
固形がんの治療効果判定のための 新ガイドライン (RECIST ガイドライン)

— 改訂版 version 1.1 —

日本語訳 JCOG 版 ver.1.0

**New response evaluation criteria in solid tumours:
Revised RECIST guideline (version 1.1)**

E.A. Eisenhauer, P. Therasse, J. Bogaerts, L.H. Schwartz, D. Sargent, R. Ford,
J. Dancey, S. Arbuck, S. Gwyther, M. Mooney, L. Rubinstein, L. Shankar,
L. Dodd, R. Kaplan, D. Lacombe, J. Verweij

***EUROPEAN JOURNAL OF CANCER* 45 (2009) 228–247**

厚生労働省がん研究助成金（国立がん研究センターがん研究開発費）

計画研究 20-15「画像によるがんの診断、治療法選択、治療効果判定に関する研究」班

主任研究者：渡辺裕一（国立がん研究センター中央病院）

指定研究 20 指 -6「多施設共同研究の質の向上のための研究体制確立に関する研究」班

主任研究者：福田治彦（国立がん研究センターがん対策情報センター）

研究協力者：木村 綾（JCOG 運営事務局）

JCOG 運営委員会承認 2010 年 6 月 14 日

No responsibility is assumed by Elsevier for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made. Original English Edition ©2008 Elsevier, Japanese Edition ©2010 Elsevier Japan K.K.

固形がんの治療効果判定のための 新ガイドライン (RECISTガイドライン) 改訂版 version 1.1

New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1)

要約

背景: 腫瘍量の変化を測定することは、がんの治療法を臨床的に評価するための重要な要素であり、がん臨床試験においては客観的な腫瘍縮小効果と原病の増悪の両方が重要なエンドポイントとして用いられてきた。2000年にRECISTガイドライン (Response Evaluation Criteria in Solid Tumors) が公表されて以来、多くの研究者、共同研究グループ、企業、政府機関等が、治療のアウトカムの評価のためにこの規準を採用してきた。しかし、広く用いるうちに多くの疑問点と問題点が指摘されてきたことを受けて、この改訂版RECISTガイドライン (第1.1版) が作成されるに至った。本特集号に同時掲載されている別稿に要約されている、今回の改訂の基となったエビデンスは、6,500例を超える大規模なデータベースでの評価やシミュレーション研究、いくつかの文献レビューにより得られたものである。

改訂版RECIST1.1のハイライト: 今回の改訂における主な変更点は以下の通りである。

- ・ **評価対象とする病変数:** 効果判定のための腫瘍量の評価に必要とされる病変数を合計で最大10から最大5に減らした (臓器ごとに最大5から最大2へ)。これは多くの試験の腫瘍縮小効果判定のデータベースの統合解析から得られた結果に基づいて決定された。
- ・ **病的リンパ節腫大の評価への組み入れ:** 短径15mm以上のリンパ節を測定可能とし、標的病変としての評価を可能とした。
腫瘍縮小効果判定に際しての腫瘍径の算出に、リンパ節の短径の測定値が含まれた。短径が10mm未満に縮小したリンパ節は正常リンパ節とみなす。
- ・ **効果の確定 (confirmation):** 腫瘍縮小効果 (奏効率) を primary endpoint とする試験では、効果の確定

(confirmation) が必要であるが、ランダム化試験においては得られた結果を適切に解釈するために必要な比較対照群があるために confirmation は必須としない。

- ・ **進行 (PD) の定義の明確化:** 腫瘍径の和がきわめて小さい場合に過剰に PD と判定することを避けるために、標的病変の PD として RECIST1.0 での定義であった腫瘍径の総和の20%増加に加え、絶対値での5mmの増加が必要とした。

さらに、RECIST1.0での混乱の原因となっていた、測定不能病変や非標的病変の「明らかな増悪 (unequivocal progression)」に関する指針も提案された。

最後に、FDG-PETによる評価の解釈を含めて、新病変の検出に関する項目が盛り込まれた。

- ・ **画像診断の手引き:** 今回の改訂版RECIST1.1には、病変の至適な形態的評価 (anatomical assessment) に関する最新の報告を付した、画像診断についての新しい付録が含まれている。
- ・ **今後の課題:** RECISTワーキンググループがRECIST1.1の作成にあたって議論した主な問題は、形態に基づく腫瘍量の1次元評価から、3次元的形態評価やPET・MRIを用いた機能的評価への移行が適切か否かということであった。しかし、ワーキンググループは、現時点では、腫瘍量の形態的な1次元評価を廃止するには、新しい方法の標準化やエビデンスが充分ではないと結論した。ただし、唯一の例外として「増悪」の判定の補助として使用するFDG-PET画像を挙げた。こうした新しい手法の使用については、本特集号の最後の論文でも詳述されているように、適切な妥当性研究が必要であると考えられる。

©2008 Elsevier Ltd. All rights reserved.

Keyword

Response criteria, Solid tumours
Guidelines

Article history

Received 17 October 2008
Accepted 29 October 2008

本稿は、EUROPEAN JOURNAL OF CANCER 45 (2009) 228–247 掲載論文、「New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1)」を、当該雑誌の出版元であるエルゼビア社の許諾を正式に得て日本語に訳したものである。翻訳作業は、厚生労働省がん研究助成金計画研究 20-15「画像によるがんの診断、治療法選択、治療効果判定に関する研究」班（主任研究者－渡辺裕一：国立がんセンター中央病院）および同指定研究20指-6「多施設共同研究の質の向上のための研究体制確立に関する研究」班（主任研究者－福田治彦：国立がんセンターがん対策情報センター）の研究活動の一環として行い、以下に示す研究班を主体とする日本臨床腫瘍研究グループ（Japan Clinical Oncology Group：JCOG）の運営委員会の審査を経て完成させた。なお、翻訳に際しては、文と文、語と語の対応よりも、原文と翻訳文の意味および論理の対応を重視し、逐語訳ではなく必要に応じて大幅な意識を加えた。解釈に迷う場合には原文の該当箇所を参照いただきたい。また、筆者らは、検討段階で気づいた原著の明らかな論理的な誤りについてワーキンググループに意見を述べたが、いくつかは修正されずに公表されている。筆者らが明らかな論理的な誤りと考える箇所については訳注として解説を加えた。さらに、無増悪生存期間のイベントについては欧米の研究グループ間によっても意見が異なるため、研究計画時や結果の公表時にはイベントの定義を明確に記述する必要がある。そうした注意が必要な箇所についても訳注として解説を加えた。

転載許諾について

原著論文の著作権はエルゼビア社（本社 Elsevier Ltd.）に、日本語訳の著作権はエルゼビア・ジャパン株式会社に帰属する。本日本語訳の転載を含む二次利用に際してはエルゼビア社（本社）とエルゼビア・ジャパン株式会社の許諾ならびにJCOG運営委員会の承認が必要である。なお、本日本語訳の本文の一部や表を加工して利用するような場合は「引用」にあたるため、原著論文を引用し、本日本語訳から一部改変したことを明記し、JCOG データセンター（webmaster@ml.jcog.jp）まで電子メールにて報告すること。その際、原著ならびに日本語訳と内容が大幅に異なる場合には、作成者に対して内容の変更を求めることがある。

JCOG 運営委員会（2010年3月現在）

JCOG 代表者 / 運営委員会委員長 / 肺がん内科グループ代表者：

田村友秀（国立がんセンター中央病院）
運営委員会副委員長： 福田治彦（国立がんセンターがん対策情報センター）
監査委員会委員長： 下山正徳（名古屋医療センター/国立がんセンター中央病院）
プロトコル審査委員会委員長： 飛内賢正（国立がんセンター中央病院）
効果・安全性評価委員会委員長： 西條長宏（近畿大学医学部）
教育研修委員会委員長： 山本精一郎（国立がんセンターがん対策情報センター）
放射線治療委員会委員長： 早川和重（北里大学医学部）
病理委員会委員長： 松野吉宏（北海道大学病院）
試料解析研究委員会委員長： 松村保広（国立がんセンター東病院）
薬事委員会委員長： 藤原康弘（国立がんセンター中央病院）
臨床試験専門職委員会委員長： 加幡晴美（国立がんセンターがん対策情報センター）
利益相反委員会委員長： 堀田知光（名古屋医療センター）
肺がん外科グループ代表者： 坪井正博（神奈川県立がんセンター）
消化器がん内科グループ代表者： 朴成和（静岡県立静岡がんセンター）
胃がん外科グループ代表者： 笹子三津留（兵庫医科大学）
食道がんグループ代表者： 安藤暢敏（東京歯科大学市川総合病院）
リンパ腫グループ代表者： 塚崎邦弘（長崎大学大学院）
乳がんグループ代表者： 青儀健二郎（四国がんセンター）
婦人科腫瘍グループ代表者： 嘉村敏治（久留米大学医学部）
大腸がんグループ代表者： 森谷亘皓（国立がんセンター中央病院）
泌尿器科腫瘍グループ代表者： 鳶巣賢一（静岡県立静岡がんセンター）
骨軟部腫瘍グループ代表者： 岩本幸英（九州大学医学部）
放射線治療グループ代表者： 平岡真寛（京都大学医学研究科）
脳腫瘍グループ代表者： 洪井壮一郎（国立がんセンター中央病院）
肝胆膵グループ代表者： 古瀬純司（杏林大学医学部）

JCOG の主体をなす研究班：

厚生労働省がん研究助成金指定研究

- 20指-1「高感受性悪性腫瘍に対する標準治療確立のための多施設共同研究」
（主任研究者－飛内賢正：国立がんセンター中央病院）
- 20指-2「呼吸器悪性腫瘍に対する標準治療確立のための多施設共同研究」
（主任研究者－田村友秀：国立がんセンター中央病院）
- 20指-3「消化器悪性腫瘍に対する標準治療確立のための多施設共同研究」
（主任研究者－島田安博：国立がんセンター中央病院）
- 20指-4「希少悪性腫瘍に対する標準治療確立のための多施設共同研究」
（主任研究者－洪井壮一郎：国立がんセンター中央病院）
- 20指-5「放射線治療を含む標準治療確立のための多施設共同研究」
（主任研究者－平岡真寛：京都大学医学研究科）
- 20指-6「多施設共同研究の質の向上のための研究体制確立に関する研究」
（主任研究者－福田治彦：国立がんセンターがん対策情報センター）

1. 背景

1.1. RECISTの歴史

腫瘍量の変化を測定することは、がんの治療法を臨床的に評価するための重要な要素であり、がん臨床試験においては腫瘍の縮小（客観的腫瘍縮小効果：objective response）と原病の増悪までの期間（無増悪生存期間：progression-free survival）の両方が重要なエンドポイントとして用いられてきた。新薬のスクリーニングを行う第II相試験において、抗腫瘍効果の証拠として腫瘍縮小効果をエンドポイントとして用いることは、長年にわたる多くの固形がんにおける共通の評価指標として既に広く認知されており、ある一定の割合の患者に腫瘍縮小をもたらす薬剤は、引き続いて行われる第III相試験において全生存期間（overall survival）または他の生存時間（time to event）を改善させる可能性が（必ずとは言えないものの）あることが知られている（文献 [1-4] 参照）。現時点で、客観的腫瘍縮小効果は他のどのバイオマーカーよりも、第II相試験における有望な薬剤をスクリーニングするための治療効果の指標として、その有用性が高く評価されている。さらに、進行期のがんを対象とする臨床試験において、新薬開発の第II相試験と第III相試験の両方で、有効性に関する結論を導く上でのエンドポイントとして無増悪期間（または無増悪生存期間）が用いられることが多くなってきているが、これらもまた腫瘍サイズの形態的な測定に基づくものである。

しかしながら、客観的腫瘍縮小効果と無増悪期間は、形態的な腫瘍量に基づく標準的な規準が広く採用され、かつ簡便に使用できる場合にのみ有用であると言える。1981年に世界保健機関（WHO）は、腫瘍縮小効果がprimary endpointである試験での使用を主たる目的とした腫瘍縮小効果の判定規準を初めて発表した。WHO規準は、病変の2方向での測定値の積を合計することで腫瘍量を全体的に評価し、治療中にベースラインからの変化を評価することにより治療効果を判定するという概念を導入した⁵。しかし、WHO規準発表後の数十年間に、この規準を使用する共同研究グループや製薬企業は、新しい技術に対応するため、またはWHO規準の原著では不明確だった部分に対処するために、WHO規準に多くの“修飾”を加えた。このことが試験結果の解釈に関する混乱を生じさせ⁶、実際に、様々な異なる効果判定規準が適用されるようになった結果、同一レジメンであるにもかかわらずその有効性に関して全く異なる結論が導き出されるようになってしまった⁷。こうした問題を解決すべく、効果判定規準を標準化し、かつ簡略化するために1990年代半ばに国際ワーキンググループ（International Working Party）が設立された。それによりRECIST（固形がんにおける効果判定規準）として知られる新規規準が2000年に公表された⁸。初版RECIST1.0の主な特徴は、測定可能病変のサイズの最小

値の定義、評価する病変数（最大10個、各臓器につき最大5個）に関する指針、腫瘍量の全般的評価方法として2方向測定ではなく1方向測定を採用、である。RECIST1.0はその後、腫瘍縮小効果や増悪をprimary endpointとする試験において、研究機関、共同研究グループ、企業により広く採用されるようになった。また各国の規制当局は、これらの評価に関する適切なガイドラインとしてRECISTを受容している。

