

介入研究に関わる疫学デザイン

研究協力者 津田 敏秀 (岡山大学医学部衛生学教室)

1、はじめに

RothmanとGreenland(1998)は、彼らの疫学の著書の中で、疫学研究の種類に関して次のように述べている。「疫学研究の種類は、科学的実験の概念にそのルーツを有している。疫学的実験が可能なきには、研究要因の比較において、そのデザインは研究要因以外の要因のバラツキを減らすという原理に従う。疫学的実験には、臨床試験(患者を対象者とする)、フィールド試験(個々の地域住民に介入を割り付ける)、そして地域介入試験(地域全体に介入を割り付ける)がある。」とし、コホート研究や症例対照研究などの観察研究(非実験的疫学研究)と区別している。

介入研究には、疫学研究者以外にもわかりやすいなどの様々な利点が存在する一方で、その具体的な実施においては様々な困難が存在し、研究者にとっては必ずしもそれらの克服は容易なものではない。まだ、研究結果の一般化(Generalizability)に関しても十分な検討が要求される。本稿は、介入研究について、主に臨床試験を中心にその特徴と長所・短所を織り交ぜながら解説する。

2、位置づけ

Flecher(1996)らは、「臨床における意志決定に際してすべての論文が等しく重要なわけではない。総説、教育論文、論説といった多くの論文は一般的に真実と信じられていることについて記述しているが、真実を確立することを目的とした原著ではない。これらの論文は要約した情報源としては便利だが、正しい知識を基本とした解釈であって、これ自体には独立した貢献性がない。さらに、ある確立した視点から書いてあるためバイアスを含んでいる可能性がある」と述べ、心筋梗塞における予防的リドカインの投与に関して、1980年までに、12編のRCTが行われ、基本的にはすべての研究において、リドカインによる治療はプラセボに比較して無効か憎悪すると報告しているに関わらず、1980年代に

出版された多くの総説や教科書では、リドカインのルーチンのあるいは選択的な使用を勧めていたことを指摘している。従って、臨床現場では、教科書や総説に頼るよりは、医師もしくは研究者自身が、直接に研究結果を読むことが重要であり、これゆえに、実際に可能であれば介入試験などの研究を行い、研究結果を定量的に明示的に示していくことが重要であると考えられる。

3、フィールドでの実施

介入研究は、患者(臨床試験)、ボランティア(フィールド試験)、もしくは communities 全体(共同体試験)に、ランダムに曝露を割り当てる。なお、臨床試験と共同体試験についてはここでは必要以上に述べない。フィールド試験には、予防であろうとスクリーニングであろうと、コホート研究と同じフィールドでの作業を含むだけでなく、募集やランダム化という特有の作業もある。参加者は、対象者にかなりの要求をするいずれの試験の利点についても納得している必要がある。

もしその試験が特定の人口集団に的を絞っているのなら、この的が対象者を集めることになる可能性のある sources(源泉母集団)を示していることもある。もし名前と電話番号もしくは住所のリストが利用できれば、研究スタッフが個別に前向きの募集方法が行えるので、募集は最も簡単である。例として、食事と類腺腫性ポリープの再発との関係を調べるための試験において、調査者は参加した消化器科医からの紹介や参加した内視鏡サービス機関の医学記録の総覧によって、参加が可能な者を定めた研究がある。

募集の可能な者のリストがないときには、調査者は、公的な掲示、ニュース、広告によって募集しなければならない。重要人物と住民組織もしくは医学組織が関心を高めることによってスポンサーになってくれるたり支持してくれることを、しばしば調査者は要請する。ほとんどの募集は、対象者自身(例えば、無料医学検査)と社会に対する試験の価値を説得することによって成功する。試験への参加にかかる駐車料金もしくは他の少額のコストを払い戻す試験もある。

もし研究による時間的拘束もしくは身体への負担が大きければ(例えば何度も採血をする)、金銭的補償が必要になるかもしれない。

ひとたび参加を募集すれば、対象者は乱数表(紙上もしくはコンピューター上の)を使って試験のどれかの群にランダムに割り当てられる。ランダム化は実験群間の参加者の類似性を保証しないので、調査者は主な臨床結果の予知因子を測定するために研究の起点での質問票を使うべきである。調査者はまた、追跡期間の終わりまで参加者の居所についての情報を集める(例えば、友人の名前と住所)。多くの介入試験では、質問票を郵送し、返事がない人には電話をかけることによって、毎年もしくは他の定期的な追跡を必要とする。もし臨床結果もしくは曝露についての新しいデータが必要ではなくても、住所については現時点のものを把握しておく必要がある(Hartge and Cahill 1998)。

