

ヒトを対象とする抗がん剤の評価に関する
ガイドラインの付録1 (CHMP/EWP/205/95 REV. 3)
登録に向けての検証的試験における、
主要エンドポイントとしての
無増悪生存期間の使用に関する方法論的考察*

Appendix 1 to the guideline on the evaluation of anticancer
medicinal products in man (CHMP/EWP/205/95 REV. 3)

Methodological considerations for using
Progression-Free Survival (PFS) as primary endpoint
in confirmatory trials for registration

欧州医薬品庁 (EMA)
欧州ヒト用医薬品委員会 (CHMP)

2008年1月24日

European Medicines Agency (EMA)
Committee for Medicinal Products for Human Use (CHMP)
London, 24 January 2008
Doc. Ref. EMA/CHMP/EWP/27994/2008
www.ema.europa.eu/pdfs/human/ewp/2799408en.pdf

訳 西川 昭子¹⁾ 手良向 聡²⁾

1) 助先端医療振興財団臨床研究情報センター研究事業 2) 京都大学医学部附属病院探索医療検証部

* 本翻訳は、EMA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EMAの許諾を必要としないものである。

序

臨床効果試験における主要エンドポイントとしてのPFS（無増悪生存期間）またはDFS（無病生存期間）の使用は、前向きに対処する必要のあるいくつかの方法論上の問題を提示している。この付録は、定義、評価の頻度および方法、確認バイアス（ascertainment bias）、逸脱および欠測データの処理、および放射線学的検査に関連して考慮すべき問題に関して、いくつかの一般的な規制指針を提供する。主要エンドポイントの選択、主要エンドポイントとしてPFS/DFSを使用することの適切性に関する指針は、この付録の範囲外である。

エンドポイントの定義

PFSは、伝統的に、ランダム化（または、非ランダム化試験における登録）から客観的な腫瘍増悪、またはあらゆる原因による死亡のどちらか早い方までの時間と定義される。DFSは、ランダム化から客観的な再発またはあらゆる原因による死亡までの時間と定義される。

疾患の増悪および再発は、通常、客観的な放射線学的知見に基づいて評価される。可能な限り、増悪の定義は確立された反応評価規準（例えば、RECISTまたはWHO規準、EBMT規準など）に従う必要がある。しかしながら、薬剤の種類、病変の部位と種類、および試験の目的によって、変更された規準がより適切である場合が認められている。例えば、追加的で客観的な臨床および生化学または放射線学的な規準が、増悪を評価するのに使用されることがある。すべての場合において、増悪イベントの定義に関する規準ができるだけ客観的であること、またその定義がプロトコルにおいて明確かつ事前に定義されていることが重要である。

データ解析において考慮すべき点

実際には、試験の転帰（outcome）についてランダム化されたすべての被験者の完全な追跡が利

用できることはまれであろう。さらに、ランダム化後に起こるプロトコル治療からの逸脱および離脱は、試験データに影響を与える可能性があり、結果として、特に、それらが治療割り付けに関連している場合には、試験の結論に影響を与える可能性もある。例をあげると、被験者は間違った試験薬を投与されるかもしれないし、全く何も投与されないかもしれないし、予定完了日前に治療から離脱する場合もあり、追跡不能になる場合もあり、増悪の証拠が得られる前に治療を変更する場合もあり、評価の欠測後に増悪の証拠を示す場合もあるかもしれない。

独立に判定される増悪の前の試験治療からの離脱というイベントは、機械的に情報を持たない（non-informative）と見なすことはできず、したがって、統計解析においてこれらのイベントを観察打ち切り（censoring）とすることの妥当性は疑わしいかもしれない。すべての抗がん剤試験において最適となるこの問題を扱う方法は存在しないが、PFS/DFSの主解析に関する解析対象集団を定義する際には、ITT（intention-to-treat）の原則が合理的なアプローチを提供し、われわれはできるだけこれに従うべきである。特に、ランダム化されたすべての患者について、転帰データは意図された評価の日程に従って収集される必要があり、増悪または再発の日付は、違反、試験薬の中止または治療法の変更に関係なく、客観的な増悪または再発の最初の証拠を得た時点に基づいて決められる必要がある。特定の試験について、異なるアプローチがより適切であると考えられる場合には、正当化が要求され、計画段階におけるCHMP Scientific Advice（ヒト用医薬品委員会科学的助言）の合意が推奨される。

