Clinical Endpoints

Michael Morse, MD, MHS
Medical Oncology
Duke University Medical Center

Endpoint Overview

- Endpoint: In clinical trials, an event or outcome that can be measured objectively to determine whether the intervention being studied is beneficial. (cancer.gov)

FDA Guidance

Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics

Guidance for Industry

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

December 2018
Clinical/Medical

https://www.fda.gov/media/71195/download

"Direct" Endpoints

- Clinically meaningful endpoints that directly measure how a patient **feels**, **functions**, or **survives**
- Endpoints that in themselves represent or characterize the clinical outcome of interest
  - Objective: survival, disease exacerbation, clinical event (e.g. MI, stroke), etc.
  - Subjective: symptom score, "health related quality of life" (validated instrument), etc.
- Customarily, the basis for approval of new drugs

Note: The term "direct" is used here to distinguish from "surrogate" endpoints, but this term is not uniformly utilized. Others may refer to these as "true" or "clinically meaningful" endpoints

EJ Sullivan, https://www.fda.gov/media/84987/download

Surrogate Endpoints

- A surrogate endpoint is a laboratory measure or a physical sign that is intended to be used as a substitute for a clinically meaningful endpoint.
- Changes induced by a therapy on a surrogate endpoint are expected to reflect changes in a clinically meaningful endpoint.
- This expectation must be supported by strong data ("validation").
  - Examples of failures of apparently reasonable proposed surrogate endpoints have led to significant skepticism.
- Ideally, the surrogate should exist within the therapeutic pathway between the drug and meaningful benefit
  - i.e. the drug results in the therapeutic benefit by virtue of its effect on the surrogate

EJ Sullivan, https://www.fda.gov/media/84987/download

FDA statement on oncology endpoints

- "…several oncology endpoints can serve different purposes (i.e., clinical endpoint that represents clinical benefit for traditional approval, surrogate endpoint to support traditional approval, surrogate endpoint to support accelerated approval) based on the specific context of use.
  - The determination is based on the specific diseases and is highly dependent upon factors such as effect size, effect duration, depth of response (e.g., number of CRs), available therapy, disease setting, location of disease, the clinical consequences of delaying or preventing disease progression or delaying administration of more toxic therapies, and the risk-benefit relationship" (https://www.fda.gov/media/71195/download)

Endpoint Families

- Primary Endpoints
  - Endpoint(s) necessary and/or sufficient to establish efficacy (define a successful trial)
- Secondary Endpoints
  - Not sufficient to establish efficacy in the absence of an effect on the primary endpoints; not required for establishing efficacy
  - Potentially could lead to additional labeling claims
- Exploratory Endpoints
  - Hypothesis generating endpoints (clinical utility unknown)
  - Variations on primary or secondary endpoints (alternate 'responder' definitions, alternate timepoints)

http://onbiostatistics.blogspot.com/2016/12/control-for-type-i-error-or-adjustment.html

Primary, secondary, tertiary endpoints

- Prespecified outcome that directly addresses the hypothesis of the trial
- The basis of trial designs, including patient enrollment

Primary endpoint

- Key secondary endpoints are examined if the primary endpoint is proven
- Sample size may or may not be sufficient to reach statistical significance
- Interpret with caution

Key secondary endpoint

Secondary endpoints

Subgroup analyses

- Statistical results are not typically presented

Post hoc analyses

Tertiary endpoint

Exploratory endpoints

Leukemia Research 2018;68:32-39

Endpoints	Definition	Advantages	Limitations
Overall survival (OS)	Time from randomization* until death from any cause	• Universally accepted measure of direct benefit • Easily and precisely measured	• May require a larger trial population and longer follow-up to show statistical difference between groups • May be affected by crossover or subsequent therapies • Includes deaths unrelated to cancer

*In nonrandomized trials, time from study enrollment is commonly used

https://www.fda.gov/media/71195/download

https://www.biooncology.com/clinical-trials/efficacy-endpoints.html

Endpoints	Definition	Advantages	Limitations
Progression-free survival (PFS)	Time from randomization* until disease progression or death	• Requires small sample size and shorter follow-up time compared with OS • Includes measurement of stable disease (SD) • Not affected by crossover or subsequent therapies • Generally based on objective and quantitative assessment	• Validation as a surrogate for survival can be difficult in some treatment settings • Not precisely measured (ie, measurement may be subject to bias) • Definition may vary among trials • Requires frequent radiologic or other assessments • Requires balanced timing of assessment among treatment arms
Time to progression (TTP)	Time from randomization* until objective tumor progression; does not include deaths		

*In nonrandomized trials, time from study enrollment is commonly used

https://www.fda.gov/media/71195/download
https://www.biooncology.com/clinical-trials/efficacy-endpoints.html

Endpoints	Definition	Advantages	Limitations
Time to treatment failure (TTF)	Time from randomization* to discontinuation of treatment for any reason, including disease progression, treatment toxicity, and death	• Useful in settings in which toxicity is potentially as serious as disease progression (eg, allogeneic stem cell transplant)	• Does not adequately distinguish efficacy from other variables, such as toxicity

