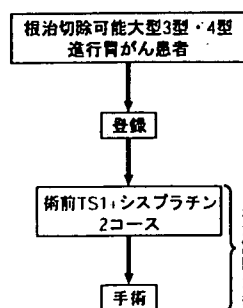


JCOG0210 (胃がん外科)



JCOG0501 (胃がん外科)

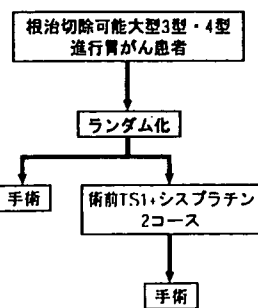
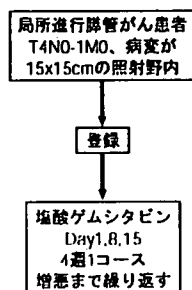


図3: JCOG 試験 JCOG0210 と JCOG0501

JCOG0506 (消化器がん内科)



第Ⅲ相試験

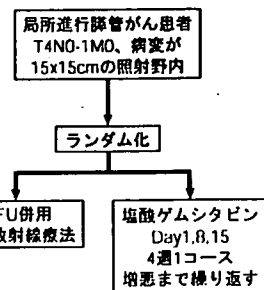


図4: JCOG 試験 JCOG0506 と計画中の第Ⅲ相試験

とすることができる。

そこで我々は、「手術単独」「術前化学療法＋手術」のどちらでも同じ規準で評価できる「根治切除割合」をJCOG0210のプライマリーエンドポイントとした。その結果、「TS-1とシスプラチンの2剤併用化学療法を術前に2コース施行してから手術を行う治療」のほうが、過去の「手術単独」の根治切除割合（予めプロトコルに判断規準として明記）よりも高かった。また、術前化学療法による手術合併症の増加等の不都合も生じなかったため、JCOG0501を計画して開始した。

ポイントは、第Ⅱ相試験のデザインにおいて、次の第Ⅲ相試験で比較する標準治療を明確化したこと、そして標準治療と試験治療の比較対象（全体か部分か）を考え、評価できるエンドポイントを選択したことである。

(3)「治療」によって「患者（選択規準）」がかわる

変則的な例ではあるが、治療によって患者集団（患者選択規準）が変わることがある。具体例で説明しよう（図4）。JCOG0506は、消化器がん内科グループによる、局所進行膵管がん患者に対する

ゲムシタビン単剤の化学療法を評価する第Ⅱ相試験である。膵がんでは、腫瘍の縮小や増大を画像検査で評価することが困難であるため、プライマリーエンドポイントを奏効割合ではなく、1年生存割合とした。また、化学療法のための試験にもかかわらず、適格規準に「原発巣とリンパ節転移をすべて15cm×15cmの照射野に含めることが可能」という条件が含まれている。これは図4の“第Ⅲ相試験”で示したように、第Ⅲ相試験で比較する標準治療が5-FUを併用した化学放射線療法であり、JCOG0506の1年生存割合を標準治療である化学放射線療法の1年生存割合と比較するには、患者を化学放射線療法を受け得る集団にする必要があった。やはりこの例でも「何と比べるのか」が重要であることが理解されよう。

ちなみにこの試験では合併症がない分、ゲムシタビン単剤療法のほうが放射線治療より患者負担が少ないことが期待される。したがって、ゲムシタビン単剤療法は生存期間で化学放射線療法を上回る必要はなく、劣らなければよりよい治療と言えることから「劣らない」ことを検証する「非劣性デザイン」の第Ⅲ相試験が計画されている。

第Ⅱ相試験デザインの進め方

筆者は第Ⅱ相試験のデザインの進め方およびその“コツ”について、以下のように考えている。

- ①常に「標準治療は何か」を念頭に置く
- ②次に行うべき第Ⅲ相試験のデザインを意識する
- ③対象集団を決める（新治療のメリットが期待できる集団の条件を具体化）
- ④試験治療レジメンを決める〔ベネフィットの最大化（maximize benefit）、リスクの最小化（minimize risk）を考慮する〕
- ⑤評価の方法を決める（エンドポイント、登録のタイミングと解析対象集団の定義、解析方法と判断規準を決める）

基本的には①→⑤の順に進めるが、実際には②～⑤は相互調整しながら、具体化することになる。

このステップについては、切除不能の肝転移を有する大腸がん患者に対する5-FU肝動脈注入療法（肝動注）の第Ⅱ相試験を仮想例として、具体的に考えてみよう。治療レジメンは、イリノテカンの静脈投与を併用する治療とした。

- （1）常に「標準治療は何か」を念頭に置く

標準治療について、最近では5-FU＋ロイコボリンにイリノテカンを加えるIFL療法やFOLFIRI療法、オキサリプラチンを加えるFOLFOX療法がある。

ここでは5-FU＋ロイコボリン＋イリノテカンの静脈投与のIFLを標準治療レジメンとする。

- （2）次に行うべき第Ⅲ相試験のデザインを意識する

標準治療はIFLなので、次の第Ⅲ相試験ではIFL療法と5-FU肝動注＋イリノテカン療法をラン

ダム化比較するデザインとなる。

- （3）対象集団を決める

試験を行うにあたっては、参加患者のベネフィットの最大化、およびリスクの最小化が重要である。まず、ベネフィット最大化の観点から、新治療のメリットが期待できる患者集団を考える。

仮想試験で考えると、肝動注療法は肝転移巣に高濃度の5-FU曝露が起きることから、特に肝転移巣に対して強い抗腫瘍効果が期待される。逆に肝以外の転移巣に対しては、静脈内投与されるイリノテカンによる肝経由で全身に回る5-FUによる抗腫瘍効果のみであり、標準治療のIFL療法に優る効果が期待できない。むしろ5-FUが肝経由で全身に回ることから、肝以外の転移巣に対する効果はIFL療法よりも劣る可能性がある。したがって、この試験の対象は、肝以外に有意な（生命予後を規定するような）病巣がない、すなわち原発巣は既に切除されていて、肺転移や脳転移を有さない患者が適切であろう。

次に、リスク最小化の観点から、新治療によって逆にリスクが高くなると思われる患者を除外する。仮想試験では、肝動注が危険となり得る門脈閉塞、肝動脈奇形、肝硬変を有する患者を除外するといった選択規準が必要となる。

- （4）試験治療レジメンを決める

まず薬剤の用量だが、第Ⅱ相試験の段階ですでに決定されているはずであるため、本稿では割愛する。第Ⅱ相試験で考慮することは、第Ⅲ相試験で比較すべき標準治療のレジメンと同じにしたほうがよいかどうかの検討である。薬物の代謝や体内動態、併用薬剤との相互作用等から、効果を最大化する用法・用量が決まっている場合はそ

れを優先させるべきだが、そうでない場合は、標準治療のレジメンに出来るだけ合わせておくことで、第Ⅲ相試験の計画を行うことが容易になる。仮想試験では、5-FUの静注、肝動注の相違以外の部分、つまり併用するイリノテカンの用法・用量、1コースの長さ、共通する支持療法等については、標準治療であるIFLと同様にしておく。それによって、第Ⅲ相試験を実施する際、施設の臨床医の混乱を防ぎ、プロトコルの逸脱を減らすことができる。その結果、本来比べたい治療法を適正に比較することにつながる。

