

表1 相ごとの方法論のエッセンス

	Phase I	Phase II	Phase III
目的	phase IIに進むか どうかを決める 安全性のスクリーニング phase IIでのレジメン (用法/用量)を決める	phase IIIに進むか どうかを決める 有効性のスクリーニング 毒性プロファイルの充実 治療変更規準などの最適化	標準治療を決める 次は日常診療 総合的なリスク/ベネフィット 評価の検証的試験 標準治療とのランダム化比較
主たる エンドポイント (ものさし)	・毒性(MTD, DLT) ・探索的な有効性 (効きそうな薬かどうか)	・有効性の短期的な 代替エンドポイント ・毒性	・有効性の長期的な 真のエンドポイント ・毒性
対象			
疾患要因 (がん種/stage)	広い (がん種を問わない)	> 狭い(特定)	= 狭い(特定)
宿主要因 (年齢/臓器機能)	狭い (正常臓器機能)	< やや広い	< 最も広い
事前の安全性情報の 蓄積と患者のリスク	ほとんどなし (リスク大)	< 少ない (リスク中)	< 多い (リスク小)
参加施設	単施設 (~少数の専門施設)	< 中規模 (専門病院主体)	< 大規模 (一般病院主体)

MTD: maximum tolerated dose, DLT: dose limiting toxicity

はなく、全体の治療開発の効率、倫理性を考慮して決定されなければならない。倫理的な要件を統計を用いて具現化したものが臨床試験における統計の方法論といえる。

3. 相ごとの方法論のエッセンス

具体的な相ごとの方法論のエッセンスを表1に示す。臨床試験の全体像を把握した上で、それぞれの試験の進め方を考えることが大切である。

II. 研究デザイン

1. 比較

臨床試験の基本は比較であり、比較の対照はいずれも標準治療である。phase IIIでは試験治療と標準治療(標準薬, 標準的な治療法)の総合的な有用性を直接同時比較し, phase IIでは過去のデータから導かれる水準や目標(奏効率や有効率), phase Iでは過去のデータから導かれる安全性の許容水準(許容される毒性の割合など)とそれぞれ比較する。よい治療を生みだすた

めに、患者の利益(有効性 [狭義のベネフィット], 安全性 [リスク/ハーム] および経済性)について標準治療との比較を行う。経済性については医療経済の仕組みに左右されるため、有効性と安全性の二つの軸が中心になっている。

2. エンドポイント(図1)

比較に用いる評価指標・評価項目のエンドポイントは患者のベネフィットを測る「ものさし」であり, true または surrogate, ハードまたはソフトというタイプがある。true endpoint は患者のベネフィットを直接反映している真のエンドポイントで, 延命効果(生存期間, 生存率)などがある。surrogate endpoint は真のエンドポイントの代わりに使われるもので, 腫瘍縮小効果(奏効率)など, より簡便で早くわかり, 少ない患者でもわかるなどの利点があり, 真のエンドポイントと評価の結果が一致するものを用いる。中間的なものとして再発抑制効果(無再発生存期間, 再発率)がある。一方, ハードなエンドポイントは生死や生存期間など誰が何回みて

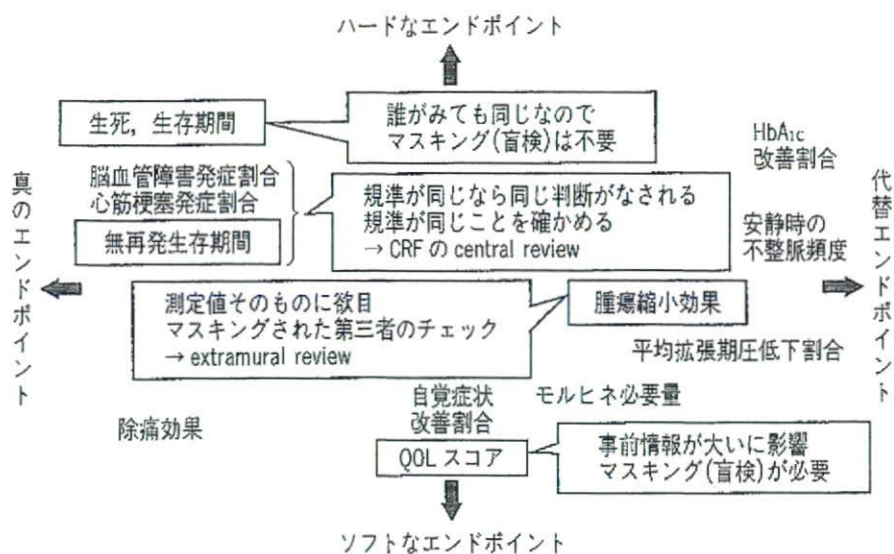


図1 エンドポイント(ハード vs ソフト, 真 vs 代替)
品質管理の方法, 結果の信頼性に関する。

も同じ指標であり, ソフトなエンドポイントは QOL スコア(概念的には QOL は true endpoint であるが, 実際に評価に用いられるアンケートによる QOL スコアは surrogate endpoint)などで観察者や観察時点によって評価が異なり, 他の要因の影響を受けやすいものをいう。ハードなエンドポイントである生死や生存期間にはマスキング(盲検)は不要であるが, 無再発生存期間や腫瘍縮小効果などでは規準が同じになるようプロトコルで定義したり, マスキングなどにより客観性を高める必要がある。

有効性と安全性のエンドポイントを用いてリスク/ベネフィットのバランスを検討し, 最終的な評価を行う。その際, 測定の目的や測定可能なエンドポイントの選択, 測定方法などについて十分検討し, 試験のデザインとエンドポイントの品質管理を具体化することが大切である。

3. 研究仮説(clinical question)

臨床試験を計画する際, たとえば「標準治療に新しい治療法を上乗せした場合に延命効果があるのか?」などの「答えがでる」質問(answerable question)として研究仮説(clinical question)を立てる。がんの治療開発の場合, 毒性が不可避であるため, 現実的かつ倫理的なデザイ

ン, すなわち「何を比べているかを明確に」を優先する hypothesis-based design よりも, 「標準治療 vs most promising regimen の比較」を優先した practice-oriented design を選択するのが好ましい。

III. 試験デザインの Tips & Pitfalls

— 第 II 相試験 —

試験デザインには患者, 治療, 評価の三つの基本要素がある。すなわち, どのような患者に対する, どのような治療を, どうやって評価するのか, ということである。これらの要素の関連性の例を図2に示した。切除不能の再発胃がんの患者に対して新治療(irinotecan+CDDPを増悪まで続ける)を標準治療(5-FU 持続静注療法)の比較対照に, 腫瘍縮小効果をエンドポイントとした。全適格例 40 例中 10 例以上が奏効したら「有望」と判断し phase III へ, 9 例以下なら「無効」と判断し排除するという判断規準(decision criteria)を設定する。患者, 治療, 評価の 3 要素に判断規準を加えた 4 要素が, 互いに関連していることをいくつか例をあげて示そう。

1. 患者が変わると評価が変わる(図3)

たとえば生存期間を primary endpoint とし

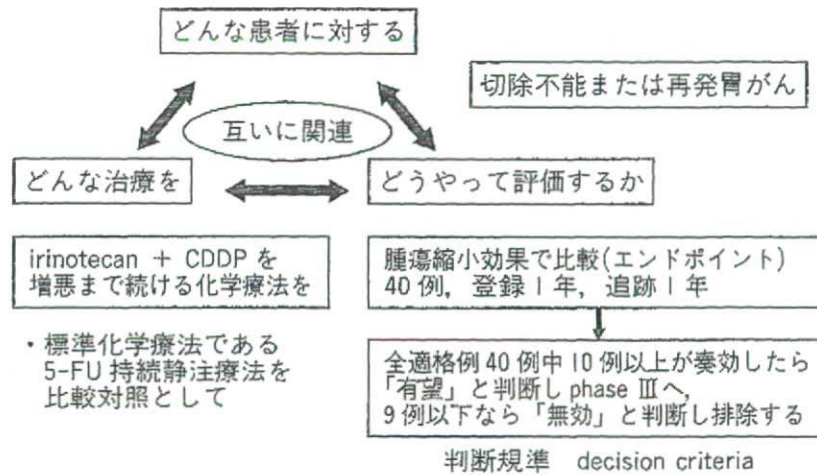


図2 試験デザインにおける三つの基本要素：phase IIの例

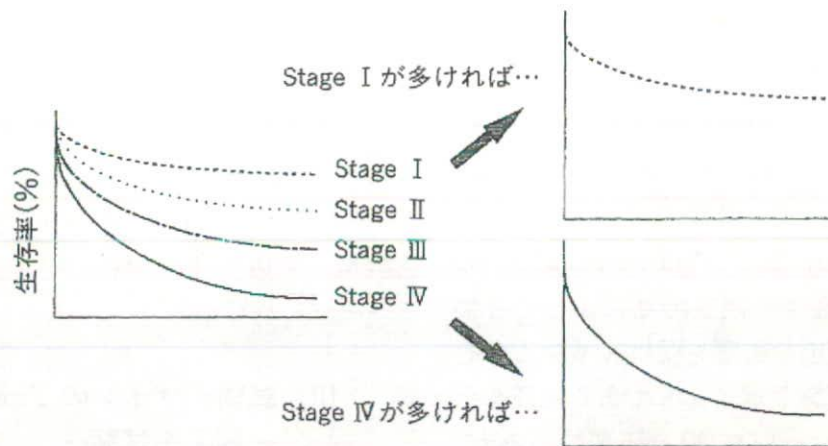


図3 対象によって判断規準は異なる
何と比べて、どれくらいよければよいという判断規準が必要。

である治療を評価しようとする非ランダム化試験を行う場合、病期の異なる患者が混在していると全体の生存期間だけでは正しい評価ができない。対象患者群に予後の良好な Stage I の患者が多ければよい結果が得られ、Stage IV の患者が多ければ悪い結果となるからである。評価したい新治療を受ける患者集団と比較することが妥当な患者集団を特定し、比較における判断規準を明確に定義する必要がある。