*In nonrandomized trials, time from study enrollment is commonly used

https://www.fda.gov/media/71195/download

https://www.biooncology.com/clinical-trials/efficacy-endpoints.html

Endpoints	Definition	Advantages	Limitations
Time to next treatment (TTNT)	Time from end of primary treatment to institution of next therapy	• For incurable diseases, may provide an endpoint meaningful to patients	• Not commonly used as a primary endpoint • Subject to variability in practice patterns

https://www.fda.gov/media/71195/download

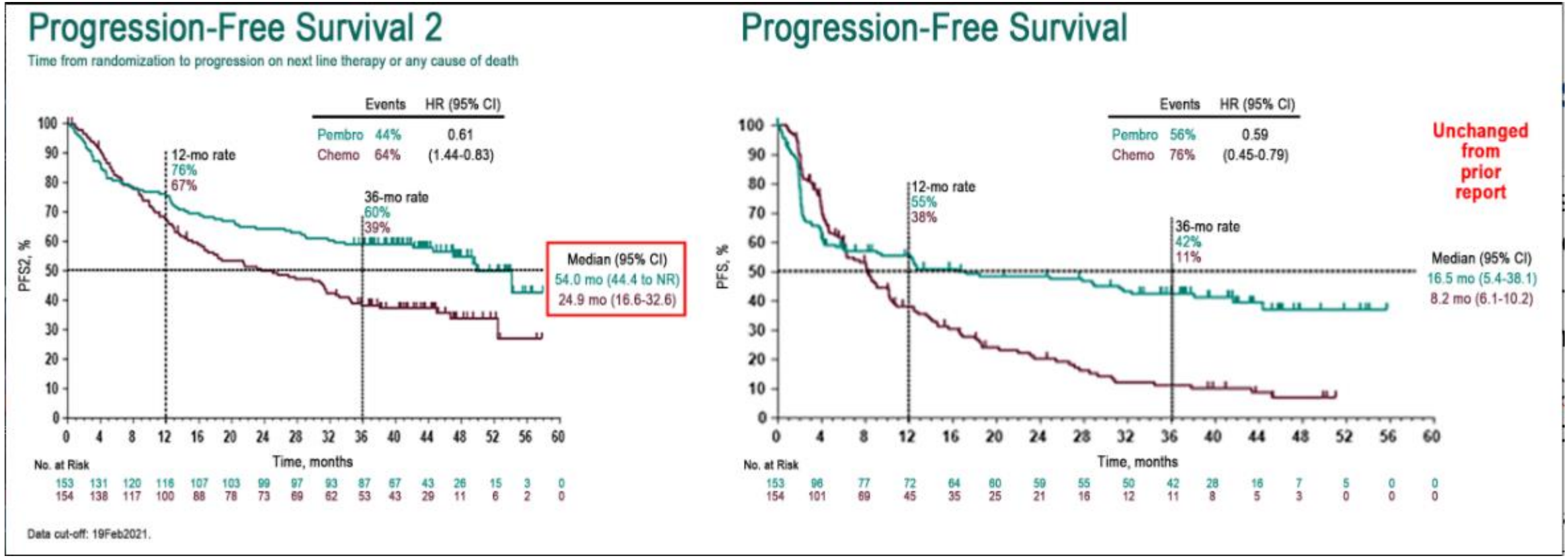
https://www.biooncology.com/clinical-trials/efficacy-endpoints.html

Endpoints	Definition	Advantages	Limitations
Objective response rate (ORR)	Proportion of patients with reduction in tumor burden of a predefined amount	• Can be assessed in single-arm trials • Requires a smaller population and can be assessed earlier, compared with survival trials • Effect is attributable directly to the drug, not the natural history of the disease	• Not a comprehensive measure of drug activity
Duration of response (DoR)	Time from documentation of tumor response to disease progression		