(5) 評価の方法を決める

(3)でも述べた考察により、肝転移しかない患者を対象とした場合は当てはまらないが、もし、肝以外にも転移巣を有する患者を対象とした場合を想定して、奏効割合について考察してみよう。

仮想試験における試験的部分は、5-FUを「肝動注」することである。研究者は「肝転移巣にどれだけ効くか」に強い興味を抱くかもしれないが、「肝転移巣の奏効割合」をプライマリーエンドポイントにしてよいだろうか。標準治療であるIFLは全身化学療法のため、肝転移巣の奏効割合について、信頼できるデータがあるかどうかは疑問である。もっと重要なことは、肝転移巣の奏効割合がIFL療法を大きく上回っていたとしても、肝以外の病変に対する奏効が劣っていれば、肝動注療法はIFLよりも有望な治療であるとは言えないであろう。

したがって、比較する相手は全身化学療法としてのIFLにおける全病変に対する通常の奏効割合であり、肝動注療法の奏効割合も同じように求めたものでなくては比較する意味がない。奏効割合をプライマリーエンドポイントにするなら、全病

変の奏効割合を用いることになる。

集学的治療や特殊技術を要する治療の第Ⅱ相試験で、意外と重要なのが患者登録のタイミングとそれによって定まる解析対象である。仮想試験で考えると、5-FUの肝動脈内注入に先立ち、肝動脈へのカテーテル留置とリザーバー(動注ポート)の埋込みがなされるが、必ずしも100%成功するとは言えない。そのため、肝動注療法の有効性に関心がある研究者は、患者登録・解析対象の設定を肝動脈へのカテーテル留置とリザーバーの留置に成功(カテ・ポート留置成功)した時点で行うかもしれない。

しかし、ここで「次に行うべき第Ⅲ相試験のデザインを意識する」技術が活きてくる。第Ⅲ相試験(図5)を行う場合、(a)、(b)のどちらになるかという、当然、(b)になる。全身化学療法を行う患者に、カテ・ポート留置は不要である。全身化学療法、動注化学療法の選択のタイミングは、カテ・ポート留置前であり、比較可能な対象は全身化学療法を選択した患者集団(実際に行ったかどうかは問わない)とカテ・ポート留置を試みた患者集団(留置の成功・失敗は問わない)である。

このように、治療方針の選択という医学的意思決定を行うタイミング(ランダム化試験の場合のランダム化時点)で患者登録を行うべきであり、タイミングを明確化する上でも第Ⅲ相試験のデザインを意識することは第Ⅱ相試験のデザインにおいて有用である。

以上を踏まえると、この仮想試験のデザインは自ずと決まる(図6)。奏効割合の分母に、カテ・ポート留置を試みたがうまくいかなかった患者を含めることで、初めて全身化学療法での奏効割合と比較できる。

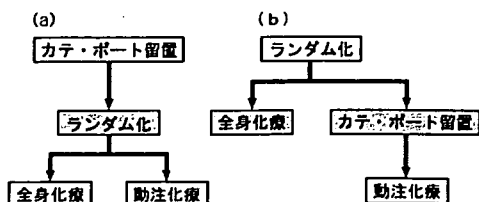


図 5：第Ⅲ相試験における登録のタイミング

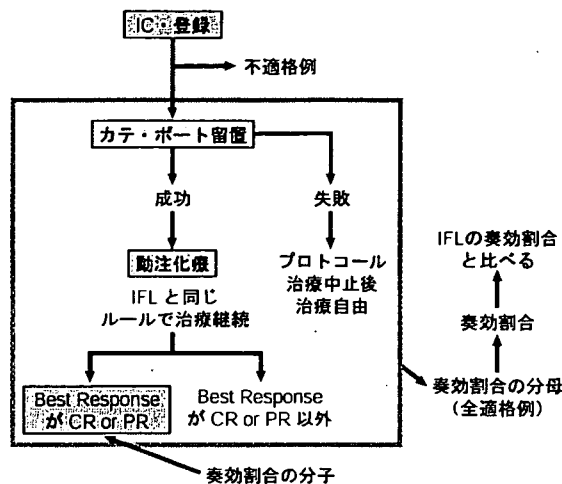


図 6：仮想第Ⅱ試験の試験デザイン

おわりに

本文中で敢えて明記しなかったが、本稿のキーワードは「基本は比較」である。

統計家やデータセンターにとって、「何を何とどう比べるのか」を明確にして提案された試験計画は、もう解かれたのも同然の試験問題であり、サンプルサイズの計算は単にプログラムを起動するだけの事務的作業である。逆に「何を何とどうやって比べるのか」が明確でない試験計画は、デザインの完成まで延々と議論を繰り返さなければならぬし、場合によっては残念ながら中止せざるを得ないこともある。

難しいと言われる臨床試験デザインのポイントは、実は「比べる」という日常ありふれた思考プロセスにあることを強調しておきたい。

連載「がん臨床試験の実践～JCOGを例に～」第2回

コンセプトの作成と審査 ――第三者に試験の意義を理解してもらう

国立がんセンター がん対策情報センター 臨床試験・診療支援部
福田 治彦

前回は臨床試験（以下、試験）の骨子を形作る過程での、デザインに関する考え方を概説したが、今回は、試験概要を文書として第三者に示すための「プロトコルコンセプト」について解説する。

プロトコルコンセプトの目的

●プロトコル作成は多段階方式が一般的

研究実施計画書（プロトコル）について、JCOGや欧米の共同研究グループ（Cooperative Group）では、最初からいきなり試験実施に必要な事項を網羅した数十ページにも及ぶfull protocol（フルプロトコル）を作成しない。その前段階として試験概要のみで構成される数ページのプロトコルコンセプト（protocol concept: コンセプト）を作成する2段階方式を採用している。米国のThe Southwest Oncology Group（SWOG）では3段階方式を採用しており、第1段階として原稿1枚で作成される「capsule summary」をSWOGトップ数名で検討し、コンセプト作成を認めるかどうかを決定している。JCOGでもグループによっては、この方式を採用することもある。

2段階（多段階）方式を採用している理由としては、フルプロトコルの作成段階で試験に大きな問題が見つかった場合、修正に多大な時間を要することが多いし、また試験の意義そのものがグ

ループ内外の審査過程で否定されて試験が中止になった場合、フルプロトコル作成にかかった労力と時間が無駄になるといったロス避けるためである。そのため、「プロトコルコンセプト」の目的は以下の2点に集約される。

- ① 試験を実施するかどうか（採択/却下）を決定する
- ② 試験の骨子の段階で試験デザインを適正化する

●試験には“正しい”結論が必要

試験は第Ⅰ相や第Ⅱ相で数十名から数十名、第Ⅲ相となると数百名から数千名の患者さんに参加いただき、年余の時間と多くの人の労力と資金を投入して実施される。したがって、試験から得られる結果・結論はそれに見合う医学的・社会的価値・倫理要件としてはscientific/social valueと呼ぶが必要になる。

