

## 循環器疾患のコホート研究と統計科学

筑波大学\* 磯 博康・嶋本 喬・山海 知子  
谷川 武・大平 哲也

(受付 1997 年 10 月 22 日; 改訂 1998 年 1 月 30 日)

### 要 旨

循環器疾患のコホート研究の主な目的は脳卒中、虚血性心疾患の発症リスクに影響を及ぼす身体・生活環境要因の同定とその影響の評価にある。わが国では主として人口の移動の少ない農村地域や 60 歳までの追跡が容易な都市の勤務者集団等において、コホート研究が行われてきた。コホート研究を行う際、重要な点は循環器疾患の発生に関わる要因を統一した方法によって集団検診等において把握し、かつ脳卒中、虚血性心疾患の発生についても統一した方法・基準により、もれなく把握することである。また、コホート研究のサンプル数の設定も重要であり、検討する発症要因の種類によって異なる。一般に脳卒中に関して血圧値等の身体所見を検討する際にはサンプル数は人年で、数千で十分であるが、アルコール摂取、喫煙、身体活動、栄養摂取といった生活習慣に関しては相対危険度が比較的低く、かつ再現性の問題もあり、より大きなサンプル数を必要とする。特に栄養摂取と循環器疾患の関連の分析ではサンプル数は 1 万人年以上必要となる場合がある。

身体・生活環境要因と循環器疾患発生との関連の分析は関連要因が複数存在することが多いため、多変量解析の手法を用い、相対危険度を算出するのが一般的である。その際、発生要因の調査におけるバラツキから regression dilution bias が生じるため、その bias を調整して相対危険度を評価する場合もある。

あるコホート研究において、疾病の発症との関連が認められても、その要因が疾病の原因であるか否か (因果関係) の保証は必ずしもできない。因果関係の立証にはコホート研究の成績が他の集団、研究者の研究によっても認められること (consistency)、生物学的な説明が可能なこと (biological plausibility) などの条件が必要となる。

キーワード：コホート研究、循環器疾患、統計科学、危険因子、サンプル数、ライフスタイル。

### はじめに

わが国における循環器疾患のコホート研究は 1960 年代初頭にさかのぼる。その当時、日本人の死因の第一位を占めていた脳卒中を減らす目的で、全国のモデル的な地域において予防対策と疫学研究が並行して進められた。対策の重点は脳卒中のハイリスクグループである高血圧を集団検診で発見し、高血圧者に対して重点的に指導を行い、ライフスタイルの改善と治療の必要な人に対する医療機関への受療の勧奨 (高血圧の二次予防) であった。結核対策以来の伝統を持つ集団検診によるスクリーニング方式が、高血圧者を効率よく把握する手段として採用されたわけである (Shimamoto et al. (1989))。

\* 社会医学系：〒305-8575 茨城県つくば市天王台 1-1-1.

さらに、1980年代に入ってからでは高血圧の管理のみでなく、高血圧の発症自体を目的とした地域住民全体に対する健康教育(減塩、栄養のバランス、節酒、過度の肥満の防止等)、すなわち、高血圧の一次予防が行われるようになった(嶋本 他(1989))。

疫学研究は、高血圧やその他の身体・生活環境要因の、循環器疾患発生への影響を定量的に分析したり、予防対策の客観的な評価を行うために長期間継続して行われてきた。本稿では循環器疾患の疫学研究の中でもコホート研究に絞り、わが国の地域においてどのように進められてきたか、その成果及び問題点を統計科学的な見地から論じる。

### 1. コホート研究の目的

循環器疾患のコホート研究の目的には様々なものがあるが、最も大きな目的は脳卒中、虚血性心疾患(心筋梗塞、狭心症等)の発生リスクに影響を及ぼす身体・生活環境要因の同定とその影響の強さの評価である。さらに、発生要因を取り除いた場合、どの程度発生リスクを減少させ得るかを推定することが可能である。コホート研究の場合、要因(factors)が高血圧、血清脂質、その他の血液検査の所見、心電図所見、眼底所見、食生活、生活環境等の因子であり、疾病(events)が脳卒中、虚血性心疾患である。

### 2. 対象集団の設定

わが国のコホート研究の多くは、脳卒中の死亡率の高い地域(農村地域)において脳卒中の予防対策と並行して進められてきた(嶋谷 他(1976)、上田(1987))。農村地域は人口の移動が都市に比べて少なく、循環器疾患の発生を把握する上で有利である点も対象集団の設定の理由として見逃せない。また、都市においても事業所の勤務者集団(主として男子)は60歳の定年までは追跡が容易であるため、農村集団との対比という意味でコホート研究が行われてきた。都市の勤務者集団は農村集団に比べて、脳卒中の発生は少ないが、虚血性心疾患の発生が比較的多く、60歳までの壮年期の発生という条件づきではあるが、虚血性心疾患の発生要因の分析には有利であった(Kitamura et al.(1994))。わが国のいくつかのすぐれたコホート研究においては地域や職域のある年齢層全員を対象に、それらの70-90%に対して要因(factor)に関する調査が行われ、循環器疾患の既往者を除いたコホートが設定された。そして、循環器疾患の発生状況に関する追跡調査が行われてきた。一般に農村地域では追跡期間中に転出し、調査から離脱する者は10年間で数%と少ない。また、職域においても、人事が安定した企業においては、定年まで追跡調査が不可能になる例は少ない。そのため、疫学研究でよく問題となる曝露要因の調査におけるselection bias(曝露要因有りもしくは無しの者が偏って、調査に参加する)や、追跡調査におけるselection bias(追跡不能者に循環器疾患の発生が多い等)は少なくすることが可能であった。

### 3. 循環器疾患発生に関与する曝露要因の把握

循環器疾患の発生要因を分析するため、集団検診等において、統一した方法により、曝露要因の調査を行っている。曝露要因としては年齢、性、安静時血圧値、body mass index(体重/身長<sup>2</sup>, kg/m<sup>2</sup>)、血清脂質(血清総コレステロール、HDL-コレステロール、中性脂肪)、安静時心電図所見、眼底所見等の身体所見や、アルコール摂取状況、喫煙状況等の生活習慣があげられる。循環器疾患のコホート研究では、多くの場合集団検診を毎年繰り返し行っているため、新しい検査項目を途中で導入し、その時点でのコホート設定をすることが可能である。近年で

は血液凝固線溶系因子(血漿フィブリノーゲン, 組織プラスミノーゲン活性抑制因子等), 血清脂肪酸, 血小板凝集能, 頸部エコー所見等が導入されている。

そして, ある要因の曝露群の発生率と非曝露群の発生率の比 (incidence ratio もしくは相対危険度) を指標として, 要因と疾病との関連の強さを評価する。要因が連続量の場合 (例えば血圧値) は, 要因を層化して, 最も低い血圧値の区分を要因なしの群 (非曝露群) として扱い, それ以上の血圧区分での発生率比を算出する。

