Weill Cornell Medicine
Meyer Cancer Center

Metrics of success in medical oncology trials

Paul Chapman, MD
Chief Medical Research Officer, Meyer Cancer Center
Medical Director, Cancer Clinical Trial Office
Professor, Weill Cornell Medical College

We're Changing Medicine.

The (NIH) Definition of Clinical Research

Patient-oriented research, epidemiological/behavioral studies, or outcomes/health services research conducted with human subjects or materials of human origin, such as tissues or cognitive phenomena, where the investigator directly interacts with human subjects.

Clinical Investigator

An individual who actually conducts a study or leads a team conducting the study (i.e. under whose immediate direction the drug is dispensed to a subject.)

Principle Investigator Responsibilities

The PI is responsible for:

- EVERYTHING
- EVERYONE'S WORK
- CONFLICTS OF INTEREST
- CONDUCTING RESEARCH WITH INTEGRITY

First steps in designing a clinical trial

- Hypothesis to be tested. Is it important enough to justify experimentation on patients?
- Scientific endpoints
- Involve biostats early
  - Can you enroll enough patients?
  - Lasagna's Law: as soon as a clinical trial begins, the supply of suitable patients becomes one-tenth of what it was said to be before the trial began.

Inclusion Criteria

List of what patient MUST have.

Patient must fulfill **all** of these requirements to be eligible.

Exclusion Criteria

List of what patients CANNOT have.

Example: number of prior therapies, concomitant conditions or medications

Informed Consent

- A PROCESS, not just a document
- Subjects must:
  - Be aware of investigational nature of study
  - Be aware that participation is optional
    - And that they will still receive the best possible alternative care available if they decline to participate
  - Be aware of risks and potential benefits of participation and not participation (alternatives)
  - Understand alternatives
  - Understand study procedures
  - Be aware that they may withdraw consent at any time
  - Have ample time to process the information (and read consent) in an appropriate environment

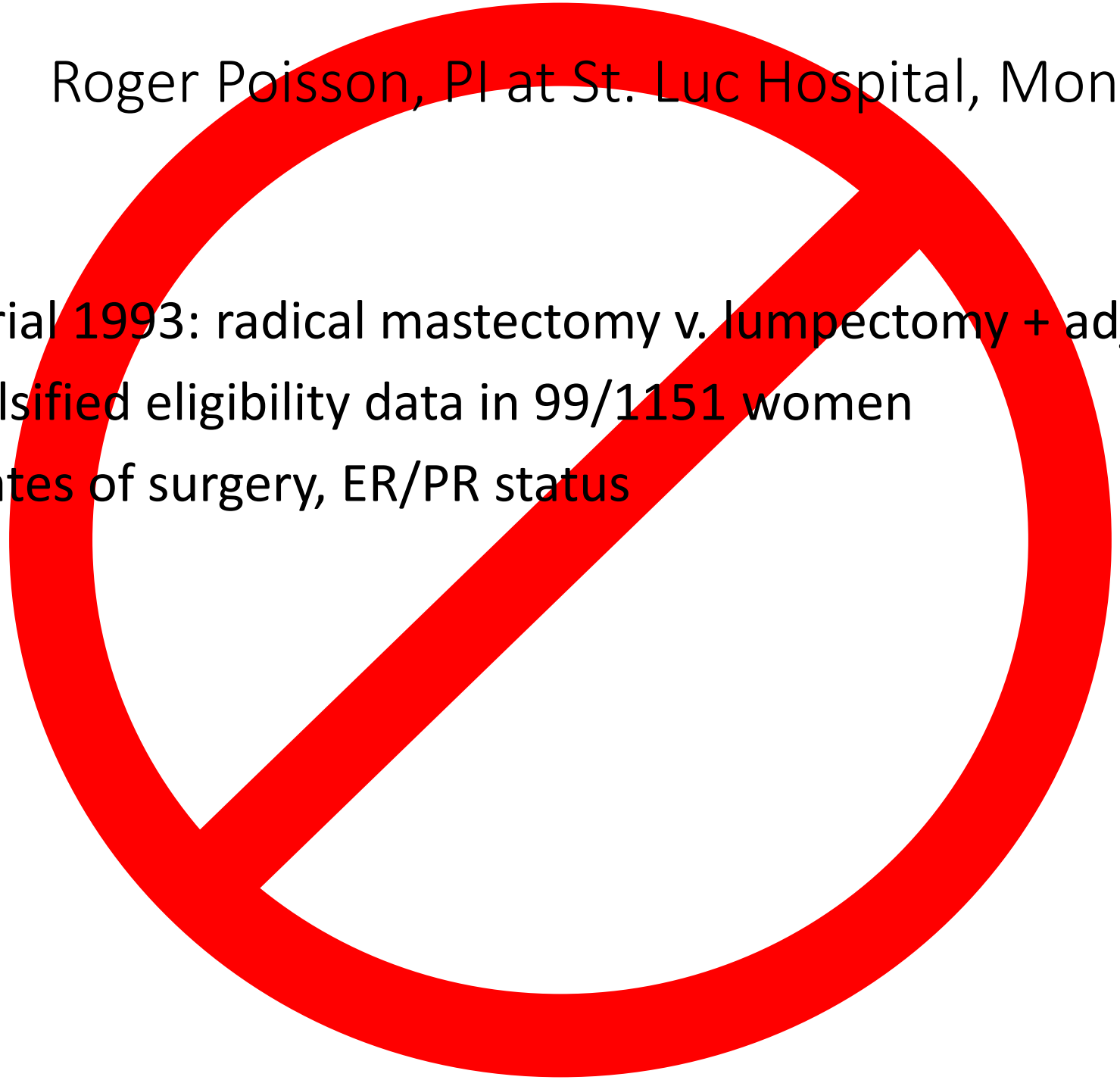
Think of protocol and consent as **contracts**

Roger Poisson, PI at St. Luc Hospital, Montreal

NSABP trial 1993: radical mastectomy v. lumpectomy + adj. chemo

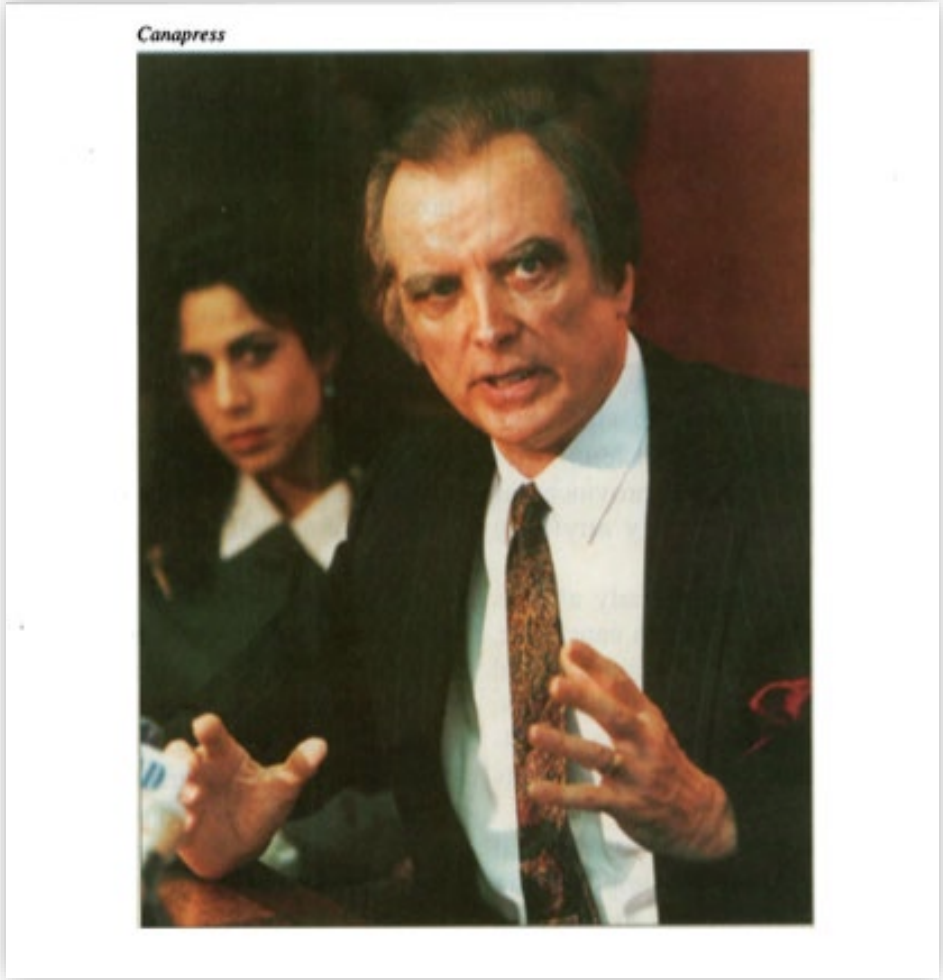
Falsified eligibility data in 99/1151 women

Dates of surgery, ER/PR status



Ethical equipoise

A state of genuine uncertainty on the part of the clinical investigator regarding the comparative therapeutic merits of each arm in a trial.



Roger Poisson, St. Luc Hospital, Montreal

"My sole concern at all times was the health of my patients. I firmly believed that a patient who was able to enter into an NSABP trial received the best therapy and follow-up treatment.