1.2. RECISTの改訂の必要性

RECISTが2000年に発表されて以来、多くの研究者が2方向測定（や3次元評価）での規準を1方向測定での規準で置き換えることの妥当性について、前向きの解析によって検証してきた（[9] 参照）。稀な例外を除き（中皮腫など）、固形がんを対象とした第II相試験において、1方向測定での規準を使うことは概ね妥当であろうと考えられている。しかし、これらの前向き解析や、使用経験が増えるに従い、いくつかの疑問や課題が提示されてきており、それらに回答し、より明確にすることの重要性が認識されるようになった。その疑問や課題とは次のようなものである。

- 患者ごとの総合効果（または試験の有効性に関する結論）に影響を及ぼすことなく、10以下の病変数で評価を行うことが可能か。
- ランダム化第III相試験で、腫瘍縮小効果ではなく増悪がprimary endpointの場合、特にすべての患者が測定可能病変を有しているわけではない場合はRECISTをどのように適用すべきか。
- FDG-PETやMRIなど新しい画像診断技術を使用すべきか、その場合はどのように使用すべきか。
- リンパ節病変の評価はどのように行うべきか。
- 効果の確定（confirmation）は本当に必要か。
- 非細胞傷害性の分子標的治療薬の試験にRECISTは適用できるか。

このRECISTガイドライン改訂版（本改訂ガイドライン）には、これらの点すべてに関する最新情報を含めた。

1.3. RECIST 1.1 作成の経緯

RECISTワーキンググループは、早期薬剤開発の専門知識を有する学術研究機関の臨床医、政府、企業、画像診断専門家、統計家から成り、定期的に会合を開いてRECIST改訂に関する議題を設定し、各種の変更を正当化するために必要なエビデンスを選定して、挙げられたエビデンスの検討を行った。改訂プロセスにおける重要課題は、企業やアカデミアのグループの試験から得られた、前向きに記録された固形がんの測定データのデータベースを構築することであった。

このデータベースは、Jan BogaertsとPatrick Therasse（本改訂ガイドラインの共著者）の指揮の下、EORTC データセンターで統合された。6,500 例を超える患者の 18,000 を超える標的病変から成り、腫瘍縮小効果と無増悪生存期間の結果に及ぼす各種の要因（必要な標的病変数、効果の確定（confirmation）の必要性、リンパ節の測定ルールなど）の影響を調べるために使用された。この結果は、RECIST ワーキンググループの評価を受けた後、本改訂ガイドラインにおける、大半の変更箇所の根拠として用いられた。その詳細は本特集号の別稿に掲載されている¹⁰。Larry Schwartz と Robert Ford（本ガイドラインの共著者）より提供されたデータベースには本改訂に反映された推論の元となる重要な情報が含まれていた¹¹。

本改訂ガイドラインでは、臨床試験における腫瘍量の評価を簡便化、最適化、標準化するための変更がなされており、この出版は時宜に叶っていると考えられる。主な変更点の要約を付録Iに掲載する。なお、腫瘍量の評価に対する基本的なアプローチは、疾患の機能的評価ではなく依然として形態的評価に基づくものであるため、本版は RECIST2.0 版ではなく 1.1 版とした。

1.4. 3次元画像評価と機能画像評価について

腫瘍縮小効果判定方法の移行、すなわち形態的評価を1方向計測によるものから、3次元画像評価や機能画像評価（ダイナミック造影 MRI、ダイナミック造影 CT、腫瘍代謝を評価する（18）F-フルオロデオキシグルコース（（18）F-fluorodeoxyglucose）によるポジトロン断層撮影（FDG-PET）など）に移行する「時」ではないのかという指摘は頻繁に寄せられる。しかし、本改訂ガイドラインで示したように、RECIST ワーキンググループ（特に画像診断研究を担当したメンバー）は、現時点ではこれらの代替評価法はその採用を推奨するほどには十分に標準化されておらず、また、十分に一般化されているとも言えないと考えた。本改訂ガイドラインでも後述するが、唯一の例外は、増悪の判定の補助として使用する FDG-PET 画像である。本特集号の論文で詳述するように¹²、我々は、（本文に記述したように、将来、形態的評価に加えたり、それに置き換えたりすることができるであろう）これらの有望な新しい方法を使用するには、適切かつ厳密な妥当性確認のための臨床研究が必要であると考えている。Sargent らにより、これらの新しい評価法のための「エンドポイント」を定義できるようにするにはどういったタイプのデータが必要であるのか、また、これらの新しい評価法/評価規準が RECIST1.0 と比べて、第II相スクリーニング試験において、真に有効な新薬を正しく選択し、真に無効な新薬を正しく棄却することに関する妥当性をより高くするための提案（どういう状況で、いつ用いるか、をどのように決めるべき

かといった考察）が解説されている。

RECIST ワーキンググループは、今後数年の間にこのようなデータが現れ、RECIST1.1 の次回の改訂において適切な変更が可能になることを期待している。

2. 本改訂ガイドラインの目的

本改訂ガイドラインは、成人および小児のがんの臨床試験において使用する、固形がんの測定の標準的な方法と、腫瘍のサイズの変化の客観的評価の定義について記述したものである。これらの効果指標はすべて腫瘍量の形態的評価および腫瘍量の変化に基づいているため、客観的な腫瘍縮小効果が primary endpoint であるすべての試験と、安定 (stable disease) や増悪、無増悪期間の評価が行われる試験において、本改訂ガイドラインは有用であると考えられる。しかしながら、ある薬剤や治療レジメンが有効かどうかを評価するこれらのどのエンドポイントに関して、その規準を満たす患者がどの程度の割合を占めるかについて、本改訂ガイドラインでは特定の仮定を置いているわけではない。そういった定義は、試験の対象となるがん種および評価する薬剤により異なるはずであるからである。プロトコールには、試験のサンプルサイズや判断規準の根拠となる有効性のパラメータを適切に定義する統計学関連の章が含まなければならない。本改訂ガイドラインには、腫瘍縮小効果の評価に関する定義と規準の提示に加え、腫瘍縮小効果をエンドポイントとして用いる試験における結果の標準的な報告方法に関する推奨事項も記載されている。

これらのガイドラインは、悪性脳腫瘍の試験に適用することも可能であるが、悪性脳腫瘍に関しては既に発表された別の効果判定規準もある¹³。また、悪性リンパ腫に関しては、効果判定に関する国際的ガイドラインが別に発表されているため¹⁴、本改訂ガイドラインは悪性リンパ腫の試験で使用されることは意図していない。

最後に、腫瘍専門医の多くは、日常診療で悪性腫瘍患者の経過観察のための画像検査による客観的な規準と、症状に基づく規準の双方に基づいて、治療継続の是非についての意思決定を行っているが、本改訂ガイドラインは、治療を担当する腫瘍医が適切であると判断する場合を除いて、このような個々の患者における治療継続の是非についての意思決定に用いられることを意図していない。

3. ベースラインにおける腫瘍の測定可能性 (measurability)

3.1. 定義

ベースライン評価 (3.2.1. 病変の測定を参照) において、

腫瘍病変およびリンパ節は、以下の規準に基づいて「測定可能 (measurable)」、「測定不能 (non-measurable)」のいずれかに分類される。

3.1.1. 測定可能 (measurable)

腫瘍病変 (tumour lesions):

少なくとも1方向で正確な測定が可能であり(測定断面における最大径(長径)を記録する)、かつ以下のいずれかのサイズ以上のもの。

- CTで10 mm (CTのスライス厚は5 mm以下、画像診断の手引きに関する付録IIを参照)
- 臨床的評価としての測径器 (caliper) による測定で10 mm (測径器により正確に測定できない病変は測定不能として記録する)
- 胸部X線写真で20 mm

リンパ節病変 (malignant lymph nodes):

病的な腫大と判断され、かつ測定可能なリンパ節は、CTで評価した短軸の径(短径)が15 mm以上 (CTのスライス厚は5 mm以下を推奨)。ベースラインおよび経過中は、短径のみを測定して評価する(本特集号のSchwartzら¹⁵を参照)。リンパ節の測定に関する情報については「標的病変および非標的病変のベースラインでの記録」に関する後述の注意も参照のこと。

3.1.2. 測定不能 (non-measurable)

小病変(長径が10 mm未満の腫瘍病変または短径が10 mm以上15 mm未満であるリンパ節病変)、および真の測定不能病変を含む、測定可能病変以外のすべての病変。真の測定不能病変とみなされる病変には次のものがある。軟膜髄膜病変、腹水、胸水または心嚢水、炎症性乳がん、皮膚や肺のリンパ管症、視触診では認識できるが再現性のある画像検査法では測定可能ではない腹部腫瘍や腹部臓器の腫大。

3.1.3. 病変の測定可能性に関して特に考慮すべき点

骨病変、嚢胞性病変、局所療法の治療歴のある病変に関しては、以下に示すコメントを参照のこと。

骨病変 (bone lesions):

- 骨病変測定の画像検査法として、骨シンチグラフィ、FDG-PET、単純X線撮影は適切ではないと考えられる。しかしこれらの手法は、骨病変の存在または消失を確認することには使用可能である。
- 同定可能な軟部組織成分を含み、CTやMRIなどの横断画像により評価できる溶骨性骨病変や溶骨性造骨性混

合骨病変は、軟部組織成分が上述した測定可能な定義を満たす場合には、測定可能病変とすることができる。

- 造骨性骨病変は測定不能である。

嚢胞性病変 (cystic lesions):

- 画像診断所見により定義される単純嚢胞の規準を満たす病変は、その定義上、当然単純嚢胞であることから腫瘍病変とみなすべきではない(測定可能病変・測定不能病変のいずれでもない)。
- 嚢胞性転移によると思われる「嚢胞性病変」が、上述の測定可能な定義を満たす場合には、測定可能病変とすることができる。しかし、同一患者で他に非嚢胞性病変が認められる場合は、非嚢胞性病変を標的病変に選択することが望ましい。

局所療法の治療歴のある病変:

- 過去の放射線治療の照射野内や、その他の局所療法が影響する範囲に存在する腫瘍病変は、病変が増悪を示さない限り、通常、測定可能とはしない。こうした病変を測定可能とする場合にはその条件をプロトコールに詳細に記載する。

3.2. 測定法ごとの詳細 (specification)

3.2.1. 病変の測定

すべての測定値はメートル法で記録する。臨床的評価(視触診)の場合は測径器を用いて測定する。すべてのベースライン評価は、治療開始前で、可能な限り治療開始に近い時期に行う。早くとも治療開始前4週以内に実施されなければならない。

3.2.2. 評価の方法

標的病変や非標的病変として報告される各病変を記録するにあたっては、ベースラインおよび観察期間を通じて、同一の評価法かつ同一の技術を用いなければならない。追跡する病変が、画像評価はできないが臨床的評価はできるという場合を除いて、常に、臨床的評価ではなく画像診断に基づく評価を行わなければならない。

臨床的病変 (clinical lesions):

臨床的病変(皮膚の小結節など)は、表在性で、かつ測径器により測定した長径が10 mm以上の場合にのみ測定可能とする。皮膚病変の場合、病変のサイズの測定ができるように、定規を写し込んだカラー写真により記録することが推奨される。上述のとおり、画像評価はより客観的で試験終了時に再検討もできるため、臨床的評価と画

像評価の両方で病変の評価が可能な場合は画像評価を優先する。

胸部X線写真：

特に新病変を発見するという点において、CTは単純X線よりも感度が高いため、増悪が重要なエンドポイントである場合は胸部CTが胸部X線写真より望ましい。しかし、病変の輪郭が明瞭で空気を含む肺で囲まれている場合には、胸部X線写真でその病変を測定可能としてもよい。詳細については付録IIを参照。

CT、MRI：

効果判定のために選択された病変を測定する方法として、現時点ではCTが最も広く利用可能で、最も再現性に優れた方法である。本ガイドラインでは、CTスライス厚が5 mm以下であるとの仮定に基づいて、CTで描出された病変の測定可能性を定義した。付録IIに記述するように、CTのスライス厚が5 mmを超える場合、測定可能病変のサイズの最小値はスライス厚の2倍とする。ある特定の状況（例：体幹部撮影など）においてはMRIも許容される。客観的な腫瘍縮小効果の評価にCTとMRIの両者を使用する場合についてのより詳細な注意事項は付録IIで記述する。CTによる放射線被曝に対する懸念がある場合には、限定された状況下でCTの代わりにMRIを使用してもよい。

超音波検査：

超音波は、病変の測定においては客観性に乏しく、客観的な腫瘍縮小効果判定のための測定法として使用すべきではない。超音波検査は、後日実施する第三者判定（independent review）の際に検査内容や結果の全体を再現することが不可能である。また検査の方法や結果が操作者に依存するものであるため、評価のたびに同一の技術が用いられて測定が行われること（再現性を保つこと）の保証ができない（より詳細に付録IIに記載した）。なお、試験の経過中に超音波で新病変が認められた場合には、超音波検査をもって新病変の出現としてもよいが、その場合であってもCTまたはMRIによる確認が推奨される。

内視鏡、腹腔鏡：

客観的な腫瘍縮小効果判定におけるこれらの検査の使用は推奨されない。ただし、生検組織の採取が可能な状況において、病理学的完全奏効を確認する目的や、または完全奏効後の再発や外科的切除後の再発がエンドポイントである試験において再発を確認する目的において、これらの検査は使用してもよい。