#### 4. 脳卒中, 虚血性心疾患の定義, 診断基準

脳卒中は脳の血管病変による脳の機械的・器質的障害を総称するものであり, 通常, 意識障害や脳の局所症状が 24 時間以内に急激に発症するものを指す。したがって, 単一の疾患ではなく, 脳出血, くも膜下出血, 脳梗塞などに大別される。また脳梗塞は脳の血管の血栓による脳血栓と, 主として心臓や大血管の血栓が飛んで脳の血管につまる脳塞栓に分類される。

虚血性心疾患は, 冠動脈の粥状硬化により心臓への血液供給の不足によって生ずる病態をさす。胸痛などの症状が一過性に出現した場合が狭心症であり, 胸痛が 20 分以上持続し, 心筋に壊死が生じた場合が心筋梗塞である。

いずれの場合も一定の臨床的な診断基準 (Computerized Tomography, Magnetic Resonance Imaging, 心電図, 心筋酵素等の検査成績を含める) により, 統一した方法により診断を行っている (Shimamoto et al. (1989), 山海 他 (1991))。

#### 5. 脳卒中, 虚血性心疾患の発生把握

脳卒中, 虚血性心疾患はいずれも急激に発症し, 症状も特異的であることが多いため, 発生時点の把握はそれほど困難ではない。しかしながらコホート研究の対象集団の中からこれらの疾患の発生をもれなく把握するためには特別の工夫が必要である。まず発生を疑われるケースを広く様々な情報源から, もれなく把握することが重要である。情報源は地域集団が対象である場合は, 1) 死亡票, 2) 国保レセプト, 3) 救急搬送録, 4) 医師からの通報, 5) 保健婦記録, 6) 全世帯アンケート, 7) 循環器検診時の問診等, である (Shimamoto et al. (1989))。勤務者集団が対象である場合は, 1) 死亡票, 2) 休業届け, 3) 組合保険請求書, 4) 産業医からの通報, 5) 循環器検診時の問診等, である (Kitamura et al. (1994))。

これらの情報源で把握されたケースに対して, ケース本人 (生存例) やその家族を訪問し, 発生時の状況を詳しく問診するとともに病院, 診療所の記録の閲覧を行う。そして, 得られた情報に基づき, 統一した診断基準により, 3~4 人の研究医師が独立して判定を行い, 診断の異なるケースについては合議して, 最終診断を下す方法をとっている (Shimamoto et al. (1989), Kitamura et al. (1994))。

#### 6. コホート研究のサンプル数

循環器疾患のコホート研究の対象は脳卒中, 虚血性心疾患が通常 40 歳以上で発症することから, この年齢層以上を対象としてきた。特に脳卒中の発生要因の分析にあたっては, 1960 年代当時, 壮年層で脳卒中を減らすことが最重要課題であったため, 曝露要因の調査も 40-69 歳男女を中心に進められた。

コホート研究の調査サンプル数は曝露群と非曝露群の人数が同数とした場合, 一般に以下の式により算出される (重松・柳川 他 (1991))。この際のサンプル数はコホートの人数に個人ご

との追跡年数を乗じた人年で表される。

$$N(\text{曝露群} \cdot \text{非曝露群それぞれの人年数}) = \frac{\{Z_\alpha \cdot (2pq)^{1/2} + Z_\beta p_1 \{1 + R - p_1(1 + R^2)\}\}^2}{\{p_1(1 - R)\}^2}$$

$p_1$ : 非曝露群における発生率

$R$ : 検出する相対危険度

$Z_\alpha$ :  $\alpha$  (第1種の過誤) に対応する標準正規分布の値

$Z_\beta$ : 検出力 (1 - 第2種の過誤) に対応する標準正規分布の値

$p = 1/2 \cdot p_1(1 + R)$

$q = 1 - p$

1990年代当初、高血圧者は非高血圧者に比べて、脳卒中発生率は約5倍 ( $R=5$ ) であった。そして、非曝露群 (非高血圧者) における脳卒中の発生率は人口1,000人当たり3 ( $p_1=0.003$ ) であった (Shimamoto et al. (1989))。これを上式にあてはめて、 $\alpha=0.05$ ,  $\beta=0.20$  とした場合の必要なサンプル数 (人年) を算出すると、

$$p_1 = 0.003$$

$$R = 5$$

$$Z_\alpha = 1.96$$

$$Z_\beta = 0.84$$

$$p = 1/2 \times 0.003(1+5) = 0.009$$

$$q = 1 - p = 0.991$$

$$\begin{aligned} N &= \frac{\{1.96 \cdot (2 \times 0.009 \times 0.991)^{1/2} + (0.84 \times 0.003)\{1 + 5 - 0.003(1 + 5^2)\}\}^2}{\{0.003(1 - 5)\}^2} \\ &= \frac{(0.262 + 0.0149)^2}{0.000144} \\ &= 532 \end{aligned}$$

すなわち、高血圧者532人年、非高血圧者532人年が必要なサンプル数となる。多くの地域集団では高血圧者が約3割を占めるため、実際は約1,800人年の住民 (40-69歳男女) の調査が必要となる。従って、研究成果を上げるための追跡期間を仮に5年とすると、約360人の住民の調査でよいことになる。

実際は循環器疾患のコホート研究では1つの町村の40-69歳男女で、1,000-3,000人のコホートを組むことが多かった。結果的には、このサンプル数で高血圧と脳卒中発生との関連の分析は十分であった。

秋田県農村におけるコホート研究 (Shimamoto et al. (1989)) では1963-1966年において男女40-69歳2,257人のコホートを設定し、10年間追跡した。また、1970-1975年において新たに男女40-69歳2,711人のコホートを設定し、10年間追跡した。このコホート研究では脳卒中を脳出血と脳梗塞の病型に分けて、発生要因を分析している。多変量解析であるロジスティックモデルを適用したところ、1960年代と1970年代いずれのコホートでも最大血圧値が脳卒中の発生リスクと有意の関連を示した (表1)。例えば1960年代のコホートでは最大血圧値20 mmHgの上昇に伴う脳出血のリスクを算出すると  $\exp(0.024 \times 20) = 1.62$  となり、脳出血のリスクの62%の上昇と関連していた。また、いずれのコホートでも眼底の高血圧性変化が脳出血のリスクと有意な関連を示した。例えば、1960年代のコホートでは眼底の高血圧性変化がある場合は無い場合に比べ、脳出血のオッズ比は  $\exp(0.989) = 2.69$  であった。

一方、脳梗塞の発生要因としては、最大血圧自体ではなく、心電図変化もしくは眼底変化が有意な要因として抽出された。このことは、脳出血が高血圧の比較的短期的な影響として発症

表1. 脳出血, 脳梗塞の発生要因の分析, ロジスティックモデルによる偏回帰係数.  
秋田農村, 男女 40~69 歳, 10 年間のコホート研究.

