

129	1. What Makes Clinical Research Ethical?	2008/7/8
130	2. Revision of Ethical Guideline for Clinical Research (2007-2008)	2008/7/8
131	3. Evaluating the Risks and Benefits of Clinical Research	2008/7/8
132	4. 包括同意——法学的倫理的検討 General Consent: Its Legal and Ethical Questions	2008/7/8
133	5. Blanket Consent in Research: What do guidelines say? What do data show	2008/7/8
研究の骨格作り（精神・神経領域）		
134	1. 臨床研究の歴史、意義、研究の定式化（2012年度版）	2014/5/22
135	2. 臨床研究のデザインと統計学（2012年度版）	2014/5/22
136	3. 文献検索のABC	2014/5/22
137	4. 横断研究をデザインする	2014/5/22
138	5. コホート研究をデザインする	2014/5/22
139	6. アウトカムの信頼性と妥当性	2014/5/22
140	7. 患者立脚型アウトカムの測定：主観的尺度の開発と検証	2014/5/22
研究の実践に必要な知識（精神・神経領域）		
141	1. 臨床試験の目的設定を考える	2014/6/19
142	2. 臨床試験の実践作法	2014/6/19
143	3. 精神科臨床試験の実践と課題	2014/6/19
144	4. 臨床研究の品質管理と品質保証	2014/6/19
145	5. 嚥下障害患者の診療が臨床研究になるまで	2014/6/19
146	6. 大規模データベースを利用したがん疫学研究	2014/6/19
研究成果の公表（精神・神経領域）		
147	1. 臨床研究論文の書き方	2014/7/17
148	2. 効果的なプレゼンテーション	2014/7/17
149	3. 色覚の多様性とカラーユニバーサルデザイン	2014/7/17
150	4. 研究者のための契約と知的財産：産学連携を行う上での留意点	2014/7/17
151	5. エビデンスの名のもとに行われる悪行の数々（偽エビデンスにだまされないために）	2014/7/17
エキスパートに学ぶ①：国内編（精神・神経領域）		
152	1. 精神保健の疫学研究：その楽しみと広がり	2014/7/31
153	2. 社会への橋渡しとしての疫学研究	2014/7/31

154	3. 公共性の自覚と臨床・研究・教育の融合	2014/7/31
155	4. 脳卒中領域における臨床研究：神経超音波と脳卒中診療体制の構築	2014/7/31
156	5. Evidence-Based Diagnostics：診断に関する研究の作り方	2014/7/31
エキスパートに学ぶ②：国際編（精神・神経領域）		
157	1. Diagnostic Entities in Depression. Modelling and Managing.	2014/7/31
158	2. The Contemporary Assessment and Measurement of Depression	2014/7/31
159	3. Seafood Deficient Diets: Neurodevelopmental and Psychiatric Risks	2014/7/31
160	4. Mind-Body Interface: Omega-3 fatty acids in physical symptoms in depression	2014/7/31
161	5. Mapping Depression Circuits	2014/7/31

II. 研究成果に関する一覧表

研究成果の刊行に関する一覧表(刊行リスト)

- 1) 山中 竹春, 山本 精一郎. 臨床試験の潮流 バイオマーカーを用いたがん第 II 相試験のデザイン. The Liver Cancer Journal. 2012. 4 巻 2 号. 108-113. メディカルレビュー社
- 2) 藤原康弘. 皆保険制度の維持と未来型医療の実現の調和 一薬事承認と保健導入の不幸な強直的なカップリングの呪縛からの解放に向けてー. 月刊基金 4 月号 p2-4 2012
- 3) 藤原康弘. 抗がん剤の臨床試験. 最新医学 67(8):1894-1898, 2012 (8 月号)
- 4) 藤原康弘. 腫瘍を対象とした臨床試験. 門脇孝 永井良三 総編集 内科学 pp324-326 西村書店, 2012 年(分担執筆)
- 5) 福田治彦: JCOG (Japan Clinical Oncology Group) の現状と課題. 薬理と治療, 2012 40: S90-93
- 6) 山中竹春. 維持療法に関する臨床試験の結果を解釈するうえで見抜くべきポイント. 日本胸部臨床. 2012;71(10):994-1005. 克誠堂出版
- 7) 山中竹春, 吉野孝之. 結腸がん Oncotype DX の現状. がん分子標的治療 2012;10 巻 2 号: 133-139 メディカルレビュー社
- 8) 福田治彦, 山本精一郎. 倫理性の担保と品質管理. ベッドサイド泌尿器科学 改訂第 4 版. 2013;5:578-81.
- 9) 吉村健一, 山本精一郎. 3. 癌の疫学 A. 疫学研究方法論. ベッドサイド泌尿器科学. 改訂第 4 版. 吉田修監修. 南江堂. 東京. 2013
- 10) 上野秀樹, 山本精一郎. 臨床試験について考える. Clinical Trials in Japan 日本の臨床試験-難治性がん克服に向けてステークホルダーからの提言-. 特定非営利活動法人パンキャンジャパン(東京). 2013.
- 11) 福田治彦. がん臨床試験における QOL 評価の問題点. 「腫瘍内科」. 科学評論社. 2013;12(4) 440-9.
- 12) 江場淳子, 中村健一, 柴田大朗, 福田治彦. Immune Related Response Criteria (irRC) ー背景, 定義, 問題点, JCOG はどう考える?. 「腫瘍内科」. 科学評論社. 2013; 12(4): 372-81.
- 13) 中村健一, 水澤純基, 柴田大朗, 福田治彦. PFS は第 III 相試験の primary endpoint となりうるか? ー知っておくべき考え方のフレームワーク. 「腫瘍内科」. 科学評論社. 2013;12(4):401-9.
- 14) 山上須賀. CRC による情報提供の重要性. 第 13 章臨床試験の情報提供とコミュニケーション. がん臨床試験テキストブック 考え方から実践まで. 大橋靖雄他責任編集. 編集: 公益財団法人パブリックヘルスリサーチセンター がん

