

産業界のためのガイダンス  
抗がん剤および生物製剤の承認のための  
臨床試験エンドポイント\*

Guidance for Industry  
Clinical Trial Endpoints  
for the Approval of Cancer Drugs and Biologics

米国保健福祉省  
食品医薬品局  
医薬品評価研究センター (CDER)  
生物製剤評価研究センター (CBER)

2007年5月  
臨床 / 医療

U.S. Department of Health and Human Services  
Food and Drug Administration  
Center for Drug Evaluation and Research (CDER)  
Center for Biologics Evaluation and Research (CBER)

May 2007  
Clinical / Medical

訳 西川 昭子<sup>1)</sup> 村山 敏典<sup>2)</sup> 小島 伸介<sup>1)</sup> 都田 桂子<sup>1)</sup>  
監訳 手良向 聡<sup>3)</sup> 福島 雅典<sup>1,3)</sup>

1) (財)先端医療振興財団臨床研究情報センター

2) 京都大学医学部附属病院探索医療臨床部

3) 京都大学医学部附属病院探索医療検証部

\* 本翻訳は、FDAのウェブページに公開された全文を訳出したものであり、翻訳掲載についてFDAの許諾を必要としないものである。

# 産業界のためのガイダンス

## 抗がん剤および生物製剤の承認のための 臨床試験エンドポイント

追加的な複写は、以下の機関より入手可能である。

*Office of Training and Communications  
Division of Drug Information, HFD-240  
Center for Drug Evaluation and Research  
Food and Drug Administration  
5600 Fishers Lane  
Rockville, MD 20857  
(Tel) 301-827-4573  
<http://www.fda.gov/cder/guidance/index.htm>*

or

*Office of Communication, Training, and  
Manufacturers Assistance, HFM-40  
Center for Biologics Evaluation and Research  
Food and Drug Administration  
1401 Rockville Pike, Rockville, MD 20852-1448  
(Tel) 800-835-4709 or 301-827-1800  
<http://www.fda.gov/cber/guidelines.htm>*

米国保健福祉省  
食品医薬品局  
医薬品評価研究センター (CDER)  
生物製剤評価研究センター (CBER)

2007年5月  
臨床 / 医療

## 目 次

I . 序 .....	110
II . 背 景 .....	111
A . 有効性のための規制要求事項 .....	111
B . 腫瘍学における過去の承認を支持するエンドポイント .....	112
III . エンドポイントに関して一般的に考慮すべき点 .....	113
A . 全生存期間 .....	113
B . 腫瘍評価に基づいたエンドポイント .....	113
1. <b>無病生存期間</b> .....	115
2. <b>客観的奏効率</b> .....	116
3. <b>無増悪期間および無増悪生存期間</b> .....	116
a. TTP と PFS .....	116
b. 医薬品承認を支持するエンドポイントとしての PFS .....	117
c. PFS 試験デザインの問題 .....	117
d. PFS の解析 .....	117
4. <b>治療成功期間</b> .....	118
C . 症状評価を含むエンドポイント .....	118
1. <b>特定の症状エンドポイント</b> .....	119
2. <b>症状データが遭遇する問題</b> .....	119
D . バイオマーカー .....	119
IV . 臨床試験デザインの考慮すべき点 .....	120
A . 単群試験 .....	120
B . 非劣性を示すようデザインされた試験 .....	120
C . 放射線療法保護剤および化学療法保護剤のための試験デザイン .....	121
V . 結 論 .....	121
付録 1 : 腫瘍測定データの収集 .....	122
付録 2 : PFS 解析で考慮すべき問題 .....	123
付録 3 : PFS 解析のための表の例示 .....	125
付録 4 : 腫瘍エンドポイントの独立審査 .....	127

拘束力のない勧告を含む

## 産業界のためのガイダンス<sup>1</sup>

### 抗がん剤および生物製剤の承認のための臨床試験エンドポイント

本ガイダンスは、このトピックに関する現在の食品医薬品局(FDA)の考えを表している。これは、いかなる人に対しても、または、関しても、いかなる権限も生じさせることはないし、または与えることもなく、FDAまたは一般市民を拘束するようなこともない。本ガイダンスに代わるアプローチが適用される法令と規制の要求事項を満たす場合には、そのアプローチが使用される可能性がある。代替的アプローチの議論が必要な場合、本ガイダンスの遂行に責任があるFDAの職員に連絡されたい。適当なFDAスタッフがつかまらない場合は、本ガイダンスのタイトル・ページに記載されている電話番号に連絡されたい。

#### I . 序

本ガイダンスは、計画中の一連の癌のエンドポイント・ガイダンスで最初のものである。これは、新薬申請(NDA)、生物学的製剤認可申請(BLA)または補足的な申請における有効性の主張を支持するために、食品医薬品局(FDA)に提出される癌臨床試験のためのエンドポイントに関して、申請者に対する勧告を提供するものである。<sup>2</sup>また、本ガイダンスでは、背景となる情報を提供し、一般的な規制原則についても議論する。本ガイダンスで議論されているエンドポイントは、癌をすでに持つ患者を治療する医薬品に関するものである。本ガイダンスは癌の発生を抑制、または低下させる医薬品のためのエンドポイントに関するものではない。

FDAは、FDAのOncologic Drugs Advisory Committee(ODAC)の前に公的なワークショップおよび議論を含む過程を通して、腫瘍学エンドポイントに関するガイダンスを開発している。<sup>3</sup>以降の各々のガイダンスは、医薬品の承認または医薬品の効能表示を支持するべく、特定の癌種(例えば肺癌、大腸癌)のためのエンドポイントに焦点を合わせるであろう。

本ガイダンスを含めたFDAのガイダンス文書は、法的に強制できるような実施責任を規定するものではない。その代わりに、特定の規定、または法的要求事項が引用されない限り、ガイダンスはあるトピックに関する現在の当局の考えについて説明するもので、単に勧告としてのみ見なされるべきである。当局のガイ

<sup>1</sup> 本ガイダンスは、食品医薬品局生物製剤評価センター(CBER)と協力して、医薬品評価研究センター(CDER)のDivision of Drug Oncology ProductsおよびDivision of Biologic Oncology Productsによって準備されたものである。

<sup>2</sup> 本ガイダンスの目的に対し医薬品とは、特別の定めのない限り、ヒトを対象とする医薬品と生物製剤の両方をさす。

<sup>3</sup> 複写は、[http://www.fda.gov/cder/drug/cancer\\_endpoints/default.htm](http://www.fda.gov/cder/drug/cancer_endpoints/default.htm) で入手可能である。

