

Pearl の Causality と行政判断

山本英二・津田敏秀*

岡山理科大学総合情報学部情報科学科

*岡山大学医歯学総合研究科社会環境生命科学専攻

(2002年11月1日 受理)

1. 集団の因果関係と行政判断

「集団で立証される因果関係は個人の因果関係を示すか」との命題に行政判断は答える必要はない。生命保険、失業保険、健康保険の考え方に見られるように個人のリスクを集団のリスクとして管理する立場からは因果関係は集団で立証されていれば良く、集団でのリスクを個人のリスクと読み替えば、個人の集まりである集団のリスクが管理できるからである。このとき、国内では外資系を除き生命保険で、喫煙者と非喫煙者を区別しないように、個々人のリスクの立証は問題にならない。これは、喫煙者、非喫煙者を込みの集団リスクを管理する立場から全体のリスクを個々に適用するため、喫煙者集団、非喫煙者集団を区別した生命保険では、当然喫煙者集団のリスクを非喫煙者のリスクに読み替えることはしない。行政の立場は、住民、国民を集団としてリスク管理する立場であるので、「集団で立証された因果関係を個人の因果関係として扱う」ことで良い。

では、科学的立場からは「集団で立証される因果関係は個人の因果関係を示すか」との命題にどのように答えるかを考える。近年の理論疫学では、以下の3つの考え方が提案され議論が深められてきている。

- Rothman 1986 Sufficient Causes
- Greenland & Robins 1986 Causal types
- Pearl 2000 Causal model

Sufficient Causes と Causal types は個人では決定論、集団ではリスク論であり、Causal model は個人リスク論である。ここでは、これら3つの考え方を Pearl2000 の Joe の例題を用いて紹介し、整理を行い、集団のリスクと個人のリスクを議論する。

2. Joe の例題

無作為化臨床試験(RCT)で処置 X ($X=1$ 処置, $X=0$ 無処置)と反応 Y ($Y=1$ 快復, $Y=0$ 死亡)が観測されるとする。今 Joe が処置を受け($X=1$)、快復した($Y=1$)とする。このとき Joe の快復は 1. 処置によって起こったのか、2. 処置に反して起こったのか、3. 処置に関係なしに起こったのか、との質問である。言い換えれば、もし Joe が事実に反して処置を受けなかったとき、Joe が快復する確率を与えよという質問である。ここで注意すべきは、Joe 個人に対して快復リスクを考えていることである。個人に対しては快復か死かのどちらかしかないとの決定論の立場を取っていないことである。一方、Sufficient Cause と Causal types は個人に対しては決定論の立場にあることに注意しよう。

今、RCTにおける集団の観測によって、 (X,Y) の同時分布が $p(x,y)=1/4$ $x=0,1; y=0,1$ であったとする。このとき、 $P(Y=1|X=0)=P(Y=1|X=1)=1/2$ となり、処置は反応と独立である。このことから Joe がもし処置を受けなかったら(counterfactual)単純に快復する確率は $1/2$ と答えて良いか否かと言うことである。求めたい確率は $Q=P(Y=1|do(X=0),(x,y)=(1,1))$ となる。 $(x,y)=(1,1)$ という evidence の下で、他の条件は変えずに(abduction)、強制的に $X=0$ と操作する(action $do(X=0)$)とき、 $Y=1$ となる確率を予測(prediction)せよという質問となる。この質問に対して、Pearl は2つの因果モデルを与え、model 1 では $Q=1$ 、model 2 では $Q=0$ という答えを与えている。

3. Rothman の Sufficient Causes

Rothman の Modern Epidemiology (1st edition 1986, 2nd edition 1998)によれば, Joe の例では以下の様に説明される. Joe が処置 X / 無処置 \bar{X} を受け, 快復 $Y=1$ するか否かはつぎの Sufficient Causes I, II, III のどれかを満足する必要がある.

