

特集 EBMの実践と画像診断・IVR研究のストラテジーⅡ

総説

IVRのエビデンスを創るための研究デザイン

曾根 美雪<sup>1)</sup>、江原 茂<sup>1)</sup>、荒井 保明<sup>2)</sup>、小林 健<sup>3)</sup>

<sup>1)</sup> 岩手医科大学 放射線科

<sup>2)</sup> 国立がんセンター中央病院 放射線診断部

<sup>3)</sup> 石川県立中央病院 放射線診断科

Study designs of clinical trials to establish the evidences of interventional radiology

Miyuki Sone<sup>1)</sup>, Shigeru Ehara<sup>1)</sup>, Yasuaki Arai<sup>2)</sup>, Takeshi Kobayashi<sup>3)</sup>

<sup>1)</sup> Department of Radiology, Iwate Medical University

<sup>2)</sup> Division of Diagnostic Radiology, National Cancer Center

<sup>3)</sup> Department of Diagnostic Radiology, Ishikawa Prefectural Central Hospital

抄録

IVRが標準的治療として広く認知され活用されるためには、“相”(phase)の概念に則った段階的な臨床試験による科学的な評価を行い、IVR以外の医師でも容易に納得できる明確なエビデンスを示す必要がある。薬物の臨床試験手法に倣って臨床試験を遂行するとともに、既存の研究手法ではカバーできないIVRに特有の問題点を明らかにして、これに対応する手法を確立することが、IVR医に求められている。

Abstract

To establish a standard treatment in interventional radiology, adequately high-level evidence established by the scientifically robust clinical trials are mandatory. Interventional radiologists play an important role not only in accomplishing clinical trials with existing methodology, but also in developing a new methodology corresponding to the characteristics of interventional procedures.

**Key words** | Clinical trial, randomized controlled trial, interventional radiology

はじめに

医療がEBM (evidence-based medicine) の時代に突入してから、すでに約20年が経過しようとしている。しかしEBMに則ったIVRを実践しようとすればする程、IVRのエビデンス不足に悩まされる

のが現状である。他方、エビデンスを創る機運も高まっており、IVRの臨床試験の施行数は増加しているものの、その内容は必ずしも十分とは言えない。臨床試験の推進には様々な障壁が存在するが、IVRの臨床試験デザインが確立しておらず、前例が少な

別刷請求先：〒020-8505 盛岡市内丸19-1

岩手医科大学放射線医学講座 曾根 美雪

TEL：019-651-5111 (内線3660) FAX：019-651-7071

いこともその一因と思われる。本稿では、IVRの臨床研究、とくに臨床試験の研究デザインについて、現状と方法論について述べる。

### 1. IVRにおける臨床研究の種類と現状

臨床研究は、目的によって、臨床試験、予後因子研究、疫学研究、妥当性研究に大別される。IVRは、その手技のほとんどが治療であり、治療法の開発に用いられる臨床試験、特にランダム化比較試験が中心となるべき領域である。臨床試験においては、研究実施前に研究計画書（プロトコル）を作成し、これに則って前向きに治療およびその評価が行われる。しかし、実際のIVRの研究は、現時点では、後向きの観察研究が多数を占める。表1に、2008年にJournal of Vascular and Interventional Radiology (JVIR)に掲載された論文のうち、臨床研究として採択された掲載論文の研究デザインの内訳を示す<sup>1)</sup>。136の論文中、前向き研究が29%、後向きが67%で、最も多い研究デザインは後向きのケース・シリーズ(63%)であった。ランダム化比較試験は、わずかに8論文(6%)で、JVIRがIVRのメジャーな雑誌であることを考えると、IVRの臨床研究において臨床試験は市民権を得ておらず、その現況に国際格差はないといえる。もっとも、これらの後向きケース・シリーズの論文は、われわれIVRの臨床医にとって参考になる内容で、有害事象の頻度や種類、手技の写真など、貴重な情報が含まれて

