図3のようになる。

これでOKをクリックすると、各症状別に人数、発現割合、さらに発現割合の95%信頼区間が出力される。出力される症状の順番を変更する事は出来ないので、報告書の順番等で出力したい場合は、1症状ずつ出力する。

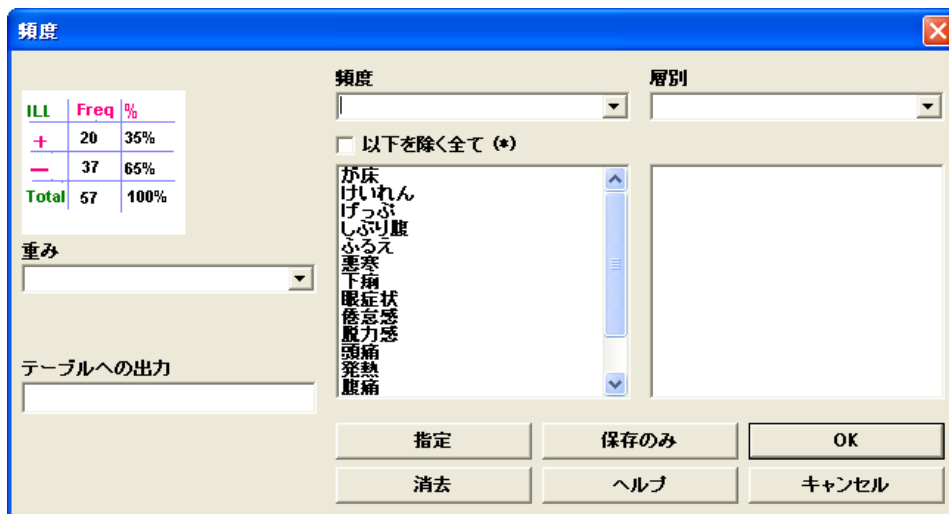


図3 頻度画面

3 平均・中央値

下痢、嘔吐回数及び発熱体温は、数値として入力されているので、これらの平均、中央値を求めるには「基本統計/平均」を使う。下痢回数の場合、対象変数に「下痢回数」を指定すると、図4の様に基本統計値が全て出力される。下痢を訴えた者は71名であるが、回数不明者が6名いるためデータ数（観測値）は65となっている。

観測値	合計	平均	分散	標準偏差
65	668.0000	10.2769	57.1721	7.5612
最小値	25%	中央値	75%	最大値 最頻値
1.0000	5.0000	9.0000	12.0000	30.0000 9.0000

図4 出力画面：基本統計量

4 変数定義・データ変換

下痢回数を1-3回、4-6回、7-9回、10回以上及び発熱体温を36°C-、37°C-、38°C-、39°C-にクラス分けした頻度分布表を作成するには、どうすれば良いだろう。下痢を例に、その手順を示す。下痢回数は最小値1回から最大値30回の整数値連続データである。これを4つのクラスに分け、度数分布表を作成する必要が生ずる。これは下痢回数の元データからクラス分けされたデータの新しい変数を作成する事になる。まず、新しく「下痢回数級別」変数を作成する。これは「変数/定義」で行う。図5定義画面で、変数名は「下痢回数級別」範囲は「標準的」追加設定の変数タイプは「テキスト」を指定し、プロンプトは必要ないので空白のままですべてのOKをクリック。

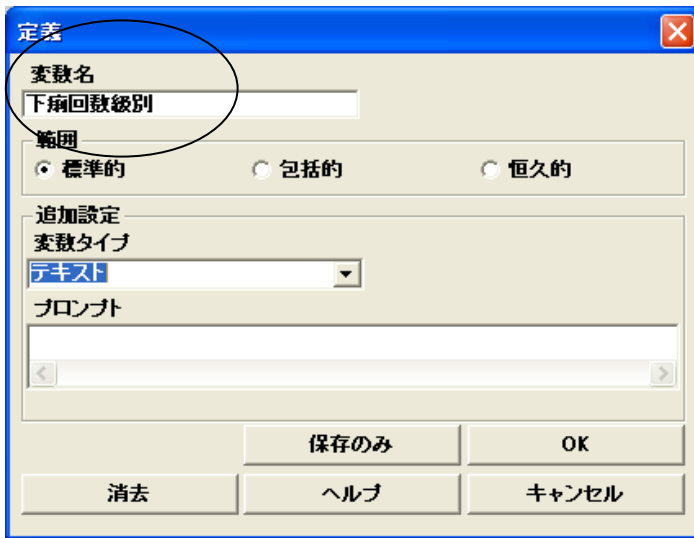


図5 定義画面

これで新変数の「下痢回数級別」が作成できた。では下痢回数データからデータ変換して下痢回数級別データを作成する。「変数 / 再コード化」で行う。図6再コード化画面で、「元変数」に下痢回数を、「再コード化変数」に下痢回数級別を選ぶ。



図6 再コード化画面 (その1)

級の中は規則的に3回なので、「範囲指定」をクリックして、「初期値」に0を、「終了値」に9を、「級間幅」に3を代入する。級下限値は含まず、級上限値は含むため、1から10でなく、0から9に指定している事に注意する。つぎの図7再コード化画面が表示される。