たとえ試験プロトコルで予見されていたとしても、試験治療から離脱した患者について増悪に関する信頼できるデータを収集することは、時には難しいかもしれない。これに対して、また他の理由に対して、観察打ち切りの規則を含めて、欠測データの処理方法を事前に定義し、正当化する必要がある。これらの方法は、試験の目的に関して妥当であると同時に、バイアスおよび情報の損失を最小限にするように選択されるべきである。

これには、PFS/DFSの解析においてイベントとして判定される増悪／再発の前の治療の中止、または変更を考慮するアプローチが含まれる可能性がある。潜在的なバイアスは常に注意を向けられる必要があり、異なったアプローチを用いて感度解析が行われる必要がある。補助的な解析は、例えば、増悪日を来院予定日とするアプローチ、区間打ち切り解析 (interval-censored analysis)、事前に定めたランダム化後のある時点に増悪を割り当てる単時点解析 (single time point analysis) などを含むかもしれない。

現在、規制観から、感度解析に対して推奨可能ないくつかのアプローチが存在する。感度解析の範囲は、試験結果が頑健なことを示すのに十分である必要があり、臨床状況および観測された試験データの性質 (例えば、患者の離脱パターン) に依存するであろう。

提示された解析の範囲から得られる結論に違いが出た場合には、説明が必要となるであろう。また、異なる解析および解析対象集団の重要性は、試験のデザイン (優越性または非劣性) にも依存するであろう。

主解析についての方針は、試験の開始前に明確に文書化される必要がある。試験の計画段階で、逸脱および欠測データの処理に対処するための感度解析を含めて、統計解析計画書に十分な考察が与えられることが重要である。盲検試験で、試験の終了時に盲検化レビューを行うことは、予見されない問題に対処できるように、選択されたデータ処理方法、および提案された解析の範囲のレビューを行う貴重な機会を提供することになるかもしれない。これらのデータ解析問題をどのように処理し、文書化するかについては、ICHトピックE9、臨床試験のための統計的原則に関する指針の注記に提供されている一般的な指針に従うべきである (CHMP/ICH/363/96)。同様のアプローチは、非盲検試験に対しては歓迎されないであろう。このような試験については、蓄積されるデータの知識を持って行われる解析の重要な点に対する修正は懸念を生じさせるので、試験プロトコルおよび統計解析計画書を記載するときに最大限の努力が要求される。

欠測データの処理に対する一般的なアプローチは、欠測データにおける考慮事項に関する利用可能な指針に従うべきである (CPMP/EWP/1776/99)。

評価の頻度および方法

PFSの評価は、その腫瘍の種類に特異的な考えられる疾患のすべての部位がベースライン時に評価されること、および理想的には同じ方法を使用して、臨床的および放射線学的に示されるように、病変部位が追跡期間中に他の部位とともに系統的に評価されることを求めている。同様に、DFSの評価は、疾患が認められそうな部位が追跡評価時に系統的に評価されることを要求するかもしれない。

したがって、評価の頻度は、期待される治療効果を検出するために適正でなければならない。増悪を評価する最適な頻度は、試験の目的、治療日程および患者の環境を考慮に入れながら、原則試験ごとに判断される必要がある。できるだけ正確に増悪を評価する必要性と、浸襲的かつ資源集成的な診断法に対する患者の曝露を最小限にする必要性の間でバランスが見出される必要がある。