https://www.fda.gov/media/71195/download

https://www.biooncology.com/clinical-trials/efficacy-endpoints.html

PFS 2

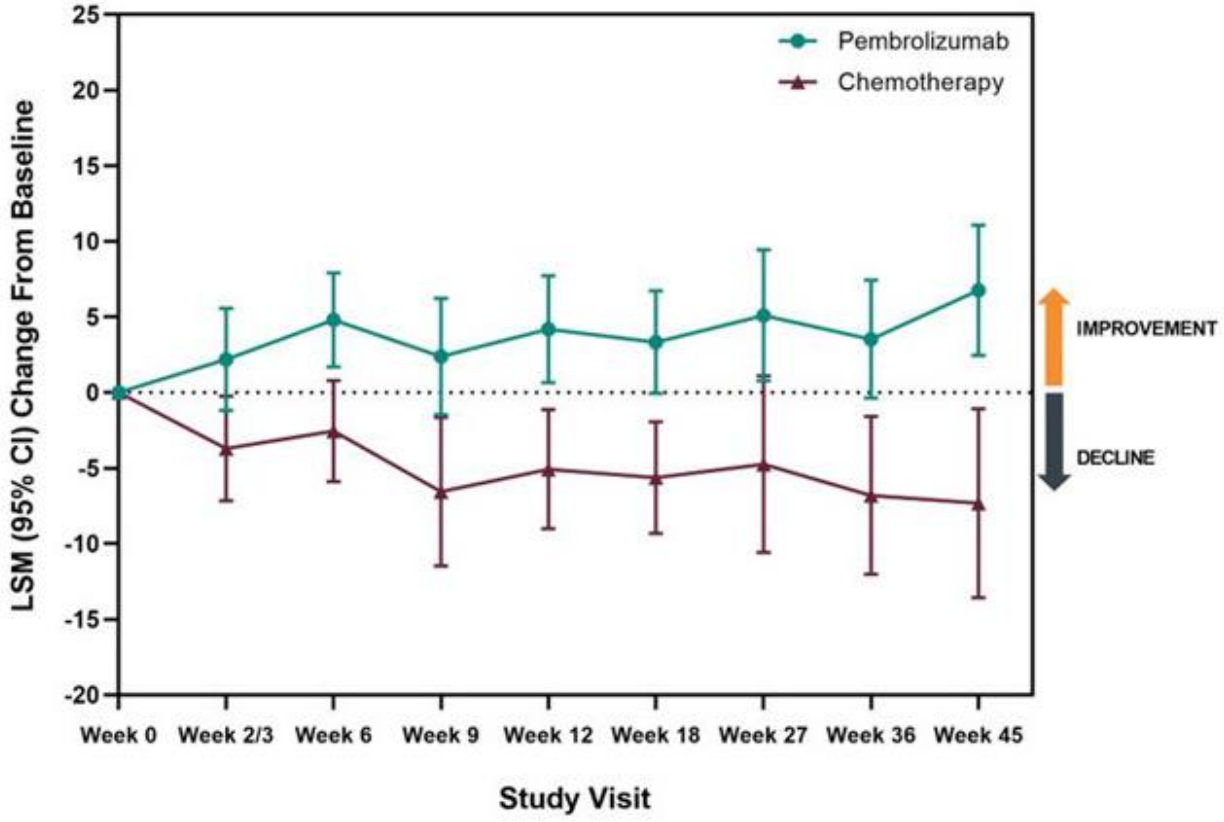


From Keynote-177 study

Other Clinical endpoints

- Composite and Co-primary endpoints (OS + PFS)
  - See: “Multiple Endpoints in Clinical Trials: Guidance for Industry” (https://www.fda.gov/media/102657/download)
- Improvement (or lack of decline) in QOL scores
- Improvement (or lack of decline) in performance status
  - E.g., time to deterioration from PS 0,1 to PS 2
- Improvement in composite scores (pain, weight loss,…)

Change in QoL scores over time



Change in score from baseline over time
EORTC QLQ-C30 GHS/QoL.

KEYNOE-177 study

Combining measures to characterize toxicities and benefits: The outcome measure patients want

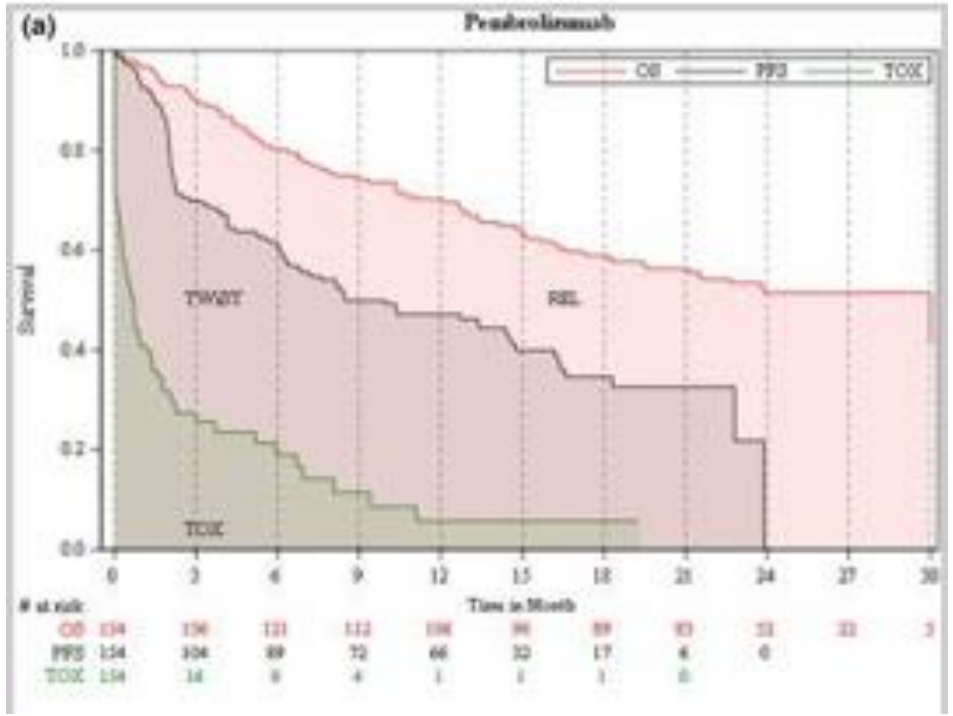
American Society of Clinical Oncology (ASCO) Value Framework (J Clin Oncol. 2016; 34:2925-34) proposed metrics that combine clinical benefit, side effects, and improvement in patient symptoms or QoL.