再コード化

元変数(連続型) 再コード化変数(文字型)
 下痢回数 下痢回数級別

Dates must be in US format

級下限値(含まず)	級上限値(含む)	再コード値(文字型)
LOVALUE		0 <=0
0		3 >0 - 3
3		6 >3 - 6
6		9 >6 - 9
9	HIVALUE	>9

範囲指定 保存のみ OK
 消去 ヘルプ キャンセル

図7 再コード化画面 (その2)

このままだと、再コード値がわかりにくいので、以下の図8の様に変更した。ここで第1列の再コード値は変更していない。この級は級数がゼロなので、頻度分布に現れないからである。また第2行から第5行の再コード値(文字型)の先頭に丸数字①②③④を付けたのは頻度分布における級の並びの順番が再コード値(文字型)の先頭の文字の順番、同値のときは2番目の文字の順番で決まるため、丸数字①②③④を付けないと図9の様な並びでなく、"1 0回以上"が先頭に来て、次が"1 - 3回"となり、順番がおかしくなるのを避ける工夫である。OKをクリックすると、下痢回数級別の再コード値(文字型)データが作成される。

再コード化

元変数(連続型) 再コード化変数(文字型)
 下痢回数 下痢回数級数

日付はUS形式であること

級下限値(含まず)	級上限値(含む)	再コード値(文字型)
LOVALUE		0 <=0
0		3 ①1 - 3回
3		6 ②4 - 6回
6		9 ③7 - 9回
9	HIVALUE	④10回以上

範囲指定 保存のみ OK
 消去 ヘルプ キャンセル

図8 再コード化画面 (その3)

「基本統計/頻度」を用いて下痢回数級別の度数分布が図9の様になる。次の5節は症例の定義の変更を行うので、症例72名の「選択解除」を行い、元の全被調査者161名に戻しておくこと。

[先へ](#)

下痢回数級数	頻度	パーセント	累積パーセント	
① 1-3回	8	12.3%	12.3%	<div style="width: 12.3%; background-color: yellow;"></div>
② 4-6回	17	26.2%	38.5%	<div style="width: 26.2%; background-color: yellow;"></div>
③ 7-9回	14	21.5%	60.0%	<div style="width: 21.5%; background-color: yellow;"></div>
④ 10回以上	26	40.0%	100.0%	<div style="width: 40.0%; background-color: yellow;"></div>
合計	65	100.0%	100.0%	<div style="width: 100.0%; background-color: orange;"></div>

図9 出力画面：下痢回数級別の度数分布

5 症例定義の変更

症例の定義を変更してデータを絞り込むには、選択/IFメニューの「選択」を用いる。変更する症例の定義は「9月19日にB旅館の朝食を食べた者のうち、9月19日から9月22日までに下痢を4回以上かつ発熱体温38℃以上の者」で、無症者の定義は「9月19日にB旅館の朝食を食べた者のうち、9月19日から9月22日までに表1のいずれの症状も呈さなかった者(データの症状が無に対応)」である。従って選択基準には以下の図10の様に入力すれば良い。

図10 選択画面

ここで「19日朝食」は「使用可能変数」リストから選べる。「= ” () > AND OR 」は選択画面からマウスでクリックすることで入力出来る。「有 無 4 38」はキーボードから入力する。**OK** をクリックすると、選択基準に合う93名のデータが選択される。「選択/IF」メニューの「選択解除」を行うと、元の161名のデータに戻る事ができる。

6 新しい症例定義による分割表

新しい症例の定義に基づく食品を食べた・食べないと、症例・無症の分割表分析は前回紹介した「基本統計/分割表」を用いればよい。例えば図11のように「曝露変数」に納豆オクラを選び、「目的変数」に症状を選ぶと、

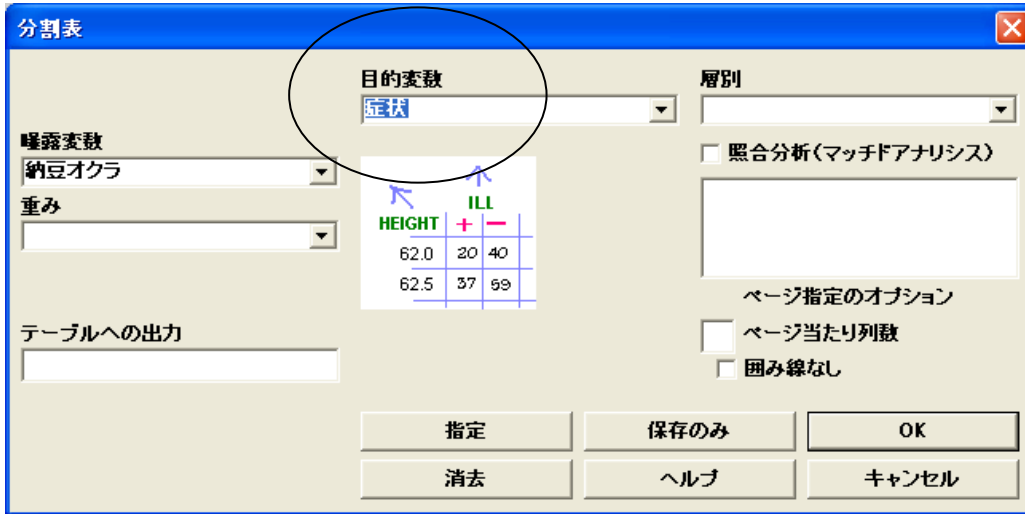


図 1 1 分割表画面

以下の図 1 2 の出力結果が得られる。分割表の右に表示された 4 つの色違いの正方形の大きさは分割表の 4 つのセルの数値に比例している。右下がり対角要素が大きいと正の関連があること、逆に右上がり対角要素が大きいと負の関連がある事が視覚的にとらえられる。

