

厚生省特定疾患の全国疫学調査による 患者数推定と推定誤差について

高木 廣文* 佐藤 俊哉** 稲葉 裕***

I はじめに

厚生省は、昭和47年以来いくつかの難病を「特定疾患」として指定している。各特定疾患に関する調査研究班を中心に、各種の研究が行われているが、そのなかの1つに、特定の難病の全国患者数を把握するために行う全国疫学調査一次調査がある（たんに「一次調査」と呼ばれている）。一次調査は、各調査研究班と難病の疫学調査研究班の事務局により、全国の特定の病院の特定診療科（ときに特定の診療所）に対して、特定年での1年間（ときには数年間）の当該疾患の受診患者数を性別に葉書に記載し、返送するように依頼するという郵送法により行われている。調査対象施設もしくは診療科数は数千に達し、その回収率は50%前後が一般的であり、未回収分の調査対象に関しては、患者数を推定する必要がある。最も簡単な方法は、報告患者数を回収率で除すというものであるが、患者のいない施設や診療科では、患者のいる施設などに比べて回答率は低いものと考えられ、ほとんどこの方法は採用されていない。

厚生省特定疾患の全国患者数推定には、「柳井らの方法¹⁾」が本誌に掲載された昭和50年以来、一次調査にもとづく標準的な推定方法としてよく用いられている²⁾。同方法は、原論文では重回帰式にもとづくものであったが、現在ではより簡単な計算方法として以下のような単回帰式にもとづく推定方式が取られている。すなわち、調査によって得られた都道府県別、もしくは適当に区分した地域ブロック別の回収率を説明変数とし、人口10万人対の有病率（報告有病率）を基準変数とする単回帰直線にもとづく推定を行うものである。しかし、同方法は回収率と報告有病率の単相関係数が負の場合、まったく推定が行えないという欠点がある。そのため、最近ではその欠点を改良するために、原点を通る回帰直線にもとづく推定を行うという「中江らの方法³⁾」、「水野らの方法⁴⁾」、さらに「中江らの方法の修正方法⁵⁾（修正中江らの方法）」などが提唱されている。

しかし、実際問題として上記の手法のうち、どの方法

によれば全国患者数をよりよく推定できるのか、また推定値はどの程度信頼できるのかという問題がある。この点に関しては、各推定法が単回帰式にもとづくが、後述するように通常とは異なる計算式を用いるために、理論的な面から客観的な議論ができないままに放置されていた。

各推定方法を比較する場合、各手法にどのような特徴や偏りがあるかを知る必要がある。最も重要な点は推定の信頼度を示す指標となる推定誤差（推定値の標準誤差）を求めて、相互に比較することである。しかし、各推定方法による推定誤差を提示し、各手法に関して理論的および実際の比較を行った報告は見あたらない。

本報告では、上記の方法による患者数の推定値の理論的な推定誤差の計算方法を初めて提示する。さらに、実際の一次調査資料による患者数推定結果を相互比較し、有用な知見を得たので報告する。最後に、厚生省特定疾患の全国疫学調査一次調査にもとづく患者数推定方法に関するいくつかの問題点に関して、理論と応用結果を踏まえて考察を行う。

II 資料および方法

1. 資料について

後述の各方法による全国患者数推定と推定誤差を検討するために、厚生省特定疾患神経皮膚症候群調査研究班疫学調査分科会と難病の疫学調査研究班事務局が、昭和61年1月～3月に実施したレックリングハウゼン病に関する全国疫学調査一次調査⁶⁾の資料を用いた。調査は200床以上の病院・施設などを対象とし、表1に示したように皮膚科、小児科、整形（および形成）外科、眼科、精神（神経）科、脳（神経）外科、耳鼻（咽喉）科の7診療科について、前述した郵送法により過去1年間での受診患者数を調査したものである。調査対象となった診療科数は7,512件、対象診療科からの回答数は4,575件、回収率は60.9%であった。報告患者数は男1,486例、女1,508例であった。

さらに、推定結果の相互比較を行うために、上記調査と同一時期に行われた厚生省特定疾患神経性食思不振症調査研究班による全国疫学調査一次調査資料^{7,8)}にもとづく推定もあわせて試みた。同調査は200床以上の病院

* 聖路加看護大学統計学教室

** 東京大学医学部疫学教室

*** 順天堂大学医学部衛生学教室

・施設などを対象とし、表2に示したように内科、精神(神経)科、小児科、産婦人科、心療内科の5診療科について、前述した郵送法により昭和60年中の1年間での受診患者数を調査したものである。調査対象診療科数は5,283件、対象診療科からの回答数は3,397件、回収率は64.3%であった。報告患者数は男110例、女2,281例であった。なお、2つの調査の同一診療科の調査対象数に若干の差があるのは、各研究班により休診や閉鎖のあった診療科に関する把握数に相違があるためである。

全国患者数の推定は、各手法ごとに診療科別・性別に行い、合計を求めることにする。また、報告有病率の計算のための地域別人口数は、昭和60年10月1日現在のものを用いた⁹⁾。

患者数推定は、地域区分を47都道府県とした場合と、9地域ブロックとした場合の2とおりについて行った。なお地域ブロックは、北海道、東北、関東、北陸、中部・東海、近畿、中国、四国、九州・沖縄の9ブロックに分けた。

2. 患者数推定と推定誤差の算出方法

まず、各方法の説明のために使用する記号を定義する。

各地域(都道府県や地域ブロックなど)で調査対象となった診療科数(もしくは施設数)を M_i とし、そのうち報告があった診療科数を N_i とする($i=1, \dots, k$; k は地域数)。また、 N_i 個の診療科から報告された患者数を t_i とする。

地域 i の調査の回収率 R_i は、

$$R_i = N_i / M_i \quad (1)$$

であり、地域 i の人口を P_i とすると、報告有病率 y_i は、

$$y_i = t_i / P_i \quad (2)$$

となる。

全国での調査対象診療科数 M 、報告診療科数 N 、回収率 R は、それぞれ、

$$M = \sum_i M_i, N = \sum_i N_i, R = N / M \quad (3)$$

となり、全国人口を P とすると、報告患者総数 t は、

$$t = \sum_i t_i \quad (4)$$

となるので、報告有病率 y は、

$$y = t / P \quad (5)$$

となる。ここで、上記の $\sum_i$ は地域についての総和を計算することを示す。

1) 柳井らの方法¹⁾

「柳井らの方法」では、まず各地域の有病率 y_i と回収率 R_i ($i=1, 2, \dots, k$)の間に、

$$y_i = \beta_1 R_i + \alpha + \varepsilon_i, \varepsilon_i \sim N(0, \sigma_1^2) \quad (6)$$