	脳出血		脳梗塞	
	初期 (n=20)	後期 (n=13)	初期 (n=45)	後期 (n=54)
男	-0.378	-0.350	-0.537*	-0.728†
年齢 (歳)	-0.008	-0.082*	0.074‡	0.061‡
最大血圧値 (mmHg)	0.024*	0.036*	0.008	0.013
血清総コレステロール値 (mg/dl)	-0.021†	0.003	0.009	0.002
心電図 ST-T変化	1.194*	-1.338	0.070	0.973*
心電図 左室高電位	-0.287	1.103*	0.804*	0.204
眼底 高血圧性変化	0.989	1.105†	0.635†	0.173
眼底 動脈硬化性変化	0.299	0.524	-0.168	0.148

\* $p < 0.05$ , † $p < 0.01$ , ‡ $p < 0.001$

初期コホート: 1963-66年 2,257人, 後期コホート: 1970-75年 2,711人.

Shimamoto et al. (1989) より表9を再掲.

するのに対し, 脳梗塞が高血圧の慢性的な影響 (高血圧性の臓器変化) と関連して発症することを物語っている.

1960年代のコホートでは血清総コレステロールの低値が脳出血のリスクの上昇と関連することが示されたが, 1970年代のコホートでは有意の関連は見られていない. この解釈については研究対象のサンプルサイズと検出力の観点から論ずる必要がある. 表1には血清総コレステロール値を連続量として扱った脳出血発生に関するロジスティック分析の成績 (Shimamoto et al. (1989)) が示されているが, 実際, 層別分析を行ってみると, 血清総コレステロール値が 140 mg/dl 未満の非常に低い区分でそれ以上の値の区分に比べて, 脳出血のリスクが約 2.5 倍であった. これを前述の式により, 必要サンプル数 (人年) を概算すると,  $p_1 = 0.0015$ ,  $R = 2.5$  として  $N = 4,224$  となる.

すなわち, 血清総コレステロール値 140 mg/dl 未満の群 (曝露群) が 4,224 人年, 血清総コレステロール値がそれ以上の群 (非曝露群) が 4,224 人年必要である. 1960年代のコホートのサンプル数は 2,298 人であり, このうち血清総コレステロール値が 140 mg/dl 未満の者は約 700 人 (約 30%), それ以上の値の者は約 1,500 人であり, 10年間の追跡調査で, それらの人年は 7,000 人年 15,000 人年であり, 必要サンプルサイズに達する.

しかし, 1970年代のコホートでは, 1960年代から1970年代にかけて住民の血清総コレステロールの平均値が 20 mg/dl 上昇し, そのため 140 mg/dl 未満を示す者の割合は約 10%と減少した. 従って 1970年代のコホートのサンプル数は 2,711 人であるが, 血清総コレステロール低値を示す者の人数は約 270 人に過ぎない. さらに, 非曝露群での脳出血のリスク ( $p_1$ ) も半減した. そのため, 血清総コレステロール低値と脳出血のリスクとの関連を分析する上での検出力が激減した. すなわち, 1970年代のコホートで血清総コレステロール低値と脳出血のリスクとの間の関連が見られなかった理由の一つに 1970年代コホートの検出力の低さがあげられる.

血清総コレステロール値が短期間で著しく上昇したのは, この間の日本人の食生活が大きく変わったことと一致している. 極端に少なかった動物性食品の摂取 (肉類, 乳製品) が流通機構の改善, 経済状態の向上等に伴い, ある程度摂取するようになったためである.

血清総コレステロールの低値と脳出血の発生リスクとの関連は日本人のコホート研究のみならず, ホノルル在住の日系男子 (Kagan et al. (1980)) や米国の白人男子のコホート研究 (Iso

表2. 血清総コレステロール区分別にみた虚血性心疾患, 確実な心筋梗塞, 脳卒中の年齢調整発生率, 相対危険度, 都市13企業勤務者, 男40~59歳, 6,408人, 7.7年間のコホート研究.

血清総コレステロール (mmol/L)	調査数	人年	虚血性心疾患 (n=46)			確実な心筋梗塞 (n=21)			脳卒中 (n=33)		
			発生数	年齢調整 相対危険度	多変量調整* 相対危険度	発生数	年齢調整 相対危険度	多変量調整* 相対危険度	発生数	年齢調整 相対危険度	多変量調整* 相対危険度
<4.50	1620	12,158	5	0.46	1.00	2	0.18	1.00	11	0.93	1.00
4.50-5.06	1579	12,047	7	0.58	1.48	3	0.26	1.53	2	0.17	0.19
5.07-5.63	1633	12,632	9	0.71	1.93	5	0.38	2.58	7	0.55	0.62
≥5.64	1576	12,200	25	2.02	4.89	11	0.89	5.14	13	1.02	1.13
p値					.001			.04			.78
偏回帰係数 (標準偏差) †					.724(.128)			.710(.186)			.177(.199)
p値					<.0001			.0001			.37

\* Cox 比例ハザードモデルにより, 年齢, HDL-コレステロール値, 最大血圧値, body mass index, 喫煙本数, アルコール摂取量を調整.

† Cox 比例ハザードモデルにより, 血清総コレステロール値を連続変量として, 上記の変量を調整して偏回帰係数を算出.  
Kitamura et al.(1994) より表3を再掲.

表 3. 血清 HDL-コレステロール区別にみた虚血性心疾患, 確実な心筋梗塞, 脳卒中の年齢調整発生率, 相対危険度 (Cox 比例ハザードモデル), 都市 13 企業勤務者, 男 40~59 歳, 6,408 人, 7.7 年間のコホート研究.

HDL-コレステロール (mmol/L)	調整数	人年			虚血性心疾患 (n=46)			確実な心筋梗塞 (n=21)			脳卒中 (n=33)		
		発生数	年齢調整 相対危険度	多変量調整* 相対危険度	発生数	年齢調整 相対危険度	多変量調整* 相対危険度	発生数	年齢調整 相対危険度	多変量調整* 相対危険度	発生数	年齢調整 相対危険度	多変量調整* 相対危険度
<1.24	1702	13	252	28	2.09	4.17	12	0.89	3.39	12	0.90	1.31	
1.24-1.44	1579	12	035	7	0.64	1.61	4	0.36	1.82	9	0.77	1.23	
1.45-1.65	1506	11	552	7	0.61	1.80	3	0.26	1.54	4	0.35	0.59	
≥1.66	1621	12	198	4	0.33	1.00	2	0.17	1.00	8	0.64	1.00	
p値					.01			.14				60	
偏回帰係数 (標準偏差) †					-2.25(.63)			-2.56(.95)				-.552(.588)	
p値					.0003			.007				.37	

\* Cox 比例ハザードモデルにより, 年齢, 血清総コレステロール値, 最大血圧値, body mass index, 喫煙本数, アルコール摂取量を調整.

† Cox 比例ハザードモデルにより, HDL-コレステロール値を連続変量として, 上記の共変量を調整して偏回帰係数を算出.  
Kitamura et al.(1994)より表 4 を再掲.

表4. 非飲酒者に対する循環器疾患の発生の相対危険度 (RR)\* と95%信頼区間 (95%CI).  
農村3集団, 男40~69歳, 2,890人, 10.5年間のコホート研究.