TABLES 納豆オクラ 症状

[次の処理](#)

[先へ](#)

症状

納豆オクラ	有	無	合計
有	40	4	44
行%	90.9	9.1	100.0
列%	95.2	8.5	49.4
無	2	43	45
行%	4.4	95.6	100.0
列%	4.8	91.5	50.6
TOTAL	42	47	89
行%	47.2	52.8	100.0
列%	100.0	100.0	100.0

単一分割表解析

	Point 推定	95% 信頼区間	
		下側	上側
パラメータ: オッズに基づく			
オッズ比(交差積)	215.0000	37.3179	1238.6820 (T)
オッズ比(最尤推定値)	184.2071	37.5378	1486.6710 (M)
		32.0357	2123.9935 (F)
パラメータ: リスクに基づく			
リスク比(RR)	20.4545	5.2604	79.5356 (T)
リスク差(RD%)	86.4646	76.0526	96.8767 (T)
(T=テイヤラー級数; C=コーンフィールド; M=Mid-P; F=フィッシャーの正確法)			
統計的検定			
カイニ乗-無補正	66.7398	0.0000000000	
カイニ乗-マンテル・ヘンツェル	65.9900	0.0000000000	
カイニ乗-補正(イエーツ)	63.3154	0.0000000000	
Mid-p 正確		0.0000000000	
フィッシャー直接法		0.0000000000	

図 1 2 出力画面 : 分割表

新しい症例定義に合致した93名中、4名の納豆オクラの喫食データが不明であるので、89名の分割表結果となる。2 x 2表データ、オッズ比の点推定値、95%信頼区間、カイ2乗値、P値が求まる。

7 層別分析

層別分析で交絡を調整するには「基本統計/分割表」で調整したい変数を「層別」で指定すれば良い。ご飯を食べた群・食べなかった群の層別解析をするには、分割表画面で、「層別」にご飯を指定する。分割表画面は図13の様になる。

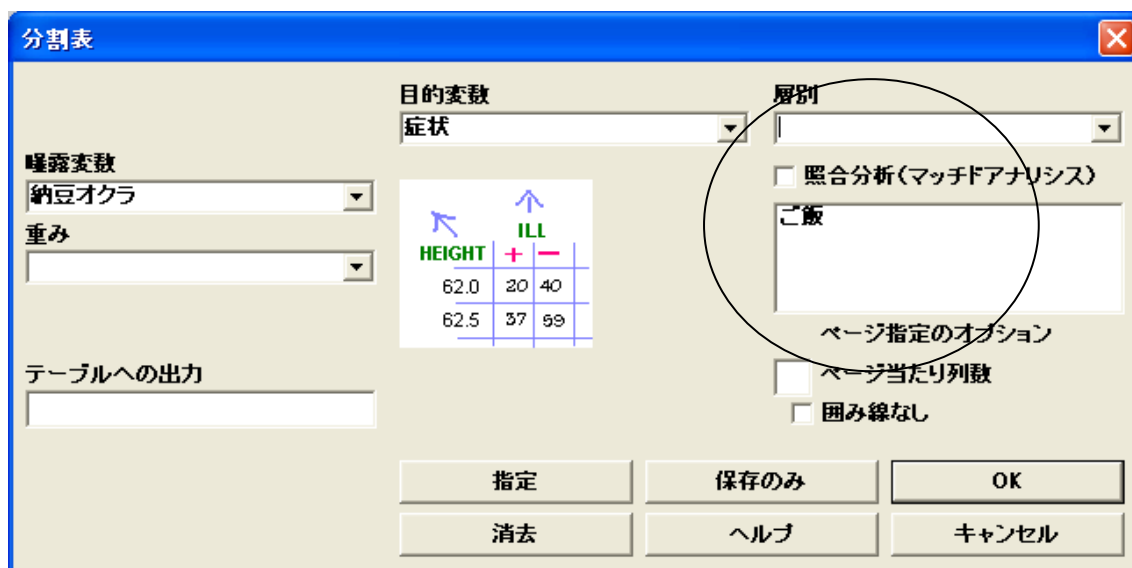


図13 分割表画面

層別分析の実行結果図14はご飯を食べた群の分割表結果、食べなかった群の分割表結果、及びこれら2群を統合したまとめにマンテル・ヘンツェル調整オッズ比の分析結果が表示される。