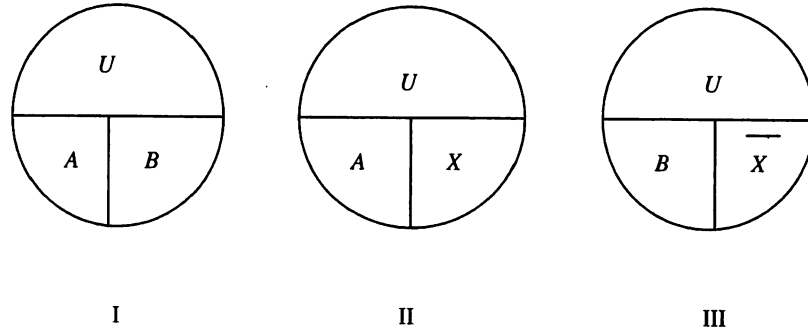


図1 Sufficient Causes I, II, III

快復 $Y=1$ が起こるのは Causal components U, A, B, X の組み合わせによるモデルである. ここでは U は全ての個人が持つ背景因子で, 組み合わせの議論では常に存在するので無視して考えて良い. A, B が存在すれば X に関わらず $Y=1$ が起こる組み合わせが I で, A, X が存在すれば B に関わらず $Y=1$ が起こる組み合わせが II である. III は B, \bar{X} が存在すれば A に関わらず $Y=1$ が起こる組み合わせである. III は X が予防的であることに注意しよう. また, 各個人にとっては快復/死亡は決定論であることに注意しよう. このとき, 各個人が A, B, X に対してどの組み合わせを持つかは $2^3=8$ 通りとなり, それぞれの組み合わせに対して, I, II, III の少なくとも一つを満たせば $Y=1$, 全てを満たさなければ $Y=0$ と反応する. これらを表にまとめると表1のようになる.

表1 Exposure pattern and Response

	Exposure			Response	Population fraction		
	A	B	X	Y	general	model 1	model 2
type 1	1	1	1	1	p_1	0.25	0
	1	1	0	1	q_1	0.25	0
type 2	1	0	1	1	p_2	0	0.25
	1	0	0	0	q_2	0	0.25
type 3	0	1	1	0	p_3	0	0.25
	0	1	0	1	q_3	0	0.25
type 4	0	0	1	0	p_4	0.25	0
	0	0	0	0	q_4	0.25	0
					1	1	1

(X, Y) の周辺分布は表2となっている. model 1 は個人が (A, B) をセットで持つか否かのどちらかの集団であり, model 2 は個人が A か B のどちらか一方を持つ集団である. さて, 今集団の大きさを N 人とする. このとき $(x=1, y=1)$ が観測された群の人数は $N(p_1 + p_2)$ である. 内訳は (A, B) を持つ個人が Np_1 人, そし

表2 (X,Y) の周辺分布

	general		model 1 and model 2	
	x=1	x=0	x=1	x=0
y=1	p_1+p_2	q_1+q_3	0.25	0.25
y=0	p_3+p_4	q_2+q_4	0.25	0.25

て A のみを持つ個人が Np_2 人である。この人々がもし、 $X=0$ となったとき反応はどうなるであろうか。
 (A,B) を持つ Np_1 人は sufficient cause I を満足するので、やはり快復 ($Y=1$) する。しかし A のみを保つ Np_2 人は I, II, III を満足しなくなり快復しない ($Y=0$)。よって集団としては、処置により快復した群が、もし処置を受けなかったら快復する確率は、

$$Q = \frac{Np_1}{N(p_1+p_2)} = \frac{p_1}{p_1+p_2}$$

となる。この確率は model 1 では $Q=1$ 、model 2 では $Q=0$ となる。Joe 個人の快復リスク Q については、この議論では直接は何も語れない。母集団分布 $(p_1, p_2, p_3, p_4, q_1, q_2, q_3, q_4)$ に関係なく、Joe が (A,B) を持っていれば $Q=1$ 、Joe が A しか持っていなければ $Q=0$ である。Joe が (A,B) 又は A のどちらを持っているか解らなければ集団での (A,B) 又は A の割合を用いて Joe の快復リスクを $Q = p_1/(p_1+p_2)$ とするのが合理的であろう。これは集団のリスクを個人のリスクに読み替えたことに他ならない。

4. Greenland & Robins の Causal types

Rothman1986 の Sufficient Causes とは別の観点から、Greenland & Robins1986 は各個人が $x=1$ と $x=0$ の暴露を受けたときの反応パターンが $2 \times 2 = 4$ タイプに分類されることを用いた議論を行っている。

