

疫学研究のデザインと相対リスクの推定

統計数理研究所 佐藤 俊 哉

(1992年5月 受付)

1. はじめに

多くの疫学研究の目的は、疾病の原因と想定されているリスク要因へ曝露されることによって、問題としている疾病の発生が増加するかどうかを調べることにある。そのための代表的な研究デザインがコホート研究とケース・コントロール研究である。コホート研究は、疾病を発生していない集団をリスク要因の曝露を受けたグループと曝露を受けなかったグループに分け(これは、一つのコホート全体の曝露状況を調べる場合もあるし、全曝露者、全非曝露者からのランダムサンプルでもかまわない)、一定期間追跡した後それぞれのグループの疾病発生状況を比較する「前向き研究」と呼ばれるデザインである。これに対してケース・コントロール研究は、研究実施時点ですでに疾病を発生している対象者からのランダムサンプルであるケースグループと、疾病を発生していない対象者からのランダムサンプルであるコントロールグループについてリスク要因への曝露状況を比較する「後ろ向き研究」と呼ばれるデザインである。

最近関心を集めている慢性疾患に対する疫学研究では、1) リスク要因への曝露から疾病発生までの期間が長い、2) 一般に疾病発生率が低い、という特徴がある。このため短期間で研究に必要なケース数を確保できるケース・コントロール研究から疾病に対する曝露の効果を評価する機会が多い。しかし、ケース・コントロール研究では、コントロール選択の際のバイアス、リコールバイアスなどの、解析ではどう対処することもできないバイアスが混入する可能性が高いことが常に問題となる (Kleinbaum et al. (1982), Rothman (1986))。

一方、コホート研究はケース・コントロール研究で問題となる種類のバイアスの影響は受けにくい、疾病がまれである場合には大規模コホートを長期間にわたって観察しなければならず、従ってコストのほとんどが健常者を追跡することに費やされてしまうという問題がある。冠状動脈疾患のリスクを下げることを目標に LDL コレステロールレベルを低下させてその効果を調べたランダム化臨床試験である Lipid Research Clinics Coronary Primary Prevention Trial (LRC-CPPT) を例にとってみよう。この研究では、3,806名の男性を平均7.4年間追跡した結果、対象とした冠状動脈疾患による死亡または非致死性の心筋梗塞の発症は342名、全対象の9%であった (Lipid Research Clinics Program (1984))。従って、LRC-CPPTのコストのほとんどは、対象とする疾病を発症しなかった3,464名の追跡に費やされたことになる。それでもコホート研究では、リスク要因への曝露から疾病の発生までを時間を追って観察できるため、疾病の自然史を調べることが可能であるし、観察の時間的順序や論理の流れが実験に最も近いという利点が存在する。

コホート研究とケース・コントロール研究の特徴は相互に補完的であるため、Mantel (1973) はコホート研究を途中からケース・コントロール研究に転換することを提案し、またケース・コントロールサンプリングから得られるオッズ比がコホートの情報をすべて用いて得られるオッズ比と同じものであることを示した。このコホート研究とケース・コントロール研究を混

合したデザインは、あくまでもコホート研究の枠組みの中で行なわれるため、ケース・コントロール研究に特有のバイアスを避けることができるし、曝露や主要なリスク要因に関する情報は全ケース（または全ケースからのサンプル）とケースでない者からのサンプルについてのみに解析すればいいので、かなりのコストの減少が期待できる。

Mantel が提案したデザインには、コホート研究では相対リスクを直接推定できるのに対し、途中からケース・コントロール研究に転換することで、推定できる効果の指標がオッズ比になってしまうという問題点がある。オッズ比は相対リスクの近似として解釈できる場合以外には、必ずしも疫学的に意味のある効果の指標とはならない (Greenland (1987)). 幸いなことに、「まれな疾病」を対象としている場合にはオッズ比が相対リスクのよい近似となり、これはケース・コントロール研究を実施する際の理論的な裏付けとなっている (Cornfield (1951)). さらに、最近の疫学理論の発展、生存解析の発展にともなって、ケース・コントロール研究でもコントロールの選択を適切に行なえば「まれな疾病である」という仮定とは無関係に、オッズ比を用いて相対リスクを（近似的にではなく）推定できることが分かってきた (Kupper et al. (1975), Liddell et al. (1977), Greenland and Thomas (1982)).