For me, it was difficult to tell a woman with breast cancer that she was ineligible to receive the best available treatment because she did not meet 1 criteria out of 22…"

Metrics of success

Don’t make important what you can measure just because you can’t measure what’s important

Endpoints of clinical trials

Primary endpoints*	Surrogate endpoints
Overall survival	Response rate
Health-related QOL	Relapse-free survival
Tumor-related symptoms	Time-to-treatment failure
Physical functioning	Metastasis-free survival

*Pazdur, Oncologist 2008:13,19-21

Response rate: RECIST 1.1

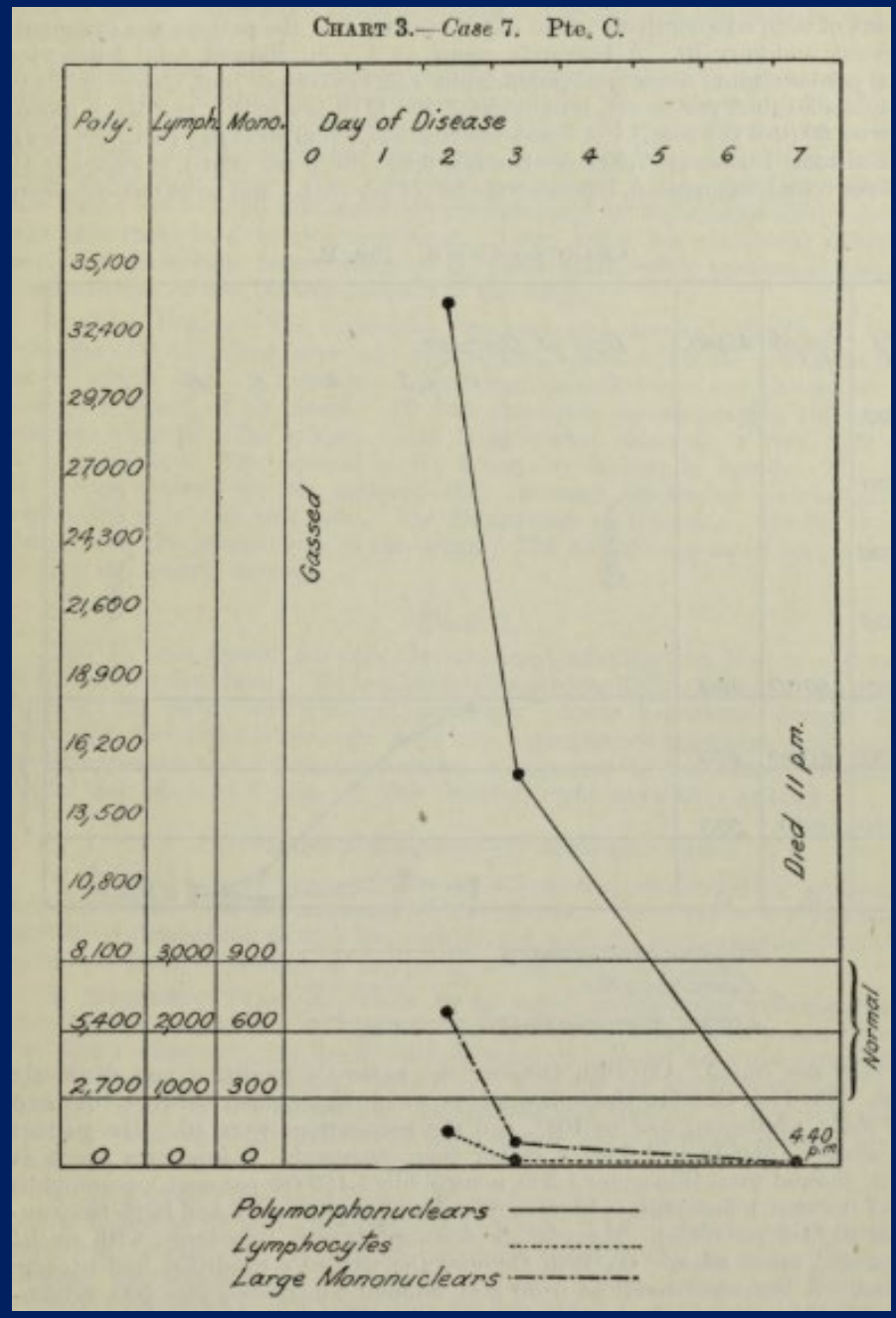
“And you might ask yourself: Well, how did I get here?”

David Byrne, Talking Heads, 1981

Nitrogen mustard gas in WWI

Case 7.

Pte. C., æt. 23, was gassed on 12th October, and sustained at the same time a through-and-through wound of the right shoulder and glancing wounds of the left shoulder and left thigh. On admission next day to No. —— C.C.S., there were early burns, with blistering, of the face, trunk and limbs, and early conjunctivitis with chemosis. The heart and lungs were normal. On admission to No. —— General Hospital on 14th October there was much œdema of the right conjunctiva and conjunctivitis of the left, with slight superficial ulceration of the lower parts of the corneæ. There were extensive multiple burns of the face, body, limbs and scrotum, with many bullæ and patches of necrotic epidermis.



Chemical Warfare Medical Comm. No.17—Report on cases of poisoning by “mustard gas” (Dichloro-ethyl sulphide), with special reference to histological changes and to alterations in leucocyte count. December 1918

SCIENCE

Vol. 103, No. 2675 Friday, April 5, 1946

The Biological Actions and Therapeutic Applications of the B-Chloroethyl Amines and Sulfides

Alfred Gilman, Major, and Frederick S. Philips, 1st Lieutenant, SnC, AUS

Pharmacology Section, Medical Division, CWS, Edgewood Arsenal, Maryland

AT THE CONCLUSION OF WORLD WAR I the theory was generally accepted that mustard gas exerted its vesicant action by releasing hydrochloric acid intracellularly. A few isolated reports appeared describing remote systemic effects of mustard gas on hematopoietic tissues (*1–3*), the gastrointestinal tract (*3–5*), and electrolyte and fluid balance (*3*). Although in the interim between wars the adverse effects of mustard gas on leucopoietic tissues (*6–8*) and on the growth of experimental tumors (*9*) received some attention, biological research on chemical warfare agents remained relatively quiescent. With the advent of World War II research on war gases was resumed, and the newer knowledge and technics of a quarter of a century of trials have also been made of the possible value of the nitrogen mustards in the treatment of neoplasms, in particular those of lymphoid tissue.

The fact that agents classified as "confidential" were involved in the above studies has heretofore precluded the possibility of presenting the results in the open literature. This report reviews briefly the contributions which have focused attention on the cellular actions of the mustard compounds and gives a general description of their systemic effects as well as a preliminary statement of their possible clinical applications. Because of space limitations important contributions of many investigators will have to go unmentioned.

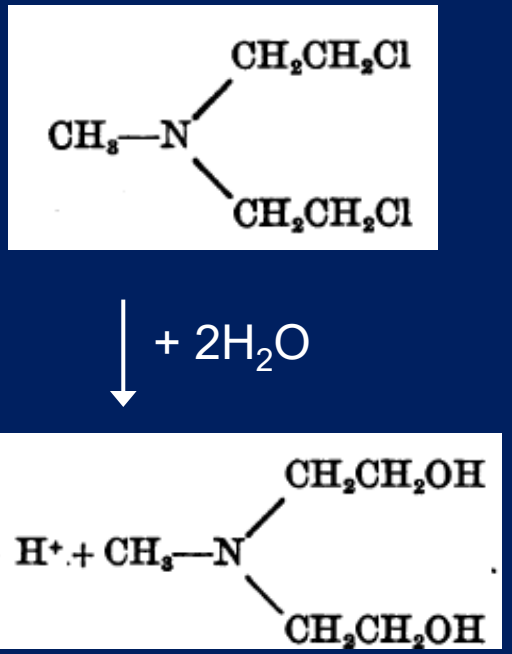


TABLE III

Results of Treatment with Nitrogen Mustard, 65 Cases

Diagnosis	Total No. of Cases	Nitrogen Mustard		Results			Patients Living		Patients Dead	
		No. of Doses	Total Dosage Mg.	Good	Fair	Poor	No.	Time since 1st Treat. Mos.	No.	Time from Rx to Death Mos.
				Per Cent						
Hodgkin's disease	28	4–40	24–193	61	18	21	15	1–26	12	¼–16
Lymphosarcoma	11	1–20	4–120	36	—	64	2	2–5	9	1.5–12
Leukemia, ch. myel.	8*	2–41	12–294	43	14	43	6	1–22	2	1–3
ch. lymph.	10	1–10	7–87	30	10	60	6	1–24	3	3–30 d.
acute	8	3–22	4–64	—	38	62	0		8	0.5–3.5

* Includes one case with inadequate follow-up.

Good result = patient returned to work or previous activities for some period of time
Fair result = “definite improvement but of a less striking or more short-lived nature”

Wintrobe, et al; Annals Int Med, 1947

Karnofsky Performance Status: First numeric outcome measure

TABLE 1

PERFORMANCE STATUS

Definition	%	*Criteria*
Able to carry on normal activity and to work. No special care is needed.	100	Normal; no complaints; no evidence of disease.
	90	Able to carry on normal activity; minor signs or symptoms of disease.
	80	Normal activity with effort; some signs or symptoms of disease.
Unable to work. Able to live at home, care for most personal needs. A varying amount of assistance is needed.	70	Cares for self. Unable to carry on normal activity or to do active work.
	60	Requires occasional assistance, but is able to care for most of his needs.
	50	Requires considerable assistance and frequent medical care.
Unable to care for self. Requires equivalent of institutional or hospital care. Disease may be progressing rapidly.	40	Disabled; requires special care and assistance.
	30	Severely disabled; hospitalization is indicated although death not imminent.
	20	Very sick; hospitalization necessary; active supportive treatment necessary.
	10	Moribund; fatal processes progressing rapidly.
	0	Dead.