腫瘍マーカー：

腫瘍マーカーは、客観的な腫瘍縮小効果の評価に単独では使用しない。しかし、ベースラインの腫瘍マーカーの値が基準値上限を上回っていた場合、完全奏効と判定するためには腫瘍マーカーの値が基準値上限を下回っている必要がある。腫瘍マーカーの意義は疾患により異なるため、プロトコルでは当該疾患に応じて測定の方法に関する指示を記載するべきである^{*}。再発卵巣癌におけるCA-125の奏効および再発前立腺癌におけるPSAの奏効については、それぞれガイドラインが公表されている¹⁶⁻¹⁸。さらに婦人科癌共同研究グループ（Gynecologic Cancer Intergroup）では、卵巣癌の初回治療に関する試験で使用するための、客観的な腫瘍縮小効果判定に組み込むCA-125による増悪の規準を作成した¹⁹。

^{*} 訳注：少なくとも研究者主導のがんの臨床試験においては、腫瘍マーカーは各施設で採用している検査法や外注業者を用いているため、測定方法を試験ごとに特定してプロトコルに記載することは現実的ではないと思われる。強いてプロトコルに記載するとすれば許容される測定方法を列記する程度であると思われる。

細胞診、組織診：

稀な症例においてではあるが、これらの検査はプロトコルで規定した場合には、部分奏効（PR：partial response）と完全奏効（CR：complete response）を区別するために使用することができる。例えば、良性病変が残存することが知られている胚細胞腫瘍のような腫瘍における遺残病変などである。体腔液の貯留が治療による有害事象であり得ることが知られている場合（タキサン系薬剤や血管新生阻害剤など）で、測定可能な腫瘍がPRまたはSDの規準を満たした場合、PR（またはSD）とPDを区別するために、治療中に出現または増悪した体腔液が、がん性であるかどうかを細胞診によって確認することは許容される。

4. 腫瘍縮小効果の判定

4.1. 全体的な腫瘍量（overall tumour burden）および測定可能病変の評価

客観的な効果または将来の増悪を評価するためには、ベースライン評価における全体的な腫瘍量（overall tumour burden）を評価し、それを経過中の測定値の対照として使用することが必要である。客観的な腫瘍縮小効果がprimary endpointである試験では、ベースラインにおいて測定可能病変を有する患者のみを対象とすべきである。測定可能疾患（measurable disease）とは、少なくとも1つの測定可能病変（3. ベースラインにおける腫瘍の測定可能性（measurability）参照）を有するものと定義される。

※訳注：測定可能疾患 (measurable disease) と測定可能病変 (measurable lesion) は論理的に不要な区別であり、区別することで却って混乱を招くと思われる。この1文を除けば本稿の他のすべての箇所において両者は交換可能な形で用いられており、原著に登場する 'measurable disease' をすべて 'measurable lesion' に置き換えても論理的にはなんら意味は変わらない。よって本日本語訳では、この1文以外の 'measurable disease' はすべて「測定可能病変」と訳した。

腫瘍の増悪（無増悪期間または特定の時期における増悪割合）が primary endpoint である試験では、対象を測定可能病変を有する患者のみに限定するのか、測定不能病変しか有さない患者も含めるのかをプロトコールに明記しなければならない。

4.2. 「標的病変」および「非標的病変」のベースライン評価での記録

ベースライン評価において2個以上の測定可能病変を認める場合、すべての浸潤臓器を代表する、合計が最大5個（各臓器につき最大2病変）までの病変を標的病変として選択し、これらについてベースライン評価での測定値を記録する（すなわち、浸潤臓器が1臓器の場合は最大で2病変、2臓器の場合は最大で4病変を記録する）。選択する標的病変を5病変のみとすることを裏付けるエビデンスについては、Bogaertsら¹⁰の論文で記述されている大規模な前向きデータベースでの解析結果を参照のこと。

標的病変は、病変のサイズ（最大径が測定可能な病変）に基づいて選択され、すべての浸潤臓器を代表するものであるべきである。さらにこれに加えて、再現性をもった繰り返し測定が可能な病変でなければならない。時には、最大の病変が再現性のある測定に適さない場合もあるが、その場合は、再現性のある測定が可能な、次に大きな病変を選択すべきである。この点については、付録IIの図3の例を参照のこと。

リンパ節は、腫瘍の浸潤がない場合でも画像により描出され得る正常な解剖学的構造物であるため、特に言及しておく必要がある。3項で記述したように、測定可能と定義され、標的病変に選択され得る病的なリンパ節腫大 (pathological nodes) とは、CTでの短径が15 mm以上であるものでなければならない。これらのリンパ節は短径のみをベースライン評価の径の和に加える。リンパ節の短径は、固形がんによる浸潤の有無を判定するために、画像診断医が通常用いている径である。リンパ節のサイズは通常、撮影面上の2方向で報告される (CTの場合、これはほぼ常に横断面である。MRIでは、横断面、矢状断面、冠状断面のいずれもあり得る)。これらの測定値のうち小さい方の値が短径である。例えば、腹部リンパ節の測定値が20 mm × 30 mmの場合、短

径は20 mmであり、測定可能なリンパ節病変とされる。この例では、リンパ節のサイズの測定値として20 mmを記録すべきである (付録II図4の例も参照)。他の病的リンパ節腫大 (短径が10 mm以上15 mm未満) はすべて、非標的病変とされる。ベースライン評価にて短径が10 mm未満のリンパ節は病的ではない (病変ではない) とみなされるため、記録または追跡すべきではない。

ベースライン評価時の全標的病変の径の和 (以下、径和。非リンパ節病変では長径、リンパ節病変では短径) を、ベースライン径和として算出し報告する。上述のとおり、リンパ節の径を径和に含める場合は短径のみを加える。ベースライン径和は、その後の客観的な腫瘍縮小効果における比較対照 (基準) として用いられる。

標的病変以外の、リンパ節病変を含む他のすべての病変 (または病変部位) は非標的病変とし、これもベースライン評価時に記録する。これらの非標的病変は測定の必要はなく、「あり」、「なし」の別、また稀には「明らかな増悪」の有無 (詳細は後述) について評価する。また、同一臓器内の複数の非標的病変を1病変として効果判定記録用紙に記録してもよい (「複数の腫大骨盤リンパ節」や「多発肝転移」など)。

4.3. 効果判定規準

本項では、客観的な腫瘍縮小効果の判定に使用する規準の定義を示す。

4.3.1. 標的病変の評価

完全奏効 (Complete Response : CR):

すべての標的病変の消失。

標的病変として選択したすべてのリンパ節病変は、短径で10mm未満に縮小しなくてはならない。

※訳注：原著がここで (whether target or non-target) となっているのは論理的に明らかな誤りであり訳出せず、「標的病変として選択した」を加えた。たとえ非標的病変のリンパ節が正常化していなくても、すべての標的病変のリンパ節が正常化していれば「標的病変の効果」はCRとすべきである (当然「非標的病変の効果」はCRとはならないため、総合効果はCRとならない)。

部分奏効 (Partial Response : PR):

ベースライン径和に比して、標的病変の径和が30%以上減少。

進行 (Progressive Disease : PD):

経過中の最小の径和 (ベースライン径和が経過中の最小値である場合、これを最小の径和とする) に比して、標的病変の径和が20%以上増加、かつ、径和が絶対値でも5 mm以上増加。

※訳注：こども、原著にある (Note: the appearance of one or more new lesions is also considered progression) は、「標的病変の効果」、「非標的病変の効果」、「新病変の出現の有無」を別々に評価して統合するという根本の論理に反するため誤りであり訳出しなかった。「新病変あり」のため「総合効果」がPDとなる。

安定 (Stable Disease : SD):

経過中の最小の径和に比して、PR に相当する縮小がなくPD に相当する増大がない。

4.3.2. 標的病変の評価に関する注意点

リンパ節:

ベースライン評価時に、標的病変として選択したリンパ節病変は、経過中に 10 mm 未満に縮小した場合でも、短径の実測値 (ベースライン評価時と解剖学的に同一の断面において測定) を常に記録する。これは、「正常リンパ節」の短径が 10 mm 未満と定義されていることから、リンパ節を標的病変に含めた場合、完全奏効の規準が満たされた場合でも病変の径和が 0 ではないことがあることを意味する。したがって、CRF などのデータ収集ツールには、CR であることが確認できるように、径和の記載欄だけでなく、各リンパ節が短径 10 mm 未満になっていることが確認できる記載欄を設ける方がよい。PR、SD、PD の場合は、各リンパ節の短径の実測値を標的病変の径和に加える。

「小さすぎて測定不能 (too small to measure)」となった標的病変:

経過中は、ベースラインで記録したすべての病変 (リンパ節病変か腫瘍病変かを問わず) は、たとえ非常に小さくなってしまった場合 (例えば 2 mm) でも、各評価時点での実測値を記録する。しかし、ベースラインで標的病変として記録された病変やリンパ節が、CT できわめてかすかにしか描出されなくなってしまうことがあり、その場合に画像診断医が正確な測定値を出せずに「小さすぎて測定不能 (too small to measure)」とレポートすることがあるかもしれない。このような場合でも、何らかの値をCRFに記録することが必要である。おそらく病変は消失したであろうと画像診断医が判断する場合には、測定値は 0 mm として記録する。かすかに視認でき、病変が残存していると判断される場合には 5 mm のデフォルト値を割り当てる。

注:リンパ節: 正常の場合でも通常認識可能なサイズがあり、また後腹膜腔のように病変が脂肪で囲まれていることも多いため、実際に画像診断医が「小さすぎて測定不能 (too small to measure)」と報告する可能性は低いと思われる。しかし、リンパ節においても、非リンパ節病変と同様に、かすかに視認でき、病変として残存していると判断される場合には、5

mm のデフォルト値を割り当てる。

このデフォルト値は CT のスライス厚の標準を 5 mm としていることによる (ただし、CT のスライス厚が他の値の場合でも、このデフォルト値は変更すべきではない)。こうした病変の測定値には再現性がない可能性が高いため、このデフォルト値を定めることにより、測定誤差に由来する効果や増悪の誤判定を回避することができる。ただし繰り返しとなるが、画像診断医が実測値を示せる場合には腫瘍径が 5 mm 未満であってもその値を記録すべきである。

治療中に分裂または融合した病変:

付録 II で記述するように、腫瘍病変が分裂して「断片」になった場合、標的病変の径和の算出に際しては、断片化した部分の長径を合わせて加算する。同様に、病変が融合しているが、病変間の境界面が同定可能な場合には、個々の病変の長径の和を径和に加える。一方、病変が真に融合して境界面の同定が不可能となった場合には、「融合病変」としての最大径となる径を長径として径和に加える。

4.3.3. 非標的病変の評価

本項では、非標的病変における腫瘍縮小効果の判定に使用する規準の定義を示す。非標的病変の中には実際には測定可能なものもあるが、それらを測定する必要はなく、各研究のプロトコールで規定されている時期に定性的な評価のみを行う。

完全奏効 (Complete Response : CR):

すべての非標的病変の消失かつ腫瘍マーカー値が基準値上限以下。すべてのリンパ節は病的腫大とみなされないサイズ (短径が 10mm 未満) とならなければならない。

非 CR/非 PD (Non-CR/Non-PD):

1つ以上の非標的病変の残存かつ/または腫瘍マーカー値が基準値上限を超える。

進行 (Progressive Disease : PD):

既存の非標的病変の明らかな増悪 (次項コメントを参照)。

※訳注：先述のごとく原著の (Note: the appearance of one or more new lesions is also considered progression) は論理的に明らかな誤りであり訳出しなかった。

4.3.4. 非標的病変の増悪の評価に関する特別の注意点

非標的病変の増悪の概念について下記に説明する。

測定可能病変を有する場合:

この場合、標的病変の効果が SD や PR であっても、非

標的病変の変化に基づいて「明らかな増悪」と判定されるためには、全体の腫瘍量の増加として治療を中止するに十分値する程度の、非標的病変の著しい増悪が観察されなければならない(付録IIの例および以下の詳細を参照)。1つ以上の非標的病変のサイズが若干「増大」しても、通常は「明らかな増悪」とは判定しない。したがって、標的病変の効果がSDやPRの場合に、非標的病変の変化のみに基づいて全体の効果として増悪と判定することは、きわめて稀である。

測定不能病変のみを有する場合：

測定可能病変を1つ以上有することが適格条件ではない第III相試験では、このような状況が起こり得る。上記と同じ原則が適用されるが、この場合、測定可能病変の評価結果を、測定不能病変の腫瘍量増大の解釈に加えることができない。(すべての病変が真に測定不能であれば定義上自明のこととして)非標的病変の増悪の定量的評価は容易ではないため、「明らかな増悪」と判定する際に適用され得る有用な方法として、測定不能病変の変化に基づく全体の腫瘍量の増加の程度が、測定可能病変のPDを判定する際に必要とされる増加量に匹敵するかどうかを判断するという方法があり得る。すなわち、腫瘍量の増加が「体積」として73%の増加に相当するかどうかを判断するのである(これは、測定可能病変では径の20%増加と同等である)。このような例としては、胸水量の「微量」から「大量」への増加、限局していたリンパ管症の広範な拡大などがあり、また、プロトコールに「治療の変更を要するに十分な増悪」として表現しておくこともできる(付録IIの図5、図6参照)。「明らかな増悪」が認められた場合、その時点で総合効果はPDとされるべきである。測定不能病変に適用する客観的な規準を定めることが理想ではあるが、病変の性質上それは不可能であり、測定不能病変における増大とは「顕著な(substantial)ものである」とせざるを得ない。