そして試験は“正しい”方法（scientific validity）によって、“正しい”結論を導かなくてはならない。そのため、米国ではコンセプトをCooperative Group自身による試験実施/非実施

※ protocol outline と呼ぶグループもあるが本稿では「コンセプト」で統一する

の判断材料にすることに加えて、スポンサーである National Cancer Institute (NCI) の臨床試験管理 / 支援部門である Cancer Therapy Evaluation Program (CTEP) でコンセプトが審査され、NCI スポンサーの研究としての採否が決められる。つ

まり、CTEP にコンセプトが承認されて初めて、Cooperative Group は第Ⅲ相試験を実施できるのである。表 1 に、CTEP のコンセプト書式を筆者が和訳したものを示す。ちなみに、第Ⅰ相試験、第Ⅱ相試験、第Ⅰ/Ⅱ相試験では、もう少し簡略化さ

表 1: NCI-CTEP の第Ⅲ相試験コンセプト書式項目 (筆者による和訳、一部省略)

I. 管理情報 (Administrative)

- ・ 研究名、研究番号、研究代表者 (氏名、住所、電話、FAX、e-mail)、共同研究者

II. 試験概要 (Science section)

1. 研究仮説

2. 目的

2.1 主目的

2.2 副次目的 (複数可)

3. 背景情報

3.1 試験方法の妥当性とデザインの妥当性

3.2 試験の重要性

3.3 すべての関連データ

4. 適格条件 (選択・除外の根拠を含む)

5. 群 / レジメン (シェーマを含む)

5.1 シェーマ

5.2 群 / レジメン

6. 統計的デザインの詳細

6.1 エンドポイント

6.1.1 プライマリ・エンドポイント

6.1.2 セカンダリ・エンドポイント

6.2 ランダム化における層別因子・割付調整因子

6.3 サンプルサイズと検出力

6.4 解析計画 (中間解析を含む)

6.5 登録見込み (予定登録患者数 / 月、総予定登録患者数 (最小～最大))

7. 実施可能性 (適格患者数見込み、同意取得期待割合、類似の試験における経験を含める)

7.1 競合する同グループ内の第Ⅲ相試験

7.2 競合する他グループの第Ⅲ相試験

7.3 競合する企業の試験

III. 薬剤情報 (Pharmaceutical section)

- ・ 表形式で、薬剤名、CTEP による薬剤配布依頼の有無、治験薬かどうか、治験届者、プラセボ対照かどうかを列記する

IV. 附随研究 (Embedded correlative study section)

1. 附随研究名

1.1 研究デザイン (検体取得方法、記録用紙、画像検査など)

1.2 研究仮説 (背景となる研究の情報を含む)

1.3 統計的デザイン (エンドポイント、サンプルサイズ、登録見込み / 月)

れた Letter of Intent (LOI) という書式が用いられている。いずれも CTEP の Web サイトで閲覧できる (<http://ctep.cancer.gov/guidelines/index.html>)。

CTEP のコンセプト 審査は、各 Cooperative Group 担当の臨床医 (各がん種の専門家)、使用する薬剤を専門にする臨床医 (臨床薬理の専門家)、生物統計家、監査担当、薬剤師、薬事担当など異なる職種の専門家集団で行われるため、試験デザインの適正化の観点からも有用と言える。特に、統計的側面については「Simon の 2-stage design」で有名な統計家の Richard Simonをはじめ、コンセプトやフルプロトコールを毎日のように審査し、経験と理論から新たな統計的デザインを開発している統計家が担当しているため、CTEP の審査を受けることで経験豊富なアドバイスが受けられる。また臨床医による審査では、直近の米国臨床腫瘍学会 (American Society of Clinical Oncology: ASCO) での発表など、最新のエビデンスも踏まえた上で科学的妥当性が検討されているため、既知の事実による無駄な試験の防止、より有用な臨床的仮説や考察の追加が行われている。

以上より、コンセプトを審査に提出する場合、表 1 と同程度の詳細さを有した情報が必要であることを理解されるであろう。臨床研究を計画する際には、外部審査を受けるかどうかは別にして、こうした書式を用いてコンセプトを作成し、研究者同士で検討することで、試験デザインが適切かどうかを確認することは有用と思われる。是非、利用していただきたいと思う。

JCOG における「コンセプト」審査

● JCOG におけるコンセプト 審査の変遷

JCOG におけるコンセプトの審査過程の変遷を簡単に紹介する。

1995 年まで、JCOG にはコンセプトのシステムはなく、筆者が JCOG データセンター (当時は JCOG 統計センター) に加わった 1996 年に、SWOG の「capsule summary」を参考にした A4 サイズ 1 枚のコンセプトを各グループから提出していただき、JCOG 運営委員会会議で試験を検討し、実施の採否を決定する仕組みを導入した。しかし運用してみると、A4 サイズ 1 枚のコンセプトに含まれる情報では、試験デザインの適正化を判断するという目的が果たせず、結局フルプロトコールを作成する段階になって大きな問題が見つかり大幅なデザインの変更を要した、プロトコールが完成しなかった等の事態が生じた。そこで 1998 年より、SWOG やヨーロッパの European Organization for Research and Treatment of Cancer (EORTC) のコンセプト書式を参考にプロトタイプを作成し、1999 年より現行に近いコンセプト書式 (表 2) を用いるようになった。

CTEP の書式では「記載量に制限はない」と記されているが、JCOG では 8 ページ以内を原則とし、10 ページ程度までを許容範囲としている。

● JCOG におけるコンセプト 作成と 審査

現在 JCOG におけるコンセプトの作成と審査は、以下の手順で進められている。

- 1) グループの研究者によるコンセプトの作成
 - 2) データセンターとの事前相談 (オプション)
 - 3) プロトコール審査委員会でのコンセプト 予備審査
 - 4) プロトコール審査委員会会議での検討
 - 5) JCOG 運営委員会での本審査と採決
- 以下、これらのプロセスを簡単に紹介する。

- 1) グループの研究者によるコンセプトの作成
- 前回紹介した試験デザインの骨子にあたる「対象 (患者)」「治療」「評価」が概ね固まり、各グルー

表 2: JCOG のコンセプト 書式(一部省略)

0. シェーマ

- ランダム化比較試験、複数のレジメンの組み合わせ、複数のモダリティの組み合わせの場合、シェーマを作成する

1. 目的

- 2～3行を目安に簡潔に試験目的を記述。対象集団(stage)と評価する治療法を明確に表現する
- primary endpoint(原則1つ)、secondary endpoint(s)を記述

2. 背景

以下の内容について、他分野の研究者が理解できる平易な表現にて明確に記述する。極力略称は用いない

- 対象疾患、疫学的事項(疾患の頻度)、疾患の特徴、対象 stage の特定
- 対象集団選択/設定の根拠: なぜこの対象としたか。主な適格規準の設定根拠を記述
- 対象患者における標準治療(state-of-the-art treatment)と予後。標準治療が確立されていない場合はその旨明記する(広く用いられている治療は何かを記述)
- 使用する薬剤の説明(単剤における主な毒性、有効性データを含む)
- 治療レジメン選択の根拠、過去の同一薬剤・類似レジメンの臨床試験データ
- 試験デザイン(phase II か phase III か。等)および endpoint の設定根拠
- phase III 以外の場合、次のステップの試験の概略(phase II の場合は、その試験で有望な結果が得られた場合に行うであろう phase III の概要)

