



Annual

**Advances and Innovations in Endoscopic Oncology
and Multidisciplinary Gastrointestinal Cancer Care**

Designing Trials with Oncology Adoption in Mind

Christina Wu, MD

Professor of Medicine

Division of Hematology/Medical Oncology

Mayo Clinic, Arizona



Disclosures

- Grant/Research Support from Pfizer.
- Consultant for DoMore Diagnostics, Exelixis, GlaxoSmithKline, Merck, Natera and Pfizer.

This presentation and/or comments will be free of any bias toward or promotion of the above referenced companies or their product(s) and/or other business interests.

This presentation and/or comments will provide a balanced, non-promotional, and evidence-based approach to all diagnostic, therapeutic and/or research related content.

This presentation has been peer-reviewed and no conflicts were noted.

The endpoint is the adoption argument

What a trial measures determines whether clinicians, payers, and guideline committees change practice. For a procedure or device, the endpoint must be credible, measurable, and attributable to the intervention itself.

Clinically significant

A difference patients feel: longer life, delayed decline, fewer symptoms, less toxicity.

Statistically significant

A difference unlikely to be chance. Necessary, but never sufficient on its own.

Regulatory relevant

An endpoint the FDA accepts as a basis for approval or label expansion.

Adoption relevant

Evidence strong enough to move guidelines, reimbursement, and referral behavior.

From tumor shrinkage to survival

ORR	PFS	OS
Response Rate Tumor burden change at a point in time <i>Single-arm interpretable; earliest signal</i>	Progression-Free Survival Time without growth or death <i>Faster, smaller trials; assessment-dependent</i>	Overall Survival Time to death from any cause <i>Unambiguous; slowest and largest</i>

Increasing certainty of patient benefit → increasing trial size, duration, and cost

ENDPOINT 1

Objective response rate (ORR)

Definition Proportion of patients with a predefined reduction in tumor burden (complete + partial response) confirmed by RECIST 1.1 or a modality-specific criterion. Duration of response (DoR) should always be reported alongside it.

Strengths

- Directly attributable to the intervention
- Interpretable in single-arm and early-phase studies
- Reads out early
- Useful for locoregional procedures with a measurable target lesion

Limitations

- Shrinkage does not reliably translate into longer life
- Sensitive to imaging interval, reader variability, and criteria chosen
- Ignores stable disease, a common benefit pattern for devices and immunotherapy
- Says nothing about durability unless DoR is reported

ENDPOINT 2

Progression-free survival (PFS)

Definition Time from randomization to documented progression or death from any cause, whichever comes first. Patients without an event are censored — so censoring rules, scan schedule, and who adjudicates progression are part of the endpoint, not footnotes.

Strengths

- Captures disease control, not just shrinkage
- Events accrue sooner than deaths — smaller, shorter, cheaper trials
- Not diluted by subsequent lines of therapy or crossover
- Delaying progression is itself meaningful when it delays symptoms

Limitations

- Assessment bias in open-label procedure trials — use blinded independent central review
- Unequal or missed scans create informative censoring
- A statistically significant gain of weeks may be clinically trivial
- Correlation with OS varies widely by tumor type and setting

ENDPOINT 3

Overall survival (OS)

Definition Time from randomization to death from any cause. Unambiguous, objectively measured, and immune to assessment bias — which is why it remains the reference standard for demonstrating clinical benefit.

Strengths

- No adjudication, no imaging schedule, no reader variability
- Captures the net effect of benefit and harm together
- Most persuasive endpoint for guidelines and payers
- Essential when a procedure carries real periprocedural mortality

Limitations

- Requires large samples and long follow-up
- Diluted by crossover and effective post-progression therapies
- Deaths from unrelated causes add noise
- Often impractical as the primary endpoint for a single-center procedure trial

Choosing among the three

	ORR	PFS	OS
What it measures	Tumor burden reduction	Time free of growth or death	Time to death, any cause
Blinding required	Central imaging review	Blinded independent review	Not required
Susceptible to bias	Moderate	High in open-label trials	Minimal
Trial size / duration	Small, fast	Moderate	Large, long
Affected by crossover	No	No	Yes

What makes intervention trials different

Blinding is hard

Sham or placebo procedures are often unethical or impractical, so open-label assessment bias falls directly on PFS and response reads.