いる。また、IVRが対象とする病態は多岐にわたり、緊急や全身状態の悪い患者が対象であることも多いため、全てのIVRについてランダム化比較試験を行うのは難しく、後向きの研究が多くなるのが止むを得ない部分もある。しかし、evidence-based medicine (EBM)の観点から見れば、どんなに努力をしても後向きの研究には多数のバイアスが混入するため、“エビデンス・レベルの低い”研究に分類されてしまうことも、また事実である。研究の理想は、臨床的な疑問に対してできるだけ真実に近い答えを得ることであるが、臨床研究においてはさまざまなバイアスが存在するため、EBMではバイアスを少なくする研究デザインがエビデンス・レベルが高いとされている<sup>2)</sup>(表2)。われわれは、臨床医であると同時に科学者であり、眼前の患者に最良の医療を提供するためには、診療と同時に、エビデンス・レベルの高い研究を行う姿勢も堅持すべきであろう。さらに、治療法の選択肢にIVRを加えるかどうかは、他領域の専門医が決定する場合が多く、IVR以外の分野の専門家にも受け入れられるような明確なエビデンスの提示を心がける必要がある。

### 2. 臨床試験における“相”(phase)の概念

治療法開発における各段階のことを“相”とよび、I-IV相からなる。一般に、第I相試験では安全性のスクリーニング、第II相試験では有効性のスクリーニングを行い、第III相試験にすすめる治療法を選択

表1. 2008年にJVIR\*に掲載された臨床研究のデザイン (n=136)

研究デザイン	論文数	%
<b>前向き</b>	<b>39</b>	<b>29</b>
ランダム化比較試験	8	6
非ランダム化比較試験	3	2
第II相試験、第I/II相試験	2	2
レジストリー研究	1	1
観察研究	7	5
その他	18	13
<b>後向き</b>	<b>91</b>	<b>67</b>
ケース・シリーズ研究	86	63
コホート研究	2	2
横断研究	3	2
<b>メタ・アナリシス</b>	<b>6</b>	<b>4</b>

表2. エビデンスのレベル<sup>2)</sup>

I a	複数のランダム化比較試験のメタ分析による
I b	少なくとも1つのランダム化比較試験による
II a	少なくとも1つの非ランダム化比較試験による
II b	少なくとも1つの他の準実験的研究による
III	コホート研究や症例対照研究、横断研究などの分析疫学的研究による
IV	症例報告やケース・シリーズなどの記述研究による
V	患者データに基づかない、専門委員会の報告や権威者の意見による

\*JVIR: Journal of Vascular and Interventional Radiology

する。第III相では、第I相、第II相試験で有望とされた新しい治療法を、既存の標準的治療と比較し、“勝った”治療法が、“標準的治療”となる。第IV相試験は、新しい治療法の長期の安全性と有効性を評価する。

臨床試験における“相”の概念は、薬物治療において確立し、研究手法が洗練されてきた。IVRでは手技が重要な要素であり、薬物治療と同じ研究手法をそのまま適用することは困難であるが、歴史があり現在も研究の中心をなすこれらの手法を理解し、できるだけ近づけることが望ましい。

### 3. 臨床試験のデザインの基本

臨床試験のデザインにおける主要な基本原則を下記に示す<sup>3)</sup>。

- ・ 目的を明確に記述すること。
- ・ 適格条件、治療内容、エンドポイントを特定すること。
- ・ 検出すべき差の大きさもしくは推定の精度を決定すること。
- ・ (ランダム化比較試験の場合) 治療法の割り付けをどのように実施するかを特定すること。
- ・ サンプルサイズの計算に用いる、分布の仮定やエラーの確率を決定すること。

以下に、研究デザインの基本構成要素と、その注意点について述べる。

#### (1) 仮説と目的の設定

日常臨床から湧き起こる臨床的疑問をもとに、研究における仮説および目的を設定する。文献検索を行い、同様の先行研究があるかどうか、またその結果やエビデンス・レベルを考慮して、研究施行の可否を決定する。仮説は、臨床医の視点から重要と思われる疑問に対して、明確な答えが得られるように設定する。さらに、得られる結果を次のステップでどのように活用するのかを想定しておく必要がある、それにより試験デザインの方向性が決まる。第I相試験、第II相試験においては、第III相試験に進めるに値する治療法を選別することが目的となる。第III相試験では、新しい治療法が既存の標準的治療と比べて優れており新たな標準的治療法となる

ことを証明するか、新治療が既存の標準的治療と同程度の有効性をもつことを証明することが目的となる。臨床研究のゴールは、日常臨床における判断の根拠となるエビデンスを創ることであり、仮説の設定においても臨床医の視点が欠かせない。また、研究を実施するのも多忙な臨床医自身であり、研究の実行可能性も考慮する必要がある。よい研究テーマの条件として、Hulleyらが提唱するFINER基準を表3に示す<sup>4)</sup>。