まとめ

戻る 先へ 現在の処理
SUMMARY INFORMATION

パラメータ	Point 95%信頼区間		
	推定	下側	上側
オッズ比推定値			
粗オッズ比(交差積)	78.7500	22.2627,	278.5630 (T)
粗(最尤推定値)	73.1322	22.2171,	296.5905 (M)
		19.9052,	358.3089 (F)
調整オッズ比 (マンテルヘンツェル)	47.0414	12.7816,	173.1309 (R)
調整オッズ比 (最尤推定量)	40.2462	11.6938,	167.9756 (M)
		10.3820,	203.6731 (F)
リスク比(RR)			
粗リスク比 (RR)	5.8594	3.0950,	11.0928
調整リスク比 (マンテルヘンツェル)	3.8926	1.9914,	7.6088

(T=テ일러級数; R=RGB; M=正確 mid-P; F=フィッシャーの正確法)

統計的検定(全体の関連)	カイ二乗	片側p値	両側p値
マンテルヘンツェル - カイ二乗-無補正	48.9185		0.0000
MH Chi-square - corrected	45.2070		0.0000
Mid-p 正確		0.0000	
フィッシャー直接法		0.0000	

次の2つの検定では、低いp値は、比が層別に異なることを示唆します

層別オッズ比が異なるか(交互作用)のためのカイ二乗	0.8145		0.3668
層別リスク比が異なるかのためのカイ二乗	0.2633		0.6078

図 1 4 層別分析まとめ出力画面

以上でデータ解析の紹介を終えて、ここから手短かに質問票作成、データ入力、地図作成を紹介する。

II 質問票作成

Epi Info を用いて質問票を作成する事が出来るので、その概要を説明する。食中毒事件での症状喫食調査用紙の作成は、トップ画面のビュー (質問票) 作成で行う。質問票はデータベースではビュー(View)と呼び、1人あたりの入力画面を意味し、質問票にあたる。まず Epi Info で使われている事例オスウエゴ Oswegoj.xls の質問票を見してみる。Epi Info 3.5.3j トップ画面から **ビュー (質問票) 作成** をクリックする。メニューバーの **ファイル** をクリックしてプルダウンメニューから **開く** を選ぶ。

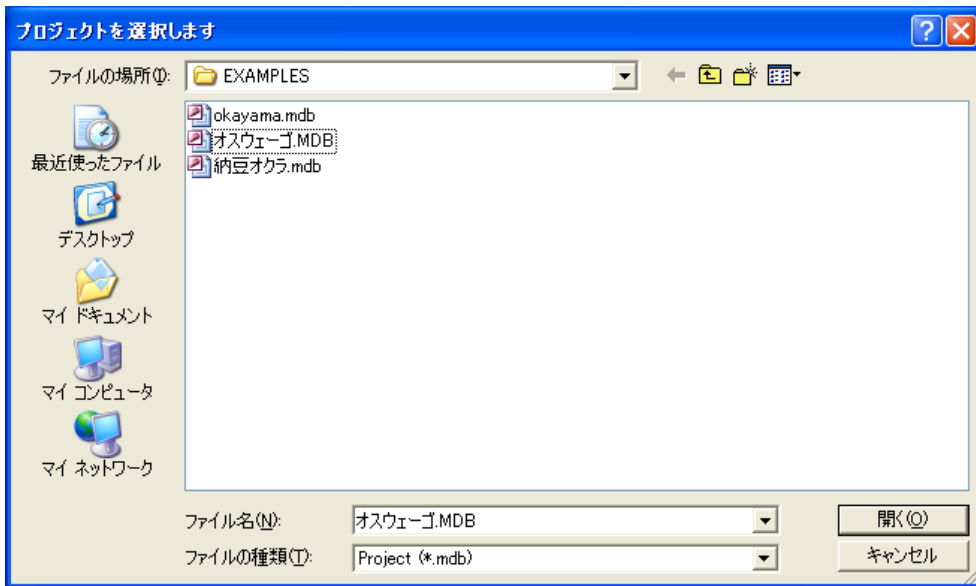


図 1 5 プロジェクト選択画面

図 1 5 フォルダ Examples 内の オスウェゴ.MDB をマウスで選択し、開く(O)をクリックすると、図 1 6 ビュー選択画面が現れる。

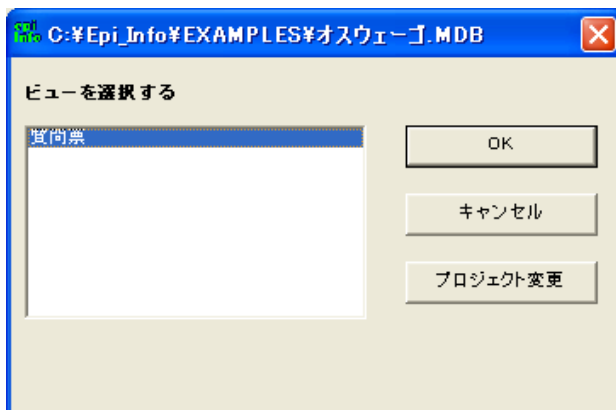


図 1 6 ビュー選択画面

データベース オスウェゴ.MDB 内にはビュー 質問票が 1 つしかないなので、これを選んで、OKをクリック。図 1 7 オスウェゴ教会会食質問票が現れる。

図 1 7 オスウェーゴ教会会食質問票

このような質問票が Epi Info のビュー（質問票）作成機能で行える。もう 1 つ、岡山市で使われている食中毒質問票を開いてみる。フォルダ Examples 内の okayama.mdb をマウスで選択し、**開く(O)**をクリック。データベース okayama.mdb 内には「ビュー 食中毒調査」が 1 つしかないのので、これを選んで、**OK**をクリック。岡山市食中毒質問票が現れる。この質問票は 1 人あたり 3 ページある。左上の 1 Page、2 Page、3 Page をマウスで選ぶと各ページが現れる。この質問票作成の詳細は「Epi Info3.3.2 日本語版による質問票作成の手引き(2007/04/01)」を参照してほしい。