ダンス中の *should* という言葉の使用は、何かが提案されている、または勧告されているという意味であって、必須であるという意味ではない。

## II . 背景

臨床試験エンドポイントは、異なった目的を取り扱っている。従来の腫瘍学医薬品開発において、初期臨床試験は安全性を評価し、腫瘍の縮小のような生物学的薬物活性の証拠を同定する。後期の相の臨床効果試験のためのエンドポイントは、一般的に医薬品が、生存の延長または症状の改善などのような臨床的有益性を提供するか否かを評価する。以下の項では、臨床効果に対する一般的な規制要求事項、および、それらがどのように抗がん剤の承認に対してエンドポイント選択に影響してきたかについて議論している。後の項では、更に詳細にこれらのエンドポイントについて記述し、それが様々な臨床の現場において疾患活動性または臨床的有益性の尺度として有用であるか否かを議論している。

### A . 有効性のための規制要求事項

新薬の有効性を示す要求事項は、連邦食品・医薬品・化粧品法に対する 1962 年の改正に基づいている。この法律は、有効性に関する実質的証拠を要求しており、この証拠は適切な対照を置き、よく管理された臨床試験から得られるものでなければならないと明示している。同様に、公衆衛生法は生物製剤が安全で、純粋で、強力なものであることを要求している。医薬品の承認を支持している臨床的有益性は、重要な臨床転帰（例えば、生存期間の延長、症状の改善）を含んでいるが、確立された代理エンドポイント（例えば、血圧、血清コレステロール）に対する効果も含んでいる。

1992 年に公布された迅速承認規制（21 CFR part 314, subpart H および 21 CFR part 601, subpart E）により、重い疾患または生命をおびやかす疾患を治療することが意図され、現在利用できる治療法よりも改善を示すか、かつてない治療法を提供するかのどちらかの医薬品または生物製剤の承認に対して、付加的なエンドポイントの使用が可能となっている。この設定において、FDA は（「疫学的、治療的、病態生理学的、または他の証拠に基づく」）臨床的有益性を合理的に予測可能な代理エンドポイントに関する効果に基づいて承認を与える可能性がある。このような代理エンドポイントは、心血管疾患に対する血圧またはコレステロールのような通常の代理エンドポイントに比べて、十分に確立されたものとはいえない。製造業者が実際の臨床的有益性を検証し記述するために臨床試験を行うという条件で、医薬品は迅速承認規制の下で承認される。市販後試験が臨床的有益性を示さない場合、または申請者が必要な試験を実施することに適正な配慮を示さない場合、その医薬品は迅速承認過程の下で市場から排除される可能性が高い。以下の議論において、通常承認という用語は、臨床的有益性の実証に基づく医薬品承認の昔ながらの方法を意味している。この用語は、有益性を合理的に予測する可能性が高い代理エンドポイントの使用と関連した迅速承認とは異なる。

臨床試験の好ましい回数などの医薬品の承認を支持するための重要な証拠は、産業界に対するガイダンスである *FDA Approval of New Cancer Treatment Uses for Marketed Drug and Biological Products*<sup>4</sup> および

<sup>4</sup> 当局は、定期的にガイダンスを更新している。ガイダンスの最新版を確実に入手するには、<http://www.fda.gov/cder/guidance/index.htm> の CDER ガイダンス・ウェブページを確認されたい。

1997年のFDA Modernization Act<sup>5</sup>の中で議論されている。ほとんどの場合において、FDAは、少なくとも2つの、適切な対照を置き、よく管理された臨床試験を推奨している。特定の場合においては、1つの試験からの証拠でも十分な場合がある(例えば、1つの多施設試験が生存期間における効果のような、重要な臨床的有益性について非常に信頼性が高く、統計学的に頑健な証拠を提供する例、および2番目の試験の結果の確認が現実的にまたは倫理的に不可能である例においては)特定の悪性腫瘍のある病期の患者の治療に対して承認される医薬品に関して、1つの試験から得られた証拠は、同じ癌の異なる病期の治療に対して、臨床効果の追加を支持するのに十分である場合がある。

#### B. 腫瘍学における過去の承認を支持するエンドポイント

通常の承認において、申請者が、臨床的有益性の直接的な証拠または臨床的有益性の確立した代理エンドポイントの改善を示すことは決定的なことである。腫瘍学における生存期間の改善は、臨床的有益性の適切な尺度と考えられている。さらにスポンサーは、抗がん剤承認のために他のエンドポイントを使用してきた。1970年代に、FDAは通常、放射線学的検査または理学所見から得られた腫瘍評価によって決定される客観的奏効率(ORR)に基づいて、抗がん剤を承認してきた。1980年代初期には、ODAC(抗がん剤諮問委員会)との議論の後、FDAは抗がん剤承認は、臨床的有益性のより直接的な証拠、例えば、生存期間の改善、患者の生活の質(QOL)の改善、身体機能の改善、腫瘍関連症状の改善に基づくべきであると決定した。これらの有益性は、必ずしもORRから予測されない、あるいはORRと関連しないかもしれない。

次の10年間では、いくつかのエンドポイントが、臨床的有益性の確立された代理エンドポイントとして使用されてきた。無病生存期間(DFS)が改善されたことにより、かなりの患者が再発時に癌の症状を呈することが予測される状況下で、選択された補助療法における医薬品の承認が支持された。完全奏効は感染症、出血、および血液製剤使用の減少と関連しており、持続的な完全奏効は、白血病の臨床的有益性の確立されたエンドポイントと考えられていた。高く実証されたORRにより、選択された固形腫瘍での通常の承認が支持される可能性は高いが、承認決定を行う際、その奏効期間、腫瘍関連症状の緩和、および医薬品の毒性が同時に考慮される必要がある(O'Shaughnessy and Wittes et al., 1991, Commentary Concerning Demonstration of Safety and Efficacy of Investigational Anticancer Agents in Clinical Trials, J Clin Oncol, 9: 2225-2232)。例えば、乳癌のためのホルモン剤に対するランダム化試験では、通常の承認を支持するエンドポイントとしてORRが使用されている。ORR、および適切な奏効期間と連結した腫瘍関連症状が改善されたことより、いくつかの臨床の場における通常の承認が支持されてきた。

迅速承認に対する代理エンドポイントは、臨床的有益性を合理的に予測する可能性が高くない(21 CFR part 314, subpart H、および21 CFR part 601, subpart E)。このような医薬品はまた、現在利用できる治療法よりも有益性を示す必要がある(21 CFR part 314, subpart H、および21 CFR part 601, subpart E)。<sup>6</sup> ORRは、迅速承認の支持において最も一般的に使用される代理エンドポイントであり、腫瘍反応は、癌治療を指導する癌専門医によって広く受け入れられている。ORRは医薬品の効果を直接反映しているため、利用できる治療法がない不応性腫瘍患者で行われる単群試験は、ORRの正確な評価を提供するものとなっている。