4.3.5. 新病変 (new lesions)

新しい腫瘍性病変の出現は疾患の増悪を示すことから、新病変の検出に関して言及しておくことは重要であろう。放射線画像での新病変の同定に関する特別な定義はないが、新病変の所見は明らかな(unequivocal)ものでなければならない。すなわち、撮影方法の相違や画像モダリティの変更による変化や、腫瘍以外の何かを示すと考えられる所見(例えば「新しい」骨病変の中には、既存の病変の治療の過程や一過性の見かけ上の増悪を見ているに過ぎないものがあり得る)であってはならない。このことは、ベースライン評価で指定した病変が部分奏効または完全奏効を示している場

合に、特に注意しなければならない。例えば、肝病変の壊死が、実際には新病変ではないのにCT読影報告書で「新しい」嚢胞性病変として報告されることもあり得る。

ベースライン評価では撮影されなかった臓器や部位において、経過の検査で病変が同定された場合、それは新病変とみなされ、増悪と判定される。このような例として、ベースライン評価では体幹部の病変が認められた患者において、試験中に脳のCTまたはMRIが実施され、転移が認められた場合がある。この患者の脳転移は、たとえベースライン評価の脳画像がない場合でもPDの証拠とみなされる。

新病変が明確ではない場合(例：サイズが小さい)、治療を続けて再評価を行うことで真に新病変であることが明らかになることがある。撮影を反復した後に新病変と判定された場合、最初の撮影の日付をもって増悪とすべきである*。

FDG-PETによる効果の評価はさらなる検証が必要ではあるが、増悪の評価(特に新病変の可能性がある時)においてCTを補完するためにFDG-PETを併用することが妥当とされる場合もある。FDG-PET画像に基づく新病変の評価は、以下の手順に従って行うことが可能である。

a. ベースライン評価でのFDG-PET陰性かつ経過時のFDG-PETが陽性*となった場合：

新病変として総合効果PDとする。

b. ベースライン評価ではFDG-PET不施行で、経過時にFDG-PETが陽性となった場合：

経過時のFDG-PET陽性が、CTで確認された新病変に対応する場合はPDとする。

経過時のFDG-PET陽性が、CTで新病変と確認されない場合は、当該部位で真の増悪か否かを判定するために、さらに経過観察後のCTの再検を要する(真に増悪だった場合、PD判定日はFDG-PETが最初に陽性を示した日とする)。CTの形態画像上は増悪と判断されなかった病変で、FDG-PETが陽性であってもそれはPDとしない。

* FDG-PET陽性病変：吸収補正画像においてFDG取り込みが周囲組織の2倍を超えるFDG集積病変を意味する。

※ 訳注：施設間や担当医間のバラツキを大きくする可能性の観点から訳者らはこの見解に賛同できない。米国のCooperative GroupであるThe Southwest Oncology Group (SWOG)や訳者らのJCOGではこれと正反対に、増悪が確定した時の検査日を増悪判定日としている。いずれにせよ、プロトコールに定義を明記し、結果の公表の際にも定義を記載すべきと考える。

4.4. 最良総合効果 (best overall response) の評価

最良総合効果とは、確定のための要件をすべて考慮にいたった上で、試験治療開始から治療終了までの間に記録された最良の客観的腫瘍縮小効果を指す。治療が終了した後は効果判定が実施されないこともあるため、最良総合効果の

判定に治療終了後の評価を加える場合には、そのことをプロトコルに明確に記載する。また、増悪前後治療が開始された場合に最良総合効果の判定をどのように行うかについてもプロトコルに詳細に記載する。最良総合効果の判定は、標的病変および非標的病変双方の所見に基づくものであり、また新病変出現の有無も考慮される。さらに、試験の性質やプロトコルの規定によっては、効果の確定のための追加の検査が必要な場合もある(4.6項を参照)。効果が primary endpoint である非ランダム化試験において「最良総合効果」をPRまたはCRとするためには、それらの確定 (confirmation) が必要である。これについては後に述べる。

4.4.1. 各時点での効果

効果の判定は各プロトコルで定められた時点ごとに行われることが想定されている。

表1は、ベースラインで測定可能病変を有する患者について、各時点での総合効果の決め方をまとめたものである。

測定不能病変(すなわち非標的病変)のみを有する場合には、表2を使用する。

4.4.2. 評価の欠損および評価不能の規定

ある時点において画像検査/測定が全く行われなかった場合、その時点の効果は「評価不能(NE)」となる。評価において一部の病変の評価しか行われなかった場合にも、その時点の効果は通常NEとする。ただし、評価されなかった病変を考慮しても、その時点の判定された効果が変わらないことが、論理的に明らかな場合はこの限りではない。これはPDの場合に生じる可能性が高いと考えられる。例えば、ベース

表1 各時点での効果：標的病変(非標的病変の有無にかかわらず)を有する場合

標的病変	非標的病変	新病変	総合効果
CR	CR	なし	CR
CR	Non-CR/non-PD	なし	PR
CR	評価なし	なし	PR
PR	Non-PD or 評価の欠損あり	なし	PR
SD	Non-PD or 評価の欠損あり	なし	SD
評価の欠損あり	Non-PD	なし	NE
PD	問わない	あり or なし	PD
問わない	PD	あり or なし	PD
問わない	問わない	あり	PD

CR：完全奏効、PR：部分奏効、SD：安定、PD：進行、NE：評価不能

ラインにおいて3病変が測定されて径和が50 mmであった場合に、経過時には2病変のみが評価されたが径和が80 mmであった場合、評価されなかった病変が仮に評価されていたとしても総合効果がPDであることには変わりがない。

4.4.3. すべての評価時点を通じての最良総合効果 (best overall response)

患者の全データが得られた時点で、最良総合効果が決定される。

完全奏効や部分奏効の確定が必要ではない試験における最良総合効果の判定：

こうした試験における最良総合効果は、全時点を通しての最良の効果と定義される(例えば、最初の評価でSD、2回目の評価でPR、最終評価でPDなら、最良総合効果はPRである)。最良総合効果をSDとする場合には、プロトコルで定められたベースラインからの最短期間の規準をも満たさなければならない。最短期間の規準が満たされないことを除いてSDが最良の効果である場合、最良総合効果はその次の評価により異なる。例えば、最初の評価でSD、2回目の評価でPDの時、SDの最短期間の規準を満たしていない場合には、最良総合効果はPDとなる。同様の患者が、最初のSD評価の後に追跡不能となった場合には、最良総合効果は「評価不能」とされる。

完全奏効や部分奏効の確定が必要とされる試験における最良総合効果の判定：

完全奏効や部分奏効は、プロトコルで定められた、次の評価時点(通常は4週後)においても、それぞれの規準が満たされた場合にのみ判定することができる。確定を要する場合、最良総合効果は、表3のように決められる。

表2 各時点での効果：非標的病変のみを有する場合

非標的病変	新病変	総合効果
CR	なし	CR
Non-CR/non-PD	なし	Non-CR/non-PD ^a
評価なしがある	なし	NE
明らかな増悪	あり or なし	PD
問わない	あり	PD

CR：完全奏効、PD：進行、NE：評価不能

a いくつかの試験では有効性評価のエンドポイントとしてSDの使用が増えており、測定可能病変がない場合にこのカテゴリーを適用することは推奨されないため、非標的病患に関しては「安定」よりも「非CR/非PD」の方が望ましい。

4.4.4. 効果判定に関する特別の注意点

リンパ節病変の短径は標的病変の径和に含められるが、そのリンパ節が「正常」のサイズ (10 mm 未満) まで縮小した場合でも画像上の測定値が報告されることがある。リンパ節のサイズが増大した際に過大評価して増悪としないために、リンパ節が正常化した場合でもその測定値を記録するべきである。前述のとおり、これは CR の場合にも CRF に記載された径和が「0」にはならない場合があることを意味する。

効果の確定が必要な試験においては、「NE」が繰り返された場合、最良総合効果の判定が複雑になると考えられる。試験の解析計画書では、効果や増悪の判定におけるデータ/評価の欠損の扱い方について言及しておかなければならない。例えば、大半の試験においては、時点ごとの効果が PR-NE-PR であった患者において、効果が確定したとみなすことは妥当である。

健康状態の全体的な悪化 (病状悪化) により、その時点での客観的な増悪の証拠が得られないまま、治療の中止が必要となった場合は「病状悪化 (symptomatic deterioration)」として報告されるべきである。治療中止後においても、増悪の客観的な証拠を得るためのあらゆる努力を払うべきである。「病状悪化」は客観的な効果の 1 カテゴリではなく、試験治療の中止理由である。このような患者における客観的な効果は、表 1~3 に記載のとおり標的病変と非標的病変の評価により判定される。

「早期増悪、早期死亡、評価不能」と定義すべき状態は試験により異なるため、プロトコールに明確に記載されるべき

である (治療期間の長さや治療コースの間隔により異なる)。

状況によっては、残存病変と正常組織の区別が困難なことがある。完全奏効とするか否かがこの区別に左右される場合には、完全奏効と判定する前に残存病変の検査 (穿刺吸引細胞診や生検) を行うことが推奨される。放射線画像における残存病変の異常が線維化や癒痕を示すものと考えられる場合、生検の場合と同様、効果の判定を CR に格上げするのに FDG-PET を使用してもよい。ただし、このような状況下での FDG-PET の使用については、予めプロトコールに記載し、その適用に際して裏付けとなる疾患特異的な医学論文を挙げるべきである。ただし、FDG-PET や生検の解像度/感度には限界があるため、いずれの方法も CR の偽陽性を導く可能性があることを認識しておかなければならない。

増悪の所見が不明確な場合 (きわめて小さく不明瞭な新病変、既存病変の嚢胞性変化や壊死など)、次に予定された評価時点まで治療を継続してもよい。予定された次の評価において増悪が確定した場合、増悪判定日は先に増悪が疑われた日*にすべきである。

※ 訳注: 「4.3.5. 新病変」の訳注で述べたように、訳者らはこの見解に賛同できない。

4.5. 腫瘍の再評価の頻度

治療中の効果判定の頻度は、治療の種類やスケジュールに応じてプロトコールごとに決められるべきものである。ただし、治療の有益な効果が現れる時期が未知である第 II 相試験においては、6~8 週毎 (各サイクルの終了時と一致した時

表 3 CR と PR の確定が必要とされる場合の最良総合効果

最初の総合評価	その次の総合評価	最良総合効果
CR	CR	CR
CR	PR	SD, PD or PR ^a
CR	SD	SD の最短規準を満たせば SD、それ以外は PD
CR	PD	SD の最短規準を満たせば SD、それ以外は PD
CR	NE	SD の最短規準を満たせば SD、それ以外は NE
PR	CR	PR
PR	PR	PR
PR	SD	SD
PR	PD	SD の最短規準を満たせば SD、それ以外は PD
PR	NE	SD の最短規準を満たせば SD、それ以外は NE
NE	NE	NE

CR: 完全奏効、PR: 部分奏効、SD: 安定、PD: 進行、NE: 評価不能

a 最初の時点で本当に CR であった場合、それ以降の時点で病変が認められると、それがベースラインとの比較で PR 規準を満たす病変であっても、その時点で PD となる (CR 後に病変が再出現したとみなすため)。その場合の最良総合効果は SD の最短期間の規準が満たされていたか否かによって決まる。ただし、いったん「CR」と判定された場合でも、その次の検査で小病変が残存していたと判断され、最初の評価時点では実際には CR ではなく PR であったということが後から考察されることがある。そうした場合には、最初の CR は PR に変更すべきであり、最良総合効果は PR となる。

期)の効果判定が必要である(特定のレジメンや状況によっては、これよりも短いまたは長い間隔での判定が適切とされる場合もある)。プロトコールには、ベースライン評価で測定すべき臓器部位 (organ sites) (対象によって高頻度に転移が起こるとされている部位)、効果判定の反復頻度(検査間隔)を明記すべきである。通常は、各効果判定時点ですべての標的病変と非標的病変の部位を撮影および測定する。特定の状況においては、ある非標的病変の臓器部位をより低い頻度で判定してもよい。例えば、骨シンチは、標的病変の効果が完全奏効である場合や骨病変の増悪が疑われる場合にのみ必要とするといったやり方である。

治療終了後も効果判定を繰り返す必要があるか否かについては、試験の目的が奏効率の評価であるか、生存時間(time to an event: 増悪/死亡)の評価であるかによって異なる。無増悪期間、無病生存期間、無増悪生存期間などの生存時間が試験の主なエンドポイントである場合、プロトコールに定められた病変部位の効果判定を決められたスケジュールで行うことが重要である。特にランダム化比較試験においては、予定された効果判定はスケジュール表に記された通りに実施されるべきであり(例えば、治療期間中は6~8週毎、治療終了後は3~4か月毎など)、効果判定の時期が治療群間で偏る原因となるような事象によって左右されるべきではない(例えば「2コース毎」のように、治療の遅延や休薬などによって評価時期が異なる決め方は避けるべきである)。

4.6. 確定のための測定 / 奏効期間 (confirmatory measurement/duration of response)

4.6.1. 確定 (Confirmation)