仮説は、帰無仮説 (null hypothesis) あるいは対立仮説 (alternative hypothesis) の形で記述される。帰無仮説とは、新治療と対照となる標準治療のアウトカムに差がないとするもので、対立仮説は、逆にこれらに差があるとするものである。通常、臨床試験では帰無仮説を仮説として採用し、これが否定される (棄却される) ことにより対立仮説を証明する形をとる。

#### (2) エンドポイント

エンドポイントは、治療によるアウトカムを評価するための指標で、患者の利益ないしは不利益を測る物差しである。仮説をもとに複数のエンドポイントを設定するが、その中で最も重要な項目 (通常一つ) を主要評価項目 (primary endpoint) とし、その他の項目を副次評価項目 (secondary endpoint) とする。エンドポイントには、真のエンドポイントと代替エンドポイントがあり、真のエンドポイントは、患者が自覚でき、臨床的に意味のあるエンドポ

表3. よい研究テーマが満たすべきFINER基準<sup>4)</sup>

<b>Feasible</b> (実施可能性)	対象者数が適切である 研究を実行できる専門性を有している かかる時間や費用が適切である 複雑すぎない
<b>Interesting</b> (真の興味)	研究者自身が真の興味を有する
<b>Novel</b> (新規性)	過去の知見を確認もしくは否定する 過去の知見を発展させる 新しい知見を加える
<b>Ethical</b> (倫理性)	
<b>Relevant</b> (必要性)	科学の進歩に貢献する 臨床医学や保険政策に貢献する 将来の研究の発展に貢献する

イントで、6Ds:Death (死亡)、Disease (疾病)、Discomfort (不快)、Disability (機能障害)、Dissatisfaction (不満)、Destitution (貧窮) で表わされることがある<sup>2)</sup>。

第I相試験では、通常、毒性(安全性)がエンドポイントとして用いられ、第II相試験では、真のエンドポイントの代りとなる代替エンドポイント(surrogate endpoint)(例:抗がん剤の試験で、画像で腫瘍の縮小程度を測定し算出する奏効率)が用いられることが多い。代替エンドポイントは、真のエンドポイントの評価には多くの患者数と長い研究期間を要することから、妥当性を担保しながらも、できるだけ短い期間で効率よく第III相試験に進める価値のある治療法かどうかを決定するために用いられる。第III相試験においては、真のエンドポイントを用いる。

### (3) 患者選択規準

適格条件(組み入れ条件、inclusion criteria)と除外条件(exclusion criteria)により、どのような患者を対象に研究を行うかを示す。適格条件は、研究の結果そのIVRが標準的治療となる場合に、最も利益を受ける対象集団を想定して決定する。除外条件は、その臨床試験に好ましくない影響を及ぼすことが予測される症例を除外するための条件として決定する。このため、除外条件は、その治療法が標準的治療となった場合に必ずしも適応外とされる条件ではない。すなわち、臨床試験に

おける症例選択規準は、研究内容の明確化と実行性とともに、研究結果の日常診療への外挿における「一般化可能性」を十分に考慮して決定される必要がある。

### (4) 治療内容

IVRの方法や用いるデバイス、薬剤を、再現性、一般化可能性を考慮しながら決定する。IVRでは、手技が大きな比重をもつため、使用する機材や機器、あるいは術者の経験や巧拙などのばらつきが結果に影響を及ぼす。このため、特に多施設共同試験においては、計画段階で手技を最適化する必要がある。“できるだけ安全に、できるだけ効果の高いIVRを行う”という原則に基づいて、無視できない手技のばらつきはコンセンサスを得て標準化する一方、個々のIVR医の裁量にまかせて良い部分は自由度をもたせるなど、バランスのとれた標準化を行うことが重要である。

### (5) 測定方法

エンドポイントの測定方法および規準の例を表4に示す。IVRの特殊性、新規性のため、研究独自の基準を用いざるを得ない場合もあるが、可能な限り、妥当性が検証された測定方法を用いる。また、誰が測定するのか、いつ測定するのかといった細かい点も、バイアスに直結するため明確に規定する。測定に伴うバイアスを減少するための方法として、第三者による中央評価や、患者が記入するQOLスコアをIVR医に