3. 患者選択規準

以下の項目について、他分野の研究者が理解できる平易かつ明確な表現を用いて記述する。

- 疾患、組織型: 組織学的・細胞学的確診を要求する場合、生検、過去の手術材料の組織診、細胞診等、許容される方法・検体等を明確に定義する。要求しない場合、臨床診断として許容される範囲を明確にする
- stage: 用いる staging criteria とともに明記する(criteriaの名称のみでよい)
- 病変の評価可能性: 測定可能病変のみ/測定可能病変または評価可能病変
- 前治療の規定: 特定の前治療からの休止期間を設ける場合、月でなく「日」もしくは「週」で規定する
- PS: ECOG performance status score を用いる
- 併用薬・併用療法に関する制限事項(ある場合)
- 臓器機能(臨床検査値): 主なもののみ

4. 治療計画

特に複数のレジメンや複数のモダリティによる治療レジメンの場合、「プロトコール治療」の定義を明確に行う。後治療との区別を明確に定義する。

ランダム化試験では群別、集学的治療の試験ではモダリティ別に、小見出しをつけてそれぞれ記述する。

4.x. プロトコール治療中止・終了規準

- プロトコール治療完了とみなすコース数、原病の増悪、治療中止とすべき毒性(有害事象)、コース開始延期の許容範囲もしくはプロトコール治療期間全体の延長許容範囲等の判断規準のうち、主なものを明確に記述する
- 減量規準等、詳細な治療変更規準の記述は不要

5. 効果判定の方法と判定規準

- 効果判定の時期(○コース毎、○コース終了後、プロトコール治療終了後○日…等)
- 評価のモダリティ(頭部CT、胸部CT、胸部X線等、必須のものと許容範囲が明確になるように記述)と判定規準
- 効果判定を行わない試験では「本試験では効果判定は行わない」と記述

6. エンドポイントと統計学的考察

- primary endpoint と secondary endpoint(s)を明記
- ランダム割付を行う場合の割付調整因子
- 第Ⅲ相試験: 標準治療群の生存時間分布の推定値(5年生存率や生存期間中央値)等とその根拠。臨床的に意味がある群間の差。 α 、 β
- 第Ⅱ相試験: 期待奏効率や閾値奏効率、 α 、 β 等
- 予定症例数と予定登録期間の概算
- 中間解析の概略について記述

7. 集積見込み

- ・年間登録数の見込みおよび予定登録期間内の登録見込み
- ・原則として同一疾患における過去のJCOG試験での登録実績に基づく。該当するデータがない場合、参加施設の年間適格患者数や手術例数等から推定した結果を記述する

8. その他特記事項

該当する場合、簡潔に記述する。

- ・病理中央診断
- ・腫瘍縮小効果の中央判定や施設外校閲
- ・凍結検体等を用いた付随的研究
- ・適応外医薬品を用いる場合その取扱について

9. 研究班

主たる研究班の名称と主任研究者。本試験の研究代表者と研究班の関係を明記。

17 指 -5「多施設共同研究の質の向上のための研究体制確立に関する研究」班以外で共同研究体制とする研究班があれば何を役割分担するかも含めて明記。

プの会議で研究者間の合意が得られれば、試験の実務責任者となる臨床医である研究事務局が表2の書式に従ってコンセプトを作成する。

また、各グループの代表者は比較的若手である臨床医を2名、医学審査員 (medical reviewer) として推薦している。医学審査員は自グループのコンセプト作成を支援する義務を負うと同時に、プロトコール審査委員会 (Protocol Review Committee: PRC) にも参加する。医学審査員は、日頃からPRCメーリングリスト (後述) やPRC会議におけるコンセプト検討に参加しているため、JCOGにおけるコンセプト検討の方法に詳しい臨床医と言える。

コンセプトを提出する際には、内容について医学審査員が承知している証左として、「コンセプト事前チェックリスト」を一緒に提出するようになっている (表3)。また、グループ代表者がコンセプトを承認している証左として、カバーレターにはグループ代表者の署名を要求している。官僚主義的と思われるかも知れないが、過去にグ

ループ代表者を初めとする当該グループの研究者が知らないまま、コンセプトが提出され混乱を来したことがあったため、やむを得ず以上の方で行っている。

2) データセンターとの事前相談

最近開始したシステムとして、PRC提出前にグループがコンセプト内容をJCOGデータセンターに相談することができる。義務ではないためすべてではないが、ほとんどのコンセプトで事前相談が行われている。

相談内容は試験デザインの骨子である (対象患者)、「治療」「評価」の中でも、特に「評価」に関してエンドポイントや統計的デザインをデータセンターがアドバイスすることによってPRCに提出するコンセプトの質を高めており、PRCでの議論の効率化に貢献していると考えられる。相談では研究事務局以外の若手臨床医の参加も許可しているため、勉強の場として活用されており、概ね好評を得ている。

表 3: JCOG におけるコンセプト 事前チェックリスト (一部省略)

体制・体裁

1. グループ内での検討

- 1) グループ班会議などで十分な (最低 1 回以上) コンセプト の検討が行われているか。
- 2) グループの医学審査員 / 委員のレビューを受けているか。
- 3) 対象集団、標準治療などについてグループ内のコンセンサスが得られているか。
- 4) プロト コール検討会への出席予定者の調整は済んでいるか。(研究代表者、事務局いずれかは必ず出席しなければならない)

2. コンセプト の書式

- 1) 章構成は書式に従っているか。
- 2) 総ページ数は多くなり すぎていないか。(原則 8 ページ程度)
- 3) 治療変更規準を細かく 書きすぎていないか。(減量規準などの詳細は不要)
- 4) 研究班の項は抜けていないか。

内容

1. 対象集団

- 1) 対象疾患の特徴・ 予後などについて十分に記載されているか。
- 2) 対象となるサブグループ (Stage 毎など) の特徴・ 予後などが十分に記載されているか。
- 3) 上記の記載と 患者選択規準の整合性が取れているか。

2. 対象に対する 標準治療

- 1) 標準治療のエビデンスは十分記載されているか。
- 2) 海外と日本での標準治療 (日常診療) の違いについて十分記載されているか。

3. 試験治療

- 1) 試験治療のエビデンスは十分に記載されているか。
- 2) 標準治療に比べての リ スク / ベネフィット は十分記載されているか。
- 3) 上記の リ スク は許容範囲内か。最小化する余地があるか。
- 4) その他、開発中の他の薬剤・ 治療法 (欧米で行われている試験など) について記載されているか。

4. 試験デザイン

- 1) 臨床的仮説と 試験デザイン、エンドポイント について十分に記載されているか。
- 2) 上記の記載と 統計記述 (α 、 β 、見込まれる差など) の整合性が取れているか。
- 3) Primary endpoint、secondary endpoints と 試験デザインの整合性は取れているか。