- 臨床研究支援事業(CSPOR)教育研修小委員会. 2013年10月15日発行.
154-8.
- 15) Tanaka N, Kawasaki Y, Kurosawa T, Toyooka L, Mieno MN, Ohotsu H, Matsuoka J. A Survey on biostatistical consulting at Japanese medical institutes. *Jpn Pharmacol Ther.* 2014. 42 Suppl. 1 s33-s44.
 - 16) 福田治彦. 【IV. 大腸癌診療のための Q&A】Q17. 臨床試験・治験とは?. *ガイドラインサポートハンドブック大腸癌* 2014年版. 356-7.2015.
 - 17) 中村健一, 柴田大朗, 福田治彦. 【免疫療法の逆襲を現実化した免疫 checkpoint の修飾】免疫療法の臨床評価に関する問題点. *腫瘍内科* 2014;14(5):480-6.
 - 18) Kawano Y, Okamoto I, Fukuda H, Ohe Y, Nakamura S, Nakagawa K, Hotta K, Kiura K, Takiguchi Y, Saka H, Okamoto H, Takayama K, Semba H, Kobayashi K, Kenmotsu H, Tsuboi M, Yamamoto N, Nukiwa T, Nakanishi Y. Current status and future perspectives of cooperative study groups for lung cancer in Japan. *Respiratory investigation* 2014;52(6):339-47.
 - 19) 中村健一, 福田治彦. 【血液腫瘍に対する臨床試験の infrastructure-現状の問題点と今後の方向性】JCOGにおける研究プロセスの効率化とアクティビティ向上のための取り組み. *血液内科* 2014;69(1):19-25.
 - 20) 金戸啓介, 高島淳生, 中村健一, 福田治彦, 松村保広. 【血液腫瘍に対する臨床試験の infrastructure-現状の問題点と今後の方向性】がん多施設共同臨床試験とバイオバンク 現状の問題点と将来. *血液内科* 2014;69(1):60-5.
 - 21) 中島章博, 小森哲志, 神浦俊文, 杉谷康雄, 木村友美, 福田治彦. ベネフィット・リスク評価の実際 研究者主導のがん臨床試験の事例紹介. *医薬品医療機器レギュラトリーサイエンス* 2014;45(7):572-9.
 - 22) 中村健一, 水澤純基, 柴田大朗, 福田治彦. 【最新がん薬物療法学-がん薬物療法の最新知見-】抗悪性腫瘍薬の臨床試験と開発状況 分子標的薬の臨床試験の現状と課題 OSかPFSか(臨床試験の primary endpoint). *日本臨床* 2014;72(増刊2 最新がん薬物療法学):637-40.
 - 23) 松岡豊: 臨床研究推進に貢献する臨床研究者育成. *精神科* 27(2), 2015年8月(印刷中)

Ⅲ. 研究成果の刊行物・別刷

Special Articles

臨床試験の潮流

② バイオマーカーを用いた がん第Ⅱ相試験のデザイン

国立がん研究センター東病院臨床開発センター室長
山中 竹春

国立がん研究センターがん対策情報センターがん情報提供研究部室長
山本精一郎

従来の殺細胞薬に対する臨床試験では比較的大きな集団が試験対象とされてきたが、分子標的薬の臨床導入が一般化するなか、適切な投与対象を意識しながら臨床試験を計画することの重要性が高まっている。本稿では、バイオマーカーの利用を前提としたがん第Ⅱ相試験のデザインについて概説する。

Key words ▶▶▶

- 分子標的薬
- バイオマーカー
- 予測マーカー

- 第Ⅱ相試験
- 臨床試験デザイン

はじめに

本稿ではバイオマーカーを組み込んだ臨床試験のうち、特に第Ⅱ相試験のデザインについて概説する。第Ⅲ相試験のデザインについては本稿では触れないが、米国国立がん研究所の生物統計家による総説¹⁾²⁾、邦文では著者らによる解説³⁾⁴⁾があるのでそちらを参照していただきたい。

■ エンリッチメントデザイン

本稿におけるバイオマーカー(以下、単にマーカーと呼ぶ)とは、治療介入の前に計測されたベースライン時のマーカーを意味するものとする。一般に臨床試験でマーカーの利用を考える際には、予後(prognostic)マーカーと予測(predictive)マーカーの区別が重要である⁵⁾。予後マーカーとは治療の有無あるいは治療の種類に関係な

②バイオマーカーを用いたがん第Ⅱ相試験のデザイン

く、疾患予後に関する情報を与える指標である。一方、予測マーカーとは特定の治療法の効果予測に関する情報を与える指標であり、簡単に言えば、その治療法に反応する患者集団を同定する。適切な予測マーカーを用いて臨床試験へ登録する患者を限定(エンリッチメント(enrichment)と呼ばれる)すれば、治療開発の成功確率は格段に上昇することが期待される。たとえば、転移性乳癌におけるトラスツズマブの第Ⅱ相試験は前臨床試験の結果に基づいてHER2過剰発現の症例のみを対象にシングルアームで実施され、高い治療効果が確認された⁶⁷⁾。転移性乳癌でのHER2陽性の割合は20~30%程度なので、エンリッチメントせずに転移性乳癌全体を対象としていた場合、トラスツズマブの効果を見逃してしまった可能性がある。エンリッチメントデザインにおける必要症例数(予測マーカーによるスクリーニングが行われる症例数)がエンリッチメントなしの通常のデザインに比べてどれだけ大きくなるかは、陽性“(+)”症例の全体に占める割合、その(+群)における治療効果の大きさ、といった要素に依存して決まる⁶⁸⁾。

エンリッチメントデザインの実施にあたっては薬剤活性と標的分子の関連に対する一定の理解が必要不可欠となる。誤った生物学的パスウェイの理解に基づいたマーカーを利用してエンリッチメントを行うと、成功確率が高まらないばかりか、対象を限定することによる集積期間の延長ならびにそれに伴う甚大なコスト増を招く羽目になってし

まう。また、予測マーカーが陰性“(−)”と診断した集団はそもそも試験に組み入れられないので、(−)群に治療効果があってもそれを見出すことは不可能である。

しばしば見過ごされている点であるが、シングルアームのエンリッチメント試験の閾値の設定には注意が必要である⁶⁹⁾。シングルアーム第Ⅱ相試験の症例数を算出する際は閾値が設定される。最低限その値を超えることが示されなければ治療開発を続ける意義に乏しいと判断できる値が閾値であり、通常は標準治療の成績を参考に設定されることが多い。ここでは全体の(つまり、unselectedな)患者集団における標準治療の奏効率を10%とし、この値を閾値に設定する。一般にエンリッチメント試験はマーカーが予測マーカーであることを期待して実施されるが、閾値を決定する際はそのマーカーが予後マーカーであるかどうかの区別も重要になる。たとえば、エンリッチされた(+群)はもともと予後良好で、あらゆる治療に反応しやすいかもしれない。そこで(+群)は標準治療でも奏効率20%が得られると仮定しよう。この場合、新薬と標準治療の間に差がないときに(すなわち、(+群)での新薬の奏効率も同じ20%であるときに)誤って新薬が有望であると結論してしまう偽陽性の確率*は格段に上昇する。本来、(+群)における標準治療の成績20%がクリアすべき基準であるのに、下駄を履かせた10%を基準としているのだから当然である。逆に治療に反応しにくい集団へエンリッチするケースでは、全体の患者集団におけ

る標準治療の奏効率を閾値として採用すると必要以上に厳しい基準となる。この例からわかるように、利用するマーカーが予後マーカーであるか否か、予後マーカーであるとするれば閾値の調整が求められる。ただし、(+群)における奏効率のヒストリカルデータが得られていることはあまり多くはない。そのようなケースでは標準治療を施行された患者の保存検体に当該マーカーに関するアッセイを実施し、(+群)における標準治療の奏効率を見積もるなどのステップが必要になる。

