

2015/10/10

JCOG臨床試験セミナー中級編

RECISTと 免疫療法の効果判定規準

国立がん研究センター中央病院 臨床研究支援室/
研究支援センター 多施設研究支援室/
JCOG運営事務局/
中村 健一

本講義の背景と目的

■ 背景

- 免疫療法の出現によりRECIST以外の効果判定規準、検定法の適用が提唱されている

■ 目的

- RECISTの本来の目的と正しい使い方を理解する
- 免疫療法で提唱されているirRC、irRECIST、Fleming-Harrington検定について理解し、適切な適用法を学ぶ

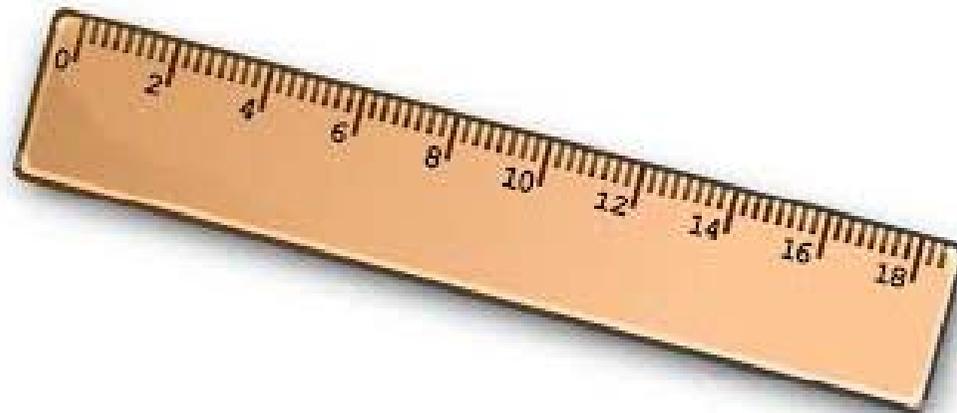
アウトライン

- RECIST
 - RECISTのよくある誤用
 - 効果判定規準と治療中止規準

- 免疫療法の効果判定規準
 - irRC (immune-related Response Criteria)
 - irRECIST

- 免疫療法の生存時間解析
 - ログランク検定の復習
 - Fleming-Harrington検定

RECISTのエッセンスとよくある誤用



RECISTあるある(誤用の例)

- PDになったら必ず治療中止と思っている
- 測定可能病変と標的病変の区別ができていない
- 標的病変の効果＝総合効果と思っている
- 総合効果と最良総合効果の違いがわからない
- プロトコールには「効果判定はRECIST v1.1に従う」とだけ書けばよいと思っている



臨床試験は「比較」である



- 個々の患者に対する治療
 - 症状が改善している間は「有効」
 - 患者がつらいと感じれば「有意な毒性」
 - 個々の患者さんの症状、病態に応じて治療を行えばよい

- 二群間での比較
 - 群間で比較のモノサシが共通であることが必要
 - 試験独自のモノサシを用いても構わない

- 他グループの試験、過去の試験との比較
 - 異なる国、グループ、施設、評価者が比較
 - 同じものさし、同じ測り方が必要

RECIST はなんのための“ものさし”か？

■ 「第III相」へのgo/no-goの指標

- 「薬剤あるいはレジメンが開発研究を続けるに値する有望な結果を示すかどうかの判断に用いられる」
- 「まさにこの状況で用いるために本ガイドラインが作成された」
- 単アームの第II相試験の結果から第III相試験に進めるかどうかの判断

「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)ー日本語訳JCOG版ー」序論より

■ 「治療」のgo/no-goを決める指標ではない

- RECISTは異なる試験どうしの「比較可能性」を優先
- 日常診療でも臨床試験でも治療の継続/中止は臨床的/総合的判断で決めるべきで、RECISTで決める必要はない(むしろ不適切)

RECISTは試験間で精密に
奏効割合やPFSを比較するための規準

PD＝治療中止ではない

- PDで治療中止は必ずしも正しくない
 - SDとPDとの線引きは比較のためのもの
 - これまで、PDだったら治療中止しても臨床的に妥当ながん種がたまたま多かった

■ プロトコルマニュアルの記載

6.2.2. プロトコル治療中止の規準

例)

以下のいずれかの場合、プロトコル治療を中止する。

1) 以下のいずれかによりプロトコル治療無効と判断

① 治療開始後に原病の増悪が認められた場合

※画像による効果判定で PD と判定されても臨床的にプロトコル治療継続が妥当と判断される場合にはプロトコル治療中止とはせず、プロトコル治療を継続する。

② Oコース終了時点までに腫瘍の縮小や症状の改善がみられない場合

注) 治療継続の可否を決める「治療無効」かどうかの判断は総合的な臨床判断で行う。画像による総合効果(CR/PR/SD/PD)はあくまでも参考とする。実際には腫瘍が縮小していても総合効果は PD となり得るし、腫瘍が増大していても総合効果は PR となり得る。総合効果が PR の時に臨床的には無効と判断してプロトコル治療中止とすることが妥当である場合もあるし、総合効果が PD であっても治療が有効と判断して治療継続が妥当である場合もある。

「PD中止」ではなく「無効中止」



PD＝治療中止ではない

■ JCOG0805 (肝胆膵グループ)

6.2.2. プロトコール治療中止規準

以下のいずれかの場合、プロトコール治療を中止する。

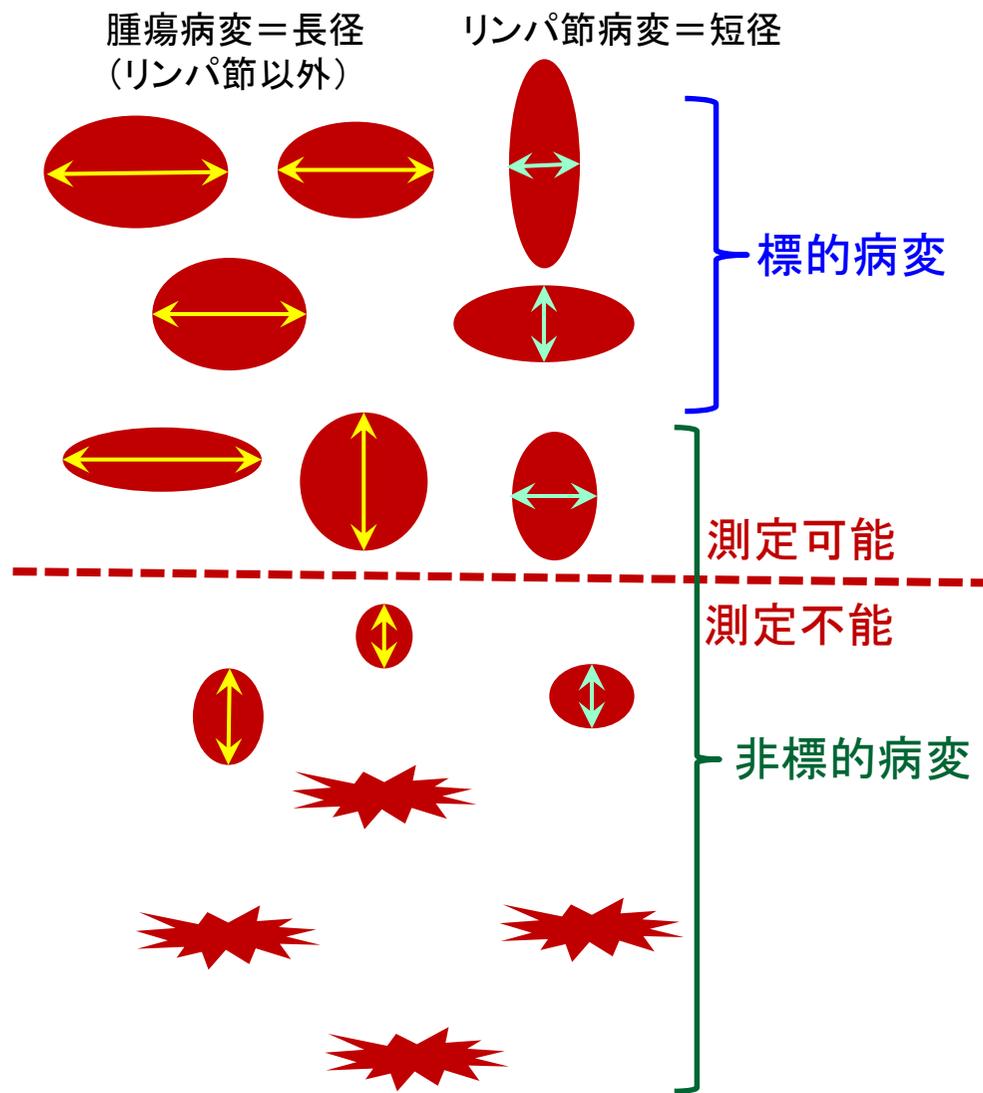
1) プロトコール治療無効と判断

- ・ 治療開始後に原病の増悪が認められた場合。原病の増悪によるプロトコール治療無効の判断は、下記の①～③を参考として、総合的な臨床判断により行う。

- ① 画像での計測上、長径(和)の増大を認め、効果判定上はPDと判定したとしても、3次元的に推定される腫瘍量の増大がない場合や、症状の改善傾向が続いている場合等には増悪とはせず、プロトコール治療を継続する。
- ② 同様に、画像にて、微小肝転移など測定不能な新病変(最大径 1 cm 未満 を目安とする)の出現を疑う所見が見られた場合にも、良性変化(ベースライン測定の際にスライスに含まれなかった微小嚢胞、反応性リンパ節腫大、炎症性変化など)を除外しきれない場合は、全身の腫瘍量や症状の改善も参考として増悪の有無は判断する。ただし、測定不能な新病変であっても画像的に良性変化を除外できる場合や、測定可能な新病変(最大径1 cm 以上を目安とする)が出現した場合は(PDかつ)増悪としてプロトコール治療を中止する。
- ③ 逆に、画像上は明らかな腫瘍の増大を認めなくても、悪液質の進行などにより臨床的に明らかな増悪と判断される場合はプロトコール治療を中止する。ただし、薬物有害反応と鑑別するために画像所見、腫瘍マーカーの推移なども参考とする。