5. 試験の実施可能性

- 1) 予定登録数および年間登録数見込みは十分に記載されているか。
- 2) 上記の記載の実施可能性は適切と判断されるか。
- 3) 保険診療 / 適応外使用 / 医療経済上の問題はあるか。

6. その他

- 1) 記載が非専門分野の研究者にも 分かる 様な平易な表現が用いられているか。
- 2) 附随研究が適切に設定・ 記述されているか。
- 3) 病理・ 効果判定などの中央判定が適切に設定・ 記述されているか。

3) プロト コール審査委員会でのコンセプト 予備 審査

PRC に提出されたコンセプト について、PRC
事務局は書面による 予備審査を行う。予備審査メ

ンバーは、当該グループ以外の医学審査員から
medical reviewer (3 名)、および primary reviewer
と secondary reviewer (各 1 名)、さらに統計外部
委員から 統計 reviewer (1 名)、最後に筆者 (DC

reviewer) の計7名である。審査項目には

1. 対象集団の選択は適切か
2. 対象に対する標準治療は適切か
3. 試験治療の選択は適切か
4. 試験デザインは適切か
5. 試験の実施可能性はあるか
6. 標準治療への貢献度はどの位か
7. 当該疾患の専門家以外の研究者にも研究の意義が理解できるか

があり、それぞれ自由に意見を述べると共に、各項目に「5(特に優れている)、4(優れている)、3(ふつう)、2(やや劣る)、1(劣る)」の5段階で点数を付ける(満点35点)。この点数はJCOG運営委員会での採否決定の参考資料となる。

審査委員のreview sheetはPRCメンバー全員で構成されるPRCメーリングリストに投稿され、PRCメンバーで共有されるとともに、当該グループの医学審査員を介して、研究事務局を含む当該グループにフィードバックされる。

当該グループはreview sheetの審査意見や質問に対し、PRC会議までに、可能な限りメーリングリストにて回答を行う。メーリングリストではreviewer以外のメンバーでも自由に発言することができる。

4) プロトコール審査委員会会議での検討

予備審査を経て、会議によるコンセプト検討(PRC会議)が行われる。PRC会議は通常18時に開始し、短い場合は2時間程度だが、3時間を超える場合もある。

PRC会議には、コンセプトを提出した研究事務局と当該グループの医学審査員1名以上の出席が義務づけられており、可能な限り当該グループの代表者も出席するよう要請している。さらに

当該グループの他の研究者も出席可能となっている。PRCメンバーはreviewer以外であっても可能な限り出席してもらい、JCOGデータセンターの統計部門とJCOG運営事務局の研究支援部門は原則として出席するため、毎回概ね30~40名の会議となる。PRC会議では原則として、研究支援部門の医師が司会を行う。進行は疾患の背景→試験の対象集団→対象集団の標準治療→試験治療の順に議論し、それらを踏まえて試験デザインの検討を行う。

PRC会議には当該グループの医学審査員、統計reviewer、データセンターのグループ担当統計家が出席しているため、たとえ当該グループでは常識となっていることでも他分野の研究者から見て疑問に思われることについては、遠慮なく第三者的な批判が行われるため、他分野の専門家同士のpeer reviewと位置付けられる。当該グループは疑問に対して、他分野の専門家が納得するrationale(論拠)示さなければならぬため、この会議は若手の臨床研究者の教育の場としても貴重な機会になっている。

書面による予備審査で未解決となっていた問題の多くは、PRC会議で直接議論することによって解決の道が見つかることが多く、当該グループ側はJCOG運営委員会での採否決定(本審査)までに、コンセプト原案における問題点の解決策を練ることができる。

PRC会議で検討された内容は、研究支援部門の医師が作成する議事録にまとめられ、PRCメーリングリストによって出席者間で内容が確認される。議事録は、JCOG運営委員会への提出資料となる。議事録には、PRC会議で当該グループ側が示せなかったデータやエビデンスを「追加コメント」として記載することができるため、JCOG運営委員会会議での検討をスムーズに進めることに役立っている。

5) JCOG 運営委員会での審査と採択

予備審査、PRC 会議で検討されたコンセプトは、年 4 回開催される JCOG 運営委員会会議の場で行われる本審査にて採否が決定する。

運営委員会の本審査では、まず、PRC に提出されたコンセプトを用いて、当該グループ代表者もしくは研究事務局が試験の概要を説明し、質疑を行う。続いて primary reviewer が PRC 会議の要約と PRC 会議で出されたアドバイス等が議事録を用いて説明される。

運営委員全員での質疑応答の後、運営委員全員が「採択」「却下」「保留」のいずれかに挙手をする。出席している運営委員の過半数が「採択」に挙手した場合、そのコンセプトが「承認」され、JCOG 試験としてのプロトコール作成が許可される。多くの場合は運営委員の大多数が「採択」に挙手することで、承認される。しかし、「採択」挙手がかろうじて過半数になることもあるし「採択」挙手が過半数に届かず、「却下」されることもある。

コンセプトで重視される「背景」

●コンセプトで重要なのは「試験の意義を示す記述—声明文」的側面

前回紹介した S. J. Pocock の「Clinical trials: a practical approach」(1984, John Wiley & Sons Ltd.) には「プロトコール」の意義として、「試験の意義を示す記述—声明文」的側面と「試験の実施手順—マニュアル」的側面の 2 つの側面があるとされている。

JCOG におけるコンセプトの取り扱いを考えると、重要なのは「試験の意義を示す記述—声明文」的側面であることが理解されよう。臨床試験を行うグループは他の分野の専門家に対して、その研究の重要性、計画の適切性、実施の可能性を示す必要があり、そのツールがコンセプトと言える。

したがって、コンセプトで最も重要なのは、表 1 における「3. 背景情報」であり、表 2 における「2. 背景」なのである。

試験によって求められる結論が臨床的に価値の高いものであれば、導かれる結果を信頼性の高いものにすることは、ある意味技術的な問題と言える。試験デザイン上の問題点はほとんどの場合、統計家を初めとするデータセンターのメンバーが知恵を絞ることによって、解決が可能であるためである。

●臨床研究者の最も重要な使命はコンセプトの記述にある

試験を計画する臨床医(臨床研究者)の最も重要な使命は、臨床的価値の高い研究仮説(clinical question)を設定すること。そしてその重要性を他分野の専門家に理解してもらうと同時に、協力が得られるようなコンセプトの記述(作文)をすることである。

前回強調した試験デザインの骨子である「患者」「治療」「評価」を「標準治療は何か？」を核として具体化するプロセスは、今回示した JCOG および CTEP の書式に従って、質の高いコンセプトを記述するプロセスそのものと言える。前回と今回の記事を併せて参照しつつ、質の高いコンセプトの作成を試みていただきたい。

新規分子標的薬剤の臨床試験及び承認のあり方

Points to consider on regulatory issues and clinical trial designs for development of molecular targeted drugs