Ⅳ バイオマーカーを組み込んだ第Ⅱ相試験

エンリッチメントデザインを採用できる程度まで薬剤作用機序に関する理解が進んでいけばよいが、そういったケースはあまり一般的でないかもしれない。その場合、分子標的薬の第Ⅱ相試験にマーカーを組み込むことの主な目的は、特定の対象に絞って治療開発を実施するというよりは、今後、a)広い患者集団を対象に薬剤開発を継続する、あるいはb)マーカーによる患者選別を実施したうえで薬剤開発を行う、いずれの方策を進めるかを決定することにあるといえる。以下ではそのような目的に沿った試験デザインについて述べる。誌面の都合により2つに絞って説明するが、いずれのデザインもマーカーに多少なりとも目処がついて、奏効率などの短期指標をエンドポイントとする場合を想定している。

*：正確には、(+群)における新薬の真の奏効率か20%のとき、データから推定される信頼区間の下限が閾値10%を上回り、新薬のほうが有望であると結論してしまう偽陽性の確率。

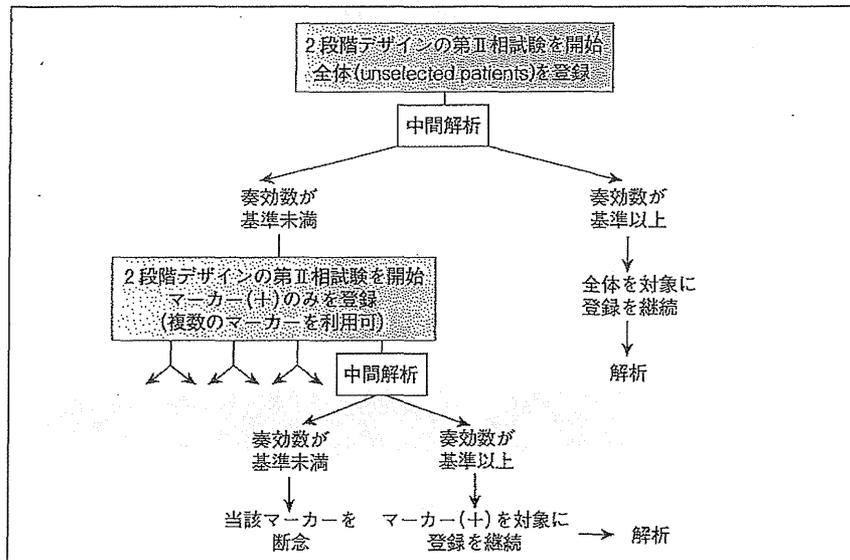


図1. タンデム・デザイン
文献10)より引用・改変

1. タンデム・デザイン

全体(unselectedな集団：たとえば、進行期〇〇癌といった広い対象)に対するシングルアーム試験を実施し、見込みがないと判断した時点でマーカーを利用したエンリッチメント試験へと移行する¹⁰⁾。2つの対象(全体, (+)群)に対する試験を“タンデム”に実施するデザインといえる。手順の概略は次の通りである(図1)。

- 最初はマーカーによる選択を実施せずに全体(unselected)を対象とした登録を行う。
- 途中、中間解析を実施して基準以上の奏効数が確認できれば、全体を対象とした登録を継続する。奏効数が基準に満たなければ、(+)群に制限した登録へと移行し、(-)群の登録を中止する。
- (+)群のみの登録へ移行した後も、2段階デザインに従って試

験の無効中止を決定する。すなわち、(+)群に一定数の症例が登録された時点で中間解析を実施して基準以上の奏効数が確認できれば、(+)群を対象とした登録を継続する。奏効数が基準に満たなければ、(+)群でも効果はないと結論する。

全体を対象に中間解析を実施する時点(シングルアーム第II相試験では20例未満ないし30例未満であることが一般的)では、(+)群に対する結論を引き出せるほど、(+)群の登録数が多くないことを想定したデザインである。症例数ならびに中間解析(全体, (+)群に対する)の無効中止基準については統計的に決定される。手順はやや複雑になるが、2つ以上のマーカーを同時に組み込むことも可能である。

2. アダプティブ・ランダム化ベイズ流デザイン

近年、ベイズ流アプローチを採用した臨床試験デザインへの関心が高まっている。通常の統計的アプローチは「データに語る」という立場であり、臨床的興味指標(たとえば、奏効率)は試験で実際に観測されたデータから推定される。事前情報は利用しない。一方、ヒストリカルデータなどによって奏効率に関する事前情報(不確実性の程度はともかくとして)が外部から得られている場合、この事前情報+観測データを組み合わせて奏効率を推測するというベイズ流の立場がある。日常的にもわれわれは、事前情報(経験に基づく予想、個人の信条など)に観測データ(現実起こったこと)を組み合わせた推測を頻繁に行っているが、ベイズ流アプローチはまさにそのような思考過程を具現

②バイオマーカーを用いたがん第II相試験のデザイン

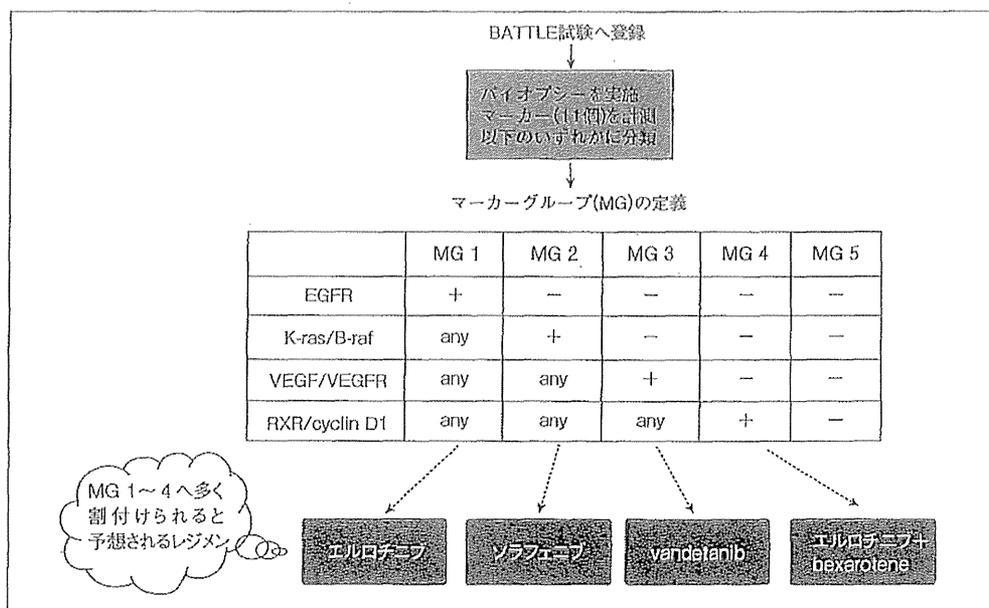


図 2. アダプティブ・ランダム化(BATTLE試験)