表4. エンドポイントの測定方法および規準の例

エンドポイント	測定方法	規準
腫瘍縮小効果	CT、MRI	RECIST (*1)
頸動脈狭窄度	血管撮影、CT、MRI、超音波	NASCET (*2)
有害事象の頻度と種類	自覚症状、血液検査、その他	CTCAEv3.0 (*3) SIR (*4) classification
QOL	患者自身によるスコア記入	妥当性の検討されたスコア(例)を用いて群間、治療前後等を比較 例: EQ-5D (*5) SF-36 (*6) EORTC (*7) QLQ-C30

\*1. RECIST : Response Evaluation Criteria in Solid Tumors

\*2. NASCET : North American Symptomatic Carotid Endarterectomy Trial

\*3. CTCAE : NCI Common Terminology Criteria for Adverse Events

\*4. SIR : Society of Interventional Radiology

\*5. EQ-5D : EuroQoL

\*6. SF-36 : Medical Outcomes Study 36-Item Short Form-36

\*7. EORTC : European Organization for Research and Treatment of Cancer

ブラインドにする、等が例として挙げられる。

(6) 必要な症例数 (サンプルサイズ) の計算

サンプルサイズの計算は、研究デザインの中でも重要な要素である。最も大切なステップは、臨床的に重要な差を推定し、これを十分に検出できる症例数を設定することである。サンプルサイズ計算に必要な項目は、下記の4項目である。

A. 効果量 (effect size)

研究で証明しようとしている、治療のアウトカムの見積もりを効果量という。効果量は、主要エンドポイントについて、新治療および対照のアウトカムを先行研究の論文等から検索したデータをもとにして決定する。第Ⅲ相試験では、先行研究で得られている標準治療の効果量と新治療の効果量の差がどのくらいであれば、“臨床的に意味がある”とするかを決める。この差が小さいほど、必要症例数は多くなる。第Ⅱ相試験では、ここまでなら許容できるという下限の閾値となる効果量を設定する。

B. データの測定精度

結果のばらつきは、サンプルサイズに影響する。ばらつきが大きい=測定精度が低いほど、データの変動が大きくなり、新治療と対照の二群でのオーバーラップが増えるため、必要な症例数は増加することになる。測定精度の高い方法を用いることにより、サンプルサイズは減少する。

C. 第Ⅰ種の過誤 (type I error :  $\alpha$ )

本当は治療の結果に差がないのに、差がある(偽陽性)としてしまう確率である。たとえば、「下大静脈フィルターが肺血栓栓塞症予防に与える影響をみる研究」を行うと仮定して、 $\alpha=0.05$  (5%)と設定すれば、“本当は下大静脈フィルターの有無による差がないのにあると判定してしまう確率”を、研究者が最大5%まで許容するという意味する。 $\alpha$ は、小さいほどより判断が厳格であることになる<sup>4)</sup>。

D. 第Ⅱ種の過誤 (type II error :  $\beta$ )

本当は治療の結果に差があるのに、差がない(偽陰性)としてしまう確率で、 $(1-\beta)$ は研究の検出力(パワー)を表す。例として、「80歳以上の圧迫骨折患者に対して椎体形成術を施行すると疼痛スコア(VAS)が4以上低下するという研究」を0.90 (90%)の検出力で行うと仮定すると、同じ研究を10回行ったら9回は差がある(VASが有意に低下する)

という結果になることを意味する。すなわち、検出力は、母集団の中から対象となる集団を選んで研究を行う際に、研究者が臨床的に意味あると考える差を検出できる確率を表しており、 $\beta$ が小さいほど検出力が高くなる。

$\alpha$ と $\beta$ は研究者自身が設定する数字であり、偽陽性を避けることが重要と考えるなら $\alpha$ を小さく、偽陰性を避けることが重要なら $\beta$ を小さくすればよいが、これらは互いにトレードオフの関係にあるため、同時に小さくすることはできない。また、 $\alpha$ 、 $\beta$ を小さくすると必要症例数が増加し、研究の実行可能性に影響をおよぼす。通常、 $\alpha$ は0.05、 $(1-\beta)$ は0.80-0.90に設定されることが多い。

