Outcome Measures (Endpoints) for Imaging Guided Intervention Trials

Michael C. Soulen, MD FSIR FCIRSE
Professor of Radiology and Surgery
Director Interventional Oncology
Abramson Cancer Center
University of Pennsylvania

1

- **Disclosures:**
  - Research grants: Guerbet, Pfizer, Sirtex
  - Consultant: Guerbet, Genentech, AstraZeneca
- **Learning Objectives**
  - To outline the range of possible outcome measures for interventional oncology trials
  - To describe methods to handle staged and repeatable therapies
  - To analyze the limitations of response assessment

2

IO Trials: Goals & Outcomes

GOAL	OUTCOME
• Cure	• Survival (OS, CSS)
• Prevention of progression as the cause of death	• Survival (PFS, DFS, OS)
• Surrogate Endpoint: disease control (CR+PR+SD)	• PFS; TTP
• Surrogate Endpoint: response (CR+PR)	• Response
• Palliation	• Pain, function, QOL

3

Survival

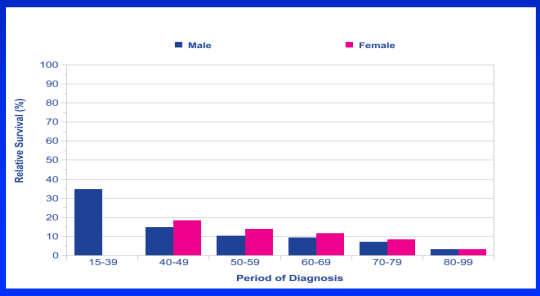
- **It's what patients care about**
- **Essential component of informed consent**
- **The gold standard for Phase III clinical trials**
- **FDA requires it**
- **Longest time (5+ years)**
- **Largest sample size (hundreds)**
- **Most expensive to measure (tens of millions$)**

4

What is Survival? - not so simple!

- Overall Survival
  - Alive or Dead
  - Appropriate for aggressive diseases where death from cancer is the expected outcome (lung, pancreatic)

Lung Cancer 5-year Survival



http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-239726

Penn IR

5

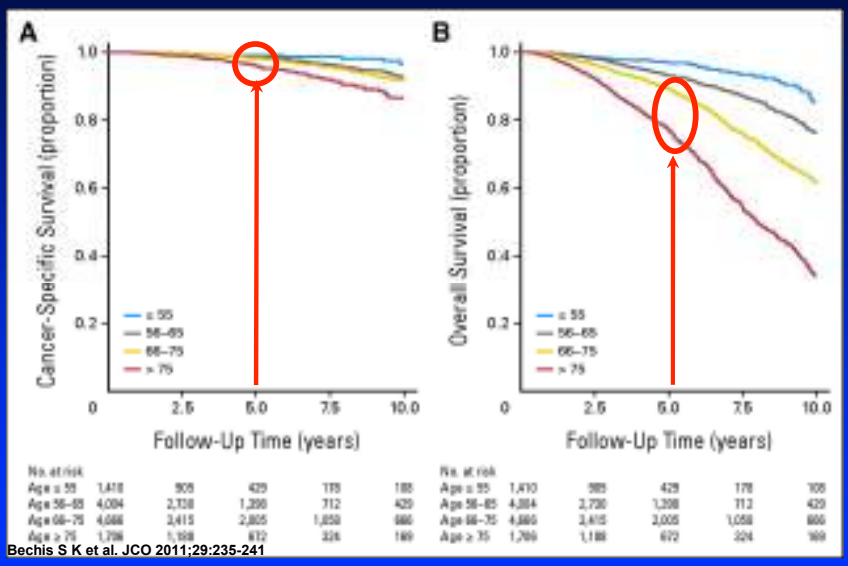
What is Survival? - not so simple!

- Disease-Free Survival
  - Alive without cancer
  - Appropriate for diseases with prolonged remission after curative therapy (breast, RCC, NET)
  - death from cancer still predominates, but OS not a practical primary endpoint
- Cancer-Specific Survival
  - Death from cancer
  - Appropriate for diseases where non-cancer related death predominates (prostate, T1 renal cell)
  - Effect measured by Cumulative Incidence Estimate rather than Kaplan-Meier, compare with Gray's Test instead of log rank test

Penn IR

6

Prostate Cancer: cancer-specific and overall survival by age stratum.



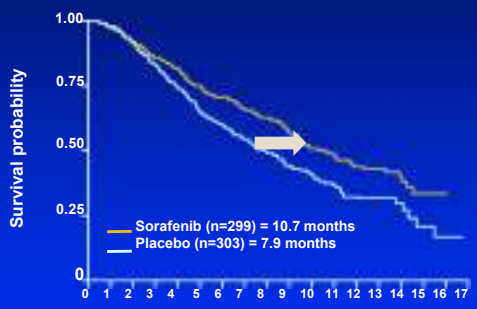
Bechis S K et al. JCO 2011;29:235-241

JOURNAL OF CLINICAL ONCOLOGY ASCO

7

How do we measure survival benefit?