文献¹⁴⁾より引用・改変

している(ベイズ流アプローチについてはBerry¹¹⁾の文献に平易な解説がある)。

「外部から得られる事前情報を試験デザインへ反映させる」という考え方を推し進めて、「試験前半に得られた情報を試験後半部分のデザインへ反映させる」ことも可能である。試験前半に得られた情報を“事前”情報として扱えばよい。たとえば、 k 個($k \geq 2$)の治療アームからなるランダム化試験において、すでに得られているデータから k 個の各群の奏効率を推定する。この推定値をもとに新規登録例の割付比率を変化させることが可能である(アダプティブ・ランダム化¹²⁾)。

BATTLE試験

第II相試験へアダプティブ・ランダム化を応用したBATTLE試験(Biomarker-integrated

Approaches of Targeted Therapy for Lung cancer Elimination)の結果が2010年に報告され、注目を浴びた¹³⁾¹⁴⁾。非小細胞肺癌の2次治療例をエルロチニブ、ソラフェニブ、vandetanib、エルロチニブ+bexaroteneの4レジメンのいずれかにランダム化し、8週時点での疾患制御率を主要評価項目とした試験である。4薬剤の作用機序に関連した11個のマーカーとして、EGFR/KRAS/BRAFの変異、EGFR/cyclin D1のコピー数増幅、VEGF/VEGFR-2/RXR受容体(α , β , γ)/cyclin D1の過剰発現が計測され、これらに関して次のようにアダプティブ・ランダム化が実施された。まず、11個のマーカーステータスをもとに、①EGFR変異 or 増幅、②K-ras and/or B-raf変異、③VEGF and/or VEGFR過剰発現、④RXR and/

or cyclin D1過剰発現の順に優先順位をつけて登録例を分類する(図2)。すなわち、EGFR変異 or 増幅のときは他のマーカーステータスにかかわらず、マーカーグループ(MG) 1へ分類する。MG 1に該当せず、かつK-ras and/or B-raf変異のときはMG 2へ分類する。MG 1, 2に該当せず、かつVEGF and/or VEGFR過剰発現のときはMG 3へ分類する。MG 1, 2, 3に該当せず、かつRXR and/or cyclin D1過剰発現のときはMG 4に分類する。以上のいずれにも該当しなければ、MG 5に分類する。5つのマーカープロファイル×4つの治療法=20通りの組み合わせが得られるが、これらに対して疾患制御率が推定される(本来は $2^{11} \times 4$ 通りの膨大な組み合わせだが、可能な組み合わせを20通りに減らす措置である)。予定登録数は約250例

で、約100例の時点までは1:1:1:1の均等なランダム化を実施、それ以降はたとえば、EGFR変異or増幅例(MG1)はエルロチニブによく反応しやすいという傾向が観測された場合、以後のMG1の症例にエルロチニブ群を高い確率で割付ける、といった割付比率の変更が行われた。仮に*n*例目に登録予定の症例がMG1に属し、かつ、すでに得られたデータからMG1における4治療の疾患制御率がそれぞれ60%、30%、20%、10%と推定されたとする。この場合、*n*例目の割付比率を0.5:0.25:0.17:0.08と設定する。割付比率のdata-orientedな変更によって、特定のマーカープロファイルをもつ症例を、プロファイル上反応しやすいと予測される治療法に高い確率で割付けることが可能になる。

アダプティブ・ランダム化は「効き目の高いと思われる治療法により多くの症例を割付けることが倫理的」という考えのもとに成立するが、一方でランダム化という操作が許容されるための前提条件である治療アーム間のclinical equipoiseが欠落しているという議論もある¹⁵⁾。また、アダプティブ・ランダム化の実施にあたってはさまざまなシナリオのもとでどのぐらいの効率性を達成できるかの統計的シミュレーションや実務レベルでの割付実施手順の確認が重要である。

おわりに

第一世代の分子標的薬(イマチニブ、トラスツズマブ、リツキシ

マブ、ゲフィニチブなど)が登場した直後は分子標的薬評価のための新しい試験デザインの必要性が盛んに議論されたが¹⁶⁾、これらの薬剤が比較的大きな成功を収めたことで、そのような議論は一旦下火になった感がある。しかし、のちに開発された分子標的薬の多くの第Ⅲ相試験がネガティブな結果に至っている事情から、第Ⅲ相試験、そして本稿で述べた第Ⅱ相試験のデザインに関する議論が再び活発化してきている。全体(unselected)の患者集団に対する分子標的薬 vs. 標準治療の第Ⅲ相試験がネガティブな結果に終わった場合、可能性としては①2治療間に本当に差がない、②分子標的薬は全体に対して有効だが、βエラーが生じた、③分子標的薬はサブグループに対して有効だが、サブグループを意識した試験デザインでなかった、などが考えられる。ネガティブな結果に終わっている理由は必ずしも明らかでないが、ゲフィチニブ¹⁷⁾、セツキシマブ¹⁸⁾、パニツムマブ¹⁹⁾などの事例からは、多くの試験で③が生じている可能性を意識させられる。

第Ⅲ相試験を成功に導くために第Ⅱ相段階からマーカーを積極的に組み込もうとするさまざまな取り組みがなされており、本稿では第Ⅱ相試験に予測マーカーをどう組み込むか、という観点から説明を行った。なお、殺細胞薬に対する奏効率をエンドポイントとした第Ⅱ相試験はこれまでシングルアームとして実施されることが多かったが、分子標的薬を評価する

うえでランダム化第Ⅱ相試験の重要性が高まっている²⁰⁾。ランダム化第Ⅱ相試験のデザインについてはRubinsteinら²¹⁾が参考になる。

References

- 1) Simon R: The use of genomics in clinical trial design. *Clin Cancer Res* **14**: 5984-5993, 2008
- 2) Freidlin B, McShane LM, Korn EL: Randomized clinical trials with biomarkers: design issues. *J Natl Cancer Inst* **102**: 152-160, 2010
- 3) 山中竹春, 山本精一郎: バイオマーカーを用いた臨床試験デザイン. *がん化学療法・分子標的治療 update*, 西條長宏, 西尾和人 編. 東京, 中外医学社, 359-363, 2009
- 4) 山本精一郎: 分子標的治療薬の評価. *日本臨床(増刊号)がん薬物療法学*. 東京, 日本臨床社, 430-437, 2009
- 5) 柴田大朗: 選択的治療を可能にする方法論 生物統計学の側面から. *Mebio Oncol* **5**: 4-9, 2008
- 6) Baselga J, Tripathy D, Mendelsohn J, et al: Phase II study of weekly intravenous recombinant humanized anti-p185HER2 monoclonal antibody in patients with HER2/neu-overexpressing metastatic breast cancer. *J Clin Oncol* **14**: 737-744, 1996
- 7) Cobleigh MA, Vogel CL, Tripathy D, et al: Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. *J Clin Oncol* **17**: 2639-2648, 1999