<sup>5</sup> <http://www.fda.gov/cder/fdama/default.htm>

<sup>6</sup> 産業界のためのガイダンス *Available Therapy* (<http://www.fda.gov/cder/guidance/index.htm>) を参照されたい。



### Ⅲ．エンドポイントに関して一般的に考慮すべき点

この項では、抗がん剤の開発における一般的な問題についての概要を提供している。一般的に使用される癌エンドポイントに関する議論の後に、これらのエンドポイントを使用する癌臨床試験デザインに関連する問題についての議論が続いている（将来のガイダンスでは、特定の治療適応エンドポイントについて議論する予定である）。この項で議論されているエンドポイントは、全生存期間、腫瘍評価に基づくエンドポイント（例えばDFS、ORR、完全奏効、無増悪期間（TTP）、無増悪生存期間（PFS））、および症状評価に基づくエンドポイントを含んでいる。表1は、抗がん剤の承認におけるエンドポイントの比較を示している。臨床効果エンドポイントの適切な解析に関する多くの問題は、産業界のためのICHガイダンスE9 *Statistical Principles for Clinical Trials* の中で言及されている。

#### A．全生存期間

全生存期間は、ランダム化から全死因による死亡までの時間と定義されており、Intent-to-treat集団で測定される。生存期間は最も信頼性が高い癌のエンドポイントと考えられており、試験が生存期間を適切に評価するために実施できる場合、これが通常は好ましいエンドポイントと考えられている。このエンドポイントは正確で測定が容易であり、死亡日によって文書により証明される。バイアスはエンドポイント測定において要因ではない。生存期間の改善は、臨床的有益性を評価するためのリスク便益分析として解析される必要がある。

全生存期間は、ランダム化比較試験で評価される必要がある。過去の臨床試験から得られたデータは、時間依存的なエンドポイント（例えば、全生存期間、PFS）に対してあまり信頼性が高いものとはいえない。既存対照と現在の治療群との間のアウトカムの明らかな違いは、患者の選択、改善された画像技術、または改善された支持療法を含めて、医薬品以外の違いから生じている可能性が高い。ランダム化試験は、直接的なアウトカムの比較を提供することによって、これらの違いの影響を最小化している。毒性プロファイルが受け入れ可能な場合、全生存期間における統計学的に有意な改善の実証は、臨床的に意義があると考えられることができ、しばしば新薬の承認を支持してきた。

生存試験を実施して解析することの難しさには、大規模な試験で追跡期間が長いこと、および生存期間解析を潜在的に交絡させる癌の後治療が関与している。

#### B．腫瘍評価に基づいたエンドポイント

この項では、腫瘍評価に基づくいくつかのエンドポイントについて議論する。これらのエンドポイントにはDFS、ORR、TTP、PFSおよび治療成功期間（TTF）が含まれる。これらの時間依存的なエンドポイントに関するデータの収集および解析は、間接的な評価、計算、および推定（例えば、腫瘍測定）に基づいている。PFSのデータ収集および解析については、付録3に示される表で補完される。

腫瘍評価エンドポイントの選択においては、2つの点を考慮する必要がある。第一に、エンドポイントが

表1 重要な抗がん剤承認におけるエンドポイントの比較

エンドポイント	許認可に必要な証拠	試験デザイン	メリット	デメリット
全生存期間	通常の承認に対する臨床的有益性	<ul style="list-style-type: none"> <li>ランダム化試験が不可欠</li> <li>盲検化は不可欠ではない</li> </ul>	<ul style="list-style-type: none"> <li>一般に受け入れられている直接的な有益性の尺度</li> <li>測定が容易</li> <li>測定が正確</li> </ul>	<ul style="list-style-type: none"> <li>より大規模な試験が必要となる可能性がある</li> <li>クロスオーバー治療および後治療に影響される可能性がある</li> <li>癌以外の原因での死亡を含む</li> </ul>
症状エンドポイント (患者の報告によるアウトカム)	通常の承認に対する臨床的有益性	<ul style="list-style-type: none"> <li>ランダム化盲検化試験</li> </ul>	<ul style="list-style-type: none"> <li>直接的な臨床的有益性の患者の観点</li> </ul>	<ul style="list-style-type: none"> <li>盲検化はしばしば困難である</li> <li>データが頻繁に欠測しているまたは不完全である</li> <li>僅かな差異の臨床的意義は知られていない</li> <li>多項目解析となる</li> <li>妥当性確認済みの測定手段が不足している</li> </ul>
無病生存期間	迅速承認または通常の承認の代理エンドポイント*	<ul style="list-style-type: none"> <li>ランダム化試験が不可欠</li> <li>盲検化が望ましい</li> <li>盲検化レビューが推奨される</li> </ul>	<ul style="list-style-type: none"> <li>生存試験と比べて、標本サイズが小規模で追跡が短期間で済む</li> </ul>	<ul style="list-style-type: none"> <li>すべての設定において、生存期間の代理エンドポイントとして統計的に妥当性が確認されているわけではない</li> <li>正確には測定できず、特に非盲検試験では評価バイアスの影響を受ける</li> <li>定義が試験間で異なる</li> </ul>
客観的奏効率	迅速承認または通常の承認の代理エンドポイント*	<ul style="list-style-type: none"> <li>単群またはランダム化試験で使える</li> <li>比較試験においては盲検化が望ましい</li> <li>盲検化レビューが推奨される</li> </ul>	<ul style="list-style-type: none"> <li>単群試験で評価可能である</li> <li>生存試験と比べて、短期に小規模な試験で評価される</li> <li>自然史ではなく、医薬品に起因する効果である</li> </ul>	<ul style="list-style-type: none"> <li>有益性の直接的な尺度ではない</li> <li>薬物活性に関する包括的な尺度ではない</li> <li>利益を得るのは患者の部分集団のみである</li> </ul>
完全奏効	迅速承認または通常の承認の代理エンドポイント*	<ul style="list-style-type: none"> <li>単群またはランダム化試験に使用してもよい</li> <li>比較試験においては盲検化が望ましい</li> <li>盲検化レビューが推奨される</li> </ul>	<ul style="list-style-type: none"> <li>単群試験で評価可能である</li> <li>持続的な完全奏効は臨床的有益性を示すかもしれない</li> <li>生存試験と比べて、短期の小規模な試験で評価される</li> </ul>	<ul style="list-style-type: none"> <li>すべての場合において有益性の直接的な尺度というわけではない</li> <li>薬物活性に関する包括的な尺度ではない</li> <li>利益を得るのは患者の小さな部分集団である</li> </ul>
無増悪生存期間 (すべての死亡を含む) または 無増悪期間(増悪前の死亡は打ち切り)	迅速承認または通常の承認の代理エンドポイント*	<ul style="list-style-type: none"> <li>ランダム化試験が不可欠</li> <li>盲検化が望ましい</li> <li>盲検化レビューが推奨される</li> </ul>	<ul style="list-style-type: none"> <li>生存試験と比べて、標本サイズが小規模で追跡が短期間で済む</li> <li>不変の測定が含まれている</li> <li>クロスオーバー治療および後治療に影響されない</li> <li>一般的には客観的で定量的な評価に基づいている</li> </ul>	<ul style="list-style-type: none"> <li>すべての設定において、生存期間の代理エンドポイントとして統計的に妥当性が確認されているわけではない</li> <li>正確には測定できず、特に非盲検試験では評価バイアスの制約を受ける</li> <li>定義が試験間で異なる</li> <li>頻繁な放射線学的評価または他の評価が必要である</li> <li>治療群間で均衡の取れた時期での評価を伴う</li> </ul>