4、ランダム割付について

ランダム割付の目的は、曝露群と非曝露群の間によるバイアスを避けるためである(他の要因が曝露群間で等しくなる)。とりわけ、未知の交絡要因unknown confounderによるバイアスを避ける方略である。しかし、小規模標本のランダム割付は、偶然の変動(チャンス)による交絡が残ることは認識しておかなければならない(Norell 1995)。しかし、このようなことも標本数が増すに連れて起こりにくくなる。ランダム割付は乱数表もしくは電子計算機が作るランダム列を用いる。二群に割り付けるには例えば0から4を介入群、5から9を非介入群として割り付ける。なお、ランダム割付はチャンスにより、介入群と非介入群の数が一致しないこともある。一致させるためにはブロック割付などでランダム化を制限する必要がある(Norell 1995)。チャンスによる交絡要因候補の介入群と非介入群間の不一致は、層別分析により避けられる可能性がある。二つ以上の介入(例えばAとB)を検証する際には、非介入、A介入、B介入、と割り付ける場合と、非介入、A介入、B介入、A+B介入の4つに割り付ける場合がある(ファクトリアルデザイン)。介入の最大の利点は、未知の交絡要因による影響を制御できる点

で、さらにこれに偽薬(プラセボ)もしくは二重盲検を加えることもある。最大の欠点は費用効果の点で、小さな違いを検出するためにはかなりの対象者が必要になる。さらに、倫理的問題や介入の際のコンプライアンスの問題など実務的難点も念頭に置かなければならない。

ランダム割付は通常個人で行われるが時には家族などの集団単位で行われる。これによってコストや便利さの利点がある。また、介入を受けなかった対象者への介入の影響を可能性を減らすことが出来る。一方、割付単位の減少は、ランダム割付の目的をとげさせる可能性を減らすことにもなる(Norell 1995)。

臨床疫学のテキスト(Flecher 1996)においては、臨床研究の妥当性決定に関する基本ガイドラインのランダム化試験に関して次のようにまとめている。

1. コホート研究における基本ガイドラインを満たしているか？

臨床試験はコホート研究である。

2. 患者を治療群と対照群にランダムに割り付けているか？

治療の比較にバイアスを全くかけない唯一効果的方法である。

3. 患者、医療者、研究者たちはどの患者がどちらの治療群に属するか分からないようにマスクされているか？

研究参加者のマスクングはバイアスがかかっていないという確信を強める。

4. その他の介入は両群で同じか？

患者に異なる治療を行うとランダム化による比較の有用性がなくなってしまう。

5. 結果は、治療の割付、または実際に受けた治療に従って記述されているか？

すべての患者が割り付けた治療を受けていない場合、2種類の解析方法がある。これらはおのおの異なる目的と科学的信頼性を持つ。intention to treat分析はランダムに構成された群の分析であり、マネジメントの決定を行う。efficacy分析は実際に治療された患者についての分析であり、介入それ自体の効果を説明

するために行うコホート研究である。

5、手順

研究が「曝露」というものを導入するのなら、介入研究参加者の全ては、最初は非曝露者でなければならない。加えて参加者は、介入や実験研究の手順をこなす能力と意志を持っていると見なされる者であるべきである。従って、事前の評価や、試行期間(run in period)を割付前に行うことで判断されることもある。参加者は交絡要因候補や修飾要因候補に関して揃えられた特定の参加資格基準を満たす者のみで、これらの情報を提供できる能力がある者のみにするべきである。

参加者に対するインフォームドコンセントは、介入研究参加前に行っておくべきである。介入研究の内容、参加への合意などは、書面で行われるのが良いであろう。参加者の主治医からも同意を得るべきであろう。参加資格者は、認識番号をつけられて参加者リストとして載せられ、最初の検診で交絡要因候補の要因や効果の修飾候補の要因を調べる検診を受ける。リストは最初に作られこれに基づいてランダム割付が行われる。対象者の登録やランダム割付は、検査・治療・フォローアップとは独立して直接タッチしない疫学・統計学者などにより行われるのが好ましい。もちろん、検査・治療・フォローアップと直接タッチする人は割り当てについて知らされない方がよい。

6、プラセボとブラインディング

プラセボ「偽薬」の役割は、参加者が研究中に介入群に属しているか非介入群に属しているかを認知させないようにするため、つまり、「プラセボ効果(薬を与えただけで直ったような気なり、薬自体の効果かどうか分からなくなる効果)」を防ぐために、介入群と非介入群において介入群に与えられた「真薬」以外は、同等にするためである。

一重のブラインディングは、プラセボなどを使うことにより、対象者が介入群に属しているか、非介入群に属しているかが分からないようにすることであ

る。二重のブラインディングは、対象者と治療および追跡のスタッフの両方が、その対象者が介入群に属しているか非介入群に属しているかが分からないようにするためである。プラセボとブラインディングは、以下のバイアスを避けるために行われる(Norelli 1995)。

1) 交絡

例えば、高血圧と心疾患において、高血圧治療者が自覚して、食事や喫煙などの他のリスクファクターをなくす努力をしたとしよう。加えて高血圧の治療がこのような他の努力により検査結果に影響を与えるようなことがある。プラセボやブラインディングは、介入群と非介入群をこのような点で同様にしておく。

2) 疾病の誤分類における異なり

例えば、症例の道程に影響を及ぼす可能性のある診療が、高血圧治療に含まれているとしよう。加えて、治療に気づくことが、参加者の受け止め方と症候の報告において、参加者に影響を与える可能性がある。プラセボとブラインディングは、differentialな誤分類を避けるためによって行われる。つまり介入群と非介入群がこれらの点において等しいようにすることによってこのような誤分類を防ぐ。

