人間栄養学講座

連載 ランダム化臨床試験をする前に
【第2回】試験計画書の作成

佐藤 俊 哉

京都大学大学院医学研究科社会健康医学系専攻医療統計学

Human Nutrition Lecture

Series : Introduction to Randomized Clinical Trials

Part 2 : Making of the Trial Protocol

Tosiya Sato

Department of Biostatistics, Kyoto University School of Public Health

試験計画書ってなに？

科学的な実験で重視されることは、結果の再現性である。仮に結果をうまく捏造できて雑誌の査読者にも気づかれずに審査を通り、論文が掲載されたとしても、その論文を読んだ研究者が論文とまったく同じ実験方法で実際に実験したにも関わらず、同じ結果が再現できなければ、結果を疑い、論文を疑い、やがては捏造であることが判明してしまう。

臨床試験も人に対する科学的な実験であるため、同じ手順で試験を実施すれば、統計的な誤差の範囲内で同じ結果が再現できなければならない。同じ手順で試験が実施できるためには、その手順を文書として残す必要がある。これが試験計画書（Protocol）である。医薬品開発では、薬事法の一部である「医薬品の臨床試験の実施に関する省令（Good Clinical Practice, GCP）」¹⁾に試験計画書の作成と、それに沿った試験の実施が規定されているし、私たち研究者が実施する一般の臨床試験でも、「疫学研究に関する倫理指針」²⁾、「臨床研究に関する倫理指針」³⁾に試験計画書の作成が定められている。

本稿では、なぜ試験計画書の作成が必要なのか、試験計画書がないとどのような不都合が起こるのか、また、試験計画書に書かなければいけないことは何なのか、について解説しよう。

なぜ試験計画書が必要なのか？

私のところには統計解析についての相談がよくあり、「簡単なことから」、「 t 検定の使い方だけですから」とご本人は言われるのだが、まず簡単に終わったためしはない。医療統計の専門家ということで、解析方法についてのアドバイスを一言二言求めてもらえるのだろうが、専門家である以上、データの氏素性を知らずして解析することはできないので、データを集めた背景と目的、実際にどうやってデータを集めたのかを事細かに聞くことになる。

そうするとほとんどの場合、研究計画書を作っていない、ということがわかり、また多くがデータの集め方がもともとの目的と合っていない、したがって解析しても望んだような結果は出ない、ということがわかる。なぜこんなことが起こるのであろうか。

実際に研究を始めてみると、最初に予想した以上にいろいろな問題や不都合が起きるものである。問題が1つ起きて、それではこの問題はこう対処しよう、また1つ不都合が起きて、じゃあここはこう妥協しよう、とその場は解決できても、次にはその場その場で対処した結果、妥協した結果のせいでまた別な問題が起きて…、と方針変更が積み重ねられて当初の目的とはだんだんずれていってしまう。もし試験計画書があれば、問題や不都合に対処しなければならない場合でも、はじめに何をしようとしていたのかを常に参照することができるので、当初の目的から大きく外れることはま

（連絡先：佐藤俊哉 〒606-8501 京都市左京区吉田近衛町 京都大学大学院医学研究科社会健康医学系専攻医療統計学
電話 075-753-4495 FAX 075-753-4487 E-mail shun@pbh.med.kyoto-u.ac.jp）

ずないはずであったのに。

このように、目的、手順を試験計画書として文書に残すことで、何をしなければならないかを常に明確にすることができるし、また変更が必要な場合でも、計画への影響を最低限に留めるような変更とすることができる。

ランダムでない臨床試験では、対象者数もそれほど多くないことから、少数の施設で実施できる。前回紹介したラクトフェリンの用量-反応試験は2施設のみで行った。しかし、ランダム化臨床試験のように、試験参加者数の多い試験となると多数の施設が共同して試験を実施することになる。このとき、全参加施設で「ラクトフェリンを12週間摂取することでC型肝炎ウイルスが減少するかどうかを調べよう」と合意して試験を開始したものの、ある施設では「うちではラクトフェリンを12週間ではなく16週間食べてもらおう」とか、また別の施設では「C型肝炎ウイルスの検査を12週後にと言われたが、うちの日常診療と合わないので16週後に測定しよう」とばらばらなことをしてしまっただけは、試験の再現性もなければ、試験の質も保証できなくなってしまう。

こんなとき実施手順を明確に定めた試験計画書があれば、すべての施設で同じ手順で試験を実施できるし、手順どおりに実施できない施設がある場合、試験計画に問題があるのであれば、計画を実施可能なように修正する、あるいは施設側の問題であるなら、その施設に改善を要求する、場合によっては参加を見合わせてもらう、といった対処が可能となる。