2. 治療が変わると評価が変わる (図4)

JCOG 第II相試験 0210 は根治切除可能な大型3型、4型進行胃がん患者を対象に術前化学療法施行後に手術を行い、根治切除割合が手術

単独のデータより高いかどうかを評価する試験である。比べる相手が手術単独なので、化学療法の奏効割合は主たる endpoint とならず、根治切除割合を primary endpoint としている。この試験で有望な結果が得られたので第III相試験 0501 を計画、同じ対象に対して primary endpoint を全生存期間として、試験治療が手術単独を有意に上回れば新しい標準治療であるという decision criteria で実施中である。

3. 治療が変わると対象が変わる (図5)

局所進行膵管がん患者を対象として化学療法を評価するこの第II相試験では、1年生存割合を primary endpoint としており、結果がよけれ

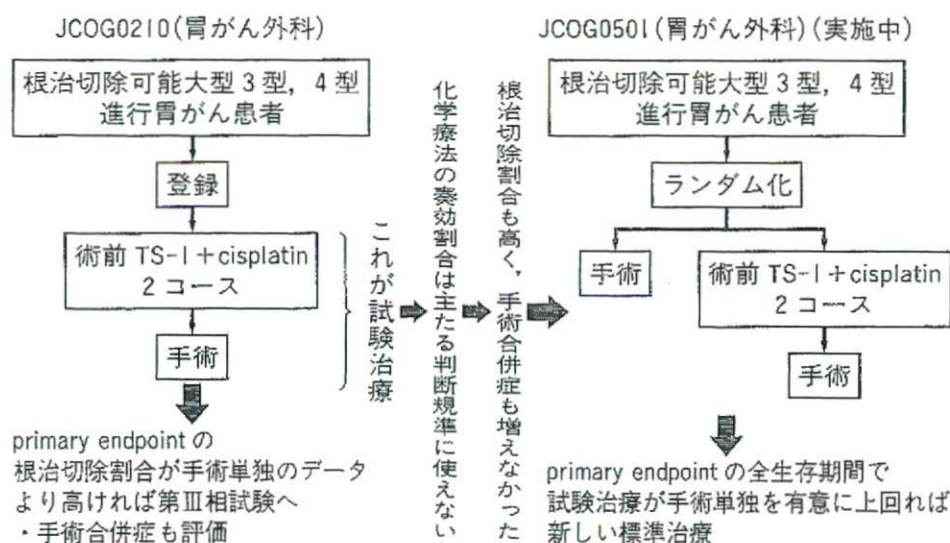


図4 JCOG 第Ⅱ相試験の例1

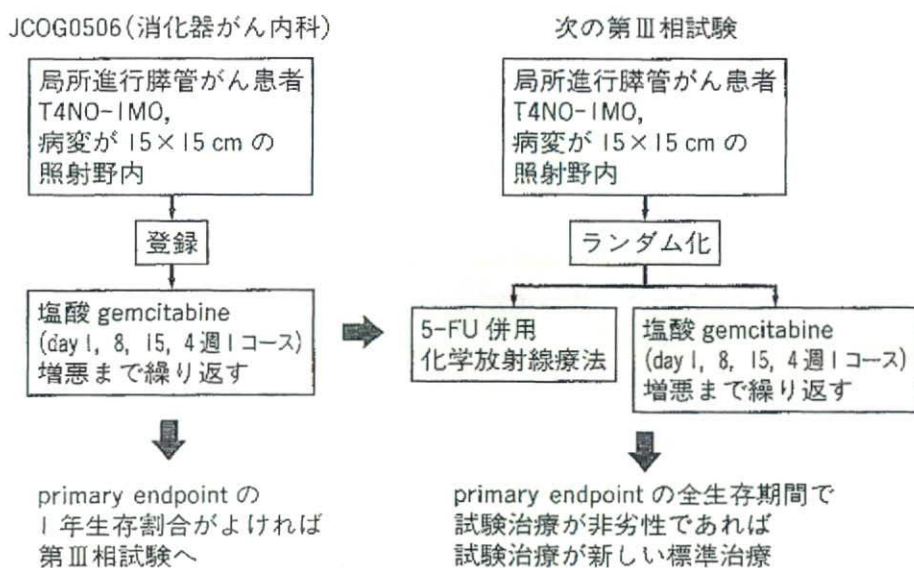


図5 JCOG 第Ⅱ相試験の例2

は第Ⅲ相試験に進めるという前提で実施された。この試験では、対象患者の適格条件に「病変が15×15 cmの照射野内に入っていること」を設定した。その理由は次の段階の第Ⅲ相試験である。第Ⅱ相試験で評価した化学療法と化学放射線療法とを比較する予定であったため、放射線治療の対象となる患者を対象に第Ⅱ相試験を実施した。すなわち、次の第Ⅲ相試験の適格患者になり得る患者を想定して、第Ⅱ相試験の適格条件を決めなければならないということである。

ある。

IV. 試験デザインの Tips & Pitfalls

—第Ⅲ相試験—

1. 優越性と非劣性

第Ⅲ相試験には大きく分けて優越性試験と非劣性試験がある。優越性試験では標準治療に対して毒性が強い新治療を比較し、非劣性試験では標準治療に対して他のメリットがある新治療を比較する。優越性試験では、たとえば標準治

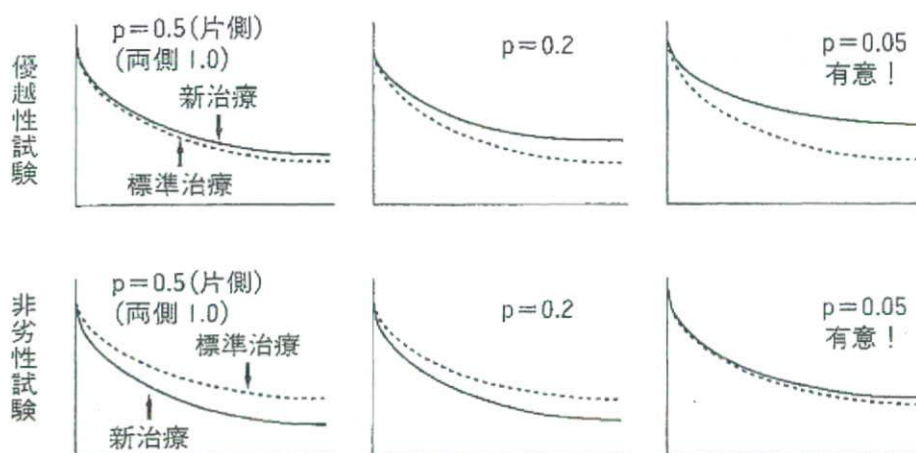


図6 優越性試験と非劣性試験の違い

療の生存曲線よりも新治療の生存曲線が優越していることを期待し、毒性の強さに見合う以上に優れた有効性を示すと仮定して差を見込む。標準治療と比較してこの程度優れていれば臨床医は新治療を選択するという線を決定し、この線と標準治療との差を臨床的に意味のある差 (clinically significant difference) として設定する。一方、非劣性試験では標準治療と比較して新治療は他にメリットがあるため、標準治療に劣っていなければよしとする設定である。たとえば、延命効果は同等であるが毒性が低い、または外来での治療が可能であるなどのメリットがある場合などである。メリットはあるがこれ以上下回るのであれば、標準治療を選択するという線を許容下限として設定する。非劣性試験では標準治療と同等か否かではなく、新治療が許容下限を上回っていることを検証する。

優越性試験と非劣性試験ではまったく手法が異なる(図6)。優越性試験では標準治療と新治療が同等の場合には有意でなく、大きく差が開けば有意であると判断される。すなわち有意に優越性が検証されてこの新治療が新しい標準治療になったという結論が得られる。非劣性試験では許容下限と新治療が同等である場合に有意でないと判断され、新治療と同等程度以上になると有意に劣っていないと判断される。すなわち有意に非劣性が検証されたことになる。第III相試験を計画する際には、まず新治療にどんな

メリットがあるのか、生存は劣っていないか、または優越性を検証しなければならないのかについて明確にし、非劣性の治療開発なのか、優越性の治療開発なのかを明確に区別する必要がある。