本研究では新しい疫学研究のデザインである「コホート研究の枠組みの中でのケース・コントロール研究」について、適切なコントロール選択の方法、相対リスクの推定に必要な条件、などに関するレビューを行なう。説明を簡単にするため、交絡や誤分類その他のバイアスは全てないと仮定し、またリスク要因への曝露は「あり、なし」の2分類、同様に疾病の発生も「あり、なし」に2分類されている場合を考える。また、相対リスクやオッズ比などの値は全て標本数が多い場合の期待値であり、曝露グループと非曝露グループの疾病発生率の比は観察期間を通じて一定であると仮定する。

次章では相対リスクの定義、基本的なケース・コントロール研究のデザインと有名な「まれな疾病の仮定」との関係について述べ、ついでネステッド・ケース・コントロール研究、ケース・コホート研究について解説する。なお、次章で述べる内容はケース・コントロール研究一般についての議論であり、「コホート研究の枠組みの中でのケース・コントロール研究」だけに成り立つわけではない。しかし、ケース・コントロール研究は基本的に「仮想的なコホート」の中で実施するものであり (Rothman (1986), p. 65), 現実のコホート研究の枠の中でケース・コントロール研究を扱うことにより、論点がより明確になると信じる。

2. コホート研究とケース・コントロール研究

相対リスクには2種類の定義があり、これは疾病発生の指標として何を用いるかに依っている (Elandt-Johnson (1975), Miettinen (1976)). 急性中毒などの曝露の短期的な影響を評価する場合には、疾病発生割合 (incidence proportion, または cumulative incidence rate と呼ばれる) が疾病発生の指標として用いられる。疾病発生割合は、閉じたリスク集団が一定期間中に疾病を発生する割合の期待値である（「閉じた」とは、疾病の発生、死亡、転出やその他の脱落でリスク集団のメンバは減少していくが、出生以外の理由、例えば転入、では補充されない状態をいう。またリスク集団は population at risk の訳語として用いる）。このとき、曝露グループと非曝露グループの疾病発生割合の比であるリスク比 (risk ratio, RR) が相対リスクのひとつの定義である。

一方、慢性疾患の疫学研究でリスク要因への曝露の長期的な影響を評価する場合や、生存解析などでは疾病発生の指標として疾病発生率 (incidence rate, または incidence density と呼ばれる) を用いる。疾病発生率とは、リスク集団の観察人-時間1単位当りの期待疾病発生数で

あり、曝露グループと非曝露グループの疾病発生率の比 (incidence rate ratio, *IRR*) が相対リスクのもうひとつの定義である。以下の議論では、疾病発生率は観察時間の経過とともに変化するが (例えば、観察時間として年齢をとれば、一般に時間の経過とともに疾病発生率は増加するだろう)、曝露グループと非曝露グループの疾病発生率の比 *IRR* は観察期間を通じて一定であると仮定する。

どちらの相対リスクも、閉じたコホートを対象としたコホート研究からは直接推定することができる。通常の人口動態——出生、疾病の発生、死亡、転入、転出——を全て許す「開いた」コホートでは、疾病発生割合を定義することはできないが、疾病発生率は求めることができるので *IRR* は推定可能である。それでは、ケース・コントロール研究から求められるオッズ比を相対リスクとして解釈するにはどのような条件が必要であり、またそれはどちらの相対リスクなのだろうか。

図1に架空のコホート研究を示す。観察開始時点では疾病を発生していないアからカまで6名のコホートを一定の観察期間追跡した結果である。○はコホートメンバがその時点で対象とする疾病を発生したことを示し、×はその時点で死亡や転出などにより観察の打ち切りが生じたことを示す。ケース・コントロール研究ではケースグループを、コホートの中で観察期間中に疾病を発生した全ケースからランダムに一定数選択する。次に、コントロールグループとしては、観察終了時のリスク集団 (観察終了時まで観察が続けられているコホートメンバ、対象とする疾病を発生した者、または他の疾病を発生、死亡、転出などの理由で観察が打ち切られた者は除かれる) から一定数をランダムに選択する。図1のコホートを例に取ると、全ケースは (ア、ウ、オ) であるから、この中からランダムにケースが選ばれ、またコントロールは観察終了時のリスク集団 (イ、カ) からランダムに選ばれる。