Karnofsky et al. The use of nitrogen mustards in the palliative treatment of carcinoma. Cancer 1(4), 1948

RESULTS FROM THE TREATMENT OF BRONCHOGENIC CARCINOMA WITH HN_2

			Status before HN_2 therapy					*Therapy*			*Results of therapy*						
				Last x-ray treatment									*Response to HN_2, first course*				
Case no.	*Sex*	*Age*	*Prob. dur. dis. mos.*	*Mos. before HN_2*	*Response, dur. mos.*	*Gen. cond.*	*PS*	*Metastases*	*No. courses HN_2*	*Total dose HN_2 mg./Kg. (total mg.)*	*Comb. ther.*	*Max. leukopenia*	*Toxic[1] effects HN_2*	*SI*	*OI*	*PS*	*Dur., mos.*	*Follow-up, mos. after HN_2 started*
								Anaplastic Bronchogenic Carcinoma									
1[2]	M	49	1½	—	—	F	30	Mediastinum, marrow	1	0.5 (38)	—	1,800	None	F	o	60	1	Died; 2
2[3]	M	60	3	—	—	F	30	Neck, nodes	4	1.5 (119)	—	1,200	Anemia[2]	G	+	60	¾	Died; 2½
3	M	64	4	—	—	G	70	Neck, nodes	1[4]	0.7 (50)	—	7,000	None	O	o	70	—	Died; 3
4	M	69	6	—	—	P	60	Neck, skull	1	0.6 (42)	—	3,800	None	F	o	60	½	Died; ½
5[3]	M	63	12	2	G—2	P	40	Nodes	1	0.8 (35)	—	2,700	Anemia, bleeding	F	++	50	1	Died; 3
6	M	45	5	—	—	G	70	Rib, femur	3	1.5 (102)	—	1,700	None	F	+	90	1	Died; 8
7	M	47	12	6	G—6	G	60	Nodes, mediastinum	1	0.6 (53)	—	5,000	None	G	++	70	½	Died; 3
8	M	46	5	—	—	G	50	Mediastinum	1	0.8 (48)	X-rays	2,900	Anemia	O	o	50	—	Died; 3½
9	M	75	6	—	—	F	60	Nodes	1	0.5 (25)	X-rays	1,300	None	F	++	70	1½	Died; 2
10[3]	M	54	4	—	—	G	50	Neck, mediastinum	2	1.3 (100)	X-rays	1,300	Anemia, bleeding	G	++	80	2	Died; 5½
11	M	60	4	—	—	P	20	Skin	3	1.5 (84)	X-rays	1,900	None	G	++	70	2½	Died; 4½
12[3]	M	64	12	12	G—6	G	60	Nodes	3	1.6 (88)	X-rays	1,000	Anemia, bleeding	F	+	60	1¼	Died; 7
13[3]	M	49	1½	—	—	G	50	Nodes	1	0.6 (41)	X-rays	3,600	None	G	++	70	2½	Died; 4½
								Epidermoid Bronchogenic Carcinoma									
14[3]	M	65	2	—	—	F	50	Pleura, liver	1	0.4 (28)[5]	—	350	None	O	o	40	—	Died; 1¼
15[3]	M	65	4	—	—	G	40	Nodes	3	1.7 (83)	—	2,300	Anemia	G	++	70	8	Died; 18
16[3]	M	62	5	2	F—1	P	50	Liver	2	1.3 (84)	—	3,600	Anemia	F	+	70	1	Died; 3
17	M	45	15	5	P	G	50	None	1	0.8 (56)	—	300	Bleeding	F	o	60	1	Died; 4
18	M	50	18	—	—	G	70	Mediastinum	1	0.4 (20)	—	15,100	None	G	++	80	8	PS 30; 10
19	M	61	36	—	—	F	70	Ribs, pleura	1	0.6 (35)	—	2,200	None	O	o	60	—	Died; 2
20	F	60	25	3	P	F	60	Nodes	1	0.4 (24)	—	700	None	F	o	60	1	PS 30; 3
21[3]	M	64	24	5	F—2	F	50	Ribs, mediastinum	2	1.1 (82)	—	2,600	Anemia	F	o	70	1½	Died; 9

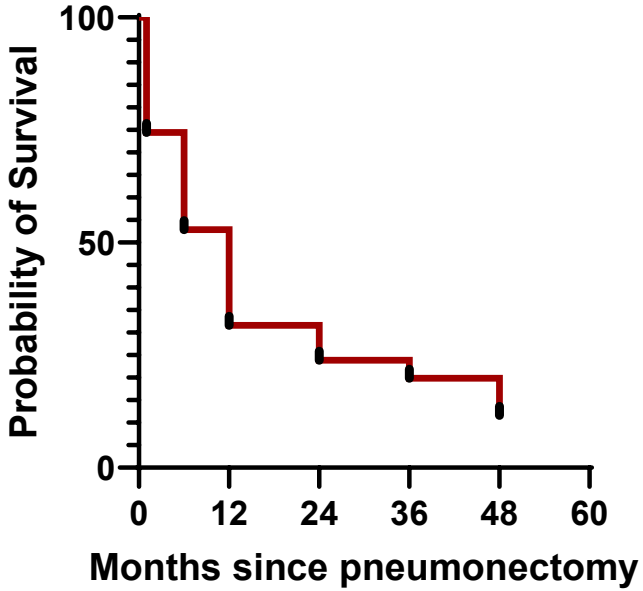
Karnofsky et al. THE USE OF THE NITROGEN MUSTARDS IN THEPALLIATIVE TREATMENT OF CARCINOMA, with Particular Reference to Bronchogenic Carcinoma. Cancer 1(4), 1948

Survival of Lung cancer after surgery 1933-1956

TABLE 2.—*Survival in 176 Patients Undergoing Pneumonectomy for Carcinoma of Lung*

Yr. After Operation	Living		Dead	
	No.	%	No.	%
<1 mo.	0	...	45	...
1-6 mo.	1	...	38	...
6 mo.-1 yr.	2	...	23	...
1	4	...	14	...
2	1	...	12	...
3	1	...	6	...
4	2	42.1	2	93.3
5-10	7	26.9	8	5.3
10-15	3	11.7	1	0.7
15-20	2	7.6	1	0.7
20-23	3	11.7	...	...
Total	26	100.0	150	100.0

Overall survival of lung cancer patients



Reinhoff Jr, et al. JAMA 1958

The effect of measuring error on the results of therapeutic trials in advanced cancer

Moertel and Hanley, Cancer 38:388, 1976

- 16 oncologists each measured 12 simulated tumor masses placed on a mattress under a ½ or 1 ½ in. layer of foam rubber
- ≥50% decrease in product of bidimentional diameters could be differentiated reproducibly
- Repeat measurements on same “tumor” revealed progression in 19% of cases whether used 25% or 50% increase

Adapted from slide courtesy of Dr. Lawrence Schwartz

2008 RECIST 1.1

Decrease # target lesions

Change LN assessments

Other minor changes

"It is not intended that these RECIST guidelines play a role in …[continuation of therapy] decision making…"

1981 WHO criteria

Standardization of assessment

Response 50% ↓ in sum of bidimensional products

Progression was ↑25% or new tumors.

"This percentage should not necessarily be regarded as influencing management of the patient."

2000 RECIST 1.0

Standards of measurability

Unidimensional measurements (PR now 30%↓)

Extended to CT scans

# of tumors to be measured

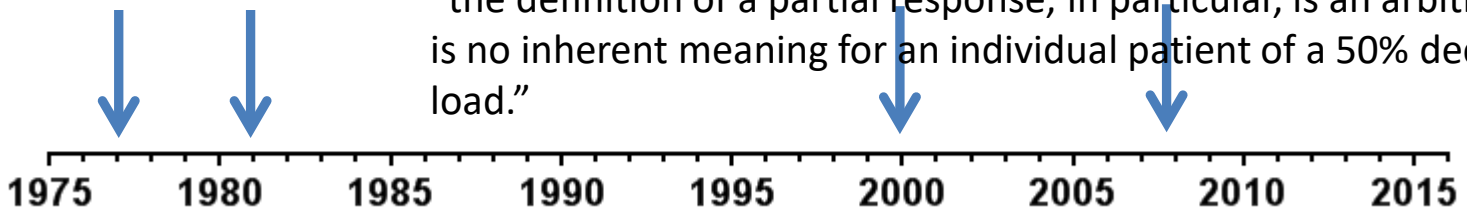
"the definition of a partial response, in particular, is an arbitrary convention—there is no inherent meaning for an individual patient of a 50% decrease in overall tumor load."