Operator learning curve

Early cases underperform. Credentialing, case minimums, and run-in phases protect the estimate of effect.

Technical success is not benefit

Ablation completeness or R0 resection is a process measure. Report it, but never substitute it for a patient outcome.

Comparator drifts

Systemic therapy improves during a long procedure trial, eroding the control-arm contrast and the OS signal.

Periprocedural harm is front-loaded

Early mortality and morbidity can offset later gains — a reason OS and safety must be co-primary considerations.

Equipoise and recruitment

Strong clinician preference slows accrual and skews who is enrolled; prespecify eligibility and track screen failures.

Endpoints patients actually feel

When an intervention is supportive, palliative, or toxicity-directed, ORR, PFS, and OS may be the wrong primary endpoints entirely.

● Toxicity incidence

Rate of protocol-specified grade ≥ 2 events over a defined window

● Quality of life

Validated instruments such as EORTC QLQ-C30, FACT, or DLQI

● Patient-reported outcomes

PRO-CTCAE symptom burden collected directly from patients

● Time to deterioration

Time to a clinically meaningful decline in function or symptom score

● Dose intensity preserved

Fewer interruptions, reductions, or discontinuations of effective therapy

● Health care utilization

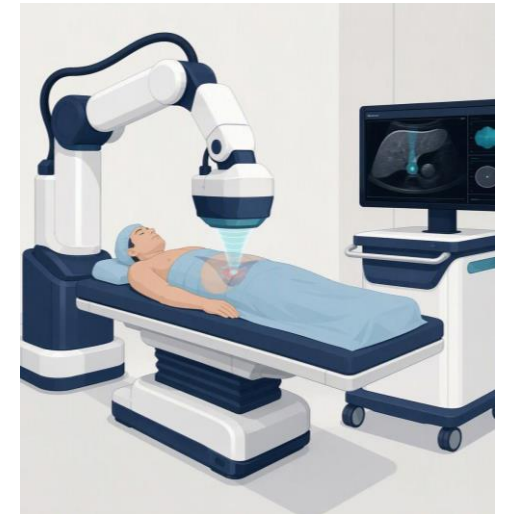
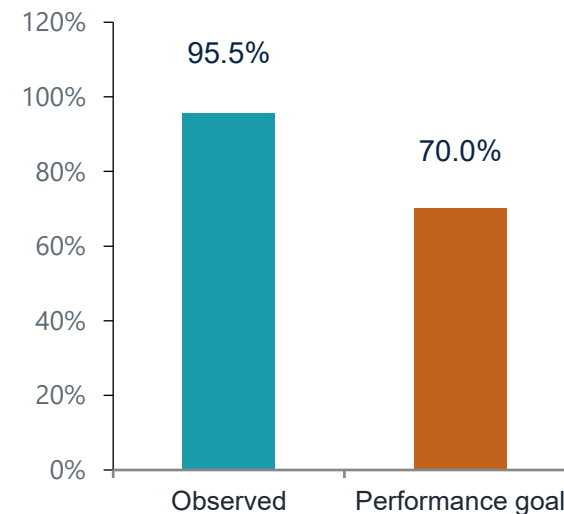
Admissions, procedures avoided, and days spent at home

CASE EXAMPLE

#HOPE4LIVER: Histotripsy

Design

- Prospective single-arm pivotal trials at 14 US and EU sites, 2021–2022; 44 patients, 49 tumours (18 HCC, 26 metastatic)
- Non-invasive focused-ultrasound histotripsy of up to three tumours under 3 cm; assessment at 36 hours and 30 days
- Co-primary endpoints: technical success 42/44 (95.5%) against a 70% performance goal, and major complications 3/44 against a 25% goal
- No ORR, PFS or OS endpoint, no comparator arm, and a 30-day horizon — so no evidence yet that patients live longer or progress later