2. 多重性

何回も比較を行うとどこかで差がでる可能性は高くなる。これを多重性という。たとえば本当は差がなくても、20回比較するとどこかで有意差がでる確率は6割を超える。

1) 多重性への対処

多重性による誤った判断を回避するためには、あらかじめ宣言した解析のみを主たる判断に用いること、最初に決めておいたサブグループ解析しか結論に用いないこと、サブグループごとに十分なサンプルサイズを確保することなどの対処が必要である。それ以外の ad hoc(後知恵)な解析はすべて探索的な結果であり、別の研究/対象で確認する必要がある。探索的な解析に基づいて検証的な結論を下さないことが原則である。

2) 多重性の調整

上記の原則では1個の臨床試験で1個しか答えがでないため、多重性を調整して、誤った結論を回避する方法がある。最もよく用いられるのが、Bonferroniの調整(Bonferroni correction)である。これは α を比較の数で割る方法であり、10個のサブグループ解析を行った場

合、有意水準 0.05 を 10 で割った 0.005 にして、 $p < 0.005$ のみを有意と考えるという方法である。

3) 中間解析における多重性の調整

中間解析の方法について国際的に確立されているものは、少数回の中間解析をプロトコールであらかじめ規定したとおりに行う、独立データモニタリング委員会の第三者だけがデータを見るということであるが、多重性の調整には様々な方法が用いられており、国際標準が決まっているわけではない。ただし、小さな p 値 (大きな群間差) の時だけ止まるというコンセンサスがあり、最も繁用されているのは O'Brien and Fleming タイプである。

おわりに

臨床試験は一人の臨床医のみで実施できるものではなく、その計画のプロセスも他の臨床医

や生物統計家、データマネージャーなど異なる職種との共同作業である。試験デザインも、その基本となる 3 要素の患者、治療、評価を共同作業で決定しなければならない。その際に明確化しなければならないことは、答えのなる質問として研究仮説を立てること、意味のある仮説であること、対象集団を明確にすること、標準治療を明確にすること、新治療が標準治療と異なる点を明確にすること、エンドポイントと判断規準を明確にすること、有効性、安全性評価の方法を決定し、実施可能性を検討、確保することである。今回、がん治療の臨床開発をモデルに臨床試験方法論の統計的な側面に焦点を絞って解説した。試験デザインにおけるポイントを明確にし、各段階において正しい判断を下すことにより、効率的かつ倫理的な臨床試験の実施と真に有用な新治療の早期開発が可能となる。

癌の臨床試験のデザインとその特徴

黒川幸典/福田治彦

*Designs and characteristics
of cancer clinical trials*

臨床試験は新薬・新治療開発のためのスクリーニングであり、大きく分けて第Ⅰ～Ⅲ相までの3段階からなる。各相にはそれぞれ標準的な試験デザインがあるが、目的に応じて試験ごとに異なったデザインを考慮する必要がある。

臨床試験の流れ

癌の分野に限らず、治療成績の向上のためにはより有効な新薬および新治療の開発が不可欠である。臨床試験は新薬・新治療開発のためのスクリーニングであり、動物などを対象とする前臨床試験を経て、人を対象とした第Ⅰ～Ⅲ相(第Ⅳ相)までの臨床試験が段階的に行われる(表)。多くの新薬・新治療は、そのスクリーニングのプロセスにおいて有効性や安全性の面でふるい落とされ、ほんの一部の新薬・新治療だけが最終的に新しい標準治療として生き残ることができる。

第Ⅰ相試験

第Ⅰ相試験の主たる目的は、抗癌剤の場合には新薬の最大耐用量(MTD)を推定し、次の第Ⅱ相試験での推奨用量(RD)を決定することである。プライマリーエンドポイントは、事前に設定した用量制限毒性(DLT)の出現の有無である。第Ⅰ相試験は、人に対して初めて医薬品が投与されるという不確実な条件下での臨床試験の場合が多いため、癌治療を専門とする単施設もしくはごく少数の複数施設に限って実施されるべきである。また、同時に複数の患者に投与するのではなく、一患者ごとに慎重な観察を経たうえで進められることが望ましい。

第Ⅰ相試験のデザインとしては、古典的なフィボナッチ変法(3例コホート法)が一般的である。これは、最も低い用量レベルから順に3人ずつ投与していく方法で、もし3人のうち1人もDLTが出現しなければ次の用量レベルへ、もし1人にDLTが出現すれば同じ用量レベルにもう3人追加登録、もし2人以上にDLTが出現すれば増量はそこで中止となる。用量の増量方法は、最初の用量を1とした場合、2番目以降の用量は2, 3.3, 5, 7, 9, 12とする。ただし、このような古典的な第Ⅰ相試験のデザインは、統計学的な根拠をもって設定されているわけではないため、MTDやRDの決定規準は試験ごとにさまざまであり、患者背景や薬剤の特性などを総合的に加味したうえで判断する必要がある。なお、抗癌剤以外の外科的手術、放射線治療、集学的治療などにおいては、標準的な第Ⅰ相試験のデザインが確立されているわけではない。

第Ⅱ相試験

第Ⅱ相試験の主たる目的は、新薬・新治療の有効性(腫瘍縮小効果)の推定と安全性の評価である。プライマリーエンドポイントとしては、RECISTという世界共通の判定規準による奏効割合(奏効率)が一般的であるが、そのほかにも病理学的完全奏効(pCR)割合や治癒切除割合などが用いられる場合もあり、試験ごとの工夫が必要である。安全性の評価には、米国NCIによって作られたCTCAEという世界共通の判定規準を用いることが一般的である。第Ⅱ相試験は、通常40~100人程度の患者を対象に行われ、第Ⅰ相試験よりも幅広い複数施設が参加することになるが、参加施設は癌

用語解説——プライマリーエンドポイント

その試験における主たる結論を下す目的で測定される最も重要な評価項目のこと。通常、プライマリーエンドポイントは1つの試験につき1つだけであり、それ以外の評価項目はすべてセカンダリーエンドポイントと呼ぶ。

用語解説—— α エラー、 β エラー

真には帰無仮説が正しいにもかかわらず、これを棄却して対立仮説を採択してしまう誤りを α エラーと呼ぶ。一方、真には対立仮説が正しいにもかかわらず、帰無仮説を棄却できない誤りを β エラーと呼ぶ。

表 臨床試験の相とその特徴

	第Ⅰ相	第Ⅱ相	第Ⅲ相
主たる目的	MTDの推定とRDの決定 (第Ⅱ相に進めるかどうかを決定)	腫瘍縮小効果の推定と 安全性の評価 (第Ⅲ相に進めるかどうかを決定)	標準治療の決定
標準的な 試験デザイン	フィボナッチ変法 (3例コホート法)	一段階デザイン, 二段階デザイン, ランダム化第Ⅱ相試験	優越性試験, 非劣性試験
標準的なプライマリー エンドポイント	有害事象(DLT出現の有無)	奏効割合	OS
参加施設	単施設(専門病院)	多施設(専門病院)	多施設 (専門病院+一般病院)
標準的な患者数	10~20人	40~100人 (閾値, 期待値, $\alpha=0.1, \beta=0.1$ から算出)	200~数千人 (帰無仮説, 対立仮説, $\alpha=0.05, \beta=0.1\sim0.2$ から算出)

治療を専門とする施設が主体となることが望ましい。

第Ⅱ相試験のデザインとしては、古典的な一段階デザインのほかに二段階デザイン、ランダム化第Ⅱ相試験などが挙げられる。二段階デザインとは、早期の段階における無効中止の判定規準をあらかじめ決めておく方法であり、ランダム化第Ⅱ相試験とは、複数の新薬・新治療の候補のなかから次の第Ⅲ相試験に進めるものを1つだけ選びたい場合に用いるデザインである。第Ⅱ相試験の患者数を決定するためには、閾値(帰無仮説に相当)、期待値(対立仮説に相当)、 α エラー、 β エラーなどの設定が必要であり、これらの計算およびデザインの選択に際しては生物統計家の協力が必要である。

第Ⅲ相試験

第Ⅲ相試験の主たる目的は、ランダム化比較にて新薬・新治療と現在の標準治療との直接比較を行い、どちらが標準治療としてふさわしいかを決定することである。第Ⅲ相試験は検証的試験であるため、プライマリーエンドポイントとしては治療による延命効果を最も直接的に評価できる全生存期間(OS)を用いることが一般的である。ただし、予後の良い患者を対象とする場合には、無再発生存期間(RFS)や無増悪生存期間(PFS)などを採用する場合もある。第Ⅲ相試験は、通常200~数千人程度の患者を対象に行われ、癌治療を専門とする施設だけでなく、実地医療に近い状況として一般病院も加えた多施設で実施されることが望ましい。

第Ⅲ相試験のデザインとしては、大きく分けて優越性試験と非劣性試験の2つがある。前者は新薬・新治療が現在の標準治療に比べて毒性が強いものの、延命効果も高いと考えられる場合に用いるデザインであり、後者は新薬・新治療が現在の標準治療に比べて延命効果では必ずしも優っているとはいえないが、毒性は明らかに軽いと考えられる場合に用いるデザインである。第Ⅲ相試験の患者数を決定するためには、帰無仮説、対立仮説、 α エラー、 β エラーなどの設定が必要であり、第Ⅱ相試験の場合と同様に、生物統計家と十分に議論を行ったうえで決定することが重要である。