3) 選択的追跡不明(もしくは選択的情報の欠損)における違い

例えば、疾病の症候は高血圧の治療において、それがあがる対象者とそれがあがらない対象者において異なって追跡不能が生じるような影響が起こる可能性がある。プラセボとブラインディングはdifferentialな選択に使われる。つまり介入群と非介入群がこれらの点において等しくなるようにする。

Double Blindが破れたときには、Double Blindを行うための目的としてきた上記のバイアスが生じていると考えるべきである。交絡については、情報を集めて、層別分析で調整する必要があるし、differentialな誤分類(情報バイアス)や選択的追跡不明については、感度分析等を用いながら考察を十分に行う必要がある。

7、バイアス

ランダム化の目的を達成するため、試行の結果は、介入群は曝露群として、非介入群は非曝露者として割り当てられた、すべての対象者に基づくべきである。従って曝露の誤分類は、実際の介入状態や(そうなった)曝露に基づくのではなく、最初の割り当てに基づくべきである。この時、もし介入群に割り当てられたはずの対象者が介入を受けないのなら(例えばコンプライアンスが保てない時に)、また介入がすべての参加者において曝露状態を変えることに失敗すれば、曝露の誤分類が生じる(Norell 1995)。加えて、同様に非曝露群に割り付けられた対象者に曝露状態が生じた時にも曝露の誤分類が生じる。これらのバイアスは、結果に相対危険度を1に向かわせるバイアスを生じ、いくつかのランダム化試験において生じた「陰性結果」を説明できる可能性がある(Norell 1995)。実際に割り当てた介入を受けた対象者のみに基づいた結果であればこのバイアスは減らせる可能性がある(「on-treatment」分析)。しかし、割付に従ってすべてのランダム化された対象者にもとづいた結果(「intention-to-treat」分析)でなければ、ランダム化の目的は達成されない。

Differential な疾病の誤分類は、プラセボの使用とブラインドによって避けられる可能性がある(ランダム化によってもある程度は避けられる)。しかし、differential な疾病の誤分類は、ブラインドの破綻もしくは副作用における違いにより招来される可能性がある。non-differential な疾病の誤分類は、ランダム化、プラセボもしくはブラインドにより招来されない(Norell 1995)。

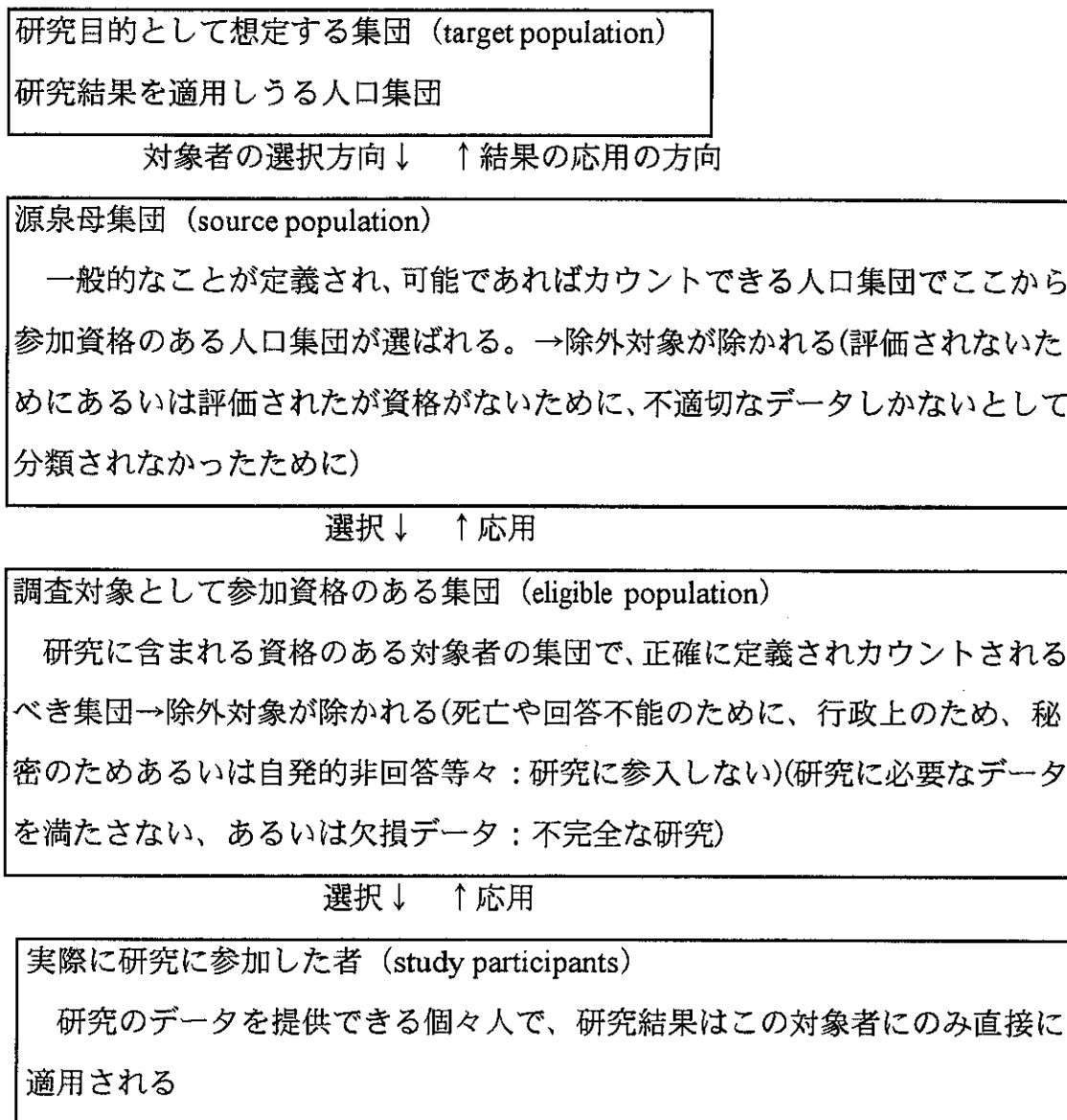
ランダム化は、交絡を防ぐためにも使われる。しかし、交絡は、群間の偶然の変動や不適切なランダム化手順のためにまだ存在している可能性はある。介入により招来された交絡は(もしくは介入に気づくことによりもたらされた交絡は)、ランダム化によっても避けられない。しかし、プラセボの使用やブラインドにより防ぐことができる可能性はある。しかし、ブラインドが破綻したり副作用に違いがあれば、このような交絡も招来しうる。観察研究におけるように、