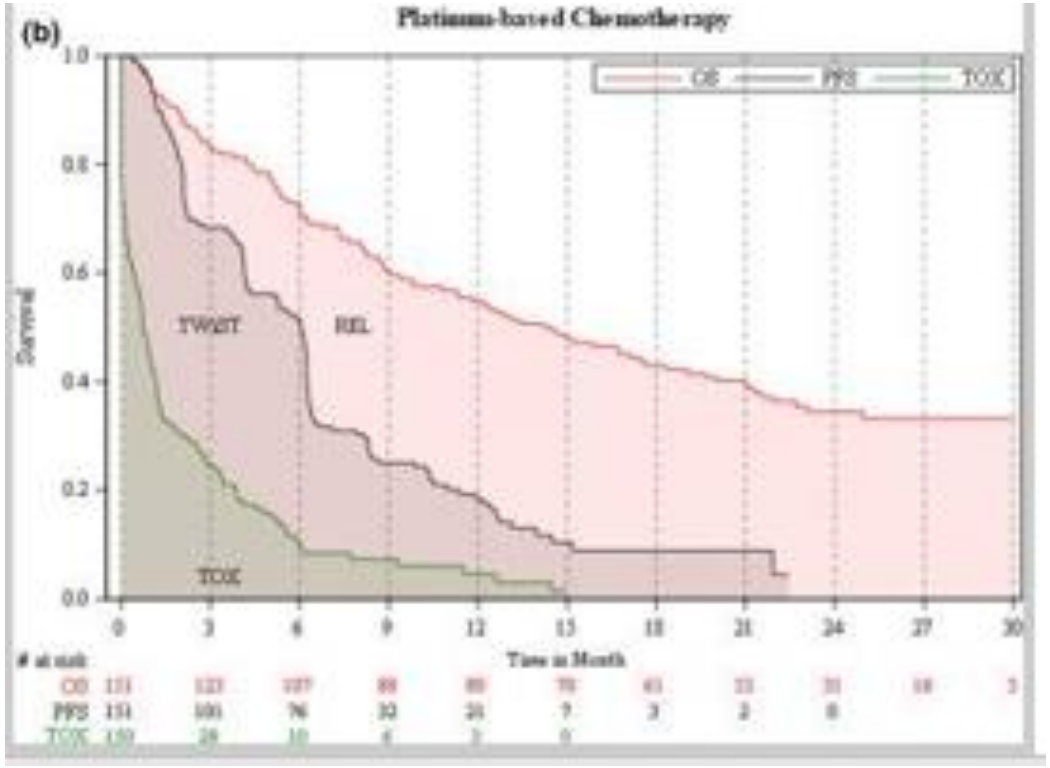
Quality-adjusted life year (QALY)
Has been supplanted by:

Quality-adjusted Time Without Symptoms of disease or Toxicity of treatment (Q-TWiST)

Patient survival is divided into: time (from randomization to progression) with grade 3 toxicity, time without symptoms or toxicities, and time from disease progression to death and utilities are given to each time.
Clinically important difference for Q-TWiST: 10-15% of OS in a study.

Q-TWIST OF PEMBROLIZUMAB VS DOCETAXEL





pembrolizumab had 2.49 months greater Q-TWiST compared to to platinum-based chemotherapy at 24 months

Pharmacoeconomics. 2019; 37(1): 105–116.

Tumor Response History

- 1981: World Health Organization (WHO) first published tumor response criteria
- Overall assessment of tumor burden by summing the products of bidimensional lesion measurements and determined response to therapy by evaluation of change from baseline while on treatment.
- Ad hoc modifications leading to confusion about actual efficacy
- International Working Party formed and new criteria--RECIST (Response Evaluation Criteria in Solid Tumors)--- published in 2000

Key advances in RECIST

- **Use of unidimensional, rather than bidimensional, measures for overall evaluation of tumor burden**
  - Works well except in mesothelioma, and except where devascularization occurs (mRECIST)
- Definitions of minimum size of measurable tumors (10mm by CT scan)
- Instructions on how many lesions to follow
  - up to 10; maximum of five per organ site
- Definitions of PR as “At least a 30% decrease in the sum of diameters of target lesions” compared with baseline.
- Widely adopted by academic institutions, cooperative groups, and industry for trials where the primary endpoints are objective response or progression.
- Regulatory authorities accept RECIST as an appropriate guideline for response/progression assessments.

RECIST1.1

- Questions arose with RECIST
  - Can fewer than 10 lesions can be assessed without affecting the overall assigned response for patients (or the conclusion about activity in trials)?
  - how to apply RECIST in randomized phase III trials where progression, not response, is the primary endpoint particularly if not all patients have measurable disease;
  - Whether or how to utilize newer imaging technologies (FDG-PET and MRI);
  - How to handle assessment of lymph nodes;
  - Whether response confirmation is truly needed;
  - The applicability of RECIST in trials of targeted non-cytotoxic drugs.