E. 片側検定と両側検定

片側検定は、IVRによるアウトカムが一方のみに向かうと仮定するもので、例えば腎血管性高血圧症の患者に対する腎動脈ステント留置術において、“ステントを留置すると、血圧は低くなるが高くなることはない”と考えての検定である。いっぽう、両側検定は、“ステントを留置すると、血圧は低くなることも高くなることもある”と考えての検定である。両側検定に比べて片側検定の方が、差が出やすくサンプルサイズは小さくなる<sup>5)</sup>が、統計学的有意差ありとの解析結果を受け入れてもらうためには両側検定が必要とされている<sup>2)</sup>。

サンプルサイズを見積もる具体的なプロセスは、下記である<sup>4,5)</sup>。

- ① 仮説を設定する
- ② 仮説に含まれる予測因子(治療の有無等)とアウトカムの変数のタイプにより、適切な統計学的検定方法を選択する(表5)。

表5. サンプルサイズの推定に使用する単純な統計学的検定法<sup>4)</sup>\*

予測因子	アウトカム	
	二区分変数	連続変数
二区分変数	カイ二乗検定**	t検定
連続変数	t検定	相関係数

\* カテゴリー変数や量一反応関係を検討する場合は、本表には合致しない

\*\* カイ二乗検定は、常に両側。t検定の場合は、片側とするか両側とするかを決定する必要がある。

- ③適切と思われる効果量を設定する。
- ④アウトカムの変動度（測定精度）を見積もる。
- ⑤  $\alpha$ 、 $\beta$ の値を設定する。
- ⑥成書にみられる表や、サンプルサイズの計算を支援するウェブサイトを用いて、サンプルサイズを決定する。
- ⑦フォローアップから脱落する症例数の見積もり、中間解析の有無による増減などにより、最終的なサンプルサイズを決定する。

#### 4. ケース・スタディ

悪性椎体骨腫瘍患者に対する経皮的椎体形成術（骨セメント）の研究を例に、実際にIVRの臨床試験研究グループ（日本腫瘍IVR研究グループ Japan Interventional Radiology in Oncology Study Group: JIVROSG）にて行われた臨床試験の実例を説明する。研究の目的と方法を詳細に記した研究計画書は数十ページに及ぶため、ここではその概要を説明する。

##### 臨床的疑問

経皮的椎体形成術は、骨転移をはじめとする椎体の悪性腫瘍患者の疼痛緩和に有用なのだろうか？

##### 研究のストラテジー

研究が立案された時点（2002年）で過去の文献を検索すると、多数の後向きケース・シリーズがあり、いずれも椎体形成術は有用と報告されていた。しかし、前向き研究はみられなかったため、安全性評価から順にエビデンスを積み上げる必要があった。そこで、まず第I相試験による安全性評価、第II相試験による有効性評価（JIVROSG-0202<sup>6)</sup>）を行い、ランダム化比較試験に進め得る治療法かどうかを前向きに検討し、安全性と有効性が確認されたなら、第III相試験による標準治療とのランダム化試験を計画することとした。

##### (1) 第I/II相試験による安全性と有効性の評価：JIVROSG-0202<sup>6)</sup>

試験デザイン：多施設共同I/II相臨床試験

主要エンドポイント：安全性の評価。

副次エンドポイント：臨床的有効性の評価。有害事象の発現頻度と程度。

患者選択規準：

適格条件

- 1) 悪性腫瘍の転移や原発性椎体腫瘍による胸椎または腰椎の病変と診断されている。
- 2) 1) による疼痛が強く日常生活の行動に制限が生じている患者であるか、転移による圧迫骨折の危険性から運動制限を医師から指示されている症例。
- 3) 椎体腫瘍が脊柱管に露出しない症例。
- 4) 主要臓器（骨髄、心、肝、肺、腎など）機能が保持されている症例。
- 5) P.S. (ECOG)：0, 1, 2, 3。
- 6) 4週間以上の生存が見込める。
- 7) 患者本人から文書による同意が得られている。

治療内容：

X線透視下またはCT透視下に骨セメント注入針を経皮的に椎体へ刺入し骨セメントを注入する。

症例登録は、はじめの9例までは3例を1段階として4週間の観察期間を設けて重篤な有害事象（有害反応）の発生頻度が30%以下であることを確認できた後、次段階に進む。この方法で順次登録された3段階9例の有害事象（有害反応）発生頻度が30%以下である場合には、10例目以後は段階なしに予定登録数まで症例を登録する。