全コホートでの疾病発生状況の期待値と、ケース・コントロール研究を行なって観察される結果の期待値を表1の記号を用いて表わすことにする。ケース・コントロール研究でのケースサンプル、コントロールサンプルは、全コホートでのケース、非ケースからのランダムサンプルである。ケース、コントロールそれぞれについて、リスク要因の曝露を受けた割合がどのくらいであるかを調べて得られるオッズ比は曝露オッズ比と呼ばれ、

$$OR_E = \frac{a}{b} / \frac{c}{d} = \frac{ad}{bc},$$

で与えられる。

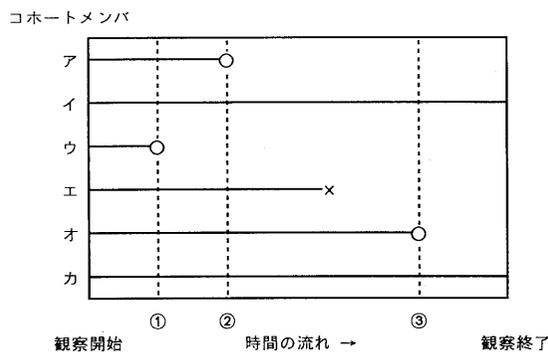


図1. コホートメンバの追跡状況。○: 対象とする疾病を発生, ×: 観察の打ち切り。

表1. ケース・コントロールサンプリング.

全コホート				ケース・コントロール			
疾病発生	曝露		合計		曝露		合計
	あり	なし			あり	なし	
あり	A	B	M_1	ケース	a	b	m_1
なし	C	D	M_0	コントロール	c	d	m_0
合計	N_1	N_0	N				

全コホートが観察できた場合、つまり観察開始時のリスク集団を曝露グループと非曝露グループに分類し、観察終了時までそれぞれのグループで疾病を発生した割合を調べて求めるオッズ比は疾病オッズ比、

$$OR_D = \frac{A}{C} / \frac{B}{D} = \frac{AD}{BC},$$

である。このとき、図1のように閉じたコホートでかつ観察打ち切りがリスク要因への曝露とも疾病の発生とも無関係に起こる場合、 $a/b=A/B$ かつ $c/d=C/D$ であるから、曝露オッズ比は疾病オッズ比と一致する。一方、リスク比は定義から、

$$RR = \frac{A}{N_1} / \frac{B}{N_0} = \frac{AN_0}{BN_1},$$

であるが、疾病オッズ比は $OR_D = A(N_0 - B) / [B(N_1 - A)]$ と書き換えられるので、 $A/N_1, B/N_0$ がともに非常に小さい場合、つまり「曝露グループ、非曝露グループともに疾病発生がまれな場合」、 OR_D は RR のよい近似となる。したがって、ケース・コントロール研究においても、まれな疾病を対象としている限り OR_E から近似的に RR を推定することができる (Cornfield (1951)).

また、観察対象が開いたコホートでかつ観察期間を通じて曝露割合（ある時点でのリスク集団中の曝露グループの割合）が一定の場合、いいかえると曝露割合に関して安定なダイナミック・コホートである場合、 IRR は、

$$IRR = \frac{A}{N_1 T} / \frac{B}{N_0 T} = \frac{A}{CT} / \frac{B}{DT} = \frac{AD}{BC},$$

となって、曝露オッズ比は疾病発生率の比に一致する (Greenland and Thomas (1982), Rothman (1986), p. 63). ただし、表1と同様に、 N_1, N_0 は曝露グループ、非曝露グループの観察開始時のリスク集団、 C, D はそれぞれの観察終了時のリスク集団、 T は観察時間である。

「まれな疾病の仮定」が必要であるのは、閉じたコホートの場合に疾病オッズ比を（従って、ケース・コントロール研究から得られる曝露オッズ比も）近似的にリスク比として解釈する場合だけである (Cornfield (1951)). また、疾病がまれな場合には疾病発生率の比 IRR も近似的にリスク比として解釈でき、しかも疾病オッズ比よりもよい近似を与える (Greenland and Thomas (1982)). ケース・コントロール研究から得られる曝露オッズ比と相対リスクの関係をまとめると、1) 疾病発生率の比として解釈したい場合には曝露に関して安定したコホートであることが重要であり、2) リスク比の近似として解釈したい場合には「まれな疾病」であるか

どうか重要である，ということになる。

3. ネステッド・ケース・コントロール研究

前章のケース・コントロール研究はコホート研究を単に途中でケース・コントロール研究に転換しただけであり，コホートメンバを一定期間追跡し続けているというコホート研究の特徴を十分には活かしていない。コホート研究を実施している以上，例えば図1中のウが疾病を発生した時点①，アが疾病を発生した時点②，などの情報を我々は利用できるはずである。