が成り立つものと仮定する。なお、 $N(0, \sigma_1^2)$ は平均0、分散 σ_1^2 の正規分布を示し、 ε_i は誤差項であり、その分布が $N(0, \sigma_1^2)$ に従うものと仮定する。

表1 レックリングハウゼン病の調査対象数、回答数および報告患者数など

診療科	調査対象数	回答数	回収率	報告患者数	
				男	女
皮膚科	919	580	0.631	586	650
小児科	1,219	843	0.692	228	190
整形外科	1,474	793	0.538	333	344
眼科	1,073	606	0.565	101	102
精神科	1,095	691	0.631	54	45
脳外科	683	438	0.641	136	130
耳鼻科	1,049	624	0.595	48	47
合計	7,512	4,575	0.609	1,486	1,508

表2 神経性食思不振症の調査対象数、回答数および報告患者数など

診療科	調査対象数	回答数	回収率	報告患者数	
				男	女
内科	1,711	1,033	0.604	28	633
精神科	1,101	732	0.665	46	838
小児科	1,222	909	0.744	26	256
産婦人科	1,106	705	0.637	0	286
心療内科	143	18	0.126	10	268
合計	5,283	3,397	0.643	110	2,281

* レックリングハウゼン病と同一診療科の対象数に若干差があるのは、休診・閉鎖などの把握数の違いによる。

母回帰係数 β_1 と定数項 α (切片)の推定値 b_1, a はよく知られているように、 $\bar{y}$ と $\bar{R}$ を報告有病率と回収率の各平均値とすると、

$$b_1 = \sum_i (R_i - \bar{R}) (y_i - \bar{y}) / \sum_i (R_i - \bar{R})^2$$

$$a = \bar{y} - b_1 \bar{R} \quad (7)$$

である¹⁰⁾。

また誤差分散 σ_1^2 の推定値 $\hat{\sigma}_1^2$ (残差分散)は、

$$\hat{\sigma}_1^2 = (1 - r_{yR}) \sum_i (y_i - \bar{y})^2 / (k - 2) \quad (8)$$

となる⁵⁾。ここで、 r_{yR} は報告有病率と回収率のピアソンの積率相関係数(以下「相関係数」と呼ぶ)であり、 k は地域数である。

回帰係数 b_1 の分散は、

$$V(b_1) = \hat{\sigma}_1^2 / \sum_i (R_i - \bar{R})^2 \quad (9)$$

である¹⁰⁾。

「柳井らの方法」では、回収率と推定有病率が無相関になるように推定有病率を求めている¹⁾。すなわち、地域 i の未回収分 $(1 - R_i)$ について、回帰係数 b_1 をかけた数値を報告有病率 y_i に加えたもの、

$$\hat{\mu}_i = y_i + (1 - R_i) b_1 \quad (10)$$

により推定有病率 $\hat{\mu}_i$ を求めている。なお、実際には回

帰式による「予測 prediction」と考えられるが、本論文では慣用に従って「推定 estimation」という用語を用いることにする。

次に、推定有病率 $\hat{\mu}_i$ の推定の信頼度を検討するために、その期待値と分散を求める。

地域 i の推定有病率 $\hat{\mu}_i$ の期待値 $E(\hat{\mu}_i)$ は、

$$E(\hat{\mu}_i) = E(y_i) + (1 - R_i)E(b_i) = \beta_1 + \alpha \quad (11)$$

と求められる。したがって、推定有病率 $\hat{\mu}_i$ の分散 $V(\hat{\mu}_i)$ は、(6)式の回帰式が成り立つと仮定できるならば、

$$\begin{aligned} V(\hat{\mu}_i) &= E\{y_i + (1 - R_i)b_i - \beta_1 - \alpha\}^2 \\ &= V(e_i) + (1 - R_i)^2 V(b_i) + 2(1 - R_i) \\ &\quad \times (R_i - \bar{R}) V(e_i) / \sum_j (R_j - \bar{R})^2 \end{aligned} \quad (12)$$

と求められる。

誤差項 e_i が先に仮定したように正規分布 $N(0, \sigma_e^2)$ に従うものとし、誤差分散 σ_e^2 の推定値として(8)式による残差分散 $\hat{\sigma}_e^2$ を用いると、推定有病率 $\hat{\mu}_i$ の分散は、

$$V(\hat{\mu}_i) = \hat{\sigma}_e^2 \left\{ 1 + \frac{(1 - R_i)(1 + R_i - 2\bar{R})}{\sum_j (R_j - \bar{R})^2} \right\} \quad (13)$$

と近似できる。

地域 i の推定患者数 $\hat{\tau}_i$ とその分散 $V(\hat{\tau}_i)$ は、推定有病率 $\hat{\mu}_i$ に人口 P_i をかければよい。すなわち、

$$\hat{\tau}_i = P_i \hat{\mu}_i, \quad V(\hat{\tau}_i) = P_i^2 V(\hat{\mu}_i) \quad (14)$$

により求められる。

全国推定患者数 $\hat{\tau}$ とその分散 $V(\hat{\tau})$ は、各地域を独立とすれば、

$$\hat{\tau} = \sum_i P_i \hat{\mu}_i, \quad V(\hat{\tau}) = \sum_i P_i^2 V(\hat{\mu}_i) \quad (15)$$

により求められる。

2) 水野らの方法⁴⁾

「柳井らの方法」による回帰直線は原点を通るとは限らないため、(6)式にもとづく回帰式で回収率が0の場合に有病率が負となることがある。また、報告有病率と回収率の相関が負の場合、事実上推定は不可能であり、理論的にも実際的にも不都合である。これらの問題を回避するために、中江らは、回収率0、有病率0というデータを恣意的に多数加え、原点の近くを通るような回帰直線を求める方法を提唱している³⁾。しかし、この方法は、データを作為的に追加するという点に大きな問題がある。そこで、水野らは調査データから直接原点を通る回帰式を求めるという、以下の方法を提唱している⁴⁾。