1977 Intl Union against Cancer

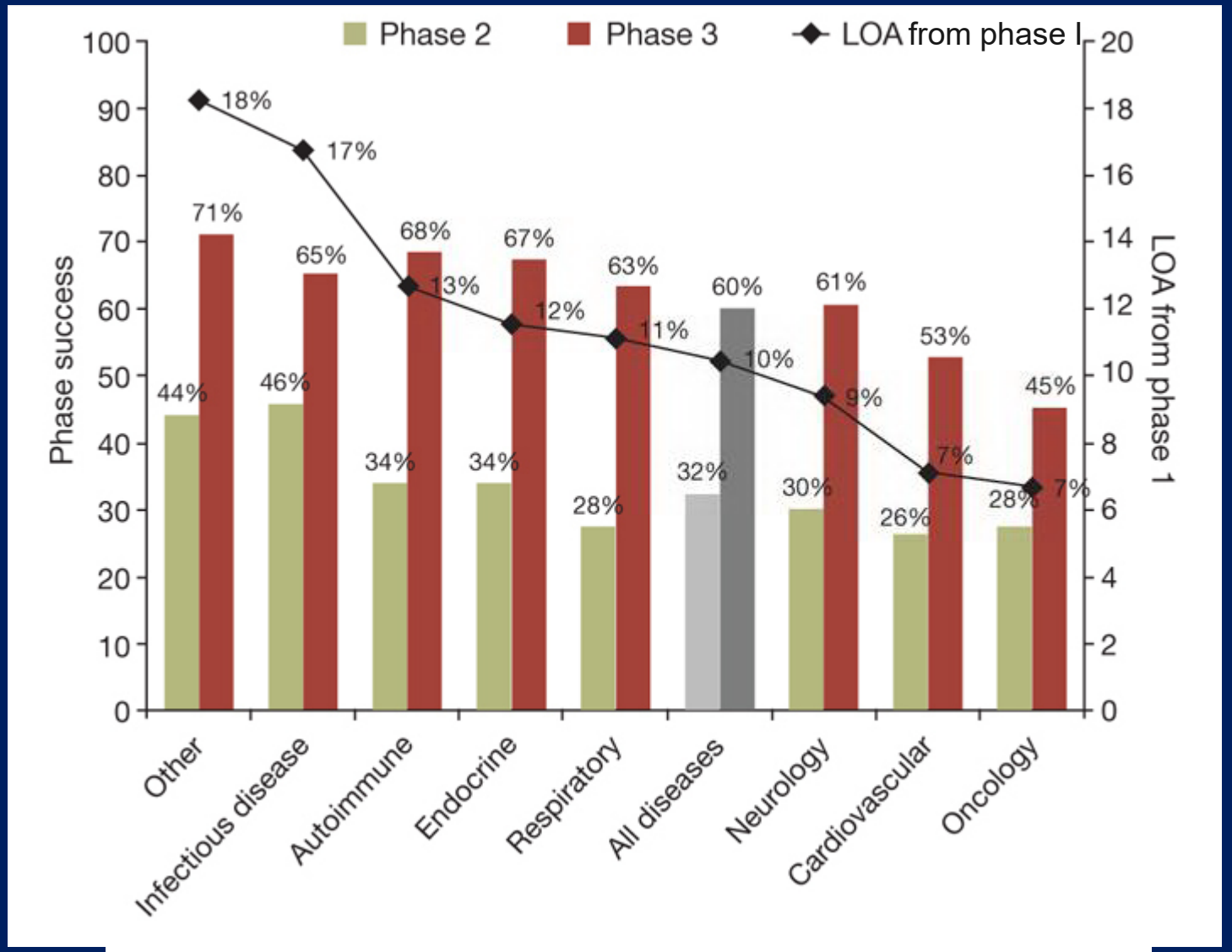
Tumors should be measured

Physical exam or X-rays

Response 50% ↓ sum of the products of the bidimensional diameters



Ted Williams: Batted 0.406 in 1941



From

Clinical development success rates for investigational drugs

Michael Hay, David W Thomas, John L Craighead, Celia Economides & Jesse Rosenthal

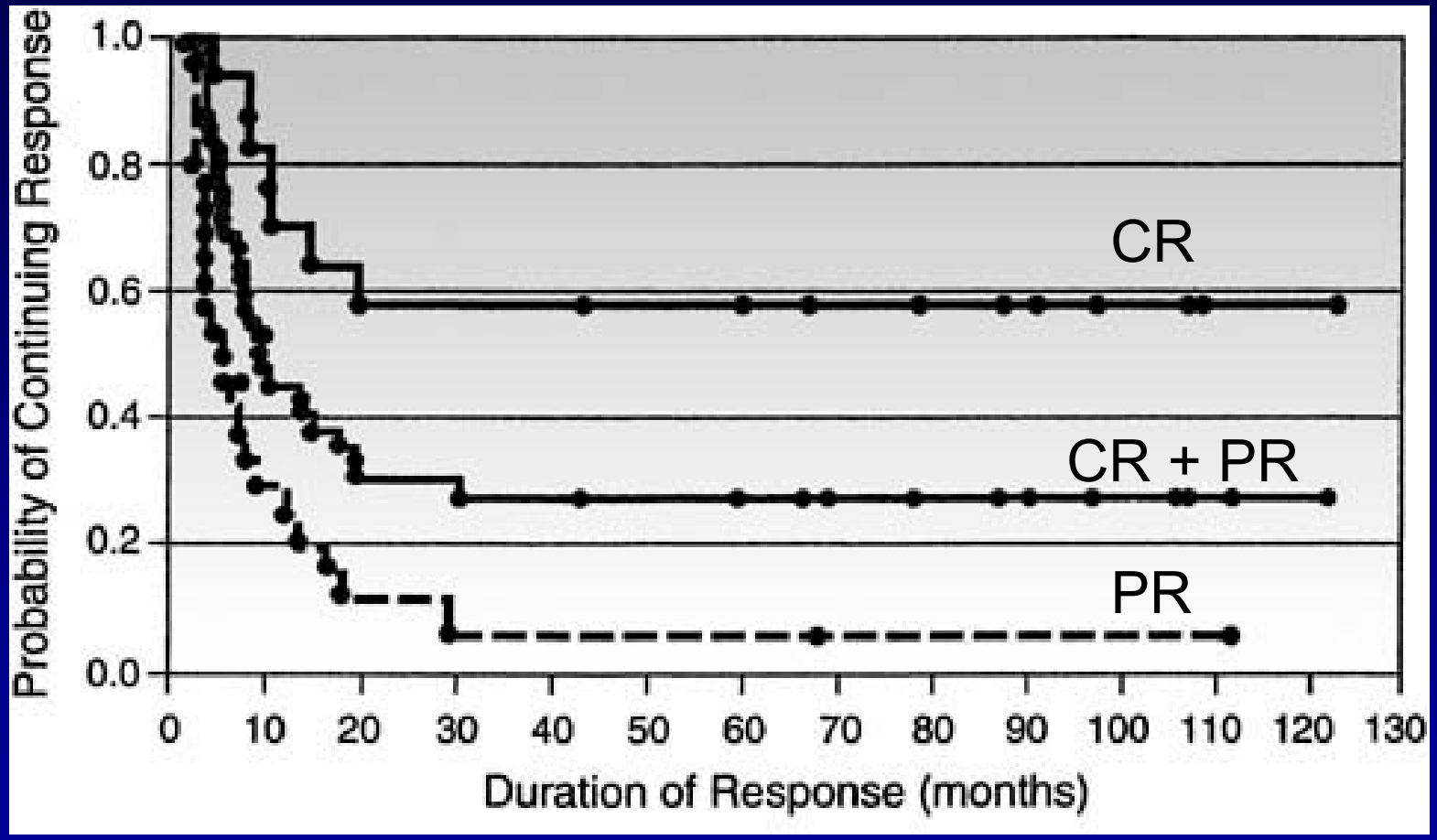
Nature Biotechnology **32**, 40–51 (2014) | doi:10.1038/nbt.2786

LOA= likelihood of approval

Measuring responses with modern therapy:

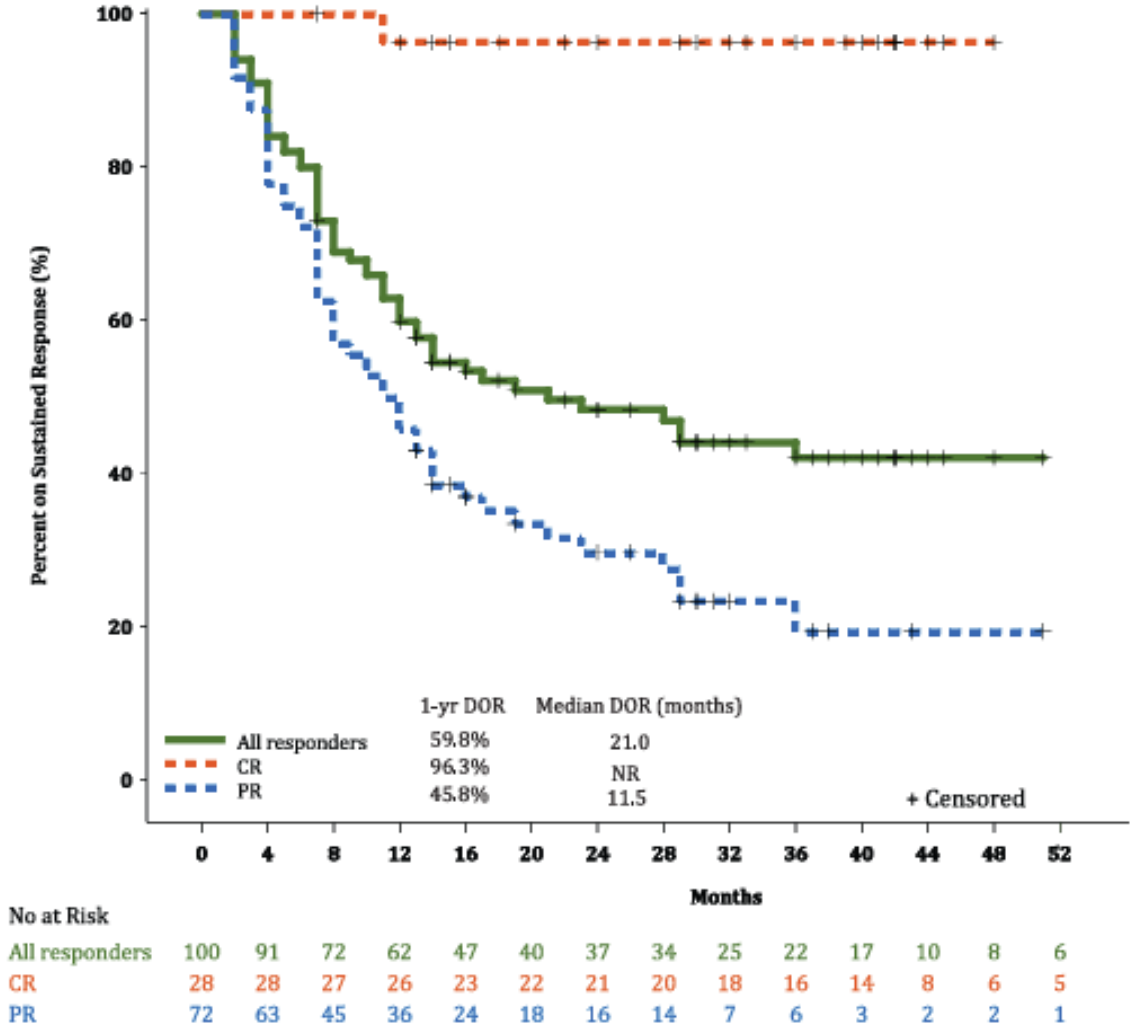
The case of melanoma and the value of a partial response