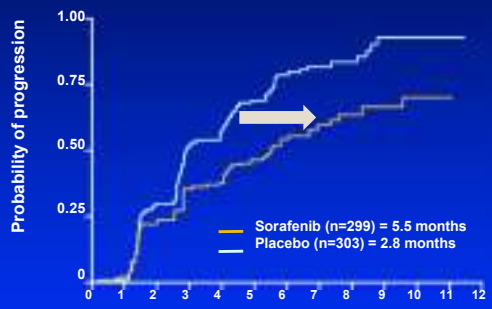
Overall survival



HR = 0.69

(95% CI: 0.55 – 0.87; p < 0.001)

Time-to-progression



HR = 0.58

(95% CI: 0.45 – 0.74; p < 0.001)

HR = hazard ratio
CI = confidence interval

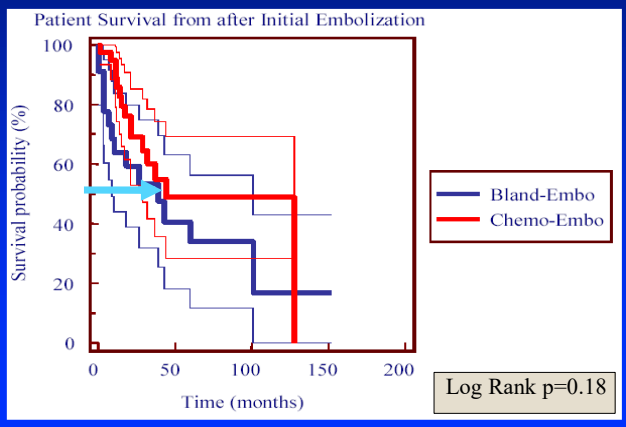
Llovet JM, et al. N Engl J Med 2008;359:378–90

Penn IR

8

Pitfall: Hazard Ratio vs. Median Survival

- No one has the median survival - could be months less but years more
- Medians mask subpopulation effects
- Should use Hazard Ratios

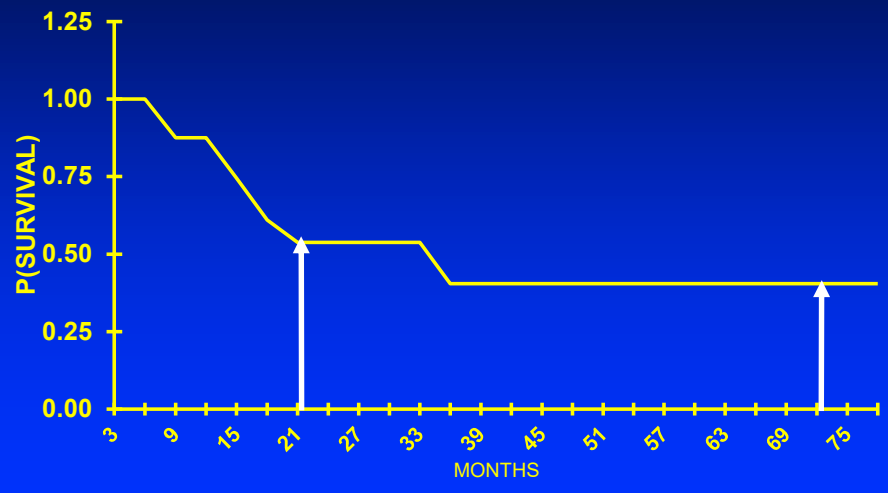


Ruutiainen AT JVIR 2007;18:847-55

Penn IR

9

Kaplan Meier survival curve for chemoembolization of sarcoma metastases



Penn IR

10

So what have we learned?

- "Survival" is usually the most important and rigorous primary outcome
- Be sure you know which type of "survival" matters
- Measure with Kaplan-Meier estimates
- Compare with Hazard Ratios

- but it's not that simple....

11

IO Trial Design Conundrums

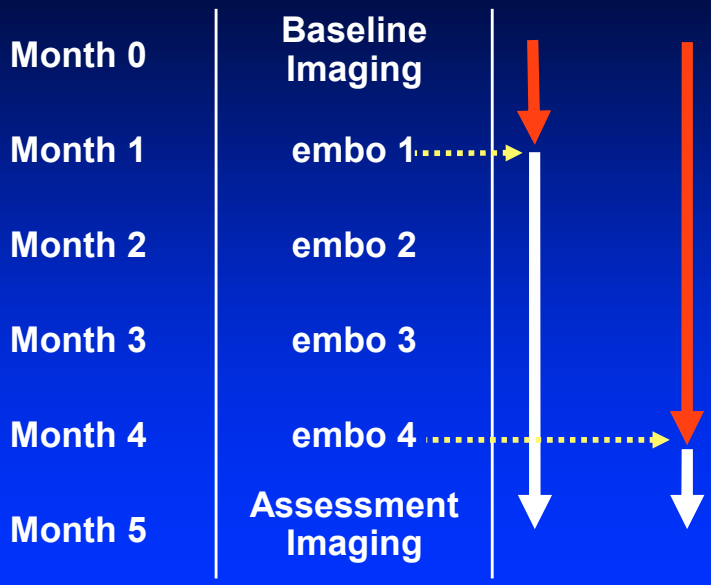
Time-based Outcomes (OS, PFS)

- for systemic therapy trials, clock starts with initiation of the drug, ends with progression of disease anywhere or death

- IO therapies may be staged: not all tumor treated at same time
- IO therapies may be repeatable: first progression may not signal failure of therapy