測定方法：

安全性：有害事象の評価にはNCI-CTC (NCI-CTC 日本語訳 JCOG-第2版（注：現在はCTCAEv3.0が用いられる）)を使用する。

有効性：治療終了後1週目の疼痛により、以下の3段階に評価する。

治療前のVAS値と治療1週目のVAS値の比較において、

- 1) 著効：治療後VAS値が0-2となる。または、治療前より5以上低下している。
- 2) 有効：治療前からの低下が2以上5未満である。
- 3) 無効：上記以外の場合。

サンプルサイズ：

重篤な有害事象については、“合併症の予測値を10%とし、30%以上なら試験中止”とした場合、 $\alpha=0.05$ 、 $\beta=0.20$ で必要症例数は30となる。一方、有効率については、“有効性の予測値を少なくとも

50%、80%以上であれば意味のある有効性”とし、 $\alpha=0.05$ 、 $\beta=0.20$ とすると、必要症例数は19となる。よって、本試験では必要症例数のより大きな、重篤な有害事象評価に必要な症例数である30例に、10%のプロトコル逸脱等を見込んで、予定登録症例数を33例とする。

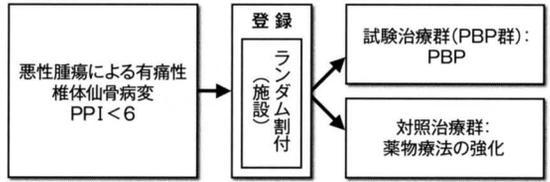
解説

IVRの第I相試験の研究方法論は、2002年にプロトコルを作成する時点では存在しなかった。薬物では“段階的に増量して安全な用量を見つける”手法が用いられることが多いが、IVRでは増量という概念は適用できない。しかし、少なくとも小人数の患者に治療を行い、安全性を確かめながら段階的に評価する手順は必須と考えられたため、“JIVROSG 3×3 method<sup>6)</sup>”を提案した。これは、少数(通常3例)の患者の治療を行って経過観察期間を設け、安全性を確認したうえで次の少数の患者を段階的に登録する方法である(図1)。

JIVROSG-0202の結果は、有効率70%、IVR施行から効果発現までの期間の中央値は1日で、手技に関連した重篤な有害事象は認められず、安全で早期疼痛緩和効果が得られる治療法と判断された。

(2) 第Ⅲ相試験による有効性の検証：JIVROSG-0804 (研究実施中)

試験デザイン：多施設共同第Ⅲ相ランダム化比較試験  
 主要エンドポイント：背部疼痛症状の推移



PPI : palliative prognosis index, PBP : percutaneous bone plasty

副次エンドポイント：包括的QOLの推移、背部痛QOLの推移、有害事象の内容と頻度、生存期間、局所疼痛緩和生存期間

患者選択規準：  
 適格条件

- 1) 悪性腫瘍の転移や原発性悪性腫瘍による有痛性胸椎、腰椎、あるいは骨盤骨の病変を有する。
- 2) 薬物療法の強化以外には骨形成術を除いて既存の疼痛緩和の方法がない。
- 3) Palliative Prognostic Index (P.P.I.) <6。