効果判定と治療中止規準は別個に考える！

RECISTの基本的ロジック



- ① すべての病変を「測定可能」か「測定不能」かに分類
 - 腫瘍病変 (リンパ節以外) 長径10mm以上
 - リンパ節病変 短径15mm以上

- ② 測定可能病変の中から「標的病変」を選択
 - 全部で5個まで (1臓器2個まで)
 - **標的病変** 以外はすべて **非標的病変**

◆ 測定可能病変と標的病変は別もの！

◆ 「評価可能病変」という用語はRECISTにはない

③各カテゴリーで効果を判定

必ずしも径和はゼロにならない！

- 標的病変の効果 ~ 径和の縮小/増大率
 - CR: すべての腫瘍病変が消失(リンパ節病変は10mm未満)
 - PR: 径和がベースラインに比べ30%以上減少
 - SD: CR/PRでもなく、PDでもない
 - PD: 径和が経過中の最小値に比べ20%以上増大
- 非標的病変の効果 ~ 消えたか増大したか
 - CR: すべての腫瘍病変が消失(リンパ節病変は10mm未満)
and 腫瘍マーカーが基準値上限以下
 - Non-CR/Non-PD: ひとつ以上の非標的病変の残存 or 腫瘍マーカーが基準値上限を越える
 - PD: 非標的病変の明らかな増大(Unequivocal progression)
- 新病変の有無 ~ 「あり」か「なし」
 - 「あり」か「なし」

④ カテゴリーを組み合わせ総合効果の判定

- 総合効果 Overall Response
 - 効果判定毎に判定

標的病変	非標的病変	新病変	総合効果
CR	CR	No	CR
CR	Non-CR/Non-PD	No	PR
PR	Non-PD	No	PR
SD	Non-PD	No	SD
PD	Any	Any	PD
Any	PD	Any	PD
Any	Any	Yes	PD

標的病変の効果と総合効果は必ずしも一致しない！

⑤ 最良総合効果の判定

■ 効果の確定 confirmation

□ Confirmationあり

- 効果の確定のためには・・・
 - 2回連続でCRであることが必要(最良総合効果CR)
 - 2回連続でPR(or CR)であることが必要(最良総合効果PR)
- 奏効率がprimary endpointである単群試験では必須

総合効果は毎回の判定ごと
最良総合効果は全体で1つ

総合効果	治療開始	SD	PR	PR	CR	PD
------	------	----	----	----	----	----

→ 最良総合効果 **PR**

□ Confirmationなし

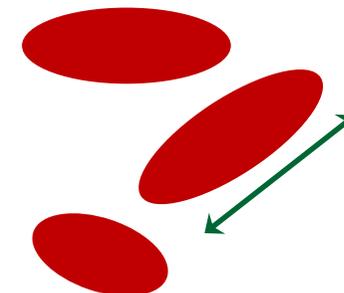
- 全コースを通じて最もよい「総合効果」=最良総合効果
- ランダム化比較試験やneoadjuvant studyでは不要

総合効果	治療開始	SD	PR	PR	CR	PD
------	------	----	----	----	----	----

→ 最良総合効果 **CR**

復習：効果判定の手順

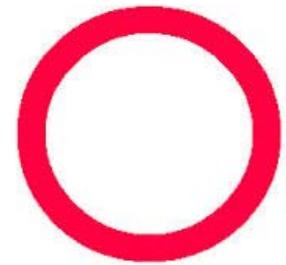
1. すべての病変について測定可能かどうかを判断
 - 腫瘍病変 長径10mm以上
 - リンパ節病変 短径15mm以上
2. 測定可能病変の中から標的病変を選択
 - 5個まで(1臓器につき2個まで)
 - 標的病変以外はすべて非標的病変
3. 各カテゴリーでそれぞれ効果を判定
 - 標的病変 CR, PR, SD, PD
 - 非標的病変 CR, Non-CR/Non-PD, PD
 - 新病変 あり, なし
4. カテゴリー毎の効果の組み合わせで「総合効果」を判定
 - 総合効果はコース毎
5. 「最良総合効果」の判定
 - 最良総合効果は全体で1つ



プロトコールには「RECISTに従う」でOK？

- ~~プロトコールに「効果判定はRECIST v1.1に従う」とだけ書いてあるものが散見される~~
- RECISTはあくまでガイドライン。詳細はプロトコールで定める必要あり
 - 用いる画像モダリティ
 - 効果判定のタイミング
 - 測定可能病変の定義
 - 非標的病変の判定に腫瘍マーカーを入れるか？
 - Confirmationの有無

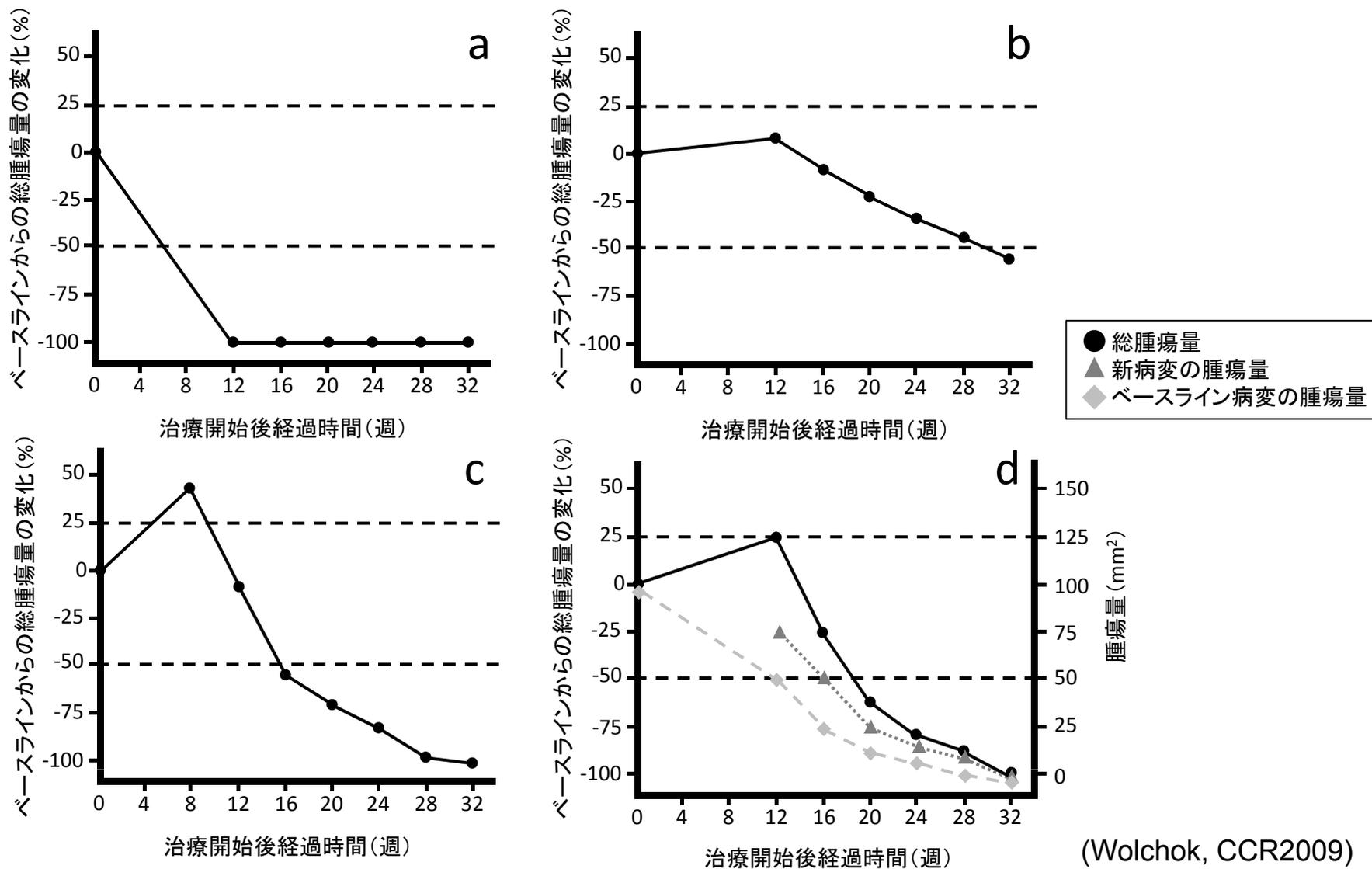
「RECISTあるある」の正解



- PDになったら必ず治療中止と思っている
 - 効果判定と治療中止規準は別個に考える
- 測定可能病変と標的病変の区別ができていない
 - 測定可能病変の中から標的病変を選択
- 標的病変の効果＝総合効果と思っている
 - 標的病変、非標的病変、新病変の3つから総合効果を算定
- 総合効果と最良総合効果の違いがわからない
 - 総合効果は毎回の判定ごと
 - 最良総合効果は全ての判定を通じて1つ
- プロトコールには「効果判定はRECIST v1.1に従う」と書けばよいと思っている
 - 用いる画像モダリティ、判定のタイミング、測定可能病変の定義、confirmationの有無などは試験毎に定義する必要あり