ネステッド・ケース・コントロール研究（シンセティック・ケース・コントロールとも呼ばれる）では疾病を発生した時点の情報を用いてコントロールの選択を行なう。コントロールの選択方法にはいくつかのオプションがあるが，最も一般的な方法は「ケースが発生した時点のリスク集団から1人またはそれ以上のコントロールをランダムに選択する」という「時点マッチング」と呼ばれる方法である（Liddell et al. (1977), Miettinen (1977), Prentice and Breslow (1978)）。ケースグループとしては全ケースを用いることが多いが，もちろん全ケースからのサンプルでも構わない。再び図1を例に取ると，ウが疾病を発生した時点①ではリスク集団は（ア，イ，ウ，エ，オ，カ）であり，アが疾病を発生した時点②では（ア，イ，エ，オ，カ），オが疾病を発生した時点③では（イ，オ，カ）がリスク集団であるから，それぞれの時点でのリスク集団からコントロールをランダムに選ぶことになる。このデザインは閉じたコホート，開いたコホートどちらにも用いることができる。

コントロール選択の際には，将来疾病を発生する者やケース本人もケースのコントロールに成り得ることに注意しなければならない。時点①を例に取ると，アとオは将来ケースとなり，ウはケース本人であるが，ウのコントロールになる可能性がある。また，同一の対象者が複数の時点でコントロールとして選択されることもあり，（イ，オ，カ）は時点①，②，③全部でコントロールとして選ばれることも起こり得る。このようにケース発生時点でマッチして収集されたケースとコントロールの解析には，通常のマッチングを行なった場合の解析法（Mantel-Haenszelの方法，あるいは条件付きロジスティック回帰など（Breslow and Day (1980), Kleinbaum et al. (1982), Rothman (1986)））をそのまま用いてオッズ比を推定し，信頼区間を求めることができる（従って，統計パッケージを用いて簡単に計算できる）。しかも，この場合のオッズ比は正確に疾病発生率の比 *IRR* を推定しており，これには「まれな疾病の仮定」を必要としない（Liddell et al. (1977), Greenland and Thomas (1982)）。

時点マッチングを行なっても，リスク集団からのランダムサンプルではなく，1) 観察期間を通じて疾病の発生や観察の打ち切りがなかった者（図1ではイとカ）のみ，2) 対象とした疾病を発生しなかった者（イ，エ，カ）のみ，あるいは3) 曝露と関係があるが研究対象ではない疾病を発生した者は除外（エだけが除外の可能性あり），をコントロール・プールとしてコントロールを選択した場合，*IRR* の推定にバイアスが生じる。最もバイアスが大きいのは2)の将来のケースをコントロールの候補から除外する場合で，1)のバイアスのほとんどは2)に由来し（つまり，打ち切り例のみをコントロールの候補から除いた場合はあまり影響しない），3)には問題となるほどの影響はない（Lubin and Gail (1984)）。

「ケースが発生した時点のリスク集団からランダムに1人またはそれ以上のコントロールを選択する」という方法では，一度コントロールとして選ばれたコホートメンバが，別の時点で（何度でも）コントロールとして選ばれる可能性がある。一度コントロールとして選択されたコホートメンバは，それ以降はコントロールの候補から除外する（ただしケースにはなり得る）ことにすると，コントロールグループの情報は若干増え，*IRR* の推定効率が改善されることが

期待できる。サンプリング方式と相対リスクの推定が複雑にはなるが、このようなコントロール選択の方法も提案されている (Robins et al. (1986, 1989), Prentice (1986b)). しかし、コントロールの選択に複雑な修正を行なっても期待されるほどの効率の増加はないようである (Langholz and Thomas (1991)).

4. ケース・コホート研究

ネステッド・ケース・コントロール研究ではケースが発生した時点でコントロールを選択するため、特定のリスク要因と複数の疾病との関係を調べたい場合、調べる疾病の数だけコントロールグループが必要になる。しかし、コホート研究では特別な考慮をしなくても、一つのリスク要因が複数の疾病に影響するかどうかを調べることができる。コホート研究の枠の中でケース・コントロール研究を実施しているのだから、複数の疾病を対象とする場合でもコントロールグループは一つでいいはずである。