Highlights of revised RECIST 1.1

- Number of lesions to be assessed: Maximum 5 (two per organ, maximum).
- Assessment of pathological lymph nodes is now incorporated:
  - Nodes with a short axis of >15 mm are considered measurable (target lesions).
  - Nodes that shrink to <10 mm short axis are considered normal.
- Confirmation of response only required for trials with response primary endpoint but NOT required in randomized studies (control arm).
- Disease progression is clarified:
  - 20% increase in sum of target lesions AND 5 mm absolute increase (avoids over-calling PD).
  - Guidance on progression of non-measurable lesions
- More guidance on optimal anatomical assessment of lesions and FDG-PET

Challenges with RECIST 1.1 in immunotherapy

- Pseudoprogression:
  - Radiologic tumor progression (new lesions, or enlarging lesions) from baseline that is not confirmed as progression on subsequent radiologic assessment.
  - Rate of 7% in melanoma
  - 1.5 – 3.0% (up to 4.7%) In non-small cell lung cancer (NSCLC) and urothelial carcinoma
  - No biomarker to predict pseudoprogression
    - PD-L1 expression level and tumor infiltrating lymphocytes have failed to correlate with the rates of pseudoprogression.
    - Circulating tumor DNA changes?

More History

- Two-dimensional immune-related response criteria (irRC) proposed in 2009 (Wolchok, Clin Cancer Res 2009;15:7412e20)
- Simplification of criteria (e.g., unidimentional) proposed in 2013, irRECIST (immune-related) (J Immunother Cancer 2016;4:30).
- RECIST working group published iRECIST, to standardize response assessment among immunotherapy clinical trials (Lancet Oncol 2017;18: el43ee152.)
  - Responses assigned using iRECIST have a prefix of “i” (ie, immune)
    - iCR, iPR, iUPD (unconfirmed PD) or iCPD (confirmed PD)

What's the same? What's different

- Still use RECIST1.1 to define measurable
- Same methods of measurement (mostly CT/MRI)
- Most tumor response assessment is the same

- iRECIST: if new lesion or enlarging lesions, assigned iUPD. Must have further progression to assign iCPD
- iRECIST: **resets the bar if RECIST 1.1 progression is followed at the next assessment by tumour shrinkage** (i.e., next progression is iUPD)
- iRECIST: if lesion remains stable, the timepoint response would again be iUPD

Comparison of response criteria

	RECIST 1.1 [10]	irRC [11]	irRECIST [12]	iRECIST [13]
Lesion measurement	Unidimensional	Bidimensional	Unidimensional	Unidimensional
Baseline lesion size	≥10 mm	5 × 5 mm	≥10 mm	≥10 mm
Baseline lesion number	5 total, 2 per organ	10 total, 5 per organ	5 total, 2 per organ	5 total, 2 per organ
CR	Disappearance of all lesions	Disappearance of all lesions	Disappearance of all lesions	Disappearance of all lesions
PR	≥30% decrease from baseline	≥50% decrease from baseline	≥30% decrease from baseline	≥30% decrease from baseline
SD	Neither PR or PD	Neither PR or PD	Neither PR or PD	Neither PR or PD
PD	≥20% increase from nadir (≥5 mm)	≥25% increase from nadir	≥20% increase from nadir (≥5 mm)	≥20% increase from nadir (≥5 mm)
Confirmed progressive disease	Not applicable	At least 4 weeks after	At least 4 weeks after and up to 12 weeks	At least 4 weeks after and up to 8 weeks
Appearance of new lesions	Always PD	Incorporate in the sum of measurement	Incorporate in the sum of measurement	Unconfirmed progressive disease, not included in the sum of measurement

RECIST, response-evaluation criteria in solid tumors; irRC, immune-related response criteria; irRECIST, immune-related RECIST; iRECIST, immune RECIST; CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease.

Int J Mol Sci. 2019 Jun; 20(11): 2674.

Hyperprogression: accelerated tumor growth rate (TGR)

- Definitions
  - Progression (RECIST) at the first evaluation and a ≥2-fold increase in TGR during ICI therapy compared with pretreatment kinetics.[1]
  - Time to treatment failure <2 months, >50% increase in tumor burden and >2-fold increase in progression pace.[2]
  - Time to treatment failure <2 months, and increase of at least 40% in the target tumor burden or at least a 20% increase and new lesions.[3]
  - Disease progression at the first evaluation with ΔTGR exceeding 50%.[4]

- Measurement: change in tumor volume ? Change in the largest diameters of target lesions?