\* 迅速承認または通常の承認の代理エンドポイントとしての妥当性は、効果の大きさ、効果の期間、および他の利用可能な治療法の有益性のような他の因子に大きく依存する。詳細については本文を参照されたい。



迅速承認と通常の承認のどちらを支持しているかに関する決定は明確にされる必要がある。第二に、エンドポイントは、腫瘍エンドポイント評価におけるバイアスまたは不確実性の可能性に対して評価される必要がある。臨床効果に関する唯一の証拠として、腫瘍測定に基づくエンドポイントに依存した試験を行っている医薬品の申請においては、第二の試験から得られる検証的な証拠を必要とする可能性がある。腫瘍を測定する際の精度は、腫瘍の設定の間で異なる可能性がある。奏効率の決定に使用される腫瘍測定は、明確な境界領域のない状況で不正確となる可能性がある（例えば、悪性中皮腫，膀胱癌，脳腫瘍）。

試験の主要エンドポイントが腫瘍測定（例えば、PFSまたはORR）に基づく場合、腫瘍エンドポイント評価は、一般的には、試験治療に関して情報の与えられていない中央判定者によって検証される必要がある（付録4を参照されたい）。研究自体が盲検化されていない場合は、この方策は特に重要なものとなる。FDAが中央判定のプロセスを検証するためには、スキャンの標本を監査することが適切であるかもしれない。データ収集に関する付加的な詳細は付録1に記載されている。腫瘍エンドポイント評価（特に、PFSまたはDFSに関する）の中央での独立した検証は、ランダム化試験が盲検化される場合（有害事象プロファイルにより、実際の試験において実質的に盲検化が解かれていない限り）、または感度分析が観察者のバイアスがないこと（特に、DFSに対する）を支持するような大規模ランダム化試験において、効果の大きさが頑健である場合には、必要ではないかもしれない。

### 1. 無病生存期間

一般的に、DFSはランダム化から腫瘍の再発または全死因による死亡までの時間と定義されている。このエンドポイントを最も頻繁に使用するのには、根治的な手術または放射線療法の後の補助療法の時である。患者の多くが化学療法で完全奏効に至る場合にも、DFSは重要なエンドポイントとなる可能性がある。全生存期間が、大部分の補助療法にとっては従来型のエンドポイントであるにもかかわらず、DFSは、生存期間が長くなる可能性があり、生存期間のエンドポイントが非実用的な状況において、重要なエンドポイントとなる可能性が高い。DFSは乳癌の術後のホルモン療法、結腸癌の術後療法、乳癌の細胞傷害性術後療法に対する承認に関して、主要な基盤を提供している。標準的な細胞傷害性療法と比較した場合、ホルモン療法の副作用は最低限のものでしかなく、良好なリスク便益関係をもたらしているといえる。DFSは臨床的有益性の代理エンドポイントである、または臨床的有益性の直接的な証拠を提供する可能性が高いといえる。この決定は効果の大きさ、そのリスク便益関係、および疾患背景に基づいている。しかしながら、生存期間の有益性がすでに確立された疾患背景においては、DFSが臨床的有益性と考えられる可能性は低い。2003年12月のODACのコンセンサスは、有益性の大きさが補助療法の毒性を上回った場合、DFSの延長が臨床的有益性を示すというものであった。2004年5月に、ODACは、DFSが外科的な補助療法における結腸癌の抗がん剤に対して、受け入れ可能なエンドポイントと考えられると勧告した。<sup>7</sup>

潜在的なエンドポイントとしてDFSを評価する際の重要な考慮すべき点には、治療効果の推定された大きさおよび標準治療の証明された有益性が含まれている。プロトコルは、DFSの定義と追跡調査および来院の日程の両方について、注意深く明確にする必要がある。予定外の評価は多くの理由が原因で起こる可能性があり、予定外の評価の起こる頻度、時期、または理由における試験群間の差異により、バイアスの生

<sup>7</sup> 複写は、[http://www.fda.gov/cder/drug/cancer\\_endpoints/default.htm](http://www.fda.gov/cder/drug/cancer_endpoints/default.htm) で入手可能である。

じる可能性がある。バイアスは、治療の割り当てに対して、患者および医師を盲検化することで、最小限にすることができる。予定外の評価が原因と考えられるバイアスの潜在的な影響は、イベントが起こった時点に関係なく、追跡期間の期間中に起こるイベントの総数の比較分析によって評価可能である。

特に死亡が、先行されるべき腫瘍増悪の文書による証明がない状態で記録される場合、DFSの定義は複雑なものとなる可能性が高い。これらのイベントは、疾患の再発または打ち切られたイベントのいずれかとして記録されうる。死亡の統計解析のいかなる方法にも若干の限界が存在するものの、すべての死亡（すべての原因からくる死亡）を再発と考えることで、バイアスを最小限にすることができる。特に観察を行わないで長期間が経過した後死亡する患者においては、DFSはこの定義によって過大評価される可能性がある。長期間におよぶ追跡来院の頻度が試験群間で異なる場合、または毒性のため脱落例がランダムでない場合には、バイアスが生じるかもしれない。解析によっては、癌に関連した死亡をDFSのイベントの中に入れて、癌に関連しない死亡を打ち切りとする場合もある。この方法では、死因の帰属におけるバイアスを生じる可能性がある。さらにまた、死亡時点であれ最終来院時点であれ、患者を打ち切るいかなる方法も、打ち切られた患者が打ち切られていない患者と再発に関して同じリスクを持つと仮定している。