交絡要因候補についての情報はデータ分析における交絡を制御するのに使われることが可能である(Norell 1995)。ある介入では、曝露の状態を変えたり、他の方略により研究目的とする疾病のリスクに影響をする可能性はある。曝露状態の変化の影響を探索するとき、後者の可能性は交絡を招来するかもしれない。例えば、観察データによれば高度の血清コレステロールは心血管疾患のリスクを増加させることを示唆している。血清コレステロールを下げる薬を使ってランダム化試験における影響を研究する際に、交絡は、研究目的とする薬が心血管疾患のリスクに影響するような他の方略により招かれることになる(Norell 1995)。

選択バイアスは、研究対象を選択する段階で生じてくるバイアスである。疫学的影響の指標に関して影響を及ぼすので、比較群間(コホート研究における曝露群と非曝露群、症例対照研究における症例群と対照群)のバランスが選択過程において崩れることにより生じうる。

実際の研究においては、研究目的として想定する集団target populationを想定し、源泉母集団source populationを定め、そのうち調査対象として参加資格のある集団eligible populationに絞り、実際に研究に参加した者study participantsが分析される(Elwood 1998)。この過程で調査対象から除かれていった者を、論文に記載し、読者に対し、選択バイアスが生じうる可能性に関する情報を与えることが好ましい(図1)。この過程は、逆の方向で実際に調査された者を分析した結果を、参加資格のある集団、源泉母集団、研究目的として想定する集団に、一般化(generalization)ができるかどうかの問題にも関係する。

図1 研究目的として想定する集団, 源泉母集団, 参加資格のある集団, 実際に研究に参加した者 (Elwood 1998から改変し翻訳した)



実際の選択バイアスの例としては、コホート研究において、調査対象の比較群を違う population から選択するために生じる健康労働者効果(healthy workers effect)や、同じくコホート研究において、比較群間で疾病の発生と関係し追跡割合が異なって生じる選択的追跡漏れ(selective loss to follow up)、また症例対照研究において、対照群が研究対照としている曝露と関係することにより生じる対

照群選択バイアス(reference selection bias)などが挙げられる。選択バイアスは、影響の指標を過大評価する方向にも過小評価する方向にも働き得るが、上記の3つの具体的な選択バイアスの例においては、影響の程度の指標を過小評価する方向に働く場合が多い。