[1]Champiat, Clin Cancer Res. 2017;23:1920-1928. [2]Kato, Clin Cancer Res. 2017;23:4242-4250.
[3]Matos J Clin Oncol. 2018;36(15 suppl):3032. [4]Ferrara, JAMA Oncol. 2018;4(11):1543-1552

Table 1 | **Studies using tumour kinetics to identify hyperprogressive disease patterns**

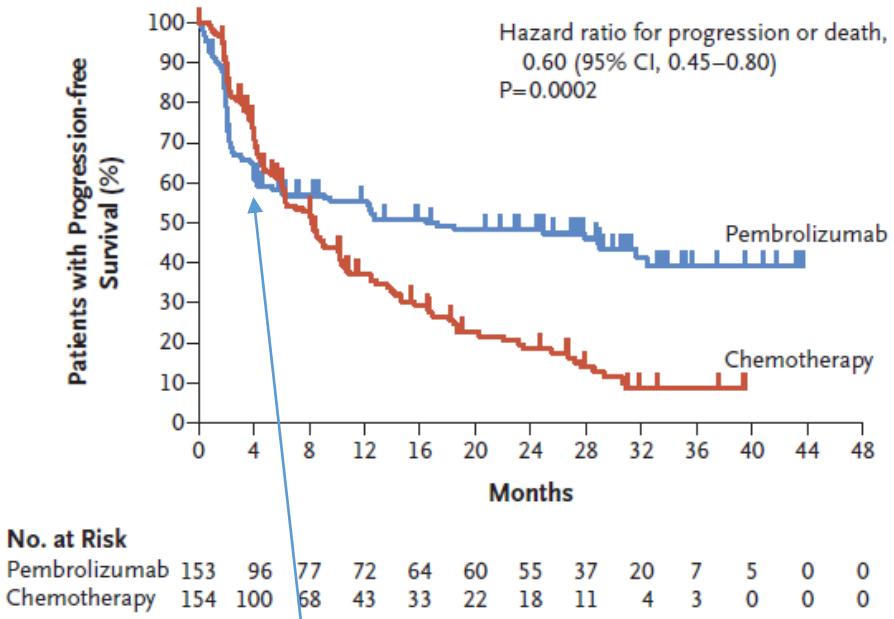
Study	Reported frequency for hyperprogressive disease	Tumour types	Factors involved	Criteria used to define hyperprogressive disease[a]	Refs
Champiat et al.	9% (12/131)	• Melanoma (9%; 4/45) • Urothelial carcinoma (25%; 2/8) • Colorectal carcinoma (12%; 1/8) • Lymphoma (14%; 1/7) • Ovarian carcinoma (40%; 2/5) • Cholangiocarcinoma (50%; 1/2) • Uveal melanoma (50%; 1/2)	Older age associated with higher risk of hyperprogressive disease (19% if ≥65 years of age versus 5% if <65 years of age; $P=0.036$)	• RECIST-defined progressive disease at first evaluation • TGR ratio ≥2 (on-treatment versus before treatment)	13
Kato et al.	6% (6/102)	• NSCLC (8%; 3/38) • Urothelial carcinoma (ND; 1/ND) • Triple negative breast cancer (ND; 1/ND) • Endometrial carcinoma (ND; 1/ND)	Presence of *MDM2* amplifications or *EGFR* alterations associated with median time to treatment failure <2 months ($P=0.007$ and $P=0.005$, respectively)	• Time-to-treatment failure <2 months • >50% increase in tumour burden compared with pre-immunotherapy (on imaging) • >2-fold increase in progression pace	14
Saâda-Bouzid et al.	29% (10/34)	HNSCC (all patients)	Regional recurrence associated with higher rate of hyperprogressive disease (90% for TGK ratio ≥2 versus 37% for TGK ratio <2; $P=0.008$)	• TGK ratio ≥2 (on-treatment versus before treatment)	15
Ferrara et al.	14% (56/406)	NSCLC (all patients)	Number of metastatic sites >2 associated with higher rate of hyperprogressive disease (19% for >2 versus 9% for ≤2; $P=0.005$)	• RECIST-defined disease progression at first evaluation • ΔTGR increase >1.5 (on-treatment versus before treatment)	16

HNSCC, head and neck squamous cell carcinoma; ND, not defined; NSCLC, non-small-cell lung carcinoma; TGK, tumour growth kinetics; TGR, tumour growth rate.
[a]TGR (assumption of 3D exponential tumour growth) and TGK (assumption of 2D linear tumour growth) are similar concepts but different methods are used to evaluate them according to investigator's preference.

Champiat., Nat Rev Clin Oncol. 2018 Dec;15(12):748-762.

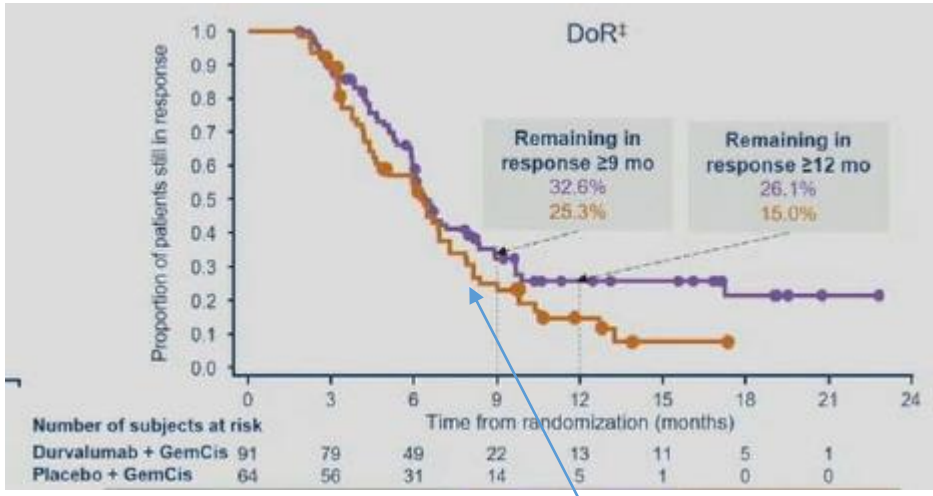
Defining benefit for challenging endpoint scenarios

