

図3 JCOG バイオバンクでの試料入庫

③ 試料の処理

試料搬送・処理業者は、血漿分離を行い血球成分からDNAを抽出し、血漿およびDNAを専用チューブに分注する。

専用チューブは、JCOG バイオバンク番号が付された2次元バーコードラベルで管理する。

JCOG バイオバンク番号は、各試料識別のため、JCOG 試験番号、JCOG 患者登録番号、試料の種類、採取時期、試料のシリアル番号を組み合わせた番号である。

(2) 試料の保管

JCOG バイオバンクでは、試料をディープフリーザー(−80℃)で保管する。試料は、JCOG バイオバンク番号、および専用チューブ底面に付された2次元バーコードで識別、管理する(図3)。

(3) 試料の提供(予定)

JCOG 内での審査機構および関連施設のIRBで承認された試料解析研究に対して、バイオバンクで保管している試料を提供する。

① 試料解析研究が「ヒトゲノム・遺伝子解析研究に関する倫理指針」の適用となる場合

JCOG バイオバンクは試料を、JCOG データセンターは必要な診療情報を、JCOG バイオバンク個人情報管理者(国立がん研究センター東病院長)へ提出する。JCOG バイオバンク個人情報管理者は、JCOG バイオバンク番号の連結可能匿名化を行う(JCOG バイオバンク番号からJCOG バイオバンク匿名化番号へ置き換える)。JCOG バイオバンク個人情報管理者は匿名化された試料および診療情報を試料解析研究実施機関へ提供する。

② 試料解析研究が「疫学研究に関する倫理指針」の適用となる場合

JCOG バイオバンクは、試料解析研究実施施設へ必要な試料を提供する。試料の識別にはJCOG 登録番号を用いる。JCOG データセンターは、必要な診療情報を試料解析研究実施施設に提供する。患者の識別番号としてはJCOG 登録番号を用いる。

4. 試料の保管期間

試料は、試料解析研究により使い切られるまで保管される予定であり、保管期間の期限は定めていない。

5. 参加研究機関

JCOG 試験参加施設のうち、JCOG バイオバンクへの参加に関して医療機関の承認が得られた施設が参加している。

JCOG バイオバンク現状の問題点と将来

1. 現状の問題点

JCOG バイオバンクの試料は、質の高い診療情報を有していることから、それらと分子生物学的異常との真の関連を正しく検出できる可能性が高く、将来JCOG バイオバンクの試料を用いた試料解析研究の成果が期待される。

しかし、より試料価値を高め、かつより円滑な運用を行うためには、(1)～(3)にあげるような課題の解決が必要である。

(1) 試料の種類、採取回数の拡充、インフラ整備

2013年7月より運用を開始したJCOG バイオバンクは、全JCOG 試験が参加しやすくすること

を優先し、実施計画書では、

- ・血液のみを採取(血漿, DNAのみを収集)
- ・採血回数: 1ポイント(原則として治療開始前)のみ

と規定している。しかし、これらの試料のみでは将来実施できる試料解析研究に限られるため、JCOGの研究者からは以下のような要望が出されている。

- ・採血回数を複数回可能としてほしい。
- ・血液から抽出したDNAのみでなく、RNAもバイオバンク試料としてほしい。
- ・保存病理組織(パラフィンブロック)や凍結組織もバイオバンク試料としてほしい。

これらをバイオバンク試料として収集することは重要だが、恒常的に収集、保管するためのバイオバンクを構築し運用するためには、JCOG内部のみならず各医療機関内での調整(液体窒素タンクの設置といったインフラ整備、臨床医や病理医へのインセンティブなど)が必要である。

また、現行の血液試料(血漿, DNA)の保管もデュープフリーザー2機で運用しており、近い将来、増設が必要である。

(2) 人員の確保

現在JCOGバイオバンクは、以下の人員構成で運用している。

- ・バイオバンクでの試料保管: 2名[国立がん研究センター東病院]。
- ・バイオバンクの運用(実施計画書/説明文書・同意書の作成、医療機関の問い合わせ窓口、試料処理搬送業者との調整、Web Entry Systemのメンテナンスなど): 3名[国立がん研究センター中央病院, JCOGデータセンター/運営事務局]。

マンパワーが非常に限られており、「(1)試料の種類、採取回数の拡充、インフラ整備」を拡充していくためには、さらなる人員の確保も必要である。

(3) 運用資金の確保

JCOGバイオバンクは2014年4月現在、独立行政法人国立がん研究センターがん研究開発費「医療イノベーション推進を目指した国立がん研究センター等バイオバンクの設立と運用」〔主任研究者: 金井弥栄(研究所副所長/分子病理分野分野

長)]の一部で運用しているが、インフラ整備や研究者へのインセンティブ、マンパワー等を確保するには十分ではない。

2. 将来—JCOGバイオバンクの将来展望—

JCOGが実施している臨床試験は、約6割がランダム化第III相試験であり、これらの試験で得られた診療情報と紐づいたJCOGバイオバンク試料を用いた試料解析研究では、単なる予後因子だけでなく、治療効果予測因子の探索も可能となる。

さらにJCOG試験では、試験実施計画書で規定された、均一性の高い集団、一定の治療方針、一定の規準による評価(有効性・安全性)が行われており、得られた診療情報もJCOGデータセンターでクリーニングされた質の高いデータである。

JCOGバイオバンク試料を用いた解析研究により、将来以下のような成果(可能性)が期待される。

- ・採血で得られたDNAや血漿タンパクにおける分子レベルの変化と、JCOG試験における毒性データとの関連を検討することで、毒性発現に関する予測因子の探索や検証が行える。
- ・毒性発現の予測因子が見出された場合、特定の患者が過度の毒性を回避し、より適切な治療を受ける治療の個別化につながる可能性がある。
- ・腫瘍組織における遺伝子やタンパク質などの分子レベルの異常と、JCOG試験によって得られた腫瘍縮小効果等の有効性データとの関連を検討することで、腫瘍の増殖や増殖抑制に関係している新たな分子標的の探索や、既知の分子標的の役割の検証が行える。
- ・新たな分子標的が見出された場合には、分子標的薬の創薬につながる可能性がある。
- ・腫瘍組織における分子レベルの異常と有効性データとの関連(交互作用)を治療別に検討することで、新たな治療効果予測因子(バイオマーカー)の探索や、既知の予測因子の検証が行える。
- ・新たなバイオマーカーが見出された場合には、バイオマーカーを用いた治療の個別化につながる可能性がある。
- ・腫瘍組織における分子レベルの異常と予後と

の関連を検討することで、新たな予後因子の探索や、既知の予後因子の検証が行える。新たな予後因子が見出された場合には、治療選択における医師や患者の意思決定に寄与できる可能性がある。

しかし、これら可能性をより確実なものとするためには、「1. 現状の問題点」に列記した課題を解決しなければならず、そのためには、何より

もまずJCOGバイオバンク運用のための恒常的な研究費確保が必要である。

文 献

- 1) Hoban CJ, Franklin W, Kopecky KJ, Baker LH. SWOG Cooperative Group Biorepository Resource : Access for Scientific Research Studies. Clin Cancer Res 2011 ; 17 : 5239.

* * *

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—

中島 章博^{*1}, 小森 哲志^{*1}, 神浦 俊文^{*1}, 杉谷 康雄^{*1}, 木村 友美^{*1}, 福田 治彦^{*2}

Benefit Risk Assessment
An Example Using Results from an Investigator-initiated Clinical Trial

Akihiro NAKAJIMA^{*1}, Tetsushi KOMORI^{*1}, Toshifumi KAMIURA^{*1}, Yasuo SUGITANI^{*1},
Tomomi KIMURA^{*1} and Haruhiko FUKUDA^{*2}

本稿は、2014年1月14日に医薬品医療機器レギュラトリーサイエンス財団主催の第149回レギュラトリーサイエンス エキスパート研修会『ベネフィット・リスク評価の実際』で、実例を用いて参加者とともに行ったワークショップの様子をまとめたものである。

研修会当日は、ベネフィット・リスク評価の基本となるフレームワーク¹⁻⁴⁾とその使い方、及び Multiple Criteria Decision Analysis (MCDA) に代表される定量的手法の基本的考え方を説明した。その後、がんの医師主導研究の二つの実例をもとに、まず演者が行ったベネフィット・リスク評価結果を一例として紹介した。その後、2-3人で1組となった参加者が、どちらかの例を選んで評価のプロセスを体験するワークショップを行った。また、限られた時間で項目の取舍選択やスコアリング、重みづけなどを体験してもらうために、あらかじめ Microsoft Excel[®] を使って作成した視覚化ツールを用意し、データも入力したうえで提供した。