表3 Causal types

Individual types			Population fraction					
			general		model 1		model 2	
			x=1	x=0	x=1	x=0	x=1	x=0
1	y=1	y=1	p_1	q_1	0.25	0.25	0	0
2	y=1	y=0	p_2	q_2	0	0	0.25	0.25
3	y=0	y=1	p_3	q_3	0	0	0.25	0.25
4	y=0	y=0	p_4	q_4	0.25	0.25	0	0
			1		1		1	

type 1 は処置にかかわらず快復する人、type 2 は処置によって快復するが無処置では死亡する人、type 3 は逆に処置によって死亡するが、無処置によって快復してしまう人、type 4 は処置に関わらず死亡してしまう人である。この集団における4つのタイプの人々が、処置を受けたか否かで、分類すると8通りとなり、 $(p_1, p_2, p_3, p_4, q_1, q_2, q_3, q_4)$ の分布を持つ。(X,Y) の周辺分布は Rothman1986 の Sufficient Causes の場合と同じ、表2となる。この集団における快復リスクは一般には $Q = p_1/(p_1+p_2)$ 、model 1 で $Q=1$ 、model 2 で $Q=0$ となり、やはり Joe 個人の Q は直接与えられない。Joe が type 1 の人間であれば $Q=1$ であり、type 2 の人間であれば $Q=0$ であるが、これは母集団分布 $(p_1, p_2, p_3, p_4, q_1, q_2, q_3, q_4)$ に寄らず確定している。Joe が type 1 か type 2 のどちらか解らなければ、集団での type 1 又は type 2 の割合を用いて Joe の快復リスクを $Q = p_1/(p_1+p_2)$ とするのが合理的であろう。これは集団のリスクを個人のリスクに読み替えたことに他ならない。

5. Pearl の Causal model

Rothman1986, Greenland & Robins1986 の議論は直接的な個人の因果推測に対する困難が存在する。これに対して, Pearl は個人に対する因果モデルを前提にして, 集団で観測されたデータを用いて, Joe の個人リスク Q を直接に推測する考え方と方法論を提案している。

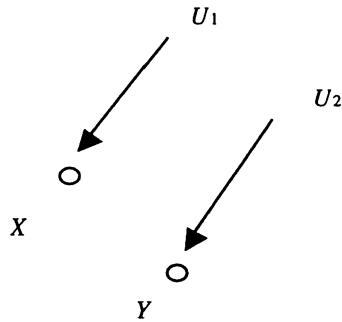


図2 model 1の因果グラフ

[model 1] model 1 の (定性的な) 因果グラフは図2の通りで, (定量的な) 構造方程式で表現すれば,

$$\begin{cases} x = u_1 \\ y = u_2 \end{cases} \quad U_1 \perp U_2 \quad \begin{cases} p(u_1) = 1/2 & u_1 = 0,1 \\ p(u_2) = 1/2 & u_2 = 0,1 \end{cases}$$

である。(U_1, U_2) は外生因子で, (X, Y) は上の構造方程式が示す因果律に従うとするモデルである。このとき (X, Y, U_2) の同時分布は表4のようになり, (x, y) = (1,1) のとき $u_2 = 1$ となっている。

表4 model 1 の (X, Y, U_2) の同時分布

$p(x, y, u_2)$	$u_2=0$		$u_2=1$		marginal	
	$x=1$	$x=0$	$x=1$	$x=0$	$x=1$	$x=0$
$y=1$	0	0	0.25	0.25	0.25	0.25
$y=0$	0.25	0.25	0	0	0.25	0.25

この状況で $x=0$ と操作したときに $Y=1$ となる確率は $Q=1$ となる。この model 1 が仮定されているとき, Joe の快復リスクは $Q=1$ である。

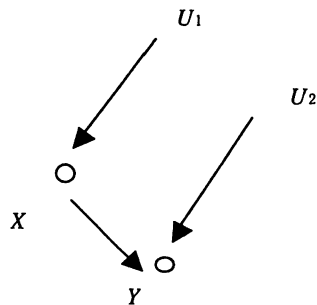


図3 model 2の因果グラフ

[model 2] model 2 の (定性的な) 因果グラフは図 3 の通りであり, (定量的な) 構造方程式で表現すれば

$$\begin{cases} x = u_1 \\ y = xu_2 + (1-x)(1-u_2) \end{cases} U_1 \perp U_2 \quad \begin{cases} p(u_1) = 1/2 & u_1 = 0,1 \\ p(u_2) = 1/2 & u_2 = 0,1 \end{cases}$$

となる因果モデルである. このとき (X, Y, U_2) の同時分布は表 5 となり, $(x, y) = (1, 1)$ のとき $u_2 = 1$ となっている. この状況で $x=0$ と操作したときに $Y=1$ となる確率は $Q=0$ となる. この model 2 が仮定されているとき, Joe の快復リスクは $Q=0$ である.