②バイオマーカーを用いたがん第Ⅱ相試験のデザイン

- 8) Simon R, Maitournam A : Evaluating the efficiency of targeted designs for randomized clinical trials. *Clin Cancer Res* **10** : 6759-6763, 2004
- 9) McShane LM, Hunsberger S, Adjei AA : Effective incorporation of biomarkers into phase II trials. *Clin Cancer Res* **15** : 1898-1905, 2009
- 10) Puzstai L, Anderson K, Hess KR : Pharmacogenomic predictor discovery in phase II clinical trials for breast cancer. *Clin Cancer Res* **13** : 6080-6086, 2007
- 11) Berry DA : Bayesian clinical trials. *Nat Rev Drug Discov* **5** : 27-36, 2006
- 12) Thall PF, Wathen JK : Practical Bayesian adaptive randomisation in clinical trials. *Eur J Cancer* **43** : 859-866, 2007
- 13) Kim ES, Herbst RS, Lee JJ, et al : The BATTLE trial (Biomarker-integrated Approaches of Targeted Therapy for Lung Cancer Elimination) personalizing therapy for lung cancer. AACR 101st Annual Meeting 2010 (abstr LB-1).
- 14) Zhou X, Liu S, Kim ES, et al : Bayesian adaptive design for targeted therapy development in lung cancer--a step toward personalized medicine. *Clin Trials* **5** : 181-193, 2008
- 15) Palmer CR, Rosenberger WF : Ethics and practice : alternative designs for phase II randomized clinical trials. *Control Clin Trials* **20** : 172-186, 1999
- 16) Korn EL, Arbuck SG, Pluda JM, et al : Clinical trial designs for cytostatic agents : are new approaches needed? *J Clin Oncol* **19** : 265-272, 2001
- 17) Mok TS, Wu YL, Thongprasert S, et al : Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* **361** : 947-957, 2009
- 18) Karapetis CS, Khambata-Ford S, Jonker DJ, et al : K-ras mutations and benefit from cetuximab in advanced colorectal cancer. *N Engl J Med* **359** : 1757-1765, 2008
- 19) Amado RG, Wolf M, Peeters M, et al : Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. *J Clin Oncol* **26** : 1626-1634, 2008
- 20) Ratain MJ, Sargent DJ : Optimising the design of phase II oncology trials : the importance of randomisation. *Eur J Cancer* **45** : 275-280, 2009
- 21) Rubinstein L, Crowley J, Ivy P, et al : Randomized phase II designs. *Clin Cancer Res* **15** : 1883-1890, 2009

月刊基金

4

2012 April



社会保険診療報酬支払基金
平成24事業年度一般会計事業計画を策定

特集1 審査情報提供事例(薬剤)を公表

特集2 平成24年度診療報酬改定の概要

特別寄稿

皆保険制度の維持と未来型医療の実現の調和

—薬事法承認と保険導入の不幸な強直的
カップリングの呪縛からの解放に向けて—

国立がん研究センター中央病院
乳腺科・腫瘍内科 科長 兼 副院長（経営担当）

藤原 康弘

臨床研究を進めにくい わが国の医療保険制度

個別化医療、テーラーメイド医療、ゲノム医療の時代が間近に迫っていると
言われて久しい。筆者の専門領域であ

る乳癌治療分野では個別化医療導入が
2000年代に入り急速に進んでいる。
しかし、既に個別化医療、ゲノム医療が
花開いている米国と異なり、低負担にあ
えいでいる国民皆保険制度下にあるわが
国で、このような未来型医療がスムーズ
に導入できるのか、筆者は疑問を持つて

いる。
日本以外の先進国の医療保険制度をみ
ると、将来、より優れた医療を生むため
には臨床研究（特に、介入を伴う前向き
な研究＝臨床試験）が重要であり、そ
の臨床研究を振興することにも配慮し
た医療保険制度の設計がなされている。
この制度の根幹をなすのがAccess with
evidence generation（AEG）という
概念である。一方、わが国では、従来か
ら、療養担当規則での規定に代表される
ように、研究的診療は好ましくないもの

であるとの考え方が医療保険の専門家や
支払側と呼ばれる方々の間には根強いよ
うに見える。
AEGという概念は、医療の進化には
数々の臨床試験の積み重ねが必要であ
り、臨床試験から生まれてくる科学的
根拠（エビデンスとも呼ばれる）の情
報量は徐々に増えてくるもの（evidence
escalatorと呼ばれることあり）で、新
しい医療技術や医薬品への医療保険（公
的保険も含め）によるカバーと患者負担
の比率は、徐々に増加してくる科学的根

拠の量に応じて決めるべきものであるという概念である。臨床開発の初期段階ではエビデンスの量が少ないため、患者負担あるいは研究費負担（アカデミアからのものもあれば、企業からのものである場合もある）で全額諸経費をカバーし、その後、研究が進展しエビデンス量が増加し、その根拠の確からしさも増えるにしたがって医療保険からの負担を増やし、患者負担・研究費負担を減らすことになる。昨年末、英国ランセット誌で日本の国民皆保険制度の特集が組まれ、改めてわが国の制度の良さがうたわれたところであり、筆者も海外の友人に胸を張れる出来事であったが、将来の医療を支える臨床研究を考慮した医療保険制度、AEGの仕組みが組み込まれてこなかった欠陥への言及が論文著者から無かったことは残念である。

今後、個別化医療の発展を日本で後押しするには当該医療を評価する臨床試験の振興が必須であり、そのためには平成20年度に導入された日本版AEGと呼べる高度医療評価制度（第3項先進医療）と先進医療制度をこれまで以上に重視する姿勢が大事であると筆者は考える。

高額薬剤・医療技術への対応

個別化医療時代、医薬品や医療技術の費用は、これまでとはケタ違いのものに

なると予想される。たとえば、昨年、米国で承認された前立腺癌患者を対象とする癌ワクチンProvenge（患者ごとに細胞を調整する個別化細胞療法）の細胞療法部分だけの価格は、人当たり9万3千ドル（約800万円）であり、国を平らげての大激論ののち、Medicareでの保険償還を昨年7月CMSは認めている。また、EJで、昨年承認された骨肉腫患者を対象とする癌ワクチンMepactは細胞療法法の価格が1万4千ポンド（約1500万円）であり、英国NICEはpatient access scheme（PAS）による値引きを条件にNHSでの保険償還を認める判断を昨年の10月に行っている。現在、わが国ではマスコミが音頭を取って癌ワクチン療法や再生医療が身近な診療をすぐにでも塗り替えるような喧伝がなされているが、前述した外国の先例を見ても明らかのように、未来型医療は個別化されてはいるが高価なのである。それを財政の破綻しかかっている現行のわが国の国民皆保険制度に導入し続けることは誰の目からみても不可能である。患者負担だけに依存する混合診療の導入は、まさに、このような個別化医療の導入時に、受けられる医療に貧富の差をもたらすことになると思われ、高額療養費制度は患者さん個人を守ってくれるが、国民皆保険制度本体の崩壊を招く恐れを