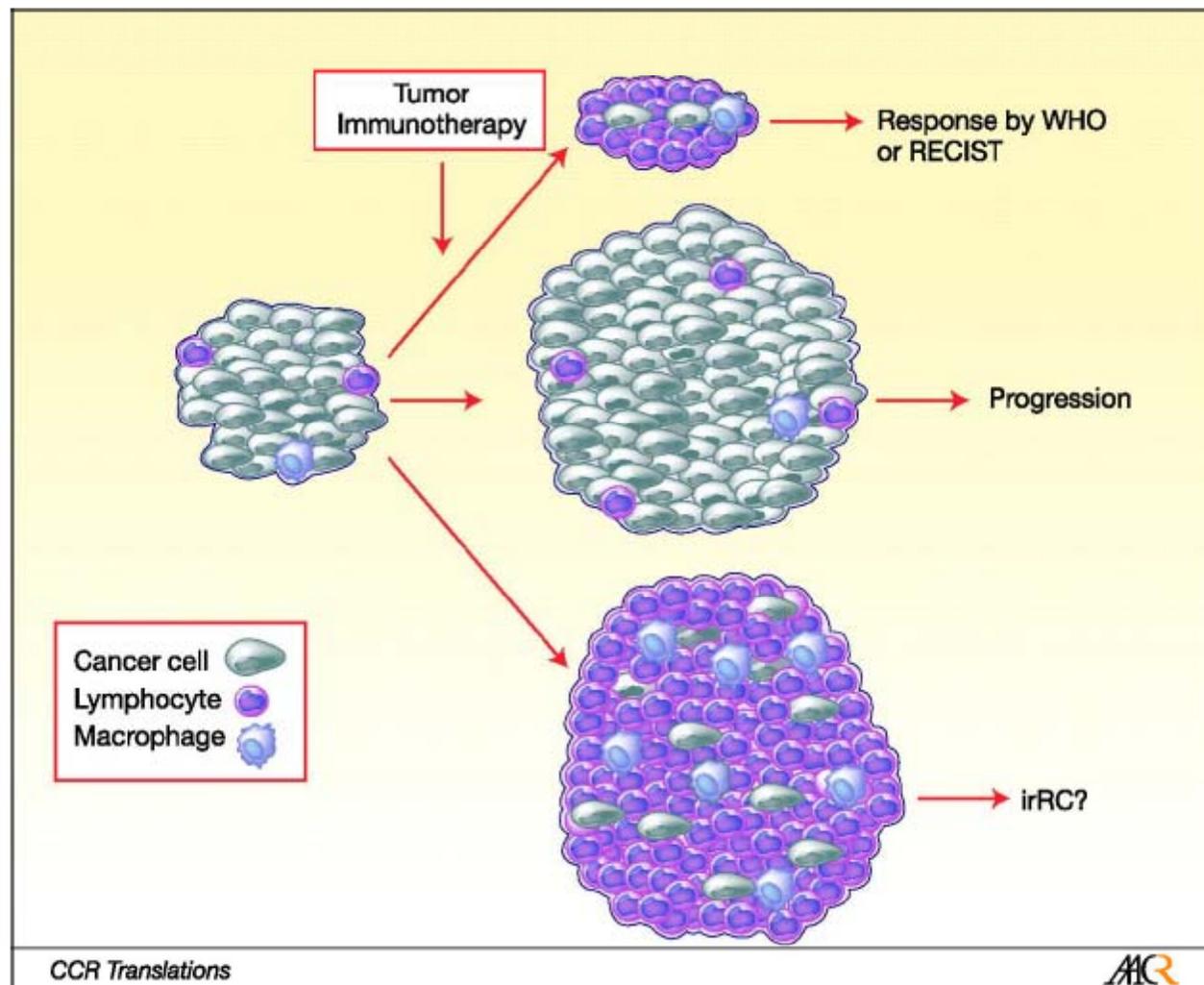
免疫療法の効果判定 — “flare effect” への対処

“Flare effect” in immunotherapy



(Wolchok, CCR2009)

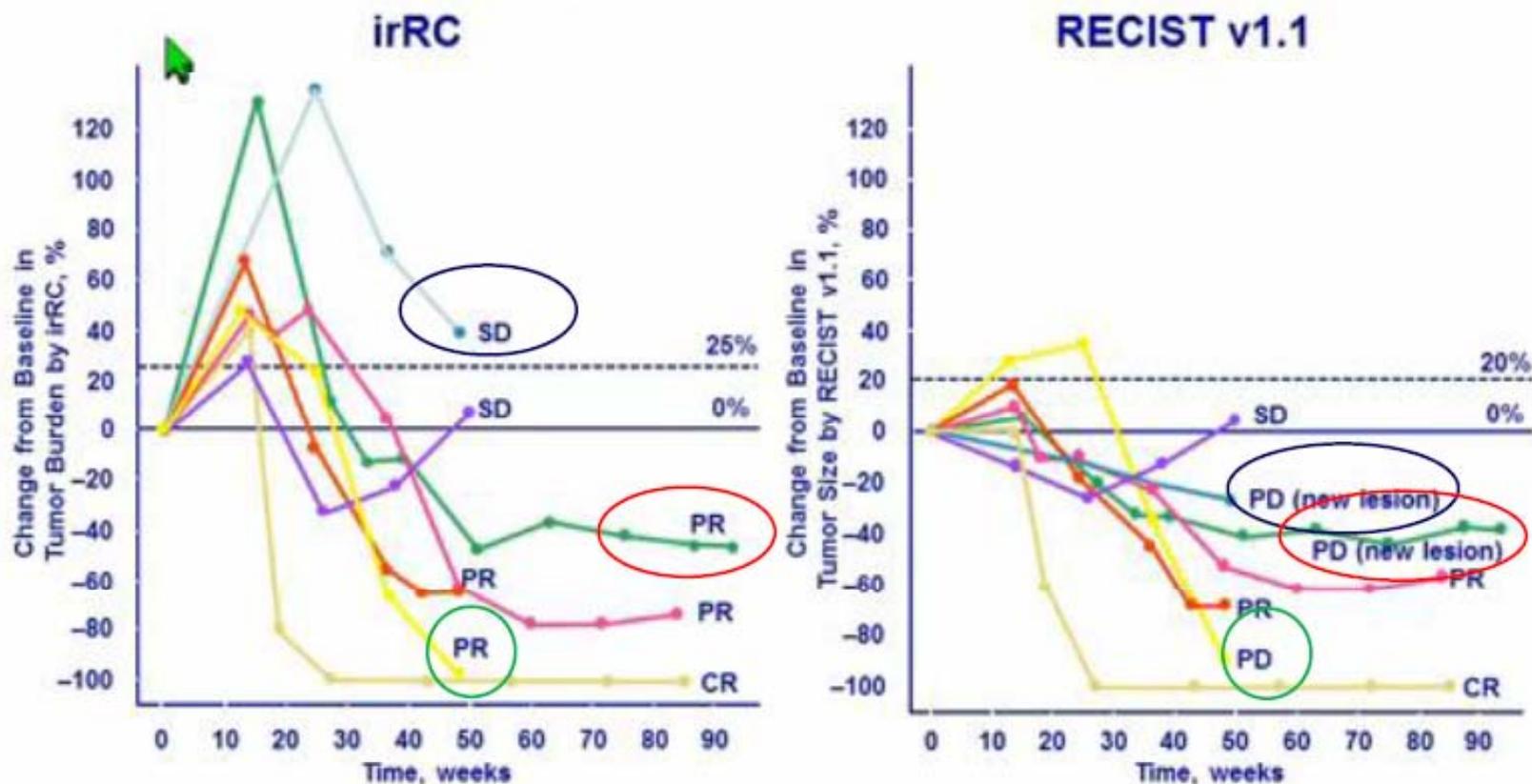
Tumor lesions may become heavily infiltrated by immune and inflammatory cells



Ribas A et al. Clin Cancer Res 2009;15:7116-7118

(横田知哉先生のスライドより)

Early pseudoprogression observed with pembrolizumab



7 of 192 patients (3.6%) showed $\geq 25\%$ increase of tumor burden at week 12 that was not confirmed as irRC PD at the next assessment

Hodi et al, ASCO 2014, abstract 3006



(横田知哉先生のスライドより)