表5 model 2 の (X, Y, U_2) の同時分布

$p(x, y, u_2)$	$u_2=0$		$u_2=1$		marginal	
	$x=1$	$x=0$	$x=1$	$x=0$	$x=1$	$x=0$
$y=1$	0	0.25	0.25	0	0.25	0.25
$y=0$	0.25	0	0	0.25	0.25	0.25

上の 2 つのモデルでは, 外生因子 (U_1, U_2) が一般に

$$\begin{cases} p(u_1) = p^{u_1} (1-p)^{1-u_1} & u_1 = 0,1 \\ p(u_2) = q^{u_2} (1-q)^{1-u_2} & u_2 = 0,1 \end{cases}$$

であっても model 1 では $Q=1$, model 2 では $Q=0$ となり, これが Joe 個人の快復リスク Q となる. (U_1, U_2) は Rothman1986 の Causal components A, B にそれぞれ対応した構造モデルの外生因子または誤差項と考えられ, また Greenland & Robins1986 の Causal types を構造化したものと考えられる.

6. おわりに

「集団で立証される因果関係は個人の因果関係を示すか」との命題に答えられるだろうか. 集団を管理する立場からは「集団で示される因果リスクをその集団に属する個人の因果リスクとして扱う」ことで上の命題に答える必要はない. 当然, 喫煙歴や性別, 年齢等で集団を狭くすれば, 真の個人の因果律により近づくが, 集団の取り方は保険の様に管理の目的による. 科学的議論での, 個人の因果関係は決定論が成立しているとする立場からは, Rothman1986 の Sufficient Causes や Greenland & Robins1986 の personal types の個人情報に完全に把握できれば, 個人因果リスクは 0 か 1 しかない. 現実には, 個人情報を完全に把握できることはないので, その分を個人因果リスクとして考える立場は, 管理できない因子による影響を誤差として扱う統計的議論とも言える. 実際にあらゆる因子を把握できない以上, 情報が不完全なため, または制御できないために確率的になるのか, 本来的に確率的な現象なのかを議論しても答えは出ない. この点で, Pearl2000 の議論は個人に確率的因果律と因果モデルを前提として個人因果リスクを求めている, 上の議論は避けているが, 前提とする因果モデルの妥当性の検証は集団的な実践による検証しか出来ないであろう. 実際, Pearl2000 は集団の因果モデルを個人の因果モデルに取り込んだとも解釈でき, このとき Rothman1986, Greenland & Robins1986 との違いはみえない.

参考文献

- [1] S. Greenland, J.M. Robins(1986). Identifiability, exchangeability and epidemiological confounding. *International Journal of Epidemiology* 15, 413-419.
- [2] K.J. Rothman(1986). *Modern Epidemiology*, Little, Brown and Company, Boston.
- [3] K.J. Rothman, S.G. Greenland(1998). *Modern Epidemiology 2nd Edition*, Lippincott-Raven.

[4] J. Pearl(2000). *CAUSALITY Models, Reasoning, and Inference*. Cambridge Univ. Press.

謝辞

この研究は統計数理研究所共同研究プログラム（13-共研-2040）医薬剤の市販後疫学調査等と行政判断（代表：津田敏秀）の援助を受けている。

Pearl's Causality and Administrative Judgment

Eiji Yamamoto and Toshihide Tsuda*

Department of Information Science, Faculty of Informatics,

Okayama University of Science,

1-1 Ridai-cho, Okayama 700-0005, Japan.

**Social and Environmental Life science,*

Graduate School of Medicine and Dentistry, Okayama University,

2-5-1 Shikata-cho, Okayama 700-8558, Japan

(Received November 1, 2002)

The proposition that causality proved in a population can drive individual causality is recently discussed in the epidemiology with terms of sufficient causes by Rothman 1986, causal types by Greenland & Robins 1986, and a causal model by Pearl 2000. The first two employ determinism in individuals and risk theory in a population and the last does individual risk theory. We arrange their discussions and show the relationship between individual causality and population causality. For administrative judgment, it is acceptable that causality proved in a population can be treated as individual causality in the population, and that individual causality could not be evaluated without proof of population based causality.