さらに臨床試験を実施する際には、人に対する実験であることから、計画が科学的、倫理的に妥当なものであるかどうか倫理審査委員会の審査を受ける必要がある^{2,3)}。審査の際には試験計画書は必須の文書となっている。では、試験計画書には具体的にどんな内容を盛り込む必要があるのだろうか。

試験計画書の実例

試験計画書には試験を再現できる内容をすべて書かなければならないわけではなく、実際には細かい手順は試験実施手順書、標準業務手順書(Standard Operating Procedure, SOP)などに規定し、さらに調査票、試験参加者への説明文書、同意文書などを作成しなければならない。試験計画書には試験の根幹に関わる内容(表1)を記載する必要がある。すべてを取り上げることとはできないが、以下では、ラクトフェリンのランダム化臨床試験⁴⁾(「ラクトフェリン試験」と略す)を例に、いくつかの項目について解説しよう。

表1 試験計画書の内容

・試験の要約	・試験参加者への倫理的配慮
・試験の背景	・目的に対応したエンドポイント
・試験の目的	・試験実施手順
・試験デザイン	・統計解析の方法
・試験実施期間、実施場所	・試験に必要な参加者数
・対象集団の設定	・組織、責任者、連絡先

1. 試験の要約、背景、目的

計画書本文はかなりの分量になるため、最初に試験全体のアウトラインを2〜3ページにまとめておくと全体像が理解しやすくなる。

本文最初には、これまでの研究の流れの中で、何がどこまでわかっていて、ここから先がわかっていないので、この試験ではこの部分を明らかにする、という研究の大きな流れを、文献や関連した先行研究を引用しながら、この試験の位置付けが明らかになるように背景を説明するとわかりやすい。ラクトフェリン試験の計画書では、前回まとめた日本でのC型肝炎の現状、ラクトフェリンについて、これまでに実施された人でのラクトフェリン摂取の安全性、用量-反応試験の結果、それを受けて有効性を検証するためにプラセボを対照とした臨床試験を実施するという内容になっている。

こういった臨床試験を実施する場合、調べたい項目はそれこそ山ほどあり、どうしてもあれもこれも、と計画に盛り込みがちであるが、有効性に関する試験の目的はしぼり込んで、その試験で必ず結論を出したい主要な目的をできるだけ1つにしたほうがよい。そして主要な目的をサポートするものや、次の研究の手がかりとなりそうなものを副次的な目的、以下、ともかく探索的に調べてみたいもの、と目的に順序を付け整理しておく、いろいろ調べてはみたものの結局なにもわからなかったという事態を避けることができる。また、安全性については、有効性とは違った側面の評価であるので、有効性とは別に常に調べる必要がある。

特に主要な目的は試験に必要な参加者数を計算する根拠となるものである、試験に参加していただく方たちの厚意をむだにしないためにも、もっとも重要なものを指定し、十分な計画の下で正しいか正しくないかはっきりと結論を出せるようにしなければならない。ラクトフェリン試験では、もっとも興味のある「ラクトフェリンを摂取することで、C型慢性肝炎患者のC型肝炎ウイルス量(Hepatitis C Virus RNA量, HCV RNA量)を減らすことができるか」どうかを主要な目的とし、副次的な目的として摂取前と比べHCV RNA量がどのくらい変化するか、ALT(alanine aminotransferase)値がどのくらい変化するか、ALT値の改善の程度を設定した。

2. 試験デザイン

ランダム化臨床試験の中でも、単純な2グループ比較から、同一対象者に時期を変えて2つの介入を行うクロスオーバー試験、同一対象者に2つの介入を同時に行う要因実験などさまざまな試験デザインがある。試験デザインは、主要な目的を達成するためにもっとも適したデザインとする必要がある。また、介入方法をランダムに割り付ける方法についても、対象者1名ずつランダムに介入方法を決める単純ランダム化、対象者を何名かブロックにして介入方法を割り付けるブロックランダム化、男女別や施設別に割り付けを行う層別割り付けなどいくつかの選択肢がある。

選んだ試験デザイン、ランダム化の方法についても計画書に記載する必要がある。ラクトフェリン試験では、試験デザインは「ランダム化二重マスク化プラセボコントロール試験」、ランダム化の方法は「施設を層別因子とするパーミューテーションブロック法」を用いることにした。これらの試験デザイン、ランダム化の方法に関する詳細は次回以降解説する。