Keynote-177: pembrolizumab vs chemotherapy
For MSI-H 1st line metastatic CRC (NEJM 2020;383:23)



Initial lack of benefit for some pembrolizumab patients

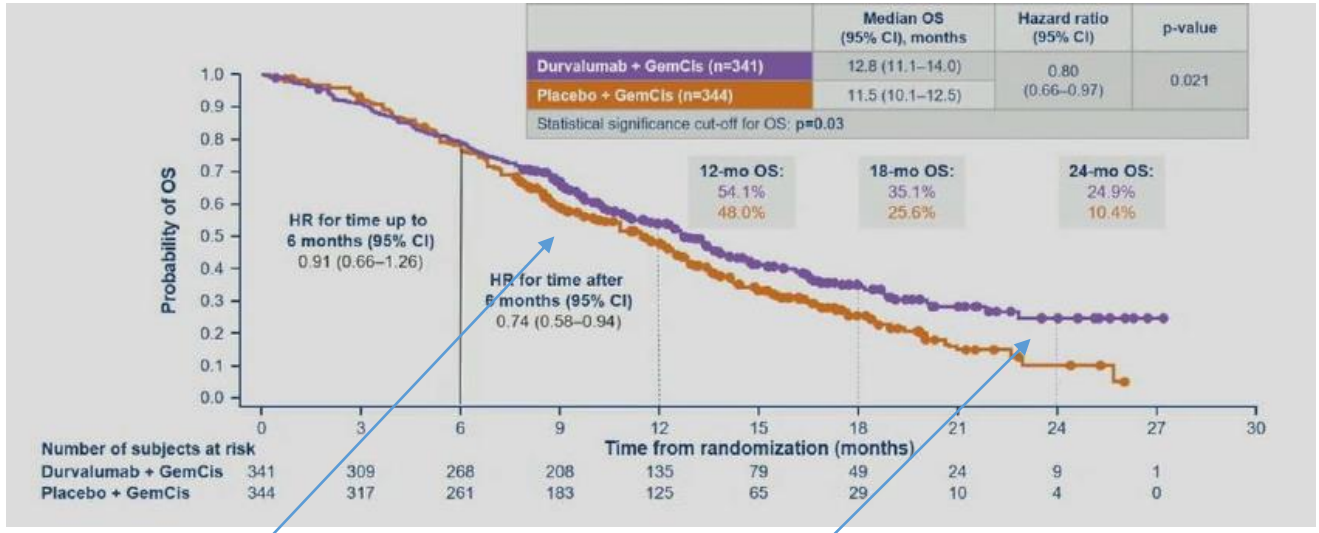
TOPAZ-1: Gemcitabine/Cisplatin +/Durvalumab
In 1st line metastatic cholangiocarcinoma (ASCOGI 2022)



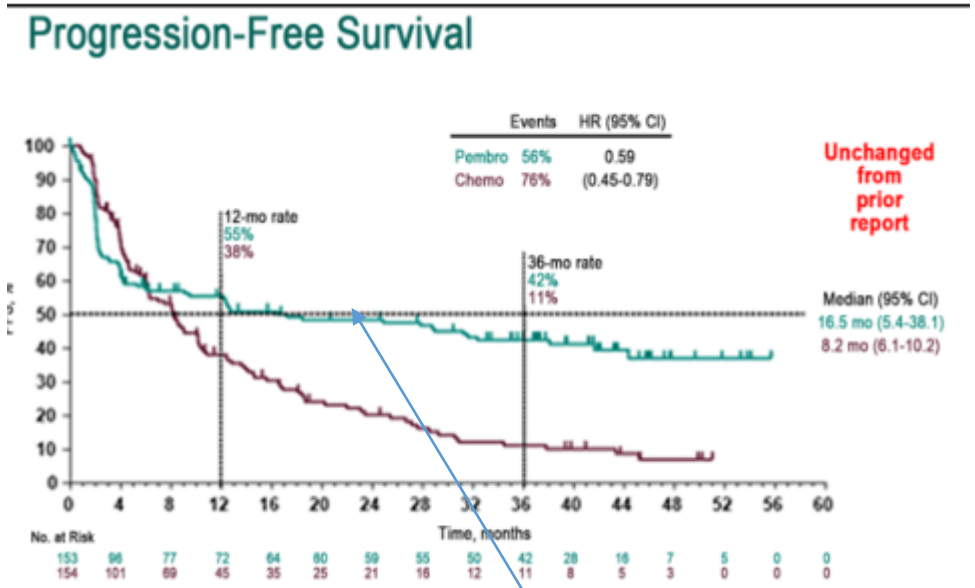
DOR curves do not separate until late.