腫瘍縮小効果が primary endpoint である非ランダム化試験においては、判定された効果が測定誤差による結果ではないことを保証するために、PR および CR の確定が必要である。このようにすることで、そうした試験において、従来、効果の確定が必要とされていたヒストリカルコントロールとの比較において結果を適切に解釈することができる(本特集号に掲載した Bogaerts らの論文を参照¹⁰⁾)。しかし、それ以外のすべての状況、すなわちランダム化試験(第II相、第III相)や、安定(SD)もしくは増悪が primary endpoint である試験においては、効果の確定は試験結果の解釈に対して価値を追加するものとはならないため、効果の確定は不要である。ただし、盲検化されていない試験においては特に、効果の確定が必要とされなくなったことにより、バイアスを回避するための中央判定(central review)※の重要性が増すと考えられる。

SDの場合、試験登録後、試験プロトコールで定義される最短の間隔(通常は6~8週間以上)を経た時点までに測定値が1回以上SDの規準を満たさなければならない。

※訳注:原著では‘central review’をより広義に用い、‘central

review’のうち試験から独立した第三者の review を意味する時に‘independent review’と使い分けているため、‘central review’を「中央判定」、‘independent review’を「第三者による再判定」と訳した。

4.6.2. 奏効期間 (duration of overall response)

奏効期間は、CRまたはPR(最初に記録された方)の測定規準が最初に満たされた時点から、再発または増悪が客観的に確認された最初の日までの期間である(試験中に記録された最小測定値を増悪の比較対照とする)。

完全奏効期間は、CRの測定規準が最初に満たされた時点から、再発が客観的に確認された日までの期間である。

4.6.3. 安定期間 (duration of stable disease)

安定(SD)は、経過中の径和の最小値を比較対照として(ベースラインの径和が最小の場合はこれを比較対照とする)、治療開始(ランダム化試験の場合、ランダム割付日)から増悪の規準が満たされた時点までの期間とする。

※訳注:原著には明記されていないが、安定期間は、最良総合効果がSD以上であった患者を対象とするエンドポイントであると解釈できる。全登録患者や全適格患者を対象とするならば、次項の無増悪生存期間と(画像によらない臨床的増悪をイベントとするか否かのみが違う)ほとんど同一のエンドポイントになってしまうためである。同様に、奏効期間は最良総合効果がCRまたはPRであった患者を対象とすることになる。

安定期間の臨床的な重要性は試験や疾患により異なる。一定期間の安定(SD)が得られた患者の割合が重要なエンドポイントであるような試験では、プロトコールに、SDを判定する2回の測定間の最短の間隔を明記する必要がある。

注:奏効期間、安定期間、無増悪生存期間は、ベースライン評価後の効果判定の頻度の影響を受ける。効果判定の頻度はプロトコールに明記すべきことであり、標準的な効果判定の頻度を定義することは、本ガイドラインが目的とするものではない。効果判定の頻度の設定には、がんの種類、病期、治療周期、標準的な日常診療など、多くの要素を考慮に入れなければならない。ただし、試験間での比較を行う場合には、評価に用いるエンドポイントの精密度に関するこうした限界も考慮しなければならない。

4.7. 無増悪生存期間/無増悪生存割合 (progression-free survival/proportion progression-free)

4.7.1. 第II相試験

本改訂ガイドラインの主たる目的は、第II相試験のエンドポイントとしての、客観的な腫瘍縮小効果の使用について解説することである。状況によっては、「奏効率※」が新規薬剤や新規レジメンの抗腫瘍効果を評価する方法として最適ではない場合がある。奏効率が最適ではない場合、「無増悪

生存期間 (PFS) や、ある特定の時点での「無増悪生存割合 (proportion progression-free)」が新規薬剤の生物学的効果に関する最初の結果を示すのに適切な代替指標となる可能性がある。しかし、無対照試験においては、有望な観察結果が観察されたとしても、それが介入によるものではなく患者選択などの生物学的因子による可能性があることから、これらの指標が批判の対象となり得ることは明らかである。したがって、これらのエンドポイントを用いた第 II 相スクリーニング試験にはランダム化デザインが最適である。例外的に、ある特定のがんでは治療反応性や予後が一貫している (一貫して予後不良である) ことから、非ランダム化試験が正当化される場合 (例えば van Glabbeke らの報告²⁰を参照) もある。しかしこうした場合にも、治療効果がない時に期待される PFS や無増悪割合を推定する根拠について慎重に記述する必要がある。

※訳注：疫学・統計学用語としての「率 rate」は、「死亡率」や「ハザード」のように分母に時間の尺度を含む、「速度」の概念を持つ指標に用いられるが、「response rate」や「survival rate」は時間の尺度を含まないため本来は「割合 proportion」が正しい。しかし、世界中で「response rate」が用いられ、我が国でも「奏効率」が定着していることから、本稿では「response rate」を「奏効率」と訳した。ただし「survival rate」は response rate ほどには定着しておらず、文献にも「survival proportion」や「proportion surviving」、「% surviving」といった表現も見られることから「生存割合」と訳した。

4.7.2 第 III 相試験

進行がんを対象とする第 III 相試験において、無増悪生存期間や無増悪期間を主たるエンドポイントに用いることが増えてきている。プロトコルで全患者が測定可能病変を有することが選択条件とされている場合には、増悪の評価は比較的明確なものになると言える。しかし、測定可能病変を有する患者のみに登録を限定することが、対象とする集団からかなりの割合の患者を除外することになるのであれば、試験結果の一般化可能性が低いという批判を受けることになる。さらに、登録を制限することにより患者登録も遅くなる。したがって、無増悪生存期間などを主たるエンドポイントに用いる場合であっても、測定可能病変を有する患者と測定不能病変のみを有する患者の両方の登録を許容する試験が増加している。このような状況下では、測定可能病変を有さない患者で「PD」と判定する根拠となる所見を明確に記述するための配慮が必要となる。

また、第 III 相試験では、測定可能病変を有する患者で登録する標的病変の最大数を5つから3つに減らすことも許容されるが、その場合はプロトコルに明記する (Bogaerts ら¹⁰、および Moskowitz ら¹¹のデータに基づく)。「増悪の評価に関する特別の注意点」にあるように、これらのガイドラインではこうした方法 (標的病変の最大数を減らす) での増悪の評価

を推奨している。さらに、利用可能ならば、卵巣がんが提示されたように増悪に関して検証された腫瘍マーカー測定値を増悪の定義に組み込むことも有用と考えられる。

薬剤開発や薬剤承認の重要な意思決定が、こうした増悪の評価に基づく試験結果によりなされるならば、「明らかな増悪 (unequivocal progression)」を確認するための、画像検査または画像 (読影) 報告書原本に対する盲検化された中央判定が必要になるだろう。

最後に、前述のとおり、増悪判定日は確認バイアス (ascertainment bias) を受けやすいため、治療群ごとの検査のタイミングを揃えるべきである。本特集号に掲載されている Dancey らによる論文²¹では、ランダム化試験における増悪の評価について、より詳細に議論されている。

4.8. 効果や増悪に関する第三者による再判定 (independent review of response and progression)

客観的な腫瘍縮小効果 (CR + PR) が primary endpoint である試験では、(観察された奏効患者数が最低限の規準を上回るかどうかに基づいて薬剤開発の重要な意思決定が下される試験では特に) 担当医判定によるすべての奏効が試験から独立した専門家により再判定 (review) されることが推奨される。ランダム化試験の場合には、判定者は割り付け群について盲検化されることが望ましい。患者の臨床情報と放射線画像を同時に検討することが最良の方法である。

増悪に関する第三者による再判定には、より複雑な問題がある。例えば、第三者による再判定にて決定した増悪判定日を、担当医判定による増悪判定日の代わりに用いる際、前者が後者よりも先になる場合には情報のある打ち切り (informative censoring) になる可能性があることから統計学的な問題が生じる^{*}。これらの要因や、第三者による再判定から得られるその他の教訓に関して、本特集号に掲載した Ford らによる論文²²に概要が記されている。

※訳注：Ford らの論文では、担当医判定による増悪が、中央判定で増悪でないと判断された場合、以降の画像検査がされていないために、PFS が担当医の増悪判定日で打ち切りとなるという例が示されている。この場合、真の増悪日の少し前で「増悪ではない」と判断されることになり、PFS が過大評価されるバイアスを生じ得る。

4.9. 最良総合効果に関する結果の報告 (reporting best response results)

4.9.1. 第 II 相試験

客観的な腫瘍縮小効果 (CR + PR) が primary endpoint であり、そのためすべての患者が測定可能病変を有している試験では、治療に関する重大なプロトコル逸脱があった場

合や評価不能であった場合でも、試験に登録されたすべての患者を結果の報告に含めなければならない。それぞれの患者は以下のカテゴリーのいずれかに分類される。

1. 完全奏効 (Complete response)
2. 部分奏効 (Partial response)
3. 安定 (Stable disease)
4. 増悪 (Progression)
5. 効果の評価不能 (Inevaluable for response)：理由を明示
(例：悪性腫瘍による早期死亡、毒性による早期死亡、腫瘍評価が反復されず/不完全、その他(具体的に))

通常、第II相試験ではすべての適格患者を奏効率算出の分母に加えるべきである(プロトコールによっては、治療されたすべての患者を分母とすることが適切な場合もある)。一般論として、奏効率の計算値には両側95%信頼区間を付すことが望ましい。試験の結論は、すべての適格患者(または治療されたすべての患者)の奏効率を根拠とすべきであり、選択された「評価可能」なサブセットでの結果を根拠とすべきではない。

4.9.2. 第III相試験

第III相試験における客観的な腫瘍縮小効果は、対象とする治療の相対的な抗腫瘍効果の指標となり得るが、ほとんどの場合、副次的なエンドポイントとして評価される。観察された奏効率の差は、対象集団に対して臨床的に意味のある治療のベネフィットを予測するものではない。第III相試験にて客観的な腫瘍縮小効果を primary endpoint として選択した場合(客観的な腫瘍縮小効果と臨床的に意味のある治療のベネフィットとの間に直接的な因果関係があることが、試験の対象集団において明確に実証されている場合に限る)は、第II相試験に適用される規準と同様の規準を用い、すべての登録患者は測定可能病変を1つ以上有していなければならない。

多くの場合、すなわち腫瘍縮小効果が副次的なエンドポイントであって、すべての試験対象患者が測定可能病変を有しているわけではない場合であっても、最良総合効果の報告の方法については事前にプロトコールに明記しておかなければならない。実際には、奏効率はITT (intent to treat) 解析(ランダム化したすべての患者を分母に加える)か、ベースライン評価で測定可能病変を有していた患者サブセットのみを対象とする解析のいずれかによるものを報告する。プロトコールには、予定しているサブセット解析を含めて、効果に関する結果の報告の方法について明記しておくべきである。

RECIST1.0では、第III相試験におけるRECISTガイドラインを「緩やかに (relaxed)」解釈してプロトコールを作成す

ることもできると提案した(例：測定する病変の数を減らすなど)。しかし、本改訂ガイドラインは、腫瘍縮小や増悪の形態の評価がエンドポイントであるすべての試験に対して、これらの規準をどのように適用すべきかを明確にするために作成されたものであり、本改訂ガイドラインではRECIST1.0で行ったような提案はもはや行わない。

■ 付録 I RECIST1.0 から RECIST1.1 への改訂における主な変更点の要約 ■

	RECIST 1.0	RECIST 1.1	Rationale 根拠	Reference in special issue (if applicable) 特集号中の 参考文献 (該当する場合)
測定可能病変の最小サイズ	CT: ヘリカル CT で 10 mm ヘリカル CT 以外で 20 mm	CT で 10 mm ヘリカル CT に関する記載を削除	大半の撮影のスライス厚 (slice thickness) は 5 mm 以下である スライス間隔 (slice interval) が 5 mm を超える場合にはスライス間隔に基づく規定の方が明確である	
	臨床: 20 mm	臨床: 10 mm (測径器で測定可能でなければならない)	測径器で測定することにより信頼性が高まる	
リンパ節: 言及せず	CT:	CT: 標的病変: 短径が 15 mm 以上 非標的病変: 10 mm 以上 15 mm 未満 10 mm 未満の場合は病変としない	リンパ節は正常な構造物であるため、病的な腫大を定義する必要がある 短径の感度が最も高い	Schwartzら ¹⁵
病変の測定可能性に関する特別の考察	—	骨病変、嚢胞性病変に関する注記を含めた	よくある質問を明記した	
全体の腫瘍量	10 病変 (1 臓器につき 5 病変)	5 病変 (1 臓器につき 2 病変)	データウェアハウスの分析から、病変数を 10 から 5 に減少しても情報量は減少しないことが明らかになっている 1 臓器につき最大 2 病変で、各病変部位を十分に表すことができる	Bogaertsら ¹⁰
標的病変の効果判定規準	リンパ節の CR には言及せず	リンパ節の CR は、短径が 10 mm 未満にならない	リンパ節の正常なサイズに従った	Schwartzら ¹⁵
	経過中の最小の径和からの 20% 増加を PD とする	経過中の最小の径和 (ベースラインの和が最小の場合はそれを基準とする) からの 20% 増加かつ 5 mm 以上の増加を PD とする	ベースラインの測定値が試験中の他の測定値よりも小さい場合、これを PD 評価の基準とすることを明確にした 径和がきわめて小さく、20% 増加が測定誤差の範囲内である場合の PD の過大判定を避けるために、絶対値で 5 mm の増加を必要とした	Schwartzら ¹⁵
非標的病変の効果判定規準	「明らかな増悪」を PD とみなす	「明らかな増悪」についてより詳細に記述し、通常はこれが標的病変の状態より優先されないようにした 1 病変の増悪ではなく、疾患の全体的な状態の変化を表すものでなければならない	RECIST1.0 では、標的病変が安定や奏効でも、非標的病変が「増悪」した場合に PD とみなされるという混乱が見られた	
新病変	—	新病変に関する新しい項目	ある病変が新病変とみなされる (したがって PD とされる) ための指針を提示した	