## 2. 客観的奏効率

ORRは、あらかじめ定められた量および最小限の期間で、腫瘍サイズが縮小している患者の割合として定義される。奏効期間は通常、最初の反応の時点から文書により証明された腫瘍増悪時までで測定される。一般にFDAは、ORRを部分奏効と完全奏効との和として定義してきた。このように定義される場合、ORRは薬物の抗腫瘍活性の直接的な尺度となり、単群試験で評価可能である。不変はORRの構成要素に含まれるべきではない。不変は疾患の自然史を反映する可能性が高く、一方、腫瘍の縮小は直接的な治療効果であるといえる。また、不変はTTPまたはPFS解析（下記を参照されたい）によってより正確に評価される可能性がある。利用可能な場合には、標準化された基準が反応を確認するのに使用されるべきである。様々な反応判定基準が適切であると考えられてきた（例えば、RECIST基準など）(Therasse and Arbuck et al., 2000, New Guidelines to Evaluate response to Treatment in Solid Tumors, J Natl Cancer Inst, 92:205-16)。反応判定基準は、試験開始前にプロトコルであらかじめ定義されている必要がある。ORRの意義は、完全奏効（腫瘍の検出可能な証拠なし）の大きさと期間、および割合によって評価される。

## 3. 無増悪期間および無増悪生存期間

TTPおよびPFSは、医薬品承認のための主要エンドポイントとして用いられてきた。TTPは、ランダム化から客観的な腫瘍増悪までの時間として定義されており、死亡は含まない。PFSは、ランダム化から客観的な腫瘍増悪または死亡までの時間として定義されている。腫瘍増悪の正確な定義は重要で、プロトコルに注意深く詳述される必要がある。

### a. TTPとPFS

TTPと比較した場合、PFSは好ましい許認可に必要なエンドポイントである。PFSは死亡を含んでおり、全生存期間に対して良好な相関性を示す。TTP解析では、死亡時または情報のある打ち切り（試験からの

脱落の非ランダムなパターン)を表す早期の来院時点で、死亡は打ち切られる。PFSでは、患者の死亡が腫瘍増悪にランダムに関連があると仮定している。しかしながら、大部分の死亡が癌に無関係である場合、TTPの方が受け入れ可能なエンドポイントかもしれない。

#### b. 医薬品承認を支持するエンドポイントとしての PFS

表1には、エンドポイントとしてPFSを用いることのメリットおよびデメリットが示されている。PFSは腫瘍の増殖を反映することができ、生存期間の有益性の決定以前に評価可能である。その決定は、後治療によって交絡されることはない。所定の標本サイズに対して、PFSに対する効果の大きさは、全生存期間に対する効果より大きくなる可能性が高い。しかしながら、存在する多くの異なる悪性腫瘍に対して、生存期間の代理エンドポイントとしてのPFSの正式な評価は困難であろう。データは通常、生存期間とPFSに関する効果の間の相関性について頑健な評価を行うには不十分である。癌試験はしばしば小規模であり、既存の医薬品の証明された生存期間の有益性は一般にはそれほど大きくない。ライセンス承認を支持するエンドポイントとしてのPFSの役割は、異なる癌の設定により変化する。PFSにおける改善により、それが直接的な臨床的有益性が臨床的有益性の代理となるか否かは、利用できる治療法と比較した場合の新しい治療法の効果、およびリスク便益の大きさに依存する。

#### c. PFS 試験デザインの問題

PFSを評価、測定、解析するための方法論は、プロトコルおよび統計解析計画書(SAP)の中で詳述される必要がある。プロトコルにおいて腫瘍増悪基準を慎重に定義することも重要である。現在、増悪を定義する許認可のための標準的な基準は存在しない。申請者は、RECIST基準を含む種々の異なる基準を使用してきた。大部分の公表されたPFS基準に示される幅広い概略は、プロトコルおよびSAPにおいて付加的な詳細で補完されるべきである。来院および放射線学的評価は、系統的なバイアスを防ぐために、2つの試験群間で対称である必要がある。可能な場合、試験は盲検化される必要がある。患者または医師の評価が増悪エンドポイントの構成要素として含まれる場合、盲検化は特に重要である。最低限、評価は情報を与えられていない独立した、一般的には放射線科医および臨床医から構成されている判定チームに従う必要がある。FDAおよび申請者は、以下の項目について前もって同意するべきである。

- 試験デザイン
- 増悪の定義
- 症例報告書(CRF)に記録される予定のデータ
- SAP
- 欠測値の処理の方法論と打ち切り法
- 適用可能な場合、独立エンドポイント審査委員会(IRC)の業務手順書(付録4を参照されたい)

#### d. PFSの解析

欠測値は、PFSの解析を複雑にする可能性がある。プロトコルでは、各々の患者に対して適切な評価のための来院が定義されている必要がある(すなわち、すべての予定される腫瘍評価が行われる来院)。解析計

画では、各々の治療群での追跡の妥当性に関する比較が概説される必要がある。不完全および/または欠落した追跡来院に関して解析を行う方法論と打ち切り法は、プロトコルに指定される必要がある。解析計画では、結果の頑健さを評価するために、主解析および1つ以上の感度分析が指定される必要がある。欠測データでのいかなる解析にも問題があるけれども、主解析および感度分析の両方における同様の結果によって結果は強化される。評価は、長期間追跡不能であった患者での死亡数を含む必要がある。このような死亡の不均衡により、追跡が不十分な治療群におけるPFSを過大評価することによって、PFS測定にバイアスがかかる可能性がある。

増悪データは複数の情報源(予定外の来院時での健康診断、および様々な種類の放射線スキャンを含む)から、また異なる時間に収集される可能性があるため、各々の評価のための来院に対するデータ収集は、予定の来院日前後の特定の短い時間間隔に限定される必要がある。データが長期間で収集される場合、イベント日と打ち切り日を決定する際に困難が生じる可能性がある。FDAは、事前に欠測している評価がない状態で増悪が観察される最も早い日に増悪日を割り当て、最終の放射線学的評価が増悪のないことを確定する日で打ち切ることを推奨している。付録3には、主解析または感度分析に使用できるPFSの潜在的な解析のための一組の表が示されている。PFSデータ収集と解析のための計画については、第II相試験終了時の会議でFDAと議論され、特別なプロトコル評価において検証される必要がある。

#### 4. 治療成功期間

TTFは、ランダム化から疾患増悪、治療毒性、および死亡を含めたすべての理由による治療中断までの時間を測定する複合エンドポイントとして定義される。TTFは、医薬品承認における許認可のためのエンドポイントとして推奨されていない。TTFは、これらの付加的な変数から臨床効果を適切に区別していない。許認可のためのエンドポイントでは、毒性、患者または医師の治療中止、または患者の不耐性から医薬品の臨床効果を明らかに区別する必要がある。