	エタノール摂取量 g/日									
	過去飲酒		1-20		21-41		42-69		70以上	
	RR	95%CI	RR	95%CI	RR	95%CI	RR	95%CI	RR	95%CI
全脳卒中 (n=178)										
年齢調整	1.4	0.7-2.8	0.9	0.5-1.6	1.0	0.6-1.7	1.5	0.9-2.3	2.7	1.7-4.3
多変量調整	1.2	0.6-2.4	0.7	0.4-1.4	0.8	0.5-1.4	1.1	0.7-1.8	1.9	1.2-3.1
出血性脳卒中 (n=58)										
年齢調整	2.4	0.6-9.9	1.7	0.5-5.8	1.9	0.6-5.7	2.0	0.7-5.5	6.2	2.3-16.5
多変量調整	1.7	0.4-7.2	1.3	0.4-4.6	1.4	0.5-4.1	1.2	0.4-3.4	3.4	1.2-9.2
非出血性脳卒中 (n=104)										
年齢調整	0.8	0.3-2.1	0.7	0.3-1.6	0.9	0.5-1.8	1.5	0.9-2.7	2.1	1.1-3.9
多変量調整	0.7	0.3-1.9	0.6	0.3-1.5	0.8	0.4-1.5	1.3	0.7-2.3	1.7	0.9-3.1
虚血性心疾患 (n=34)										
年齢調整	1.2	0.3-4.6	0.9	0.3-3.0	0.7	0.2-2.2	1.1	0.4-2.8	0.9	0.3-3.1
多変量調整	1.2	0.3-4.6	0.9	0.3-3.1	0.7	0.2-2.1	0.9	0.4-2.5	0.8	0.2-2.8
急性死 (n=19)										
年齢調整	3.4	0.2-55.0	3.5	0.3-38.5	4.5	0.5-39.9	4.0	0.5-34.4	9.9	1.2-83.8
多変量調整	3.4	0.2-55.2	3.5	0.3-39.0	4.3	0.5-39.0	3.6	0.4-31.7	8.8	1.0-76.1
循環器疾患 (n=225)										
年齢調整	1.5	0.8-2.7	0.9	0.5-1.6	1.1	0.7-1.7	1.5	1.0-2.3	2.7	1.7-4.1
多変量調整	1.3	0.7-2.4	0.8	0.5-1.5	0.9	0.6-1.5	1.2	0.8-1.8	2.0	1.3-3.1

\*年齢, 高血圧区分, 血清総コレステロール値, 喫煙区分, 糖尿病の有無を調整.  
Iso et al.(1995)より表3を再掲.

et al.(1989))においても認められている。それらのコホート研究はそれぞれ8千人, 35万人とコホートの人数が大きいことに特徴がある。特に白人においては日本人に比べ, 血清総コレステロールの低値の者の割合が低く, かつ, 脳出血の発生も少ないことから, 非常に大きなサンプル数が必要であった。

以上は, 脳卒中発生の要因に関する分析であったが, もう一つの主要な循環器疾患である虚血性心疾患の発生要因の分析に関しても, サンプル数を考慮する必要がある。虚血性心疾患の発症要因について都市の勤務者集団において行った例を表2に示す。都市の13事業所の40-59歳6,408人を対象に平均7.7年間追跡し, 血清脂質(血清総コレステロール, HDL-コレステロール)と虚血性心疾患の発生との関連を分析した(Kitamura et al.(1994))。

血清総コレステロール値が5.6 mmol/L (218 mg/dl)以上の者は4.50 mmol/L (174 mg/dl)未満の者に比べて虚血性心疾患発生の相対危険度が5(表2)であり, この場合の必要なサンプルサイズを算出すると,  $\alpha_1=0.003$ ,  $R=5$ として $N=532$ 人年となる。すなわち, 血清総コレステロール5.64 mmol/L以上の群, 4.50 mmol/L未満の群でそれぞれ532人年必要となる。実際上記のコホートは両群とも約12,000人年であり十分なサンプル数であった。

また, HDL-コレステロール1.24 mmol/L (48 mg/dl)未満の者は1.66 mmol/L (64 mg/dl)以上の者に比べて虚血性心疾患発生の相対危険度は4(表3)であり, この分析のための必要なサンプルサイズを計算すると,  $\alpha_1=0.003$ ,  $R=4$ で $N=782$ となる。すなわち, HDL-コレステロール1.24 mmol/L未満の群, 1.66 mmol/L以上の群はそれぞれ782人年必要となる。実際上記のコホートでは両群とも約12,000人年であり, この分析の場合も十分なサンプル数であっ



た。

## 7. 生活習慣と循環器疾患の発生

前述した身体要因である血圧値、血清脂質等は循環器の発生との関連において、相対危険度が2～5と比較的強い関連を示す。一方、アルコール摂取や喫煙については循環器疾患発生との関連の相対危険度が一般に1.5～3と比較的小さい。この理由の一つとして、アルコール摂取や喫煙といった生活習慣は自己申告によるものであり、その再現性が身体所見のようなより客観的な指標に比べて、一般的に低いことがあげられる。すなわち regression dilution bias (後述)が生じるためである (MacMahon et al. (1990))。

このバイアスはその他の生活環境、例えば身体活動状況や栄養摂取状況の場合となると一般にさらに大きくなる。従って生活環境と循環器疾患発生との関連を分析する場合、曝露要因の調査の再現性を高めるとともに、コホートのサンプル数 (人年) を大きくする必要が生じてくる。

アルコール摂取量と脳卒中発症との関連を分析した例を示す (Iso et al. (1995))。農村3集団の男40-69歳 2,890人の10.5年間のコホート研究である (表4)。アルコールを一日70g以上 (日本酒換算で3合以上) を摂取する者はこれまで全くアルコールを摂取していない者に比べ、脳卒中の相対危険度は2であった。

そこで必要なサンプル数は、 $k_1=0.006$ ,  $R=2$   $N=2,127$  人年となる。すなわち、アルコール70g/日以上以上の群、全く飲まない群、それぞれ2,127人年必要である。実際、上記のコホートでアルコール70g/日以上以上の群は439人、アルコールを全く飲まない群は585人であり、10.5年の追跡ではそれぞれ6,585人年、6,143人年であり、十分なサンプル数であった。

しかし、これが栄養摂取量となると、一般に脳卒中との関連はさらに小さいため、サンプル数をさらに大きくする必要がある。表5に、3農村、1都市近郊集団の男女40-69歳 5,931人の8年間のコホート研究の成績を示す (磯 (1996))。緑黄色野菜の摂取が1日平均22g未満の群が1日80g以上の群に比べて全脳卒中の相対危険度が1.41 (1/0.71)であった。この場合の必要なサンプル数を算出すると、 $k_1=0.005$   $R=1.4$   $N=12,553$  人年となる。すなわち、緑黄色野菜の摂取が22g/日未満の群と80g/日以上以上の群でそれぞれ12,553人年必要である。本研究では緑黄色野菜の摂取が22g/日未満の群と80g/日以上以上の群の人数は約2,000人で8年間の