増悪／再発の推定時点が評価の頻度の結果として単に変化し、より頻度が少ない評価に起因して、より長いPFS/DFSとなる可能性がある。したがって、ランダム化試験において腫瘍評価の方法および頻度は、治療サイクルが異なる長さのときでさえ、試験群にわたって対称である必要がある。患者の臨床状態の変化が予定外の疾患状態の評価を促す際、適切に盲検化されていない臨床試験は特に確認バイアスの危険にさらされる。プロトコルで定義された日程を遵守することは不可欠であり、逸脱は報告される必要がある。来院日程の遵守は解析時点で記述的に調べられる必要があり、試験結果へのいかなる影響も検討される必要がある。

予定された評価からの逸脱の処理

主解析のためには、2つの予定された腫瘍評価

の間で検出される増悪の時期は、例えば、どちらか一方の予定された評価時点に基づくべきではなく、増悪が文書化された時点に基づいて割り当てられるべきである。予定された評価時点に基づく代替的な解析は、感度解析として含むことが可能である。予定された評価時期からの逸脱は、検出バイアス (detection bias) を生じる可能性があるため注意が必要である。

異なる追跡来院の日程および／または追跡の異なるパターンによるバイアスの問題は、適切な試験デザインおよび実施によって最小限にされる必要がある。異なる追跡来院の日程が避けられない場合、生じたバイアスを検討する方法は試験プロトコルの中で概説されていなければならない。

例えば、術前補助療法の研究において、手術は疾患の再発または増悪が観察される可能性のある主要な時点である。しかしながら、手術の時点が2つの治療群で異なることは避けられないかもしれない。このような場合に考えられる1つの解決策は、最後に予定された手術の直後のある固定された時点を選ばず、(手術の時点でのイベントを含めて) その前に起こったすべてのイベントを評価する時間として、その固定された時点を利用することであるかもしれない。

増悪の推定時間の精度を保持しながらバイアスを最小限とすることを目的として、補助的な解析における追跡パターンの予期されなかった違いをどう処理するかについても、様々なアプローチが提案されており、このような解析を事前に特定することについて考慮が与えられるべきである。しかしながら、事後のデータ解析は、バイアスのかかったデータ収集に対して補正するという点で限られた価値しかもたない。一般に、治療群間での追跡来院の日程に大きな乖離が存在する場合、いずれのアプローチも重要なバイアスがないことを納得のいくように立証することはないであろう。

感度解析は、プロトコルまたは統計解析計画書において計画され記述される必要があり、いかなる変更も研究報告書において正当化されなければ

ならない。

独立したレビュー

疾患または増悪の評価は、特に適切に二重盲検化されていない臨床試験、および殊に多くの病変を追跡する必要のある進行性疾患の試験で、研究者バイアス (investigator bias) を潜在的に被る可能性がある。これらの状況、つまりPFSまたはDFSが主要エンドポイントである状況において、症例報告書に報告されているような臨床的腫瘍増悪データの独立かつ盲検化された放射線学的評価および調査は、研究者バイアスを疑う理由がある場合には、データの質に関して何らかの再保証を提供するかもしれない。このバイアスの源は、PFSまたはDFSが主要エンドポイントである適切に二重盲検化されたランダム化試験においては主要な問題ではないかもしれない。独立した評価を通して割り当てられた結果の役割は、プロトコルに前もって詳述される必要がある。一般に、独立した評価が適切である場合には、主解析は独立した評価を通して割り当てられた結果に基づく必要がある。独立し、盲検化された評価からの試験結果が、研究者の評価と重要な程度に至るまで異なる場合、PFS/DFSに関するデータはより説得力があるものとなるであろう。独立したレビューの手順は、臨床試験文書に前向きに定義され記述されなければならない。

規制指針

DFSまたはPFSが、販売認可申請を支持する検証的試験で主要エンドポイントとして使用される状況においては、PFSまたはDFSの定義と評価、および逸脱の処理に関連する方法論上の問題に関して、CHMP Scientific Advice (ヒト用医薬品委員会科学的助言) を考慮することは適切であるかもしれない。