以下では当日紹介した二つの試験のうち、進展型小細胞肺癌の治療法を比較した JCOG (Japan Clinical Oncology Group: 日本臨床腫瘍研究グループ) 9511 試験でのベネフィット・リスク評価の事例を紹介する⁵⁾。

なお、医薬品のベネフィット・リスク評価の基本については『ベネフィット・リスク評価入門』⁶⁾、定量的手法である MCDA については『ベネフィット・リスク評価 中級編』⁷⁾

(いずれも日本製薬工業協会ウェブサイトにて無料で公開中)をご参照いただければ幸いである。

1. 事例の紹介

JCOG9511 試験は、進展型小細胞肺癌に対するエトポシド+シスプラチン (EP) 療法とイリノテカン+シスプラチン (IP) 療法の多施設共同ランダム化第 III 相比較試験である。予定登録数を 230 例として 1995 年に開始したが、中間解析の結果、多重性を考慮しても統計学的に有意な差が生存期間に認められたため、登録数 154 例で早期中止され、IP 療法が進展型小細胞肺癌に対する新たな標準治療であると結論付けたものである。生存期間の中央値は IP 群が 12.8 か月、EP 群が 9.4 か月であった ($p=0.002$: 多重性調整しないログランク検定)。重篤な骨髄抑制は IP 群よりも EP 群で頻度が高く、重篤な下痢は EP 群よりも IP 群で頻度が高かった。

以下、代表的なベネフィット・リスク評価のフレームワークに沿って演者側から提示した内容を説明する。フレームワークは考える道筋や順番を提案するガイダンスで、いくつかのステップからなる。すなわち、①背景の定義、②ベネフィットやリスクの項目の列挙、③データソースの特定、④項目の選択、重みづけ、⑤結果のまとめと意思決定である。

^{*1} 日本製薬工業協会臨床評価委員会データ・サイエンス部会タスクフォース 1 東京都中央区日本橋本町 3-4-1 (〒103-0023)

^{*2} 国立がん研究センター多施設臨床試験支援センター/JCOG データセンター 東京都中央区築地 5-1-1 (〒104-0045)

1.1 背景の定義

ベネフィット・リスク評価の結果は、その目的や意思決定など背景情報によって異なり得る。すなわち疾患の重篤度やアンメット・ニーズによって、同じデータでもその医学的価値は異なる可能性がある。したがって、疾患や医薬品の情報など、意思決定の背景情報を定義することが極めて重要である。Table 1 に JCOG9511 試験のデータをもとにした意思決定の背景情報をまとめた。

当日のワークショップでは、評価する立場をスポンサー(例えば企業など)だけでなく、規制当局、患者や医療従事者など自由に選んで評価を行ってもらった。

1.2 データソースの特定、ベネフィットやリスクの項目の列挙と選択

通常なら、ここで複数の臨床試験からいくつかを選んだり、あるいはその全てを利用したりするが、今回は JCOG9511 試験の情報のみを参考にした。ここで、ベネフィットやリスクの項目を列挙し、選択した。BRAT (Benefit Risk Action Team) はデータソースを特定する

際、試験デザインの特徴、試験の質の尺度(例えば、Jadad Scale)などの情報も含めたデータソーステーブルを作成することを提案している⁹⁾。意思決定に影響を与える項目として、ベネフィットは生存期間、リスクは治療関連死、米国 National Cancer Institute (NCI) による共通毒性規準 (Common Terminology Criteria for Adverse Events (CTCAE)) での Grade 4 の非血液毒性 (生命を脅かす有害事象)などを挙げ、バリュー・ツリーを作成した (Fig. 1)。バリュー・ツリーとは、項目をベネフィットとリスクにわけて階層構造状のツリーで表示したものである。

ワークショップでは、付箋紙にベネフィットやリスクの項目を一つずつ記載し、それらに優先順位をつけて並べ替えたり、取捨選択したりしながら、それぞれのグループのバリュー・ツリーを作成してもらった。今回は JCOG9511 の試験結果を提示した上で、項目の列挙と選択を同時に行ったが、データソースを特定する前に項目を列挙したほうが、データに影響されることなく、重要な項目を列挙できるだろう⁹⁾。

Table 1 JCOG9511 試験の背景情報

ベネフィット・リスク評価を行う立場：	試験実施者である JCOG (スポンサー)
意思決定の目的 (=JCOG9511 の試験目的)：	進展型小細胞肺癌に対するよりよい治療法の選択 <ul style="list-style-type: none"> ・エトボシド+シスプラチン (標準治療) ・イリノテカン+シスプラチン (新治療)
対象疾患： 進展型小細胞肺癌	増殖速度が速く、予後が不良 <ul style="list-style-type: none"> ・3年生存割合：10%程度
(当時の) 標準治療の効果	<ul style="list-style-type: none"> ・奏効割合：75%程度 ・完全奏効割合：10%程度
罹患率	肺癌の年間発症数は男性で約 65,000 人、女性で約 30,000 人。 そのうち、小細胞癌は 10~15%。
アンメット・ メディカルニーズ	<ul style="list-style-type: none"> ・生存時間の延長 ・重篤な有害事象の軽減 ・有効な分子標的薬なし
対象集団	<ul style="list-style-type: none"> ・組織診又は細胞診で小細胞肺癌の診断 ・進展型の診断 ・前治療なし ・測定可能又は評価可能病変を有する ・PS*：0, 1, 2 ・少なくとも3か月以上の生存が可能と考えられる ・年齢：70歳以下 ・骨髄、肝、腎機能が保持されている ・重篤な心肺合併症、3年以内に活動性の重複癌がない
医薬品の情報	用法・用量
	<ul style="list-style-type: none"> ・EP (エトボシド+シスプラチン) <ul style="list-style-type: none"> ➢ エトボシド 100 mg/m², day1,2,3 ➢ シスプラチン 80 mg/m² day1 ➢ 3週ごとに4コース ・IP (イリノテカン+シスプラチン) <ul style="list-style-type: none"> ➢ イリノテカン 60 mg/m² day1,8,15 ➢ シスプラチン 60 mg/m² day1 ➢ 4週ごとに4コース

*Eastern Cooperative Oncology Group (ECOG) performance status

JCOG9511の Value Tree

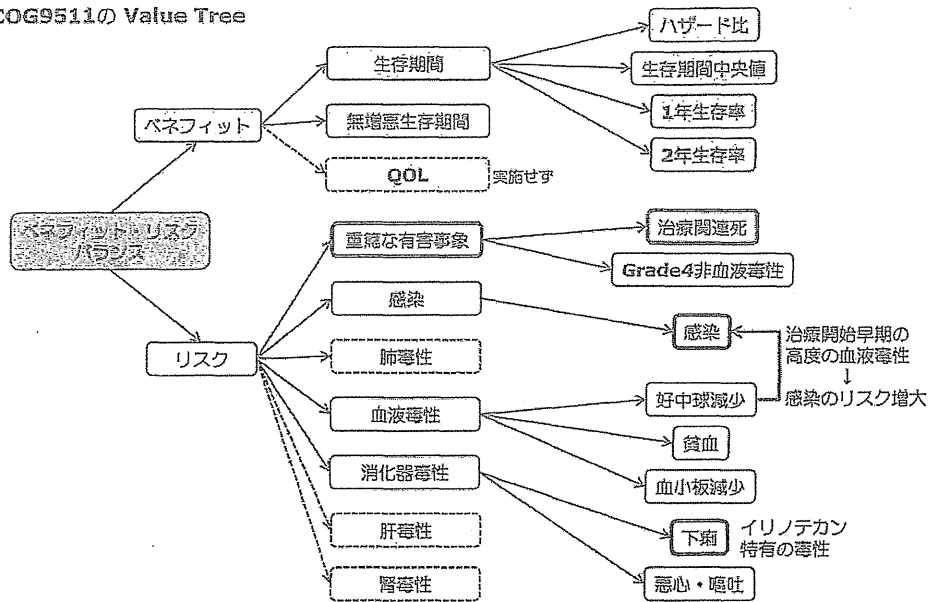


Fig.1 バリュー・ツリーの例

1.3 重みづけと結果のまとめ(視覚化)

演者からは定量的手法を用いた重みづけは提示せず、定性的に優先順位をつけた例を提示した。最後に結果をまとめ、半定量的手法⁹⁾を用いてフォレスト・プロットによる視覚化を行った例を示した (Fig. 2)。各項目のリスク差の点推定値と95%信頼区間をプロットし、上に位置する項目ほど優先順位が高いと考えた。