III データ入力

質問票作成で作成した質問票を用いてデータを入力できる。単票（カード）形式のデータ入力なので、エクセル等の表の各行への入力より、容易であるため、誤りも少なく出来る。ノート PC を使えば現場で直接入力も可能となる。

オスウェーゴのビュー 質問票 にデータを入力してみる。Epi Info 3.5.1j 初期画面から **データ入力** をクリック。メニューバーの**ファイル**をクリックしてプルダウンメニューから **開く**を選び、フォルダ Examples 内のオスウェーゴ.MDB のビュー質問票を指定する。図 1 8 データ入力画面が現れ、

すでに2人分のデータが入力してあるのが分かる。

オスウェーゴ教会会食質問票

コード 名前 年齢 性別

症状の有無 食事時間 発症時刻

喫食履歴(Yes or No)

食品	
焼きハム	<input type="text" value="Yes"/>
ほうれん草	<input type="text" value="Yes"/>
マッシュポテト	<input type="text" value="Yes"/>
キャベツサラダ	<input type="text" value="No"/>
ゼリー	<input type="text" value="No"/>
ロールパン	<input type="text" value="Yes"/>
トーストパン	<input type="text" value="No"/>
フルーツサラダ	<input type="text" value="No"/>

飲み物	
ミルク	<input type="text" value="No"/>
コーヒー	<input type="text" value="Yes"/>
ウオター	<input type="text" value="No"/>

デザート	
ケーキ	<input type="text" value="No"/>
パニラアイス	<input type="text" value="Yes"/>
チョコレートアイス	<input type="text" value="No"/>

図18 データ入力画面

左にある新規ボタンをクリックして、3人目からのデータが入力できる。最後はファイルから保存、終了。この質問票によるデータ入力の詳細は「Epi Info3.3.2 日本語版によるデータ入力の手引き(2007/04/01)」を参照してほしい。

IV 地図作成

Epi Info のトップページの背景に、ジョン・スノーによる1954年夏ロンドン・ブロードストリート周辺のコレラ症例分布を示したスポットマップが使われているように、疫学データの空間的広がりを示す地図表示は有益な情報を提供する。記述疫学の3要素：時間・場所・人の「場所」の情報をメインにした分析である。現在は疾患の地域分析、感染症の発生動向監視等に地理情報システムGISは有力な技術、分析法になっている。Epi Infoにも、保健医療のための地理情報システムGISが地図作成機能として提供されている。

ここではこの地図作成機能の紹介に、サンプルとして準備してある全国都道府県白地図に都道府県名と腸管出血性大腸菌O157発症数の色分け地図を描いてみる。Epi Info 日本語版トップページから **地図作成** をクリックする。