危惧する。全国民に必須の通常の医療については、より負担を充実し強固堅牢にした公的保険でカバーしつつ、高額な未承認医療は、国民の経済力による極端な格差の排除と価格に見合う医療上の効果に関する科学的評価の仕組みを組み込んだ新たな保険制度を検討すべきである。

適応外薬問題

欧米ではさまざまな疾患に広く使われている医薬品が、日本では強直化した薬事制度のため、一部の疾患にしか使えないという非科学的かつ患者不在の「適応外使用問題」は、抜本的な改革が必要である。がん診療の領域で顕著なことであるが、欧米では医薬品が日常の保険診療の中で使用できるからといって、規制当局がその適応に対して薬事法（各国で日本の薬事法にあたる法律）に基づき承認をいちいち与えているわけではない。甚で「世界の標準治療が日本でのみ承認されておらず使えない」と喧伝されるが、実は欧米においても当該効能の薬事法上の承認は無いという例には枚挙にいとまがない。当センターの研究者の調査では、米国で保険償還されている抗癌剤の95.1%適応のうちFDAが承認してい

るのは33%に過ぎない。米国では、薬事法の承認は物質として品質を担保することにより品質が置かれ、最初に承認した適応以外については、以降のエビデンスの集積を支払目が評価しつつ、償還の可否をダイナミックに判断している。このような薬事法に基づく承認と保険償還を別立てで考えて運用する仕組み（米国のコンベンティアと呼ばれる制度が有名）英国のNHSでも同様の仕組みがある）をわが国でもしつかり整備する必要がある。すなわち、適応外薬については、信頼性の高い臨床試験結果が複数存在していると第三者機関（米国では社会保険法の中で当該機関を指定している）が判断した場合、薬事法の承認が無くとも、その使い方は保険適用可とできる仕組みを整備すべきである。日本でも再審査期間を満了した医薬品については、いわゆる55年通知の規定が適用され、社会保険診療報酬支払基金の審査情報提供検討委員会が「保険診療の中で原則として認め得る適応外薬」の品目リストを数年に一度開示しているが、今後はこの判断根拠の公開を一層進めることにより、科学的信頼性を高める方策を検討すべきである。このため、この制度を再審査期間内の医薬品へも対象を拡充し、保険償還可能とした判断根拠を医薬食品局が所管している「医療上の必要性の高い未承認

薬・適応外薬検討会議」のワーキンググループの出す報告書並し公表すべきだと思ふ。評価は日本医学会加盟の学会に任せれば良い。これにより、適応外使用問題の大半は解決に向かうと思われる。

コンパニオン診断薬の保険診療への導入を巡る障壁

最後に個別化医療実現のカギを握っている遺伝子診断技術や診断キット（体外診断薬と日本では薬のように呼ばれている）を巡る問題を紹介したい。投薬や技術実施の可否判断に重要な役割を果たす体外診断薬をCompanion in vitro diagnostic devices（コンパニオン・ダイアグノスティックスあるいはコンパニオン診断薬）と呼ぶ。このコンパニオン診断薬、特に遺伝子診断技術やキットの保険診療上の取扱いが日本では立ち後れている。昨年8月、米国では肺癌と悪性黒色腫の抗癌剤が相次いで承認され、その添付文書の効能・効果欄にFDAが同時に承認したコンパニオン遺伝子診断キットを使用して投薬する旨が記載されている。今後、米国では個別化医療の領域の薬剤（分子標的薬）の承認にあたっては、このような方式が用いられると思われるが、わが国でこれに倣った場合、体外診断薬の薬価収載スケジュールを新薬発売に合わせる措置がまず必要になる。

さらに困った問題なのが、Oncotype Dx（乳癌の術後の抗癌剤治療の要否を判定する遺伝子診断で、既に世界中の日常診療で使われている）のようなキットが存在しない遺伝子診断技術（事実、この診断法にはFDA承認が無く、CLIAという臨床検査機関の検査品質を保証する法律で質の担保されたカリフォルニアにある一か所の検査機関で診断が行われ、保険償還も米国を中心に広く認められている）を国民皆保険制度に組み込む手段を我々が持ち合わせていないことである。日本では一回当たり40万円弱の検査費用を患者が自費診療下で全額負担しているのが現状である。今後、出てくる遺伝子診断キットや手法は高価であり、現行の診療報酬点数では低すぎるため、メーカーは薬事承認・薬価収載をためらい、高価格の製品を公的保険外で使用せざるを得なくなることは不可避である。この他にも遺伝子診断を行う検査機関の認証・評価制度が無いことも将来に向けた懸念である。個別化医療の開花に向けて早急な整備が必要である。

以上みてきたように、個別化医療、ゲノム医療の本格導入を目前にしてわが国の保険医療制度や薬事制度の見直しは待たなしの状況であることを訴えて、稿を終えたい。

臨床研究のススム

16

抗がん剤の臨床試験

藤原 康弘 国立がん研究センター 企画戦略局長
同 中央病院 乳腺・腫瘍内科 科長抗がん剤の臨床試験の計画と実施、
成果公表までの流れと種類

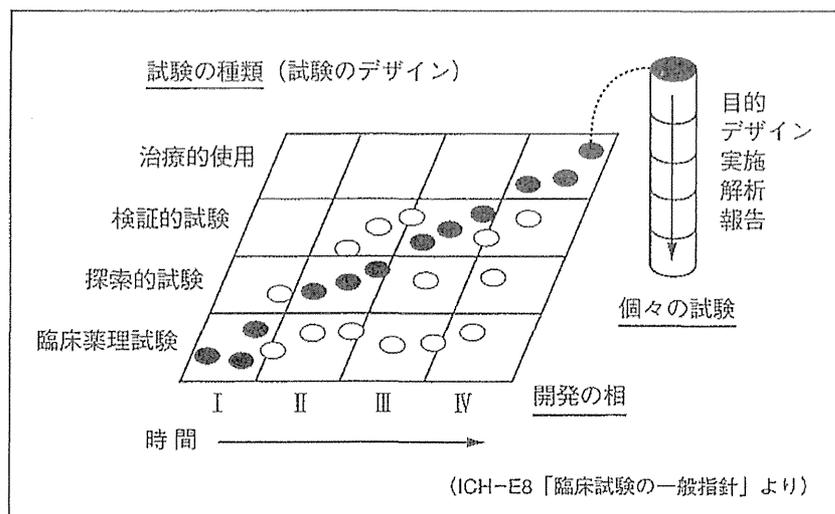
臨床試験は、解決したい臨床的課題を目的として明確にしたうえで、研究計画書（プロトコール）や同意説明文書、収集するデータと収集方法、まとめ方を、生物統計家、データマネージャー（DM）、臨床研究コーディネーター（CRC）と相談しながら作成することから始まる。さらに準備段階では、実施体制〔研究協力者、検査手順や外部委員会（効果・安全性評価委員会、病理中央判定委員会などが抗がん剤の臨床試験、特に第Ⅲ相試験では必要となることが多い）〕の整備も必要である。準備の最終段階で、UMIN（外国への日本のプレゼンスを示す意味合いから、米国の clinicaltrials.gov への登録も望ましい）への臨床試験登録と倫理審査委員会での審査・承認を経て、実施の段階に移る。実施段階では、インフォームド・コンセント（同意説明文書を用いて試験の意義や試験の持つ利害への理解を踏まえての文書同意）、被験者登録、診療のみならず、発生する有害事象への対応、得られる各種データ（画像診断、臨床検査値）の品質チェック（症例報告書やデータベースに正確にカルテの内容が集積されているかを確認する作業：モニタリングと呼ばれる）をDM、CRCの協力のもとでリアルタイムに行うことが必要である。そして、臨床試験の最終段階は、解析と結果の公表である。つまり収集したデータのクリーニングを行ったうえで解析