表5. 野菜摂取と脳卒中の発生の相対危険度 (95%信頼区間)\*。

3農村, 1都市近郊集団, 男女40~69歳, 5,931人, 8年間のコホート研究。

		全脳卒中 n=271	出血性脳卒中 n=90	非出血性脳卒中 n=181
全野菜 (g/日)	<140	1.00	1.00	1.00
	140~239	0.85(0.65, 1.12)	0.63(0.38, 1.03)	1.01(0.72, 1.41)
	≥240	0.76(0.56, 1.04)	0.68(0.40, 1.14)	0.81(0.54, 1.19)
緑黄色野菜 (g/日)	<22	1.00	1.00	1.00
	22~79	0.89(0.66, 1.19)	0.83(0.51, 1.35)	0.93(0.64, 1.35)
	≥80	0.71(0.53, 0.95)	0.45(0.27, 0.77)	0.87(0.61, 1.24)
その他の野菜 (g/日)	<80	1.00	1.00	1.00
	80~159	0.82(0.62, 1.10)	0.90(0.54, 1.51)	0.80(0.56, 1.13)
	≥160	0.94(0.70, 1.27)	1.18(0.71, 1.95)	0.84(0.58, 1.23)

\*Cox 比例ハザードモデルにより、性、年齢、最大血圧、血清総コレステロール値、アルコール摂取量、肥満度、糖尿病の有無、ECG上の左室肥大の有無を調整。

磯(1996)より表2を再掲。

表6. 脳出血及び脳梗塞の病型別発生要因の分析, Coxハザードモデルによる偏回帰係数, 農村3集団 (人口8,717人), 7年間のコホート内症例対照研究, 男女40歳以上.

変 量 (単位)	脳梗塞														
	脳出血				穿通枝系				皮質枝系						
	回帰係数	標準誤差	p値	回帰係数	標準誤差	p値	回帰係数	標準誤差	p値	回帰係数	標準誤差	p値			
最大血圧 (mmHg)	0.019	0.007	0.006	0.020	0.008	0.010	0.013	0.011	0.222	-0.014	0.018	0.428	0.061	0.027	0.076
血清総コレステロール(mg/dl)	-0.004	0.005	0.366	0.001	0.004	0.873	-0.003	0.008	0.717	-0.014	0.012	0.256	0.027	0.020	0.185
心電図ST-T異常 (1=有、0=無)	0.252	0.557	0.651	-0.093	0.524	0.859	1.804	0.615	0.003	3.700	1.247	0.003	1.118	1.135	0.325
心房細動 (1=有、0=無) 注 <sup>1)</sup>	-	-	-	2.144	1.316	0.103	3.201	1.201	0.008	注 <sup>2)</sup>	-	-	-	-	-
例数 (人)		48		50		33		165		17		9		45	
	症例		250												
	対照														

注<sup>1)</sup> 脳出血例に心房細動を有する者がなかったため検討成績を示さず.

注<sup>2)</sup> 心房細動は皮質枝系脳梗塞と塞栓型と血栓型の分類基準項目となっているため塞栓型, 血栓型に分けた場合は検討せず. 山海他 (1992)より表7を再掲.

追跡であるため、16,000人年となり、サンプル数は充足していた。

## 8. コホート内症例対照研究

コホート研究では一般にベースライン時に循環器疾患の既往のない者を全員、分析の対象とするが、調査項目の中に生体試料を用いた高価な検査がある場合には研究費用が膨大となる。そこで循環器疾患が発生した時点で、コホート集団の中から、対照者を抽出(症例：対照=1：1～3等)し、症例と対照との間で比較を行う方法が考案されている(Breslow and Patton (1979))。この方法を用いると、通常の症例・対照研究で問題となる曝露要因に関する様々な bias がないため、精度を落とすことなく分析が可能である。一例として血清の脂肪酸構成と脳卒中発生との関連を分析した研究を示す(磯 他(1995))。血清の脂肪酸構成の測定は比較的高価で、その測定に時間がかかる。そのため、コホート対象者全員に対して測定する方法をとらずに、3農村集団男女40-79歳約1万人の血清を-80°Cで凍結保存した。そして脳卒中の発生があった時点で地域、年齢、性、血清保存年をマッチさせた健常者を対照(1：3マッチング)として選び、両者の血清脂肪酸構成を測定し、比較を行った。平均約5年間の追跡期間で、111人の脳卒中発生者を認め、対照333人との間で比較したところ、血清中のリノール酸の割合が高いほど、脳卒中の発生リスクが低くなる関連を認めた(磯 他(1995))。

また、コホート内症例対照研究の手法は分析しようとする疾病がすべて把握しきれない場合にも使われる。例えば脳卒中の病型をCT所見によって正確に病型分類した場合、約20～30%がCT検査を受けていないため、CT検査を受けた症例のみを選び、それに対して対照を選んで、発生要因の分析を行う場合である(山海 他(1992))。

表6にその成績の例を示す。3農村集団(人口8,717人)の平均7年間の追跡により脳卒中が246例発生し、そのうちCT検査を受けた181例(74%)を分析対象とした。この181例中、CT上異常所見がなかった症例19例とくも膜下出血31例を除いた131例に関して、対照を症例1例につき5例を性、年齢、地域をマッチングして抽出した。その結果、脳出血、穿通枝系脳梗塞、皮質枝系脳梗塞の間で、発生要因が異なることが示された。

## 9. 多変量解析

循環器疾患の発生には、一般の感染症とは異なり、複数の要因(危険因子)が関与していることが多い。そこで個々の危険因子の疾病発生への影響を検討する際、他の要因の影響を考慮して分析する必要がある。すなわち、これまでの知見に基づく交絡因子の影響を調整した上で、ある要因の疾病発生に対する相対危険度を算出する。問題となる交絡因子(問題とする危険因子以外で、疾病発生と関連のある因子)が1つか2つであれば交絡因子で層別解析して相対危険度を算出することは可能であるが、交絡因子が多くなると層別解析は大規模な調査でない限り困難である。

そこで統計的手法として多変量解析が適用される。循環器疾患の危険因子に関する多変量解析は米国のフラミンガム研究において初めて適用された(Dyer (1980))。その際の統計解析モデルはロジスティックモデルであった。このモデルは疾病の発生確率を  $Y$ 、危険因子を  $X_1$ 、交絡因子を  $X_2 \cdots X_i$  とすると

$$\text{logit}(Y) = \ln\{Y/(1-Y)\} = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_i X_i$$

すなわち、 $Y$  のロジット変換した値は  $X_1 \sim X_i$  の一次回帰式で表される。これを書き換えると

表7. 血清総コレステロール区分別に見た脳卒中中の年齢調整死亡率, 相対危険度, 米国男子, 35~57歳, 350,977人, 6年間のコホート研究.