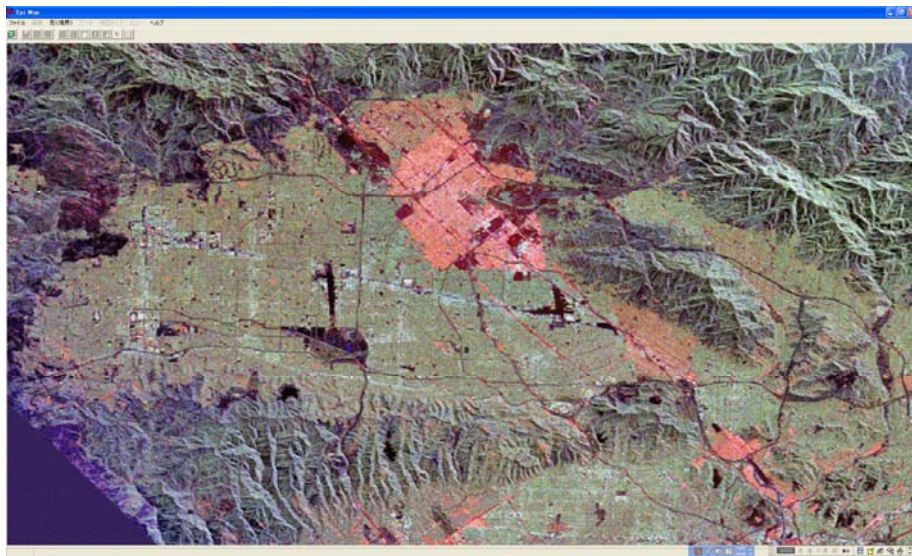


図19 地図作成画面

図19 米国西海岸地方の写真が表示され、ESRI社のロゴがスポットで表示される。ファイルのプルダウンメニューから **地図管理** を選択する。

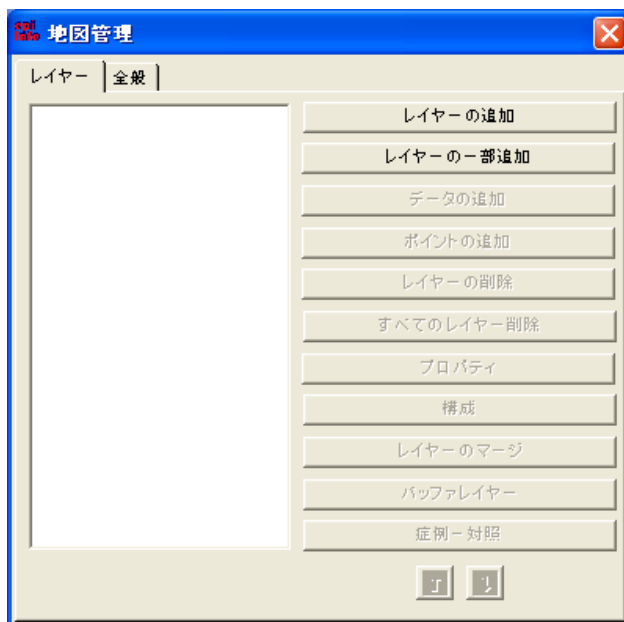


図20 地図管理画面

この図20 地図管理メニューから **レイヤーの追加** を選択する。

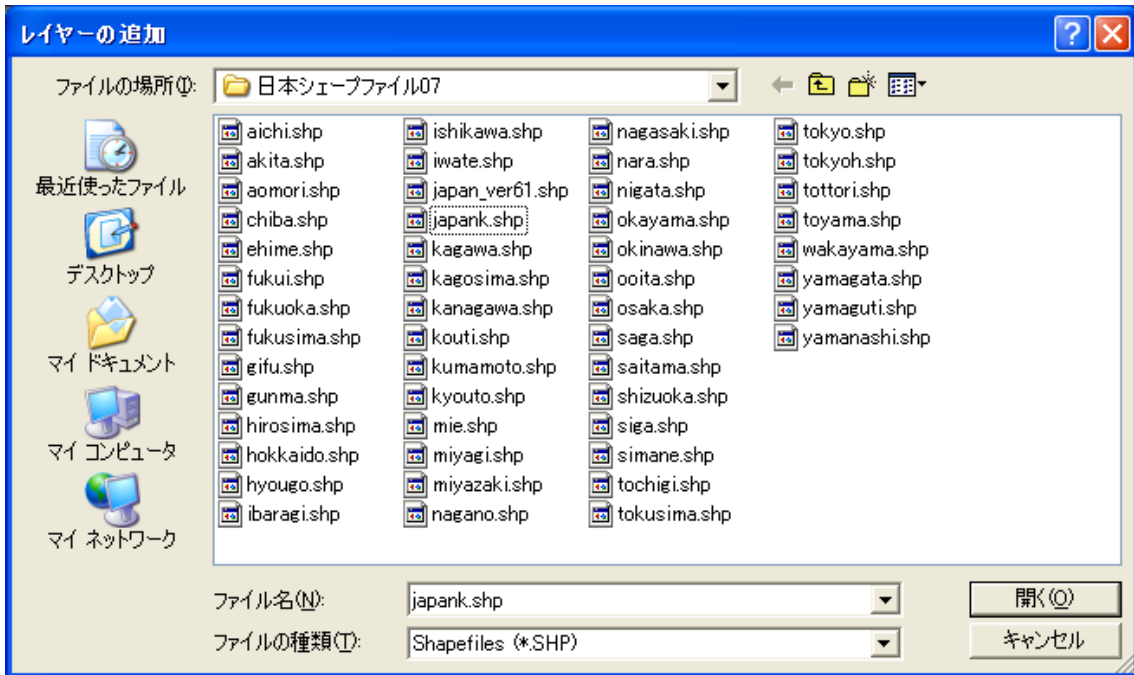


図 2 1 ファイル選択画面

C:\Epi_Info フォルダの日本シェープファイル 07 フォルダには図 2 1 のようにサンプルとして全国地図各県地図のシェープファイルを準備してある。この中の全国市町村界データ `japan_ver61.shp(.shx,.dbf)` の著作権は ESRI ジャパン (株) に帰属している。各県データはそれを加工して作成したものである。ユーザによるデータ改変に伴う障害については、ESRI ジャパン (株) は保証しないことを了解すること。それではファイル名に `japank.shp` を選んで **開く** をクリック。