12

Staged Therapies

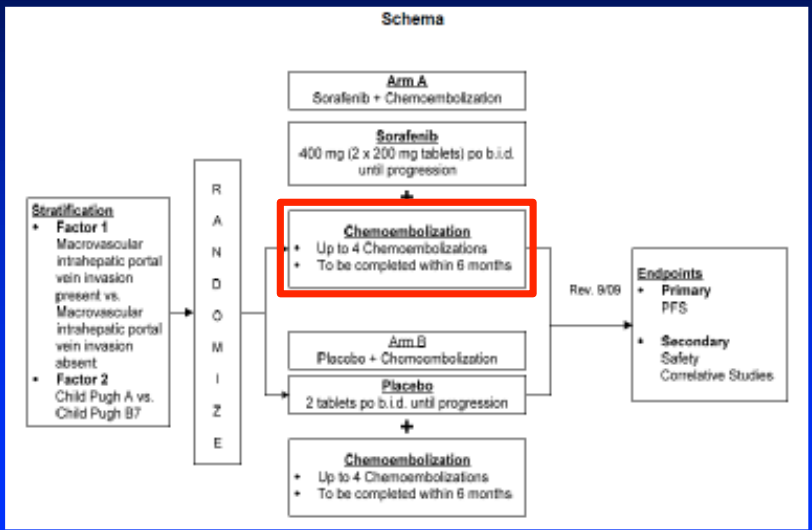


Penn IR

13

Trial Design for Staged Therapies

ECOG 1208: TACE +/- sorafenib



Penn IR

14

Trial Design for Staged Therapies

ECOG 1208: TACE +/- sorafenib
Scheduled Imaging

5.3.7 Tumor imaging studies for the purposes of determining PFS are to be performed as follows:

- Baseline (chest CT, Abdomen/pelvis CT or MRI)
- 4 months after first chemoembolization (chest CT, Abdomen/pelvis CT or MRI)
- 8 months after first chemoembolization (chest CT, Abdomen/pelvis CT or MRI)
- Every 2 months beginning at 10 months post baseline

Penn IR

15

SPACE: Scheduled Therapy & Imaging

- First DEB-TACE performed 3-7 days after start of sorafenib or placebo
- Subsequent DEB-TACE performed on day 1 (±4 days) of cycles 3, 7, and 13, and every 6 cycles thereafter
- Patients allowed optional DEB-TACE sessions between cycles 7-13 and 13-19

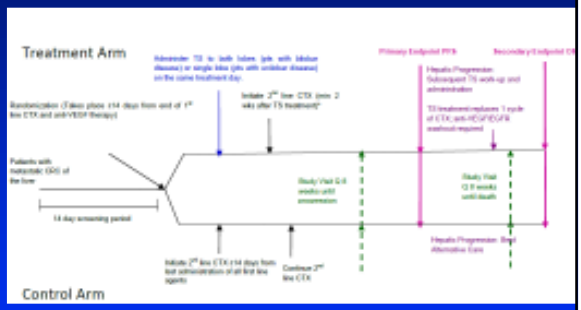
Lencioni R et al. ASCO GI 2012

Penn IR

16

EPOCH: single-session whole-liver therapy

- Colorectal metastases progressing on 1st-line systemic chemotherapy
- Randomized to 2nd-line chemotherapy +/- Y90 radioembolization
- required whole-liver radioembolization



Penn IR

17

Surrogate Measures of Survival

- Time-to-Progression (TTP) or Progression-Free Survival (PFS)
  - proposed endpoint for Phase II trials
  - faster and cheaper than survival
- Limitations:
  - Imaging-based outcome (not clinical)
  - doesn't necessarily correlate with survival
  - problematic to measure with staged therapies
  - relevance if therapy is repeatable (ablation, embo)?
  - can be surprisingly hard to measure (RADIANT-2)

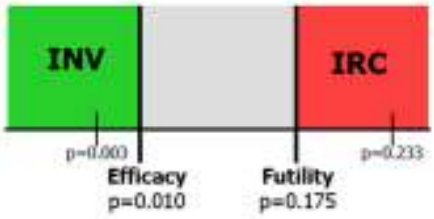
Penn IR

18

FDA Advisory Committee evirolimus vs. placebo for carcinoid

Discordant Results

Carcinoid Trial (C2325) 2nd interim analysis demonstrated discordance of PFS between INV and IRC

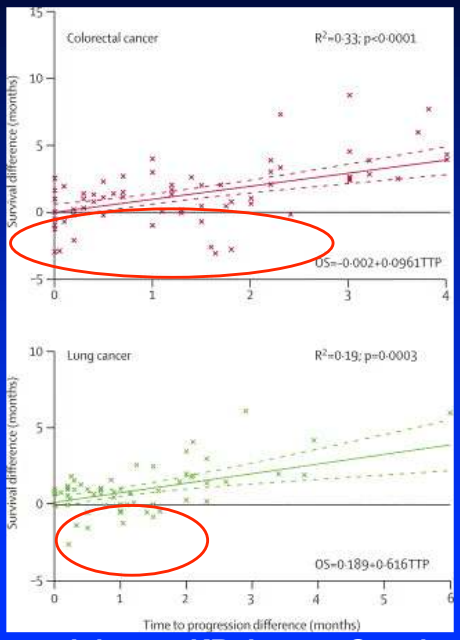


- **IRC** PFS (p=0.233) crossed the **Futility Boundary** (p = 0.175)
- **INV** PFS (p=0.003) crossed the **Efficacy Boundary** (p = 0.010)

Penn IR

19

TTP Limitations



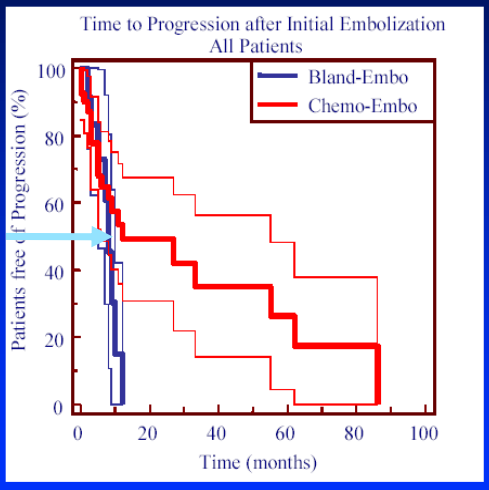
- 337 randomized trials. Johnson KR, Lancet Oncology 2006