このような考え方から提案されたのがケース・コホート研究 (case-cohort studies (Prentice (1986a)), あるいは hybrid epidemiologic designs (Kupper et al. (1975)), case-base studies (Miettinen (1982)) とも呼ばれる) である。ケース・コホート研究では、コントロールは観察開始時のリスク集団、つまり全コホートメンバからランダムに選択される (このコントロールグループをサブコホートと呼ぶ)。ケースグループは全ケースを用いても、ケースのサンプルを用いてもよい。明らかに、このデザインでもコントロールグループの中に将来疾病を発生する者が含まれる可能性があり、解析の際には同一人物がケースグループとコントロールグループに同時に属するということが起こり得る。図1の例では、ケースグループは (ア, ウ, オ) またはそのサンプルで構成され、コントロールグループは全コホートメンバ (ア, イ, ウ, エ, オ, カ) からランダムに選択される。

このデザインのもとで観察される結果の期待値を表2の記号で表わすことにする。ケースサンプルはケース・コントロール研究と同様に全ケースからのランダムサンプルであるが、コントロールサンプルは全コホートからのランダムサンプルとなっている。ケース・コホートデータから曝露オッズ比を求めると、

$$OR_E^* = \frac{a}{b} \frac{n_1}{n_0} = \frac{an_0}{bn_1},$$

となる。このとき、閉じたコホートでかつ観察打ち切りがリスク要因への曝露とも疾病の発生とも無関係に起こる場合、 $a/b = A/B$, $n_1/n_0 = N_1/N_0$ であるから、曝露オッズ比は正確にリスク比に一致する (Kupper et al. (1975), Miettinen (1982)). また、観察対象が開いたコホート

表2. ケース・コホートサンプリング.

全コホート				ケース・コホート			
疾病発生	曝露		合計		曝露		合計
	あり	なし			あり	なし	
あり	A	B	M ₁	ケース	a	b	m ₁
なし	C	D	M ₀				
合計	N ₁	N ₀	N	サブコホート	n ₁	n ₀	n

でかつ観察期間を通じて曝露割合が一定の場合、曝露オッズ比は疾病発生率の比に一致する ($OR_E^* = IRR$, Greenland et al. (1986)).

このどちらの場合も「疾病がまれである」という仮定とは無関係に成立する。信頼区間を求めるために必要な OR_E^* の推定値の近似的な分散は、コホートそのものを推測の対象としている場合 (Kupper et al. (1975)), コホートがもっと大きい標的集団からのサンプルである場合 (Miettinen (1982)), について与えられている。また、 OR_E^* の最尤推定量とその近似的な分散が最近提案されている (Sato (1992b)). [Kupper et al. (1975) の層別解析は誤っているので注意が必要である。ケース・コホート研究での正しい層別解析の方法は Greenland(1986), Sato (1992a) に述べられている。] また、競合危険 (研究対象ではない疾病を発生し、リスク集団から除かれる場合) による影響を補正してリスク比を推定する方法も提案されている (Flanders and Louv (1986), Flanders et al. (1990)).

ネステッド・ケース・コントロール研究 (あるいはもっと直接には生存解析) と同様な考え方から、ケースとケースが発生した時点のサブコホート中のリスク集団を比較することも可能である。ケース発生時点の情報を利用するケース・コホート研究は、閉じたコホートの場合について詳しく検討されている (Prentice (1986a)), このようにして得られるオッズ比は正確に疾病発生率の比に一致し、しかも「まれな疾病の仮定」を必要としない。ただし、相対リスクの推定は複雑であり、現在のところ汎用統計パッケージには組み込まれていない (Prentice (1986a), Wacholder et al. (1989)).

5. 効率の比較

ネステッド・ケース・コントロール研究もケース・コホート研究も、ともにコホート研究の枠の中でケース・コントロール研究を実施するための研究デザインであり、曝露に関する情報、交絡要因その他のリスク要因に関する情報はケースと (コントロールとして) コホートからのサンプルについてのみ必要であるため、コストの節約が期待できる。しかし、当然のことではあるが、これらの研究デザインから得られる相対リスクの推定精度はコホートの情報を全て使って得られる相対リスクの推定精度よりも悪くなる。従って、推定精度が悪くなる分に見合ったコストの減少がなければ、ネステッド・ケース・コントロール研究やケース・コホート研究を実施する意味はないことになる。

表3はリスク比の推定に関して、コホート研究とケース・コホート研究の相対効率を示したものである (Sato (1992b) の結果の一部)。相対効率とは、全コホートの情報を用いて推定する