Recommended Readings

- ① 福田治彦ほか 訳：米国 SWOG に学ぶがん臨床試験の実践—臨床医と統計家の協調をめざして、医学書院、東京、2004
- ② 日本臨床腫瘍学会 編：新臨床腫瘍学—がん薬物療法専門医のために、南江堂、東京、2006
- ③ 大橋靖雄ほか 編：臨床試験の進め方、南江堂、東京、2006

関連事項

- レシスト▶▶ 48頁
NCI-CTC▶▶ 50頁
臨床腫瘍医が必要とする
生物統計学の基本用語▶▶ 52頁



国立がんセンターがん対策情報センター
多施設臨床試験・診療支援部 部長

福田 治彦



レジデント・ノート

— 専門医がわかりやすく解説 —

Q uestion

臨床試験について教えてください。

A nswer

臨床試験とは、よりよい治療法や診断法の開発を目的として、患者さん(がん以外の第Ⅰ相試験では健常人)を対象に行われる研究です。特にがんの治療開発では研究者主導臨床試験が大きな役割を担っています。

□ 「相(phase)」と「試験の種類」について

一般的に医薬品開発は、以下の第Ⅰ相~第Ⅳ相という段階を経て行われます。

- 第Ⅰ相：人に安全に投与できるかどうかを探索する段階
- 第Ⅱ相：当該疾患に対する有効性があるかどうかを探索する段階
- 第Ⅲ相：標準治療よりもよい治療であるかどうかをランダム化比較にて検証する段階
- 第Ⅳ相：市販後の日常診療で広く使われてはじめてわかるような、長期投与による毒性やまれな毒性を調べる段階

各相で行われる臨床試験の種類は、典型的には第Ⅰ相-臨床薬理試験、第Ⅱ相-(有効性の)探索的試験、第Ⅲ相-検証的試験ですが、各相で複数の臨床試験が行われることも多く、抗がん剤以外の医薬品開発においては、第Ⅱ相や第Ⅲ

相で臨床薬理試験を行って用量を決定することもあり、治療開発の「相」と試験の「種類」とは必ずしも1対1対応しません。ただし、がんの治療開発においては、第Ⅰ相で増量しながら最大耐用量と推奨用量を決める試験を行い、推奨用量での腫瘍縮小効果を第Ⅱ相で評価し、第Ⅲ相で標準治療とのランダム化比較試験を行うという方法が定着していますので、試験の種類を「相」で表現する「第Ⅰ相試験」「第Ⅱ相試験」「第Ⅲ相試験」という呼び方が一般的となっています。

ちなみに「検証的研究」とは、その結果に基づいて日常診療の治療方針を決めることができる研究を指し、基本的にはランダム化比較第Ⅲ相試験のみが該当します。「探索的研究」とは、次の研究のために役立つ結果を出す研究のことであり、探索的研究の結果をそのまま日常診療に活用してはいけません。

□ 臨床試験は「スクリーニング」であること

すべての医薬品が第Ⅰ相、第Ⅱ相、第Ⅲ相と順当に進ん

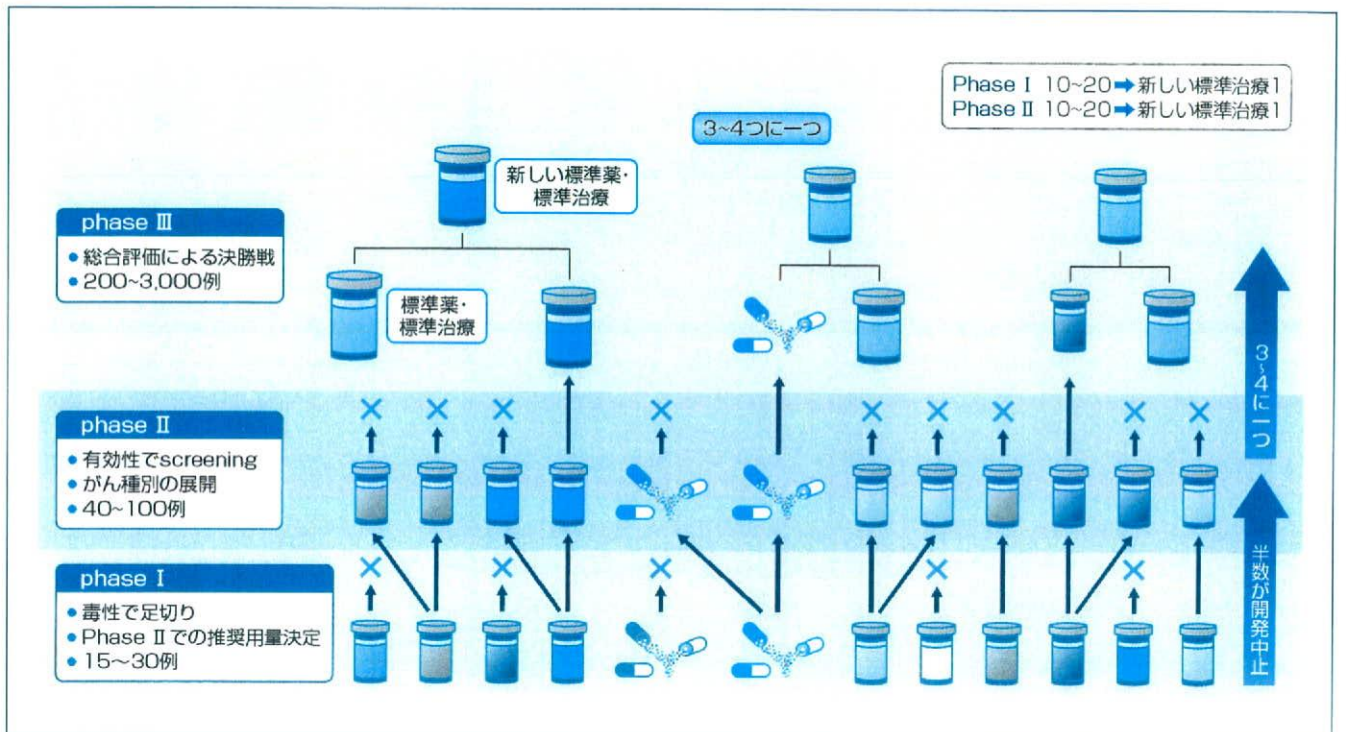


図1 臨床試験はスクリーニング

で、新薬や新しい標準治療が次々に出てくるのであれば、もっと急速に医学(治療学)は進歩していくはずですが、残念ながら治療開発の成功確率は高くありません。第I相に入った医薬品候補が最終的に規制当局の承認を得て市販薬になる確率は、医薬品全体では約10%です。最も確率の高い循環器用薬は20%で、5つに一つの医薬品候補が実際に薬となります。ところが抗がん剤のこの確率は5%しかなく、20の候補からわずか一つの抗がん剤しか生まれません。裏返せば、第I相試験に入った抗がん剤候補20のうち19までは抗がん剤にはならないということです。ですので、臨床試験による治療開発とは、全体として、有効でない医薬品(治療法)の候補をふるい落としていく「スクリーニング」の仕組みとすることができます。図1はその全体イメージと各相での成功確率を示したものです。抗がん剤の場合、第I相でおおむね半分が落とされます。安全であることがわかった抗がん剤候補は第II相に進むときに複数のがん種に展開されます(例:肺癌と乳癌)ので、世の中で行われているがんの第I相試験の数と第II相試験の数はおおむね同じになります。有望であると判断されて第II相から第III相に進むのがおおむ

ね3つか4つに一つ。第III相で標準薬(標準治療)に優って新しい標準薬(標準治療)になるのもやはり3つか4つに一つです。以上より、第I相試験がはじまった段階、第II相試験がはじまった段階ともに、その抗がん剤候補が新しい標準薬になる確率は20分の1程度であり、成功確率がやや高いと思われる分子標的薬の開発や、市販薬を用いた集学的治療の開発でもおそらくせいぜい10分の1程度と思われます。この低い成功確率という厳しい現実があるが故に、十分な根拠もなくやみくもに、また単なる思いつきで安易に臨床試験を行うことは厳に戒められなければなりません。また、無駄な繰り返しが行われなためにも、臨床試験を行った以上は、その結果が期待はずれのものであったにせよ、必ず公表しなければなりません。正しい方法論に基づいて個々の臨床試験を行い、適切に結果を公表することが、結局は、より早くよりよい治療を患者さんに届ける道であるといえます。