「柳井らの方法」と異なり、定数項 α を除いて、報告有病率 y_i と回収率 R_i ($i=1, 2, \dots, k$) の間に、

$$y_i = \beta_2 R_i + e_i, \quad e_i \sim N(0, \sigma_e^2) \quad (16)$$

が成り立つと仮定する。ここで、 e_i は「柳井らの方法」と同様に誤差項であり、正規分布 $N(0, \sigma_e^2)$ に従うものとする。

母回帰係数 β_2 の推定値 b_2 は、

$$b_2 = \sum_i R_i y_i / \sum_i R_i^2 \quad (17)$$

である¹¹⁾。この場合、誤差分散 σ_e^2 の推定値 $\hat{\sigma}_e^2$ は、

$$\hat{\sigma}_e^2 = (\sum_i y_i^2 - b_2 \sum_i R_i y_i) / (k - 1) \quad (18)$$

となり、推定値 b_2 の分散は、

$$V(b_2) = \hat{\sigma}_e^2 / \sum_i R_i^2 \quad (19)$$

である¹¹⁾。

「水野らの方法」は、回収率を1とした場合の推定有病率を求め、それに全国人口 P を乗じて全国患者数を推定するものである。したがって、推定有病率 $\hat{\mu}$ は回帰係数 b_2 と一致する。

上記のことから、推定有病率 $\hat{\mu}$ の期待値と分散は、以下のように求められる。

推定有病率 $\hat{\mu}$ の期待値は、

$$E(\hat{\mu}) = E(b_2) = \beta_2 \quad (20)$$

と求められ、また推定有病率 $\hat{\mu}$ の分散は、

$$V(\hat{\mu}) = E(b_2 - \beta_2)^2 = V(b_2) \quad (21)$$

と求められる。したがって、推定有病率 $\hat{\mu}$ の分散は、回帰係数 b_2 の分散に一致することがわかる。

また、全国患者数 τ の推定値 $\hat{\tau}$ とその分散は、

$$\hat{\tau} = P \hat{\mu}, \quad V(\hat{\tau}) = P^2 V(\hat{\mu}) \quad (22)$$

により求められる。

3) 修正中江らの方法⁵⁾

回帰直線は「水野らの方法」と同様に、原点を通るように求めるが、地域別の調査対象の未回収分についての補正方法は、「柳井らの方法」と同様に行うという、上記2方法の中間的方法を考えてみよう。

報告有病率と回収率の回帰直線が原点を通るように定める点までは、「水野らの方法」と同一式、すなわち(16)～(19)式を用いるが、地域 i での推定有病率 $\hat{\mu}_i$ は「柳井らの方法」と同一式、すなわち、

$$\hat{\mu}_i = y_i + (1 - R_i) b_2 \quad (23)$$

により求めることにする⁵⁾。ここで回帰係数 b_2 は「柳井らの方法」と異なり「水野らの方法」と同一である。

地域 i の推定有病率 $\hat{\mu}_i$ の期待値は、

$$E(\hat{\mu}_i) = E(y_i) + (1 - R_i)E(b_2) = \beta_2 \quad (24)$$

と求められ、「水野らの方法」の推定有病率の期待値と一致する。また、推定有病率 $\hat{\mu}_i$ の分散は、

$$\begin{aligned} V(\hat{\mu}_i) &= E\{y_i + (1 - R_i)b_2 - \beta_2\}^2 \\ &= V(e_i) + (1 - R_i)^2 V(b_2) \\ &\quad + 2(1 - R_i)R_i V(e_i) / \sum_j R_j^2 \end{aligned} \quad (25)$$

と求められる。上式に(19)式を代入することで、

$$V(\hat{\mu}_i) = \hat{\sigma}_e^2 \{1 + (1 - R_i^2) / \sum_j R_j^2\} \quad (26)$$

と推定有病率 $\hat{\mu}_i$ の分散が求められる。

地域 i の推定患者数 $\hat{\tau}_i$ および全国推定患者数 $\hat{\tau}$ とそれぞれの分散は、「柳井らの方法」と同様に(14)式、(15)式により与えられる。

表3に、上記の3方法による各統計量の推定式をまとめて示した。なお、以下では「柳井らの方法」、「水野ら

の方法」,「修正中江らの方法」を,それぞれ「方法 Y」,「方法 M」,および「方法 N」のように略記する。

ところで,方法 Y では,報告有病率と回収率の相関係数が負の場合,推定有病率が報告有病率より小さくなり不都合である。そのため今回は,相関係数が負の場合は(7)式による回帰係数 b_1 を 0 とし,推定有病率の分散は報告有病率の分散により代用することにした。

III 結 果

1. レックリングハウゼン病の患者数推定

1) 都道府県別資料による患者数推定

レックリングハウゼン病について,表4に各診療科別,性別に都道府県別資料による報告有病率(人口10万

対)と回収率の各平均値・標準偏差,および両者の相関係数を示した。また,方法 Y で用いる単回帰直線の(7)式による回帰係数と切片,および方法 N と方法 M で用いる原点を通る回帰直線の(17)式による回帰係数を併せて示した。なお,精神科の男,耳鼻科の男女の3件で,報告有病率と回収率の相関係数が負となり,前述したように方法 Y での単回帰直線の回帰係数を 0 とした。

表4に示した各回帰式を用いて,各方法別に性別・診療科別に推定患者数と推定誤差を求め表5に示した。表5の「誤差」とは(15)式と(22)式による各推定患者数の分散の平方根をとったもの,すなわち推定値の標準誤差(推定誤差)を示す。したがって,誤差分布に正規性が仮定できれば,ほぼ[推定値 $\pm 2 \cdot$ 誤差]の範囲が患者