表3. コホート研究とケース・コホート研究のリスク比の相対効率。

リスク比 疾病発生割合*		1		1.5		3	
		0.05	0.10	0.05	0.10	0.05	0.10
サブコホートの サンプル割合	0.05	0.51 (0.10)*	0.34 (0.15)	0.47 (0.11)	0.30 (0.17)	0.41 (0.15)	0.26 (0.24)
	0.1	0.69 (0.15)	0.53 (0.19)	0.65 (0.16)	0.48 (0.21)	0.60 (0.19)	0.43 (0.28)
	0.2	0.83 (0.24)	0.71 (0.28)	0.81 (0.25)	0.68 (0.30)	0.77 (0.28)	0.63 (0.36)

* 非曝露グループの疾病発生割合。

* カッコ内は全コホートに対するケース・コホート研究の対象者の割合。

表4. 曝露の影響がない場合 ($IRR=1$) の相対効率 (Langholz and Thomas (1990), Table 1 より引用).

ケース・コントロール		ケース・コホート		有効な対象者の割合
マッチング	相対効率	サブコホート	相対効率	
1:1	.50	.050	.51	.10
1:3	.75	.143	.77	.19
1:5	.83	.226	.85	.26

リスク比の分散とケース・コホート研究で推定するリスク比の分散の比を、コホートのサイズが非常に大きい状況のもとで計算したものである。リスク比を (1, 1.5, 3), 非曝露グループの疾病発生割合を (0.05, 0.1), コホートからサブコホートを選ぶ割合を (0.05, 0.1, 0.2) と変化させ、ケースは全ケースを用いることにし、コホート中の曝露グループの割合は 0.5 に固定した。また、カッコ内にはケース・コホート研究で有効な対象者 (全ケース+サブコホート-サブコホートでケースとなった者) の全コホートに対する割合を示した。相対効率は、リスク比の増加、疾病発生割合の増加とともに減少するが、全体としてみると、全コホートの 1/3 から 1/4 の情報で 60% から 80% の相対効率が得られている。

疾病発生率の比 IRR の推定に関しては、ネステッド・ケース・コントロール、ケース・コホートどちらのデザインも用いることができるので、両者の比較がなされている (Prentice (1986a), Langholz and Thomas (1990))。比較の際には、ネステッド・ケース・コントロール研究でマッチさせるコントロール数、ケース・コホート研究でのサブコホートのサンプル割合の決め方に注意が必要であり、有効な対象者数 (ケースとコントロールの両グループにまたがったり、複数回コントロールとして選ばれた対象者は全て 1 と数える) が同じ数になるように設定しなければならない。この点を考慮して $IRR=1$ のときの相対効率を比較した結果を Langholz and Thomas (1990) が与えているので、その一部を表 4 に引用した。表 4 には観察期間中の疾病発生割合を 0.05 とした場合のみを再掲した。例えば表 4 の 2 行目は、ネステッド・ケース・コントロール研究でケース 1 例に対して 3 例のコントロールをマッチさせた場合、全コホートに対する相対効率は 0.75 であり、それに対応するケース・コホート研究はサブコホートのサンプル割合が 0.143, 相対効率が 0.77 であり、どちらも全コホートに対する有効な対象者の割合は 2 割である、ということを示している。どの場合もケース・コホート研究の方が若干効率が高い。しかし差はわずかであり、どちらのデザインでも全コホートの 2 割程度の情報で 70% 以上の相対効率が得られていることが分かる。

6. 考 察

「コホート研究の枠組みの中でのケース・コントロール研究」の代表的なオプションとしてネステッド・ケース・コントロール研究とケース・コホート研究を紹介した。これらのデザインはコホートメンバの曝露情報、主要なリスク要因に関する情報がすでにデータベース等に登録されているような場合、簡単に実施できる。例えば、Brandt and Nielsen (1992) ではデンマークの病院外来登録を用いてケース・コホート研究を行なっているし、職業に関連した有害物質への曝露の影響を調べる研究のように、曝露記録が特別に保存されている場合にもあてはまるだろう。また、血中の微量な成分を測定しなければならないような場合 (特に、測定のコストが高いとき)、もし血液の凍結保存が可能であれば、ケースとコントロールについての血液のみ解

凍して測定を行なえばよいので、有効なデザインとなる (Prentice (1986a), Breslow and Day (1987)).