Penn IR

20

TTP Hazard Ratio vs. Median

Same issues as for survival



Ruutiainen AT JVIR 2007;18:847-55

Penn IR

21

Repeatable Therapies

TIME	PATIENT 1	PATIENT 2
0	A0 ablation	A0 ablation
6 months	(-) recurrence	(+) recurrence A0 ablation
12 months	(-) recurrence	(-) recurrence
PFS	12 months	6 months
DFS/CSS/OS	12 months	12 months

Penn IR

22

Trial Design for Repeatable Therapy: Landmark Analysis

MULTICENTER FEASIBILITY STUDY OF PERCUTANEOUS RADIOFREQUENCY ABLATION OF HEPATOCELLULAR CARCINOMA IN CIRRHOTIC PATIENTS

SCHEMA

ELIGIBILITY	REGISTRATION	TREATMENT	FOLLOW UP
Cirrhosis 1-3 HCC ≤ 3.0 cm or A single tumor >3.0 cm but ≤ 5cm		Percutaneous RFA (Repeat RFA permitted for 15 months after initial ablation session)	CT scans within 1st week after initial RFA and every 3 months post ablation for 18 months*

SPECIFIC AIMS/OBJECTIVES

The primary aim for this trial is to estimate the proportion of patients undergoing solitary or repetitive percutaneous RFA treatment sessions whose livers have no identifiable tumor by CT scan at 18 months following initiation of therapy.

Since there is no direct comparison rate in this study, it was sized to detect a difference between a disappointing and promising rate. Based on the literature, the surgical success rate is approximately 65% @ 18 months.[25] It is believed that RFA could have a considerably better success rate of approximately 85%. A sample size of 35 participants will provide at least 85% power to detect the proposed alternative RFA success rate (assuming a null hypothesis of 65%

Penn IR

23

Trial Design for Repeatable Therapy

- Time-to-Treatment-Failure (TTTF)
- Time to Untreatable Progression (TTUP)
- Progression when test therapy can no longer be performed
- pre-determined objective criteria for untreatable progression
- technical (can't get there) vs. clinical (declining liver function or performance status) vs. progression
- who decides?
- are criteria reproducible?
- appealing concept but hard to do
- FDA does NOT accept this

Penn IR

24

25

Response: the 4 Essentials

- **Accuracy** -- what you measure is truth
- **Precision** -- you and everyone else get the same measurement, every time
- **Simple & Generalizable**-- anyone can do it in daily practice without limiting workflow
- **Useful** -- informs management decision or prognosis

Penn IR

26

Response Examples

- RECIST/WHO
- Necrosis ("EASL")
- mRECIST
- Lipiodol retention
- Functional imaging
  - PET, diffusion, etc.

27

RECIST 1.1
What Med Oncs Use

- Sum of single longest diameters of index tumors (2 per organ, up to 5)
- CR = gone
- PR = 30% reduction in sum of LD's
- Progression = 20% increase in sum of LD's or new tumors
- Stable = neither PR or progression

28

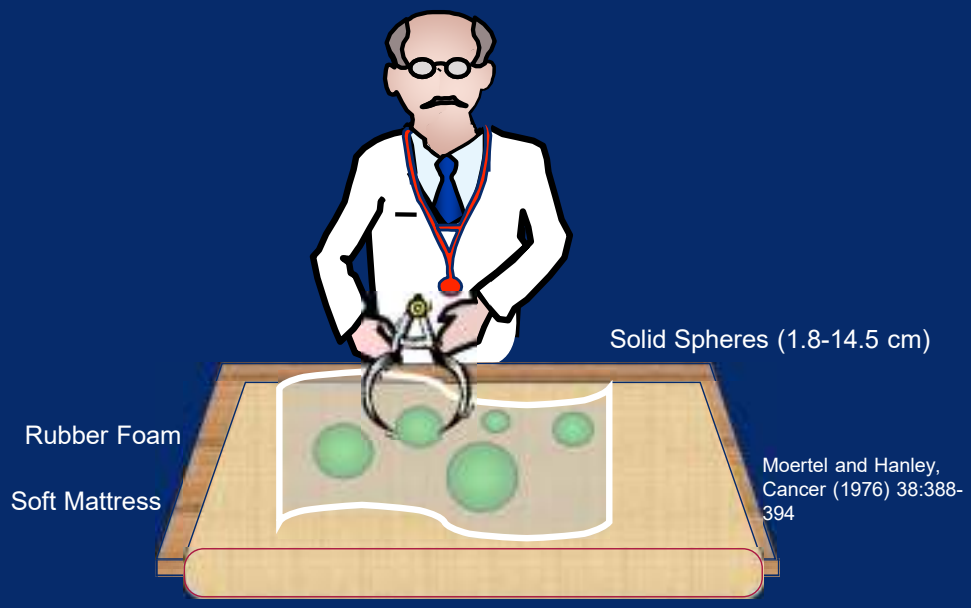
What is the basis for RECIST?