#### C. 症状評価を含むエンドポイント

症状の改善は、臨床的有益性と考えられる。FDAの医薬品の承認では、患者の症状評価および/または症状の改善を表す身体的徴候(例えば、体重の増加、滲出液の減少)が主要な臨床効果エンドポイントとして使用されてきた。しかしながら、総合的な健康関連の生活の質(HRQL)の尺度は、腫瘍学における医薬品承認で主要な臨床効果エンドポイントとして用いられてはいない。HRQL測定手段およびそれらの妥当性確認は、産業界のためのガイダンスの草案である*Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims*で議論されている。<sup>8</sup> 抗がん剤承認を支持する主要なエンドポイントとして使用されるべき徴候および症状の改善、またはQOL評価について、腫瘍の症状における改善と薬物毒性の減少とをFDAが区別できるようにすべきである。総合的なHRQL測定手段に基づく見かけ上の有効性のメリットは、有効性というよりは単に毒性が少ないということを示している可能性がある。

<sup>8</sup> 最終化されれば本ガイダンスは、このトピックに関する現在の食品医薬品局(FDA)の考えを表すであろう。ガイダンスの最新版に関しては、<http://www.fda.gov/cder/guidance/index.htm>のCDERガイダンス・ウェブページを確認されたい。



## 1. 特定の症状エンドポイント

TTPと類似したエンドポイントである**癌の症状の無増悪期間**は、潜在的な代理エンドポイントというよりも臨床的有益性の直接的な尺度である。ただし、以下のような問題が考えられる。先に議論したように、増悪を測定する際の問題（例えば、評価の欠測）はまた、症状の増悪までの時間を評価する際にも存在する。ほとんどの癌試験は盲検化されていないので、評価に偏りが生じる可能性がある。また、腫瘍の増悪と癌の症状の発生との間に遅延が見られることがある。しばしば、代替治療は症状のエンドポイントに到達する前に開始され、この解析を交絡させる。多くの癌試験は、軽微な癌の症状を有する患者で実施される。さらに、腫瘍の症状は薬物毒性と区別するのが難しいこともある。

**複合症状エンドポイント**は、類似した臨床的重要性の構成要素を備えている必要があり、その結果は1つの構成要素にのみに排他的に帰されるべきではない。例えば、複合有益性エンドポイントに基づいて、骨に転移した癌患者の治療に対する医薬品が承認されてきた。骨に関連したイベントは、病理学的骨折、骨に対する放射線療法、骨に対する手術療法、および脊髄圧迫と定義されている。

適切な集団の選択は、症状の有益性を文書により証明するために重要であろう。試験のベースライン時に症候性の患者は、カテゴリカルな症状反応解析で評価されることができる。ベースライン時に無症候性の患者は、最初の症状までの時間の解析が使用されうる。患者が試験薬を中止する、または新薬を開始しても、最初の症状が文書化されるまで追跡が続けられるならば、症状の増悪はまだ評価可能である。

## 2. 症状データが遭遇する問題

欠測データおよび頻度の少ない評価は、特に非盲検試験において症状データの評価を複雑にする可能性がある。薬物毒性または腫瘍増悪のための治療中止は、症状データが欠測する1つの原因である。理想的には、患者が治療を止めても、解析のためにデータ収集用紙で情報を収集し続ける必要がある。複数の症状に関するデータ収集は、多重性に関して前向きに対処し、必要な統計的調整はSAPに指定される必要がある。

### D. バイオマーカー

血液または尿において測定される**パラプロテイン値**は、骨髄腫の反応判定基準の一部として使用されてきたけれども、通常、血液および体液から検査される**バイオマーカー**は、抗がん剤承認のための主要なエンドポイントとして使用されてこなかった。利用できる検査の妥当性を確立し、バイオマーカーの改善が臨床的有益性を予測可能か否かを決定するために、更なる研究が求められている。

FDAは、複合エンドポイントの要素として**腫瘍マーカー**を受け入れてきた。CA-125の著しい増加と連結した特定の臨床イベント(PSの顕著な低下や腸閉塞症)の発生は、卵巣癌患者における増悪と考えられていた。あるいは、バイオマーカーは予後因子の同定および試験デザインで考慮される患者や層別因子の選択に有用であろう。

#### IV．臨床試験デザインの考慮すべき点

21CFR 314.126 について、規則で述べられているように、FDA は、「適切な対照を置き、よく管理された試験」から得られる臨床効果の実質的証拠に基づいて医薬品を承認している。試験は対照に対して妥当な比較を行い、医薬品の効果の定量的評価を提供しなければならない。臨床効果を示すための最も信頼性の高い方法は、盲検化されたランダム化対照試験で、臨床的に意味のあるエンドポイントにおいて、統計学的に有意な改善を示すことである。以下の項では、癌試験のデザインに関連したいくつかの問題について議論する。

##### A．単群試験

利用できる治療法がなく、主要な腫瘍の退縮が試験薬に帰するものであると仮定できる設定において、FDA は迅速承認を支持する実質的証拠として単群試験で観察される ORR および奏効期間を時として支持してきた。奏効率は、完全奏効が輸血の必要性の減少、感染症の減少、生存期間の延長と関連のある通常の承認に関する急性白血病のような設定で使用されてきた。単群試験は、生存期間、TTP、または PFS のようなイベントまでの時間というエンドポイントを適切に特徴づけることはない。多くの種類の癌では自然史が変化に富むため、イベントまでの時間というエンドポイントを評価するにはランダム化試験が必要とされる。

##### B．非劣性を示すようデザインされた試験

非劣性 (NI) 試験は、新薬が事前に特定された量 (非劣性の限界) までは標準治療 (実薬対照) に比べて劣っていないことを示すことによって、新薬の有効性を示すべきである (Temple and Ellenberg, 2000, Placebo-Controlled Trials and Active-Control Trials in the Evaluation of New Treatments, Part 1: Ethical and Scientific Issues, Ann Intern Med, 133(6): 455-63)。<sup>9</sup> この非劣性の限界は、実薬対照薬の効果より小さい臨床的に受け入れ可能な損失である必要がある。標準治療は、よく特徴づけられた臨床的有益性 (生存に対する有益性) を有する必要がある。新薬が実薬対照に比べて非劣性の限界以上に劣っている場合、それは有効性がないと見なされる。

NI 試験は、実薬対照の治療効果の大きさを決定する外部の (過去の) データに依存する。癌試験において、この効果はしばしば適切に特徴づけられることがなかった。NI 試験はまた、恒常性の仮定にも依存する。この仮定は、実薬対照の効果が過去の試験と現在の試験の間で一定のままであることを含んでいる。これは、現在の試験と実薬対照の効果が生じた過去のデータの間において、患者集団の特徴、支持療法、および評価技術の恒常性を仮定する。実薬対照の治療効果の推定された大きさは、過去の試験の包括的なメタアナリシスに基づく必要がある。これらの試験は、プラセボ群と比較した場合の実薬対照の効果を確実に再現している必要がある。NI 試験を行う際の難しい点には、実薬対照の効果の推定、および留保されるべき効果の量 (NI 限界) の決定などがある。NI 試験は通常、優越性試験と比較して大きな標本サイズを必