3. 対象集団の設定

試験の対象となる集団を明確にすることは、試験結果をどこまで一般化できるかを考察する際に重要である。そのためには対象者の選択基準、除外基準（両者合わせて適格基準という）を明らかにし、基準を設定した根拠を述べる必要がある。

ラクトフェリン試験では、日本人のC型肝炎患者の治療を目標としているが、実際に対象者を選ぶ集団は、参加者の登録期間である2001年から2004年にラクトフェリン試験に参加した7施設を訪れたC型肝炎患者であり、そこからさらに試験で定めた適格基準に合致し、試験内容を十分理解して自由意思で同意した（インフォームド・コンセント）試験参加者が実際の対象集団となる。

ターゲットとしている集団でラクトフェリンの効果が確かめられるかどうかを、内部妥当性（Internal Validity）というが、ランダム化臨床試験では、全試験参加者をランダムな要素にもとづいて2つのグループに分けることで、全試験参加者が全員ラクトフェリンを摂取した場合と、全員がラクトフェリンを摂取しなかった場合の理想的な状況に近い比較ができるため、全試験参加者をターゲット集団と考えると内部妥当性を満たしている。しかし、その結果をもっと広い集団に一般化できるかどうか（一般化可能性 Generalizability あるいは外部妥当性 External Validity）については判断が難しい。

一般化できるかどうかは、今回の試験の適格基準やラクトフェリンの作用機序から考えることになるだろうが、通常は1つの試験の結果だけで一般化できるこ

とはまずありえない。様々な対象、様々な地域、様々な状況で実施された複数の試験結果から、どの程度の一般化が可能かを判断することになる。そのためには、個々の試験でどのような選択基準、除外基準を設定したかを明確にする必要がある。

4. 倫理的配慮

臨床試験は人に対する実験であるので、科学性よりも倫理性が優先されなければならない。そのためにはヘルシンキ宣言、疫学研究に関する倫理指針、臨床研究に関する倫理指針などの関連するガイドラインを遵守して試験を実施しなければならない。倫理的配慮の項には、これらの倫理ガイドラインに沿って試験を計画し実施すること、また最終版には、倫理審査委員会の審査を経ていることを明示すべきである。

疫学研究に関する倫理指針、臨床研究に関する倫理指針をよく読むと、健康食品の効果を調べる試験は、ラクトフェリン試験のように患者を対象として臨床現場で実施するものであっても疫学研究の倫理指針の対象となっていたり、どちらのガイドラインに沿って実施したらよいのかが不明確であったり、さらにはどちらの指針にもあてはまらない、といったケースがありうる（現在、疫学研究に関する倫理指針は見直しが行われている途中である。<http://www.mhlw.go.jp/shingi/2007/03/s0313-6.html>）。

しかし、ガイドラインの抜け道を探して臨床試験を実施することが倫理的に許されるはずはないし、そもそもこれらのガイドラインは達成しがたい理想を掲げたようなものではなく、現実に関心を実施するのであれば、最低限こういった内容は守って実施してほしい、というものである。ラクトフェリン試験であれば疫学研究に関する倫理指針、臨床研究に関する倫理指針、どちらも十分満たすように計画しておくほうがよい。このラクトフェリン試験を開始したのはどちらのガイドラインも公布される以前であったが、GCPを満たすように計画したので、幸いなことにどちらのガイドラインも満たして実施することができた。

試験参加者に対する倫理的配慮で重要なことは、試験の目的、意義、内容を対象者に理解できるように説明し、他の選択肢、試験に参加する利益、不利益、負担などを含めて対象者が十分に納得した上で、自由意思で試験に参加してもらうインフォームド・コンセントと、試験参加者のプライバシー、個人情報情報を十分に保護する対策の2点である。

対象者は医学や臨床試験の専門家ではないので、口頭で一度説明を聞いただけで試験の内容が理解できるはずがない。いくら口頭でわかりやすく試験内容を説明してもそれだけで理解するのは無理なので、専門家

ではない人が後から読んで試験の内容が確認できるような説明文書を作成して対象者に渡し、じっくり考えてから試験に参加するかどうかを判断できるような配慮が必要となる⁵⁾。ラクトフェリン試験では、ランダム化の説明などにわかりやすいイラストを付け、文字の大きさや字体も考えて A4 で 9 ページの説明文書を用意し、説明を受けた次の来院時に参加するかどうかを決めてもらった。