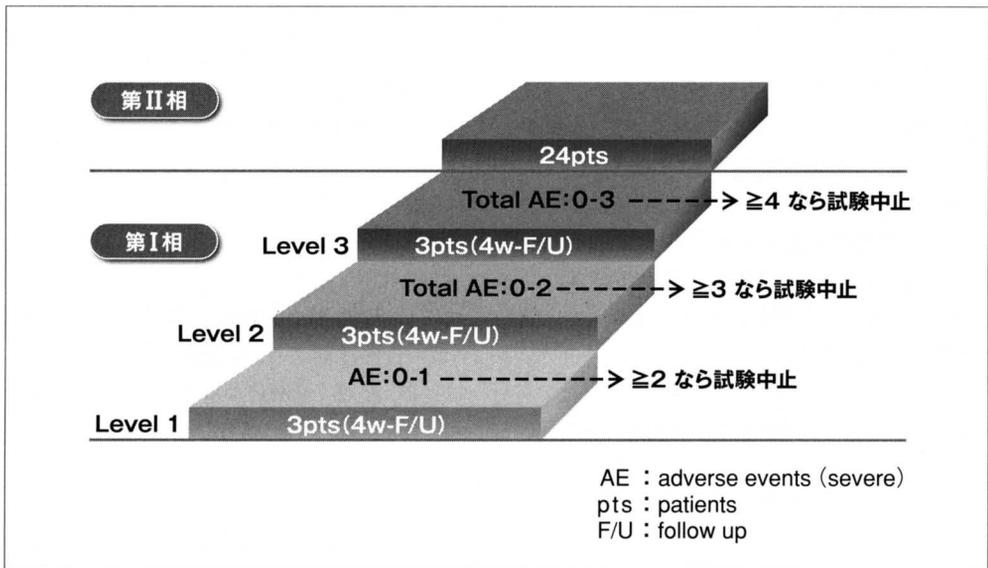


図1. JIVROSG 3×3 method

- 4) 患者本人から文書による同意が得られている。
- 5) 背部痛に伴う症状と包括的QOLが測定されている。

#### 治療内容：

試験治療群にはPBPを、対照治療群には薬物療法の強化を行う。

#### 測定方法：

主要エンドポイント：登録後4週間の背部疼痛症状の推移をVAS (visual analogue scale) で経時的に測定し、これより算出された局所疼痛緩和の総合評価値 (背部疼痛症状測定曲線の曲線下面積：Area Under the Curve (AUC) に該当) を2群間で比較する。副次エンドポイント：包括的QOL (EQ-5DとSF-8) の推移と背部痛のQOL (Roland-Morris Disability Questionnaire：RDQ) の推移、有害事象の内容 (CTCAE v3.0にて評価) と頻度、生存期間、局所疼痛緩和生存期間。包括的QOLについては登録後4週間の総合評価値 (包括的QOL曲線のAUC) を2群間で比較する。背部痛のQOLについてはRDQを用いて最高24点の素点を求め疾患特異的尺度を求め評価する。症状およびQOLの記載は患者本人が行い、説明はIVR施行医以外の医療従事者、送付は患者自身ないしはIVR施行医以外の者が行う。

#### サンプルサイズ：

本試験では、主要エンドポイントである“有痛性悪性椎体、仙骨病変による疼痛症状の総合評価値”について、ベースライン値を共変量とした共分散分析を行い、試験治療群 (PBP群) の対照群に対する優越性を検証する。

JIVROSG-0202の結果、これまでの経皮的椎体形成術に関する報告、ならびに有痛性骨転移の疼痛緩和の臨床的現状より、主要エンドポイントの総合評価値の差は約30、標準偏差は約40と推測した。総合評価値の分布に正規分布を仮定し、両側有意水準10%の下で、1群あたりの症例数を18例とすれば、検出力70%を確保できる。早期脱落例を考慮し、1群あたりの症例数を20例、両群合わせて40例と設定した。

#### 解説

本試験は、緩和の領域においてIVRという侵襲性を伴う治療法を試験群とするランダム化比較試験であり、あまり前例のない臨床試験としてもチャレンジなものと言える。目的は、既存の方法である手術療法ないしは放射線治療の適応とならない患者や、これらの治療を行っても疼痛が緩和しないもしくは再燃した患者を対象として、薬物療法の強化と比較して経皮的骨形成術が有用かどうかを評価することとしている。すなわち、“経皮的椎体形成術ないしは骨形成術は、悪性骨腫瘍の患者の疼痛緩和における標準的治療となり得るか”という臨床的疑問に答えを出す研究と位置付けることができる。

対象は、予後の限られたがん終末期の患者であることから、倫理的配慮より、任意の時期に患者の意思により割付けられた治療法を中止することができ、かつ、治療法の交差 (クロスオーバー) を許容することとした。背部疼痛による症状と包括的QOLについて、割付治療継続期間中 (最長4週間) の総合評価値を比較するデザインを用いた。

ランダム化は、試験治療と対照治療に患者をランダムに割りつけることにより、結果に影響を及ぼすかもしれないすべての因子 (交絡因子) が、ランダムに双方の群に含まれることを保証する。バイアスを減少するには最も強力な研究デザインであり、治療法の検証には必須の手法である。また、本研究では、疼痛およびQOL評価をIVR医にブラインド化しているが、これは、IVR施行医が説明や聞き取りを行うことで主観が入り込む“測定者バイアス”を減少させるための手法である。さらに、臨床試験における重要な概念である、“intention-to-treat (ITT) 解析”を用いているが、これは、研究に参加した患者はすべて“臨床的に治療を行うと企図した”とみなして結果の解析に含める手法である。臨床試験に同意し参加した患者は、全員が予定された治療をうけるわけではない。中には途中で気が変わって治療をうけない場合もあるし、全身状態の変化により他の治療に変更される場合もあり得る。これらの患者を脱落例として解析から除外すると、バイアスを減少するために前向き検討やランダム化を行った意義が薄れることになってしまう。また、患者な