表3 患者数推定方法のまとめ

統計量	柳井らの方法 (方法 Y)	修正中江らの方法 (方法 N)	水野らの方法 (方法 M)
①推定有病率	$\hat{\mu}_i = y_i + (1 - R_i) b_1$	$\hat{\mu}_i = y_i + (1 - R_i) b_2$	$\hat{\mu}_i = \hat{\mu} = b_2$
②回帰係数	$b_1 = \Sigma_i (R_i - \bar{R}) (y_i - \bar{y}) / \Sigma_i (R_i - \bar{R})^2$	$b_2 = \Sigma_i R_i y_i / \Sigma_i R_i^2$	$b_2 = \Sigma_i R_i y_i / \Sigma_i R_i^2$
③残差分散	$\hat{\sigma}_1^2 = (1 - r_{yR}^2) \Sigma_i (y_i - \bar{y})^2 / (k - 2)$	$\hat{\sigma}_2^2 = (\Sigma_i y_i^2 - b_2 \Sigma_i R_i y_i) / (k - 1)$	$\hat{\sigma}_2^2 = (\Sigma_i y_i^2 - b_2 \Sigma_i R_i y_i) / (k - 1)$
④推定有病率の期待値	$\beta_1 + \alpha$	β_2	β_2
⑤推定有病率の分散	$V(\hat{\mu}_i) = \hat{\sigma}_1^2 \left\{ 1 + \frac{(1 - R_i)(1 + R_i - 2\bar{R})}{\Sigma_i (R_i - \bar{R})^2} \right\}$	$V(\hat{\mu}_i) = \hat{\sigma}_2^2 \{ 1 + (1 - R_i^2) / \Sigma_i R_i^2 \}$	—
⑥推定患者数	$\hat{\tau}_i = P_i \hat{\mu}_i$	$\hat{\tau}_i = P_i \hat{\mu}_i$	—
⑦推定値の分散	$V(\hat{\tau}_i) = P_i^2 V(\hat{\mu}_i)$	$V(\hat{\tau}_i) = P_i^2 V(\hat{\mu}_i)$	—
⑧全国患者数	$\hat{\tau} = \Sigma_i \hat{\tau}_i$	$\hat{\tau} = \Sigma_i \hat{\tau}_i$	$\hat{\tau} = P b_2$
⑨全国値の分散	$V(\hat{\tau}) = \Sigma_i P_i^2 V(\hat{\mu}_i)$	$V(\hat{\tau}) = \Sigma_i P_i^2 V(\hat{\mu}_i)$	$V(\hat{\tau}) = P^2 \hat{\sigma}_2^2 / \Sigma_i R_i^2$

記号は本文中で使用しているものに従っている。

表4 レックリングハウゼン病の都道府県別資料による回帰直線のあてはめ

診療科: 性	平均* 有病率 (SD)	平均 回収率 (SD)	相関係数	単回帰直線 回帰係数	切片	原点を通る 回帰直線**
皮膚科: 男	0.853 (0.576)	0.657 (0.169)	0.192	0.6528	0.4244	1.2590
女	0.947 (0.630)	"	0.178	0.6639	0.5112	1.3940
小児科: 男	0.455 (0.421)	0.695 (0.140)	0.144	0.4345	0.1529	0.6459
女	0.311 (0.245)	"	0.176	0.3082	0.0968	0.4421
整形外科: 男	0.529 (0.391)	0.542 (0.112)	0.172	0.6013	0.2035	0.9613
女	0.518 (0.431)	"	0.128	0.4912	0.2515	0.9361
眼科: 男	0.108 (0.158)	0.555 (0.169)	0.354	0.3303	-0.0757	0.2055
女	0.147 (0.209)	"	0.414	0.5103	-0.1365	0.2852
精神科: 男	0.086 (0.109)	0.634 (0.122)	-0.108	0.0000	0.0864	0.1279
女	0.085 (0.124)	"	0.134	0.1361	-0.0014	0.1340
脳外科: 男	0.233 (0.197)	0.623 (0.191)	0.316	0.3264	0.0292	0.3692
女	0.216 (0.201)	"	0.349	0.3672	-0.0128	0.3484
耳鼻科: 男	0.079 (0.133)	0.607 (0.162)	-0.037	0.0000	0.0788	0.1193
女	0.092 (0.182)	"	-0.205	0.0000	0.0919	0.1261

* 有病率は各都道府県別人口(昭和60年10月1日現在)10万人対。

** 原点を通る回帰直線の回帰係数。

数の95%信頼区間を与える。ただし、信頼区間の下限が報告患者数未満の場合、分布に偏りがあると考えられるので、そのような信頼区間を採用すべきではない。

表5に示したように、推定患者数は方法Yで男2035.5、女2112.9と最小となり、方法Mでは男2192.2、女2257.7、方法Nで男2344.0、女2399.7と最大となった。

推定誤差は方法Nで男104.5、女107.6と最小であり、方法Yで男109.7、女112.9、方法Mで男117.0、女125.5と最大であった。なお、方法Yでは報告有病率と回収率の相関が負の場合、推定患者数と推定誤差は低めに計算されている。

診療科別の推定患者数と推定誤差をみても、上記の合計の結果とはほぼ同様の傾向が認められる。とくに推定誤差については、相関係数が負の場合以外は方法Nが最小であった。また、小児科の男以外は、方法Mの推定誤差がすべて最大であった。

2) 地域ブロック別資料による患者数推定

表4と同様に、表6に各診療科別、性別にレックリングハウゼン病の地域ブロック別資料による報告有病率と回収率の平均値・標準偏差、および相関係数を示し、あわせて患者数推定のための回帰係数などを示した。

報告有病率と回収率の相関係数が負となる場合は、小児科の男女、整形外科の男女、精神科の男、耳鼻科の男

表5 レックリングハウゼン病の都道府県別資料による方法別・診療科別・性別患者数の推定値と誤差

診療科	(1)柳井らの方法 (方法Y)		(2)修正中江らの方法 (方法N)		(3)水野らの方法 (方法M)	
	男 推定値(誤差)	女 推定値(誤差)**	男 推定値(誤差)	女 推定値(誤差)	男 推定値(誤差)	女 推定値(誤差)
皮膚科	732.6(70.9)	803.7(78.7)	868.7(69.6)	972.7(77.5)	748.4(74.3)	858.5(84.7)
小児科	309.1(52.6)	249.5(30.8)	348.6(50.5)	275.4(29.6)	384.0(51.7)	272.3(31.0)
整形外科	497.7(53.9)	483.3(60.6)	596.2(47.3)	609.5(53.2)	571.4(61.3)	576.5(70.5)
眼科	186.5(19.0)	239.2(24.8)	154.2(18.2)	178.7(24.0)	122.1(22.6)	175.6(30.4)
精神科	54.0(13.0)*	76.7(16.3)	82.9(13.5)	76.2(15.1)	76.0(15.1)	82.5(17.3)
脳外科	207.7(23.4)	213.5(23.9)	217.0(22.7)	209.2(23.1)	219.4(25.2)	214.6(26.2)
耳鼻科	48.0(15.9)*	47.0(22.0)*	76.3(16.4)	78.0(23.1)	70.9(18.9)	77.7(27.1)
合計	2035.5(109.7)	2112.9(112.9)	2344.0(104.5)	2399.7(107.6)	2192.2(117.0)	2257.7(125.5)