なお、臨床的にはハザード比が生存時間の評価指標とし

てよく使用されるが、今回は2年生存率及び1年生存率の差を用いて示している。演者の解釈として、IP療法により享受できる生存期間延長(ベネフィット)は、治療関連死とGrade4非血液毒性という重篤な有害事象のリスク増加よりも大きいと考えること、下痢、感染は注意深く観察し、支持療法の充実を今後図る必要があること(註:その後支持療法の進歩により、下痢のコントロールが可能となった)などを紹介した。

JCOG9511の Forest Plot

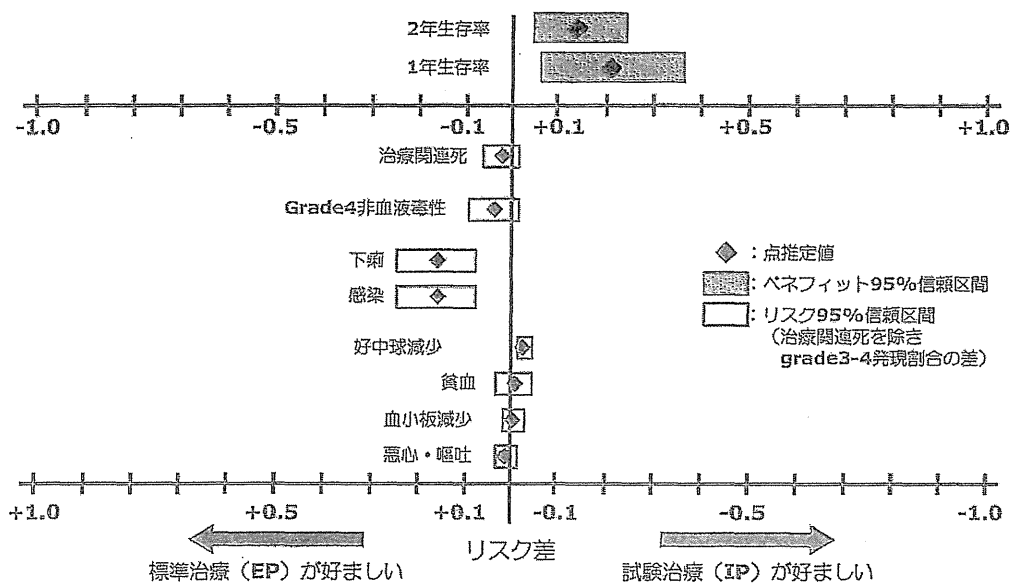


Fig.2 フォレスト・プロットの例

2. 定量的手法体験ワークショップ

以上のように、演者側からは意思決定の背景、試験の概要及び半定量的(定性的)方法を用いて行ったベネフィット・リスク評価の事例を紹介した。実際はこれで十分な場合が少なくない⁶⁾。しかし、このワークショップではMCDAに代表される定量的手法を試してみたい参加者のために、簡単な説明とツールを提供した。MCDAの基本的な説明はタスクフォースの報告書⁷⁾に譲るが、希望者には以下の三つの手順で考えてもらった。

2.1 スコア化

1.2項で挙げた個々の項目に対し、その項目が取る値の大きさを医療上の価値の大きさに翻訳する作業である。例えば、ベネフィットの項目の一つである2年生存率が0%から30%の範囲で変化し得ると考え、生存率ごとの価値の大きさを評価する。2年生存率が0%とき価値は最小で0点、また30%のとき価値は最大で100点等と定めよう。Fig. 3 Aの左側のグラフでは、2年生存率が大きいほど比例的に価値も大きいと評価している。一方、右側のグラフでは、2年生存率が10%より小さいか大きいかで、価値が増加する程度(直線の傾き)が異なると評価している。

同様に、例えばリスクの項目の一つである治療関連死の発現割合について、やはり発現割合の値と医療上の価値との関係を検討する。Fig. 3 Bのグラフのように、発現割合が大きいほど治療上の価値は比例的に小さくなる、というのも一つの評価である。また、Fig. 3 Bの右のグラフのよ

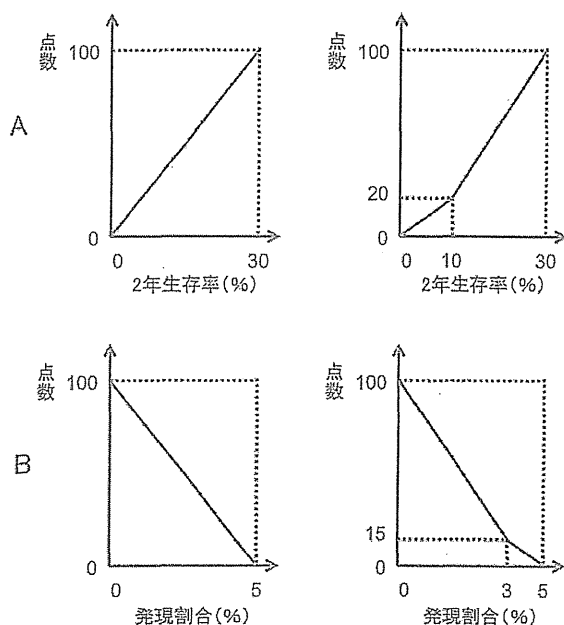


Fig.3 2年生存率(A)及び治療関連死(B)のスコア化の例

うに、発現割合が3%より小さいか大きいかで比較して、医療上の価値の減少の程度が異なると判断する場合もあるかもしれない。

このように、スコア化を通じて、個々の項目を自分ならどのように医療上の価値と結び付けているかを明確にする。

2.2 重みづけ

データと医療上の価値との関係をだまかに決めたら、次に複数の項目間の相対的な重要性の違いを0から1までの値を取る「重み」として定める。今回は、全ての項目の重みの合計が1となるように定めたツールを配布した。先の例では、2年生存率については0%の場合を0点、30%の場合を100点と定め、また治療関連死については5%の場合を0点、0%の場合を100点と定めていた。

すなわち、2年生存率については生存者が30%増えることが100点の差として表されており、また治療関連死については死亡者が5%減ること(あるいは生存者が5%増えること)が100点の差として表されている。どちらも100点の違いではあるが、これらが医療上等価であるとは考えにくい。両者を天秤にかけたときに、どちらがどれくらい重いかという判断を、それぞれの項目対ごとに考えてもらった。これは、ある項目における100点という価値を、別の項目における100点と相対比較する作業であり、一方が他方に対して何倍価値があるか、というトレード・オフの関係を定めることでもある。中には、ある項目の相対的重要性を減らしたところ、結果が思わぬ方向に変わったので驚いたという参加者がいた。その理由を探ってみたところ、その項目の重みを小さくした分、他の項目の重みが相対的に大きくなったためであった。

このように、重みの評価は、複数の項目を同時に考慮しながら絶対的な評価を定める作業であり、こうして定めた重みを通じて、その評価者がどの項目をどの程度重視しているかを明確にする。

2.3 合計スコア

ワークショップで配布したツールは、項目ごとの「スコア×重み」を全て足し合わせて合計スコアを算出するものであった。また、全ての項目の合計スコアだけでなく、ベネフィットとリスク、それぞれの総合スコアを算出することもできた。このような計算結果の例をFig. 4及びTable 2に示した。

Best, Worstは変化する範囲の最良値(価値が最大となる値)と最悪値(価値が最小となる値)であり、項目ごとにこの範囲でスコア化を行う。項目間の相対的な重要度(重み)を検討し、Table 2のようなEffect Tableにまとめ、

Table 2 総合スコア算出

項目	変化する範囲		重み	試験結果		単位
	Best	Worst		イリノテカン +シスプラチン(IP)	エトポシド +シスプラチン(EP)	
ベネフィット	合計					
生存期間	中央値					
	15	8	0	12.8	9.4	month
1年生存率	70	20	100	58.4	37.7	%
2年生存率	40	0	100	19.5	5.2	%
無増悪生存期間	中央値					
	8	3	0	6.9	4.8	month
6か月無増悪生存率	80	20	0	65.3	35.6	%
12か月無増悪生存率	15	5	0	12.5	7.9	%
リスク	合計					
重篤な有害事象	治療関連死					
	0	5	30	4.0	1.8	%
Grade4 非血液毒性	0	15	10	6.7	2.6	%
感染	感染					
	0	30	5	5.3	3.9	%
血液毒性	好中球減少					
	0	100	2	65.3	92.2	%
貧血	0	50	1	26.7	29.9	%
血小板減少	0	30	1	5.3	18.2	%
消化管毒性	下痢					
	0	30	5	16.0	0.0	%
悪心・嘔吐	0	30	1	13.3	6.5	%