いしは研究者が、希望しない治療に割りつけられたら治療を中止してしまい、希望する治療に割りつけられたときのみその治療を続けるような“選択バイアス”もあり得るので、ITT解析は必須である。

### おわりに

眼前の患者に最良の医療を提供するために、EBMは欠くことのできない現代医学の基本理念となっ

ている。われわれが行うIVRが真に広く活用されるためには、EBMの方法論に則ってエビデンスを示す必要がある。IVRにおいては技術が大きな部分を占めるため、薬物療法の臨床試験方法論をそのまま適用することは困難である。しかし、長い歴史を経て確立され洗練された薬物療法の臨床試験手法を倣うことは必須であり、さらにIVRに特有な問題点を明確にして研究手法を構築していく必要がある。

### 参考文献

1. Journal of Vascular and Interventional Radiology : web site. <http://www.jvir.org/>
2. 福井次矢, 編. 臨床研究マスターブック. 東京 : 医学書院, 2008.
3. Green S, Benedetti J, Crowley J, ed. 福田治彦, 新美三由紀, 石塚直樹 監訳. 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—. 東京 : 医学書院, 2004.
4. Hulley SB, ed. 木原雅子, 木原正博 監訳. 医学的研究のデザイン : 研究の質を高める疫学的アプローチ. 東京 : メディカル・サイエンス・インターナショナル, 2006.
5. Gallin JI, ed. 井村裕夫, 竹内正博, 藤原康弘, 渡辺亨 監訳. NIH 臨床研究の基本と実際. 東京 : 丸善株式会社, 2004.
6. Kobayashi K, Arai Y, Takeuchi Y, et al. Phase I/II clinical study of percutaneous vertebroplasty (PVP) as palliation for painful malignant vertebral compression fractures (PMVCF) : JIVROSG-0202. Ann Oncol 2009 Jul 1. [Epub ahead of print]

ダウンロードされた論文は私的利用のみが許諾されています。公衆への再配布については下記をご覧ください。

### 複写をご希望の方へ

断層映像研究会は、本誌掲載著作物の複写に関する権利を一般社団法人学術著作権協会に委託しております。

本誌に掲載された著作物の複写をご希望の方は、(社)学術著作権協会より許諾を受けて下さい。但し、企業等法人による社内利用目的の複写については、当該企業等法人が社団法人日本複写権センター（(社)学術著作権協会が社内利用目的複写に関する権利を再委託している団体）と包括複写許諾契約を締結している場合にあっては、その必要はございません（社外頒布目的の複写については、許諾が必要です）。

権利委託先 一般社団法人学術著作権協会

〒107-0052 東京都港区赤坂 9-6-41 乃木坂ビル 3F FAX：03-3475-5619 E-mail：info@jaacc.jp

複写以外の許諾（著作物の引用、転載、翻訳等）に関しては、(社)学術著作権協会に委託致しておりません。

直接、断層映像研究会へお問い合わせください

Reprographic Reproduction outside Japan

One of the following procedures is required to copy this work.

1. If you apply for license for copying in a country or region in which JAACC has concluded a bilateral agreement with an RRO (Reproduction Rights Organisation), please apply for the license to the RRO.

Please visit the following URL for the countries and regions in which JAACC has concluded bilateral agreements.

<http://www.jaacc.org/>

2. If you apply for license for copying in a country or region in which JAACC has no bilateral agreement, please apply for the license to JAACC.

For the license for citation, reprint, and/or translation, etc., please contact the right holder directly.

JAACC (Japan Academic Association for Copyright Clearance) is an official member RRO of the IFRRO (International Federation of Reproduction Rights Organisations).

Japan Academic Association for Copyright Clearance (JAACC)

Address 9-6-41 Akasaka, Minato-ku, Tokyo 107-0052 Japan

E-mail info@jaacc.jp Fax: +81-33475-5619