* 報告有病率と回収率の相関が負のため、回帰係数を0とし、誤差は報告有病率の標準偏差にもとづいて計算。

**「誤差」とは、推定値の分散の平方根を示す。

表6 レックリングハウゼン病の地域ブロック別資料による回帰直線のあてはめ

診療科:性	平均* 有病率 (SD)	平均 回収率 (SD)	相関係数	単回帰直線 回帰係数 切片		原点を通る 回帰直線**
皮膚科:男	1.009(0.308)	0.646(0.048)	0.428	2.7569	-0.7726	1.5679
女	1.062(0.404)	"	0.509	4.2966	-1.7148	1.6576
小児科:男	0.480(0.180)	0.693(0.037)	-0.012	0.0000	0.4805	0.6914
女	0.347(0.245)	"	-0.138	0.0000	0.3469	0.4982
整形外科:男	0.592(0.170)	0.531(0.037)	-0.487	0.0000	0.5918	1.0981
女	0.625(0.234)	"	-0.523	0.0000	0.6247	1.1545
眼科:男	0.135(0.089)	0.560(0.060)	0.373	0.5556	-0.1765	0.2443
女	0.155(0.084)	"	0.383	0.5364	-0.1457	0.2794
精神科:男	0.098(0.060)	0.649(0.069)	-0.068	0.0000	0.0984	0.1492
女	0.105(0.081)	"	0.299	0.3480	-0.1213	0.1632
脳外科:男	0.263(0.135)	0.645(0.095)	0.682	0.9713	-0.3633	0.4200
女	0.238(0.095)	"	0.770	0.7734	-0.2612	0.3771
耳鼻科:男	0.069(0.049)	0.587(0.081)	-0.385	0.0000	0.0690	0.1110
女	0.073(0.053)	"	-0.007	0.0000	0.0728	0.1215

* 有病率は地域ブロック別人口(昭和60年10月1日現在)10万人対。

** 原点を通る回帰直線の回帰係数。

女の7件と増加した。

表7に各方法による推定患者数と推定誤差を示した。男の推定患者数は、方法Yで2444.6と最小であり、方法Nで2470.3、方法Mで2545.2と最大となった。女の推定患者数は方法Nで2532.9と最小であり、方法Mで2618.2、方法Yで2901.0と最大であった。

表5に示した都道府県別の場合と比べると、相関係数が負の場合を除いて、地域ブロック別による推定患者数の方が大きかった。しかし、方法Nでは、両者の推定患者数の差は他の2方法に比べ比較的小さかった。

推定誤差は、方法Nで男119.8、女141.5と最小であり、方法Mで男142.7、女174.7、方法Yで男229.9、女281.7と最大となった。また、性別・診療科別に各方法の推定誤差を比較すると、相関係数が負の場合以外は方法Nで最小であった。

推定誤差は、表5の都道府県別の場合と診療科別・性別に比較すると、相関係数が負の場合を除いて地域ブ

ック別で大きい傾向がある。ただし、脳外科の女で方法Nと方法Mの誤差は小さくなっていた。表4と表6の脳外科の女の相関係数を比べると、地域ブロック別によることで、0.349から0.770と数値がきわめて大きくなっており、逆に報告有病率の標準偏差は0.201から0.095と半減していた。また、方法Nでは眼科の女でも誤差は減少しており、この場合、相関係数の値は0.414から0.383と若干地域ブロック別の方が小さいが、報告有病率の標準偏差は0.209から0.084と小さくなっていた。

2. 神経性食思不振症の患者数推定

1) 都道府県別資料による患者数推定

表8に各診療科別、性別に都道府県別資料による報告有病率と回収率の各平均値・標準偏差、および相関係数を示した。また、患者数推定のための回帰係数などを併せて示した。なお、小児科と産婦人科の女の2件で、報告有病率と回収率の相関係数が負となった。

表9に推定患者数と推定誤差を示した。男の推定患者

表7 レックリングハウゼン病の地域ブロック別資料による方法別・診療科別・性別患者数の推定値と誤差

診療科	(1)方法 Y		(2)方法 N		(3)方法 M	
	男 推定値(誤差)	女 推定値(誤差)**	男 推定値(誤差)	女 推定値(誤差)	男 推定値(誤差)	女 推定値(誤差)
皮膚科	1192.4(206.4)	1628.1(261.9)	930.9(80.2)	1027.3(106.1)	932.0(92.1)	1020.8(124.2)
小児科	228.0(47.2)*	190.0(28.7)*	355.2(50.6)	284.9(31.6)	410.9(55.2)	306.8(35.1)
整形外科	330.0(44.4)*	344.0(62.4)*	632.0(57.0)	670.2(77.9)	652.7(76.1)	711.0(106.2)
眼科	244.4(61.0)	245.6(58.0)	164.0(24.7)	176.8(23.3)	145.2(31.6)	172.0(30.5)
精神科	54.0(15.7)*	126.2(42.2)	87.7(17.4)	83.1(22.4)	88.7(19.8)	100.5(26.0)
脳外科	344.8(43.9)	302.2(27.4)	226.3(31.4)	214.0(20.4)	249.7(36.0)	232.2(23.8)
耳鼻科	48.0(12.8)*	47.0(14.2)*	74.2(15.2)	76.7(15.8)	66.0(18.8)	74.8(19.9)
合計	2444.6(229.9)	2901.0(281.7)	2470.3(119.8)	2532.9(141.5)	2545.2(142.7)	2618.2(174.7)

* 報告有病率と回収率の相関が負のため、回帰係数を0とし、誤差は報告有病率の標準偏差にもとづいて計算。

** 「誤差」とは、推定値の分散の平方根を示す。

表8 神経性食思不振症の都道府県別資料による回帰直線のあてはめ

診療科：性	平均有病率 (SD)	平均回収率 (SD)	相関係数	単回帰直線 回帰係数	切片	原点を通る 回帰直線**
内科：男	0.059(0.122)	0.622(0.112)	0.091	0.0997	-0.0026	0.0956
女	1.019(1.398)	〃	0.244	3.0446	-0.8741	1.6832
精神科：男	0.054(0.086)	0.669(0.119)	0.083	0.0606	0.0130	0.0795
女	1.049(0.877)	〃	0.340	2.5073	-0.6288	1.5966
小児科：男	0.036(0.066)	0.763(0.137)	0.161	0.0776	-0.0229	0.0485
女	0.397(0.309)	〃	-0.006	0.0000	0.3968	0.5034
産婦人科：女	0.361(0.441)	0.645(0.139)	-0.071	0.0000	0.3605	0.5238
心療内科：男	0.012(0.035)	0.106(0.168)	0.467	0.0972	0.0012	0.1004
女	0.266(0.683)	〃	0.418	1.7013	0.0855	1.9314