<sup>9</sup> 産業界のための ICH ガイダンス E10 *Choice of Control Group and Related Issues in Clinical Trials* (<http://www.fda.gov/cder/guidance/index.htm>) も参照されたい。



要とし、臨床試験の結果の追試を伴う。さらにまた、後治療および実薬対照群へのクロスオーバーは、すべてのNI解析を交絡させる可能性がある。生存期間以外のエンドポイントを用いたNI試験には問題がある。

### C. 放射線療法保護剤および化学療法保護剤のための試験デザイン

放射線療法保護剤および化学療法保護剤は、治療法の毒性を改善するようデザインされた医薬品である。これらの試験には通常2つの目的がある。第一の目的は、癌治療の毒性の改善を評価することである。第二の目的は、抗がん活性が保護剤によって損なわれるか否かを決定することである。第二の目的では通常、全生存期間よりは代理エンドポイント、例えばORRまたはTTPなどを評価する。

## V. 結論

本ガイダンスで概説される一般原則によって、申請者が販売申請のためのエンドポイントを選択できるようになるべきであるが、FDAは、NDAまたはBLA販売申請を支持することが意図されるプロトコルを提出する前に、申請者がFDAと話し合いの場を持つことを推奨する。FDAは、これらの会議に癌専門医、統計学者、臨床薬理学者、およびしばしば外部の専門家顧問からなる集学的なFDAチームを必ず参加させるであろう。申請者は、これらの会議後、プロトコルを提出することができ、医薬品の販売申請を支持するエンドポイントの受容性とプロトコル・デザインの確認を提供する特別なプロトコル評価を要請することができる。<sup>10</sup>最後に、言うまでもなく販売承認は臨床試験のデザインのみには依存するわけではなく、医薬品の販売申請におけるすべての試験からの結果およびデータに関するFDAの審査にも依存している。

\* \* \*

<sup>10</sup> 産業界のためのガイダンス *Special Protocol Assessment* (<http://www.fda.gov/cder/guidance/index.htm>) を参照されたい。

## 付 録 1

### 腫瘍測定データの収集<sup>11</sup>

以下の事項は、腫瘍測定データについて考慮すべき重要な点である。FDAは、以下の事項を推奨する。

- CRFおよび電子データにより、治療前のベースライン来院中に同定された標的病変が記録されること。ベースライン病変に関する遡及的な同定は信頼性に劣ると考えられる。
- 腫瘍病変には固有の識別文字または数字が割り当てられること。この割り当てにより、ある解剖学的部位に生じている複数の腫瘍間での区別、およびベースラインで測定される腫瘍と追跡中に測定される腫瘍の対応が明確になる。
- 追跡中の重要な時点における完全なデータ収集が確実になされるような仕組みが整備されていること。すべての標的病変がベースライン時に評価され、同一の撮影および測定方法が、ベースラインおよび追跡で必要なすべての検査に対して使用されることをCRFが保証すること。
- CRFは、スキャンが各来院で実施されたか否かを記録するデータフィールドを含むこと。
- 病変が完全に消失した場合にはゼロが記録されること。そうしなければ、病変の消失は欠測値と区別することができない。
- 追跡検査は、当初のもしくは新たな疾患部位の新規病変の随時の検出につながる。新規病変の発生および部位については、CRFおよび提出される電子データの中に記録されること。

---

<sup>11</sup> この付録において、腫瘍データは画像ではなく、SAS 移送ファイルにおけるデータのことを指している。一般的に画像はNDAまたはBLAに提出されることはないが、審査過程にFDAにより監査を受ける可能性がある。

## 付 録 2

## PFS 解析で考慮すべき問題

プロトコルおよびSAPにはPFSの主解析が詳述される必要がある。この解析は、エンドポイントの詳細な説明、腫瘍を評価するための適切な方法、およびIRCについての手順のようなバイアスを最小限にするための手順を含む必要がある。試験の実施において予測される問題を評価し、結果が頑健であるか否かを評価するために、1つか2つの副次解析を指定する必要がある。その際、以下の重要な要因が考慮されるべきである。

- 増悪日の定義 .生存期間解析においては正確な死亡日が得られる .PFS解析においては正確な増悪日は得られない .PFS解析で用いられる記録された増悪日 (*PDate*) の定義として以下の2つの方法が使用できる .

1 .増悪を断言することができる最初の時点に割り当てられる *PDate* .

- 新規病変に基づく増悪については、*PDate*はその新規病変が検出された最初の観察日である .
- 標的病変測定値の和に基づく複数の評価が異なる時点で行われる場合、*PDate*は標的病変測定値の和について予め定義された増加を示した標的病変の直前の観察日または放射線学的評価日である .

2 .すべての放射線学的評価（増悪をまとめて文書化する）が行われた直後の、プロトコルで予定されている医療機関への来院日としての *PDate* .

- 打ち切り日の定義 .打ち切り日は、追跡期間の終了または脱落前に増悪の記録のない患者において定義される .これらの患者において、打ち切り日は、しばしば増悪状態が適切に評価された最後の日として定義される .1つの受け入れ可能なアプローチでは、最後の評価実施日が使用される .しかしながら、複数の放射線学的検査が増悪の決定の際に評価される可能性がある .2つめの受け入れ可能なアプローチでは、これらの放射線学的評価に対応する医療機関への来院日が使用される .
- 適切なPFS評価の定義 .増悪の証拠のない患者において、PFSについての打ち切りは、しばしば最後の適切な腫瘍評価の日に依存する .適切な腫瘍評価を規定する慎重な定義には、非標的病変を評価して新規病変を検索することを考慮した標的病変評価の妥当性および放射線学的検査の妥当性を含む .

- 部分的に欠測している腫瘍データの解析 .解析計画には ,適切な腫瘍評価のための来院のデータが部分的に欠測している場合の増悪状態の算出方法が記述される必要がある .
- 完全に欠測している腫瘍データ .データの収集されない評価来院の後に死亡または増悪を示した評価来院が続く場合がある .また ,引き続き評価は増悪を示さない場合がある .後者の場合 ,治療を続け増悪に対してモニタリングを続けることは ,適切かもしれない .しかしながら ,このアプローチは以降のイベント発生に依存して欠測データを異なって取り扱い ,情報のある打ち切りを表す可能性がある .もう1つの可能なアプローチは ,引き続きPFS評価からのデータを含めることである .評価の頻度が高い場合 ,および1回の追跡来院だけが欠けている場合には ,これは適切かもしれない .2回以上の来院が欠けている場合は ,直前の適切な腫瘍評価時点で打ち切りとすることがより適切でありうる .SAP では ,欠測データの潜在的効果を評価するために ,主および副次PFS解析が詳述される必要がある .脱落例の理由は ,打ち切りおよび増悪状態を決定する手順に組み込まれる必要がある .例えば ,主解析については ,文書により証明されていない臨床的増悪 ,癌治療法の変更 ,またはPSの低下が原因で試験から外れる患者は ,直前の適切な腫瘍評価時で打ち切りとすることができる .副次感度分析は ,増悪イベントとしてこれらの脱落例を含めるであろう .増悪に関して来院が欠けることは問題となりうるが ,欠測となった来院回数にかかわらず ,疾患増悪について患者を追跡し続けるためにあらゆる努力を尽くす必要がある .
- 測定不可能な疾患の増悪 .適切ならば ,各々の評価方法(例えばCTスキャン ,骨スキャン)に対して ,増悪基準が記述される必要がある .測定不可能な疾患に基づく増悪を文書により証明するスキャンは ,情報が与えられていない審査委員会によって審査され ,FDA による検証に対して利用可能とすべきである .
- 疑わしい病変 .解析時の増悪状態の割り当てに関して ,不確定な病変を評価して追跡するためのアルゴリズムが提供される必要がある .