血清総コレステロール (mmol/L)	調査数 (mg/dl)	くも膜下出血			脳内出血			非出血性脳卒中		
		死亡数	年齢調整	多変量調整*	死亡数	年齢調整	多変量調整*	死亡数	年齢調整	多変量調整*
<4.14	21,185	5	2.73	1.00	11	6.53	1.00	3	1.91	1.00
4.14-5.15	107,657	16	1.57	0.58	20	1.99	0.32	17	1.68	0.96
5.16-6.18	138,426	21	1.49	0.53	30	2.15	0.32	33	2.36	1.21
6.19-7.22	65,081	12	1.88	0.58	16	2.26	0.32	27	3.92	1.81
≥7.23	18,628	1	0.47	0.16	6	3.04	0.38	12	5.91	2.57
計	350,977	55	1.57	-	83	2.36	-	92	2.62	-
p値				0.40			0.05			0.06
偏回帰係数† (p値)				-0.0041(0.26)			-0.0055(0.06)			0.0063(0.002)

\* Cox 比例ハザードモデルにより, 年齢, 最小血圧, 喫煙本数, 人種を調整.

† Cox 比例ハザードモデルにより, 血清総コレステロール値を総変量として, 上記の共変量を調整して偏回帰係数を算出.  
Iso et al.(1989)より表2を再掲.

$$Y = \frac{1}{1 + \exp[-(\alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i)]} \quad \beta_1 \sim \beta_i : \text{偏回帰係数}$$

危険因子 ( $X_1$ ) の曝露無しの疾病の一定期間の発生確率を  $Y_0$ , 曝露ありの疾病発生確率を  $Y_1$  とすると

$$\begin{aligned} \text{logit}(Y_0) &= \alpha + \beta_1 \cdot 0 + \beta_2 X_2 + \dots + \beta_i X_i, \quad \text{logit}(Y_1) = \alpha + \beta_1 \cdot 1 + \beta_2 X_2 + \dots + \beta_i X_i \\ \beta_1 &= \text{logit}(Y_1) - \text{logit}(Y_0) = \ln[Y_1/(1 - Y_1)] - \ln[Y_0/(1 - Y_0)] \\ &= \ln \frac{Y_1}{1 - Y_1} - \ln \frac{Y_0}{1 - Y_0} \\ &= \ln [\text{オッズ比}] \end{aligned}$$

すなわち, 疾病のオッズ比  $= \exp(\beta_1)$  で求められる。

コホート研究において, 調査対象者の追跡期間が同じで, かつ疾病の発生頻度が小さい場合はオッズ比は相対危険度に近似できる。しかし, 各人で追跡期間が異なる場合や転出等により追跡不可能者がある場合は疾病の発生確率は人年法により算出することが望ましい。この場合に用いられることの多い多変量解析として Cox の比例ハザードモデルがある。

追跡期間中の  $T$  時点におけるハザード (ある一瞬の追跡期間における疾病の発生数をその中間時点の調査対象で除した値) についての関数を  $\lambda(t)$ , 下式の  $X_1 \dots X_i$  がすべて平均値の場合の  $t$  時点におけるハザード関数を  $\lambda_0(t)$  とする。危険因子を  $X_1$ , 交絡因子を  $X_2 \dots X_i$  とする。  $t$  は時間とする。

$$\lambda(t)/\lambda_0(t) = \exp[\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i]$$

すなわちハザード比は  $X_1$  から  $X_i$  の一次回帰式を指数関数変換した値として表される。危険因子 ( $X_1$ ) の曝露なしの疾病のハザード比は

$$\lambda(t)_{X_1=0}/\lambda_0(t) = \exp[\beta_2 X_2 + \dots + \beta_i X_i]$$

曝露有りのハザード比は

$$\lambda(t)_{X_1=1}/\lambda_0(t) = \exp[\beta_1 + \beta_2 X_2 + \dots + \beta_i X_i]$$

従って疾病発生の相対危険度は

$$\begin{aligned} &\lambda(t)_{X_1=1}/\lambda(t)_{X_1=0} \\ &= \lambda(t)_{X_1=1}/\lambda_0(t) \div \lambda(t)_{X_1=0}/\lambda_0(t) \\ &= \exp[\beta_1 + \beta_2 X_2 + \dots + \beta_i X_i] / \exp[\beta_2 X_2 + \dots + \beta_i X_i] = \exp(\beta_1) \end{aligned}$$

となり, 疾病発生の相対危険度  $= \exp(\beta_1)$  といった簡潔な式で求められる。

このモデルを白人 35-57 歳男子 350,977 人の平均 6 年間の追跡調査成績 (Iso et al. (1989)) に適用した例を示す (表 7)。血清総コレステロール値と脳内出血の死亡率との間には, 負の関連が, 非出血性脳卒中の死亡率との間には, 正の関連が見られた。比例ハザードモデルを適用したところ, 血清総コレステロール値 (mg/dl) の偏回帰係数 ( $\beta_1$ ) は脳内出血で  $-0.0055$ , 非出血性脳卒中で  $0.0063$  であった。そこで, 血清総コレステロール値が  $40$  mg/dl 低下すると, 疾病発生の相対危険度は, 脳内出血で

$$\exp(-0.0055 \times -40) = 1.25$$

非出血性脳卒中で

$$\exp(0.0063 \times -40) = 0.78$$

すなわち、血清総コレステロール値が 40 mg/dl 低下は脳内出血の死亡率の 25% の増加、非出血性脳卒中の 22% 低下と関連すると解釈できる。

## 10. Regression dilution bias とその調整

曝露要因に関するデータは本来ならその個人の平均的な曝露状況を代表するデータであることが望ましいが、個人内での変動や測定上の誤差から、その平均的な値からのバラツキが存在する。疫学研究を実施するにあたり、調査の精度を高めてこのバラツキを少なくする努力を惜しんではならないが、多くの曝露要因においてはある程度のバラツキは避けることができない。曝露要因のバラツキがランダムに起こった場合には、曝露要因と疾患発生との関連は弱まる。バラツキが大きい場合、実際に関連が存在する場合でもその関連が検出されないこともあり得る。このバイアスは regression dilution bias と呼ばれている。表 8 は米国のフラミンガム市のコホート研究の成績 (MacMahon et al. (1990)) で、ベースライン調査時の最小血圧値で 5 つのグループに分けた際、それぞれのグループの平均値が算出できる。2 年後、4 年後での最小血圧値のデータをベースライン調査時の 5 グループごとに、それぞれの平均値を算出し、ベースライン調査時での平均値と比べてみる。すると第 2 番目の最小血圧 80~89 mmHg での平均値には大きな差はないが、第 1 番目の最小血圧 79 mmHg 以下の平均値では、2 年後、4 年後上昇し、第 3 番目以降の最小血圧 90 mmHg 以上の平均値では、2 年後、4 年後低下している。すなわち、平均値への回帰が起こっている。そのため、最小血圧の平均値の範囲 (range) はベースライン調査時の 47.7 mmHg から、2 年後は 31.6 mmHg、4 年後 25.5 mmHg とそれぞれ 66%、60% 減少している。