免疫療法の専門家の主張

- PD後にresponseが生じ得る (flare effect)
 - PD後にも治療継続が適切な場合あり
 - 臨床的に明らかに増悪でなければ治療継続すべき
- 長期間持続するSDは、抗腫瘍効果を意味する
- Cytotoxic drugに比べ治療効果の出現が遅い (delayed effect)

免疫療法にはRECISTは使えない(?)
irRC (immune-related Response Criteria) を提案

irRCの実際: ベースライン

■ 測定可能病変

- 径5mm x 5 mm以上(2方向)
- 各臓器5個まで、合計10個まで

RECIST

腫瘍(非リンパ節)病変 長径10mm以上
リンパ節病変は短径15mm以上

RECIST

各臓器2個まで、合計5個まで

■ 総腫瘍量(TMTB, total measured tumor burden)を計算

- 測定可能病変の直交する2方向積和(SPD: sum of the product of the two diameters)
- 測定不能病変はTMTBに含めない

RECIST

1方向の径和

irRCの実際：各評価時点で総腫瘍量を計算

「標的病変」は定義せず

■ 総腫瘍量 = SPD (測定可能病変) + SPD (新病変)

□ 「新病変 = PD」ではない

■ 5mm x 5mm以上の新病変は総腫瘍量の計算に組み入れ

□ 新病変として組み入れるのは合計10個まで(各臓器5個まで)

■ 5mm x 5mm未満の新病変は無視

□ irCRの判定にのみ用いる

RECIST

標的病変、非標的病変、新病変を
独立に評価

■ 総腫瘍量の変化でirCR, irPR, irSD, irPDを判定

□ 測定不能病変はirCRの判定にしか用いない

irRC: 総合効果の判定

- 総合効果 Overall Response
 - コース毎に判定

総腫瘍量 TMTB ベースラインからの変化	測定不能病変 (ベースラインで存在)	測定不能の 新病変	総合効果
-100%	消失	なし	irCR
-100%	あり/増大	Any	irPR
-50-99%	Any	Any	irPR
-50%~+24%	Any	Any	irSD
+25%以上	Any	Any	irPD

- 最良総合効果
 - irCR, irPR, irPDの判定にはconfirmationを要する
 - PDで治療を早期に中止させないための配慮
 - 4週間以上の間隔で同じ総合効果が継続して「確定」

irRC:まとめ

- 新病変の出現はirPDとはせず、総腫瘍量に含める
- 総腫瘍量の変化で総合効果を判定
- 総腫瘍量が25%以上増加していなければirPDではない
- irPDの判定にはconfirmationを要する

できるだけPDと判定させないために
涙ぐましい努力が見られる

irRCの問題点

- 従来のcytotoxic drugとの比較可能性なし
 - RECISTの奏効割合とirRCの奏効割合/PFSは比較できない
 - 従来の薬剤ではirPDのconfirmationは受け入れ難い
- 効果判定規準と治療中止規準を混同
 - 早期PD (flare effect) でも治療継続が妥当な場合がある、、、であれば治療中止規準を工夫して、効果判定はRECISTを使えばよい
- なぜ2方向積和？
 - irRCは、皮膚腫瘍領域の研究者が提唱
 - 一方向の径和を用いるirRECISTが新たに提唱

効果判定規準としてはirRCは推奨されない！



効果判定と治療中止規準を切り分けた例①

ORIGINAL ARTICLE

Nivolumab and Ipilimumab versus Ipilimumab in Untreated Melanoma

Michael A. Postow, M.D., Jason Chesney, M.D., Ph.D., Anna C. Pavlick, D.O., (N Engl J Med 372;21, 2015)

■ Primary endpoint

- BRAF V600 wild-typeに対する奏効割合 (RECIST v1.1)

■ 治療中止規準

- Treatment was continued as long as clinical benefit was observed or unacceptable side effects occurred.
 - Clinical benefitが見られている間は治療を続けてよい
- Patients who had investigator-assessed disease progression could be treated beyond progression...
 - 研究者が増悪と判断したとしても増悪後に治療継続してもよい

効果判定と治療中止規準を切り分けた例②

ORIGINAL ARTICLE

Pembrolizumab versus Ipilimumab in Advanced Melanoma

KEYNOTE-006
(N Engl J Med 372;26, 2015)

Caroline Robert, M.D., Ph.D., Jacob Schachter, M.D., Georgina V. Long, M.D., Ph.D.,

- Primary endpoint
 - PFS (progressionはRECIST v1.1による増悪)
 - SecondaryとしてRECIST v1.1での奏効割合(中央判定)、irRCによる奏効割合(施設判定)を収集
- 治療中止規準
 - After initial evidence of radiologic progression, patients whose condition was clinically stable could continue to receive study treatment until imaging that was performed approximately 4 weeks later confirmed progression.
 - 初回に画像的増悪と判断しても臨床的に安定していれば次回4週後の画像判定時まで治療継続してよい

効果判定と治療中止規準を切り分けた例②

ORIGINAL ARTICLE

Pembrolizumab versus Ipilimumab in Advanced Melanoma

Caroline Robert, M.D., Ph.D., Jacob Schachter, M.D., Georgina V. Long, M.D., Ph.D.,

- プロトコールに治療継続の詳細あり
 - irRCは“治療中止規準”として用いる
 - 測定可能病変の定義はRECIST v1.1
 - irPDでもconfirmationまで治療継続
 - ただし以下の場合には治療中止すべし
 - 検査値の悪化など臨床的増悪を示唆する所見あり
 - ECOG PSが悪化
 - 急速な腫瘍の増大や症状の悪化で別の治療を要する

2.4.2 Radiological Tumor Assessment

The Immune Related Response Criteria (irRC) will be used for assessment of tumor response for the purposes of managing patients on protocol treatment and decision making for discontinuation of study therapy due to disease progression. These disease assessments will be performed by the investigator with site radiology reading. One exception is that for baseline tumor assessments, at least one measurable lesion according to RECIST 1.1 criteria must be present on a bi-dimensional imaging study (CT or MRI) at baseline. A measurable lesion is defined as measuring at least 10 mm in longest diameter or twice the slice thickness whichever is greater; the exception being that lymph nodes must measure 15 mm in short axis for lymph nodes. Copies of tumor images will be collected for independent central review using RECIST 1.1. Refer to Section 2.8.1. Efficacy Measurements, for additional information regarding radiological assessments.

The irRC and RECIST 1.1 response assessments are specifically described in the IOM. For the purposes of the efficacy endpoints of the study, response assessment based on RECIST 1.1 by independent central review will be applied as the primary measure and irRC by investigators will be evaluated as a secondary assessment of the efficacy endpoints.

If imaging shows a complete response (CR) or partial response (PR), tumor imaging should be repeated at least 4 weeks later to confirm response, per RECIST 1.1 and irRC recommendations. Patients will then return to regular scheduled imaging every 6 weeks, starting with the next scheduled imaging time point. Patients who obtain a confirmation scan do not need to undergo scheduled imaging assessment 2 weeks later (e.g. If a patient obtains a scan at Week 16 to confirm a Week 12 response, they will not also be required to complete the scheduled Week 18 scan).

If imaging shows PD, it is at the discretion of the investigator to keep a patient on study treatment or to stop study treatment until imaging is repeated approximately 4 weeks later in order to confirm PD, as described in the irRC recommendations. If there is an increase in tumor burden at week 12, according to irRC guidelines, the baseline sum of perpendicular diameters is to be reset to week 12, for subsequent assessment of response if all of the following conditions are met (see IOM for further details): 1) The time point is within the first regularly scheduled time point (+/-7 days); 2) The time point has the largest SPD from baseline; 3) The time point has no index lesions that have an assessment of NE and patients remain on treatment. [75] Patients that are deemed clinically unstable or who have biopsy proven new metastatic lesions are not required to have repeat imaging for confirmation. This decision will be based on clinical judgment of a patient's overall clinical condition, including performance status, clinical symptoms, and laboratory data. At a minimum, patients must meet the following criteria for continued treatment on study after disease progression is identified at a tumor assessment:

- Absence of symptoms and signs (including worsening of laboratory values) indicating disease progression
- No decline in ECOG performance status
- Absence of rapid progression of disease or of progressive tumor at critical anatomical sites (e.g., cord compression) requiring urgent alternative medical intervention.

If repeat imaging shows an objective response or stable disease relative to baseline, or reset baseline if applicable treatment with MK-3475 will continue/resume and the next studies will be conducted every 6 weeks as previously scheduled. If repeat imaging confirms PD, patients will be discontinued from study therapy.

In patients who discontinue study therapy early without documented disease progression, every effort should be made to continue monitoring their disease status by radiologic imaging following the guidelines described in Section 3.2.5.21 (Duration of Follow-up).