	RECIST 1.0	RECIST 1.1	Rationale 根拠	Reference in special issue (if applicable) 特集号中の 参考文献 (該当する場合)
総合効果	標的病変と非標的病変の組み合わせによる表	2つの表：1つは標的病変と非標的病変の組み合わせによる表、もう1つは非標的病変のみの表 特別な注意： リンパ節の評価法および測定法 遺残組織がある場合のCR 「明らかな」増悪についての考察	PFSがエンドポイントで、必ずしも全患者がベースラインで測定可能（標的）病変を有していない試験でも、RECISTが使用されているという現状を考慮した これらのトピックスについて多くの質問があった	Danceyら ²¹
確定のための測定	CRおよびPRについて： 最初に記録されてから4週以降に、再び規準が満たされなければならない	効果がprimary endpointである非ランダム化試験についてのみこの要件が残る	データウェアハウスの検討では、確定を行わない場合に奏効率が高くなることが示されているが、これが重要になるのは、同時比較対照がなく、効果がprimary endpointである試験においてのみである	Bogaertsら ¹⁰
無増悪生存期間	一般的なコメントのみ	第II相試験のエンドポイントとしてのPFS（または無増悪生存割合）の使用についてより詳細にコメント 第III相試験におけるPFSの評価についてより詳しく記述した	第III相試験におけるPFSの使用が増えているため、測定不能病変しか有さない患者におけるPDの評価についての指針が必要となった	Danceyら ²¹
効果に関する結果の報告	第II相試験の結果報告について、9つのカテゴリーを提案	第II相と第III相試験に分けて記述 9つのカテゴリーを5つにまとめた 第III相試験については、効果の報告についての指針を掲載	第II相と第III相のデータを一貫性のある形で報告する方法を簡略化し明確化した	
第III相試験における効果	プロトコールに明記されていれば、より緩やかな指針も可	この項目は削除し、上記の項で言及した 第II相と第III相で異なる規準を設ける必要はない	効果がprimary endpointではないランダム化試験では、病変数を減らしかつ確定を不要としたことで、効果の評価が簡便となり、個別の「ルール」は不要となった	
画像に関する付録	付録I	付録II：MRIとPET/CTの使用に関する詳細な指針を加えて更新した 他の実践的な指針も含む	新しいモダリティの使用が広まっていること に対処した よくある質問や画像の中央判定の経験を踏まえて、指針を強化した。	
新しい付録		付録I：RECIST1.0と1.1の比較 付録III：よくある質問		

利益相反

特になし

謝辞

RECIST ワーキンググループは、RECIST の今回の改訂における変更役に役立つ解析が行えるようデータベースを提供してくれた以下の組織に深謝する。

Amgen;
AstraZeneca;
Breast Cancer International Research Group (BCIRG);
Bristol-Myers Squibb;
European Organisation for Research and Treatment of Cancer (EORTC) Breast Cancer Group and Gastrointestinal Group;
Erasmus University Medical Center, Rotterdam, The Netherlands;
Genentech;
Pfizer;
RadPharm;
Roche;
Sanofi Aventis.

また、改訂ガイドラインのドラフトに対して有用なコメントを寄せてくれたアカデミア、政府、製薬企業の、以下の方々にも深謝したい。

Ohad Amit, Phil Murphy, Teri Crofts and Janet Begun, GlaxoSmithKline, USA;
Laurence H. Baker, Southwest Oncology Group, USA;
Karla Ballman, Mayo Clinic, USA;
Charles Baum, Darrel Cohen, and Mary Ashford Collier, Pfizer, USA;
Gary J. Becker, American Board of Radiology, Tucson, USA;
Jean-Yves Blay, University Claude Pertrand, Lyon France;
Renzo Canetta, Bristol-Myers Squibb, USA;
David Chang, Amgen Inc., USA;
Sandra Chica, Perceptive Informations Inc. (PAREXEL), USA;
Martin Edelman, University of Maryland Greenbaum Cancer Centre, USA;
Gwendolyn Fyfe, Genentech, USA;
Bruce Giantonio, Eastern Cooperative Oncology Group, USA;
Gary Gordon, Abbott Pharmaceuticals, USA;
Ronald Gottlieb, Roswell Park Cancer Institute, USA;
Simon Kao, University of Iowa College of Medicine, USA;
Wasaburo Koizumi, Kitasato University, Japan;
Alessandro Riva, Novartis Pharmaceuticals, USA;
Wayne Rackhoff, Ortho Biotech Oncology Research and Development, USA;
Nagahiro Saijo, President Japanese Society of Medical Oncology, Japan;
Mitchell Schnall American College of Radiology Imaging Network, USA;
Yoshiki Shimamura, PAREXEL International Inc., Japan;
Rajeshwari Sridhara, Centre for Drug Evaluation and Research, Food and Drug Administration, USA;
Andrew Stone, Alan Barge, AstraZeneca, United Kingdom;
Orhan Suleiman, Centre for Drug Evaluation and Research, Food and Drug Administration, USA;

Daniel C. Sullivan, Duke University Medical Centre, USA;
Masakazu Toi, Kyoto University, Japan;
Cindy Welsh, Centre for Drug Evaluation and Research, Food and Drug Administration, USA.

最後に、本稿に共著者として挙げているワーキンググループの恒常メンバーではないが、7年にわたって何度もワーキンググループミーティングに出席して貴重な貢献をしてくれた以下の諸氏にも感謝申し上げたい。

Richard Pazdur, Food and Drug Administration, USA;
Francesco Pignatti, European Medicines Agency, London, UK.

■ 付録II 標準的な形態学的放射線画像診断に関する仕様 ■

CT および MRI の撮影に関する以下のプロトコール (撮影手順) は、RECIST による評価が行われる臨床試験における推奨事項をまとめたものである。試験内での被験者間の比較可能性および試験間の結果の比較可能性を最適なものとするためには、画像に関する要件と撮影に関するパラメータを標準化することが理想である。これらの推奨事項は、撮影に関する最適なプロトコールと、世界各国のあらゆる放射線画像の撮影施設において実施可能な技術との間でのバランスを取るようデザインされたものである。ただし、これらの指針は、機能画像の撮影技術および腫瘍サイズの 3 次元評価には適用されない。

撮影装置の品質管理は強く推奨される。製造業者と施設による標準的な保守点検スケジュールに従って、市販のファントムを用いた品質管理が実施されるべきであり、RECIST の 1 方向測定で再現可能な測定値を得るためには、この方法が適切であると考えられる。CT 画像の品質管理には、画像ノイズと画像の均質性、CT 値、空間分解能の分析が含まれる。品質管理のための分析の頻度は様々であってよいが、臨床的に意味のある撮影条件に焦点を当てるべきである。線量分析は常に重要であり、撮影を行う際には ALARA (As Low As Reasonably Achievable) 原則に従うべきである。ALARA 原則とは、被曝線量が線量限度を可能な限り下回るように維持するために、あらゆる適切な努力を行うことを意味する。

※ 訳注：本項の文章構成がわかりにくいいため、以下、訳者らの判断で項番号と項タイトルを付加した。

I. 具体的な注意事項 (Specific Notes)

1) 胸部 X 線

肺実質に囲まれた病変の胸部 X 線での測定を行ってもよいが、この測定値は密度の和を表すものであるため腫瘍径の測定値としてはふさわしくない。さらに、胸部の新病変の検出感度において X 線は CT に劣る。したがって、肺実質病変および縦隔病変の測定は、胸部 CT を用いて実施するべきである。胸部 MRI は特別な事情がある場合にのみ用いるべきである。経静脈性造影剤が投与できない場合 (例：造影剤に対するアレルギー) でも、MRI や胸部 X 線よりも胸部単純 CT の方が望ましい。

2) CT

胸部、腹部、骨盤の CT では、対象の解剖学的な領域全体を連続的に撮影する。原則として、ベースラインにおける測定可能病変の最小サイズは、スライス厚の 2 倍以上かつ 10 mm 以上でなければならない (スキャナーのスライス厚が

5 mm を超える場合の最小サイズについては、下記を参照)。病変サイズと部分体積効果との関係に関する正確な物理学的解釈は複雑であるが、10 mm 未満の病変では正確かつ再現性のある測定は困難であると言える。この原則はベースライン評価での撮影に適用されるものであるが、経過中の CT では病変が縮小して上記の最小サイズを下回る可能性もある。その場合でも径の測定は行うべきであり、さらに「小さすぎて測定不能 (too small to measure)」と報告された病変の径は、視認可能と判断された場合、既定値の 5 mm とする。

RECIST を用いた最適な腫瘍評価のための CT 撮影において最も重要なパラメータは、解剖学的な撮影範囲、造影剤投与の有無、スライス厚、再構成間隔である。

a. 解剖学的な撮影範囲 (anatomic coverage):

多くの固形がんにおける最適な解剖学的撮影範囲は、胸部、腹部、骨盤である。撮影範囲は、評価対象の疾患において転移が起きやすいことが知られている全領域を網羅し、かつ個々の患者の徴候や症状に基づいて浸潤が疑われる部位も追加で含める必要がある。ベースライン評価時に撮影されなかった部位において、事後的に検出された病変は、増悪を表す新病変とみなされることから、ベースライン評価および以降の経過での画像撮影の範囲の決定は慎重になされる必要がある。これにより、腫瘍の測定のみならず新病変の検出という点でも、より高い一貫性が得られる。

b. 経静脈性造影剤の投与 (IV contrast administration):

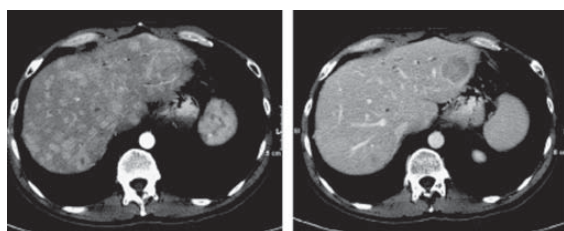
固形がんの転移を最適な形で描出し測定するには、経静脈性造影剤の投与 (用量と投与速度) と撮影のタイミングの一貫性が必要である。典型的には、大半の腹部画像は門脈相で撮影され、(理想的には) 各検査において造影剤の注入後同じ時間で撮影される (経静脈性造影時の時相の違いが病変の計測に及ぼす影響については図 1 を参照)。大半の固形がんでは造影剤投与後の 1 種類の時相で撮影される。3 時相 CT 撮影は、他の種類の多血管性腫瘍でも病変の描出能を上げる目的で使用されることがあるが、一貫性と均一性を得るために、特に肝細胞腫瘍と神経内分泌腫瘍では、これらの腫瘍に対して既に一般診療において標準プロトコールとなっている 3 時相 CT 撮影を用いることが推奨される。3 時相 CT 撮影による時間分解能の向上により画像診断医は一貫性があり再現性に優れた病変測定を行うことができる。

経静脈性造影剤の用量と投与速度は、CT 撮影装置、CT 撮影プロトコール、使用する造影剤の種類、利用可能

な静脈経路、患者の医学的状態によって決められる。このように、経静脈性造影剤の投与法は様々であるため、造影剤の投与法や投与量に関する厳密な規則を設けるよりも、転移巣が最も明瞭に描出されるように、個々の患者の経過中の検査で一貫した方法が使われるように、適切な造影剤が適量投与されることを促していく方が賢明であろう（理想的には、それらの方法がプロトコールに明記されるか施設ごとに設定される方がよい）。

さらに、個々の患者においてベースライン評価と経過中の検査で同じ方法を用いることがきわめて重要である。それによって腫瘍測定の新規性が大いに高められる。登録前に患者がアレルギーや腎不全のために経静脈性造影剤を用いた CT 撮影を実施できないことが予め明らかな場合、ベースライン評価および経過中に、単純 CT、造影 MRI、単純 MRI のいずれを用いるかは、試験対象の腫瘍の種類や疾患の解剖学的部位を考慮して決定する。また、ベースライン評価時の造影 CT 実施後に造影剤が禁忌となった場合も同様に、単純 CT、造影 MRI、単純 MRI のいずれを用いるかは、試験対象の腫瘍の種類や疾患の解剖学的部位に基づいて決定し、もし可能であればその時点までの検査との比較ができる最適なものを選択するべきである。このように他の方法での代替が可能か否かは、症例ごとに画像診断医と相談して判断する。代替が不可能な場合にはその患者でのその時点以降の効果は評価不能 (NE) とすべきである。

以上のように、異なるモダリティを用いた標的病変の測定、および非標的病変や新病変の解釈においては、新しいモダリティを用いた場合に同じ病変のサイズが異なっ



動脈相（左）と門脈相（右）における多血管性 (hypervascular) 転移巣。描出される病変数が造影剤投与後の 2 つの時相の間で大きく異なること、およびそれによって病変の測定も大きく異なることに注意が必要である。最適で再現性に優れた腫瘍測定のためには、造影剤投与後の時相も含めて、一貫した方法での CT 撮影が重要である

図 1 経静脈性造影剤投与時の時相の違いによる測定 / 描出に関する相違

描出されることがあるため注意が必要である（図 2：同一病変の CT 像と MRI 像の比較を参照）。また、腹腔内臓器の描出と臓器間の区別が重要である場合には経口造影剤の使用が推奨される。

c. スライス厚および再構成間隔 (slice thickness and reconstruction interval):