ベースライン調査時の 5 つのグループでの最小血圧値の平均値を横軸に、脳卒中、虚血性心疾患の発生率を縦軸にプロットして、その傾きを算出する場合には後者の最小血圧の平均値の幅が 60% 減少しているため、後者の傾きは前者に  $1/0.6 = 1.67$  を乗じた値となる。

この regression dilution bias を調整する概念は多変量解析 (ロジスティックモデル、Cox 比例ハザードモデル等) の成績にも適用できる。すなわち偏回帰係数に 1.67 を乗じることによって、調整することができる。換言すると、最小血圧と脳卒中、虚血性心疾患発生との関連 (相対危険度) は regression dilution bias により過小評価されており、実際は算出された相対危険

表 8. ベースライン調査時の最小血圧値の区分別に見た 2 年後、4 年後の最小血圧の平均値。米国フラミンガム市、男女 3,776 人。

ベースライン 最小血圧区分 mmHg	調整数	最小血圧の平均値		
		ベースライン	2 年後	4 年後
1: ≤ 79	1719	70.8	75.7	76.2
2: 80-89	1213	83.6	83.0	83.9
3: 90-99	566	93.5	91.2	91.3
4: 100-109	186	103.4	99.2	98.5
5: ≥ 110	92	116.4	107.3	104.7
最小血圧の平均値の範囲		47.7	31.6	28.5

MacMahon et al. (1990) より表 3 (改変) を再掲。

度の1.67倍の関連を有することになる。

Regression dilution biasの調整は、個人の平均的な曝露状況と疾病発生との関連を明らかにするための統計的処理である。しかしながら関連が有意でなかった場合、regression dilution biasを調整することによって、もしかしたら関連は存在するかもしれないという議論は控えるべきであり、あくまで因果関係が確定している関連について、その関連の強さを調整する手段として用いるのが妥当である。

## 11. コホート研究と因果関係の判断

疫学研究において要因と疾病発症との関連が確認されたとき、要因が疾病発生の原因であると判断するには一般に以下のような条件が提示されている。まず、「要因が疾病発生の前に存在する(temporality)」ことである。コホート研究の場合、当該疾病の既往歴のない人々を対象としてコホートを設定するため、この条件は自動的に保証される。第2点は「関連が一致している(consistency)」ことである。この条件は異なる集団、異なる研究者においても同様な関連が見られることで、循環器疾患の主要な研究は国内外で数多く存在するため、この条件に合うかどうかの検討は比較的容易である。第3点は「強固な関連(strength)が存在する」ことであるが、相対危険度がどの程度以上が強固な関連であるかの基準はない。また、曝露要因の種類によっても異なり、例えば血圧値等の身体要因であれば2-3以上というようなおおまかな経験上の判断しかできない。しかし、生活・環境要因については相対危険度が1.5-2くらいであっても因果関係を否定できるとは限らない。従って強い関連があれば、因果関係がより強く示唆されるというくらいに考えるべきである。

第4点は「量・反応関係(dose-response relationship)が存在する」ことである。曝露量の多いほど、疾病の発生率が高いことを指し、因果関係の判断のための有力な条件である。しかし、曝露量がある量以上にあがらないと疾病の発生率が増加しない場合もあるため、この条件にあてはまらないからといって、因果関係を否定できるわけではない。

以上の第1～第4の条件はコホート研究の中で一括して検討でき、このことがコホート研究の優れた点である。しかしながら、得られた関連の「生物学的な説明(biological plausibility)」には、コホート研究のみでは不十分であり、動物実験、病理学、細胞・分子の生物学的レベルでの成績が必要となることが多い。

## おわりに

これらの循環器疾患のコホート研究の多くは国内外を問わず、地域や職域の集団で数千人を対象として5～20年間の追跡を行い、数々の成果を上げた。海外で最も歴史が古く、有名な研究として、フラミンガム研究があるが、その他にもプエルトリコでの研究、ホノルルの日系人を対象とした研究、シカゴ市の企業従事者の研究、英国の企業従事者の研究、フィンランドのノースカレリアでの研究、ノルウェーのオスロ市での研究等が挙げられる(MacMahon et al. (1990))。これらの欧米諸国の研究は虚血性心疾患のリスクファクターの解明が主な目的であったが、これらの研究に共通して、高血圧、高コレステロール血症、喫煙が虚血性心疾患の発症リスクを高めることが示された(Dyer (1980), MacMahon et al. (1990))。わが国においても、大阪の企業従事者の研究(Kitamura et al. (1994))、広島・長崎の原爆被爆者の研究(Kodama et al. (1990))で、上記の危険因子の関与が確かめられている。わが国では虚血性心疾患に比べ、脳卒中の発生が多いことから、脳卒中の危険因子に関するコホート研究が主として行われてきた。本稿で例を挙げた地域住民のコホート研究以外にも、福岡の久山町での研究(上田(1987))、

熊本・福岡の牛深・田丸丸での研究 (Kimura et al. (1977)), 広島・長崎の原爆被爆者の研究 (Szatrowski et al. (1984)), 新潟の新発田市での研究 (Nakayama et al. (1997)), 愛媛の赤羽町・旭町での研究 (Okada et al. (1976)) 等があり, 脳卒中の危険因子として高血圧の関与が大きいこと, 脳卒中の病型により, 脂質代謝異常, 心電図や眼底所見の異常の関与が異なることが示されている。これらのコホート研究は本稿で述べた, 対象の設定, 曝露要因の把握, 循環器疾患の発生の把握, サンプル数, 等の統計学上の問題点をクリアしており, 信頼できる疫学データを提供してきた。近年, 動脈硬化の進展や循環器疾患のコホート研究の血液凝固・線溶系の因子の関与が示唆され, 検討され始めている (Meade et al. (1986), Hamsten et al. (1987))。また, 様々な生活習慣が循環器疾患の発生にどうかかわるのかについても, 検討されつつある。その際, 必要なサンプル数が従来よりも多くなるため, 幾つかのコホート研究が同一の方法により疫学データを収集し, 分析する方法がとられたり (ARIC Investigators (1989)), より大きなサンプル数のコホート研究が立案, 実行されている (Stampfer et al. (1985))。