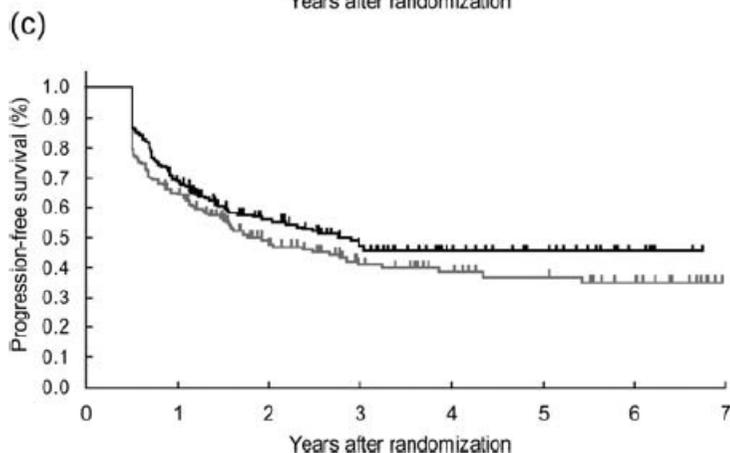
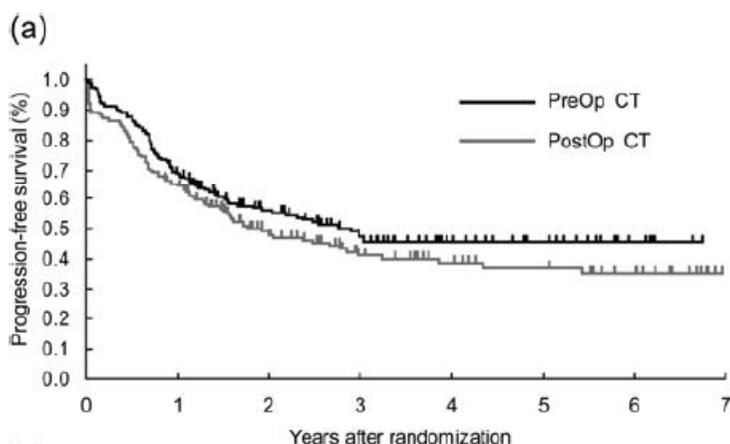
irRECIST

**NEW**

- irRCとRECISTのハイブリッド
 - 測定可能病変、測定不能病変のRECISTの定義を採用
 - 標的病変の選択規準もRECIST準拠
 - 1臓器2個まで、全部で5個まで
 - 1方向の径和を採用
 - 総腫瘍量＝径和(標的病変)＋径和(新病変)
 - irPDのconfirmationはflare effectが疑われる場合のみ採用

- どこまで広まるかは不明
 - 1方向にしてRECISTに出来るだけ合わせたことは評価できる
 - 従来のcytotoxic drugとの比較性はない
 - 免疫療法どうしの比較ではsecondary endpointとして使用？

もうひとつのflare effectの解決法: Landmark method



- 効果判定はRECISTで評価
 - 治療中止規準はirRCの考えを導入
- PFSの評価法
 - 登録後の一定期間=landmark timeの間は何が起きてもイベントとしない
 - Landmark time時点で以下があればPFSのイベント
 - 径和の20%以上の増大
 - 出現した新病変が残存
 - Landmark timeまでに死亡

- ◆ 早期のPDではイベントにならない
- ◆ 従来の薬剤であっても適用可能

(Nakamura, JJCO 2012)

Short summary

- 免疫療法では早期PD後に効果が見られることがある（flare effect）
- flare effectを考慮して治療を継続してもよいが、効果判定規準は比較可能性を重視してRECISTを使うべき
- irRCやirRECISTは治療継続/中止規準として使用可
- 登録後一定期間の増悪はイベントとしないLandmark methodの適用も考慮

免疫療法の生存時間解析 — “delayed effect” への対処

Delayed effectを考慮した解析

日本免疫治療学研究会

がん免疫細胞療法用臨床試験ガイダンス

—FDAガイダンスを踏まえた日本免疫治療学研究会の考え方—

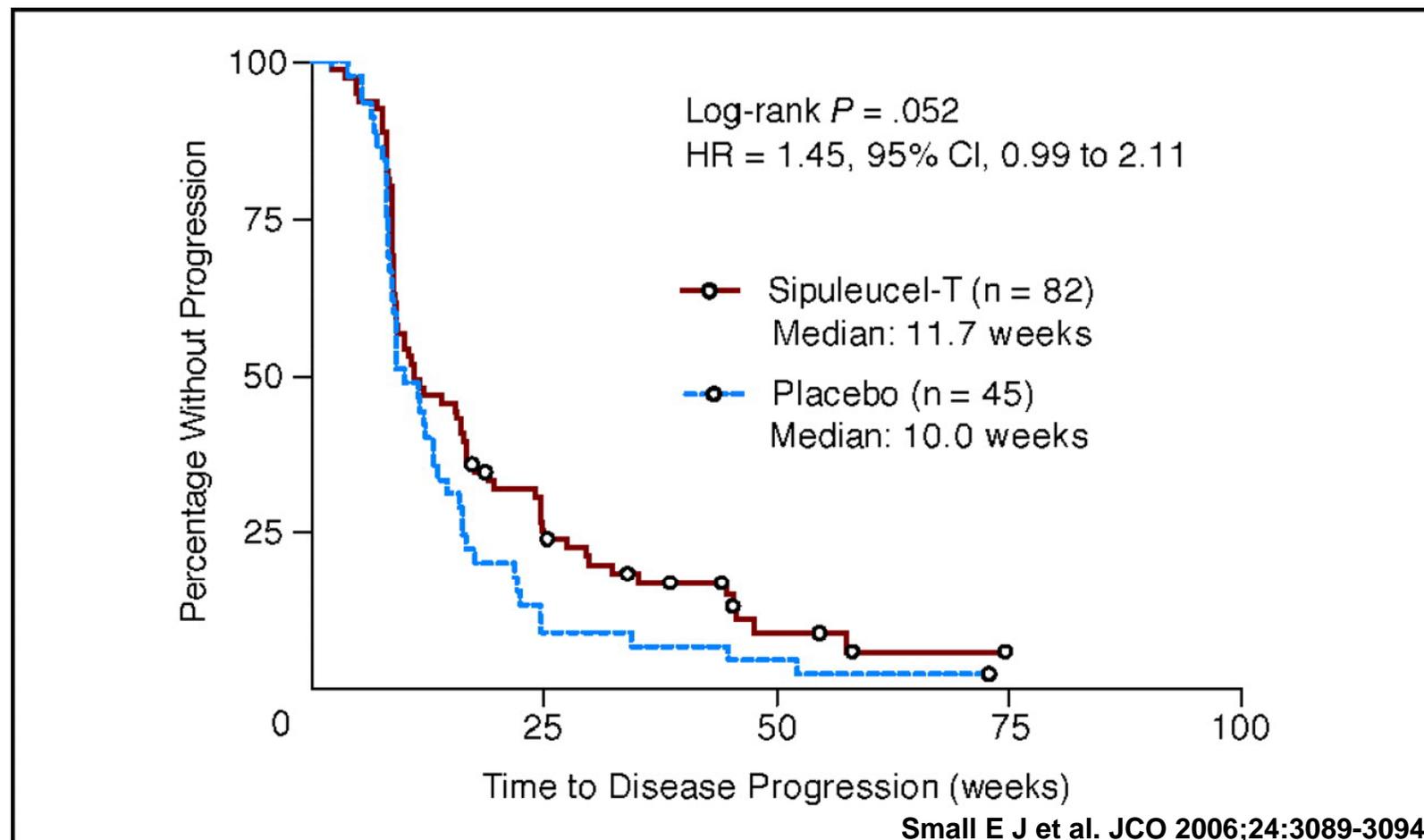
5-4-5. 有効性の評価

第Ⅲ相試験において有効性は主要評価項目を中心に評価される。この解析においては、生存期間中央値や生存割合をKaplan-Meier法によって推定し、また、群間比較にはLog-rank検定が用いられることが一般的である。

しかし、がん免疫細胞療法においては、治療施行から免疫反応が誘導されるまでに一定の時間を要し、遅発性の効果発現が観察される場合がある。このようながん免疫療法に特有の免疫学的作用機序が、科学的合理性をもって事前に想定される場合は、試験群と対照群のハザード比が時間を通じて一定である場合（比例ハザード性）に検出力がもっとも高くなるように計画されているLog-rank検定の代わりに、Harrington-Fleming法で提唱されているような、観察期間の後期に重みを付けた統計解析の必要性も想定されるところである。ただし、Harrington-Fleming検定は、Log-rank検定と異なり、パラメータにより特徴づけられるパラメータ族から構成されるので、事前に、合理的なパラメータ設定の下に適切な検定を計画しなければならない。

