

## 細分課題5

### サーベイランス法，とくにマーカーに関する研究

#### 5・1 情報理論によるマーカーの格付け

新潟大学医学部衛生学教室

渡 辺 巖 一  
高 山 洋 一

#### ま え が き

Internatinal Clearinghouse for Birth Defects Monitoring (以下ICBDMと略す)へ1977年までに参加しているモニタリング・プログラムは、12カ国、14であった。これに加え1978年後半から、東京都日赤病院のプログラムが登録された。1979年スペイン、1980年イタリーが、それぞれ病院を基盤とするプログラムでICBDMへ参加する予定である。さらに、デンマーク、ニュージーランドの2国は、国家規模で先天異常の登録とモニタリングを行うべく、目下準備中である。

このような事情で、わが国が先天異常のモニタリング・プログラムを、より本格的に推進するにあたっては、つぎに挙げる諸点を、事前に十分検討しておく必要がある。その第1は、モニタリング・システム(以下M.S.と略す)の適切な大きさ(年間と出生数と資料収集の地域的拡がり)である。第2は、M.S.に必須な主要先天異常の種類(マーカー)の決定である。なお、ICBDMでは従来12のマーカーを設定していた。ただし、1978年度より先天股脱を除外し、上肢と下肢の減奇形は一つに纏めて報告することとし、さらに臍ヘルニヤ(皮膚におおわれていないもの)を新しく追加することとした。これらが果して、適切なものであるか否かを、検討しておく必要がある。第3は、この国で警報発令をする際、如何なる基準、水準に達したとき、それをなすべきかの問題である。以上の諸点について、本研究班が一定の見解をもつことは、重要なことと思料される。

## 研 究 目 的

まえに述べた3点について、各国のM.S.は未だ模索中であり、一致した基準と方法を採用する合意に達していない。したがって、各M.S.相互間の比較が、必ずしも容易でない。

いまマーカーの選択をとりあげてみるに、発生件数の多いものは、誤差を少くする点で信頼度を増すが、鋭敏度において欠けるところがある。逆に、発生件数の少ない奇形は、その増減判定に誤差が多い半面、鋭敏度は優れている。このように、マーカーの選択には、二律相反する矛盾なしとしない。

また、M.S.の大きさについて云えば、年間出生数の少ない小規模のものが不適なことは明らかである。さりとて、規模が大きければ、それだけ良質なものとなる、とは言い難い。年間出生数を拡大すれば、資料収集の地域は広範となり、地域特性が平均化によって隠れてしまう。また、M.S.へ参加する医師、助産婦の数が増加するため、診断基準の確保が困難になる。ここに自づと適正規模の問題がある。

本研究は、以上の2点について、情報理論の手法を用い解析を試みたので、ここに報告をする。

## 研 究 方 法

熱力学におけるエントロピーの概念は、今日情報科学の中核をなし、情報の質、換言すれば、信頼度、确实性を測る尺度として用いられている。

情報理論において、エントロピー(H)は、つぎのように定義される(単位はbit)。

$$H \equiv - \sum_{i=1}^N P_i \log_2 P_i, \text{ ただし } \sum_{i=1}^N P_i = 1$$

またその最大値は

$$H_{\max} = \log_2 N$$

であり、その差

$$\Delta H \equiv H - H_{\max}$$

は、情報の不確実さを知る量である。そして、分布に対しての適用を考えると、全くの random 状態、すなわち一様分布の際、不確実さは最大であるとし、そのときのエントロピーが  $H_{\max}$  となる数学モデルが一般である。しかし、実際の適用にあたっては、数値を分布に変換する操作が若干繁雑である。この点を解消するため、上記と逆に、観測値の変動が全くないとき、 $H = H_{\max}$  になるモデルを考案した。すなわち、ICBDM の四季毎を 1 期とし、 $N$  期にわたる資料について、各記号をつぎのようにする。

$A_i$        $i$  期における奇形 A の発生件数

$L_i$        $i$  期における観察した出生数

(ただし、 $A_i \in L_i$ )