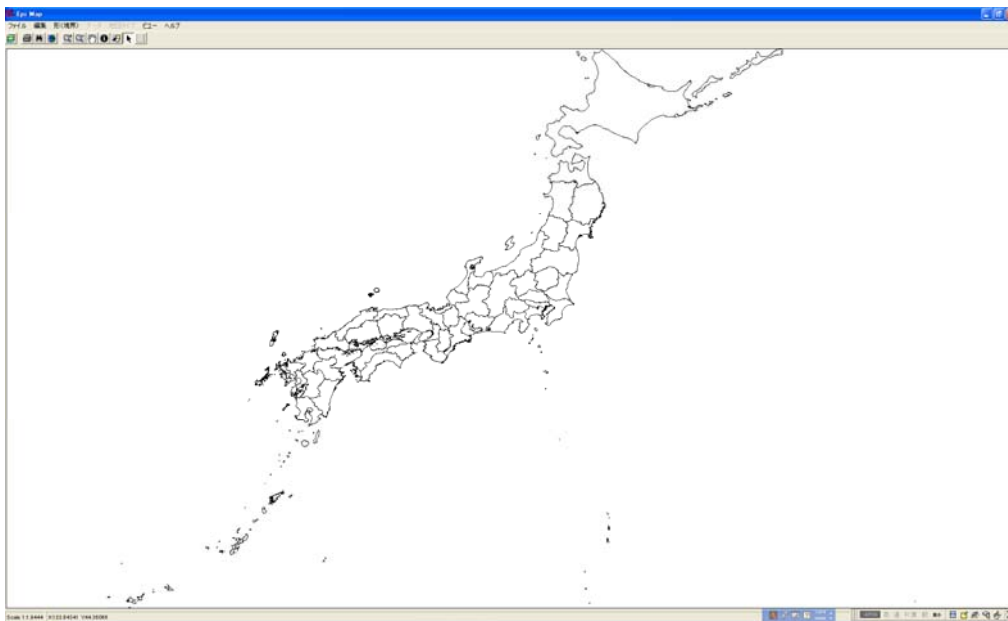


図 2 2 出力画面：都道府県白地図

図 2 2 都道府県白地図が表示さる。次に地図管理メニューから **プロパティ** を選択する。

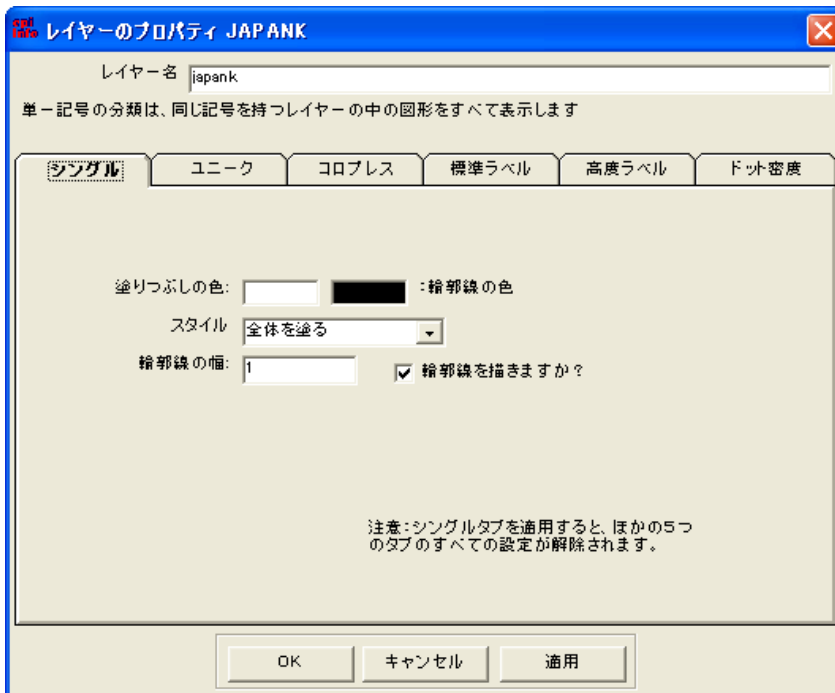


図 2 3 プロパティ画面

この図 2 3 プロパティメニューで、高度ラベルをマウスでクリックして選択すると、図 2 4 高度ラベル画面が選択出来る。。

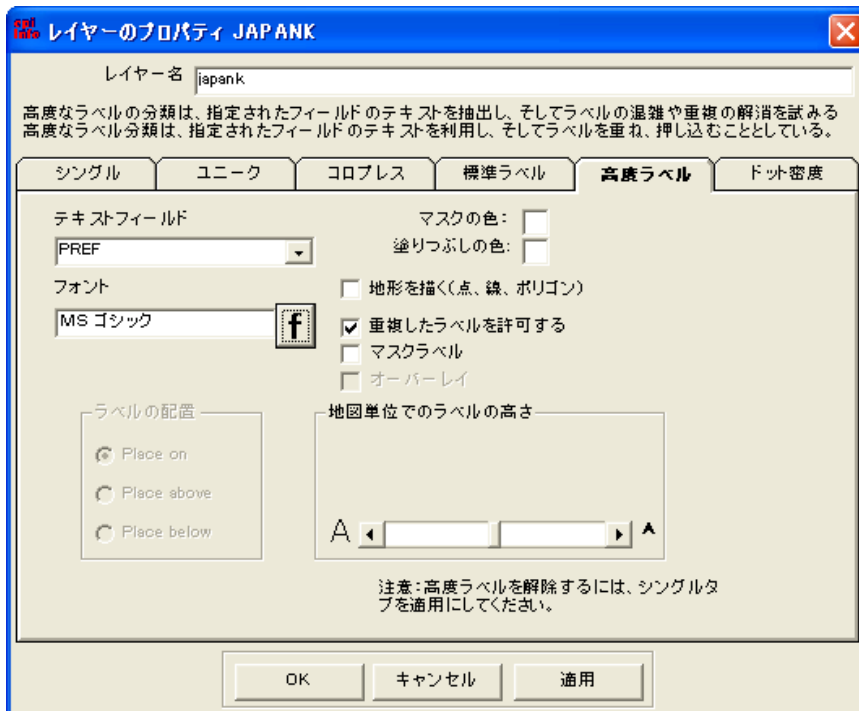


図 2 4 プロパティ・高度ラベル画面

テキストフィールドに **PREF** [県名] を選び、フォントにはMSゴシックを選択する。適用 をクリックすると県名が表示される。また **コロプレス** を選んで、図2 5コロプレス画面の