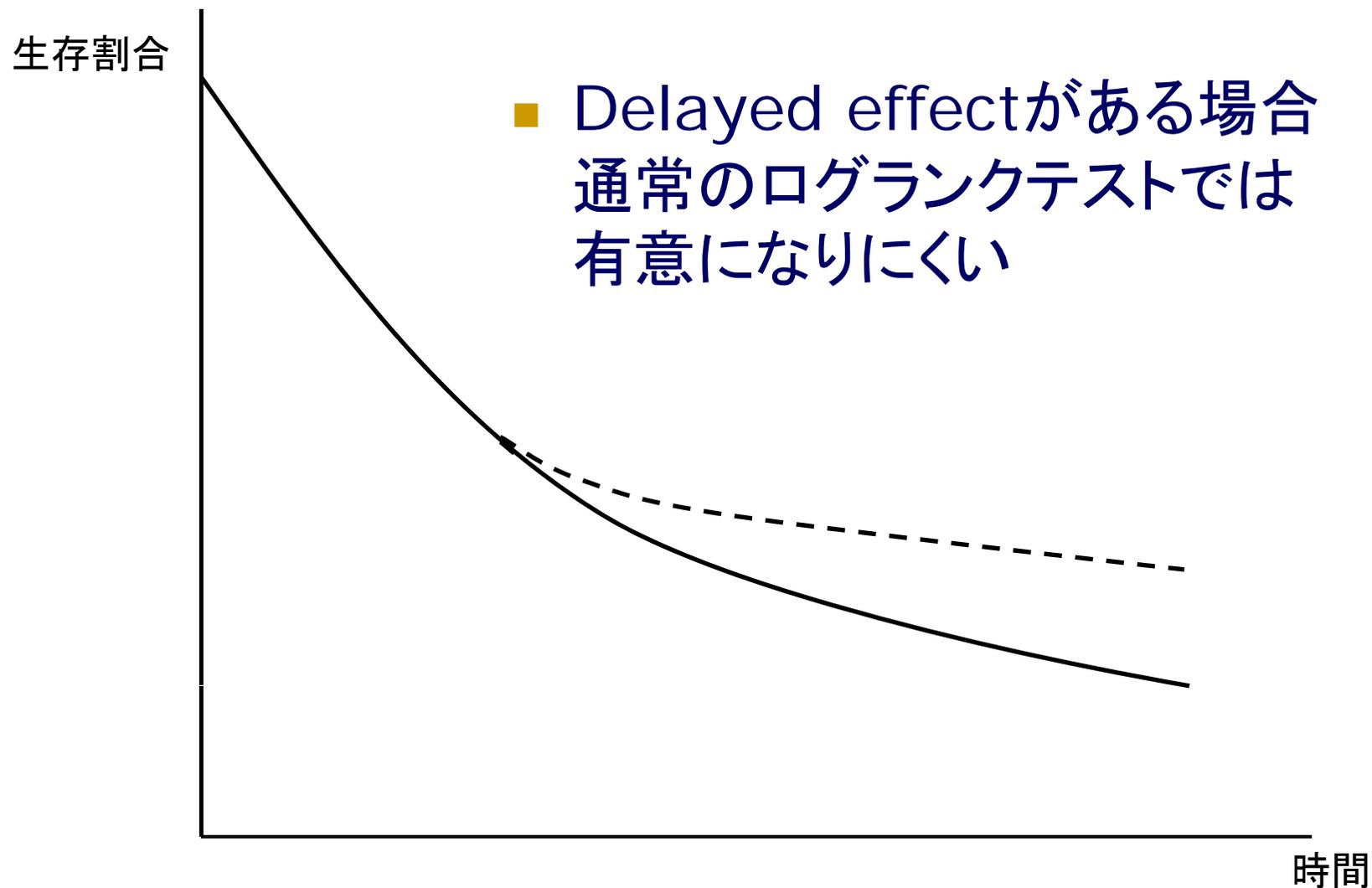
免疫療法における“delayed effect”

- Sipuleucel T (Provenge)のプラセボ対照試験
 - 前立腺癌に対するがんワクチン

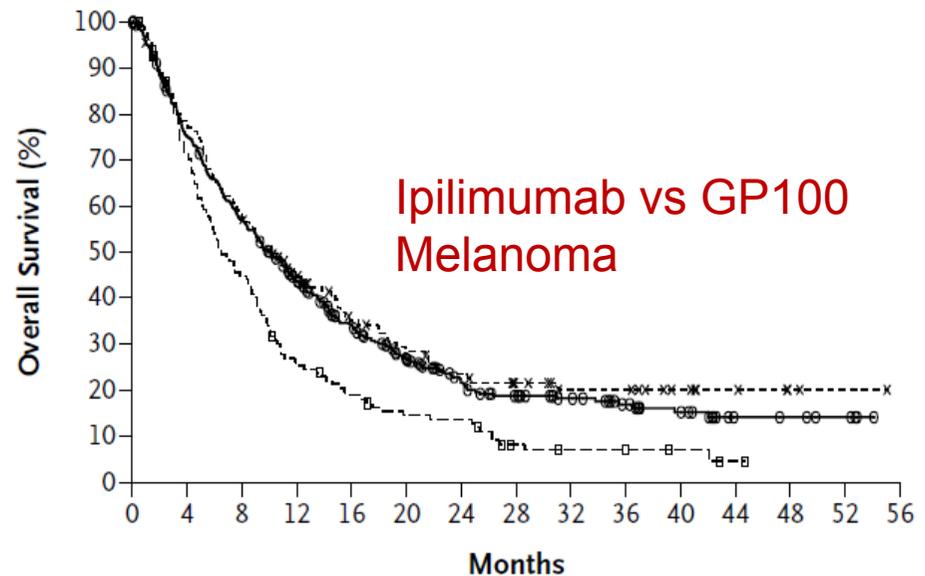
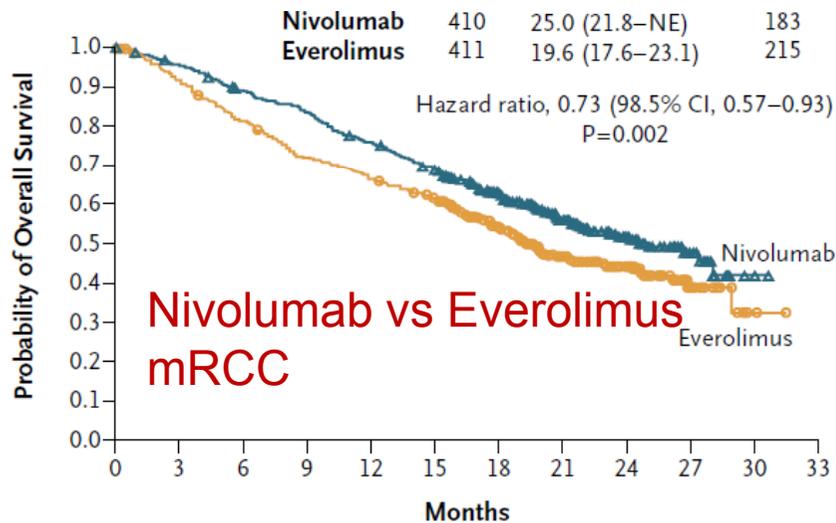
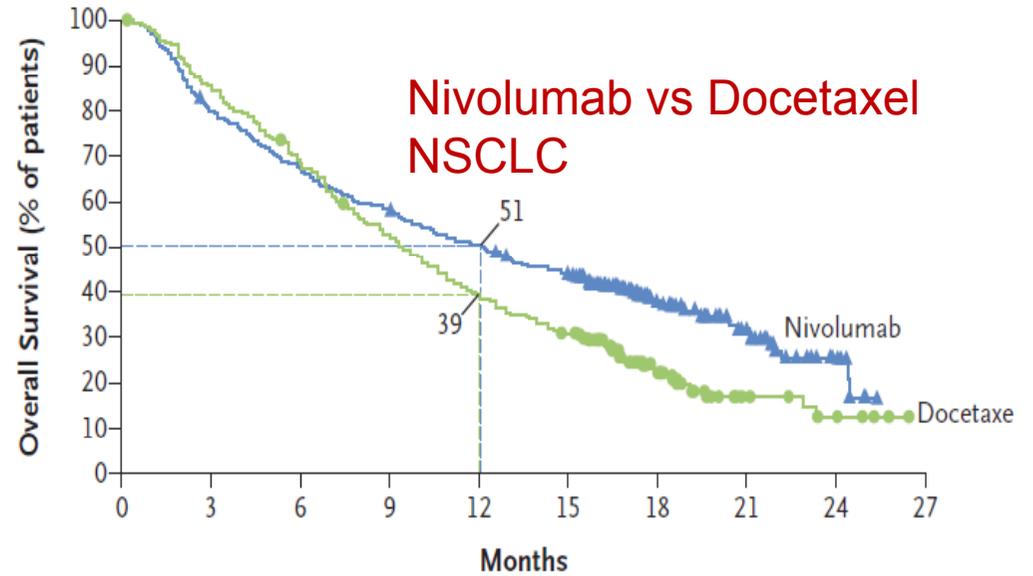
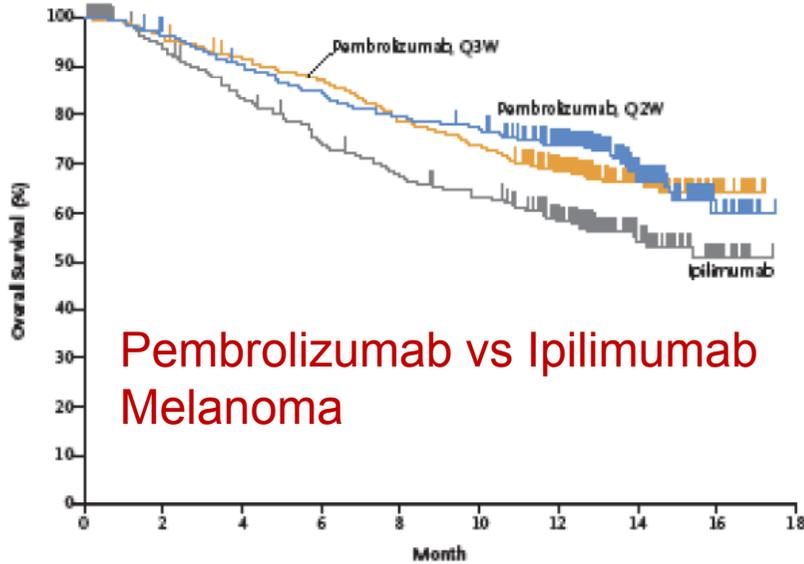