プライバシー、個人情報の保護についても、「あなたの個人情報は守られます」といった、保護の方法も理解できないし、誰が守るのがわからない書き方ではなく、具体的に述べる必要がある。ラクトフェリン試験の説明文書では、「あなたのカルテや病院記録など、プライバシーの保護には十分に配慮いたします。この臨床試験を通じて得られたあなたに関する記録は、当院の他に臨床試験研究班のデータセンターに保管されますが、あなたのお名前はわからないようになっておりますし、担当医師と試験の管理者、専任のデータ管理者以外の目に触れることはありません。また、あなたのお名前や個人が特定できるような情報が、試験結果の報告や発表に使用されることはありません」と記した。

5. エンドポイント、統計解析、必要参加者数

適切なエンドポイント（試験の目的となる臨床結果）を選ぶのはなかなか難しい。食事療法が血圧に与える影響をみたい場合、試験終了時の血圧値がエンドポイントになることもありうるし、試験開始時から終了時までの血圧の変化（終了時と開始時の血圧の差や比）がエンドポイントとなることもあるだろう。また、運動療法と心筋梗塞の予防を調べる試験であれば、心筋梗塞の発生割合や発生率（割合と率の違いは文献 6 を参照）、心筋梗塞発生までの時間がエンドポイントとなるかもしれない。また、ADL (Activities of Daily Living)、健康関連 QOL (Quality of Life)、痛みなどのスケールがエンドポイントとして適切な場合もあるだろう。

エンドポイントとしては、測定が可能で、測定精度も高く、試験目的に合ったエンドポイントを選ぶことが重要である。ラクトフェリン試験では、各施設で安定して HCV RNA 量が測定できるようになったことから、ラクトフェリン摂取開始前から摂取 12 週後に HCV RNA 量が半減以下となった参加者の割合を「ウイルス学的効果」と定義し、ラクトフェリン群、プラセボ群で比較することとした。

エンドポイントが連続的な値か、健康イベントの発生割合か発生率か、あるいはイベント発生までの時間が決まると、それに応じて統計解析の方法もほぼ決まるが、試験計画書では「*t* 検定を行う」とか「カイニ乗検定を行う」といった解析の枝葉末節ではなく、な

にを比較するのか、比較の際に考慮すべき要因は何か、といった主要な方針を記載するだけで十分である。

ラクトフェリン試験では、ラクトフェリン群、プラセボ群別に参加者の性別、年齢、試験開始時の HCV RNA 量といった背景要因を比較し、ウイルス学的効果に関する割合の差の検定と推定、HCV RNA 量と ALT 値の変化、そして 12 週間のラクトフェリンまたはプラセボ摂取中と 4 週間の観察期間の HCV RNA 量の推移をグラフにすることを定めた。これらの解析の詳細な方法については、医療統計専門家が「解析計画書」を作成し、実際の解析に当たることになる。またラクトフェリン試験では、必要な試験参加者数の半数の結果が得られたところで正式な中間解析を実施することにしたので、中間解析を実施することについても試験計画書に記載した。

主要な目的に対応する主要エンドポイントとその解析方法が決まると、研究に必要な参加者数を統計的に計算することができる。とは言ってもいくつかの値を見積らないといけないが、割合の差の検定であれば、表 2 に示す、コントロール介入群および試験介入群のイベント発生割合、アルファレベル、検出力を設定する必要がある。

コントロール介入群については文献や過去の試験結果などから情報を得ることができる。試験介入群のイベント発生割合の見積りは、本来ならコントロール介入群のイベント発生割合をこの程度上回ったら効果があると考えられる最小の値とすべきであるが、小さい差ほど確実に検出しようとする、とんでもなくたくさんの試験参加者数が必要となる。このため試験介入群の割合も過去の試験結果などから見積ることになるが、コントロール介入と違い、試験介入の効果は対象者数があまり多くない試験から得られているため、その推定精度も考慮して少し割り引いて見積るほうが安全であろう。

表 2 試験参加者数の計算に必要な情報

コントロール介入群のイベント発生割合	文献や過去の試験結果から
試験介入群のイベント発生割合	過去の試験結果から (本来は、効果があると考えられる最小の値とすべき)
アルファレベル (有意水準)	本当は試験介入の効果がコントロール介入と同じ場合に、間違って「差がある」と判断する可能性
検出力	本当は試験介入の効果がコントロール介入を上回る場合に、正しく「差がある」と判断する可能性