図2 5 プロパティ・コロプレス画面

数値フィールドに **O1572008** [2008年O157患者数] を選び、**凡例を置き換える** をクリックして、**適用** をクリックすると、図2 6のようにO157患者数での色分け地図が描かれる。

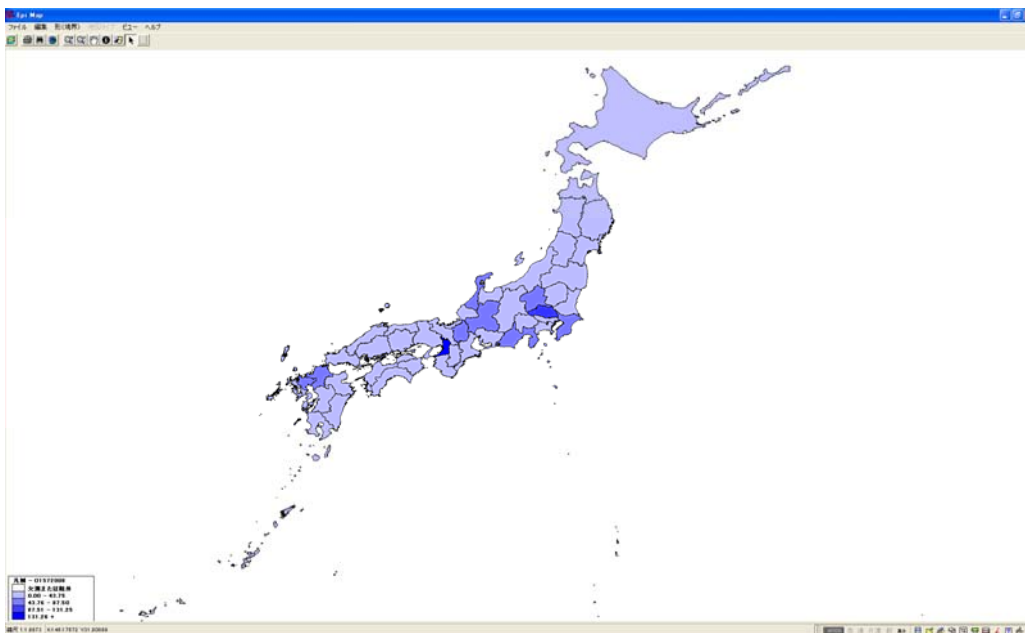


図2 6 出力画面：人口数色分け地図

コラム：

地図作成で使うユーザが準備するシェープファイルは、図形データ (*.shp)、インデックスデータ (*.shp)、属性データ (*.dbf) の3種類で構成される地図データファイル群(ファイルの集合体) のことである。図形データとインデックスデータが各地域の図形情報を与え、属性データがそれら地域の属性情報を与える。このシェープファイルは、GISの世界市場でトップシェアを誇る米国ESRI社 (Environmental Systems Research Institute) が提唱したもので、一般に広く公開されており、GIS業界の標準フォーマットと言われている。ESRI・LGGI製品はもちろん、多くのGISソフトウェアで利用が可能である。Epi InfoではESRIジャパン(株)が無償で配布している全国市町村界シェープファイルをサンプルとして提供している。

http://www.esri.com/gis_data/japanshp/japanshp.html

属性データ (*.dbf) は、各地域の任意の属性また図形データ・インデックスデータと結合するためのキーを格納しているファイルで、データ・フォーマットは多くの表計算アプリケーションで使われているdBASE形式である。これはエクセルで編集できるので、ユーザが自分のデータを準備すれば地図上にそのデータを表示できる。

以上で簡単な地図作成の紹介を終わる。詳細な地図作成マニュアルは「Epi Info3.3.2 日本語版による地図作成の手引き(2007/04/01)」を参照してほしい。

最後に前回を含めて、Epi Info 日本語版の利用手引きが Epi Info のHP

感染症情報センターHP：<http://idsc.nih.gov/jp/index-j.html> 画面の Epi Info 画面からダウンロードできる。Epi Info の旧バージョン様に作成したものであるが、内容は現バージョンに対応出来ているので参照出来る。これらのマニュアルを参考に、疫学調査に世界標準といえるCDC提供の疫学ソフト Epi Info をどしどし使って欲しい。我々も日本語版の改善を今後も続けていくつもりである。

[Epi Info 3.3.2 日本語版によるデータ解析の手引き\(2007/04/01\)](#) (PDF ファイル・652KB)

[Epi Info 3.3.2 日本語版による調査票作成の手引き\(2007/04/01\)](#) (PDF ファイル・530KB)

[Epi Info 3.3.2 日本語版によるデータ入力の手引き\(2007/04/01\)](#) (PDF ファイル・161KB)

[Epi Info 3.3.2 日本語版による地図作成の手引き\(2007/04/01\)](#) (PDF ファイル・302KB)