### 参 考 文 献

- ARIC Investigators (1989). The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. *American Journal of Epidemiology*, **129**, 687-702.
- Breslow, N. E. and Patton, J. (1979). Case-control analysis of cohort studies, *Energy and Health* (eds. N. E. Breslow and A. S. Whittemore), 226-242, SIAM, Philadelphia.
- Dyer, T. R. (1980). *The Framingham Study. The Epidemiology of Atherosclerotic Disease*, Harvard University Press, Cambridge.
- Kodamda, H., Sasaki, H. and Shimizu, Y. (1990). Trends of coronary heart disease and its relationship to risk factor in a Japanese population: a 26-year follow up, Hiroshima Nagasaki Study, *Japanese Circulation Journal*, **54**, p.414.
- Hamsten, A., De Faire, U., Walldius, G., Dahlen, G., Szamosi, A., Landou, C., Blomback, M. and Wiman, B. (1987). Plasminogen activator inhibitor in plasma: risk factor for recurrent myocardial infarction, *Lancet*, **2**, 3-9.
- 磯 博康 (1996). 栄養と病型別にみた脳卒中に関する日米疫学共同研究, 上原記念生命科学財団研究報告集, **10**, 14-15.
- Iso, H., Jacobs, D. R. Jr., Wentworth, D., Neaton, J. D. and Cohen, J. D. (1989). Serum cholesterol levels and six-year mortality from stroke in 350,977 men screened for the multiple Risk Factor Intervention Trial, *New England Journal of Medicine*, **320**, 904-910.
- Iso, H., Kitamura, A., Shimamoto, T., Sankai, T., Naito, Y., Sato, S., Kiyama, M., Iida, M. and Komachi, Y. (1995). Alcohol intake and the risk of cardiovascular disease in middle-aged Japanese men, *Stroke*, **26**, 767-773.
- 磯 博康, 嶋本 喬, 工藤美奈子, 山海知子, 今野弘規, 大平哲也, 梅村詩子, 小池和子, 佐藤眞一, 飯田稔, 小町喜男 (1995). 血清脂肪酸構成と脳卒中発症との関連——コホート内症例対照研究——, 日本公衛誌, **42** (Suppl), p.721.
- Kagan, A., Popper, J. S. and Rhoads, G. G. (1980). Factors related to stroke incidence in Hawaiian Japanese men, the Honolulu heart study, *Stroke*, **11**, 14-21.
- Kimura, N., Nakayama, Y. and Tashiro, H. (1977). The 15-year follow-up population survey on cerebrovascular disease in rural Kyushu, Japan, *Japanese Journal of Medicine*, **16**, 142-144.
- Kitamura, A., Iso, H., Naito, Y., Iida, M., Konishi, M., Folsom, A. R., Sato, S., Kiyama, M., Nakamura, M., Sankai, T., Shimamoto, T. and Komachi, Y. (1994). High density lipoprotein cholesterol and premature coronary heart disease in urban Japanese men, *Circulation*, **89**, 2533-2539.
- MacMahon, S., Peto, R., Cutler, J., Collins, R., Sorlie, P., Neaton, J., Abbott, R., Godwin, J., Dyer, A. and Stamler, J. (1990). Stroke and coronary heart disease. Part 1, prolonged differences in blood pressure: prospective observational studies for the regression dilution bias, *Lancet*, **335**, 765-774.
- Meade, T. W., Mellow, S., Brozovic, M., Millen, G. J., Chakrabarti, R. R. and North, W. R. (1986). Haemostatic function and Ischemic heart disease: principal results of the Northwick Park heart



- study, *Lancet*, **2**, 533-537.
- Nakayama, T., Date, C., Yokoyama, T., Yoshiike, N., Yamaguchi, M. and Tanaka, H. (1997). A 15.5-year follow-up study of stroke in a Japanese provincial city. The Shibata study, *Stroke*, **28**, 45-52.
- Okada, H., Horibe, H., Ohno, Y., Hayakawa, N. and Aoki, N. (1976). A prospective study of cardiovascular disease in Japanese rural communities, Akabane and Asahi, part 1: evaluation of risk factors in the occurrence of cerebral hemorrhage and thrombosis, *Stroke*, **7**, 599-607.
- 山海知子, 宮垣武司, 磯 博康, 嶋本 喬, 飯田 稔, 谷垣正人, 内藤義彦, 佐藤眞一, 木山昌彦, 北村明彦, 小西正光, 寺尾敦史, 土井光徳, 小町喜男 (1991). CT 所見を中心とした脳卒中の疫学的研究——農村集団における病型別にみた発生割合の検討——, *日本公衛誌*, **38**, 901-909.
- 山海知子, 磯 博康, 嶋本 喬, 宮垣武司, 飯田 稔, 谷垣正人, 内藤義彦, 佐藤眞一, 木山昌彦, 北村明彦, 小西正光, 寺尾敦史, 土井光徳, 児島三郎, 小町喜男 (1992). CT 所見を中心とした脳卒中の疫学的研究——コホート内症例対照研究による脳出血ならびに脳梗塞の病型別発生要因の検討——, *日本公衛誌*, **39**, 410-419.
- 重松逸造, 柳川 洋 (1991). 『新しい疫学』, 日本公衆衛生協会, 東京.
- 鳴谷亮一, 小町喜男, 渡辺 孝 (編) (1976). 『日本人の栄養と循環器疾患』, 保健同人社, 東京.
- Shimamoto, T., Komachi, Y., Iida, M., Doi, M., Iso, H., Sato, S., Nakanishi, N., Terao, A., Naito, Y. and Kojima, S. (1989). Trends for coronary heart disease and stroke and their risk factors in Japan, *Circulation*, **79**, 503-515.
- 嶋本 喬, 稲田 敏, 土井光徳, 磯 博康, 福内恵子, 佐藤眞一, 北村明彦, 宮垣武司, 笹本知子, 吹野洋子, 小町喜男 (1989). 茨城県協和町における循環器疾患対策『循環器疾患の変貌』(小町喜男 他編著), 449-463, 保健同人社, 東京.
- Stampfer, M. J., Willett, W. C., Colditz, G. A., Rosner, B. A., Speizer, F. E. and Hennekens, C. H. (1985). A prospective study of postmenopausal estrogen therapy and coronary heart disease, *New England Journal of Medicine*, **313**, 1044-1049.
- Szatrowski, T.P., Peterson, A. V. Jr., Shimizu, Y. et al. (1984). Serum cholesterol, other risk factors, and cardiovascular disease in a Japanese cohort, *Journal of Chronic Diseases*, **37**, 569-584.
- 上田一雄 (1987). 久山町における循環器疾患の予防および管理に関する研究, *日循協誌*, **5**, 1838-1840.

## Cohort Study on Cardiovascular Disease and Biostatistics

Hiroyasu Iso, Takashi Shimamoto, Tomoko Sankai, Takeshi Tanigawa  
and Tetsuya Ohira

(Institute of Community Medicine, University of Tsukuba)

The main purpose of cohort studies for cardiovascular disease is to identify and examine the magnitude of constitutional and lifestyle factors in relation with risk of cardiovascular disease. Cohort studies have been conducted in rural communities where population is stable, and urban employees in which follow-up is easy to be conducted until the age of 60. It is important to examine risk variables in health examinations and to conduct complete surveillance for cardiovascular disease using standardized methods. To determine a sample size is also important. The sample size depends on risk variables examined. In general, a few thousand person-years are required to examine the relation between constitutional risk factors such as blood pressure, and cardiovascular disease, while a larger sample size in case of lifestyle variables. For example, then thousand person-years are needed to examine the relation between nutrition and cardiovascular disease. Use of multivariate analyses is common to examine relative risks because multiple constitutional and lifestyle variables may cause cardiovascular disease. Adjustment for regression dilution due to variations of risk variables can be conducted when estimating relative risks. Existence of the relation between risk factors and cardiovascular disease does not necessarily mean causal relationship. To establish causality, we need consistency of the results of cohort studies and biological plausibility.