資料) 文献 7

アルファレベルはできるだけ小さくしたいが、これも小さくすればするほど、たくさんの試験参加者数が必要となる。検証的なランダム化臨床試験では、両側5%水準が長年標準となっていて、この慣例を変えるのは難しいのが実情である。検出力も差がある場合に確実に検出しようと高く設定すると参加者数が増えるが、検証的なランダム化臨床試験では80%以上に設定するのが普通であり、もう二度とその試験の実施ができないような場合は90%以上とすることもある。

ラクトフェリン試験では、用量-反応試験の結果から、ラクトフェリン群でのウイルス学的効果を15%、通常とは逆にプラセボ群の情報がなかったが、なにもしなくても自然変動で5%はウイルス学的効果がみられると仮定し、アルファレベルは慣例どおり両側5%とした。通常ならここで検出力を80%などと設定して必要な試験参加者数を計算するのであるが、事前に7施設に試験への登録が可能な患者数を尋ねたところ、1年で90名がぎりぎりであった。このため試験参加者の登録期間を3年間、250名の登録を目標とすることにあらかじめ決め、ここから逆に検出力を計算したところ、75%は確保できることから、これで試験を実施することにした（参加者数、検出力の計算は教科書⁷⁾を参照してほしい）。

6. 組織

ラクトフェリン試験は参加者数250名と、それほど規模の大きい試験ではないが、それでも7施設の共同研究であることから、しっかりした組織がないと試験の実施、試験の質の管理が滞ってしまう。共同研究は、単なる各施設の研究者の集まりで実施できるわけではなく、試験を運営する上でのそれぞれの責任の所在を明らかにし、試験を円滑に進めるために、試験の代表者、試験の運営に責任をもつ運営委員会、データセンター、さらに中間解析実施のためには、運営組織とは独立した第三者組織である独立データモニタリング委員会の設置が必須である。

ラクトフェリン試験の試験組織としては、代表者（国立がんセンター中央病院肝胆膵内科 岡田周市）、各施設の責任医師および実務担当者、医療統計専門家、医療倫理の専門家などを役割に応じて、また試験の運営をつかさどる運営委員会、調査票の回収とデータベース作成を行うデータセンター（臨床試験支援会社に業務を委託）、そして中間解析を実施するための独立データモニタリング委員会のメンバーを記載した。

*

試験計画書は専門家だけが読むものではない。たとえば倫理審査委員会では、その分野の専門家以外の研究者や、対象者の権利を代表する委員が計画書を読み、審査するのであるから、専門家が読めば分かるという書き方をしてはならない。介入方法の説明などでは専門用語を避けるわけにはいかないが、必ず説明をつけ、少しでもわかりやすくする努力をするべきである。

文 献

1) 厚生大臣：医薬品の臨床試験の実施の基準に関する省令。平成9年3月27日、厚生省令第28号（最終改正平成18年3月31日、厚生労働省令第72号）
<http://law.e-gov.go.jp/htmldata/H09/H09F03601000028.html>

2) 文部科学省、厚生労働省：疫学研究に関する倫理指針。平成14年6月17日、14文科振第123号、科発第0617001号（平成16年12月28日全部改訂、14文科振第930号、科発第1228002号）
<http://www.niph.go.jp/wadai/ekigakurinri/ekigakuririh170401/shishin-all.pdf>

3) 厚生労働省：臨床研究に関する倫理指針の施行等について。平成15年7月30日、医政発第0730009号（平成16年12月28日全部改訂、医政発1228001号）
<http://www.mhlw.go.jp/general/seido/kousei/ikenkyu/rinri/0504sisin.html>

4) Ueno, H., Sato, T., Yamamoto, S., Tanaka, K., Ohkawa, K., Takagi, H., Yokosuka, O., Furuse, J., Saito, H., Sawaki, A., Kasugai, H., Osaki, Y., Fujiyama, S., Sato, K., Wakabayashi, K. and Okusaka, T.: Randomized, double-blind, placebo-controlled trial of bovine lactoferrin in patients with chronic hepatitis C, *Cancer Science*, **97**, 1105-1110 (2006)

5) 佐藤恵子：人体実験から臨床試験へ、これからの臨床試験/椿 広計、藤田利治、佐藤俊哉編, pp.35-48 (1999) 朝倉書店、東京

6) 佐藤俊哉：宇宙怪人しまりす医療統計を学ぶ, pp. 5-16 (2005) 岩波科学ライブラリー 114, 東京 <http://www.iwanami.co.jp/moreinfo/0074540/top.html> で立ちよみができる

7) Friedman, L. M., Furburg, C. D. and DeMets, D. L.: *Fundamentals of Clinical Trials*, 3rd ed., pp.94-129 (1996) Mosby, Missouri