RECIST1.1 に従った測定は臨床的に最も多く使用されているスライス厚で行われることになる。CT 撮影は 5 mm 以下のスライス厚で連続的に行われることが推奨される。また本ガイドラインは、測定可能病変の定義において最小 5 mm のスライス厚が用いられることを想定している。実際、スライス厚の変化は病変の測定や新病変の検出に影響を及ぼし得るが、被曝量を最小限にする配慮も必要である。スライス厚を 5 mm とした場合、ベースライン評価における 10 mm 以上の病変が測定可能とされる。施設によっては、5 mm を超えるスライス厚を医学的に許容可能としているかもしれないが、その場合には、ベースライン評価における測定可能病変の最小サイズは、撮影に用いたスライス厚の 2 倍とすべきである。最新の CT 装置の大半はマルチスライス CT であり、撮影パラメータを変えることによって多くの種類のオプション画像が得られる²³。システムに特化した疑問がある場合には、機器の販売元に問い合わせるか装置のマニュアルを参照すべきである。

d. 代替造影剤 (alternative contrast agents):

多くの新しい造影剤があり、一部は臓器特異的である²⁴。それらの中には、例えば肝病変の評価やリンパ節の性状評価など²⁵、日常診療の一部として使用されているものもあるが、やはり臨床試験ではまだ使用すべきではない。

3) FDG-PET

FDG-PET は、いくつかのがんにおいては、病変の検出、病期分類、病期の再分類 (restaging) のための有用なツールとして受け入れられている。多くの研究がなされてはい

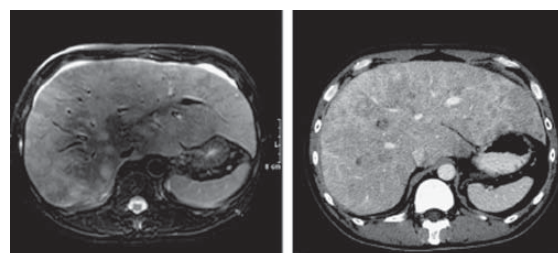


図 2 測定法の違いのみで見かけ上「増悪」と判断される同一病変の MRI 像(左)と CT 像(右)

るが、第 II 相試験における腫瘍縮小効果の形態的評価に FDG-PET を組み込む（またはこれに切り替える）ための規準は未だ確立していない。にもかかわらず、FDG-PET は多くの薬剤開発試験において、治療の有効性および増悪の評価手段として現在使用されている。FDG-PET をプロトコールに含める際には、画像撮影前に 60 分の FDG 取り込み時間を置くことが悪性腫瘍患者での撮影に最適であるとのコンセンサスが得られている²⁶。FDG-PET の全身撮影は、対象とすべき全領域を撮像することができ、また新病変が出現した場合にそれが前回の検査以降に生じた増悪であると判定できる点で、重要であると言える。頭蓋底から大腿中央部までの画像が注入後 60 分に撮影されるべきである。PET カメラの仕様は様々であり製造業者により異なるため、同じ患者の連続撮影では同じ装置または同じモデルの装置を使用しなければならない。全身撮影は、吸収補正下の 2 次元モードまたは 3 次元モードのいずれかで実施してもよいが、選択する方法は同じ臨床試験においてはすべての患者の一連の検査において一貫している必要がある。

4) PET/CT

PET/CT など、複数のモダリティを統合した検査が日常診療として普及しつつあり、急速な進化を遂げているモダリティ/技術となっている。したがって、本稿における推奨事項も時間と共に急速に変化していくと思われる。現在のところ、形態に基づく有効性評価における PET/CT の低線量 CT 像や吸収補正用 CT 像の使用は限られたものであるため、RECIST の形態に基づく測定において指定している診断用造影 CT の代替としてこれらを使用することは推奨されない。しかし、PET/CT の一部として実施される CT が、（静注や経口の造影剤を用いた）診断用 CT と同等の診断上の品質を有するものであることを施設が証拠書類をもって証明できる場合には、RECIST の測定に PET/CT の CT 像を使用することができる。ただし、ルーチンに使用されていない場合や一連の検査として実施されていない場合には、PET/CT の PET 像は研究者にバイアスを与え得る付加情報を追加する可能性に注意が必要である。

5) 超音波検査

超音波検査は、主観的で術者により異なることが避けられないため、臨床試験における腫瘍の縮小や病変の増悪の評価に用いるべきではない。これにはいくつかの理由がある。後日行われる第三者審査用に検査全体を再現することができないため、それが真実か否かに拠らず、利用可能なハードコピーフィルムがイベントを真に正確に反映したものであると仮定せざるを得ない。さらに、例えば、唯一測定可能な病変が腹部の傍大動脈領域にあり、腸内ガスが病変を覆っ

ている場合には、超音波ビームはガスを通過しないため病変は描出されない。そのため、その患者の病期（治療評価の場合は再病期評価）は正確ではなくなる。

6) 理学所見

理学所見による病変評価の再現性も限定的であるが、病変が表在性で 10 mm 以上であり、測径器 (calipers) による測定が可能な場合は許容される。ただし通常は、臨床試験対象患者が、CT で測定可能な他の病変を 1 つ以上有している場合にのみ表在性病変を測定可能病変とすることが望ましい。皮膚病変やその他の触知可能病変についても、理学所見として測定して標的病変とすることができる。

7) MRI

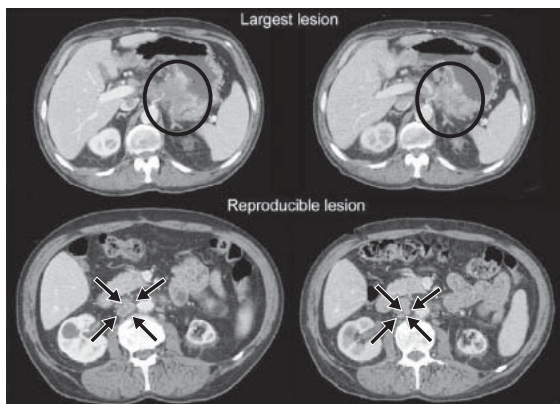
MRI の使用は依然複雑な問題を含んでいる。MRI のコントラスト分解能、空間分解能、時間分解能はきわめて優れているが、MRI の撮影条件には多くの変数があり、それらが画質、病変の描出能や測定値に大きな影響を及ぼす。さらに、MRI の利用可能性は世界各国で異なる。MRI を用いる場合には、CT を用いる場合と同様に、使用する撮影シーケンスの技術的仕様を疾患の種類と部位の評価に合わせて最適化する必要がある。さらに、これも CT の場合と同様であるが、経過中の評価で用いるモダリティはベースライン評価で用いたものと同一にし、同じパルスシーケンスで測定 / 評価する必要がある。通常は、T1 強調画像、T2 強調画像、ガドリニウム造影検査画像による腹部と骨盤の横断画像を撮影する。撮像野 (field of view)、マトリックス、加算回数、位相エンコードステップ、脂肪抑制法や高速シーケンスの使用の有無は、撮影する部位の特性や使用する装置に応じて最適化する必要がある。すべての装置、身体部位、疾患に対応するように、特異的な MRI パルスシーケンスパラメータを規定することは、本稿や付録の範疇を超えるが、理想は、同じ型の装置を使用して、前回の検査で用いた撮影プロトコールに可能な限り厳密に従うことである。さらに、体幹の撮影においては、可能であれば息止め下の撮影を行うべきである。

8) 標的病変の選択

通常は、浸潤臓器を代表する最大の病変（各臓器につき 2 病変まで、計 5 病変まで）を選択し、標的病変として追跡する。しかし、最大の病変の測定が容易ではなく、またその形状が経過観察に適していないことがある。このような場合は、再現性に最も優れた病変のうち最大のものを選択することが推奨される。図 3 に、最大の病変が再現性の最も高い病変ではなく、他の病変を選択し追跡する方がよい実例を示す。

II. 病変の測定

選択した病変の最大径は、画像が得られた断面において測定すべきである。体幹部CTの場合、これは横断面となる。等方性再構成 (isotropic reconstructions) を行う場合には、再構成された画像上で測定を行うことができるが、すべての放射線画像検査施設で等方性再構成が可能なのではないことに注意を要する。そのため、ある評価時期では横断面で測定し、次の評価では異なる断面で測定するといった好ましくない状況が起こる可能性がある。一部の腫瘍、例えば脊柱近傍の病変では、冠状断面や矢状断面で測定した方がよいものもある。これらの病変における冠状断面や矢状断面の再構成画像が等方性的の場合、またはこれらの断面方向で画像がMRIで撮影される場合には、これらの面での病変測定も許容される。評価の際に同じ方向の断面を用いる時、ベースライン時の測定とは異なるスライスレベル、異なる位置関係、異なる方向で病変測定を行うことになったとしても、常にその断面での各標的病変の最長となる径を測定する。腫瘍の辺縁をなぞって最長径を算出するソフトウェアツールを使用してもよく、それによって測定のバラツキが小さくなる



最大の病変が再現性が最も高いとは限らない。そのような場合は、再現性が最も高い病変を標的的病変として選択すべきである。この例では、胃の原発巣 (上段: ベースライン (左) と経過時 (右) の2画像で丸で囲んだ部分) は、薄いスライス厚での3次元再構成CT画像を用いれば、ベースラインと経過時で同程度に胃を拡張させて測定することは可能かもしれない。しかし、これを多施設共同試験で再現するのは困難であり、試みる場合にも慎重な撮影条件の入力と分析が必要となる。最も再現性に優れた病変はリンパ節 (下段: ベースライン (左) と経過時 (右) の2画像で矢印で囲んだ部分) である。

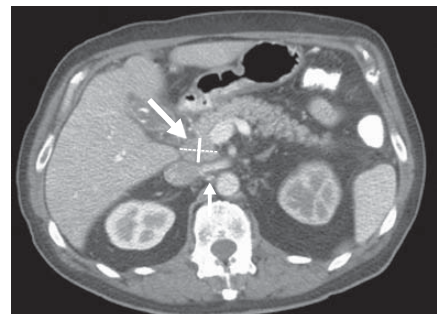
図3 胃癌原発巣とリンパ節の例

ことが期待される。

各断面における最大径 (長径) を用いて、測定可能病変とすること (longest diameter rule) に対する唯一の例外はリンパ節の測定である。リンパ節病変の同定は短径により行われることから、このことは、病的腫大か否かの判定のみならず標的的病変の径和に加える測定値についても適用する指針となる。図4はこの点について説明したものである。大きな矢印はリンパ節病変を示す。最長径と直行する短径は15 mmであり、これが径として記録される。近傍 (小さい矢印) には正常のリンパ節がある。長径は10 mmを超えているが、短径は10 mmを明らかに下回っており、このリンパ節は病的ではないとみなす。

ある病変が消失し、後に再び出現した場合には、測定を継続する。ただし、病変が再出現した時点での効果は、他の病変の状態により異なる。例えば、CRに至った後に病変が再出現した場合は、再出現の時点でPDと判定される。一方、PRまたはSDの場合には、一度消失した1つの病変が再出現した場合、その病変の径が効果を算出するために残りの病変の径和に加えられることになる。すなわち、多くの病変が残存する状態では、1つの病変が見かけ上「消失」した後に再出現したとしても、そのみでPDとは判定せず、全病変の径和がPDの規準を満たした場合にPDと判定するのである。このように判定することとするのは、大半の病変は真に「消失」するわけではなく、使用した画像モダリティの分解能の限界によって描出されないだけであるという認識があるためである。

病変の精密な境界の特定は、特に、肝臓、脾臓、腎臓、副腎、脾臓などコントラストが腫瘍と類似する臓器に存在する病変においては困難な場合がある。さらに、腫瘍周囲の浮腫が病変を取り囲んでいることがあり、一部のモダリティでは、この浮腫と実際の腫瘍との区別が困難なことがある。



大きな矢印はリンパ節病変を示しており、実線で示された短径を測定し追跡する。小さい矢印は、短径が10 mm未満の、病的腫大とはしないリンパ節を示す。

図4 リンパ節評価

実際、浮腫領域内にある腫瘍細胞の存在形態は病理学的にも多彩である。したがって、ベースライン評価からその後の全経過時点において、再現性に優れた方法で測定を行うことが最も重要なのである。これは、同一の画像モダリティを一貫して使用することの強い根拠のひとつでもある。

病変が「断片化 (fragment)」した場合は、各病変の径を加算して標的病変の径和に加える。同様に、病変が融合した場合、病変間の境界面が維持されていれば、病変ごとの最大径の測定を行う。病変が真に融合して境界面の同定が不可能となった場合には、融合した病変の径が最大となる方向の径を「融合した病変」の長径とする。

Ⅲ. 非標的病変の増悪

「明らかな増悪 (unequivocal progression)」と判定するには、標的病変の効果がSDやPRであってもなお治療の変更が必要と担当医が感じる程度の、明らかで著しい非標的病変の増悪がなければならない。明らかな増悪の例を図5および図6に示す。

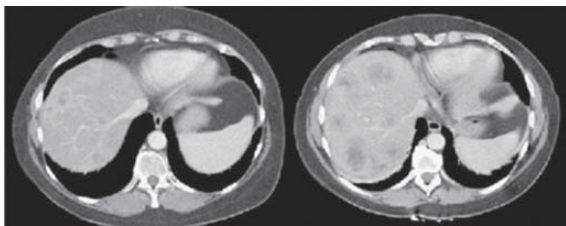


図5 非標的病変における明らかな増悪の例 (肝)