$R_i$        $i$  期における奇形発生率

ここで、以下のように  $P_i$  を定義すると、この  $P_i$  ( $i=1, \dots, N$ ) はエントロピー・モデルを満足する。

$$P_i = \frac{R_i}{\sum R_i}, \quad \text{ただし } R_i = \frac{A_i}{L_i}$$

これは、エントロピー・モデルとして最も単純な型であるが、その特徴は明らかで、 $A_i$  の変動のうち、 $L_i$  の変動によるものは、 $L_i$  に対して補正されている。そのため母数の変動を考慮せず、各奇形の母数  $L_i$  によらない変動を、そのエントロピーによって同列に比較することが可能である。なお、資料としては、1973～77 年にわたり ICBDM が 14 プログラムについて作成したモニタリング・レポート（1 年を 4 期に分け集計している）を用いた。

## 研 究 成 果

$\Delta H$  は、変動の大小を表わすので、 $\Delta H$  の大きい順（変動の少ない順）に 12 奇形をならべ、各プログラム間の比較を行った。ただし、 $\Delta H$  値そのものは、プログラムによって各奇形のベースラインが異なるため、直接の比較に耐えない。しかし、各プログラムにおける各種奇形のエントロピー順位は、そのプログラ

ム内における共通の事象として、ある程度同じような基盤をもつと考えられる。そこで、その順位につき、奇形の分類をノンパラメトリックに行った。結果は、表1に示すとおりである。

表1から分るように、エントロピーによれば、尿道下裂、口唇裂、ダウン症候群などは、変動の小さい群へ入る。また、食道、および直腸肛門の狭窄・閉鎖、上肢と下肢の減奇形などは、変動の大きい組へ分類された。このような分類になった要因として考えられるものを挙げれば、つぎのとおりである。1) 発生率自体の変動、2) 診断の難易による人為的変動、3) 発生件数の多寡による効果、などがそれである。1)と2)については、資料のみから弁別することができない。しかし3)については、2、3の奇形がこの組分に当てはまらない。したがって、一般に想定されるよう、標本数を大にすれば、プログラムの質が改善されるとする立場は、この場合必ずしも正しくない。例えば、先天股脱は、各プログラムとも最多発件数であるにも拘らず、その順位は上位でない。

一方、外表に出現し、生下時診断の容易な口唇裂、尿道下裂は上位に、逆に生下時直ぐ診断をつけ難い消化管の狭窄・閉鎖は下位にある。そして、診断基準が程度により不確実になりがちな上肢・下肢の減奇形も、下位に分類されている。このことは2)の人為的要因の関与を容認するものと解することができる。この場合、生下時診断困難とされる先天股脱が、下位に組分けされて然るべきであるのに、いづれとも確定されない群に入っている。また、同じ群に入っている無脳症は、生下時診断が比較的容易である。この矛盾する両奇形は、1)と2)、すなわち、発生率自体の変動が、発生件数の多い効果に影響されて、いづれとも確定しえない組分に属してしまっただけであろう。これらの諸点については、さらにモデルのたて方を吟味する必要がある。

つぎに各奇形のエントロピーを合せ、全体としてこの変動の大小を、各プログラムにつき比較したのが表2である。

表2の結果へ影響を与えるものとして、ここでも表1の場合と同じように、1)各マーカーの発生率自体の大小、2)各国のプログラムにおける診断基準や資料収集法の差、3)標本集団の大きさの差異、などを挙げるができる。そして表2の場合、第3の要因が支配的であるように思われる結果となった。そこで、図1に、1期あたり平均標本数(出生数)に対して、各プログラムにおけ

る各奇形の $\Delta H$ の総和を図示してみた。果せるかな、総 $\Delta H$ が標本数に依存する有様をみる結果となった。

ここで興味あることは、標本数1万前後で曲線がほぼ plateau に達し、その後標本数がたとえ増加しても、総 $\Delta H$ は、ごく僅かしか増さない事実である。換言すれば、標本数(3カ月間の出生件数)を1万以上にしても、さほど情報の改善はみられないことになったのである。