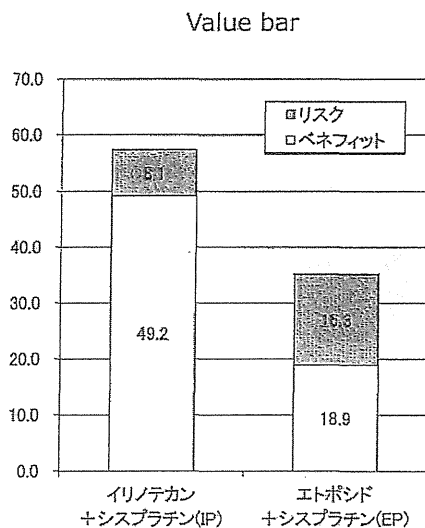


Fig.4 総合スコア視覚化の例

合計スコアを算出する。合計スコアは大きいほど価値が高く、Fig.4のような Value bar を用いて視覚化する。

3. 感度分析

ワークショップは2人(あるいは3人)1組で行ってもらったが、その中でも各項目の相対的重要性や解釈に違い

が見られた。当日はツールを提供したので、それぞれの組内で意見が割れた場合には、両方の意見を試してみること(=感度分析)も体験してもらった。演者側でもあらかじめスコアリングや重みづけに以下のようないくつかのバリエーションを仮定して、それぞれがベネフィット・リスク評価で得られた結論にどのような影響を与えるのかを確認する感度分析を実施し、ワークショップの後に紹介した。

3.1 重要な項目の重みを0~100%まで変化させる

リスクの項目のうち、治療関連死を特に重要な項目として重みを変化させ、結果の変化をFig.5aに示した。治療関連死の重みが37%以上となった場合に結果が逆転し、EP群のベネフィット・リスク・バランスの方が優れている結果となった。とは言え、あらためて治療関連死の重みを37%以上とすることの意味を考えてみると、これは、例えば治療関連死が5%増えることを、2年生存率や1年生存率が50%~70%程度減少することと同程度と評価することに相当する。このような評価が妥当と考えられるケースは少ないと考えられるため、結論が逆転することはなさそうだ、と言えそうな結果である。

3.2 重要な項目の結果の値を、信頼区間の下限から上限まで変化させる

今度はベネフィットの項目のうち、各群の2年生存率の

値を信頼区間下限から上限まで変化させた場合の結果の変化を Fig. 5 b に示した。変化させた範囲内のどのような値をとった場合でも、結果が逆転することはなかった。

3.3 バリュー・ツリーに一項目追加する
重要性や一貫性の観点からバリュー・ツリーから落としていた項目があった場合に、それらを追加することが結果

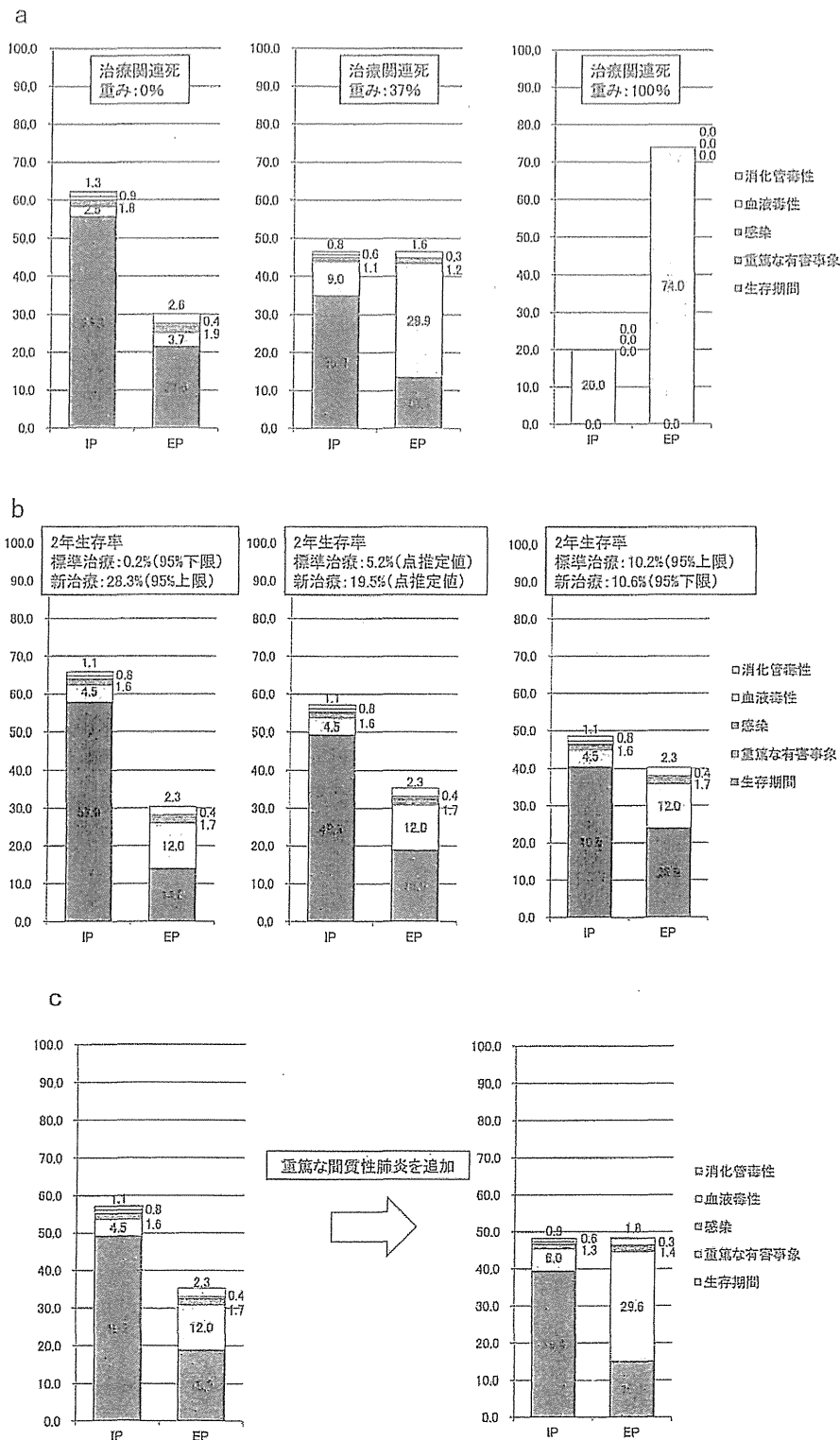


Fig. 5 感度分析の事例

に影響を与えるかを検討するものであり、今回は結果に影響を与えるものはなかったが、仮想的な例として「重篤な間質性肺炎」がIP群で4.4%発現し、EP群では0%であったという情報があり、Bestを0%、Worstを5%としてスコア化し、重みを20%とした場合、結果がどのように変化するかをFig. 5cに示した。重篤な間質性肺炎を追加した結果、IP群とEP群でベネフィット・リスク・バランスがほぼ同じになった。ベネフィット・リスク評価にこれほど大きな影響を与える項目がもし漏れていたなら、こういった項目はバリュー・ツリーに追加すべきである。

ここで取り上げた例以外にも、ベネフィット全体の重みを0～100%に変化させたり、スコア化に影響するBest・Worstの値を変化させる感度分析などもある。特に、不確かさが大きい要素について感度分析を行う必要があるだろう。

4. 考察

ベネフィット・リスク評価は新しく古い話題である。最近欧米の複数のグループから、医薬品のベネフィット・リスク評価に応用可能なフレームワークが相次いで提案されたことから、それぞれの医薬品に関して具体的な取り組みが行いやすくなった背景がある。今回のワークショップではその中の一つ、米国PhRMAから提案されたBRATのフレームワークを主に参考にした定性的(半定量的)な手法を用いて、日本のがんの臨床試験の実例を用いたベネフィット・リスク評価の事例を紹介し、更にその評価をもとにMCDAを中心とした定量的手法にも触れてもらった。

今回のワークショップでは、がん領域の事例を紹介したが、これにはいくつかの理由がある。まず、がん領域ではベネフィットやリスクの項目が国際的に標準化されており、複数の試験間の比較が行いやすい。例えばベネフィットとしては通常生存期間の延長、リスクとしてはNCI共通毒性規準を用いた評価項目が、国や試験によらず広く用いられている。ただ、通常ベネフィットの指標として用いるハザード比をリスクと同じスケールのフォレスト・プロットで表現しにくいなどの限界点もある。更に対処可能な有害事象やリスク最小化策の効果が期待できる安全性懸念事項に関しては、利用するデータの幅やその重みづけに工夫が必要である。また、フォレスト・プロット上で表現されないアンメット・メディカルニーズなどの重要な背景情報についても、総合的に判断して最終的な評価を行う必要がある。