ネステッド・ケース・コントロールとケース・コホート2つのデザインのうち、どちらがより好ましいデザインであるかは研究の目的により異なる。もし我々の興味がリスク比を調べることにあるのならば、ケース・コホート研究から「まれな疾病の仮定」を必要としないでリスク比を推定することができる。また、実際に対象としている疾病の発生がまれであるならば、2章に示したケース・コントロール研究を行なってもいいだろう (もちろん、「疾病がまれ」な場合は疾病発生率の比もリスク比のよい近似となるのでネステッド・ケース・コントロールやケース発生時点の情報を使ったケース・コホート研究からもリスク比を推定できる)。

興味のある効果の指標が疾病発生率の比である場合は、効率のよさ (5章) だけからみると、ネステッド・ケース・コントロール研究もケース・コホート研究もそれほどの差はない。従って、どちらのデザインを選ぶかは研究目的に依存するが、対象とする疾病が1つである場合は、解析の単純さを考えてネステッド・ケース・コントロール研究が好ましいと筆者は考える。複数の疾病を対象とする場合は、1つのコントロールグループですむケース・コホート研究がコストの点から優れているだろう。また、ケース・コホート研究は、1) サブコホートを用いて疾病発生率や (介入研究の場合には) 副作用をモニターすることができる (Prentice (1986a)), 2) 職業コホートのようにコホート全体が多少とも職業上の曝露を受けている場合にはコホート外の曝露を全く受けていないグループがコントロールとして必要であるが、このときSMRなどの標準化した効果の指標を求めることができる (Wacholder and Boivin (1987)), といった特長も持っているので、こういったことも考慮して研究デザインを選ぶ必要がある (この選択については Wacholder (1991) のまとめが役にたつであろう)。

本論文では、コホート研究の枠組みの中でケース・コントロール研究を実施するという疫学研究のデザインについて、特にコントロールの選択方法と相対リスクの推定についてまとめを行なった。このデザインはコホート研究のコストを減少させ、しかもケース・コントロール研究で問題となるバイアスを避けられるため有用である。実際、16,000人の追跡調査を実施している Atherosclerosis Risk in Communities Study (The ARIC Investigators (1989)) ではコストを考慮して、解析にはネステッド・ケース・コントロールを用いている。また、既に終了しているコホート研究のデータを他の目的に利用する場合や、直接の研究目的ではないが科学的に興味のある補助情報を手にいれたいという場合にもこれらのデザインを適用することが可能 (Davis (1990)) なので、今後ますます利用されていくであろう。

謝 辞

貴重なコメントをいただいた東京大学疫学教室中井里史先生に心から感謝致します。

参 考 文 献

- Brandt, L.P.A. and Nielsen, C.V. (1992). Job stress and adverse outcome of pregnancy: a causal link or recall bias?, *American Journal of Epidemiology*, **135**, 302-311.
- Breslow, N.E. and Day, N.E. (1980). *Statistical Methods in Cancer Research, Vol. I — The Analysis of Case-Control Studies*, Oxford University Press, New York.
- Breslow, N.E. and Day, N.E. (1987). *Statistical Methods in Cancer Research, Vol. II — The Design and Analysis of Cohort Studies*, Oxford University Press, New York.
- Cornfield, J. (1951). A method of estimating comparative rates from clinical data, *Journal of the*