35

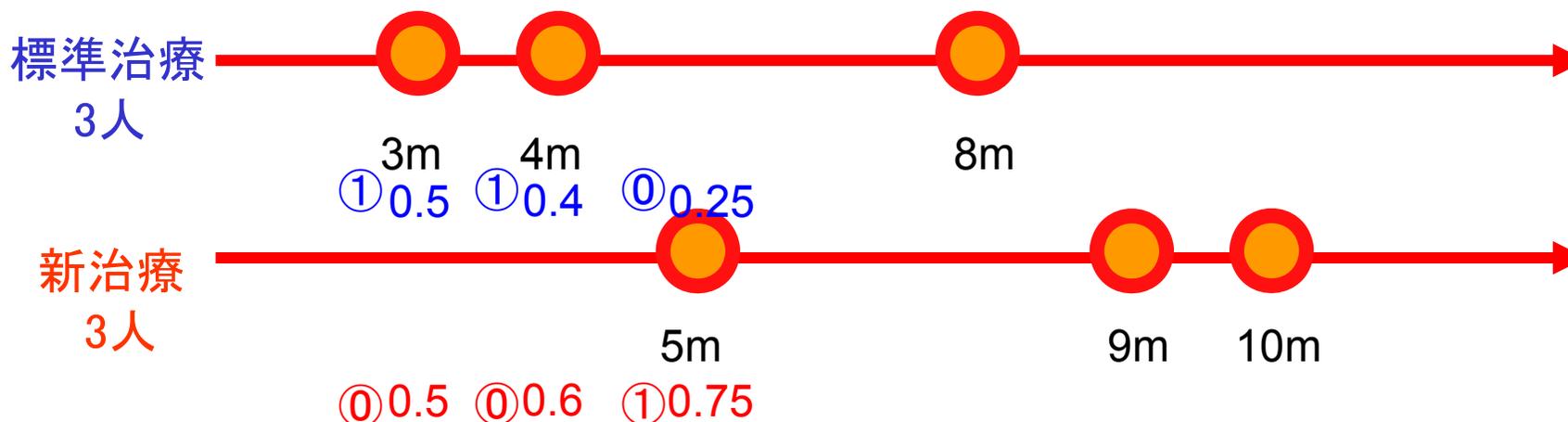
免疫チェックポイント阻害剤でも同様の懸念？



Delayed effect?



ログランク検定の復習



1. 帰無仮説(生存時間に差がない)の下で、各イベント発生時点での期待イベント数を計算する
 - ◆ 各群の"Number at risk"に比例
2. 各時点での観察イベント数(0 or 1)と期待イベント数の"ズレ"を計算
3. 各イベントが生じた時点の"ズレ"を合計
4. "ズレ"の合計以上に極端な値が出る確率(=p値)を計算

(水澤のスライドを改変)³⁸

全ての時点のズレを合計

	標準治療群			新治療群		
	観測 イベント数	期待 イベント数	ズレ	観測 イベント数	期待 イベント数	ズレ
3m	1	0.5	0.5	0	0.5	-0.5
4m	1	0.4	0.6	0	0.6	-0.6
5m	0	0.25	-0.25	1	0.75	0.25
8m	1	0.33	0.67	0	0.67	-0.67
9m	0	0	0	1	1	0
10m	0	0	0	1	1	0
合計	3人	1.48人	1.52人	3人	4.52人	-1.52人

期待されるよりも
1.52人死亡が多かった

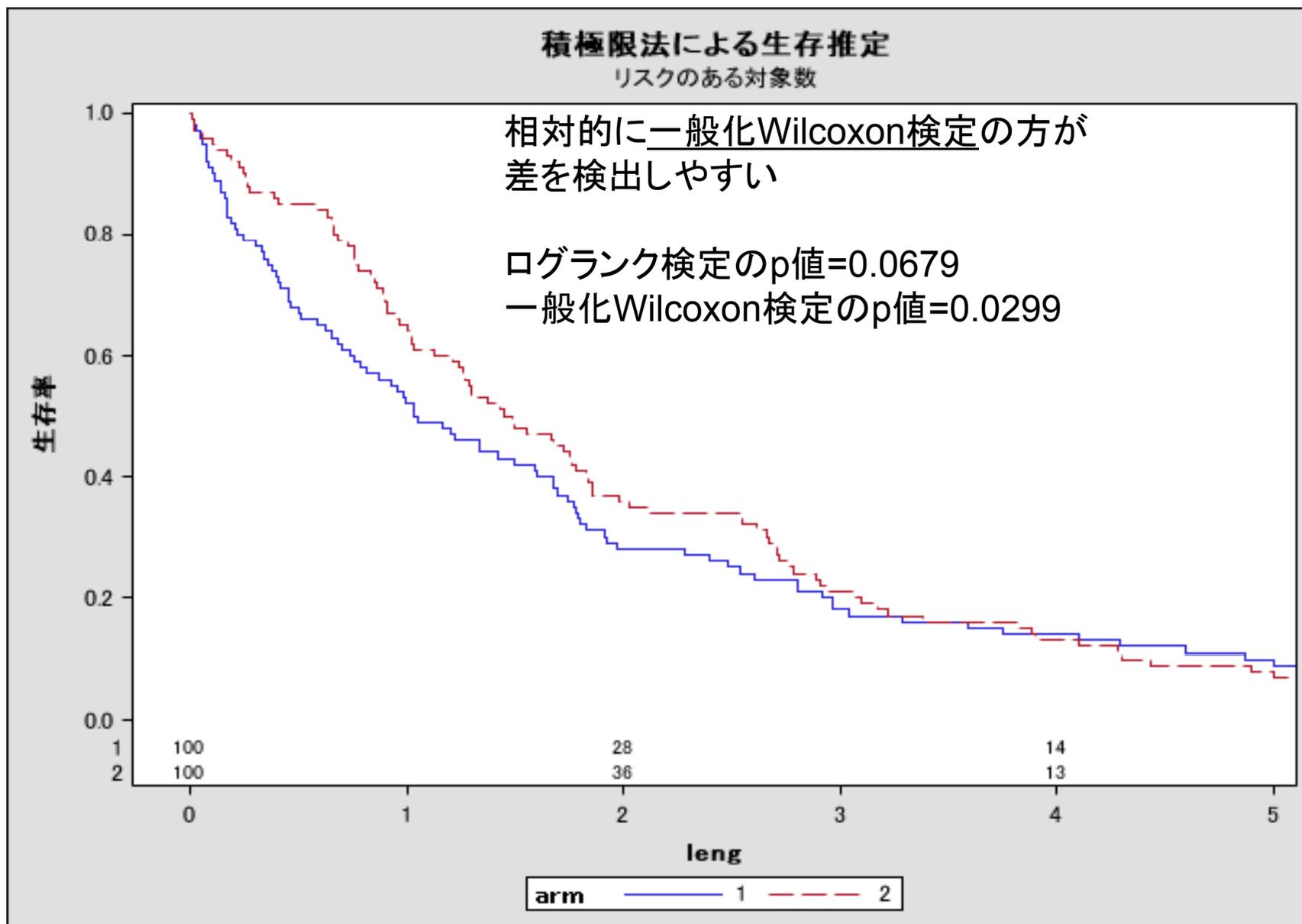
期待されるよりも
1.52人死亡が少なかった

(水澤のスライドを改変)