なお、当時はQoLの評価項目を取得していなかったが、当日のワークショップではこの評価項目を重要視してベネ

フィット・リスク評価を行った参加者が多かった。

MCDAに関しては、上述の通り必須ではないし、かなりの人手や時間がかかる。更に、定量的な方法を用いることでより客観的な評価ができるはずだ、などと考えるのは短絡的である。むしろ、本来主観的に行うほかないベネフィット・リスク評価の中身を、可能な範囲で、あるいは最小限必要なところだけ定量化して見やすく、もしくは「透明に」している、と考えた方が実情に近い。もし行くとすれば、例えば同時に複数の項目を評価する場合、突出して「明らかに重要」といえるような項目がなく項目間の重要性が競合する場合、議論を重ねても評価者によって判断が異なる場合、不足情報等がありデータの確からしさが疑わしい場合などに効果があるかもしれない。

また、スコアリングのモデルなどの「正しさ」に不安が残る場合もあるかもしれないが、そもそも「正しい」関数を完璧につくることが本来の目的ではない。逆にこうした「関数」を申請時に作成しておけば、市販前から市販後まで通してベネフィット・リスク評価の履歴管理ができる。有効性・安全性の情報を更新する必要が生じたら、関数に入力していた値を再検討・変更したり、ある評価項目の医学的な重要性を見直す必要が生じたら、関連する評価項目の重みを再検討・変更したり、未知の重篤な有害事象の情報が得られたら、評価項目を一つ追加して(スコア化等も行い)関連する評価項目の重みも再検討・変更したりすればよい。透明性を高めるために、審査時に使った関数を公開し、市販後のベネフィット・リスク再評価でも運用することも考えられる。更に、3章で述べた感度分析は現状の結果について実施したものであったが、近い将来に新しいデータが得られた場合に、どのようにベネフィット・リスク評価の結果が変わるかを感度分析によって検討することも可能である。例えば重要な有害事象が新たに見つかった場合に、どのように変化するかを、3.3項と同様の方法で検討が可能であるし、もし医療環境の進歩により下痢がコントロール可能になったら下痢の重みももっと低くなるかもしれないが、その場合に評価結果がどのように変化するかを、3.1項と同様の方法で検討できる。こういった検討を行うことで、新しい情報が得られる前に先手を打った対応がとれる場合もあるかもしれない。

当日のワークショップには80名近くの参加者がおり、全ての組で行われた評価を把握することはできなかったが、それでもいろいろな意見が出たことはうかがえた。もっと大人数で行った時には、更に多くの、そして大きな差異がみられることだろう。したがって定性か定量かによらず、感度分析を常に検討すべきである。

3章で紹介した以外にも様々な感度分析の方法があり、例えば別の立場の方々にもベネフィット・リスク評価を実施

してもらったり、今回のワークショップで体験してもらったように、検討した項目の取舍選択や相対的重要性が変わるとどのように結果が変わるのか、なども感度分析の一種である。様々な感度分析を行うことで、自身が実施したベネフィット・リスク評価の意味をより深く理解することに大きな意味がある。

MCDCA を用いた定量的手法では、実際には価値判断がより細かく繰り返して実施される分、感度分析の重要性が更に上がることになる。Hiviewなどのツールを用いることで、重みづけの調整などが評価の結論にどのように影響するかを視覚化することもできる。こういった様々な視覚化ツールを使うことは感度分析の負担を下げ、ステークホルダー間の相互方向の合意形成に役立つだろう。

おわりに

医薬品のベネフィット・リスク評価は2014年3月現在、新たなICHのテーマ候補としても提案されており、今後当局や各企業、又はアカデミアなどで継続して取り組まれていく課題であると考えられる。定量的手法を用いた解析を行う場合でも、患者や医療従事者の定性的(質的)な評価が各項目のスコアリングや重みづけに必須であることから、事前の選考に関する調査を積極的に行う必要がある。また、がん領域のように比較的国際的標準化が進んだ領域ばかりではないことも考慮する必要がある。

このように臨床試験計画の段階から申請時のベネフィット・リスク評価を念頭に置いて計画することが重要で、ここで紹介した事例がその一助となれば幸いである。

文 献

- 1) Coplan, P.M., *et al.*, Development of a framework for enhancing the transparency, reproducibility and communication of the benefit-risk balance of medicines. *Clin Pharmacol Ther.* 2011, 89 (2), p. 312-315.
- 2) Liberti, L., McAuslane, N., Walker, S. Progress on the development of a benefit/risk framework for evaluating medicines. *Regulatory Focus.* 2010, p. 1-6.
- 3) European Medicines Agency, Benefit-risk methodology project. Work package 4 report: Benefit-risk tools and processes. [http://www.ema.europa.eu/docs/en_GB/document_library/Report/2012/03/WC500123819.pdf]. May 9 2012.
- 4) FDA, Structured approach to benefit risk assessment in drug regulatory decision making. Draft PDUFA V Implementation Plan. <http://www.fda.gov/downloads/ForIndustry/UserFees/PrescriptionDrugUserFee/UCM329758.pdf>, February 2013.
- 5) Noda, K., *et al.*, Irinotecan plus cisplatin compared with etoposide plus cisplatin for extensive small-cell lung cancer. *N Engl J Med.* 2002, 346 (2), p. 85-91.
- 6) 日本製薬工業協会データサイエンス部会, ベネフィット・リスク評価 入門. 2013. <http://www.jpma.or.jp/about/board/evaluation/allotment/pdf/benefitrisk.pdf>.
- 7) 日本製薬工業協会データサイエンス部会, ベネフィット・リスク評価 中級編 多基準決定分析への招待—その理論と事例—. 2013. http://www.jpma.or.jp/information/evaluation/allotment/pdf/benefitrisk_mid.pdf.
- 8) Levitan, B.S., *et al.*, Application of the BRAT framework to case studies: observations and insights. *Clin Pharmacol Ther.* 2011, 89 (2), p. 217-224.
- 9) Levitan, B., Evaluating benefit-risk during and beyond drug development: An industry view. *Regulatory Reporter.* 2012, 9 (6).

日本臨床 72 卷 増刊号 2 (2014 年 2 月 20 日発行) 別刷

最新がん薬物療法学

—がん薬物療法の最新知見—

VII 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)

中村 健一

水澤 純基

柴田 大朗

福田 治彦

VII 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)

Primary endpoint for cancer clinical trials to establish standard treatment:
Overall survival or progression-free survival

中村 健一 水澤 純基 柴田 大朗 福田 治彦

Key words : overall survival, progression-free survival, true endpoint, surrogate endpoint, hard endpoint

はじめに

従来、がんの第III相試験の primary endpoint の gold standard は全生存期間 (overall survival: OS) であったが、最近では、① OS は後治療の影響を受けやすい、② OS の延長に従ってサンプルサイズが増大する、といった問題から無増悪生存期間 (progression-free survival: PFS) が用いられる機会が増えている¹⁾。臨床試験において PFS を primary endpoint として設定しうるか、という問題に対する答えは、試験の相や目的によって異なりうるが、本稿では標準治療の確立を目的とした第III相試験において、PFS を primary endpoint としうるかどうかを考えるうえで、知っておくべき基本的な考え方のフレームワークを述べる。

1) Trueness と hardness

米国 NCI の統計家である Richard Simon は、エンドポイントとは「患者のベネフィットを測るものさし」であると定義している。この定義に従えばエンドポイントは、患者のベネフィットを反映し、かつ、ものさしとして正確である必要がある。本稿では前者の「患者のベネフィット

を反映しているかどうか？」を trueness と呼び、後者の「きちんと測れているか？」を hardness と呼ぶ。

1) True 軸

trueness の観点からみた際のエンドポイントには2種類存在する。患者の真のベネフィットを反映する true endpoint と、true endpoint の「代替 (surrogate)」として用いられる surrogate endpoint である。前者の代表は OS で、後者の代表は奏効割合である。腫瘍縮小自体は患者のベネフィットとはいえないが、腫瘍縮小が延命につながり、かつ、OS に比べて結果を早くに知ることできることから、奏効割合は古くから surrogate endpoint として広く用いられてきた。つまり良い surrogate endpoint であるためには、① true endpoint とよく相関し、② surrogate endpoint を用いることで何らかのメリットが存在することが必須である。