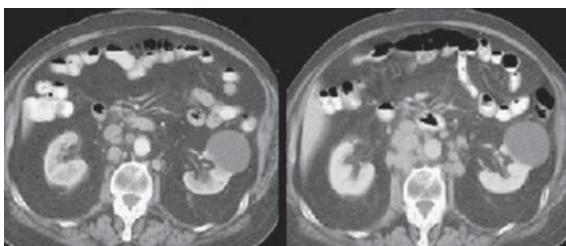


図6 非標的病変における明らかな増悪の例 (リンパ節)

■ 付録Ⅲ よくある質問 (frequently asked questions) ■

質 問	回 答
ベースラインでは独立していたいくつかの病変が経過観察時の評価で融合していた場合、どうすべきか?	融合した塊の最長径を測定し、径和に加算して記録する。
増悪とみなされる新病変は、どれくらいの大きさが必要? 1 cm 未満の小さな病変も対象となるのか、病変は測定可能でなければならないのか?	新病変とするのに「測定可能規準」を満たす必要はない。(同じ方法で行われた検査での) 以前の画像において、病変が存在しなかったことが明らかで、その後に出現した明らかな病変があれば、増悪とみなす。(撮影の方法や条件が原因で) 疑問がある場合には、次回予定の評価まで治療を継続することが推奨される。通常は、その時点ですべてが明確になる。その病変が増大した場合は新病変となるが、増悪日は最初に疑われた日とする(訳注:本文 4.4.4. の訳注参照)。病変が消失した場合には、画像診断医の確認のもとアーティファクトとすることもある。
1つの病変が後の検査で2つに分裂した場合はどのように測定するのか?	それぞれの病変の長径を測定して径和に加算する。
増悪の定義は、全標的病変の状態によるのか、それとも1病変のみによるのか?	RECIST1.1ガイドラインでは、増悪とするには、全標的病変の径和の20%増加かつ径和の絶対値の5 mm以上の増加が必要である。
規制当局はRECIST規準を受容しているのか?	RECIST1.0の準備と採択にあたっては、多くの共同研究グループや製薬企業のメンバーが加わった。作成に際してFDAへの相談がなされ、その回答は「要求はしないが、RECISTの使用を支持する」とのことであった。このほか、欧州とカナダの規制当局も参画しており、現在RECISTは欧州における抗癌剤開発の指針に関する通知に組み込まれている。多くの製薬会社でも使用されている。RECIST1.1も同様に公表前に広く配布された。
CTのスライス厚が5 mmを超える場合の測定可能病変の規準は何か?	RECIST1.1では、CTの最大スライス厚は5mm、測定可能病変の最大のサイズはその2倍、すなわち10mm(スライス厚が5 mm未満の場合も同様)を推奨している。スライス厚が5 mmを超える装置を用いる場合には、病変の最小の長径は実際のスライス厚の2倍とする。
標的病変がきわめて小さくなり、「測定可能な」サイズである10 mm未満となった場合、何を記録すべきか?	標的病変の測定可能性はベースライン時に決定される。以降は、たとえ10 mm未満になっても、実際の測定値を記録する。病変がきわめて小さくなって、画像診断医によっては「小さすぎて測定不能」とすることがある。このような場合(病変が依然存在すると考えられる場合)には、既定値として5 mmを適用することを推奨する。画像診断医が病変は消失していると判断した場合は、既定値として0 mmを記録する。
いくつかの病変が縮小してPR規準を満たし、その時に1つの病変が消失していた場合、後にその「消失」した病変が再出現した時にはPDとするのか?	PR(またはSD)の場合は、径和がPD規準を満たさない限り、1つの病変が再出現してもPDとはしない。その病変は単に径が径和に加算されるだけである。CRであった場合には、消失していた病変が明らかに再出現すればPDとする。
標的病変の長径を測定する際に、病変の形状が変化して、最大径の方向が変わった場合においても、最初に測定した方向で径を測定するのか?	実際の最大径の方向が最初の測定(または経過中の異なる時点)で用いた方向とは異なる場合でも、常にその時点での最大径を測定する。 これに対する唯一の例外はリンパ節である。RECIST1.1では、常に短径を測定することとしており、腫瘍病変*の場合と同様、短径の向きが経過中に変化することもある。 ※訳注:原著では「target lesion」とあるが、明らかに、リンパ節以外の病変をさすはずであることから腫瘍病変とした。

質 問	回 答
<p>ベースラインで標的病変を選択し追跡したが、これらの標的病変のうち1つがその後評価不能となった場合（異なる検査方法を用いたなどによる）、他の標的病変および総合効果にどのような影響が及ぶか？</p>	<p>このような場合、以下のいずれかの処理を行う：</p> <p>(a) まだ治療が継続されている場合、施設に連絡して今後の評価は必ずベースライン検査と同じ方法を用いて行うよう求め、少なくとも一部の経過が完全に評価可能になるようにする。</p> <p>(b) (a) が不可能な場合、経過観察で用いたものと同じ方法の検査がベースラインの検査で使用されていないかを確認する。同じ方法によるベースラインの測定値が得られれば、その病変は評価可能となる。</p> <p>(c) (a) も (b) も不可能な場合、すべての症例報告用紙からその病変を削除する、または、その病変の全体への影響がきわめて重要であるとしてその病変が評価不能となったことにより総合効果の判定も評価不能とする、のいずれかを選択する。このような決定は判定委員会（review panel）で議論されるべきである</p> <p>こうした病変をベースラインの径和には含め、その後の経過の径和には含めないことは、効果に有利なバイアスとなるため推奨されない。</p>
<p>理由が何であれ、1つの非標的病変が中央判定できない場合はどうなるのか？すべての評価を無効とするのか？</p>	<p>他の病変がCRになっている場合に、1つの非標的病変が大きく影響することが時に起こり得る。そのような状況では、1つの非標的病変を評価できないことによりCRの判定が不可能になる。また非標的病変の増悪が著しい時に、標的病変の効果にかかわらずPDとなることもある。ただし、特に残りの測定可能病変がSDやPRの場合に、こうしたことが起きることはきわめて稀である。</p>
<p>径和が2サイクル目で32%、4サイクル目で28%、6サイクル目で33%減少した。このような場合、その後のスキャンによりPRを確定する必要があるのか？、またはこの場合はPRが確定されたことになるのか？</p>	<p>腫瘍縮小割合が30%前後で推移することは稀ではない。このような場合、全体の経過を考慮してPRが確定されたとみなされることが大半である。PRの効果が見られた2つの時点の間で、2～3回の非PRが観察された場合、最も保守的な方法はこれをSDとすることである。</p>
<p>乳癌に対する術前化学療法の試験において、病変の評価にマンモグラフィは使用されないのか？この状況でもCTの方が好ましいのか？</p>	<p>この状況では、CTもマンモグラフィも最適ではない。術前化学療法では、乳房病変を経過観察する上で好ましいモダリティはMRIである。</p>
<p>臨床的評価とCTの両方で測定可能な病変を持つ患者の場合、どちらを選択すべきか？</p>	<p>CTを選択する。診断画像は事後的な中央判定や第三者による確認が可能のため、両方使用できる場合には常に画像診断により評価する。</p>
<p>ベースラインでは充実性だった病変の中心部が壊死になった。どのように測定すべきか？</p>	<p>病変全体の最大径を測定する。治療に反応した壊死病変は最終的には縮小するためである。一部の薬剤（血管新生阻害剤など）ではこのような効果が見られることがあるため、このような現象が頻繁に見られた場合には、試験結果を報告する際に報告した方がよい。</p>
<p>病変の経過観察にMRIを使用する予定の場合、測定可能な最小のサイズはどのくらいか？</p>	<p>部位によっては造影CTをMRIで代替することができるが、肺に対してはできない。撮影のスライス厚が5mmでギャップがない場合、測定可能な最小のサイズはCTと同じ（10mm）である。MRIのスライス厚がこれよりも厚い場合には、ベースラインにおける測定可能病変のサイズは、スライス厚の2倍とする。スライス間のギャップがある場合には、ベースラインで測定可能病変のサイズを決定する際にそれを考慮する必要がある。</p>

質問

回答

RECIST ではPET/CTを使用することはできるか?

現在のところ、PET/CTの低線量CT像または吸収補正用CT像は、RECISTの測定で用いる上でその診断上のCT像の品質が必ずしも最適とは限らない。しかし、PET/CTの一部として実施するCTが、(経静脈性造影剤や経口造影剤を用いた)診断用CTと同等の診断上の品質を有することが施設で実証されれば、RECISTの測定にPET/CTを用いることができる。ただし、日常的に一連の検査としては実施していない場合には、PET/CTのPET像は研究者にバイアスを与えかねない付加的なデータを持ち込む点に注意が必要である。

文献

- Paesmans M, Sculier JP, Libert P, et al. Response to chemotherapy has predictive value for further survival of patients with advanced non-small cell lung cancer: 10 years experience of the European Lung Cancer Working Party. *Eur J Cancer* 1997;33:2326-32.
- Buyse M, Thirion P, Carlson RW, et al. Relation between tumor response to first-line chemotherapy and survival in advanced colorectal cancer: a meta-analysis. *Meta-analysis group in Cancer. Lancet* 2000;356:373-8.
- Goffin J, Baral S, Tu D, et al. Objective responses in patients with malignant melanoma or renal cell cancer in early clinical studies do not predict regulatory approval. *Clin Cancer Res* 2005;15:5928-34.
- El-Maraghi RH, Eisenhauer EA. Review of phase II trial designs used in studies of molecular targeted agents: outcomes and predictors of success in phase III. *J Clin Oncol* 2008;10:1346-54.
- Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981;47:207-14.
- Tonkin K, Tritchler D, Tannock I. Criteria of tumor response used in clinical trials of chemotherapy. *J Clin Oncol* 1985;3:870-5.
- Baar J, Tannock I. Analyzing the same data in two ways: a demonstration model to illustrate the reporting and misreporting of clinical trials. *J Clin Oncol* 1989;7:969-78.
- Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors (RECIST Guidelines). *J Natl Cancer Inst* 2000;92:205-16.
- Therasse P, Eisenhauer EA, Verweij J. RECIST revisited: a review of validation studies on tumour assessment. *Eur J Cancer* 2006;42:1031-9.
- Bogaerts J, Ford R, Sargent D, et al. Individual patient data analysis to assess modifications to the RECIST criteria. *Eur J Cancer* 2009;45:248-60.
- Moskowitz CS, Jia X, Schwartz LH, Gönen M. A simulation study to evaluate the impact of the number of lesions measured on response assessment. *Eur J Cancer* 2009;45:300-10.
- Sargent D, Rubinstein L, Schwartz L, et al. Validation of novel imaging methodologies for use as cancer clinical trials end-points. *Eur J Cancer* 2009;45:290-9.
- Macdonald DR, Cascino TL, Schold Jr SC, Cairncross JG. Response criteria for phase II studies of supratentorial malignant glioma. *J Clin Oncol* 1990;8:1277-80.
- Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol* 2007;10:579-86.
- Schwartz LH, Bogaerts J, Ford R, et al. Evaluation of lymph nodes with RECIST 1.1. *Eur J Cancer* 2009;45:261-7.
- Rustin GJ, Quinn M, Thigpen T, et al. Re: New guidelines to evaluate the response to treatment in solid tumors (ovarian cancer). *J Natl Cancer Inst* 2004;96:487-8.
- Bubley GJ, Carducci M, Dahut W, et al. Eligibility and response guidelines for phase II clinical trials in androgen-independent prostate cancer: recommendations from the Prostate-Specific Antigen Working Group. *J Clin Oncol* 1999;17:3461-7.
- Scher H, Halabi S, Tannock I, et al. Design and end points of clinical trials for patients with progressive prostate cancer and castrate levels of testosterone: recommendations of the Prostate Cancer Clinical Trials Working Group. *J Clin Oncol* 2008;26:1148-59.
- Vergote I, Rustin GJ, Eisenhauer EA, et al. Re: new guidelines to evaluate the response to treatment in solid tumors [ovarian cancer]. *Gynecologic Cancer Intergroup. J Natl Cancer Inst* 2000;92:1534-5.
- Van Glabbeke M, Verweij J, Judson I, Nielsen OS. EORTC Soft Tissue and Bone Sarcoma Group: Progression-free rate as the principal end-point for phase II trials in soft-tissue sarcomas. *Eur J Cancer* 2002;38:543-9.
- Dancey JE, Dodd LE, Ford R, et al. Recommendations for the assessment of progression in randomised cancer treatment trials. *Eur J Cancer* 2009;45:281-9.
- Ford R, Schwartz L, Dancey J, et al. Lessons learned from independent central review. *Eur J Cancer* 2009;45: 268-74.
- Catalano C, Francone M, Ascarelli A, Mangia M, Iacucci I, Passariello R. Optimizing radiation dose and image quality. *Eur Radiol* 2007;17(Suppl 6):F26-32.
- Low RN. Abdominal MRI advances in the detection of liver tumours and characterization. *Lancet Oncol* 2007;8(6):525-35.
- Barrett T, Choyke PL, Kobayashi H. Imaging of the lymphatic system: new horizons. *Contrast Media Mol Imaging* 2006;1(6):230-45.
- Shankar LK, Hoffman JM, Bacharach S, et al. National Cancer Institute. Consensus recommendations for the use of 18F-FDG PET as an indicator of therapeutic response in patients in National Cancer Institute Trials. *J Nucl Med* 2006;47(6):1059-66.