□ **がんの治療開発と研究者主導臨床試験について**

がん治療が他の疾患と異なる点として、致死的な疾患であること(患者さんの生命予後が悪い)や、毒性(副作用)が

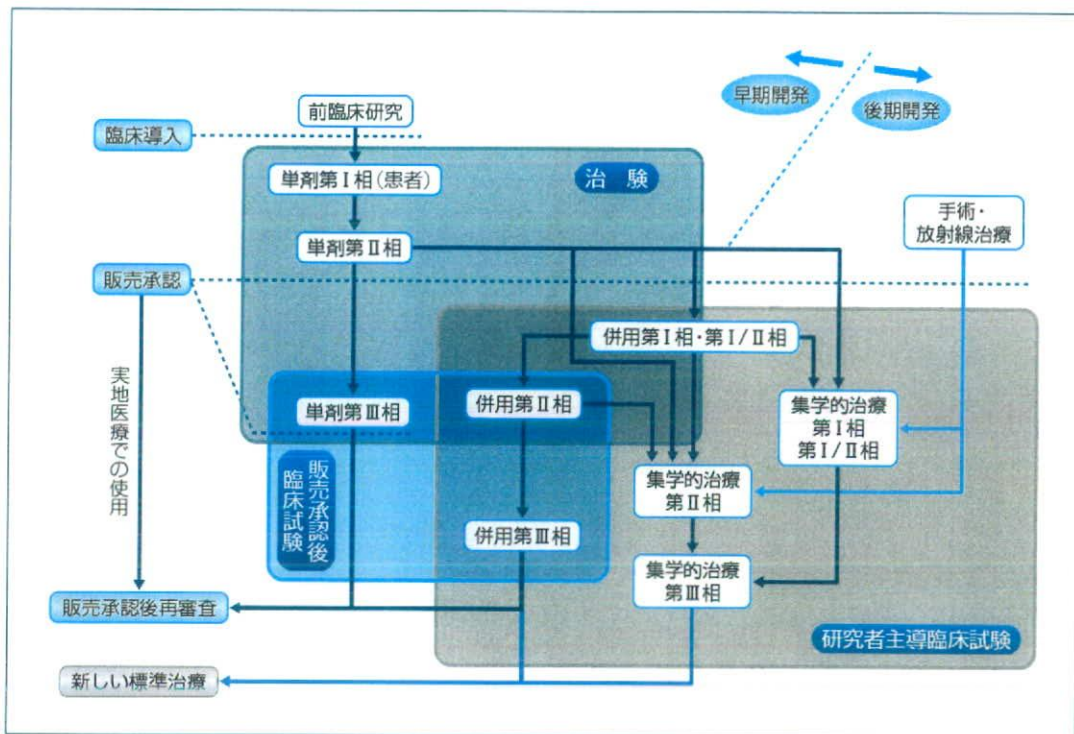


図2 がんの治療開発

必発であることがあげられますが、治療開発としての最大の特徴は、研究者主導臨床試験の占めるウェイトが大きいということです。化学療法、手術、放射線治療が、がん治療の3本柱といわれますが、これらそれぞれ単独で治癒が得られるがん種はごく限られており(例：早期胃癌の手術、頭頸部癌の放射線治療、小児血液腫瘍の化学療法など)、これらの2つ以上を組み合わせる集学的治療(multi-modality therapy)が多くのがん種で標準治療となっています。そのため、図2に示したように、製薬企業による抗がん剤の治療開発だけでは必ずしも標準治療の進歩は達成されず、製薬企業が行わない、集学的治療としての治療開発を研究者主導で行わなければなりません。

がんの治療開発にとって研究者主導臨床試験は“不可欠”なのです。

□ 集学的治療の臨床試験の特徴について

多くが集学的治療の評価が目的である研究者主導臨床試験では、化学療法のみでの評価が目的である治験では用いない試験デザイン上やデータマネジメント上のノウハウが必要

となります。たとえば、標準治療である手術単独に対して術前化学療法を評価する臨床試験では、術前化学療法の奏効率のデータは比べる相手である手術単独療法には存在しません。そのため手術単独でも術前化学療法後の手術でも同じように比べることができ、かつ延命効果に関連していると考えられる根治切除割合(完全切除割合)を試験のprimary endpoint(主たる評価指標)に用いたりします。また、術前化学療法と術後化学療法のどちらがよいかを調べる臨床試験で、増悪と死亡をイベントとする無増悪生存期間(progression-free survival)をprimary endpointに用いたりすると、ランダム化してからの手術の時期が両群間で異なるために非根治切除に終わった場合の解析上の扱いによっては、群間の公平な比較ができなくなり、結局はすべての死亡をイベントとする全生存期間(overall survival)をprimary endpointにすべきであるといったノウハウも集学的治療の研究者主導臨床試験ならではのものです。

集学的治療を評価する臨床試験は方法論的には難しいですが、研究者主導臨床試験ならではの取り組み甲斐のある研究でもあります。

臨床試験におけるデータセンターの役割

Roles of data coordinating center in multicenter clinical trials

木村 綾 福田治彦

Key words : 臨床試験, データセンター, がん

はじめに

‘臨床試験には臨床医がいればよい’という時代は終わって久しいといえるだろう。臨床試験、特に多施設共同臨床試験の遂行に、臨床医の支援機構としてのデータセンターが不可欠であるとの認識は、ここ数年でかなり定着してきた。この認識の変化は、主としてデータの品質管理の必要性和試験結果の信頼性確保の観点からの意識の変化といえるだろう。しかし、多施設共同臨床試験におけるデータセンターの役割として、品質管理・品質保証といった‘管理的側面’のみならず、研究の円滑な遂行を促す‘支援的側面’も重要であるが、モニタリングや監査といった管理的側面の方法論が大筋として確立されてきたのと異なり、データセンターが実際にかかわっていくことが円滑な遂行を助けることとなるのか、という支援的側面については、依然、国内の研究グループがそれぞれ試行錯誤を続けている段階にある。

多数の医療機関が参加する、後期治療開発を中心とする多施設共同臨床試験では、試験の責任者である研究事務局は、参加医療機関の臨床医やCRCらと緊密な連絡を取ることが不可欠であるが、データセンターは研究事務局に代わって、もしくは研究事務局を支援して、臨床試験に関する情報を発信し、施設から報告される

様々な情報を一元管理している。それらを通じて研究事務局の負担を軽減し、試験のより円滑な遂行を支援することが可能であると考えられる。

データセンターが担う役割は、臨床試験におけるデータ入力・管理と解析のみ、と考える諸兄も多いと推察するが、実際にはデータセンターはそれより遙かに広範囲の役割を担っている。データの管理・集計・解析の信頼性は臨床試験の結論の信頼性に大きくかかわり¹⁾、いうまでもなく大変重要であるが、本稿では、臨床試験を円滑に遂行するために必要なリソースとしてのデータセンターの支援的側面の役割を中心に述べる。

JCOGには、その中央支援機構として‘データセンター’と並んで‘運営事務局’が設置されている(図1)²⁾。‘データセンター’と‘運営事務局(Operations Office)’という中央支援機構の二頭体制は欧米の共同研究グループでは一般的であるが、本稿では両者の中央支援機構を一括して‘データセンター’として紹介することとする。また、中央支援機構には、管理的側面の強い有害事象報告を扱う部門と監査を扱う部門も必要であるが、今回は割愛する。

1. データセンターの構成と役割

データセンターで支援的側面を担う部門およ

Aya Kimura, Haruhiko Fukuda: Clinical Trials and Practice Support Division, Center for Cancer Control and Information Service/JCOG Data Center 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/JCOG データセンター

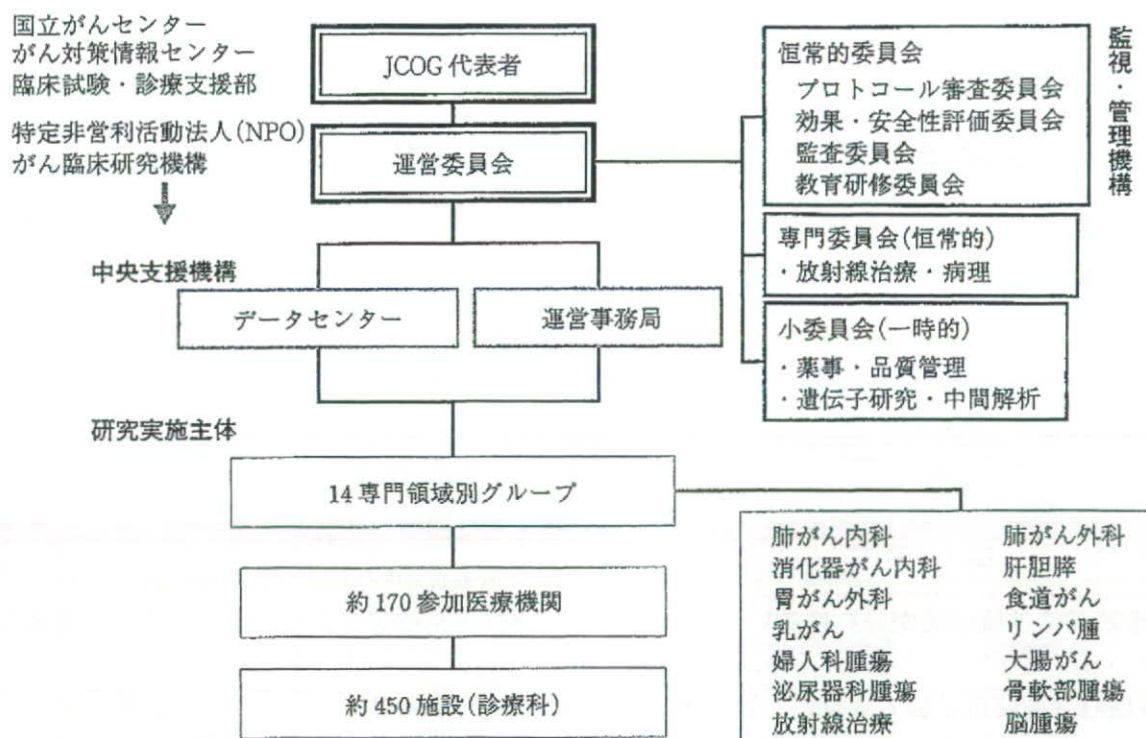


図1 JCOG 組織図

びスタッフの職種は、以下のようにまとめられる³⁾。

a. 統計部門：生物統計家

臨床試験で正しい結論を導くためには、プロトコル作成段階から、統計学的な記述に関して生物統計家の関与が必要である。計画時には試験デザイン(サンプルサイズの計算など)を検討し、試験開始後は中央モニタリング、中間解析、最終解析における統計解析を担う。統計家は多くの臨床試験にかかわるため、臨床試験の運用、方法の専門家としての役割をもつ。