High dose IL-2 –Durability of responses



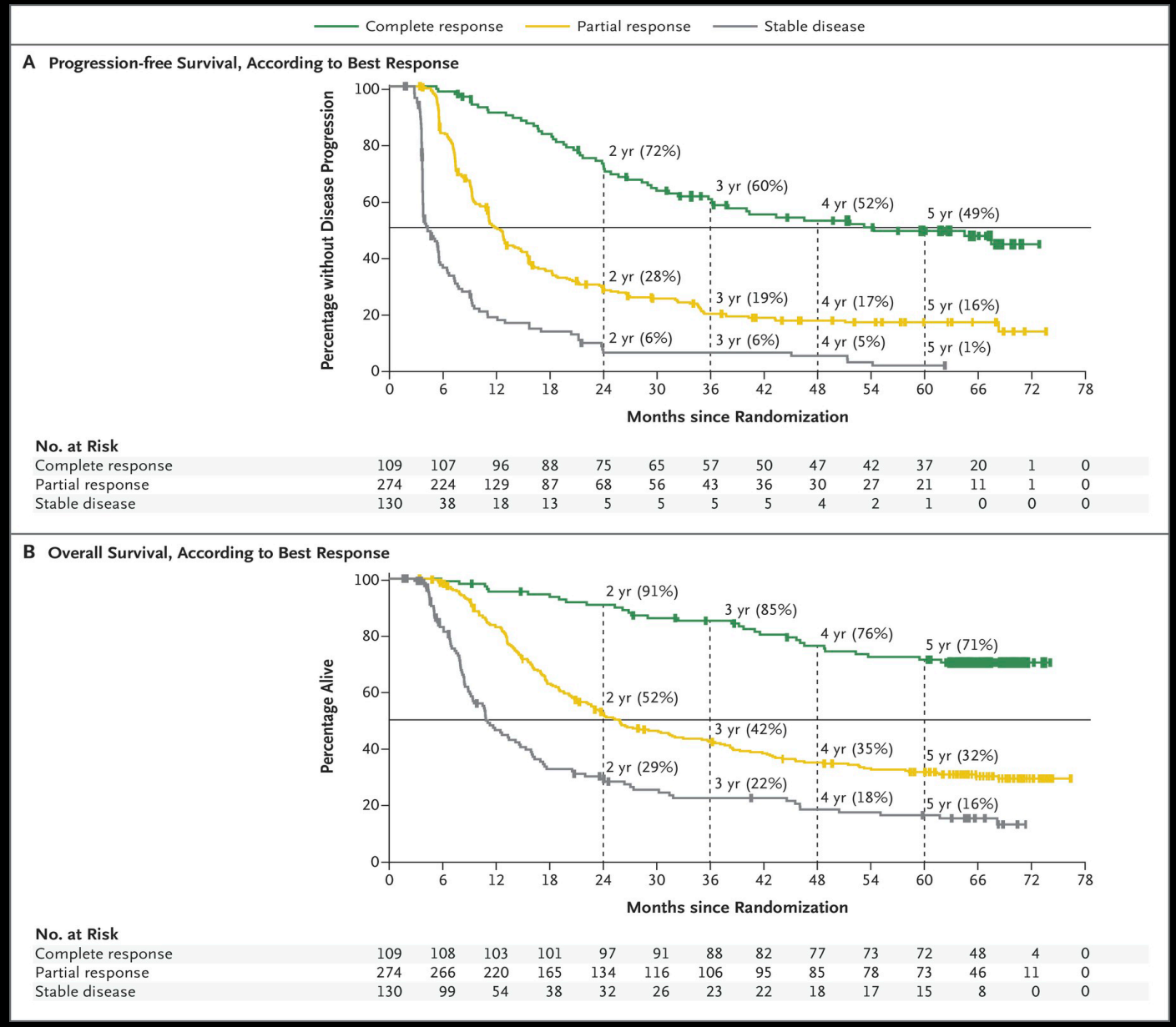
Atkins et al. JCO 17:2105, 1999

Duration of TIL responses: Meta-analysis through 2018



Dafni et al. Annals of Oncology 30: 1902–1913, 2019

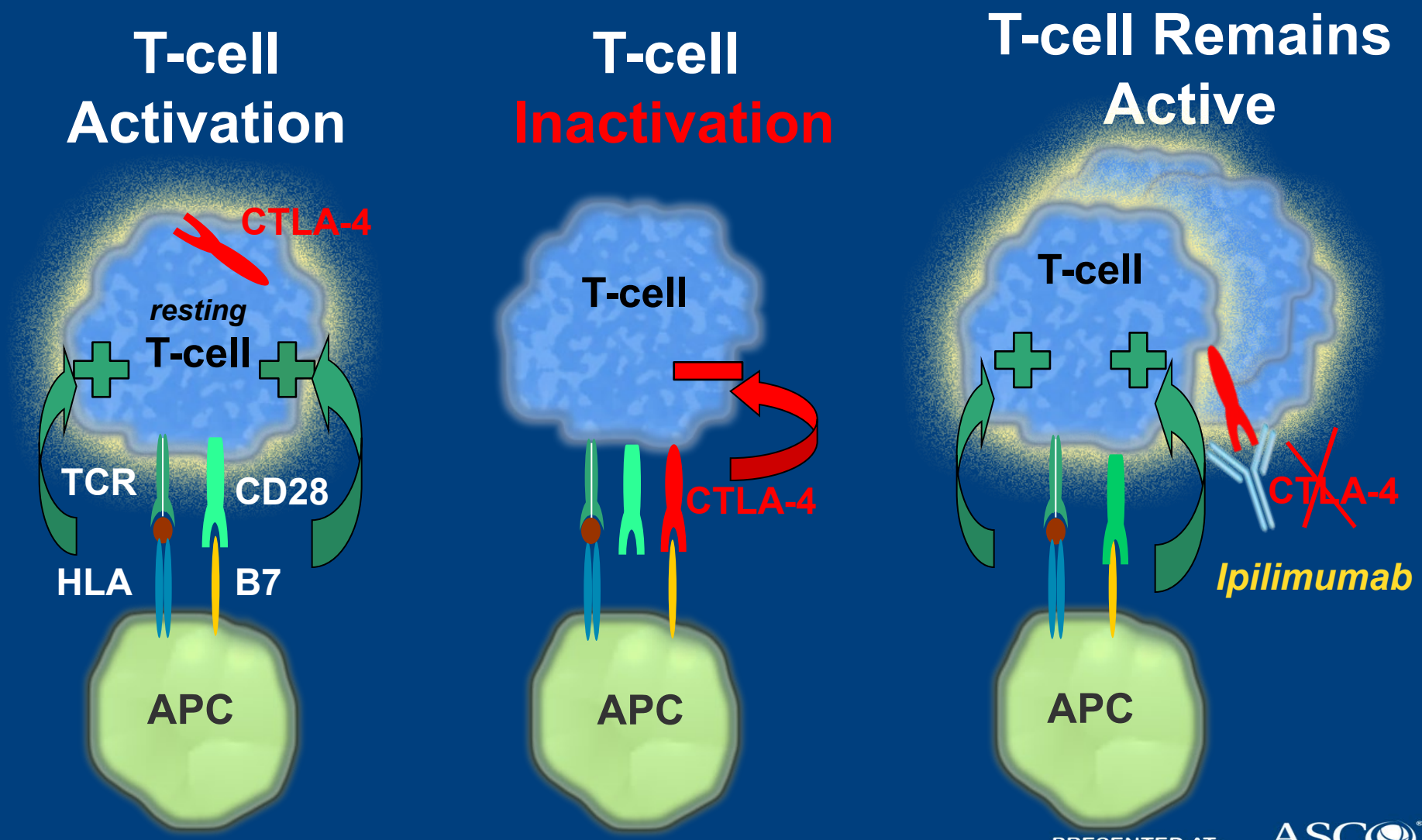
Dabrafenib + Trametinib: 5 yr outcomes in melanoma. PFS and OS According to the Best Confirmed RECIST Response.