資料の異常検出力に関して、必要以上の広範に過ぎる調査は、地域差の相殺化と平均化によって、鋭敏さを減殺すると考えることができ、図1の結果は納得のいくものである。すなわち、適切なM.S.の大きさは、年間出生数4万前後となり、出生率を15対1,000とするとき人口250万と計算される。

標本の大きさに関しては、奇形発生がPoisson分布に適合するとの観点から計算し、出生数1万/年が適当であるとする遠藤班員の知見がある。本研究でえた4万/年の数値は、曲線の肩付近におけるプロットが少ないため、精度が多少低いことも考えらぬではない。いづれにしても両者の相違に関しては、さらに検討をする必要があるであろう。

## 要 約

International Clearinghouse for Birth Defects Monitoring 様式の Monitoring System をわが国へ導入することを前提として、すでに実施されているプログラムにつき、情報理論による解析を試みた。その結果、奇形のマーカーに関しては、主に診断基準に係わる要因が影響し、プログラムについては主に標本数が強い効果を発揮する、との知見をえた。このことから、プログラムの各期(3カ月)出生数は、1万前後が適正な規模である、と結論することができる。

表1 エントロピーによる奇形の分類

エントロピー	奇形の種類	P*
各国とも 変動小	尿道下裂	0.032
	口唇裂	0.032
	ダウン症候群	0.091
いずれとも 確定せず	先天股脱	0.19
	二分脊椎	0.33
	無脳症	0.67
	口蓋裂	0.67
	水頭症	0.18
各国とも 変動大	食道狭窄・閉鎖	0.018
	直腸肛門狭窄・閉鎖	0.018
	上肢減奇形	0.018
	下肢減奇形	0.018

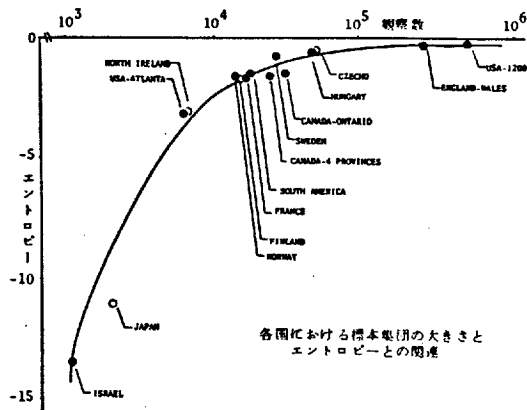
\* 上表におけるPは、12奇形のうち、変動の小さい方から5位以内とそれ以下に、等しく出現したとするときの分布からの偏りを、Fisherの直接確率計算法により算出したものである。

表2 エントロピーによる各国資料の比較

エントロピー	国名	P
各奇形とも 変動小	ENGLAND-WALES	0.018
	USA-1200 HOSPITALS	0.07
	CZECHOSLOVAKIA	—
	HUNGARY	0.018
各奇形とも 変動中	SWEDEN	0.181
	CANADA-ONTARIO	0.67
	FRANCE	0.50
	NORWAY	0.50
	SOUTH AMERICA	0.50
	CANADA-4 PROVINCES	0.19
各奇形とも 変動大	NORTH IRELAND	—
	USA-ATLANTA	0.032
	JAPAN	—
	ISRAEL	0.0062

表2のPは、表1と同様である。

図1



↓ **検索用テキスト** OCR(光学的文字認識)ソフト使用 ↓  
論文の一部ですが、認識率の関係で誤字が含まれる場合があります

まえがき

Internatinal Clearinghouse for Birth Defects Monitoring(以下 ICBDM と略す)へ 1977 年までに参加しているモニタリング・プログラムは, 12 カ国, 14 であった。これに加え 1978 年後半から, 東京都日赤病院のプログラムが登録された。1979 年スペイン, 1980 年イタリーが, それぞれ病院を基盤とするプログラムで ICBDM へ参加する予定である。さらに, デンマーク, ニュージーランドの 2 国は, 国家規模で先天異常の登録とモニタリングを行うべく, 目下準備中である。