b. データ管理部門：データマネージャー

データセンターの中核をなす部門であり、各データマネージャーは、臨床試験を実施する研究グループ(専門領域別グループ)を担当し、登録業務、データ管理業務を行う。研究事務局に代わって、臨床試験進捗情報を施設の臨床医などに向けて発信し、施設からの問い合わせに対応する。グループの会議に出席し、担当する臨床試験の進捗を報告する。中央モニタリングではレポート作成において中心的な役割を担う。

c. システム部門：システムエンジニア、システムアドミニストレーター

データセンターで使用するシステム(ハードウェア、ソフトウェア、ネットワーク、データベース)の保守管理を担う。

d. 研究支援部門：医師、メディカルライター

JCOGでは研究支援部門がプロトコル作成支援を行っている。プロトコル本体と患者用説明文書の作成支援を行っており、更に試験開始後のプロトコル改訂も支援する。その他、Cooperative Groupでは不可欠の施設・研究者情報管理も担っている。

2. 医療機関とデータセンターのかかわり

多施設共同臨床試験において、データセンターが情報のやりとりを行う相手には、大きく分けて、試験に参加する医療機関の臨床医と、研究事務局や研究代表者という試験の責任者の二者が存在する。まず医療機関とデータセンターのかかわりについて述べる。

患者登録や登録後のCRF(case report form)

記入データの不明点の問い合わせにはデータマネージャーがあたり、研究者の異動や医療機関として試験に新規参加する場合の手続きなどには、施設・研究者情報担当者が対応する。施設側の責任者として JCOG では、施設研究責任者と施設コーディネーターを選出し、特に施設コーディネーターが、データセンターの各部門との連絡窓口を担っている。研究者の所属以外の施設情報としては、施設で使用している臨床検査データの施設基準値を定期的に調査して更新している。

a. 登録開始と治療開始

1) IRB 承認の確認

各試験のプロトコルの施設 IRB 承認(正確には IRB 承認に基づく医療機関の長の承認)が得られると、施設コーディネーターは JCOG データセンターに承認書のコピーを FAX で送付する。以前はこの IRB 承認書は研究事務局が管理をしていたが、試験数が増加したことや、IRB 承認書の書式が施設ごとに異なり、IRB 承認書の確認が次第に複雑になってきたことから、JCOG データセンターで管理することとした。送付される IRB 承認書の中には、条件付承認である場合や、試験計画の変更を要求するものである場合が時折みられる。プロトコル改訂が必要となるような指摘である場合には改訂を行うこともあるが、追加の説明で IRB の承認が得られることがほとんどであるため、そうした場合には必要に応じて、研究事務局とともにデータセンターが、施設 IRB に対する追加説明の文書を作成したりして、施設の研究者を支援している。

2) 登録業務について

患者登録は主に電話または FAX にて行われる。WEB でのオンライン登録を開始したが、まだ 1 つの試験での限定運用のため、それ以外の試験の患者登録は、登録オペレーターが電話または FAX を受けて患者登録業務を行っている。JCOG では登録オペレーターはデータマネージャーが当番制で担当している。登録に際しては適格条件を一つ一つ確認し、すべての適格条件が確認されたらデータベースに患者を登録

する。登録完了後、登録確認通知を FAX し、当該患者用の CRF 一式を施設コーディネーター宛に郵送する。

3) 治療開始

マルチモダリティの臨床試験では、試験登録患者の診療に携わる医師が複数の診療科に存在することになる。informed consent (IC) を取得し登録を行った医師はその患者の‘担当医’として登録され、治療終了後に行われる追跡調査も登録した医師が実施することとなる。その間のプロトコル治療にあたる医師は、登録した医師である場合もあれば、モダリティにより手術担当医や内視鏡医、放射線治療医、化学療法担当医である場合もあり、施設内でのプロトコルの周知徹底が容易ではない。JCOG では、常に最新版のプロトコルを施設の担当者が参照できるように、研究事務局が製本して配布するプロトコル冊子のほか、最新版のプロトコルを研究者専用サイトで公開している(図 2)。

b. 試験実施期間中(治療期間中)

プロトコル治療開始後、施設研究者は CRF に治療経過を記載していくこととなる。CRF 記入に関する質問のほか、それぞれの試験に関する質問の多くも、まずデータセンターに寄せられることが多い。データセンターでは CRF 記入に関する質問についてはデータマネージャーが回答するが、それ以外の治療に関する質問などはすべて研究事務局に報告している。

中央モニタリングの期間(詳細は後述)に入ると、CRF の回収と不完全 CRF に関する問い合わせを行う。検査値の未記入や誤記と思われる項目の問い合わせのほか、後述する研究事務局による CRF レビューにて指摘された項目について、再度施設研究者に問い合わせを行う場合がある。この段階では臨床医・CRC など施設側研究者とのやりとりが多く発生する。主にデータマネージャーが対応するが、この期間の情報共有は臨床試験遂行上非常に重要である。データマネージャーは必要に応じて、研究者から寄せられた情報や CRF にて判明した疑問点などをまとめ、研究事務局に報告する。データマネージャーにとって、臨床的な判断を要する疑問

The screenshot shows the JCOG website interface. On the left is a vertical menu with categories like '肺癌内科' (Lung Cancer Internal Medicine), '肺癌外科' (Lung Cancer Surgery), etc. The main content area is titled 'プロトコル等関連文書' (Protocol-related Documents). Below the title is a table listing documents with columns for 'ファイル名' (File Name), 'ファイルタイプ' (File Type), and 'ファイルサイズ' (File Size). Below the table are two notes: one about file size and another about CRF forms.

	ファイル名	ファイルタイプ	ファイルサイズ
メモランダム/覚え書き、付表	AppendixJCOG0110_070712.pdf	PDF ファイル	59K
プロトコル承認書/改訂承認書	ApprovalJCOG0110v2_1.pdf	PDF ファイル	329K
説明同意文書モデル	ICJCOG0110_070712.doc	ワードパッドドキュメント	66K
プロトコル改訂内容	LetterJCOG0110v2_1.pdf	PDF ファイル	204K
CRF記入マニュアル	ManualJCOG0110v1_2.pdf	PDF ファイル	414K
プロトコル	ProtocolJCOG0110v2_1.pdf	PDF ファイル	515K

【注意】
コンテンツ内のファイルはサイズが大きいものがあるため、デスクトップ等に保存してから閲覧するようになっています。
そのまま閲覧しようとするとファイルが開かない可能性があります。

「CRF記入マニュアル」に接続しているCRFは、必ずしも最新のCRFとは限りません。ご了承ください。

図2 プロトコル関連文書ダウンロードページ

や不明点に遭遇し、それを適切に処理することは、その後の解析に影響を及ぼすこともあるため、かなり難易度が高い。その場合、データマネージャーは研究支援部門に所属する医師に相談し、研究事務局への報告手段や報告の内容、研究事務局への報告前に担当医に詳細を尋ねておくべきか否かなどについて指示を仰いでいる。データマネージャーとデータセンターの医師との連携も重要である。

c. プロトコル治療終了後

治療終了後、プロトコルに規定された期間の追跡調査が行われる。第III相試験と第II相試験が中心のJCOG臨床試験では、プライマリー・エンドポイントとセカンダリー・エンドポイントのいずれかに、必ず生存期間が含まれるため、年余の長期に渡って追跡調査が行われる。治療後に、登録した施設以外の医療機関に患者が転院することもあるため、転院先への照会も施設研究者の役割となるが、この期間中に担当医が異動することも少なくない。そのため施設コーディネーターは、JCOG施設情報担当者に漏れなく担当医の異動連絡をすることが必要である。施設コーディネーターは試験開始時から施設の調整役を担っているが、タイムリーな連