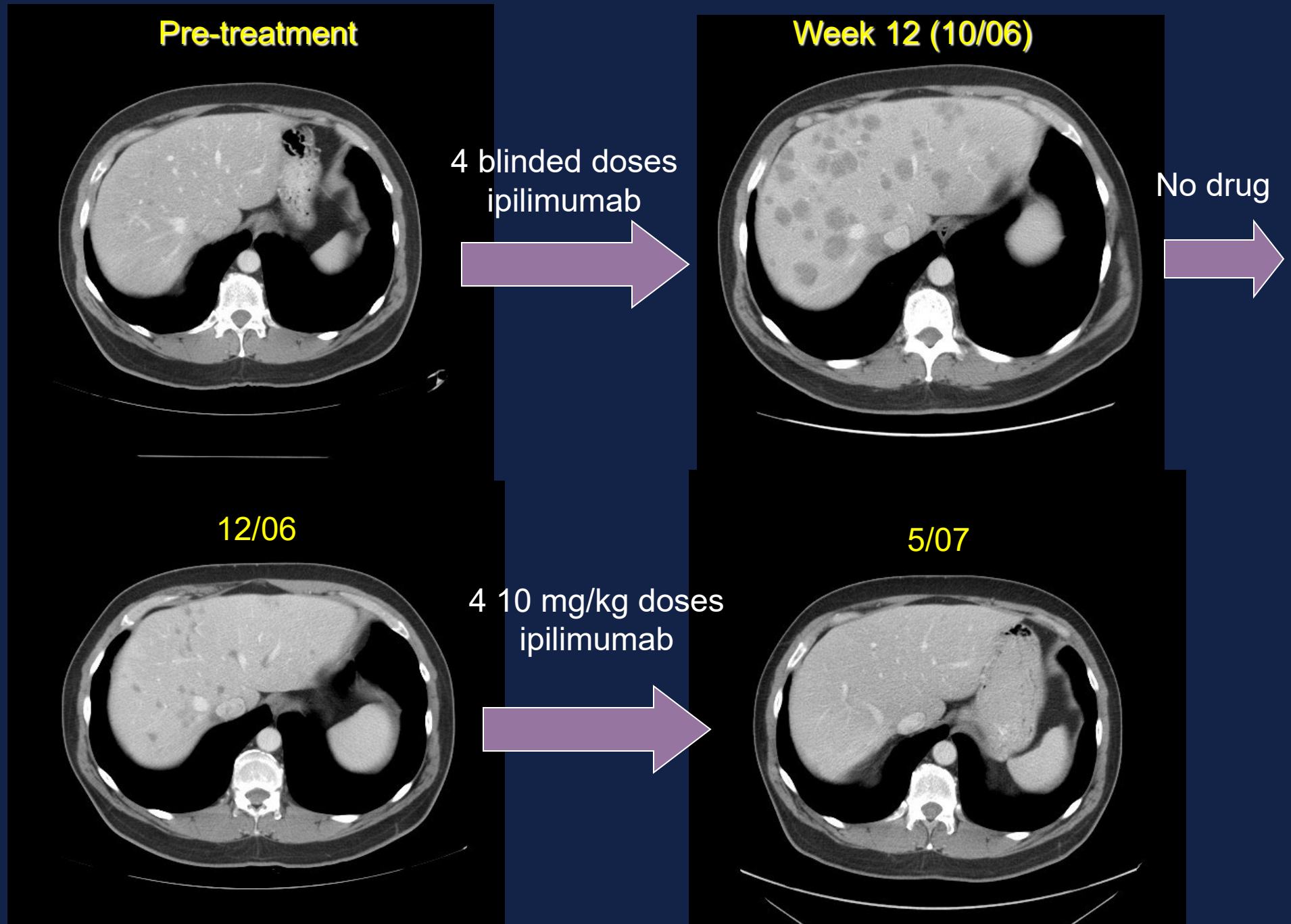
Robert C et al. N Engl J Med2019;381:626-636

The NEW ENGLAND JOURNAL of MEDICINE

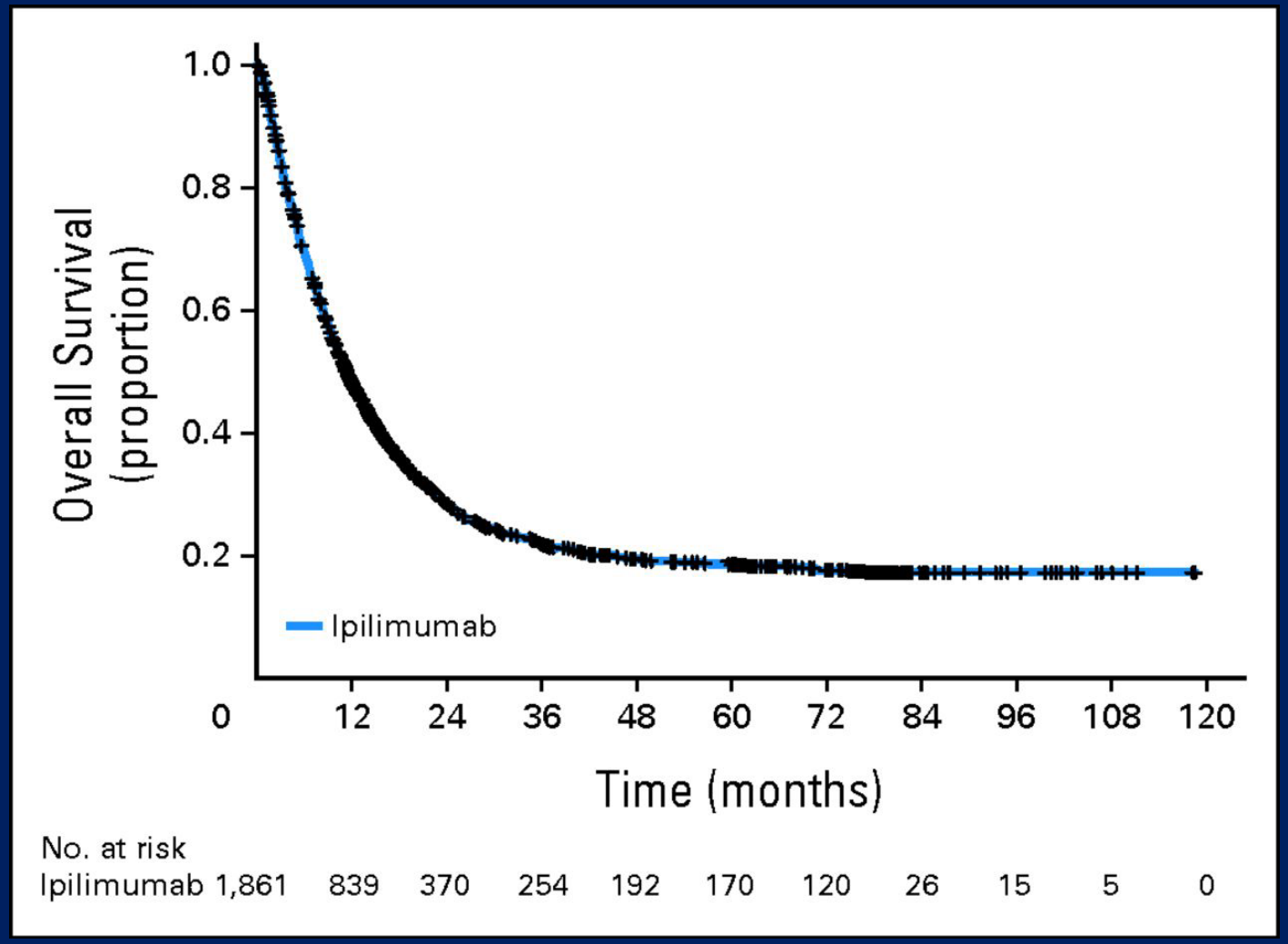
Checkpoint inhibitors: Ipilimumab



Korman, Peggs and Allison: Adv. In Immunol. 2006;90:297-339


Pre-treatment
4 blinded doses
ipilimumab
Week 12 (10/06)
No drug
12/06
4 10 mg/kg doses
ipilimumab
5/07

Ipilimumab: Pooled overall survival data.