選択バイアスの影響

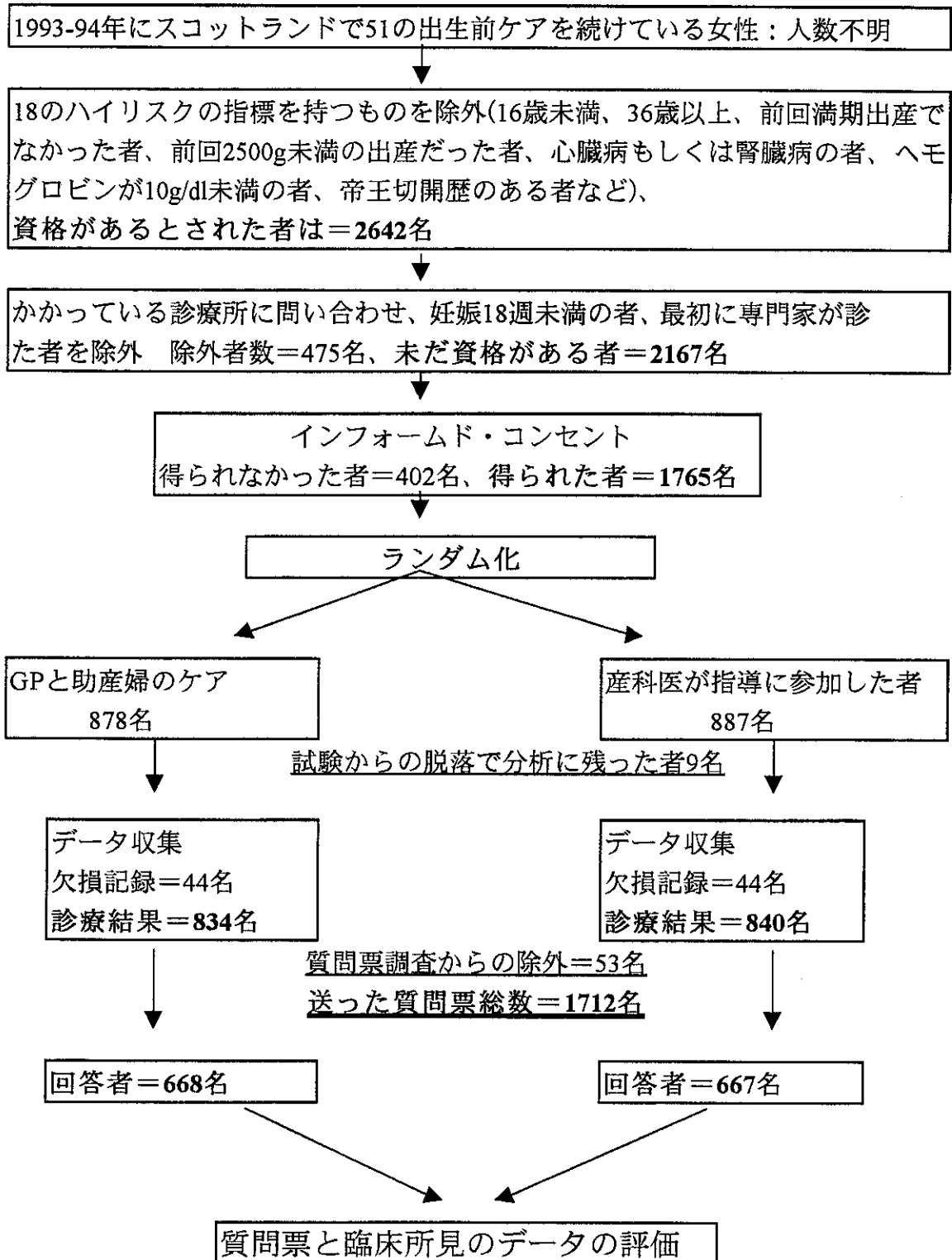
選択バイアスは研究について3つの影響を持ちうる。

- 1、内的妥当性に影響を与える。
 - 2、外的妥当性(一般化)に影響を与える。
 - 3、検証しようとする仮説を修飾する。
-

Elwood(1998)より改変

もし追跡不能(もしくは情報の喪失)が研究中の疾患により影響されるなとき、選択バイアスが招来される。differentialな選択はプラセボとブラインドにより(そしてある程度ランダム化により)避けることが可能である。しかし、ブラインドの破綻もしくは副作用における違いにより招来される可能性がある。ブラインドの破綻は、ブラインドにより防ぐことを目的としていたバイアスを招くことにつながる。このようなバイアスについては、上記にて説明した。Non-differentialな選択(絶対影響の過大推定もしくは過小推定になる)は、ランダム化、プラセボ、ブラインドにより影響を受けない(Norell 1995)。図2に介入試験の具体的なデザインについて記した。

図2 介入試験の具体的デザイン(Tuckerらのデータ：1996、Elwoodより改変)



8、介入研究のプランニングの時に考慮すべき諸点(Lee 1992)

1) プランニングの時期

数ヶ月もしくは1年は、研究プロトコール、もしくは介入試験の目的を特定し達成する計画案を作成する時間として与えるべきである。プロトコールを準備するガイドラインは後に詳述する。もし介入試験が2つ以上の期間を巻き込むのであれば、さらに計画のための時間(1年以上)を与えるべきである。

2) 介入群の数

介入試験に含まれる群の数は、1年に研究に入ると予想されるのが可能な患者の数に密接に関連する。それぞれの介入群における十分な数の患者を保証するために、介入試験は比較のための介入群の数を少なくするべきである。

3) 試験期間

試験にかかると見積もられる期間は、患者の参入のための期間と反応(事象もしくは損失)と生存の観察のための追跡期間が含まれる。GeorgeとDesu(1974)は、患者の参入がポアソン分布に従い、生存時間(もしくは損失時間)が指数分布に従うと仮定して介入試験に必要な期間を得ようとした。一般的に利用できるルールはないが、もし介入試験を長い期間にわたって拡大するときは、別の治療法が比較の候補として現れる可能性がある。實際上、5年以上に拡大される介入試験は、十分に根拠づけられねばならない。

4) 患者の比較可能性

比較介入試験は、治療群間の相違の唯一の説明が、治療の結果であるというふうに計画されるべきである。患者は予後因子などの点で比較可能でなければならず、そうでないと、その結果は非常に読み違えられる可能性があるだろう。介入参入の時期において患者の比較可能性を達成する技術は、正式なランダム化である。しかし、ランダム化は、すべての予後因子の点において、治療群間における患者が比較可能であることを保証するわけではない。しばしば調整を分析段階で行うことが可能ではあるものの、比較可能性は、分析がなされる前に、チェックされて確かめられねばならない。