柴田大朗、福田治彦

国立がんセンターがん対策情報センター
臨床試験・診療支援部、薬事・安全管理室長
同、臨床試験・診療支援部長

はじめに

分子標的治療薬の開発方法は、特殊であるか？答えはイエスでもありノーでもある。細胞傷害性の薬剤の開発とは趣の異なる点がある一方で、癌の領域に限らず一般の医薬品の開発という観点からみると特殊とはいえない点もある。癌の分子標的治療薬を評価するうえで、何が特殊で何が特殊でないかを知っておくことは重要である。

開発の流れ

一般的な抗癌剤開発の基本的な流れは、おおよそ以下のとおりである。第Ⅰ相で用量規制毒性の発現状況とMTD(Maximum Tolerated Dose：最大耐用量)を把握し、用法・用量と対象となる癌種を規定し、第Ⅱ相で主として奏効割合を通してその効果を把握する。その段階で効果が十分期待できる場合には、第Ⅲ相で

既存の標準治療との比較を行う。

通常、医薬品は、薬剤としての効果が存在することの確認を経て、現在の治療体系のなかに組み込まれる臨床的な有用性が存在するかを確認するという流れで開発されるが、癌以外の領域に比べて、癌に対する治療が多くの場合集学的に行われるものであるために、医薬品の承認を取得するという観点に絞ってみると第Ⅲ相試験の組み方が難しく感じられる場合もある。また、分子標的薬の開発にあたって選択される試験デザインのなかには、従来の抗癌剤の一般的な開発ではあまり用いられなかったものも含まれる。そのような新しいタイプの試験については当然のことながら批判的な評価がなされることもあり、標準治療を確立するという観点からは、分子標的薬の臨床試験として実施されることがある試験デザインのなかには、たとえばRandomized Discontinuation Design(RDD)¹⁾のようなデザインでは被験薬の臨床的な位置付けを明確することに役に立たないのではないかと疑問を感じる

向きもあるだろう。

医薬品の開発は step by step で行われるものであるという原則に立ち返ってみると、ひとつひとつの臨床試験を取り上げてそのデザインの是非を議論するだけでなく、その試験が開発のどの段階で実施されたものであるのか、何を示すことを目的として計画されたものであるのかを含め、開発の各段階(相, phase)においてそれらの臨床試験がどのような役割を果たしているのかを考慮したうえで議論する必要がある²⁾。この視点を抜きにして、たとえば「ランダム化」が行われているという共通点のある、一般的な第Ⅲ相比較試験、RDD による比較試験、ランダム化第Ⅱ相試験を相互に比較してみても意味がない。RDD による比較試験は標準治療を確立するために実施する第Ⅲ相比較試験とは試験の目的が異なり、また、ランダム化第Ⅱ相試験は決して検出力の低い第Ⅲ相比較試験ではない。

試験デザイン

抗癌剤の開発に関しては、多くの経験を経て一般的に受け入れられている開発方針が存在する。その開発方針は、細胞傷害性の薬剤に最適化されているところがあり、分子標的薬の開発にあたっては必ずしもその開発方針が適切である場合ばかりとは限らない。ただし、分子標的薬であっても腫瘍縮小効果が明確に観察できるような薬剤の場合は、第Ⅱ相・第Ⅲ相の開発方針は細胞傷害性の薬剤と同様な考え方で実施できることもあり、分子標的薬の開発に従来の開発方針は役に立たないという紋切り型の理解は不適切である。ポイントは、開発しようとしている分子標的薬がどういう観点で従来の抗癌剤と異なるか、何によって効果を測ることができるかを明らかにし、それらをふまえた適切な試験デザインと開発戦略を選択することである。これは癌以外の領域の医薬品開発においては通常行われていることであり、分子標的薬であるから特別な考え方を採用しなければならないわけではない。

細胞傷害性の薬剤に関して腫瘍縮小効果の有無をみ

るということは、ある意味で Proof of Concept とも考えられる。一方で、直接的な腫瘍縮小効果が期待できない薬剤については、奏効割合によって「効果の手応え」を把握することができないために、ほかのなんらかの評価を行う必要性が出てくる。増悪までの期間を評価する、Stable Disease の割合をみるなどもひとつの見方である。しかし、これら进行评估するには、縮小を評価する場合以上に疾患の自然経過や観察期間などの時間の要因について着目する必要がある。一般に、こういった時間が影響を与える変数をヒストリカルコントロールと比較することは、奏効割合の場合以上に困難である。そのため、ランダム化第Ⅱ相試験や RDD による比較試験などによって、用法・用量の選択を行う際の参考とする、あるいは第Ⅲ相に進むに値する効果があるか否かを判断する、といったケースが出てくることになる。

分子標的薬の開発時に用いられることがある試験デザインはさまざまあるが、ここではランダム化第Ⅱ相試験³⁾、RDD による比較試験、薬剤の投与対象の絞り込みを意図した enrich を取り入れた試験⁴⁾の概略を簡単に紹介する。これらは、必ずしも目新しい試験デザインというわけではなく、癌以外の領域では従来も用いられてきた例がないわけではない。

ランダム化第Ⅱ相試験とは、第Ⅲ相に進む可能性のある複数の候補治療のなかから、1つの治療を選択する目的で実施される試験である。よく用いられるものとしては、Simon, Wittes and Ellenberg の選択デザイン⁵⁾が挙げられる。被験者を複数の候補にランダム化し、結果として1番成績のよかった群を第Ⅲ相の試験治療として選ぶという方法である。この方法は通常、候補のなかのどの群にも安全性や利便性などの面で大きな違いがないという状況下で行われる。逆にこのような試験デザインの性格を利用して、分子標的薬の開発にあたって、毒性に大きな違いのないような用法・用量の候補のなかから第Ⅲ相に進めるものを選択する際に用いられることもある。ただし、ランダム化第Ⅱ相試験にも多くのバリエーションがあるが、多くの場合統計学的には第Ⅰ種の過誤、第Ⅱ種の過誤はコント

ロールされていない。そのため、結果として群間で統計学的に有意な差が観察されたとしても、その結果をもって第Ⅲ相試験の代替とみなすことは困難である。

RDD による比較試験は、全例に実薬を投与し、薬剤に対して反応のなかった被験者と著しく効果のあった被験者を除外し、ある程度の幅の反応のあった被験者を実薬投与継続かプラセボ投与かにランダム化する方法である。この試験によって、「ある程度の反応」が被験薬によって引き起こされたものであるのかについて結論を下すことができる一方で、観察された群間差は無効例や著効例を除外したうえでの差であるために、実際の臨床現場で用いた際の被験薬の効果を示すものではない。本試験はあくまで薬効が存在することを示すための試験デザインであり、標準治療の確立などの薬剤の臨床的な位置付けを明確にするデザインではないと考えることが妥当であろう。

Enrich を取り入れた試験は、被験者の分子標的に関する状況を選択基準に含めた試験である。被験薬が効果を発揮することがないと考えられる集団を除外して実施する試験であるため観察される群間差が大きくなることが期待され、enrich を行わない試験と比べ被験者数は少なくなる。ただし、将来、この試験の対象外となった被験者においても当該薬が使われる可能性があるのであれば(あるいはそのような可能性を残して開発を進めたい場合には)、enrich を行った試験は治療開発全体の流れを俯瞰してみると必ずしも適切とはいえない場合がある⁹⁾。臨床現場において実際に使われるであろう部分集団に対して、有効性や安全性の情報が確認されないままになるためである。そのような場合はランダム化前にすべての被験者の分子標的に関する状況を測定・記録しておき、分子標的の状況と治療との交互作用を調べるデザインが適切である場合もある。この試験デザインの実例としては、米国 FDA で承認を取得したのち、市販後に実施されている erlotinib の試験がある¹⁾。