Schadendorf D et al. JCO doi:10.1200/JCO.2014.56.2736

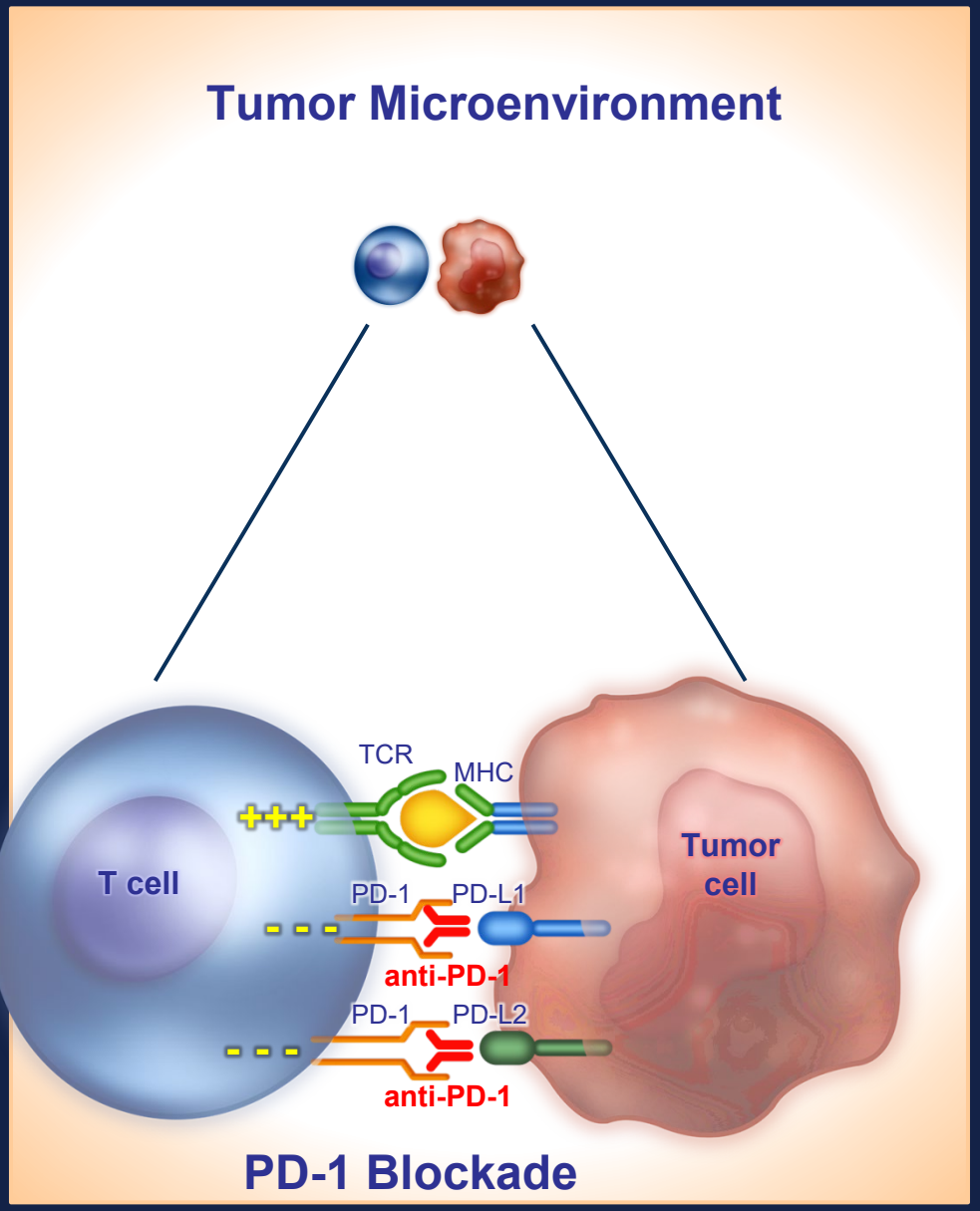
irRECIST

Table 1. Comparison of Key Differences in RECIST v1.1 and irRC

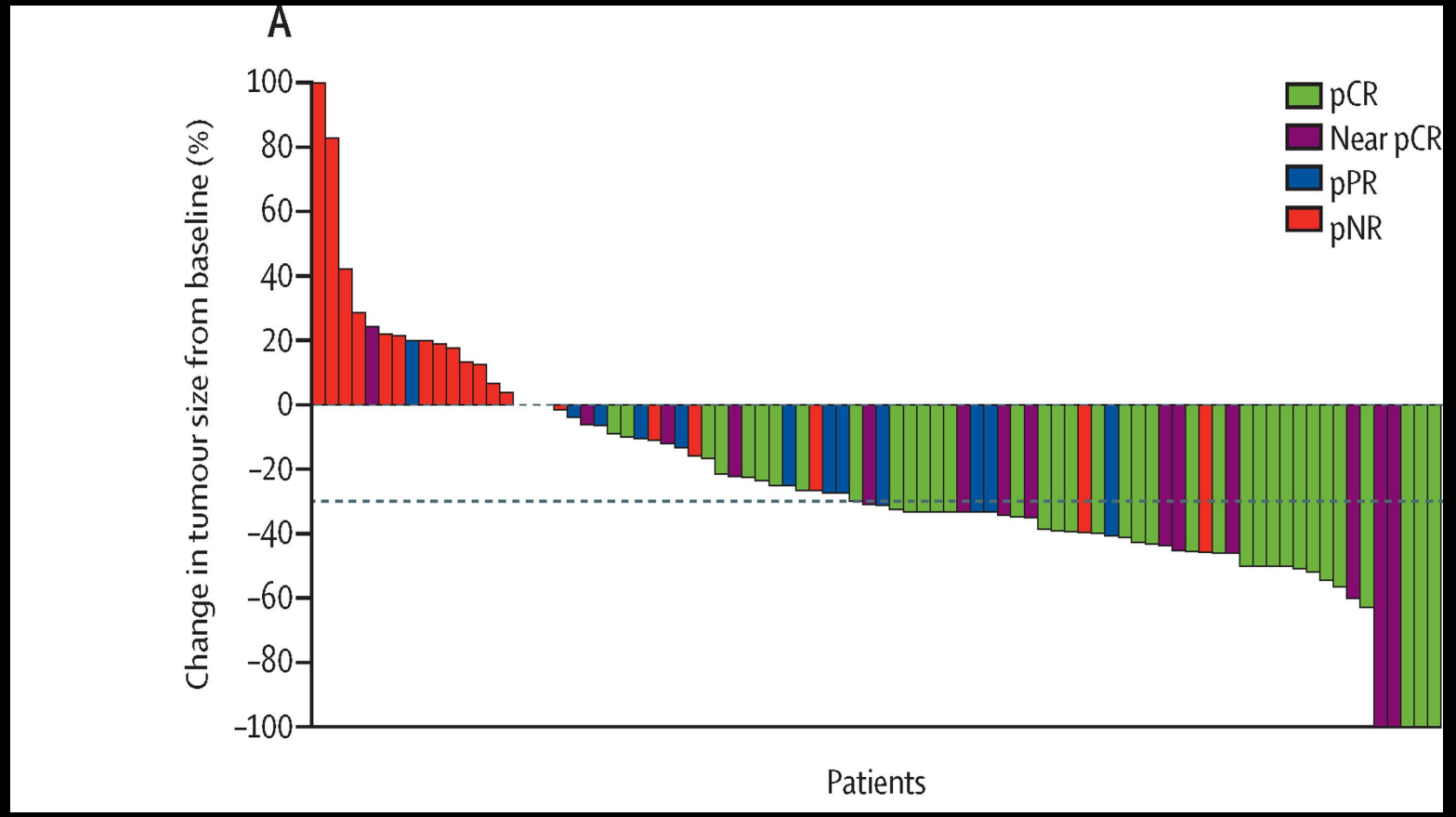
Category	RECIST v1.1	irRC
Measurement of tumor burden	Unidimensional	Bidimensional
Target lesions	Maximum, 5*	Maximum, 15 index lesions
New lesion	Results in progressive disease at first appearance	Up to 10 new visceral lesions and 5 cutaneous lesions may be added to the sum of the products of the two largest perpendicular diameters of all index lesions at any time point
Complete response	Disappearance of all target and nontarget lesions Nodes must regress to < 10 mm short axis No new lesions Confirmation required	
Partial response	≥ 30% decrease in tumor burden compared with baseline Confirmation required	≥ 50% decrease in tumor burden compared with baseline† Confirmation required
Progressive disease	≥ 20% + 5-mm absolute increase in tumor burden compared with nadir Appearance of new lesions or progression of nontarget lesions	≥ 25% increase in tumor burden compared with baseline, nadir, or reset baseline† New lesions added to tumor burden Confirmation required
Stable disease	Neither partial response nor progressive disease	

Hodi et al. *J Clin Oncol* **34**, 1510-1517(2016)

Checkpoint inhibitors: Anti-PD1 antibodies



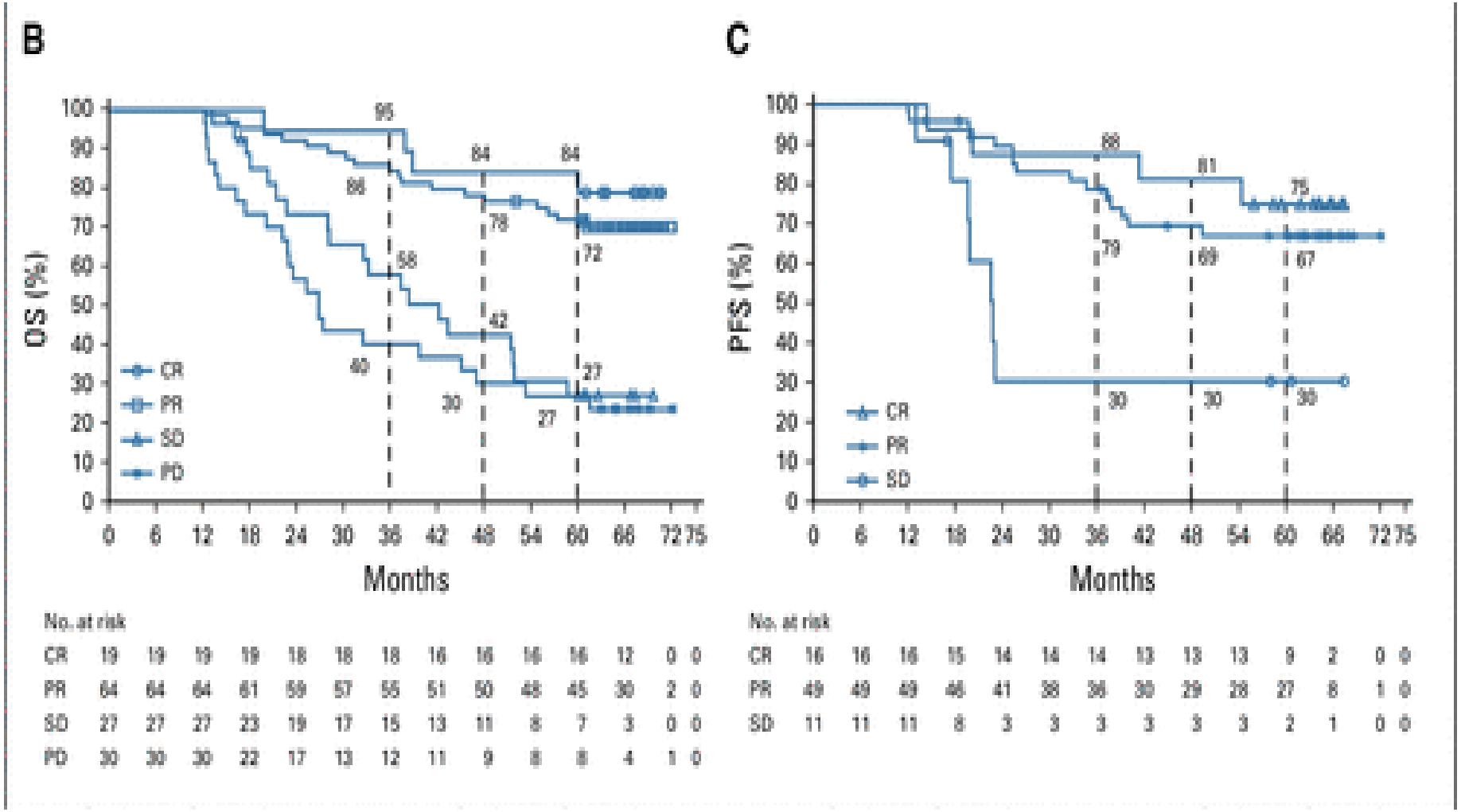
Opacin trial: Neoadjuvant Ipi/nivo in stage III melanoma



Rozeman et al.