5) 介入割付の比

二つの群に患者を等しく割り付けることが通常行われる。しかし、新しい治療法と古い治療法を比較したり、一つの治療が非常に困難だとか高価なときは6対4とか67対33(2対1)とかの割合で患者をランダムに割り付けることがアドバイスされる可能性がある。二つの治療群間における真の違いを検出する可能性は、2対1より比が大きくなならない限りは減らない(Peto 1976)。

6) 対照群の設定

ほとんどの臨床試験が基本的に要求されるのは、「対照」であり、これは介入群の特定の特徴を共有しながらも介入されなかった患者の一群のことである。問題は同時的に得られる対照を用いるべきかどうか、もしくは過去から得られる対照群で充分かである。倫理的問題は、有効であろう治療を与える患者にそれを見合わせる事が適切かどうかである。もちろん、治療の有効性が判明していないことである；もし有効性が判明しているのなら、試験の必要性などないであろうが。一方、Phase IやII(第1段階や第2段階)において、その治療が介入試験を見送られたのなら、試験を正当化する何らかの基本がなければならない。問題の厳しさは、何がかかっているか、に依存する。例えば試験における治療が通常頭痛を早くなおしたい目的ならば、厳格な介入試験の倫理性はそんなに厳しく問われることはないであろう。しかし、命がかかっているとか死や予後影響が厳しい病気(癌など)の危機に瀕しているための何らかの治療を差し引かえることは、たとえ一時的にでも、大変重要なことであろう。一方、新しい治療が確かに、しかしいつもベストではなく、決して危険からいつも解放されているわけではないことはきちんと認識しておくべきである。例えば、ある抗生物質やホルモンがいつも無害というわけではない。従って、介入試験ごとに倫理的なバランスで、それ自身の環境とそれ自身の問題に従ってそれぞれで、徹底的に重みづけしなければならない。

過去の対照を利用することは、討論の対象となり続けてきた。特に癌の介入試験において議論されてきた。過去の対照を利用することは、全ての現在の患

者が試験において新しい治療を受けることができる。そして結果は、以前に標準的な治療を受けた患者の結果と比較される。過去の対照を用いることに伴う主な問題は、患者の群間で比較可能性と評価の方法を確かめることができないことである。

7) 介入の運営

同じ様式でそれぞれの介入処方において患者を管理することは重要である。反応と毒性の定義は、それぞれの介入群における患者に関して、精確に同じであるべきである。試験から患者を除くときの決定も、それぞれの介入群において等しく適用されるべきである。ある調査者が介入の一方を好んで、患者がそれについてより長く介入され続け、非常に厳しいときのみ毒性があると分類されるのなら、結果はバイアスされる可能性があるだろう。一方、より毒性があるとされる介入を割り当てられた患者は、悪い影響を報告しやすい可能性がある。医師のもしくは患者のバイアスの可能性や治療管理の相違を避ける一つの方法は、二重盲検法を行うことである。二重盲検法において、治療は同じ形式で準備されるので、患者も試験を行う医師も、特定の介入のどちらがなされたかには気づかない。しかし、二重盲検法試験は、組織することが難しく、バイアスを減らすことにおいては期待するほど有効ではない。例えば、白血病患者に関して、化学療法と非治療を比較したいとき、薬が錠剤であるとき、それは患者の両群に錠剤を与える必要がある(対照群にダミーの錠剤)。そして、白血球数を測定するために、それぞれの群に血液サンプルを取ることが必要である。これは煩雑な管理である可能性がある。また、医師は予想される毒性の型に気づき、観察される毒性により、どちらの患者がそれぞれの治療を受けているか、見つけだすことができる。加えて、患者とその家族に二重盲検法試験を説明することは難しい。従って、二重盲検による研究は、注意深い配慮が与えられるべきである外科手術法もしくは放射線治療のような治療を用いる症例においては、それらは適用することはできない。

8) 倫理的な配慮

もし医師が、特定の患者に関して、ある治療が、もう1つの治療より良いと感じるなら、彼もしくは彼女は治療をランダム化して割り付けることができない。医師がベストと信じている手順で患者を治療することは、非倫理的ではない。従って、医師は患者に対して治療の一つがより良いものであると考えるのなら、そして研究中の何らかの時に1つの治療がベターであるという明瞭な徴候があるのなら、患者のランダム化はストップされるべきであろう。

何らかの統計的有意差が得られる前に医師のバイアスのために介入試験が終わるのを防ぐために、試験の結果は、諮問委員会により試験が中止されるかどうかの決定が下るまで参加した医師には知らせるべきではない。この諮問委員会に諮ると言うだけでは完全には倫理的問題を解決したとは言えない。しかし、この方法は、「参加した医師だけにのみ、なされるより、さらに客観的な判断が与えられる。

倫理的問題は試験の開始において生じる可能性があるし、試験の進行の間においても生じる可能性がある。例えば、患者の状態が悪化したと考えられるとき、研究から患者を引き上げさせるかどうかの倫理的問題が存在する。患者を研究から引き上げさせる可能性があることを考慮すべき厳しい状態を考慮に入れることは重要であり、またそれらを試験が開始される以前に研究プロトコールに詳しく記述していることは重要である。