- -30% and +20% diameter change is equivalent to a 50% change in tumor volume
- -30% and +20% diameter changes statistically correlate with patient survival
- Single linear diameters are more precise and accurate than bidimensional (WHO) measurements
- A bunch of medical oncologists palpating a foam mattress

Penn IR

29

History of response criteria



30

History of response criteria

- Sixteen oncologists determined the diameter of 12 spheres (1.8-14.5 cm)
- The measured size of identical spheres differed
  - by at least 25% in 25% of the measurements
  - by at least 50% in 6.8% of the measurements

  („false-positive rate for response“)

Moertel and Hanley, Cancer (1976) 38:388-394

31

2 cm lung nodule
3 radiologists re-reading same scan and scan done 15 min later
% difference in measurement

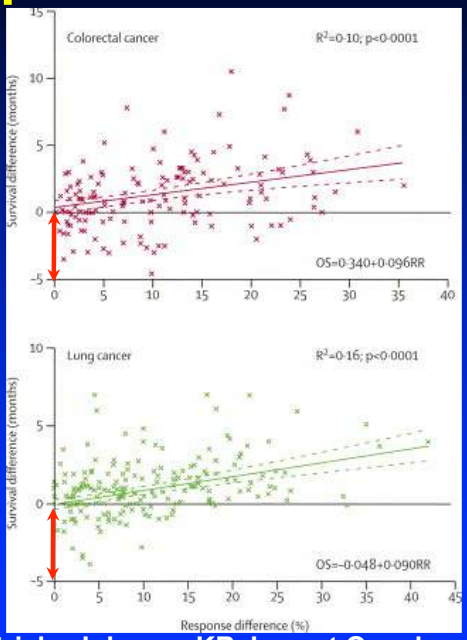
Measurement, Comparison, and Reader	Tumor Size on Scan 1	Example	
		±2 cSDs	Difference*
Unidimensional (cm)	2.00		
Scan 1 vs scan 2			
1		1.66, 2.40	−16.8, 20.1
2		1.59, 2.52	−20.6, 26.0
3		1.59, 2.51	−20.4, 25.6
Scan 1 repeat			
1		1.81, 2.21	−9.5, 10.5
2		1.65, 2.42	−17.4, 21.0
3		1.57, 2.55	−21.6, 27.5

20%-30% difference in measuring same lesion on same scan or concurrent scan by same reader

Zhao, Radiology 2009;252:263-272

32

Response vs. Survival

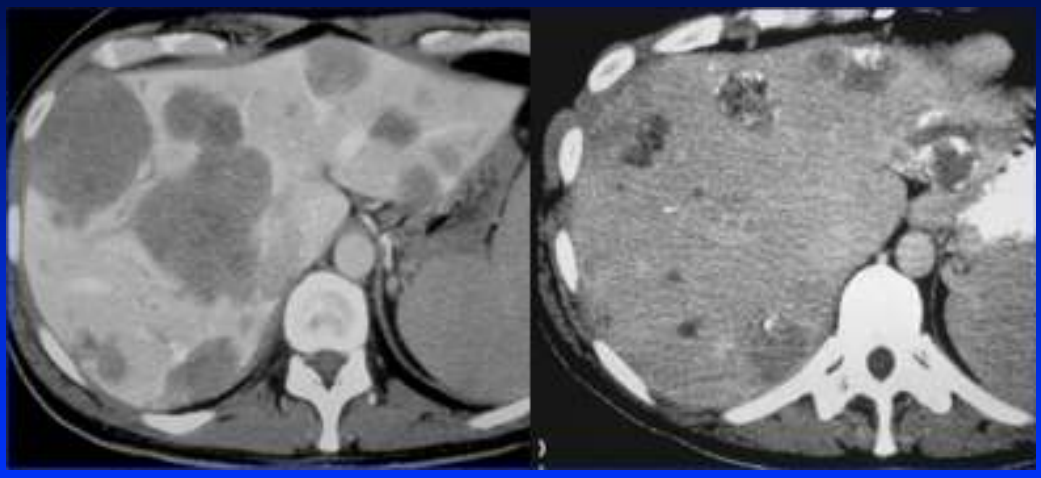


337 randomized trials. Johnson KR, Lancet Oncology 2006

Penn IR

33

Colon Cancer: chemoembolization

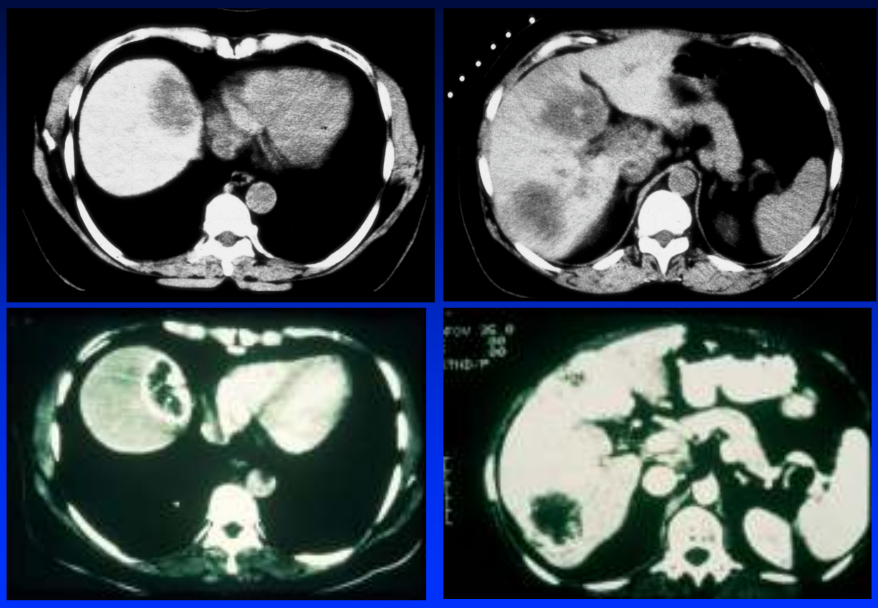


Partial response, died 6 months

Penn IR

34

Colon Cancer: chemoembolization



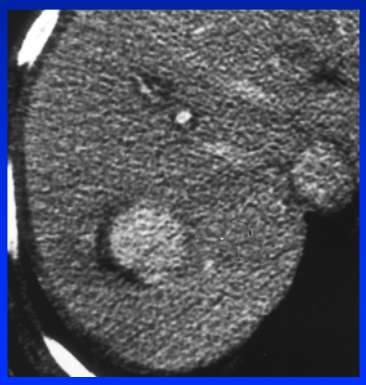
Stable disease at 1 year, CEA down 90%, ECOG 0