#SITCWinterSchool

Where landmarks may more accurately demonstrate clinical benefit



1. Median difference is small or negligible
2. When benefits become greater over time



3. When there is unstable mPFS

TOPAZ-1 trial (ASCOGI 2022)

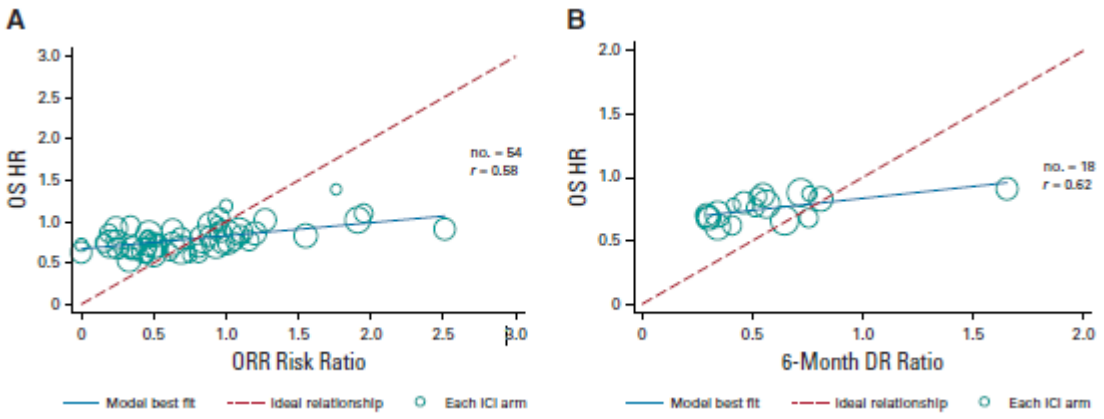
Choice of clinical endpoint for specific studies?

Endpoint	Type of Endpoint			Study Design Recommendations		
	Clinical Endpoint	Surrogate Endpoint for TA*	Surrogate Endpoint for AA**	Randomized	Single-Arm	Independent Blinded Review
Overall Survival	X			X		
Symptom Endpoints (patient-reported outcomes)	X			X		
Disease-Free Survival or Event-Free Survival	X	X	X	X		X***
Objective Response Rate	X	X	X	X	X	X
Complete Response	X	X	X	X	X	X
Progression-Free Survival or Time to Progression	X	X	X	X		X***

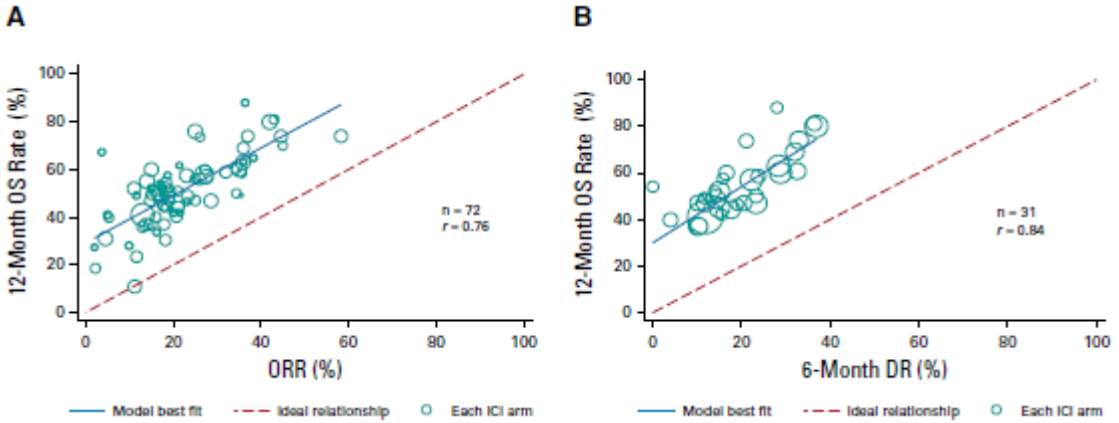
EJ Sullivan, https://www.fda.gov/media/84987/download

How to more effectively use clinical endpoints in IT studies?

ORR, 6 mo Duration of response benefits are not good surrogates for OS benefit



ORR, 6 mo Duration of response are "prognostic" for OS



The effect of the IT on RR or 6 mo DR (relative to another therapy) does not correlate with the OS benefit

OS remains the most important and clinically meaningful end point

The 12 mo OS (regardless of therapy) is associated with the ORR or 6 mo DOR

ORR/DOR can be used in drug development and screening of agents for phase III studies

Summary

- OS is still the gold standard but harder to prove as more therapies become available (or cross-over allowed)
- ORR, duration of response may not be adequate surrogate endpoints for beneficial effect on OS achieved by immunotherapies, but can be used for screening drugs for phase III studies.
- RECIST1.1 is still the standard, but iRECIST provides insight on atypical responses
- Is hyperprogression a true phenomenon and will it become an important endpoint
- Are we measuring all the endpoints that patients think are important?