絡が最も必要となるのは、この追跡期間中である。

3. 研究事務局とデータセンターのかかわり

続いて、各臨床試験の責任者である研究代表者・研究事務局とデータセンターがどのようにかかわっているかを述べる。

a. 試験の計画からプロトコル完成まで

1) JCOGにおけるプロトコル作成の手順

JCOGでは欧米のCooperative Groupに倣って、プロトコル作成段階を‘(プロトコル)コンセプト’と‘フルプロトコル’の2段階としている(図3)⁴⁾。研究事務局が立案したコンセプトは、プロトコル審査委員会(Protocol Review Committee: PRC)の検討(予備審査)を経て、運営委員会会議にて審査を受ける。コンセプトが運営委員会で承認されると、フルプロトコルの作成が許可される。フルプロトコルはプロトコルマニュアルに示す章構成に沿って作成され、研究支援部門のプロトコルコーディネーターが窓口となって、章ごとに研究事務局とのやりとりを重ねていく。また、プロトコル作成の初期段階からデータセンターの

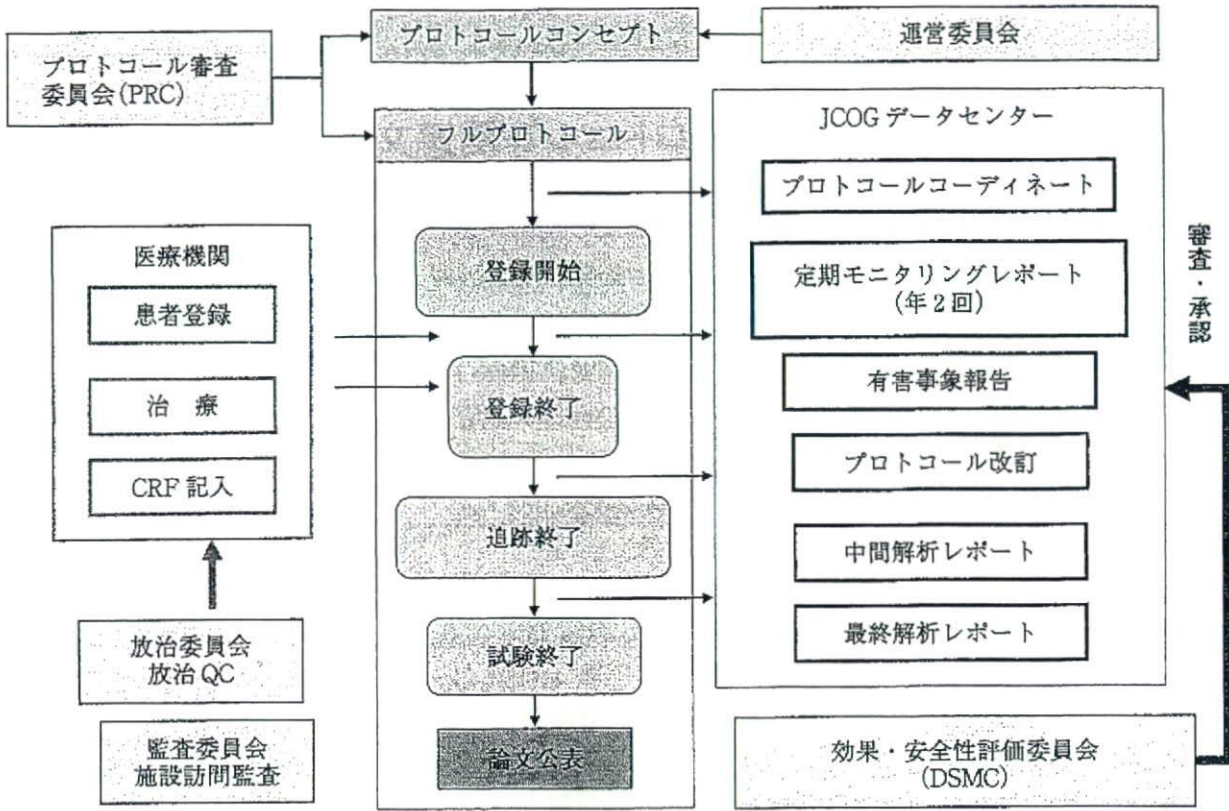


図3 プロトコル作成

生物統計家加わる。生物統計家が初期段階からかわることで、エンドポイントの解析に必要なデータがどのデータであるのか、またそのデータが確実に得られるかといった、実施可能性を含めたプロトコルレビューを行うことができる。プロトコル作成が進むとデータマネージャーも参加し、プロトコルレビューを行う。それらのレビューにより最終プロトコル案が完成する段階になるとメディカルライターによるマイクロエディットを行う。科学的な記載に加えて、日本語としての記述や章ごとの整合性を確認し、体裁を整える。

2) Case report form (CRF) の作成

研究事務局はプロトコルコーディネーターと協力して、CRFの項目を列記した‘CRFドラフト’をプロトコルと並行して作成する。CRFドラフトは、当該臓器がんに共通する背景因子情報や当該臨床試験に特化した臨床データなどを含み、プロトコルの一次審査の審査対象となる。一次審査終了後に、CRFドラフトを基にDM部門にてCRFを作成する。CRFが完成する

と、二次審査に提出されるとともに、並行して、参加施設のCRCや研究者の協力の下、CRF記入テストを行って、書きにくい箇所などの修正を加えてCRFを完成させる。

3) 患者説明文書

臨床試験の説明要件を統一することは、患者が知りうる情報を均一化し、不都合な情報が伝わらないことを防ぐために必要である。JCOGでは患者説明文書にテンプレートを用意し、均質な患者説明文書作成に努めており、また、プロトコルと同様、研究支援部門の担当者が直接加筆して研究事務局を支援している。

b. 試験実施期間中

1) 中央モニタリングと研究事務局によるCRFレビュー

JCOGデータセンターでは、登録開始後の全試験について、原則半年に一度のタイミングで中央モニタリングを実施している。データマネージャーが中心となり、施設からの登録状況、IRB承認状況、CRF回収状況、不適格の可能性、プロトコル治療中の逸脱などを報告し、追跡

調査による生存期間の更新などを行う。これらを掲載したモニタリングレポートが、研究代表者・研究事務局などに送付される。

モニタリングレポートを研究事務局に送付した後、研究事務局がデータセンターを訪れ、モニタリングレポートで指摘された問題点の検討を中心とするCRFレビューが行われる。データマネージャーが発見する問題点や不明点はモニタリングレポートに掲載し検討されるが、研究事務局によるCRFレビューにて発見される問題点も重要である。研究事務局はCRFレビューにより、プロトコル治療の進捗状況を把握し、問題点を直接施設の研究者にフィードバックする。中央モニタリングと研究事務局によるCRFレビューが、試験実施期間中において、最も重要な品質管理活動である。

2) 研究進捗会議

年に数回、研究進捗会議(グループ班会議)が開かれる。臨床試験を円滑に遂行するうえで会議が果たす役割は大きい。会議を利用してモニタリングレポートを参加施設に配布し、問題点の共有を図る。多施設共同臨床試験では、1例ずつの治療経験に加えて、データセンターに集められるデータでの傾向をつかむことも非常に重要である。普段は研究者向けのメーリングリストを通じた情報共有を図るが、会議に参加し研究者の意見を直接聞くことで当事者意識が育ち、以降の施設でのプロトコル遵守やCRF記入に与える影響も大きい。データマネージャーが気付いた情報を、正しく研究事務局に伝え、モニタリングレポートや会議の場で議論されることが、問題点の早期のフィードバックと試験データの質的向上、登録患者の安全性確保につながる。

c. 最終解析、公表、総括報告

すべてのデータが回収されると、最終解析となる。ここでは、プロトコルに規定されているエンドポイントの解析に加え、サブグループ解析も行う。最終解析や統計学的事項については、他稿を参照されたい。解析に用いたデータセットの所有権は研究事務局とグループに帰属するが、CRF、データベースは引き続きJCOG

データセンターに保管される。

学会発表活動の支援として、JCOGデータセンターでは、JCOG試験を発表する際に用いるポスターテンプレートとスライドテンプレートを作成し2008年1月より提供を開始した。JCOG試験の発表が増えてきたことで、統一されたポスターデザインのニーズが高まったことによる。

おわりに

臨床試験に携わる多くの研究者は日常診療を行いながら、臨床試験を立案、計画、実行している。特に臨床試験の責任者である研究事務局には、一定以上の臨床試験の知識と経験が必要であることはいうまでもない。しかし、臨床試験にのみ従事する‘臨床試験専門医’といった存在が臨床試験の責任者になればいいかというところではなく、また、臨床試験は治療法のみでのプロトコルがあれば成立するものではないことも、今更特記すべきことではない。ところが、現実には臨床試験を実施するうえでの理解と実態に解離があることは否めない。それはどこに起因しているのかと考えると、ツールとしてのデータ管理システム、統計解析プログラムがあればことが足りると思われていることが一番大きいのではないかと感じている。

プロトコル作成における適格規準の設定、治療レジメンのデザイン、治療変更規準の設定、試験開始後の想定外の事象に対する臨床的な判断などにおいては、臨床的なセンスとともに当該疾患や当該治療に対する十分な知識と経験が不可欠である。しかし、臨床試験の責任者たる臨床医は、日常診療の‘片手間’で臨床試験に従事せざるをえないという不可避的な矛盾を有する。しかし、だからといって、携わる者すべてが‘片手間’であるようなプロジェクトがうまくいくとも思えない。臨床試験も、やはり‘片手間’でなく毎日その試験のことを考えて働く専任の人間が居なければ成功は望めないであろう。システムとしての臨床試験を稼働させるには、専任の人間が臨床試験全体を管理する必要があると考える。