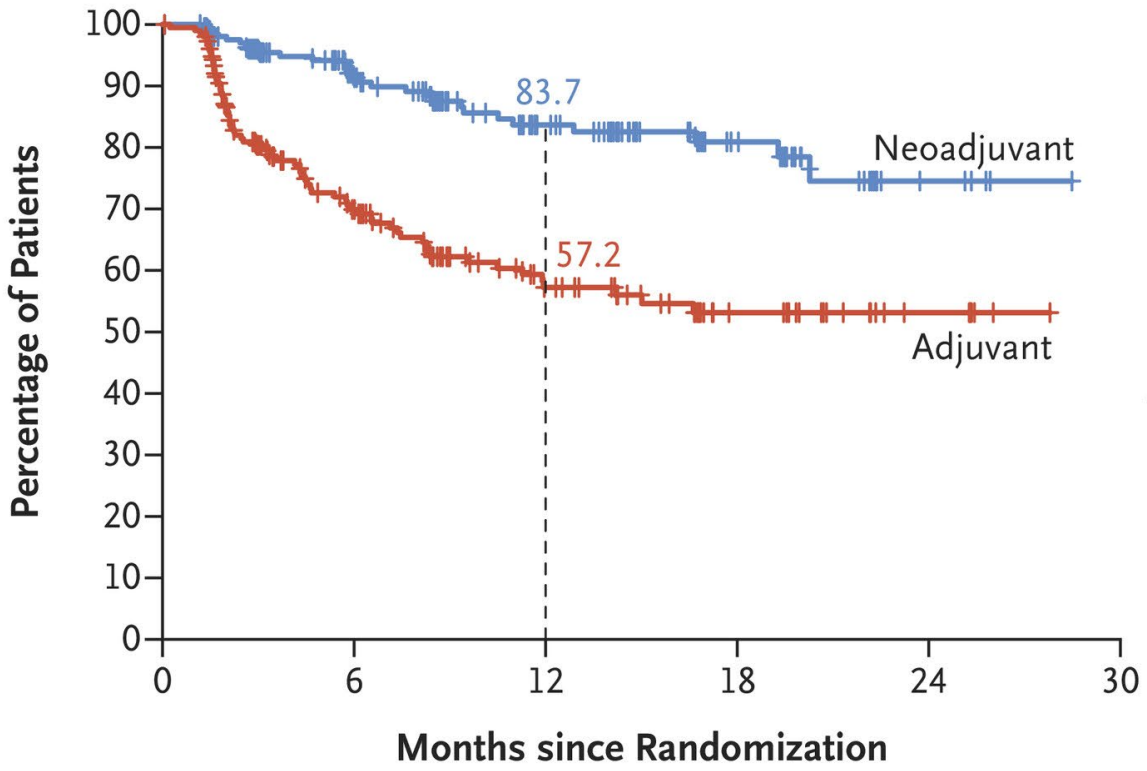
The Lancet Oncology 2019 20948-960DOI: (10.1016/S1470-2045(19)30151-2)

PRs have equivalent outcomes as CRs in melanoma patients treated with Nivolumab



Robert; et al. J Clin Oncol 2020 09-30

NADINA trial: Neoadj ipi/nivo vs. Adj nivo



No. of Events/ Total No. of Patients	
Neoadjuvant	28/212
Adjuvant	72/211

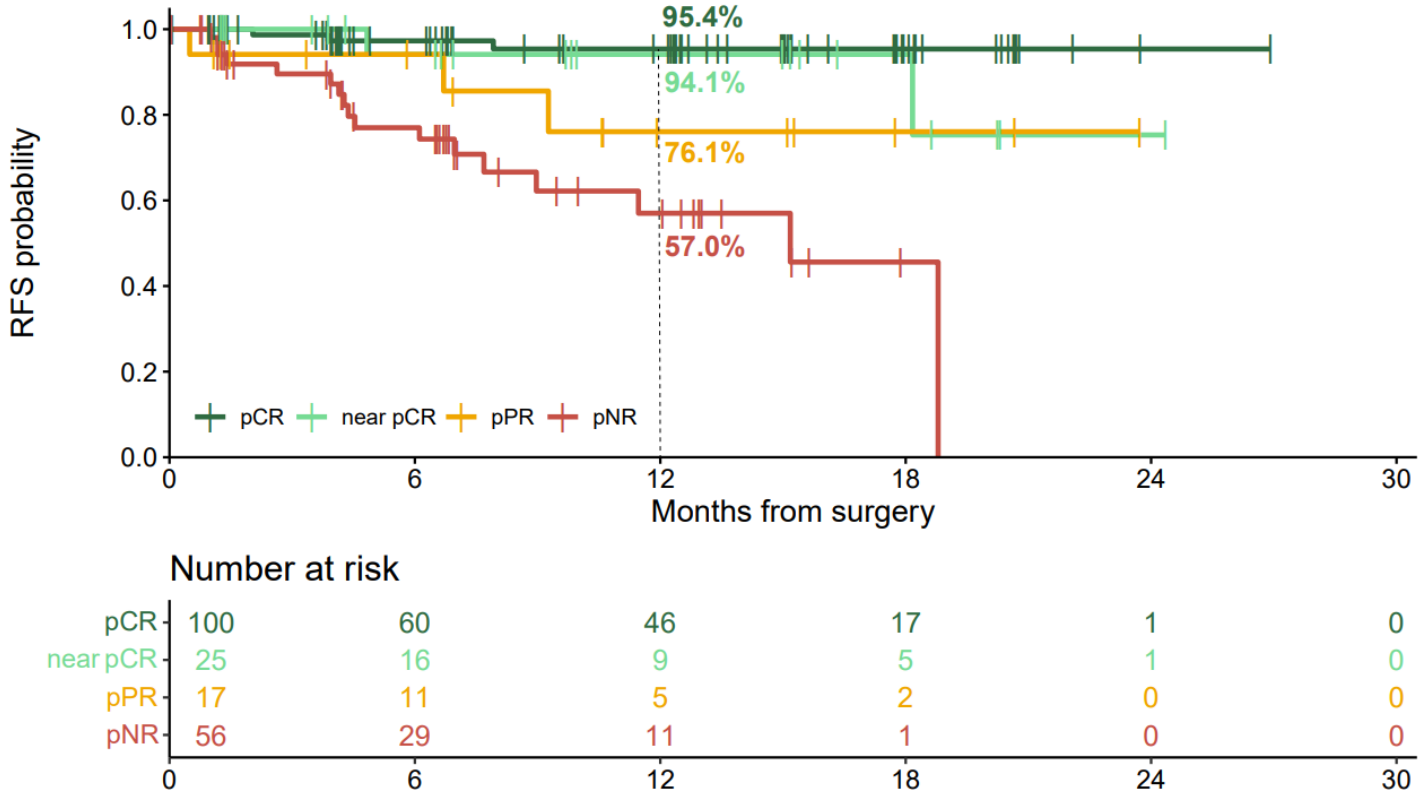
Adjusted difference in restricted mean survival time, 8.00 mo (99.9% CI, 4.94–11.05); P<0.001

Hazard ratio for progression, recurrence, or death, 0.32 (99.9% CI, 0.15–0.66)

No. at Risk (no. censored)					
Neoadjuvant	212 (0)	126 (71)	77 (111)	34 (152)	5 (179)
Adjuvant	211 (0)	100 (57)	53 (89)	23 (116)	6 (133)

Blank et al. NEJM 2024

NADINA – RFS According to Pathologic Response



Evolution of response evaluation

Performance status → WHO measurements → RECIST → irRECIST → path responses

Correlated with outcomes

Adverse event

Any undesirable experience associated with the use of a medical product in a patient.

Adverse event attribution. How likely is it related to study drug?

RELATIONSHIP	ATTRIBUTION	DESCRIPTION
Unrelated to investigational agent/intervention[1]	Unrelated	The AE *is clearly* ***NOT*** *related* to the intervention
	Unlikely	The AE *is* ***doubtfully related*** to the intervention
Related to investigational agent/intervention[1]	Possible	The AE ***may be related*** to the intervention
	Probable	The AE ***is likely related*** to the intervention
	Definite	The AE ***is clearly related*** to the intervention

CTCAE v5.0 grading of AST elevation

Grade	Description
1	<3x upper limit normal.
2	3-5x upper limit normal.
3	5-20x upper limit normal.
4	>20 x upper limit normal
5	

CTCAE v5.0 grading of pneumonitis

Grade	Description
1	Asymptomatic or mild symptoms; intervention not indicated
2	Moderate symptoms limiting age-appropriate ADLs
3	Severe symptoms limiting self care ADL
4	Life-threatening
5	Death

CTCAE v5.0 grading of motor neuropathy

Grade	Description
1	Asymptomatic or mild symptoms; intervention not indicated
2	Moderate symptoms limiting age-appropriate ADLs
3	Severe symptoms limiting self care ADL
4	Life-threatening
5	Death

Toxicities on clinical trials

- Attribution
- CTCAE grade: Data management not clinical care
- Take care of the patient