2) Hard-soft 軸

一方、エンドポイントは「ものさし」であるため、ものさしとしての機能が優れている必要がある。優れたものさしでは、誰が見ても同じ結果を導き出せる必要があり、このようなものさしを hard なエンドポイントという。例えば死亡日は誰が見ても変わらないため OS は hard であ

VII

Kenichi Nakamura, Junki Mizusawa, Taro Shibata, Haruhiko Fukuda: Multi-institutional Clinical Trial Support Center, National Cancer Center; JCOG Operations Office/JCOG Data Center 国立がん研究センター多施設臨床試験支援センター, JCOG 運営事務局/JCOG データセンター

るが、QOLなどは人によって評価が異なりうる soft なエンドポイントの代表である。試験の primary endpoint として設定するには (true か surrogate かにかかわらず) hard であることは必須の条件である。

2 Trueness からみた PFS

エンドポイントを考える際に必要なフレームワークとして、trueness と hardness という 2 つの観点を述べたが、この trueness から考えた際には、PFS が primary endpoint として用いることのできる条件は 2 つしかない。 すなわち、① PFS が OS の surrogate endpoint となることが示されている、② PFS 自体が患者にとっての clinical benefit をもつ、のいずれかである。

1) PFS が OS の surrogate endpoint である

メタアナリシスなどで PFS が OS の surrogate endpoint であることが示されている場合には、PFS 自体が clinical benefit をもつことを示す必要はない。これは、surrogate endpoint である以上、PFS が positive であれば、clinical benefit をもつ OS も positive となることが前提であるからである。各種のがんのメタアナリシスで OS に対する PFS の surrogacy を検証しようとする試みがなされているが、現在のところ一定の surrogacy が示されているのは大腸がん²⁾、頭頸部がん³⁾のみであり、反対に胃がん、乳がんなどでは surrogacy が示されていない。

ただ、最近では大腸がんの分子標的薬を含む試験も含めた大規模なメタアナリシスで、PFS が OS の surrogate endpoint とはならなかったという発表がなされており、二次治療以降で有効な薬剤が存在するようながん種では、PFS の OS に対する surrogacy を示すことは今後は難しくなるであろう⁴⁾。

2) PFS 自体が clinical benefit をもつ

PFS を primary endpoint に設定しうる、もう一つの条件は、PFS 自体が true endpoint であること、つまり何らかの clinical benefit をもつことである。PFS ではないが、頭頸部がんにおける喉頭温存生存期間は OS 以外の true endpoint

の好例である。喉頭摘出することになれば、発声、味覚といった点で患者の QOL に大きな障害が生じることから、仮に OS に差が全くないとしても喉頭温存生存期間の延長は患者のベネフィットを直接反映すると考えられる。

一方、通常の PFS がもつ clinical benefit については controversial で、OS が延長するという保証がない状況で、増悪までの期間がわずかに延長したとしても、それは毒性や QOL の低下に見合わない可能性もあり、また同じ 20% の径増大といっても、小さな病変が 20% 増大したとき、大きな病変が 20% 増大したときでは意味が異なる点も clinical benefit を正確に反映しているとはいえない⁵⁾。clinical benefit をもつかどうか疑わしい状況で PFS が多少延長したとしても、PFS に比べて増悪以降の survival が相対的に長くなってきている現状では、PFS のわずかな延長がもつ意味は更に薄れているといえよう。

以上より、標準治療を決定する第 III 相試験において PFS を primary endpoint として用いるための必須条件は、① PFS が OS の surrogate endpoint となることが示されている、② PFS 自体が患者にとっての clinical benefit をもつ、のいずれかを満たし、かつ、hard であること、となる。NCI の統計家である Korn らも、'clinical benefit を直接的に測る中間的なエンドポイントが存在しないがん種では、OS を primary endpoint とすべき' と結論している⁶⁾。

3 増え続けるサンプルサイズへの対応策

しかし、現実問題として、OS が 20 カ月以上になるようながん種では、OS を primary endpoint としてランダム化比較試験を組むことは、多くの場合で登録数の観点から困難である。稀少がんなどで十分な患者数が存在しないような場合にも同様の問題が生じるが、このような状況への対応策は、次に挙げるような方策が考えられる。

1) 大規模なサンプルサイズに対応できる体制を構築する

近年、国際共同試験や国内共同試験の必要性が叫ばれているが、共同試験の重要な動機の一つは、OSをprimary endpointとしたときに増え続けるサンプルサイズである。ただし、国際共同試験ではしばしば国別に治療効果の差がみられるというデメリットも存在するため、共同試験の相手方は対象や治療により慎重に検討する必要がある。

2) デザインの工夫によるサンプルサイズの調整

a. 試験の精度を落とす

通常の優越性デザインを取るかぎり、サンプルサイズを減らすために緩められるパラメータは、 α (有意水準)、 β (本当は有効な治療を捨ててしまう確率、 $1-\beta$ が検出力)、 Δ (臨床的に意味のある差)の3つしかない。 Δ は試験治療の毒性などのデメリットに見合ったアウトカムの幅であり、臨床的な意思決定と試験結果を一致させる観点からは Δ を緩めることはお勧めできない。通常の第III相試験では、 α は片側2.5%か両側5%、 β は10-20%に設定されるのが世界標準であるが、対象疾患の患者数が限られている場合には β を20-30%程度、 α についても片側5%や、かなりの稀少がんである場合には片側10%も許容されるかもしれない。もちろん試験の精度を落とすことは望ましくないが、精度を多少落としたとしても試験を行う価値があると考えられる場合には、やむを得ない選択肢とはなりうる。

b. 特殊な非劣性試験

ある特殊な状況では、NCIのFreidlinらが‘hybrid design’と呼んでいる非劣性試験の設定を行うことによりサンプルサイズを減らすことができる場合がある⁷⁾。通常、非劣性試験では標準治療と試験治療で見込まれる予後は同じに設定するが、このデザインでは、試験治療群が少し勝つという見込みがある場合にそれを前提として計画することでサンプルサイズを劇的に減らすことができる。詳細は他稿に譲るが⁸⁾、試験治療が標準治療に比べ有効性以外の何らか

のメリットを有し(非劣性デザインが正当化できる根拠があり)、かつ、試験治療が少し勝つという見込みが根拠に基づき説明可能な場合には適用できる可能性がある。

3) PFSとOSのどちらをもdecision ruleに組み込む

3つめのオプションとしては、PFSでサンプルサイズ設定を行うことによって登録数は現実的な範囲におさめ、試験のdecision ruleとしては、PFSが有意でかつOSで試験治療が標準治療に一定以上上回る(例:OSのハザード比の点推定値が0.85を下回る=検定は行わない)場合に‘positive’と結論づけるというデザインである。こうすることで、試験治療の真のOSが標準治療の真のOSを下回っている際に、間違っただけで試験治療を取ってしまう確率を一定以下に抑えることができる。多くのPFSでは、一定のtruenessをもっていたとしても、そのtruenessは価値観に左右される程度のものであり、OSで一定以上上回らないことには、コミュニティが試験レジメンを新たな標準治療として受け入れることはないだろうという考えがこのデザインの背景にある。ただしJCOGでは、まだこのデザインの実例はなく、まだ広く受け入れられているデザインではないため、設定に際しては生物統計家への相談が必須である。

おわりに

これまで主にtruenessの観点からPFSがprimary endpointとして用いることのできる条件を述べてきたが、PFSがtrueであれ、surrogateであれ、hardであることは必須条件である。PFS評価には様々なバイアスが入りうるため、これらのバイアスを最小化することはPFSをprimary endpointとして使用する場合には必須である。少なくとも、標準治療を確立する第III相試験では本稿で述べたようなtruenessやhardnessを十分考察のうえで、PFSをprimary endpointとすべきかどうかを決定すべきであり、本稿で示したフレームワークやデザインの工夫がそれらの議論の一助となれば幸いである。

VII

謝辞 なお、本稿の内容は、独立行政法人国立がん 田治彦)に基づく研究成果によるものである。
研究センターがん研究開発費 23-A-16(主任研究者: 福

文献

- 1) Broglio KR, Berry DA: Detecting an overall survival benefit that is derived from progression-free survival. *J Natl Cancer Inst* 101: 1642-1649, 2009.
- 2) Buyse M, et al: Progression-free survival is a surrogate for survival in advanced colorectal cancer. *J Clin Oncol* 25: 5218-5224, 2007.
- 3) Michiels S, et al: Surrogate endpoints for overall survival in locally advanced head and neck cancer: meta-analyses of individual patient data. *Lancet Oncol* 10: 341-350, 2009.
- 4) Shi Q, et al: Individual patient data (IPD) analysis of progression-free survival (PFS) versus overall survival (OS) as an endpoint for metastatic colorectal cancer (mCRC) in modern trials: Findings from the 16,700 patients (pts) ARCAD database. *J Clin Oncol* 31: (Suppl: Abstr 3533), 2013.
- 5) Fleming TR, et al: Issues in using progression-free survival when evaluating oncology products. *J Clin Oncol* 27: 2874-2880, 2009.
- 6) Korn EL, et al: Overall survival as the outcome for randomized clinical trials with effective subsequent therapies. *J Clin Oncol* 29: 2439-2442, 2011.
- 7) Freidlin B, et al: Randomized clinical trial design for assessing noninferiority when superiority is expected. *J Clin Oncol* 25: 5019-5023, 2007.
- 8) 中村健一ほか: PFSは第III相試験の primary endpoint となりうるか? 知っておくべき考え方のフレームワーク. *腫瘍内科* 12: 401, 2013.