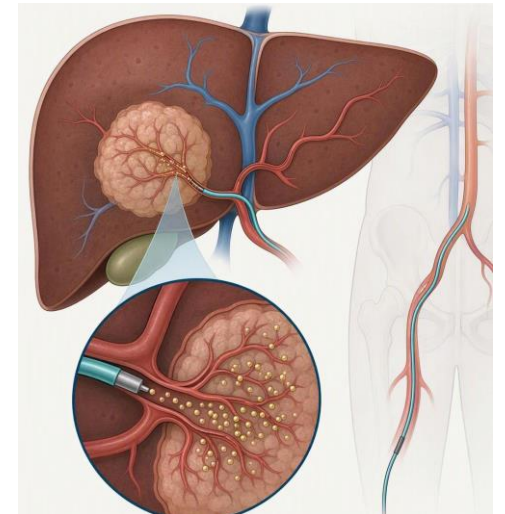
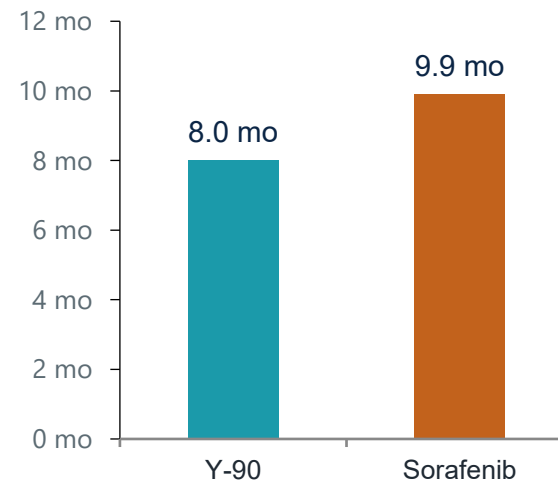
Endpoint lesson: Technical success is a device endpoint, not a patient endpoint. It cleared FDA de novo review but leaves the oncologist's question unanswered, which is why uptake has lagged the headline 95%.

CASE EXAMPLE

SARAH: local control without a survival gain

Design

- 459 patients with locally advanced or inoperable hepatocellular carcinoma, randomized 1:1
- **Yttrium-90 resin microsphere radioembolization** vs sorafenib, an open-label procedure-versus-drug comparison
- Primary endpoint: overall survival — median 8.0 vs 9.9 months (HR 1.15, $p = 0.18$); superiority not shown
- Radioembolization was better tolerated, with fewer treatment-related adverse events and better quality of life — but no survival gain followed



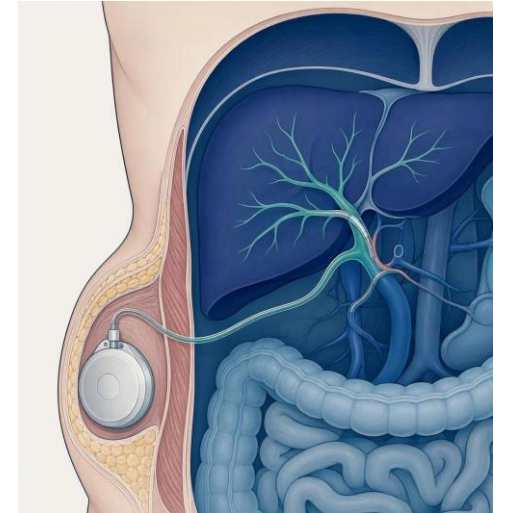
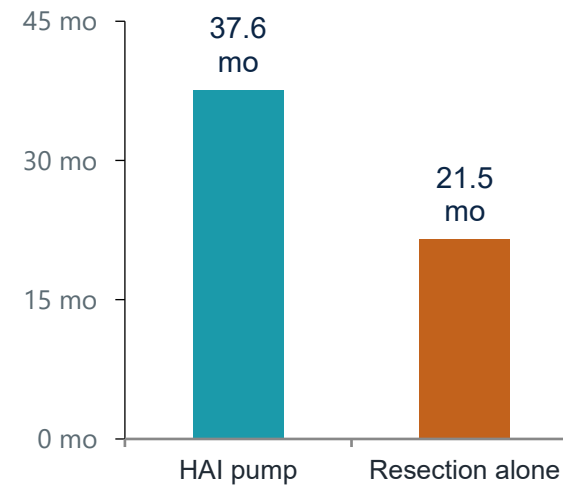
Endpoint lesson: Delay in hepatic progression did not persuade oncologists, because overall survival did not change. A local endpoint earns practice change only when it translates into survival or quality of life.