そのほか、治療個別化のためのバイオマーカーの妥当性を評価する試験デザインについては、Sargent らの総説がくわしい⁸⁾。

承認のされ方

分子標的薬は従来の抗癌剤と異なり、患者の分子標的に関する状況に応じて投与の可否を判断することができる場合がある。そのため、医薬品が承認される際に投与対象が特定の条件を満たす患者のみに絞られるケースがある。これについては、本邦に限らず欧米においても方針が明文化されているわけではないが、実態としては作用機序と得られているエビデンスとの関係をふまえて判断が下されている。

分子標的と薬効との関係は、理論的に関連があるか否かという作用機序の面からの議論と、それによって投与対象となる患者を絞り込むことが適当か否かという臨床的な判断をとともなう議論とを切り分けて考える必要がある。投与対象を厳格に絞る、あるいは投与対象を絞ることを推奨するといった判断をすべての場合において規制当局が下すべきか否かは立場によってさまざまな考え方があるだろうが、ひとつの考え方として、米国 FDA における添付文書中のバイオマーカーの扱いの整理を紹介しておく (Table of Valid Genomic Biomarkers in the Context of Approved Drug Labels▶ http://www.fda.gov/cder/genomics/genomic_biomarkers_table.htm)。ここでは理論的な背景とエビデンスとをふまえて、個々に “test required”, “test recommended”, “information only” というように、その扱いの軽重がつけられている。分子標的薬という区分は、万能の医薬品の代名詞ではなく、また、どのような状況下でも常に患者に対して個別化した薬剤の投与を可能とするものでもない。個々の薬剤の性質と得られた試験結果をふまえた評価・扱いがなされることが重要である。

References

- 1) Rosner GL et al: *J Clin Oncol* 20: 4478-4484, 2002
- 2) 厚生省医薬安全局審査管理課長通知:「臨床試験の一般指針」について、医薬審第 380 号, 平成 10 年 4 月 21 日, [ICH E8▶ http://www.pmda.go.jp/ich/e/e8_98_4_21.pdf]
- 3) Rubinstein LV et al: *J Clin Oncol* 23: 7199-7206, 2005

- 4) Simon R et al : *Clin Cancer Res* 10 : 6759-6763, 2004
- 5) Simon R et al : *Cancer Treat Rep* 69 : 1375-1381, 1985
- 6) 宇山佳明 : *Cancer Frontier* 7 : 195-205, 2005

- 7) Cohen MH et al : *Oncologist* 10 : 461-466, 2005
- 8) Sargent DJ et al : *J Clin Oncol* 23 : 2020-2027, 2005

連載「がん臨床試験の実践～JCOGを例に～」第3回

プロトコールの作成と審査(前編)

——異なる背景を持つ人が使うコミュニケーションツール

国立がんセンター がん対策情報センター 臨床試験・診療支援部

福田 治彦

第1回では試験デザインを、前回は試験デザインを文書として第三者に示すツールである「プロトコールコンセプト」を取り上げた。今回は臨床試験のメインツールである「プロトコール」を取り上げる。プロトコールの内容は大部となるため、2回に分けて掲載する。前編の今回は、プロトコールとは何か、プロトコールは誰が読むか等を確認し、後編ではプロトコールの標準化および各章で注意すべき点を述べる。

1. プロトコールとは何か？
(What is protocol?)

●プロトコールの定義

一般的には「研究実施計画書」、治験では「治験実施計画書」のことを「プロトコールまたはプロトコル(protocol)」と呼ぶ。そもそも、プロトコールとはどういう意味なのか、広辞苑には「①(条約の)原案。議定書。②外交儀礼。③コンピュータ・システムで、データ通信を行うために定められた規約。～以下省略～」と記載されている。一方、Oxford 現代英英辞典では、以下のように記載されている(カッコ注、下線は筆者による)。

「1. a system of fixed rules and formal behaviour used at official meetings, usually between governments, 2. the first or original version of an agreement, especially a treaty(協定、条約) between countries, etc., 3. a set of rules that control the way data is sent between computers, 4. a plan for carrying out a scientific experiment or medical

treatment.]

つまり研究実施計画書をプロトコールと呼ぶのは原義からの転用であって、古くは、条約原案や外交儀礼書がプロトコールと呼ばれていたことがわかる。さらにフリー百科事典『ウィキペディア(Wikipedia)』では、以下のように記載されている(下線は筆者による)。

『プロトコル(protocol、プロトコールとも)とは、複数の者が対象となる事項を確実に実行するための手順等について定めたもの。日本語では、場合に応じて規定、議定書、儀典などと訳される。

国際儀礼上のプロトコルとは、国際的な実務、交流の場における公式なルールや慣習などを指すものである。主に、列席者の序列、国旗の取扱についての原則、パーティーなどの場における進行や服装・マナーについての一般的な運用方法を示すものだが、これといった厳格な定義はない。日本では、外務省大臣官房儀典官室(英語名"Protocol Office")が「外交官及び領事官の接受・差遣、外国人に対する栄典の授与に関する推薦、外交上の儀礼」などといった「プロトコール」を

所掌している。なお、コンピュータ用語などでは "Protocol" は「プロトコル」と表記・発音されることが多いが「国際儀礼」の意味で使う場合は「プロトコール」とされることが一般的である。これは、前者が英語由来の外来語であるのに対して、後者がフランス語に由来する外交用語であることが理由であると思われる。』

以上の引用には、プロトコルの本質を表すいくつかのキーワードが含まれている。1つは「複数の者、国際、国家間、政府間」というユーザー（または読者）に関する特性であり、もう1つは、「formal、official、公式な」という位置づけに関する特性である。すなわちプロトコールとは「異なる背景や文化を有する複数の（異なる職種の）人間が共同で作業を（臨床試験を行う）上でのルールや手順を記述した公式文書」と定義することができる。ちなみに、S. J. Pocock の『Clinical trials: a practical approach』（1984, John Wiley & Sons Ltd.）では「formal document specifying how the trial is to be conducted: どのように試験が行われるかが記述されている公式文書」と「公式」という言葉がきちんと入っている。これに対して、類義語であるレジメン（regimen）は、「a set of rules about food and exercise or medical treatment that you follow in order to stay healthy or to improve your health」（Oxford 現代英英辞典）とある。「なにか身体に良いことをするための一連の手順」がレジメンであるから、筋トレプログラムもレジメンの1つと言えよう。ただし、ダイエットプログラムを「レジメン」と呼ぶかどうかは定かではない。