ほとんどの研究所は、全ての人間を対象として含む計画に関して研究所の公式の審査委員会(IRB)による了承を要求している。通常、同意形式に沿った完全な提案が、研究費が得られる時にはIRBに提出されなければならない。臨床試験の開始の前に、プロトコールにおいてなされる何らかの試行も、さらに承認のためにIRBに提出されなければならない。IRBの責任は、介入試験が行う価値があることを確かめることである。便益がリスクに勝っており、可能性のある対象には十分に情報が与えられていることが確かめられていることである。一般的に、次に挙げる情報が合意形式において含まれなければならない。

1. 臨床試験の目的
2. 調査する薬、装置、もしくは手順
3. 使う手順の全てを含む臨床試験の記載
4. 参加者への可能性のある便益
5. 参加者へのありうるリスク
6. 障害もしくは悪影響の際の医学的治療と補償に関する情報
7. 秘密保持と秘密を守るために使われる方法

加えて、対象の参加が自発的であることが確かめられねばならない。そして合意が取り消され、試験からの離脱が何の罰もなく、利益の損失もなくいつでも許されることが可能である。特に重要な記載は、合意の形式が、医学専門用語を用いず素人の用語で書かれていなければならない。このことはしばしば調査者によって無視される。

9) その他

良い研究プロトコルを準備することは非常に重要である。特に試験が協力を要するときには。すべての参加調査者もしくは研究所は、詳細な研究プロトコルを持たなければならない。介入試験がうまくはこび続けるための第一の責任者が誰かというのも重要である。しばしば、これは臨床医よりも統計学者になる。その主な理由はこの実効ある責任を遂行するためには時間が要求されるからである。

9、介入研究のデザインの種類

比較臨床試験の実験デザインには二つの型がある：固定標本試験(fixed-sample trials)と連続試験(sequential trial)である。固定標本試験においては、二つ(もしくはそれ以上)の治療群に割り付けられた患者の数が、研究が始まる前に固定され

ている。連続試験においては、新しい患者を採用し続けられるかどうかの決定は、その時点で集められた結果によって決められる。以下のように、まずいくつかの固定標本試験について短く述べ、それから連続試験について述べる。

1) 単純なランダム化デザイン (Simple Randomized Design)

この最も単純な例では、患者はランダムに二つ(もしくはそれ以上)に、患者の特徴を考慮することなしに割り当てられる。通常、ランダム化は固定した数の患者に制限され、それぞれの治療群は等しいサイズになるであろう。このデザインの主な利点は、その単純さと重要な予後因子が未知で、対象となる患者の特徴に関して均一であるという点である。これの欠点は、治療群間で比較可能性がなく、バイアスされた分析である。

このデザインの一つの例としては、躁鬱病の患者における維持療法に関する血清リチウムレベルの標準的なレベルと低いレベルを比較する二重盲検研究がある(Galenberg et al. 1989)。94名の患者が維持療法に関してリチウムの二つの異なった投与量をランダムに割り当てられた：標準用量すなわち0.8-1.0mmol/lの血清リチウム濃度を算出するように調整された群、低用量すなわち0.4-0.6mmol/lの血清濃度を示すような用量の群である。著者らは、低容量群においては標準用量群に比較して2.6倍(95%信頼区間、1.3-5.2)再燃のリスクが高かったことを報告した。結論は、0.8-1.0mmol/lの血清リチウム濃度を算出するような用量がそれより低い血清リチウム濃度になる用量より効果があったというものであった。しかし、悪い副作用が、より高い用量の群では、より高い頻度であった。

2) 層化ランダム化デザイン (Stratified Randomized Design)

重要な予後因子が分かっている、患者群をさらに予後カテゴリーもしくは「層(strata)」サブグループ化することができる時、治療群間の比較可能性が達成し得るだろう。最も単純な例は、ペア比較デザインであり、これは層が患者のペアの時の話であり、ランダムに選んだ一人には治療Aが与えられ、もう一人には治療Bが与えられる(二つの治療の比較の場合に)。このペアは一つもしくはそれ以上の予後因子によりマッチされ、治療A、B以外はできるだけ似たように治療さ

れ、同時的に観察されねばならない。この方法の主な困難性は、同時に二人の患者がいることがそんなにあることである。結果としてこのデザインは、患者のプールがかなり利用可能で特定の時期にそろえることができる慢性病にしばしば限られる。マッチされた予後因子は、よく考え抜かれ、事前に定義されねばならない。例えば、50歳から60歳の間のいかなる年齢も同じ層に分類されるだろう。

多くの臨床試験はこのデザインを使う。例えば、大動脈の狭窄の外科的修復の後の、術後高血圧に対するプロプラノロールの影響に関する研究が層化ランダム化デザインを使って行われた(Leenen 1987)。狭窄摘除術を受けた4歳から16歳の23人の患者(7人の少女と16人の少年)を対象に研究した。対象者は、年齢4歳から10歳、10.1歳から16歳で層別化された。それぞれの年齢層内で、患者はランダムに割り付けられ、プロプラノロールとプラセボを投与された。

3) クロスオーバーデザイン (Crossover Design)