* 有病率は各都道府県別人口（昭和60年10月1日現在）10万人対。

** 原点を通る回帰直線の回帰係数。

数は方法 *M* で192.6と最小となり、方法 *Y* で208.2、方法 *N* で208.3と最大となった。女の推定患者数は、方法 *M* で3841.9と最小となり、方法 *N* で4261.1、方法 *Y* で4472.8と最大となった。

推定誤差は方法 *N* で男20.3、女230.6と最小であり、方法 *Y* で男22.0、女246.2、方法 *M* で男27.2、女420.6と

最大であった。

診療科別の推定患者数と推定誤差をみても、レックリ
ングハウゼン病と同様に、上記の推定患者数の合計の結果とほぼ類似した傾向が認められる。推定誤差については、相関係数が負の場合と小児科の男以外は、方法 *N* が最小であった。小児科の男では、方法 *M* の推定誤差

表9 神経性食思不振症の都道府県別資料による方法別・診療科別・性別患者数の推定値と誤差

診療科	(1)方法 Y		(2)方法 N		(3)方法 M	
	男 推定値(誤差)	女 推定値(誤差)**	男 推定値(誤差)	女 推定値(誤差)	男 推定値(誤差)	女 推定値(誤差)
内科	52.1(16.3)	1391.9(184.0)	51.1(14.8)	1052.5(167.8)	56.8(16.9)	1036.6(196.0)
精神科	58.5(11.2)	1372.2(108.2)	62.4(10.4)	1178.1(101.9)	47.2(11.1)	983.2(111.1)
小児科	38.0(8.1)	256.0(37.3)*	33.5(7.9)	336.5(38.7)	28.9(7.4)	310.0(37.2)
産婦人科	0	286.0(53.3)*	0	405.5(55.5)	0	322.6(62.2)
心療内科	59.6(5.1)	1166.8(104.1)	61.3(4.9)	1288.4(100.3)	59.7(16.6)	1189.5(347.7)
合計	208.2(22.0)	4472.8(246.2)	208.3(20.3)	4261.1(230.6)	192.6(27.2)	3841.9(420.6)

* 報告有病率と回収率の相関が負のため、回帰係数を0とし、誤差は報告有病率の標準偏差にもとづいて計算。

** 「誤差」とは、推定値の分散の平方根を示す。

表10 神経性食思不振症の地域ブロック別資料による回帰直線のあてはめ

診療科：性	平均 有病率* (SD)	平均 回収率 (SD)	相関係数	単回帰直線 回帰係数 切片		原点を通る 回帰直線**
内科：男	0.066(0.054)	0.622(0.047)	0.709	0.8070	-0.7726	0.1103
女	0.957(0.506)	〃	-0.055	0.0000	0.9566	1.5247
精神科：男	0.081(0.044)	0.691(0.089)	0.026	0.0131	0.0720	0.1155
女	1.344(0.558)	〃	0.167	1.0504	0.6180	1.9296
小児科：男	0.045(0.026)	0.754(0.039)	0.223	0.1503	-0.0686	0.5952
女	0.376(0.154)	〃	0.195	0.7750	-0.2088	0.4988
産婦人科：女	0.439(0.239)	0.655(0.042)	-0.213	0.0000	0.4391	0.6631
心療内科：男	0.009(0.013)	0.102(0.087)	0.273	0.0412	0.0048	0.0687
女	0.266(0.347)	〃	0.349	1.3906	0.1238	2.0926

* 有病率は地域ブロック別人口（昭和60年10月1日現在）10万人対。

** 原点を通る回帰直線の回帰係数。

表11 神経性食思不振症の地域ブロック別資料による方法別・診療科別・性別患者数の推定値と誤差

診療科	(1)方法 Y		(2)方法 N		(3)方法 M	
	男 推定値(誤差)	女 推定値(誤差)**	男 推定値(誤差)	女 推定値(誤差)	男 推定値(誤差)	女 推定値(誤差)
内科	221.2(30.0)	633.0(134.7)*	54.4(14.4)	1010.1(149.4)	65.6(17.0)	939.0(179.7)
精神科	48.7(18.4)	1061.5(233.6)	69.8(12.6)	1248.6(157.7)	68.7(13.6)	1188.3(173.5)
小児科	49.2(16.4)	379.6(99.0)	35.2(7.0)	335.6(42.1)	35.4(7.1)	307.2(43.7)
産婦人科	0	286.0(63.5)*	0	436.8(70.7)	0	408.4(81.7)
心療内科	31.3(12.7)	1010.7(330.5)	45.4(9.0)	1385.6(236.1)	40.8(20.4)	1288.8(542.7)
合計	350.4(40.9)	3370.9(442.5)	204.9(22.3)	4416.7(331.2)	210.5(30.6)	4131.6(604.6)

* 報告有病率と回収率の相関が負のため、回帰係数を0とし、誤差は報告有病率の標準偏差にもとづいて計算。

** 「誤差」とは、推定値の分散の平方根を示す。

が最小であり、この場合、相関係数は0.161と小さいが、回収率は0.763と大きく、また報告有病率の標準偏差は0.066と小さかった。しかし、心療内科の男では相関係数は0.467と大きく、報告有病率の標準偏差も0.035と小さいが、回収率は0.106ときわめて低く、方法Mによる推定誤差は他の方法より大きくなっていた。

2) 地域ブロック別資料による患者数推定

表8と同様に、表10に各診療科別、性別に地域ブロック別資料による報告有病率と回収率の平均値・標準偏差、および相関係数を示し、あわせて患者数推定のための回帰係数などを示した。

報告有病率と回収率の相関係数が負となる場合は、内科と産婦人科の女の2件であった。

表11に推定患者数と推定誤差を示した。男の推定患者数は、方法Nで204.9と最小であり、方法Mで210.5、方法Yで350.4と最大となった。女の推定患者数は方法Yで3370.9と最小であり、方法Mで4131.6、方法Nで4416.7と最大であった。

表9に示した都道府県別の場合と比べると、方法Nでは両者の推定患者数の差は比較的小さいが、他の2方法では推定患者数の差は大きかった。

推定誤差は方法Nで男22.3、女331.2と最小であり、方法Yで男40.9、女442.5、方法Mで男30.6、女604.6であった。診療科別・性別に各方法による推定誤差を比較すると、これまでの結果と同様に、相関係数が負の場合以外はすべて方法Nで最小となった。