を行い、結果の総括をしたうえで学会や学術論文として公表することで、一連の流れは終了する。

臨床試験は目的・実施時期により、臨床薬理試験、探索的試験、検証的試験、治療的使用と大きく区分される。一方、医薬品の臨床開発は4つの逐次的な相（第Ⅰ相から始まり第Ⅳ相まで）を経て進んでいくと考えると理解しやすいため、第Ⅰ相試験、第Ⅱ相試験、第Ⅲ相試験などと大別されることが多い。しかし、相という概念は一種の記述表現であり、この開発の相と臨床試験の種類との関係は、ICH-E8「臨床試験の一般指針」（平成10年4月21日 厚生省医薬安全局審査管理課長通知 医薬審第380号、http://www.pmda.go.jp/ich/e/e8_98_4_21.pdf）の図1で示すように、必ずしも1対1対応しない関係にあると理解すべきである。

抗がん剤の臨床試験は、抗がん剤の製造販売承認取得のために実施される治験のみならず、前述の「治療的使用」に分類される種類の研究者主導の臨床試験（医薬品、医療機器をおのおの単独で用いるのみならず、手術や放射線照射、免疫療法とのさまざまな組み合わせがあるので、それを検討する集学的治療の臨床試験）が広範に実施され、エビデンスが積み重なり、診療内容が進化していく。日々の診療は、過去の幾多の治験と臨床試験の積み重ねによって成り立っていることを強く認識しておかねばならない。

図1 開発の相 (phase) と臨床試験の種類 (type)



抗がん剤の臨床試験の種類

抗がん剤を用いた臨床試験は、ローマ数字の大文字でI, II, III (IVは製造販売承認後の企業主導の臨床試験を意味すること)の3段階で、開発の相(「フェーズ」と呼ぶことも多い)が進んでいくと考えると理解しやすい。高血圧や糖尿病の医薬品開発における第I相試験では健康成人を対象として行われるのとは異なり、抗がん剤の第I相試験はがん患者を対象として行われ、主として安全性が15~30人程度の規模で検討される。第II相試験では安全性に加えて腫瘍縮小効果などの有効性を100人未満の規模で検討され、第III相試験では延命効果などを中心とした臨床的有用性が100から数千人規模で検討される。安全性と腫瘍縮小効果については、世界共通の評価指標が用いられている。安全性に関しては、米国国立がん研究所が作成したCommon Terminology Criteria for Adverse Events [CTCAE (有害事象共通用語規準) v4.0. 日本語訳 JCOGについては <http://www.jcog.jp/doctor/tool/ctcae4.html> を参照] という評価規準に従い、被験薬との因果関係を問わず発生した好ましくないイベントや臨床検査値異常(「有害事象」と呼ばれる)の内容および重症度が評価される。有害事象のうち、被験薬との因果関係がある、または否定できないものを、副

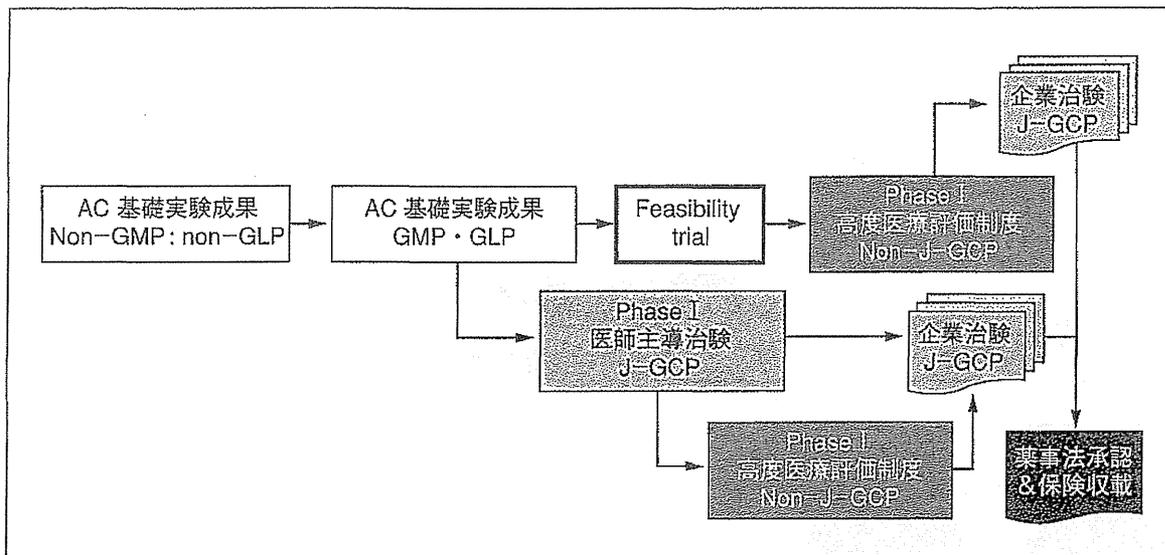
作用(薬物有害反応)と呼ぶ。有効性については、固形がんの腫瘍縮小効果判定では Response Evaluation Criteria In Solid Tumors (RECIST) の使用が一般的である(RECIST 規準や第III相試験における各種のエンドポイントの定義は JCOG プロトコルマニュアル v2.2 を参照。 <http://www.jcog.jp/doctor/tool/manual.html>)。一方、白血病などの血液腫瘍においては RECIST 以外の判定規準も用いられ、complete remission (CR) のみを有効とすることがある。