●プロトコルの位置づけ

本題に戻ろう。以上のことから、「プロトコー

ル」は治療レジメンの上位概念であり、治療レジメンを含めた、研究全体の手順を記した公式文書と位置付けることができる。少なくとも臨床試験においては、研究実施計画書全体をプロトコールと呼び、治療内容の一連の規定を治療レジメンと呼ぶと混乱しないであろう。ちなみに「プロトコールをたくさん走らせる」のように、研究・試験そのものを指す場合もあるが、混乱を避けるためにも、臨床試験そのものは臨床試験と呼び、その手順を記した文書はプロトコールと呼ぶことを推奨する。

2. 誰が読むのか？

（Who reads protocol？）

●メモや覚書はプロトコールではない

例えば、実験方法を先輩から教えてもらう場合、多くの基礎研究者は試薬の調製方法、実験機器の使い方、反応させる時間等を聞いて記録し、それを基に自分なりのマニュアルを作成するであろう。このマニュアルは「作業を確実に行うため、一連の手順を書いた文書」という意味ではプロトコールに似ているが、自分で読む文書、すなわち公式文書ではないため、プロトコールとは言えない。

同様に同じ実験室に在籍する、同じ程度の知識と技術を持った、同じ専門用語で会話する基礎研究者間で共通の作業を行うための文書やメモ書き等も、おそらくプロトコールとは言えないだろう。しかし異なる知識・技術を有する実験助手、または機器や試薬のメンテナンスを担当する助手と手順を共有する目的で作成された「〇〇研究室××遺伝子解析作業手順」といった文書であれば、プロトコールと呼べられると思われる。

●プロトコールにおける3つのユーザー

ところで、プロトコールは、誰が読み・使うの
であろうか？ 筆者はローカルユーザー、セント
ラルユーザー、レビューアーの3つに分けて考え
ている。

1) ローカルユーザー (local users)

ローカルユーザーとは患者・被験者の診療に当
たる医療現場でプロトコールを使う臨床医、CRC
(臨床研究コーディネーター)、看護師、薬剤師、
臨床検査技師などを指す。多くの試験では、職種
や対象疾患に関係なく、ユーザーになることが考
えられる。したがって、プロトコールは当該疾患
を専門とする臨床医には常識であることについて
も、非専門家のユーザーにも理解できるように詳
しく書かれていなければ、その機能を果たし得な
いことになる。

こうした認識は、既に20年以上の歴史を有す
るJCOGであってもまだ浸透しているとは言い難
く、未だに「施設の担当医にとっては自明である
から、プロトコールに記載する必要なし」と考え
る臨床医もまだ数多く存在する。

2) セントラルユーザー (central users)

多施設共同臨床試験を行う際、施設のリソース
を除いた中央機構のことを「セントラル」と呼ぶ。
セントラルユーザーについて、JCOGを例にする
と、JCOGデータセンターのデータマネージャー、
統計家、JCOG運営事務局のプロトコールコー
ディネーターや試験進捗情報管理担当者、各種委
員会の事務局担当者が該当する。

その他にも企業治験におけるモニターや、施設
訪問監査の担当者もセントラルユーザーと言え
るだろう。

3) レビューアー (reviewers)

上記1と2は試験実施者側のスタッフであるた
め、「ユーザー」と位置付けたが、ユーザーとは
異なる用途でプロトコールを読むレビューアーが
いる。

レビューアーはローカルレビューアーとセント
ラルレビューアーに分けて考えることができる。
ローカルレビューアーには、施設の倫理審査委
員会の委員や事務局があげられる。一方、セン
トラルレビューアーは、プロトコール審査委員
会 (Protocol Review Committee: PRC) や効果・
安全性評価委員会 (Data and Safety Monitoring
Committee: DSMC) の委員や事務局、規制当局
で許認可のために試験結果を審査する審査官があ
げられる。

これら1~3を通じて言えることは、実際にプ
ロトコールを読み、使用する人を考えると、当該
疾患の専門家であっても、プロトコールを作る側
の責任者である臨床医 (JCOGでは研究事務局と
呼ぶ) は、人数の上ではごくわずかに過ぎないとい
うことである。したがって、その人にしか理解
できない書き方でプロトコールを作成し、「事足
れり」と考えることは如何に浅薄であるかがわか
ると思う。

3. 何が書かれるべきか？

(What should protocol
include?)

●プロトコールとプロトコールコンセプトの 違い

前回、プロトコールの意義として「試験の意義
を示す記述・声明文」的側面と「試験の実施手順
マニュアル」的側面があり、プロトコールコン

セプトでは声明文的側面がより重要であると述べた。しかし、プロトコールコンセプトはあくまでも研究組織が試験実施の是非を検討するための内部文書であり、たとえ質の高いコンセプトが作られたとしても、プロトコールで声明文的側面を省略できるわけではない。

●プロトコール記述における2つの側面

「プロトコールに何が書かれるべきか？」については、 Pocock の『 Clinical trials 』に則って、「試験の意義を示す記述—声明文」的側面と「試験の実施手順—マニュアル」的側面から概説する。

1) 試験の意義を示す記述—声明文的側面

声明文的側面としてプロトコールに書かれるべき内容は、以下の通りである。

- ① なぜ、今、この試験を行う 価値があるのか (rationale)
- ② 研究者相互の認識の共通化に必要な情報
- ③ 試験実施機構の意思決定 (その試験を行うかどうか) の判断材料となる情報
- ④ 施設が受け入れるかどうか (施設 IRB 審査) の判断材料となる情報
- ⑤ 患者さんへの説明に必要な情報

上記のうち、プロトコールコンセプトのシステムを有している研究組織では、①～③をプロトコールコンセプト、およびプロトコールでカバーしているが、④と⑤はプロトコールのみでカバーしなければならない。また、④と⑤の内容はプロトコールの「背景」に記載されるべき内容である。

2) 試験の実施手順—マニュアル的側面

マニュアル的側面としてプロトコールに書かれるべき内容は、以下の通りである。

- ① 対象患者の特定、登録の手順
- ② 治療の内容、評価の方法
- ③ データ収集と管理の方法、解析の方法
- ④ 責任体制、連絡先

前章で述べたように、これら試験の意義や実施手順は、非専門家の臨床医および臨床医以外で試験に携わる人が理解できる文章で書かなければならない。当然、日本語として正しく、かつ読み易いことが必要であることは言うまでもない。

4. どういう構造であるべきか？

(What should protocol be structured?)

●プロトコールの“お作法”とは

プロトコールの章構成について、誰がいつどうやって決めたのか定かではないが、『 Clinical trials 』には、表 1 のような例が載っている。プロトコールの章構成はそれぞれの組織が各々決めてきたが、いくつかを比べてみると一定の“お作法”があることに気付く。表 2 に米国 The Southwest Oncology Group の例を示したので、まずは表 1 と表 2 をじっくりと見比べてみていただきたい。

章構成についてはヨーロッパの European Organization for Research and Treatment of Cancer (EORTC) や日米欧三極合同での ICH-GCP も概ね同様である。こうした共通の“お作法”というべき慣習を整理してみたのが表 3 である。“お作法”の例を 1 つ述べると「試験の意義を示す記述—声明文」的側面は試験の意義 (背景、目的) に書かれており、それ以外の章は基本的には「試験の実施手順—マニュアル」的側面が書かれていることに気づくだろう。