Copyright © 2014, Igaku Shoin Ltd. All rights reserved. Printed in Japan.

特集Ⅱ

今後の臨床研究の方向性

臨床研究推進に貢献する
臨床研究者育成*

松岡 豊**

Key Words : 患者志向型研究, 臨床試験, 臨床研究教育, 臨床疫学

はじめに

わが国ではドラッグ・ラグが社会問題化して以降, 積極的な臨床研究・治験活性化計画が進められてきたが, その期待を担うはずの人材育成についてはあまり語られてこなかった。生物医学分野で世界との競争力にわが国が最も劣っている領域は, 臨床試験の結果集積に基づき, 医療を徐々にではあるが確実に進歩させていく患者志向型研究(patient-oriented research : POR)であると指摘されている¹⁾。米国NIHは, 臨床研究・橋渡し研究における次世代の研究者を育成することを本務と考え, 若手研究者の給料や研究費を助成する K グラントや若手を指導するメンターに対するグラントも整備している²⁾。米国が生物医学分野で世界のリーダーになっている背景には, 膨大な基礎研究分野への投資に加えて, 大規模な臨床試験や疫学研究を確実に実施できるインフラ整備と人材育成を行ってきた歴史がある¹⁾。わが国の研究を推進するためには, 再生医療やオーダーメイド・ゲノム医療等の最先端医療を実施する体制作りだけではなく, その成果を日常診療に橋渡しするPORが重要であ

る。

科学技術立国を標榜するわが国では, 疾病の制圧と健康な社会の構築を目標とする医学研究の益々の発展が求められる。「科学技術イノベーション総合戦略2014(平成26年6月24日閣議決定)」では臨床現場で活用される医療技術の開発を5つの政策課題の1つにあげている。そして科学技術イノベーションに適した環境創出に必要な重点課題として, 世界で活躍する人材に対する「挑戦」の機会提供と人材育成に取り組む大学や研究開発法人の機能強化を掲げた。また「健康・医療戦略(平成26年7月22日閣議決定)」においても, 臨床研究および治験の効率的・効果的な推進のための人材育成・確保, 臨床研究および治験関連業務に従事する職員に対する臨床研究および治験に係る教育機会の確保・増大を明記している。本稿ではPORに関する教育の現状と課題について検討し, 筆者らが平成22年度から国立精神・神経医療研究センターで取り組んできた臨床研究者育成プログラムについて紹介する。そして臨床研究を推進していくうえでの教育のあり方について若干の私見を述べる。

患者志向型臨床研究が求められている

医学研究は一般に基礎研究と臨床研究に大別される。臨床研究には病気の成因を明らかにす

* The development of physician scientist for promoting clinical research.

** Yutaka MATSUOKA, M.D., Ph.D.: 独立行政法人国立精神・神経医療研究センタートランスレーショナル・メディカルセンター情報管理・解析部 [〒190-0014 東京都小平市小川東町4-1-1] ; Department of Clinical Epidemiology, Translational Medical Center, National Center of Neurology and Psychiatry, Kodaira, Tokyo 190-0014, JAPAN

る疾患志向型研究(disease-oriented research : DOR)と患者を対象として診断, 治療法を開発する患者志向型研究(POR)がある。DORは組織や遺伝子をはじめとした患者の生体試料を用いて疾患の病態解明, 診断法や治療法の開発を目的に行われることが多いので, 基礎研究の延長線上にあると捉えることもできる。PORは, 対象が個性や感情を持つ「まごとの人間」であり, 臨床疫学や生物統計学の知識が必須で, 心理学・行動医学的アプローチも必要となる⁹⁾。カナダ国立健康研究所ではPORを広義にとらえ, ヒトにはじめて行う研究から臨床試験, アウトカム研究, そしてその知見をヘルスケアシステムや臨床実践に橋渡しすることまで含めている。

エビデンスレベルの高いPORを掲載する臨床医学誌(New Engl J Med, Lancet, JAMA)の2008年から2011年の4年間に掲載された国別原著論文数を調査した報告によると, わが国は25位であった⁹⁾。しかし, 同時期におけるDOR等を掲載する基礎系雑誌(Nature Medicine, Cell, J Exp Med)の結果は4位と健闘していた。これまでわが国にはPORを推進する機運, 人材, 体制, 財源等が不足していたわけであるが, この低迷から脱するようになることは不可能ではない。前述の閣議決定を根拠に, 今こそ現場の臨床家が中心となって, 価値あるPORを世界に発信する時機が到来している。そのためにも有意な人材育成とそれを支える環境整備が必要である。研究推進に求められる狭義の人材は, 生物統計家, CRC, データマネジャー, 研究倫理専門家, 法律家などであるが, 以下あまり語られていないprincipal investigator (PI)育成に絞って論じる。

臨床医の臨床研究教育に対する期待

PORを推進するためには, 臨床の最前線を担う臨床医における臨床研究実施への関心や教育へのニーズを把握しておく必要がある。三品らの30歳以上45歳未満の中堅臨床医310名を対象にしたインターネット調査によると, 臨床研究に関心を持つ医師は85%, 教育の受講希望者は78%であった⁹⁾。教育は臨床研修終了後にセミナー・研修会の受講生として受けたいとの意見が多く, ある程度の臨床経験を得たのち, 臨床医の仕事

を中断することなく学べる環境が望まれていた。学位(博士)取得希望者の90%が受講を希望していたが, 学位取得に興味のない医師やすでに学位を取得している医師の75%も受講を希望していた。井手野らが, 群馬県内の勤務医と開業医595名を対象に行ったアンケート調査によると, 統計解析(70%), 研究計画(10%), 医学英語(9%)について学ぶ機会を求める意見とともに個別相談の必要性についても明らかになっている⁹⁾。また, 時間がない, 指導者がいないなどの研究推進を妨げる環境因子についても指摘された。研修を受けたい内容は「臨床研究や疫学データの解釈」が最多で, 研修を受けるのに好まれる時間帯は土曜日であった。Takahashiらが「新医師臨床研修制度の評価に関する研究班」において1124名の研修医を対象に実施した調査によると, 学術活動経験は臨床研修の満足度増加に関連していた⁷⁾。学術活動を妨げる因子として時間不足, 指導医がいない, 支援がないことが指摘された。これらの実態調査から, 臨床医の臨床研究に対するニーズは高いことが示唆される。研修病院では, 生涯教育プログラムの一つとして医師の臨床研究参加への機会提供や臨床研究研修会の開催が求められるであろう。

わが国の研修病院における 研究教育の取組み

大学院の取組みとして京都大学大学院医学研究科社会健康医学系専攻臨床研究者養成コース(2000年), 東京大学大学院医学系研究科公共健康医学専攻(2007年), 大阪大学大学院医学系研究科医科学修士社会人コース(2007年)長崎大学大学院国際健康開発研究科国際健康開発専攻(2008年)等で人材育成が始まっている。しかし, 多忙な臨床医が年単位で職場を離れて大学院で学ぶのはハードルが高い。そこで, 現場の実情にあった臨床研究教育のあり方を考えるため, 先駆的な取り組みを行っている大学以外の研修病院の実情を紹介する。

聖路加国際病院は, 2003年に全職員の臨床研究を促し支援する目的で臨床実践研究推進センター(現:臨床疫学センター)を開設した³⁾。スタッフは日本人医師5人(MPH取得者4人), 米国人