- National Cancer Institute*, **11**, 1269-1275.
- Davis, C.E. (1990). Efficient means of studying ancillary questions in clinical trials, *Statistics in Medicine*, **9**, 97-100.
- Elandt-Johnson, R.C. (1975). Definition of rates: some remarks on their use and misuse, *American Journal of Epidemiology*, **102**, 267-271.
- Flanders, W.D. and Louv, W.C. (1986). The exposure odds ratio in nested case-control studies with competing risks, *American Journal of Epidemiology*, **124**, 684-692.
- Flanders, W.D., DerSimonian, R. and Rhodes, P. (1990). Estimation of risk ratios in case-base studies with competing risks, *Statistics in Medicine*, **9**, 423-435.
- Greenland, S. (1986). Adjustment of risk ratios in case-base studies (hybrid epidemiologic designs), *Statistics in Medicine*, **5**, 579-584.
- Greenland, S. (1987). Interpretation and choice of effect measures in epidemiologic analysis, *American Journal of Epidemiology*, **125**, 761-768.
- Greenland, S. and Thomas, D.C. (1982). On the need for the rare disease assumption in case-control studies, *American Journal of Epidemiology*, **116**, 547-553.
- Greenland, S., Thomas, D.C. and Morgenstern, H. (1986). The rare-disease assumption revisited: a critique of "Estimators of relative risk for case-control studies", *American Journal of Epidemiology*, **124**, 869-876.
- Kleinbaum, D.G., Kupper, L.L. and Morgenstern, H. (1982). *Epidemiologic Research: Principles and Quantitative Methods*, Van Nostrand Reinhold, New York.
- Kupper, L.L., McMichael, A.J. and Spirtas, R. (1975). A hybrid epidemiologic study design useful in estimating relative risk, *J. Amer. Statist. Assoc.*, **70**, 524-528.
- Langholz, B. and Thomas, D.C. (1990). Nested case-control and case-cohort methods of sampling from a cohort: a critical comparison, *American Journal of Epidemiology*, **131**, 169-176.
- Langholz, B. and Thomas, D.C. (1991). Efficiency of cohort sampling designs: some surprising results, *Biometrics*, **47**, 1563-1571.
- Liddell, F.D.K., McDonald, J.C. and Thomas, D.C. (1977). Methods for cohort analysis: appraisal by application to asbestos mining, *J. Roy. Statist. Soc. Ser. A*, **140**, 469-490.
- Lipid Research Clinics Program (1984). The lipid research clinics coronary primary prevention trial results: I. Reduction in incidence of coronary heart disease, *Journal of the American Medical Association*, **251**, 351-364.
- Lubin, J.H. and Gail, M.H. (1984). Biased selection of controls for case-control analysis of cohort studies, *Biometrics*, **40**, 63-75.
- Mantel, N. (1973). Synthetic retrospective studies and related topics, *Biometrics*, **29**, 479-486.
- Miettinen, O.S. (1976). Estimability and estimation in case-referent studies, *American Journal of Epidemiology*, **103**, 226-253.
- Miettinen, O.S. (1977). The author replies (Letter to the Editor), *American Journal of Epidemiology*, **105**, 498-502.
- Miettinen, O.S. (1982). Design options in epidemiologic research: an update, *Scandinavian Journal of Work and Environmental Health*, **8** (Suppl. 1), 7-14.
- Prentice, R.L. (1986a). A case-cohort design for epidemiologic cohort studies and disease prevention trials, *Biometrika*, **73**, 1-11.
- Prentice, R.L. (1986b). On the design of synthetic case-control studies, *Biometrics*, **42**, 301-310.
- Prentice, R.L. and Breslow, N.E. (1978). Retrospective studies and failure time models, *Biometrika*, **65**, 153-158.
- Robins, J.M., Gail, M.H. and Lubin, J.H. (1986). More on "Biased selection of controls for case-control analysis of cohort studies", *Biometrics*, **42**, 293-299.
- Robins, J.M., Prentice, R.L. and Blevins, D. (1989). Designs for synthetic case-control studies in open cohorts, *Biometrics*, **45**, 1103-1116.
- Rothman, K.J. (1986). *Modern Epidemiology*, Little, Brown and Co., Boston.
- Sato, T. (1992a). Estimation of a common risk ratio in stratified case-cohort studies, *Statistics in Medicine*, **12**, 1599-1605.
- Sato, T. (1992b). Maximum likelihood estimation of the risk ratio in case-cohort studies, *Biometrics*, **48**, 1215-1221.

- The ARIC Investigators (1989). The atherosclerosis risk in communities (ARIC) study : design and objectives, *American Journal of Epidemiology*, **129**, 687-702.
- Wacholder, S. (1991). Practical considerations in choosing between the case-cohort and nested case-control designs, *Epidemiology*, **2**, 155-158.
- Wacholder, S. and Boivin, J.F. (1987). External comparisons with the case-cohort design, *American Journal of Epidemiology*, **126**, 1198-1209.
- Wacholder, S., Gail, M.H., Pee, D. and Brookmeyer, R. (1989). Alternative variance and efficiency calculations for the case-cohort design, *Biometrika*, **76**, 117-123.

Epidemiologic Study Designs and Estimation of Relative Risks

Tosiya Sato

(The Institute of Statistical Mathematics)

In cohort studies, subjects are followed from exposure to a putative risk factor to incidence of the disease under study. One can observe the natural history of the disease in these studies, and make a close causal inference when experiments are conducted. If the disease under study is "rare", it is necessary to follow a large-scale cohort for a long time and almost all costs are spent to follow a large number of healthy subjects. While case-control studies are an appropriate design for the rare disease, they are vulnerable to various kinds of bias.

Recently, several case-control-within-cohort designs have been proposed. It is well-known that the nature of cohort and case-control studies are complementary. Combining elements of both designs, the study costs will be reduced and biases vulnerable in case-control studies can be avoided. This paper reviews control selection procedures and required assumptions in the estimation of relative risks in these new epidemiologic designs.