クロスオーバーデザイン(Jones 1989)は、単純なランダム化デザインとそれぞれの患者がそれ自身の対照を持っているペア化されたデザインの組み合わせである。このデザインを使用する通常の方法は、一連の治療Aの後に続いて治療Bを半分の患者に、治療Bの後に治療Aを残りの患者に割り当てることである。一人の患者はランダム割付によって二つの治療の一つを割り当てられ、それぞれの治療は患者の病気の状態が比較可能な時に割り当てられる。二つの治療の三つの比較が可能である：異なった患者間で最初とその次の段階で比較し、そして同じ患者内で比較する。このデザインは患者が再燃か軽快状態にある時に使用される。だんだん悪くなる病気や急速に重篤状態に進行する病気に関しては患者内での比較は利点がない。例えばもし目的が緩和状態の長さを比較することなら、患者は二つの連続した緩和状態を持つことが不可能か二つの緩和状態が異なった段階である可能性がある。クロスオーバーデザインには他にも実際上の困難を伴っている。もし目的が治療AとBを同じ患者内で比較するなら、患者の中には第二の治療を受けるほど十分に生き残っていないものもいる可能性

がある。もう一つには、患者が第一の治療に鋭い反応を示したので、第二の治療の前に長い時間がかかる可能性がある。同じ患者において治療AとBを比較することが、異なった患者において治療AとBを比較する結果を固めるが、このデザインでは、生存についての治療の影響を知ることはできない。

クロスオーバーデザインは症候の短期間の回復を測定するのに良い。それは治療に関するさらに精確な比較が可能になる。時に、それぞれの患者が両方の治療を受ける機会を持っているので、より倫理的であると見なされる。調査者は、患者内比較の可能性はバラツキがより小さいのでより小さいサンプルサイズを使うことができる。しかし、議論もある。最大の問題は、治療の影響が持ち越されないことを仮定していることである。この仮定が侵された時、このデザインは治療影響を検出するには敏感ではない。クロスオーバーデザインを考慮するにおいて、調査者はいくつかの問いを発しなければならない。まず、病気は充分安定しているか？第二にそれぞれの段階において治療の影響が見られるのに十分な時間はあるか？第三に、第一の治療と第二の治療の間に持ち越し影響が消えるのに十分な時間があるか(Matthew)？第四に、患者の協力とコンプライアンスが二つの段階を通して十分に得られるかどうか？

患者の状態が安定していることを確かめるために、「試走(run-in)」期間が、試験の開始時において、ランダム化が行われる前になされることが役に立つかもしれない。同様に、「洗い出し(wash-out)」期間が、第二の治療が始まる前に一般的に要求される。

クロスオーバーデザインは調査者によって広く使われている。例えば、無作為割付で二重盲検の研究が、酸素100%の部屋の空気でプロのサッカー選手がどう振る舞うかを比較するために行われた(Winter 1989)。

他の例として、三つの治療期間を持つ多クロスオーバーデザインがある。それぞれは、薬効のある薬による治療(tetrahydroaminoacridine (THA)) とプラセボ治療からなり、ランダムに二重盲検の手順で割り付けられたアルツハイマー病の患者であった(Molloy 1991)。

4) ファクトリアルデザイン **Factorial Design**

$r \times s$ のファクトリアルデザインにおいては、二つの治療の内、一つが r レベルで投与されるもう一つは s レベルで投与される。目的は、様々な治療影響と、共同作用の可能性を検証することである。最も単純なファクトリアルデザインは、 2×2 表であり、ここでは二つの治療がその反応の関係に関して研究され、それぞれに二つのレベルが与えられる。つまり高濃度と低濃度、もしくは本当の薬とプラセボである、治療の数とレベルの数が大きい時、多くの患者は必要とされ、結果は解釈しがたいものになる可能性がある。

医師の健康研究(Stampfer 1985, Steeing Committee of Physicians' Health Study Research Group 1989)はランダム化、二重盲検、プラセボ対照 2×2 表ファクトリアルデザインを、低容量のアスピリンが心血管疾患による死亡率を減少させるかどうか、そしてベータカロチンが癌の罹患率を減少させるかどうかを決定させるためにデザインされた。総勢22,071人の医師が最初にランダムに、実効性のあるアスピリンとプラセボに割り当てられた、そしてそれぞれのグループ内で対象者はさらに実効性のあるベータカロチンとプラセボに割り当てられた。従って、四つの等しい対象者数のサイズのグループが形成された。それぞれの群は異なる薬の組み合わせを受けた：(1)アスピリンとカロチン、(2)アスピリンとカロチンプラセボ、(3)アスピリンプラセボとカロチン、そして(4)アスピリンプラセボとカロチン、の4つの組み合わせであった。Stamperらは、このデザインを用いて次のことを指摘した。それぞれの対象者は関心のある疑問に対して貢献してくれる結果となり、カロチン仮説がアスピリン研究に対し物質的に影響していることなく検証することができる。加えて、対象者全体のわずか4分の1しかプラセボ投与しか受けないことは、この研究は参加者に参加しやすい研究となっている可能性がある。

ファクトリアルデザインを用いた他の研究は、Hungら(1988)による 3×3 の例、Manninenら(1988)による $2 \times 2 \times 3$ のデザイン、Elaad adn Ben-Shakherら(1989)による 2×4 のデザイン、Davisら(1989)による 3×3 のデザイン等がある。