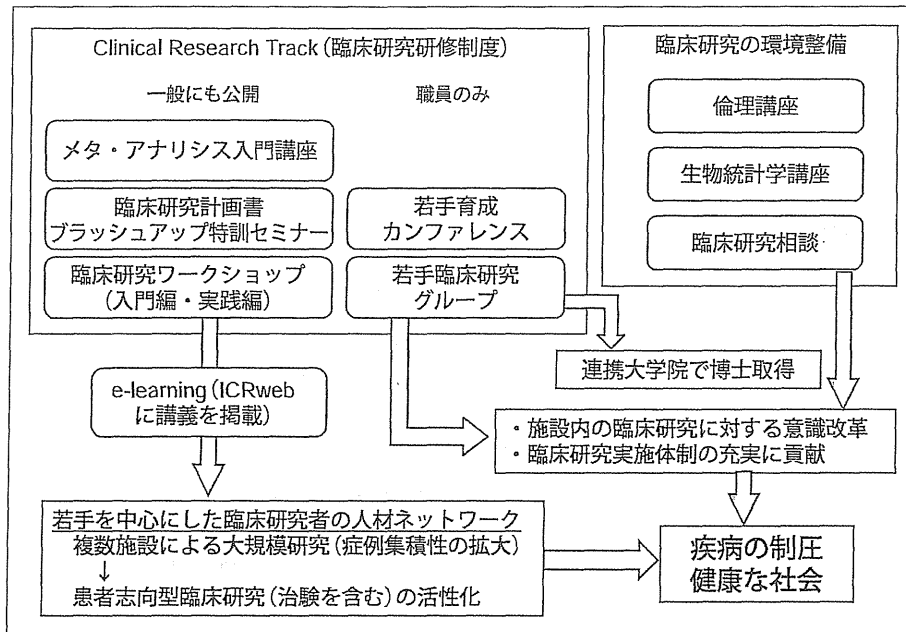


図1 ●●●●

医師1人，看護師1人，研究員1人，事務員1人の9人である。その内容は，個別臨床研究の支援・メンタリング(研究デザイン，統計学的方法，プレゼンテーション方法，論文作成サポート)，専門研修医や医療従事者の受け入れ，臨床疫学や生物統計学に関する勉強会やワークショップ，公衆衛生大学院MPH留学相談，独自の臨床研究プロジェクトを実施している。講義としては，1・2年次研修医を対象にした臨床研究勉強会を20回(臨床疫学と生物統計学を各10回)行い，2年次は研究計画を自ら立ててそれを臨床研究として実践することが義務づけられている。2年次研修医は6月までにテーマに関する専門性を有する臨床指導医1人と研究指導者として臨床疫学センターのメンバーから1人を決定し，8月までに研究計画書作成，9月に倫理審査委員会に提出承認を受けたのち，3か月間かけてデータ収集・分析を行い，12月に抄録提出，1月に学会形式の口演発表を行っている。発表会では厳密な審査が行われ最優秀者は表彰されるシステムである。研修医時代に臨床と研究の両面でメンタリングを受ける体制が整備されており，大変魅力的である。

国際医療研究センターでは，2004年に臨床医

のオプション研修コースの一つとして臨床研究研修コースが開始された。総合診療科をローテーションする研修医を対象にした6回の研修医コースにおいて，EBMの意義と実践，臨床疑問の解決方法，疫学の基礎と応用，データマネジメント，医療倫理と倫理審査，文献検索の講義を受けることになっている。その到達目標は，臨床疑問を自ら研究という形で答えを見出せるためのスキルを養うこととしている。レジデントに対しては，3か月間病棟勤務を離れて集中的に臨床研究の方法論を学び，スキルを高める集中研修コースが準備されている。この期間に自らの研究のデータ解析，論文作成などを一気に進めることができる。その他，統計学セミナーや統計ソフトウェアの操作に関する講習会が開催されている。

国立精神・神経医療研究センターでは，先の臨床研究に関する倫理指針改正に対応する形で，2009年に臨床研究に関するセミナーを16回実施した。2010年には若手研究者ならびに研究に携わる全医療者の教育・研修を促進するためにClinical Research Track(CRT：臨床研究研修制度)を開始した⁹⁾。専任スタッフは臨床疫学を学んだ医師2人，生物統計家1人(2014年から欠

員), 事務員1人である。臨床研究に興味をもつ医療者がどの段階からでも, 他医療機関からでも参加できるよう, さまざまなプログラムを徐々に整えていった(図1)。「若手臨床研究グループ」では, 競争的外部資金を獲得する一歩手前の医療者による萌芽的PORの研究提案を毎年10課題程度採択し, 運営費交付金を財源とする少額の研究費を助成し, 専任スタッフ2名が研究計画から論文執筆までの2~3年間を一貫指導する体制を築いた。グループ代表者は年に8回開催される「若手育成カンファレンス」において研究の進捗状況や成果を発表する。「臨床研究ワークショップ」は夏期の入門編と冬期の実践編を準備, いずれも他医療機関の希望者も参加可能な形で開催している。入門編では臨床疑問からPECO/PICO (Patient, Exposure/Intervention, Comparison, Outcome)の作成と研究倫理の基本原則を理解すること, 実践編では将来PIを目指す者を対象に臨床試験あるいは観察研究のプロトコル作成を目標に, それぞれ少人数による演習と模擬ピアレビュー委員会を行う。ワークショップの総受講者数は272人になった。「臨床研究計画書ブラッシュアップ特訓セミナー」は受講者が文部科学省・学術振興会の科学研究費に応募できる水準の研究計画書を実際に完成させることを目標とした個人塾のような形式で実施しており, すでに成果を上げている。最近, メタ・アナリシス入門講座等も開始しPORの統合研究推進にも力を注いでいる。

おわりに

高木兼寛が兵食改善で脚気を予防した洋上実験をLancetに報告⁹⁾してから128年が経過した。高木の研究こそPORであったではないか。高木はビタミンB₁欠乏という原因解明までは至らなかったが, 日清・日露戦争における海軍の脚気死亡者を激減させることに成功した。現代においても, 目の前で苦しむ患者の丁寧な観察から疾患の新たな診断法・治療法のヒントが見つかるに違いない。将来の標準医療に貢献するためにはPORの推進が必須である。医学における基礎研究の貢献度から察しても, 日本人は欧米人と比較して遜色ない研究力を持っているはずで

ある。現場医師の臨床研究に対する興味と関心は高いし, 政府も人材育成の重要性を指摘している。今こそ, アカデミアが臨床研究を基礎研究と同等に評価する文化を育むチャンスである。全国の臨床研究中核病院, グローバル臨床研究拠点, そして国立高度専門医療研究センターとMPHを授与する大学院大学が連携し, それらを中心としたオールジャパンの研究環境を構築するとともに, 研究マインドを持つ一人一人の医療者が世界をリードするPIとして成長することを積極的に支援することが期待される。

謝辞: 本稿執筆に際し, 国立精神・神経医療研究センターのCRTとともに作り上げてきた同志の中川敦夫氏(現・慶應義塾大学医学部クリニカルリサーチセンター講師)から助言を得た。筆者はCRTの立ち上げに際して, 新保卓郎先生(前・国立国際医療研究センター医療情報解析研究部長, 筆者が研修医1年次の指導医)の温かい支援を受けた。

文 献

- 1) 藤原康弘. トランスレーショナルリサーチを成功させる秘訣—臨床研究のインフラストラクチャー整備. 医学のあゆみ 2002; 200: 544.
- 2) Fleming M, Burnham EL, Huskins WC. Mentoring translational science investigators. JAMA 2012; 308: 1981.
- 3) 高橋 理, 大出幸子, ゴーダム・A・デシュパンデ, ほか. 研修病院における臨床研究とその支援体制. 日本内科学雑誌 2010; 99: 1662.
- 4) 辰巳邦彦. 主要基礎・臨床医学論文掲載数の国際比較. 政策研ニュース 2012; 35: 48.
- 5) 三品浩基, 横山葉子, 川上浩司, ほか. 臨床医を対象とした臨床研究への関心および教育のあり方についての調査: 中堅臨床医の臨床研究への関心は高いか? 医学教育 2009; 40: 105.
- 6) 井手野由季, 菊地麻美, 田村遵一, ほか. 医師の生涯教育のニーズ—臨床研究と研修に関して—. 医学教育 2014; 45: 349.
- 7) Takahashi O, Ohde S, Jacobs JL, et al. Residents' experience of scholarly activities is associated with higher satisfaction with residency training. J Gen Intern Med 2009; 24: 716.
- 8) 松岡 豊, 中川敦夫, 中林哲夫: 臨床研究の教育