表9の都道府県別の場合と推定誤差を比較すると、相関係数が負の場合を除いて、地域ブロック別で誤差が大きい傾向がある。ただし、小児科の男で方法Nと方法Mによる推定誤差は、地域ブロック別の場合の方が小さかった。表8と表10の小児科の男の相関係数を比べると、地域ブロック別にすることで0.161から0.223と数値は若干大きくなっており、報告有病率の標準偏差も0.066から0.026と減少していた。また、方法Nでは内科の男でも推定誤差は減少していた。この場合、相関係数は0.091から0.709ときわめて増大し、報告有病率の標準偏差も0.122から0.054と減少していた。

IV 考 察

厚生省特定疾患の全国疫学調査にもとづく全国患者数推定方法についての簡単な解説と、推定誤差の計算方法を提示し、実際の調査資料により推定結果と推定誤差の比較を行った。ここでは、上記の結果を踏まえて、いくつかの点に関して考察を行う。

なお、本報告で示した数値は、あくまでも一次調査にもとづくものを取り上げたに過ぎない。厚生省特定疾患全国疫学調査は、本例のような一次調査だけでなく、一

次調査の結果を踏まえて「二次調査」を行うのが一般的な手順である。「二次調査」は、一次調査で患者ありと報告のあった診療科を調査対象とし、患者の性、年齢、職業、診断の確実性、日常生活、臨床症状などを調査するために設定された質問紙を郵送し、主治医の記載後に返送するように依頼するという、郵送法による調査である。

難病の患者数推定には、上記の二次調査にもとづく重複例の除去、疑い例の検討、未受診者の考慮などを加える必要があることは言うまでもないだろう。

1. 地域区分の相違による推定患者数と推定誤差の変動について

レックリングハウゼン病に関する表5と表7、および神経性食思不振症に関する表9と表11に、それぞれ示した全国患者数推定結果の比較では、方法によらず都道府県別の場合の方が、地域ブロック別の場合に比べ推定誤差は小さくなっていた。また、都道府県別と地域ブロック別による推定患者数は、すべての方法で2疾患とも比較的大きな差が認められた。しかし、方法Nでは2疾患ともに、他の2方法に比べて差は小さかった。この理由として、前述のように、方法Nは方法Mと方法Yの中間的な形式を採用しているためとも考えられるが、必ずしも理論的に明らかなものではなく、実際例で今後さらに検討する必要があるだろう。

ところで、推定誤差の大小は、表3の③残差分散と⑤推定有病率の分散の各式から、報告有病率の分散もしくは2乗和、報告有病率と回収率の相関係数もしくは積和、標本数（地域区分の個数）、そして各地域の回収率の各影響を受けることがわかる。簡単にいえば、相関係数は正で大きいほど、報告有病率の変動は小さいほど、各地域の回収率は1に近いほど、そして標本数は大きいほど、推定誤差は小さくなると考えてよいだろう。しかし、これらの数値は相互に関係しており、すべてを満たすように都道府県別資料を地域ブロック別にまとめることは、必ずしも容易ではないと思われる。実際、応用例でみたように、地域ブロックにまとめたため相関係数がきわめて大きくなる場合もあるが、逆に負の相関を示す場合もある。報告有病率と回収率の標準偏差は、地域ブロック別の場合の方がすべての場合で小さくなったが、性別・診療科別の推定誤差が減少したものは少数であった。推定誤差の大小は上記の各項目の関係に依存しているため、1つの項目のみを取り上げるのは適切ではないだろう。しかし、応用例からも、報告有病率の標準偏差が小さいこと、相関係数が正で大きいこと、回収率が全体に大きいことなどは、推定誤差を小さくするための必要条件であることは明らかと思われる。

本報告で扱った疾患に関しては、都道府県別資料を地

域ブロック別にまとめても、患者数推定が良好に行えるとは限らなかった。この点からいえば、地域ブロック別による推定は、都道府県別による場合よりも推定誤差が小さく、かつ相関係数が正の場合にのみ採用すべきであろう。各調査資料にもとづいて、両者での分析を一通り行ってから、最終的にどちらを選択するかを決める必要があるものと考えられる。

地域区分をどのようにすれば、より推定誤差の少ない患者数推定が行えるかは、患者数推定の理論的方法論とともに、今後の重要な問題であろう。

2. 各地域の患者数推定について

本報告では、都道府県別や地域ブロック別の推定患者数を示さなかった。方法M以外は、地域別の推定患者数を求めて、その合計から全国患者数を算出している。したがって、各地域別の推定患者数を比較することはきわめて容易である。しかし、個々の地域での推定患者数もしくは推定有病率を求めて比較することは、以下の理由から意味があるとは考えられない。

方法Yは、原論文では推定有病率と回収率の相関がなくなるように、その推定式が設定されたと述べられている¹⁾。これは理論的事実であるが、より重要な点は表3で明らかなように、有病率に関して、「各地域の推定有病率の期待値は一定であり、各地域の真の有病率は等しい」ことが暗黙に仮定されていることである。この有病率に地域差がないという非明示的な仮定は、他の2方法についても同様である。

真の有病率が地域によらず一定ならば、推定有病率と回収率が無相関になることは当然期待されることである。上述のように、真の有病率に地域差がないと仮定される場合に、各地域の推定有病率の比較に意味があるとは考えにくい。また、現実問題として各地域の推定誤差は、表に示していないが無視できるほど小さくはない。

ところで、真の有病率の推定方法として、真の有病率に地域差がなければ、各地域の推定有病率の平均値により推定するのが最も簡単な方法と考えられる。しかし、今回は各方法の比較を行うことに重点を置いており、この問題には触れない。しかし、有病率の地域差がないという仮定の妥当性の問題とともに、今後の課題として検討を加える必要があるだろう。

3. 理論的問題点と他の推定方法

本研究で取り上げた各手法は、特定疾患の患者数推定方法として比較的よく用いられていると思われるが、これらの方法に共通する問題もある。すなわち、方法Yでは前述のように、報告有病率と回収率の相関が負の場合、推定は困難であり、患者数の推定結果は意味があるとはいえなくなる。負の相関を示した場合、本報告では報告患者数と推定患者数が等しくなるようにしたが、こ

れは便宜的な方法であり、一般的なものではない。他の2方法では、相関係数が負の場合であっても患者数推定は形式的には可能である。しかし、推定が可能であることと推定結果が妥当であることは別の問題である。確かに、原点を通る回帰直線を用いるので、相関が負の場合の推定上の不都合を避けることはできるが、そのような場合に回帰直線を求めることに妥当性があるとは思えない。回帰分析は予測のための手法であり、2変数間の相関係数が小さければ予測値の誤差は急激に増加する。さらに、相関係数が正でなければ原点を通る直線を求めても、それは単に形式的意味しかない。本報告では、方法Nによる患者数推定が最も良い結果を与えていると考えられるが、今後は回帰によらない推定方法も考慮すべきであろう。