Penn IR

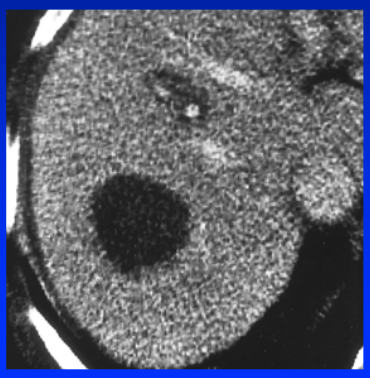
35

Post RF Ablation: It Grew!

- Contrast enhanced CT
  - Pre RFA



- Contrast enhanced CT
  - 6 month post RFA



Penn IR

36

The "EASL Criteria" Myth

- EASL 2000 Consensus Document
- "should take tumor necrosis into account"
- no criteria!!!

Bruix, J Hepatol 2001;35:421-30

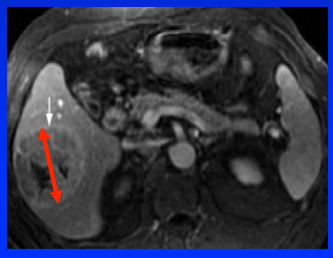
Penn IR

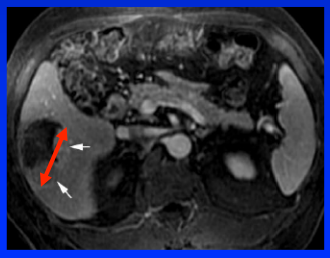
37

Necrosis Response Criteria

mRECIST

- Single longest diameter of enhancing (viable) tumor
- CR = none, PR 30% reduction from baseline, PD 20% increase from baseline or new lesions





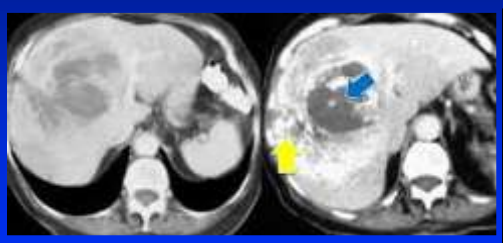
Lencioni, Semin Liver Dis 2010;30:52-60

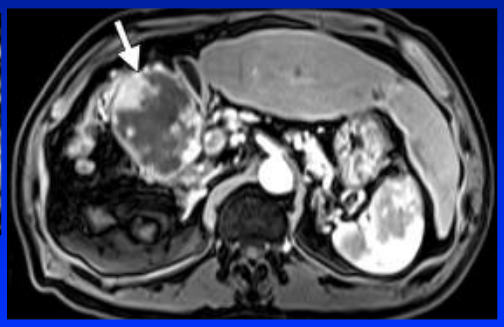
Penn IR

38

Chemoembolization of HCC: mRECIST reader confidence

60% of measurements of residual viable tumor diameter classified as "not confident"





Penn IR

39

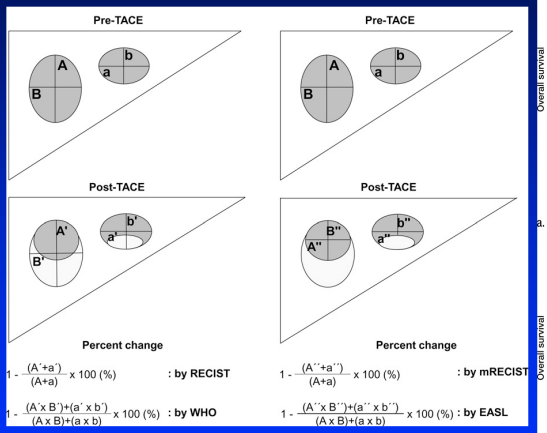
NWU Necrosis Criteria

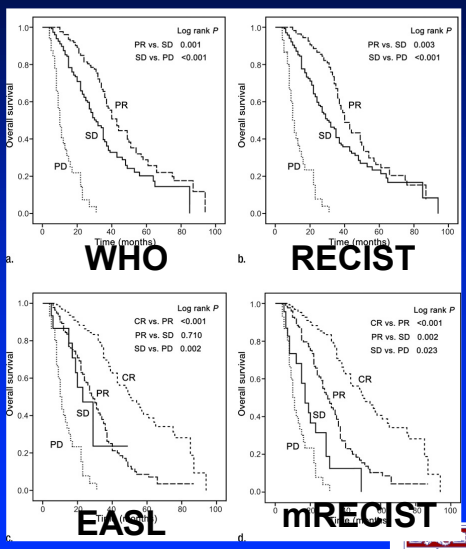
Complete Response (CR)	100% decrease in amount of enhancing tissue in index lesion
Partial Response (PR)	≥50% decrease in amount of enhancing tissue in index lesion
Stable Disease (SD)	<50% decrease in to ≤25% increase in amount of enhancing tissue in index lesion
Progressive Disease (PD) **Simple visual estimate**	>25% increase in amount of enhancing tissue in index lesion New lesions or metastases New vascular invasion New enhancement in previously treated index lesion warranting further locoregional therapy