## 付 録 3

## PFS 解析のための表の例示

Ⅲ . B . 項で議論されているように、感度分析はPFS解析が頑健なものか否かを決定するのに役立つ可能性が高い。しかしながら、これらの感度分析は主解析の結果に対して探索的で補助的であり、臨床効果が感度分析のみに基づいて主張される可能性は低いといえる。異なる感度分析は、増悪イベントおよび増悪データの打ち切り日が、どのように割り当てられるかについて指定する表の形式で記述される。以下の3つの表は、3つの異なる感度分析の例を記述したものである。

a . 表Aは、文書で十分に立証された、検証可能な増悪イベントのみを対象とする感度分析を表している。他のデータは打ち切りとされる。表Aでは、増悪日は以下のとおりである。

- IRCによって検証された放射線学的評価のみに基づく。臨床的増悪は、増悪エンドポイントとは見なされない。
- 腫瘍増悪が記録された最初の時点に割り当てられる。
- 患者が綿密に追跡されている場合の死亡日。しかしながら、2回以上の来院がない後に発生した死亡は最後の来院時で打ち切りとされる。

表 A PFS 1 (文書により証明された増悪のみを対象)

状 態	増悪または打ち切り日	アウトカム
ベースラインの腫瘍評価がない	ランダム化	打ち切り
予定されている来院間での文書化された増悪	次のうち最も早い日 ● 新規病変を認めた放射線学的評価日 (増悪が新規病変に基づく場合) または ● 直前の測定病変の放射線学的評価日 (増悪が、測定病変の和の増加に基づく場合)	増悪
増悪なし	最後の測定病変の放射線学的評価日	打ち切り
文書化されていない増悪による治療の中止	直前の測定病変の放射線学的評価日	打ち切り
毒性または他の理由による治療の中止	直前の測定病変の放射線学的評価日	打ち切り
新たな抗がん治療が開始された	直前の測定病変の放射線学的評価日	打ち切り
最初のPD評価以前の死亡	死亡日	増悪
適切な評価のための来院間の死亡	死亡日	増悪
2回以上来院が欠けた後の死亡または増悪	直前の測定病変の放射線学的評価日	打ち切り

表Bの感度分析においては、予定されていた来院日によりのみ打ち切り日およびイベント日を割り当てることにより、腫瘍評価のための追跡スケジュールにおける潜在的なバイアスが修正される。しかしながら、増悪が最後の来院により近い時点で起こる場合、特に非盲検試験では、このアプローチはバイアスを生じさせる可能性がある。このアプローチは盲検化、ランダム化試験で適切でありうる。

表B PFS 2 (統一的な増悪および評価日)

状態	増悪または打ち切り日	アウトカム
ベースラインの腫瘍評価がない	ランダム化	打ち切り
予定されている来院間での文書化された増悪	次に予定されている来院	増悪
増悪なし	最後の適切な評価を伴う来院日	打ち切り
文書化されていない増悪による治療の中止	直前の適切な評価を伴う来院日	打ち切り
毒性または他の理由による治療の中止	直前の適切な評価を伴う来院日	打ち切り
新たな抗がん治療が開始された	直前の適切な評価を伴う来院日	打ち切り
最初のPD評価以前の死亡	死亡日	増悪
適切な評価のための来院間の死亡	死亡日	増悪
2回以上来院が欠けた後の死亡または増悪	直前の適切な評価を伴う来院日	打ち切り

b. 表Cの感度分析では、医師の評価によってPFSが評価される。しかしながら、増悪が最後の来院により近い時点で起こる場合、特に非盲検試験では、このアプローチはバイアスを生じさせる可能性がある。このアプローチは、盲検化、ランダム化試験で適切でありうる。

表C PFS 3 (医師の判断を対象)

状態	増悪または打ち切り日	アウトカム
ベースラインの腫瘍評価がない	ランダム化	打ち切り
予定されている来院間での文書化された増悪	次に予定されている来院	増悪
増悪なし	最後の適切な評価を伴う来院日	打ち切り
医師が臨床的増悪を認める場合	予定されている来院 (または来院間の場合、次に予定されている来院)	増悪
毒性または他の理由による治療の中止	直前の適切な評価を伴う来院日	打ち切り
増悪を認めないで新たな抗がん治療が開始された	直前の適切な評価を伴う来院日	打ち切り
最初のPD評価以前の死亡	死亡日	増悪
適切な評価のための来院間、または患者が評価のための来院を1回逃した後の死亡	死亡日	増悪
追跡不能の期間が長かった(2回以上の評価の欠測)後の死亡	直前の適切な評価を伴う来院日	打ち切り



## 付 録 4

## 腫瘍エンドポイントの独立審査

医薬品の承認を支持する臨床試験の結果は、申請者およびFDAによって検証可能であるべきである。単群試験で決定される ORR では、限られた数の画像を審査することにより評価可能である。医薬品の承認が PFS の測定に基づく場合、慎重な計画を立てることによって、バイアスを最小限にし、申請者およびFDAによる結果の検証が可能になる。IRC の活用によって、放射線学的知見の放射線造影の解釈、および評価における独立した審査におけるバイアスが最小限なものとなるであろう。IRC の機能および運営（独立審査権）の概要を述べる審査権が明確に記載された計画書は、試験の開始前に FDA と合意されている必要がある。この計画では、委員会の独立性の保証、および結果の収集、保管および移送の手続きが記述される必要がある。この審査権には、データの最終的な解釈と監査手順において、臨床データの解釈と組み入れにおける相違の解決法が含まれる必要がある。IRC の活用に関しては、産業界のためのガイダンス *Developing Medical Imaging Drug and Biological Products, Part 3: Design, Analysis, and Interpretation of Clinical Studies* でさらに議論がなされている。