人に依存したシステムは精度が低いと思われる。ところが、臨床試験は、まさに人を対象とし、人が治療をし、人がデータを発生させている。いかに優れたデータ管理システムや解析プログラムを導入・構築しても、完全な臨床試験システムが構築できるはずがない。臨床試験は人を対象とする以上、想定外の事象は必ず起きるものであり、著者らの経験でも、そうした想定外のできごとへの対応を誤ると試験そのものを失敗に導いてしまうこともある。

多施設共同臨床試験のデータセンターの最も大きな役割は、機械的にデータを収集して解析結果を吐き出すことではない。むしろ、想定外の事象への対応を含めて‘人’の手に寄らざるをえない部分の臨床試験の支援であるといえる。

特に、日常診療の‘片手間’で臨床試験に従事しつつ、よりよい治療を1日も早く届けることを念じて臨床試験に携わる試験責任者の研究事務局と施設の臨床医の間の情報共有やコミュニケーションの媒体となることこそが、‘人’の手になるデータセンターが果たすべき重要な役割であると著者らは考える。臨床医間のコミュニケーションを支援し促進することは、有害事象情報の共有や、より適切な支持療法のノウハウの共有を通じて、臨床試験に参加して下さった患者さんの安全性を確保することにもつながると考える。臨床試験においても、大事なことは人と人とのコミュニケーションであることを強調して稿を終える。

圖文 献

- 1) 福田治彦：研究者主導のがん多施設共同臨床試験におけるデータマネジメント。日本小児臨床薬理学会雑誌 16(1): 76-82, 2003.
- 2) 佐藤暁洋, 福田治彦：JCOGについて。分子細胞治療 6(4): 55-59, 2007.
- 3) 福田治彦：エビデンスをつくる人々。EBM ジャーナル 2(3): 100-112, 2001.
- 4) 福田治彦：「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」。The MEDICAL Oncologists 3(3): 25-32, 2007.

臨床試験プロトコールの書き方^{1*}中村 健一^{**} 福田 治彦^{**}

Key Words : protocol, clinical trial, medical writing, quality control, cancer

はじめに

今号から数回にわたり、臨床試験プロトコールの書き方について概説する。筆者らは、がんの集学的治療を開発するための研究者主導臨床試験を行っている「日本臨床腫瘍研究グループ (Japan Clinical Oncology Group : JCOG)」(<http://www.jcog.jp/>)のデータセンター、運営事務局の一員として、JCOG試験の支援や管理に携わっている。JCOGデータセンターとJCOG運営事務局は、計約40名のスタッフにより、計約90の試験を支援・管理しているが、そのうち計画中の約20試験については、JCOG運営事務局の研究支援部門の5名のスタッフ(うち3名が医師)が中心となり、プロトコール作成やプロトコール改訂の支援を行っている。

本シリーズでは、研究支援部門の日常業務における経験に基づいて、プロトコール作成にかかわる最低限の知識とコツを示していきたい。

プロトコールとは?

「プロトコール」の定義であるが、臨床試験の古典的な教科書である Pocock の *Clinical Trials*¹⁾には、「formal document specifying how the trial is to be conducted : どのように臨床試験が行われるのかが書かれた公式文書」と定義されている。

一般的な日本語訳としては、「研究実施計画書」があてられるが、治験においては「GCP(医薬品の臨床試験の実施の基準に関する省令)」で「治験実施計画書」、臨床研究に関する倫理指針では「臨床研究計画書」とされている。いずれにしても、その研究の手順を詳述した「計画書」をプロトコールと呼ぶ。慣習的には、「当科では現在〇本のプロトコールを走らせています」、あるいは「当科のプロトコールではシスプラチンを隔週で投与しています」など、「研究/試験」、あるいは「レジメン/治療法」の意味で用いられることもある。しかし、これらは誤りとまでは言えないが、“protocol”の原義が「外交儀典書」や「条約原案」といった「(公式)文書」であることから、混乱を避ける意味では、臨床試験やレジメンをプロトコールと呼ぶのではなく、文書である「計画書」をプロトコールと呼ぶことが推奨される。

臨床研究に関する倫理指針

被験者保護、および研究者が円滑に臨床研究を行えることを目的に定められた「臨床研究に関する倫理指針」(以下、臨床指針)が2008年7月に改正され、2009年4月より施行されることとなった。臨床指針では、あらゆる臨床研究において「研究責任者は、臨床研究計画において、臨床研究の実実施計画および作業内容を明示しなければならない」と規定されており、さらに、それらの臨床研究計画が臨床指針に適合しているか否かを、あらかじめ倫理審査委員会に審査を行わせ

* How to write a clinical trial protocol.

** Kenichi NAKAMURA, M.D. & Haruhiko FUKUDA, M.D.: 国立がんセンターがん対策情報センター多施設臨床試験・診療支援部(〒104-0045 東京都中央区築地5-1-1); Clinical Trials and Practice Support Division, Center for Cancer Control and Information Services, National Cancer Center, Tokyo 104-0045, JAPAN

なければならないと記述されている。つまりは、プロトコルを書くことなく臨床研究を行ってはならず、プロトコルは倫理審査委員会で必ず審査されなければならないと明記されている。データの単純な集計にとどまる観察研究の場合の扱いはQ&Aに盛り込まれる予定であるため、詳細は厚生労働省のホームページ(<http://www.mhlw.go.jp/general/seido/kousei/i-kenkyu/index.html#4>)を参照いただきたいが、少なくとも、なんらかの介入を伴う臨床研究(すなわち臨床試験)の場合には、必ずプロトコルを作成し、倫理審査委員会の審査を受けることがこれまで以上に強く求められることになる。実際、改正後の臨床指針には、「3(11)厚生労働大臣等の調査への協力」として、臨床研究機関において臨床研究が臨床指針に従って行われているかどうかの適合性調査(いわゆる「監査」)が行われることが新たに盛り込まれており、実際に施設への訪問調査も行われる見込みである。倫理指針が刑事罰もあり得る連邦法である米国と異なり、わが国の臨床指針は法ではなく「ガイドライン」であって、罰則規定は設けられてはいない。しかし、臨床指針への重大な違反が発覚した場合には、社会的な批判は免れず、民事裁判における判断の根拠となることは容易に予想されることから、実質的にはかなりの強制力をもつと考えられる。

これまでプロトコルと言えば、治験で製薬企業が作成するもの、あるいは、大規模臨床試験でデータセンターが作成するものであって、「プロトコルを書くこと」は自らとは無縁のものと考えてきた臨床医も多いであろうが、今後は、これまでプロトコルなしでも事実上許容されていたと思われる単施設の研究においても、臨床医が自らの手でプロトコルを作成する必要があるであろう。そのため、プロトコルに何を、どのように書くべきか、という知識は臨床研究に携わる医師に欠くべからざるものになったと言える。

プロトコルのもつべき機能

1. 多施設、多職種間のコミュニケーションツール
がん領域の治療開発において、第I相試験や

早期第II相試験といった「早期開発」の臨床試験は、通常がんセンターや大学病院などの専門病院が中心となり、単施設～少数施設の研究として行われる。早期開発の段階では、評価する治療法の開発経緯や過去の安全性情報など事前にわかっている情報が少ないため、プロトコルに盛り込むべき情報は比較的少ない。そして、未知の有害事象が生じる可能性が高いため、治療変更や有害事象に対する対処などは、同じ施設内で情報共有を密にして、臨機応変に対応すべき場面も多い。このような段階の試験では、治療変更標準や支持療法に関するプロトコルにおける取り決めはシンプルかつ柔軟にしておき、むしろ、早期開発の専門家としての担当医の臨床的判断や、担当医間の直接のディスカッションを優先することの方が、登録患者の安全性を確保するためには重要である。

これに対して、後期第II相試験や第III相試験といった「後期開発」の臨床試験は、多数の病院かつ一般病院も参加する多施設共同試験の形をとる。早期開発の段階とは異なり、担当医同士がリアルタイムに口頭で密に情報共有を行うことは不可能であるため、文書による情報共有が基本となる。後期開発では、これまでの治療開発経緯や早期開発段階での安全性情報などの情報量は格段に多くなるし、早期開発で得られたノウハウを治療変更標準や支持療法、検査スケジュールとして具体化することができるため、後期開発のプロトコルに含めるべき内容は必然的に多くなる。さらに、後期開発のプロトコルでは、より多くの施設のより多くの研究者がプロトコルの内容を一義的に理解する必要があることから、内容の正確性や整合性、客観性がより強く求められることとなる。そのため、(プロトコルは分厚ければ良いというものではないが)JCOGの第III相試験のプロトコルは、通常70～100ページ程度のボリュームとなっている。

このように、とくに後期開発ではプロトコルは「多施設」の研究者に読まれることを念頭において作成する必要があるが、プロトコルはさらに「多職種」間のコミュニケーションツールでもある。プロトコルの読み手としては、参加施設の研究者(医師)、臨床研究コーディネー