検定手法ごとに"ズレ"を足す重みが違う

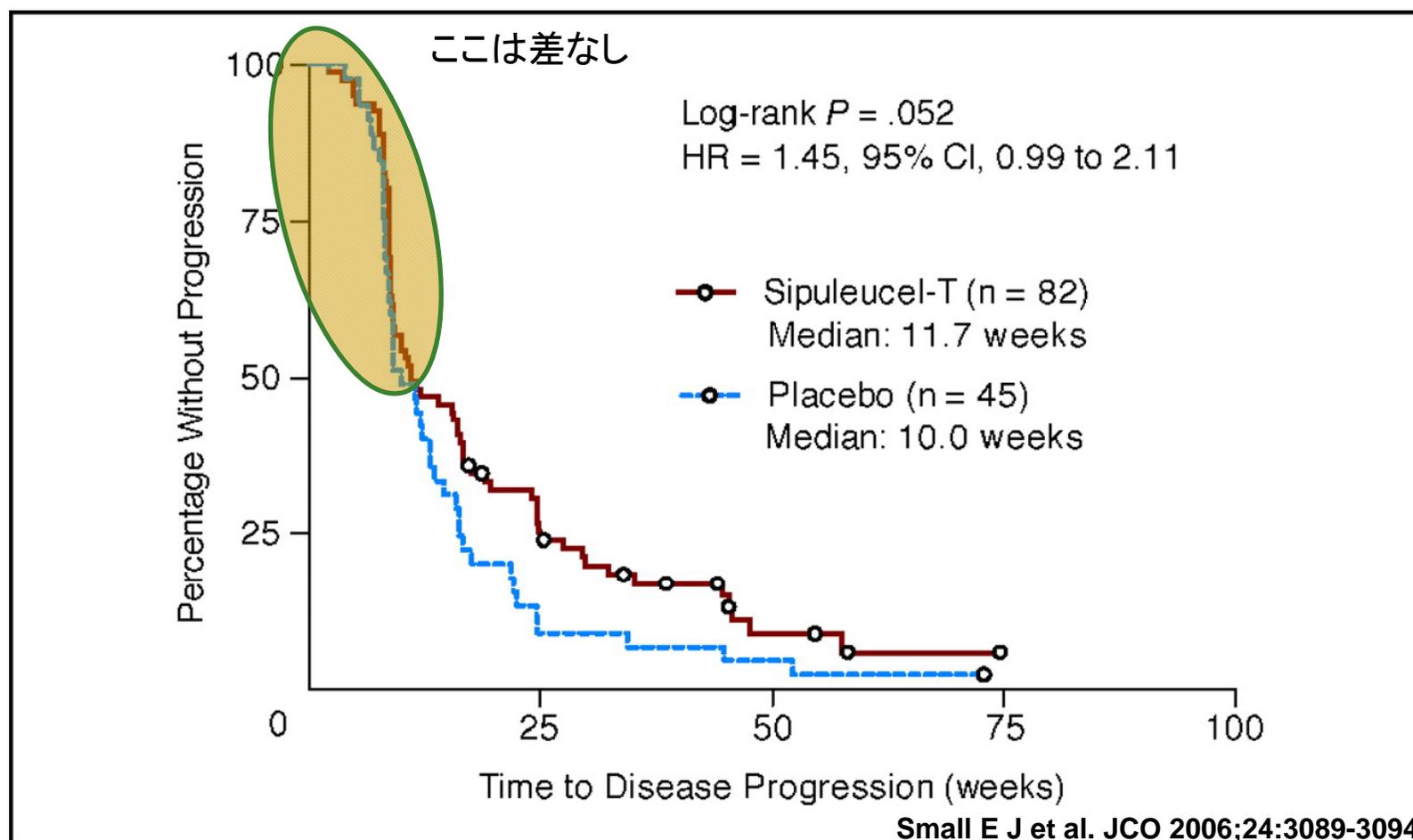
- ログランク検定
 - 重み: 1
 - どの時点のズレも同じ重み
- 一般化Wilcoxon検定
 - 重み: Number at risk (リスク集合)
 - Number at riskが多い早期のイベントを重視
- Fleming-Harrington検定
 - 重み: $S(t_{(f-1)})^p \times [1 - S(t_{(f-1)})]^q$
 - $S(t_{(f-1)})$ は、ほぼその時点での生存割合に相当
 - $p=0, q=0 \rightarrow$ 重み1 (ログランク検定)
 - $p=1, q=0 \rightarrow$ 重み $S(t_{(f-1)}) \div$ 時点生存割合 (早期を重視)
 - $p=0, q=1 \rightarrow$ 重み $1 - S(t_{(f-1)}) \div$ $1 -$ 時点生存割合 (晩期を重視)
 - 免疫療法では通常、晩期を重視した設定を採用
 - p と q は必ず事前にプロトコールに定める必要あり
 - p と q の設定法にコンセンサスはない

時点の前の方の差が大きい場合



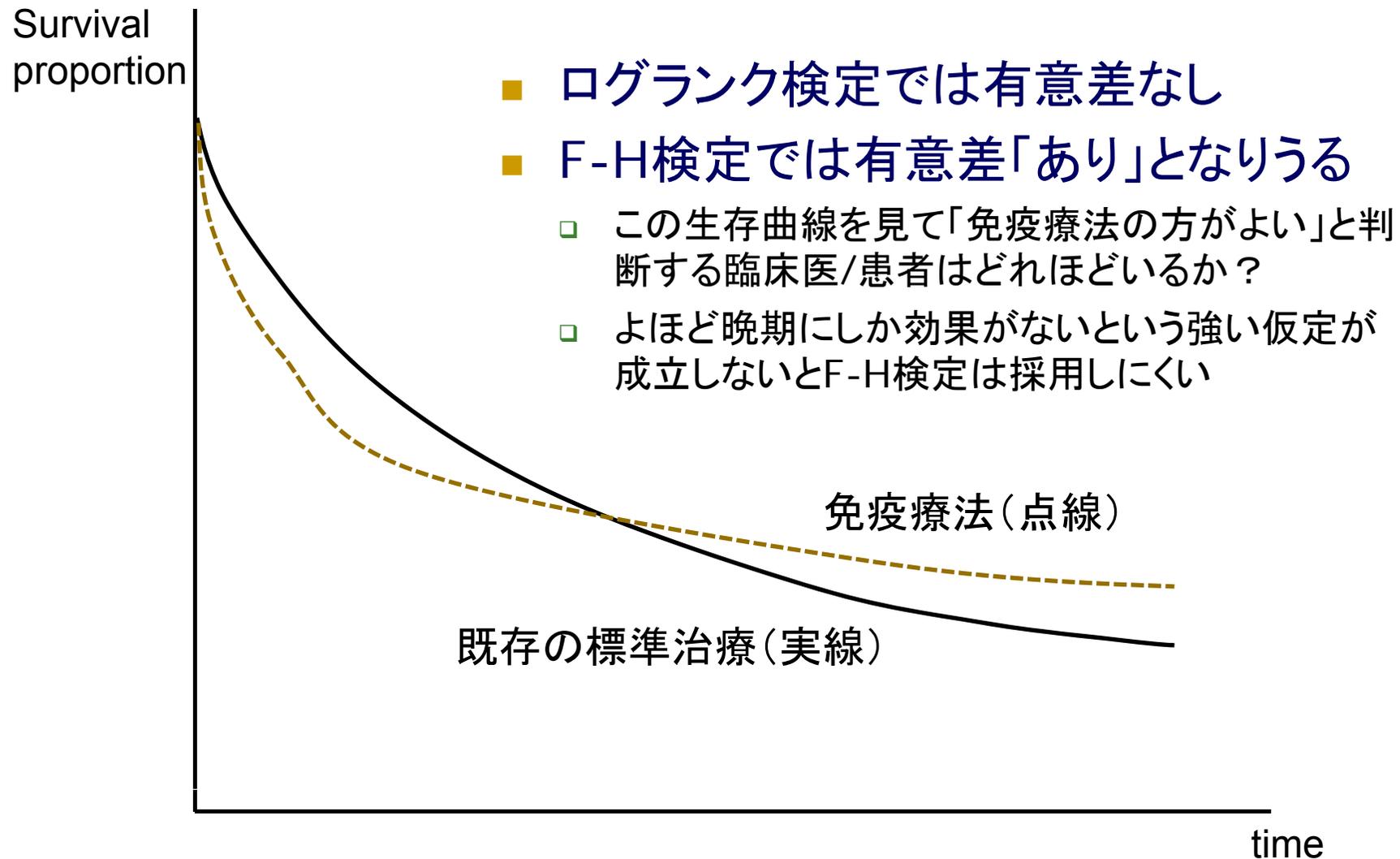
時点の後の方が大きい場合

- 事前に時点の後ろの方が差を重視した検定(ex: F-H検定)をしておけば統計的に有意となったかもしれない



42

生存曲線がクロスする場合



どの検定手法を選択すべきか？

- 時点毎のイベントの重みを変える
 - 特に晩期に「薬効があるかどうか」を確かめたい・・・ならF-H検定は正当化される
 - 標準治療を決める状況では多くの場合に正当化されない
 - エンドポイントとは「患者のベネフィットを測るものさし」

- どの時点で死ぬのがイヤか？
 - ログランク検定: どの時点で死ぬのもイヤ
 - 一般化Wilcoxon検定: 早期に死ぬ方がもっとイヤ
 - F-H検定: 晩期で死ぬ方がもっとイヤ

- ログランク検定は最も前提が少ないニュートラルな手法
 - 強い理由がない限りはログランク検定が推奨される

Short summary

- 免疫療法では晩期に生存曲線が開くことがある (delayed effect)
- 免疫チェックポイント阻害剤でdelayed effectがあるかどうかはcontroversial
- delayed effectを考慮してF-H検定を使うのは、晩期しか効果がないという強い仮定が成り立つ薬効探索の段階のみ
- 従来薬剤との比較、標準治療を確立する段階では最も仮定が少ないログランク検定が推奨される

Take home messages

- 免疫療法では従来の薬剤と異なる薬効の現れ方が見られる (flare effect, delayed effect)
- 臨床試験は「比較」である
 - 従来の薬剤と比較する場合には同じものさしで比べる必要あり
- 効果判定の観点では従来のRECISTやログランク検定で十分に比較可能

Thank you for your kind attention !!