臨床試験では、年齢や臓器機能、全身状態[ECOG の performance status (PS)] などについて、症例選択基準や除外基準で制限を設けて実施されるため、試験結果を日常診療に応用するに際しては、一般化可能性に注意を払う必要がある。すなわち、全身状態の悪い患者(PS 3, 4)、高齢者や合併症のある患者、併用薬の多い患者などへの試験結果の応用は慎重になる必要がある。

1. 第I相試験

製剤の安定性や品質が担保され、非臨床試験成績(薬効薬理試験、毒性試験、動物での薬物動態試験など)が終了して初めて、ヒトに試験薬が投与される段階となる。非臨床試験で観察された事象に基づき、用量に依存した試験薬の

図2 アカデミア (AC) 発の基礎研究成果から承認まで — 早期開発に焦点を当てて —



安全性を検討するのが第Ⅰ相試験の主な目的である。抗がん剤の第Ⅰ相試験では、標準的治療に無効となった、あるいは標準的治療の存在しないがん患者が対象となる。初回投与量は、マウスに対する10%致死量(LD₁₀値)の1/10量が一般的に使用される。第Ⅰ相試験では、薬物動態学的検討(試験薬の体内での吸収、分布、代謝、排泄に関する諸性質)ならびに薬力学的検討(薬物動態パラメーターと毒性出現との関係)を行い、適切な投与量および投与間隔を決めるための参考とする。

2. 第Ⅱ相試験

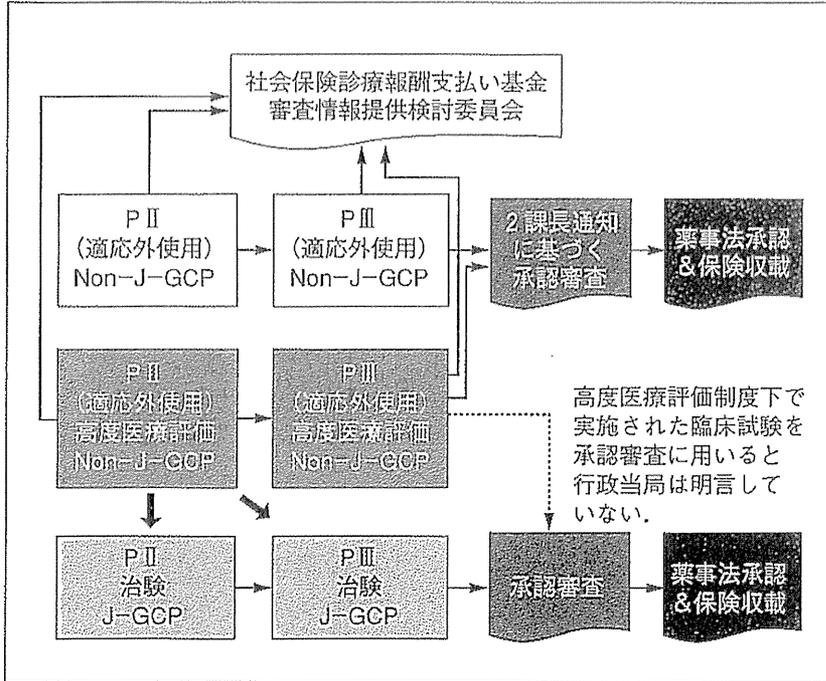
第Ⅱ相試験の主たる目的は、特定のがん腫に対する抗腫瘍効果の有無の評価である。第Ⅰ相試験により決定された用法・用量に従って、対象とするがん腫における臨床的意義のある治療効果[腫瘍縮小効果(RECIST規準に基づく奏効割合。奏効率とも言う)や生存率の上昇、臨床的意義が公知である腫瘍マーカーの変化、QOLの改善など]と安全性について、比較対照群を置かず(シングルアームで)評価する。例外的に対照群を置く、ランダム化第Ⅱ相試験というものもある。試験計画は、治療効果の評価するために適切なデザインで、かつ、科学的

に十分な精度で評価を行うことが可能な症例数となるよう、生物統計学的検討に基づいて立案することが重要である。

3. 第Ⅲ相試験

第Ⅲ相試験は、その時点で標準と考えられる治療法と新治療法とを比較することで、より優れた治療法を確立するために行われる臨床試験である。通常は2群比較で行われ、被験者は各群にランダムに割り付けられる。主な評価指標(エンドポイント)は、全生存期間(overall survival: OS)、無増悪生存期間(progression free survival: PFS)などである。QOLも評価使用として用いられる場合もあるが、解釈や解析が難しくなることが多く、主要な評価指標とされることは少ない。第Ⅱ相試験同様、デザイン、症例数や試験期間の設定には生物統計家の参画が必須であり、試験データの品質管理・品質保証を担うデータセンターを設けるとともに、試験の安全かつ倫理的な進捗に、試験実施に直接関与しない第三者の専門家によって構成される効果安全性評価委員会(独立データモニタリング委員会とも呼ばれる)による監視も重要である。

図3 アカデミア (AC) 発の臨床試験成績の承認への利用
—後期開発に焦点を当てて—



がん臨床試験の成果の応用と保険診療下での実施における留意点

我が国において未承認の抗がん剤を用いて臨床試験を実施する際には、治験の枠組みで実施することが基本である。一方、適応外の抗がん剤を用いて研究者が主導の臨床試験を実施する場合には、高度医療評価制度を利用すると、混合診療問題を気にすることなく試験を実施できるメリットがある。ただし、高度医療への申請に当たっては、自施設で数例程度の当該抗がん剤を用いた使用経験が要求される。その場合には、保険診療下での実施は原則認められないという点には注意が必要である。

さて、この臨床試験実施者が特に頭を悩ませる国民皆保険制度・保険診療と臨床試験との兼ね合いの問題を考察したい。我が国では、保険診療下で臨床試験（研究的診療）を実施するには評価療養と認定されている必要がある。すなわち、薬事法令に則って行う企業治験、医師主導治験、2008年4月から制度導入された先進医療（第2項）および高度医療〔先進医療（第

3項〕] においてのみ研究的診療が可能なのである。日本以外の欧米各国では、診療に係るエビデンスは早期臨床開発研究の段階から種々の臨床試験を経て保険診療となるまで徐々に成熟していくものであり、日本のように薬事承認を境に研究から診療へとスイッチがいきなり入る（保険償還が薬事承認の範囲にすべて可能となる）ことが可能な社会ではない。研究費のみで実施する臨床試験に始まり、健康保険からの支払いの割合を徐々に増やし、患者自己負担の割合を徐々に減らしていく access with evidence generation (AEG) [access with evidence development (AED) と称される場合もある] という概念で医療の進歩を支えようという考え方が、世界共通の認識である。我が国でも、AEG の概念に近い先進医療制度の活用に、今後努めるべきであると筆者は考えており、図2、3に示すようなフローでの各種臨床試験の薬事承認および保険診療への活用が望ましいと考える。

臨床試験は医療を巡るさまざまな制度〔薬事承認審査、オーファンドラッグ制度、コンパッションエートユース、保険診療、薬価、health