Takagi¹²⁾は報告のあった各施設の患者数の分布に着目し、多項分布にもとづく患者数の最尤推定法を提示しており、別法として多重超幾何分布による最尤推定法についても触れている。本報告と若干資料の扱いは異なるが、レックリングハウゼン病に関する多項分布にもとづく全国患者数の推定値と誤差は、都道府県別資料による方法Nの結果とはほぼ近い値であった¹²⁾。しかし、理論的な関連については明らかではなく、今後検討する必要があるものと考えられる。また、全国疫学調査によらない他の既存資料にもとづく患者数推定の手法との比較も、当然今後の課題となるであろう。

実際に推定値の良さを問題にする場合、推定誤差とともに、その偏りについても評価する必要があるだろう。しかし、本研究で扱った患者数推定に関しては、患者の発見率（患者が対象施設を受診する確率に関する分布など）に適当な仮定を置かないと、推定上の偏りを評価できないため、この点に関しては触れていない。

また実際問題として、我々は真の患者数を知ることにはできない。したがって、より信頼性の高い調査法を考案し、実施するとともに、理論的に最も妥当と思われる方法を考える必要がある。問題を抱えた方法を適当に修正するだけでは、結果が相互に類似しているからといって、疫学研究上満足することはできないだろう。今後は報告有病率などにどのような仮定を置けばよいか、また未回収が起こるメカニズムをどう捉えるかを検討し、それらを踏まえて、偏りが小さく推定誤差も小さい患者数推定方法を開発する必要があるだろう。

V 要 約

厚生省特定疾患全国疫学調査一次調査による地域別の報告有病率と回収率にもとづく患者数推定方法のうち、「柳井らの方法」、「修正中江らの方法」、「水野らの方法」について簡単な説明とともに、全国患者数の推定値の推

定誤差の理論式を初めて提示した。

実際例として、昭和61年に実施した厚生省特定疾患であるレックリングハウゼン病および神経性食思不振症の一次調査資料を用い、都道府県別と地域ブロック別の2種類の地域区分により、各方法による患者数推定値とその推定誤差を求め比較した。

2疾患の患者数推定結果は、いずれも「修正中江らの方法」で推定誤差が最小となり、また地域区分の相違による推定患者数の変化も最小であり、安定した推定結果を与えた。さらに、2疾患とも地域区分として都道府県別による全国患者数推定の方が、すべての方法で地域ブロック別による場合に比べ推定誤差が小さかった。

今後の課題として、回帰式による患者数推定に付随するいくつかの問題点を指摘し、併せて他の患者数推定方法に関しても若干の考察を加えた。

本論文の一部は第46回日本公衆衛生学会総会（1987、長崎市）において発表した。

なお、本研究の一部は、昭和60、61、62年度厚生省特定疾患事業補助金による。（受稿 昭62. 12. 16）

文 献

- 1) 柳井晴夫，他：疫学調査における有病率補正の方式，日本公衛誌，22，71～81，1975
- 2) 西川秋佳，他：特発性門脈亢進症の有病率に関する疫学的検討，日本公衛誌，34，85～88，1987
- 3) 中江公裕，他：膠原病四疾患第一次全国疫学調査成績；全国患者数の推定，厚生省特定疾患難病の疫学調査研究班，昭和59年度研究業績報告書，135～150，1985
- 4) 水野正一，他：全国調査成績を用いての有病数の一推計方法，日本公衛誌，33(10)，255，1986
- 5) 高木廣文，他：レックリングハウゼン病と結節性硬化症の全国受診患者数の推定について，厚生省特定疾患神経皮膚症候群調査研究班 昭和61年度研究報告書，15～19，1987
- 6) 新村真人，他：神経皮膚症候群全国実態調査・第一次調査，厚生省特定疾患神経皮膚症候群調査研究班 昭和61年度研究報告書，5～9，1987
- 7) 稲葉 裕，他：神経性食思不振症の患者数の推定—昭和60年全国調査を資料として—，厚生省特定疾患神経性食思不振症調査研究班 昭和61年度研究報告書，55～61，1987
- 8) 稲葉 裕，他：神経性食思不振症の全国調査による患者数の推定，厚生省特定疾患難病の疫学調査班 昭和61年度研究業績報告書，22～24，1987
- 9) 厚生統計協会：国民衛生の動向，33(9)，357，1986
- 10) 豊川裕之，柳井晴夫編著：医学・保健学の例題による統計学（初版），50～59，現代数学社（京都），1982
- 11) Snedecor, G. W. and Cochran, W. G.: Statistical Methods (6th edition), the Iowa State Univ. Press, Ames., 1967；畑村又好，他共訳：スネデカー，コ克蘭 統計的方法 原書第6版，131～165，岩波書店（東京），1972
- 12) Takagi, H.: A mode prediction for the number of patients with an incurable disease in the nationwide epidemiological survey, Bulletin of the Biometric Society of Japan, 8(2), 51～60, 1987

ESTIMATION ERRORS WITH RESPECT TO THE ESTIMATED NUMBER OF PATIENTS IN THE NATIONWIDE EPIDEMIOLOGICAL SURVEY ON INTRACTABLE DISEASES

Hirofumi TAKAGI

(Department of Statistics, St. Luke's College of Nursing)

Tosiya SATO

(Department of Epidemiology, School of Health Sciences, University of Tokyo)

Yutaka INABA

(Department of Hygiene, Juntendo University School of Medicine)

For the purpose of comparing methods used in estimating the number of patients with a specific disease based on reported prevalence and recovery rates for each regional division in the nationwide epidemiological survey on intractable diseases, we developed theoretical equations to calculate variances of the estimates (estimation errors and/or standard errors) by using the methods of Yanai et al. and Mizuno et al., and the modified method of Nakae et al., respectively.

Using the above three methods, we estimated the number of patients with Recklinghausen disease, and also with anorexia nervosa respectively, and evaluated the corresponding estimation errors. The data were treated in one of two ways according to the location of each hospital: i.e., by dividing it into 1) forty-seven prefectures, or into 2) nine districts.

Examination of the results obtained by the three methods showed that the modified method of Nakae et al. gave the minimum estimation error. It also provided the least variation of the estimates between the different data treatments according to regional divisions.

Furthermore, it was found in all cases that using the data of forty-seven prefectures provided smaller standard errors than using that of nine districts.

Finally, several common problems with estimation methods based on the linear regression and other methods that are possible for our study are discussed.