Penn IR

40

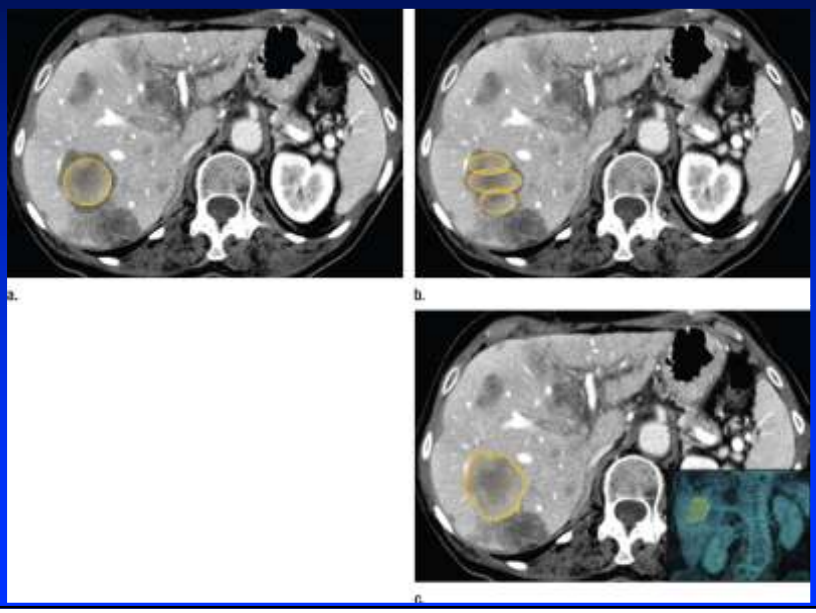
Shim RAD 2012;262;708-718





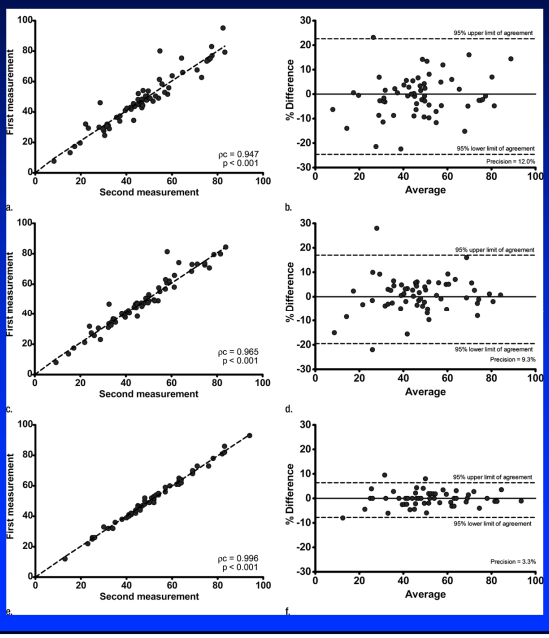
41

Precision: single ROI, mult ROI, volumetric ROI



42

Precision: single ROI, mult ROI, volumetric ROI



Precision was 12.0% for the single-ROI method

9.3% for the average-measurement-in-three-ROIs method

3.3% for the volumetric method.

Chalian, RAD 2012;262:853

43

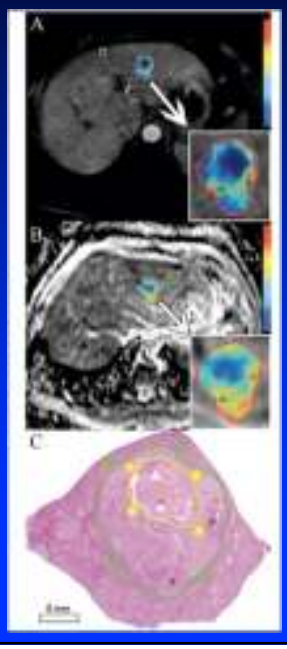
Radiology. Dec 2014; 273(3): 746–758.

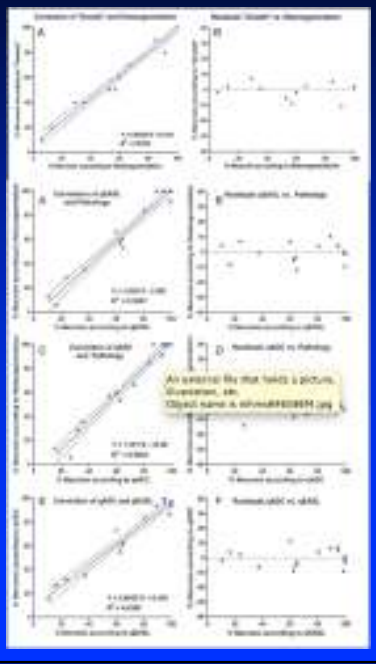
Radiologic-Pathologic Analysis of contrast-enhanced and Diffusion-weighted MR imaging in Patients with HCC after TACE: Diagnostic Accuracy of 3D Quantitative Image Analysis