CASE EXAMPLE

PUMP: hepatic control without a PFS gain

Design

- 241 patients randomized across 12 Dutch centres, 2018–2026; 220 in the modified intention-to-treat analysis
- Resection of colorectal liver metastases plus an **implanted hepatic arterial infusion pump** delivering floxuridine, vs resection alone
- Primary endpoint: progression-free survival — median 15.0 vs 15.5 months (HR 0.92, $p = 0.66$); no benefit shown
- Hepatic PFS nearly doubled, 37.6 vs 21.5 months (HR 0.82, $p = 0.32$)



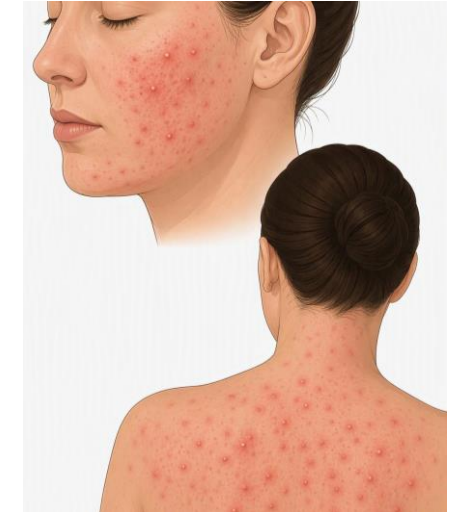
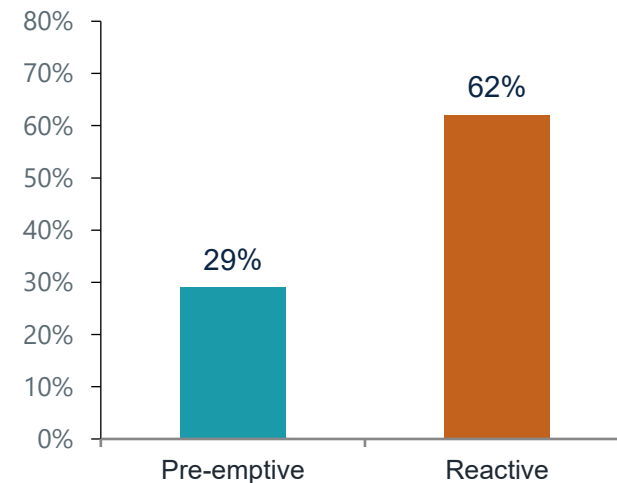
Endpoint lesson: The pump did what it was designed to do inside the liver, but the trial was powered on overall PFS. An organ-specific win that the primary endpoint cannot register will not move practice.

CASE EXAMPLE

STEPP: when the endpoint is improving toxicity

Design

- 95 patients on panitumumab-containing therapy, randomized 1:1
- Pre-emptive skin regimen (moisturizer, sunscreen, topical steroid, doxycycline) vs reactive treatment after toxicity appeared
- Primary endpoint: incidence of protocol-specified grade ≥ 2 skin toxicity during the 6-week treatment period
- Quality of life measured with the Dermatology Life Quality Index



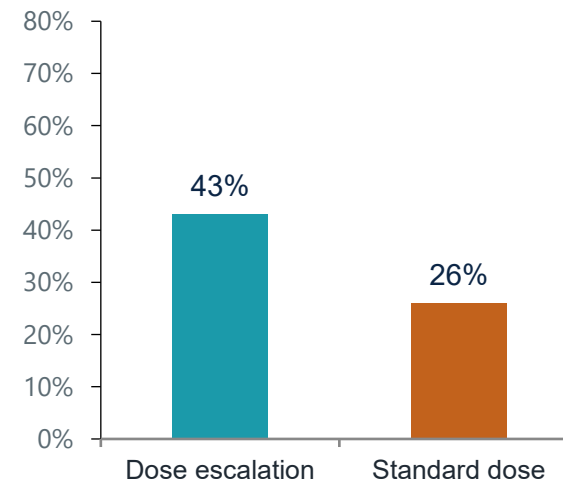
Endpoint lesson: A supportive-care intervention earned adoption on toxicity and quality of life, with no requirement to show an ORR, PFS, or OS advantage.