17 HCC resected after TACE

3D segmentation for contrast enhancement and ADC

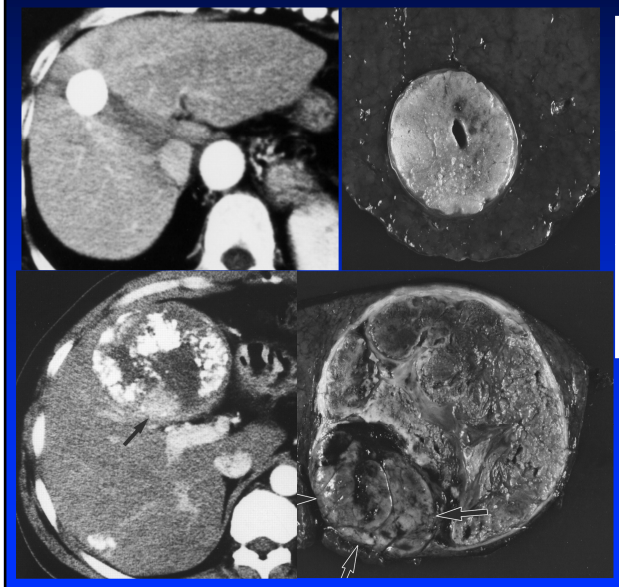
Slice-by-slice correlation with histology



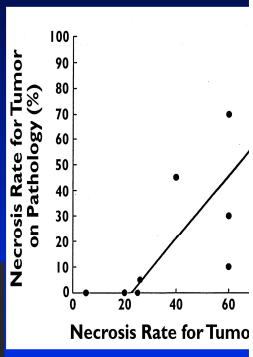


44

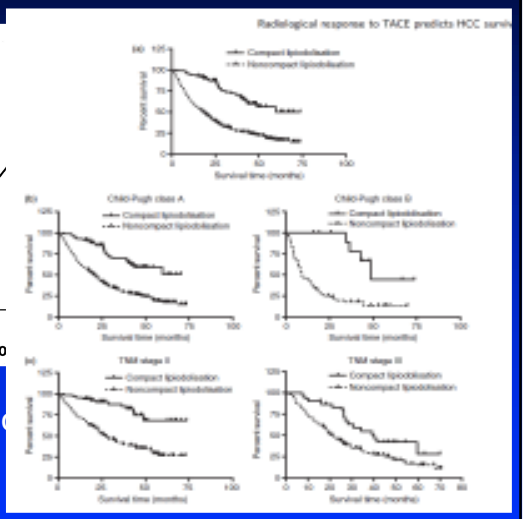
Old Concept: Lipiodol as an imaging biomarker for response



Takayasu, K. et al. Am. J. Roentgenol. 2000;175:699-704



Path validati



Kim, AP&T 2012;35:1343

AJR

Penn IR

45

New Concept: Index lesion

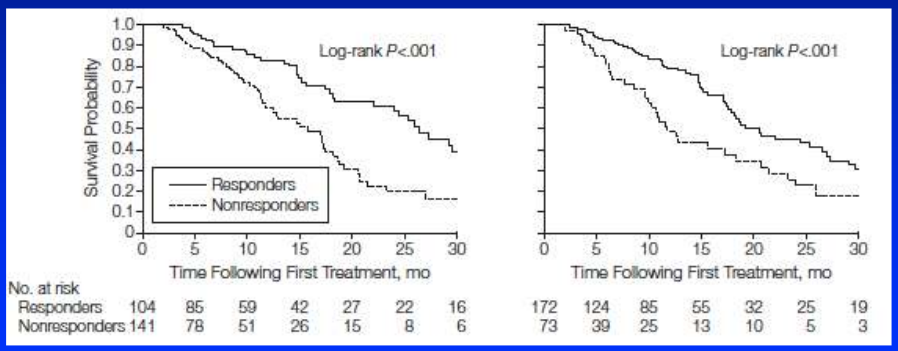
JAMA®

Online article and related content current as of March 23, 2010.

Imaging Response in the Primary Index Lesion and Clinical Outcomes Following Transarterial Locoregional Therapy for Hepatocellular Carcinoma

Ahsun Riaz; Frank H. Miller; Laura M. Kulik; et al.

JAMA. 2010;303(11):1062-1069 (doi:10.1001/jama.2010.262)



WHO Criteria NWU Necrosis Criteria

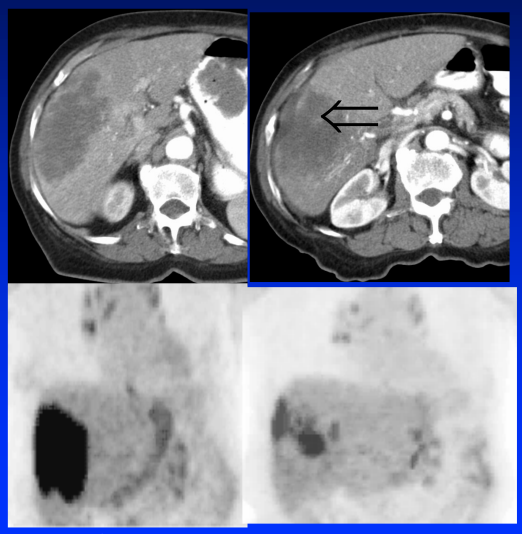
Penn IR

46

PET post Y-90

Stable on CT

mprovement on PET



courtesy R. Salem

⇐	Response CT	No Response CT
Response PET	5	13
No Response PET	0	1

European Journal Nuclear Medicine-March 2002

Penn IR

47

Take Home Points

- Image-guided therapies pose multiple challenges in clinical trial design
- Time-based outcomes such as survival and disease control should allow for staged and repeated therapy
- Imaging-based endpoints surprisingly difficult to measure reliably
- More robust response criteria for IO needed

Penn IR

48