CASE EXAMPLE

ReDOS: an endpoint that measured deliverability

Design

- 116 evaluable patients with refractory metastatic colorectal cancer at 39 US sites, randomized 1:1
- Weekly dose escalation from 80 mg to 160 mg/day vs standard 160 mg/day dosing
- Primary endpoint: proportion of patients starting cycle 3 — 43% vs 26% (one-sided $p = 0.043$)
- Median OS 9.8 vs 6.0 months (HR 0.72, $p = 0.12$); PFS 2.8 vs 2.0 months; quality of life preserved



Regorafenib toxicity

CORRECT phase 3; Grothey A, et al. Lancet. 2013;381(9863):303–312.

Hand-foot skin reaction

47% any / 17% grade 3

Fatigue

47% any / 10% grade 3–4

Diarrhoea

34% any / 7% grade 3

Hypertension

28% any / 7% grade 3

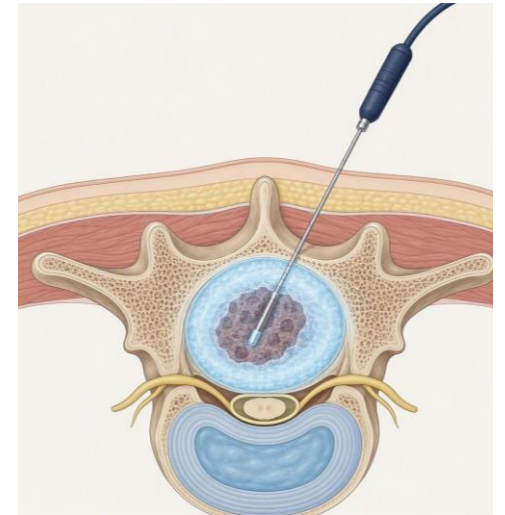
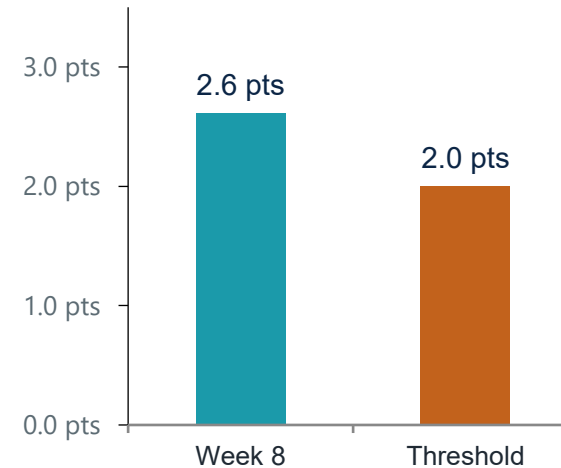
Endpoint lesson: A feasibility endpoint — staying on therapy — changed practice, while the OS and PFS trends supported it rather than carried it.

CASE EXAMPLE

MOTION: cryoablation judged on pain relief

Design

- Prospective single-arm phase 2 trial across 11 centres (7 US, 4 EU); 66 patients treated, 65 completed follow-up
- Percutaneous cryoablation of one painful bone metastasis per patient; median hospital stay 26.6 hours
- Primary endpoint: change in BPI-SF worst pain from baseline to week 8 — mean -2.61 points (95% CI -3.45 to -1.78)
- Relief held through 24 weeks, with reduced opioid use and improved quality of life as secondary endpoints



Endpoint lesson: no ORR, PFS or OS endpoint could capture this benefit. Where the aim is palliation, speed and durability of symptom relief — not tumour shrinkage — are the outcomes that define success.

TAKEAWAYS

Matching the endpoint to the question

- | | | |
|-----------|--|---|
| 1 | Start from the claim | Write the sentence you want in the guideline first, then pick the endpoint that earns it. |
| 17 | Pair a surrogate with a hard endpoint | Power on PFS or ORR if you must, but collect OS and safety to defend the result. |
| 3 | Protect the measurement | Blinded central review, fixed imaging intervals, and prespecified censoring rules. |
| 4 | Measure what the intervention plausibly changes | A procedure that reduces toxicity should be judged on toxicity and quality of life. |
| 5 | Prespecify clinical significance | Define the minimum difference that would change your own practice before the data arrive. |

Thank you!



wu.christina@mayo.edu