Endpoint Selection in Phase II Oncology trials

Anastasia Ivanova

Department of Biostatistics
UNC at Chapel Hill
aivanova@bios.unc.edu

Primary endpoints in Phase II trials

Recently looked at journal articles published in leading oncology journals in 2005 and 2010

- Country
- Cancer site
- Single center vs multicenter
- Design (single stage, Simon, Fleming etc)
- Number of arms (1 or 2),
  if more than 1 arm, was it randomized?
- Primary endpoint, secondary endpoints
- Was H_0 rejected? Was new treatment recommended?
- Were stats clearly described?

Primary endpoints in Phase II trials

- Statistical design

	2005	2010
Single stage	**32%** of articles	**46%** of articles
Simon's	**58%** of articles	**46%** of articles
Fleming's	**8%** of articles	**8%** of articles

Primary endpoints in Phase II trials

- Size of the trial

	2005	2010
≤ 40	**74%** of articles	**48%** of articles
41-60	**18%** of articles	**26%** of articles
< 60	**8%** of articles	**26%** of articles

Primary endpoints in Phase II trials

- Statistical design is **NOT** clearly described

2005	2010
18% of articles	**5%** of articles

Primary endpoints in Phase II trials

- Response rate (RR) is the most frequent primary endpoint in Phase II oncology trials (preliminary results)

Response (RR) = Complete Response + Partial Response

	2005	2010
Response rate	**81%** of articles	**54%** of articles

PFS and PFS6

- The second most frequently used primary endpoint is

PFS (Progression Free Survival)	**15%**
PFS6, PFS3, PFS2, PFS1.5	**15%**

- Trials with PFS6 use **Simon's design** to allow for early stopping for futility.

- Data are summarized using a **Kaplan-Meier curve** rather than **proportion** of patients **with PFS6**

CR, PR and SD

- Other primary endpoints include
  - **Complete Response** (CR)
  - **Disease control** = CR + PR + Stable Disease
  - **Clinical benefit** = CR + PR + Stable Disease >3 months

Growth modulation index (GMI)

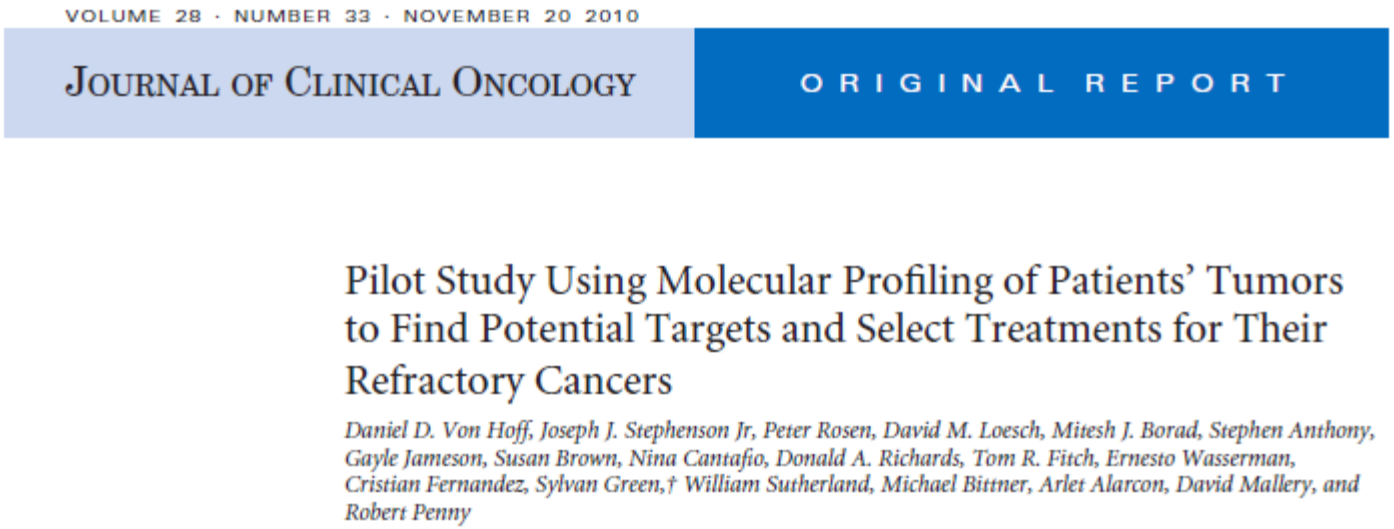
- Each patient receives **two successive treatments**, standard first and then experimental (usually a biomarker driven)

- **TTP(1)** - time to progression on the first treatment
  TTP(2) - time to progression on the second treatment

- **GMI = TTP(2) / TTP(1)**

Growth modulation index (GMI)

- GMI was introduced by Von Hoff (1998)
- Von Hoff et al. (2010)



VOLUME 28 · NUMBER 33 · NOVEMBER 20 2010

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Pilot Study Using Molecular Profiling of Patients' Tumors to Find Potential Targets and Select Treatments for Their Refractory Cancers

Daniel D. Von Hoff, Joseph J. Stephenson Jr, Peter Rosen, David M. Loesch, Mitesh J. Borad, Stephen Anthony, Gayle Jameson, Susan Brown, Nina Cantafio, Donald A. Richards, Tom R. Fitch, Ernesto Wasserman, Cristian Fernandez, Sylvan Green,† William Sutherland, Michael Bittner, Arlet Alarcon, David Mallery, and Robert Penny

- Cousin (2013) showed that **GMI of higher than 1.33** was associated with significant improvement in OS

Proportion of successful Phase II trials

	NO	YES
Null hypothesis rejected	82%	18%
Drug recommend for future investigation	44%	56%

Successful Phase II trials

- If H_0 was not rejected based on RR, however PFS or OS was good, and therefore the drug is recommended

- Publications with stats not clearly described were significantly more likely to recommend the drug for further study

Tumor Response versus Disease Control

CR = Complete response

PR = Partial Response

SD = Stable Disease

PD = Progressive Disease

Tumor Response (**TR**) = **CR + PR**

Disease Control (**DC**) = **TR + SD**

Tumor Response versus Disease Control

- Historically the RR was a frequently used endpoint to assess novel cytotoxic agents which induced tumor regression

- Many targeted agents are **cytostatic** and inhibit tumor growth rather than cause tumor regression.

- In situations such as this, when RR is modest, clinicians may be interested in disease control (DC) since it may better predict clinical outcome.

Tumor Response versus Disease Control

- An unblinded single arm trial may be **susceptible to bias** in the assessment of tumor response by the investigator and lacks a "control arm" for comparison.

- Issues in determining tumor response have included variation in the assessment of response or variability in tumor measurements on repeat imaging.

Example: Phase II trial in NSCLC, testing TR and DC simultaneously

Phase II trial in patients with advanced non-small cell lung cancer who have experienced disease progression after platinum-based therapy

Principal Investigator Tom Stinchcombe

Simon's minimax design for DC

$p_{0D} = 0.25$, $p_{1D} = 0.50$, $\alpha = 0.05$

24 patients are needed for power of 80%

Simon's minimax design for **RR**

$p_{0T} = 0.05$, $p_{1T} = 0.25$, $\alpha = 0.05$

16 patients are needed for power of 80%

Testing TR and DC (TR + SD) simultaneously

Consider two hypotheses:

about tumor response (TR) rate

and about disease control (DC) rate, DC = TR + SD

$$H_{0T}: \; p_T < p_{0T} \quad \text{versus} \quad H_{1T}: \; p_T > p_{0T}$$

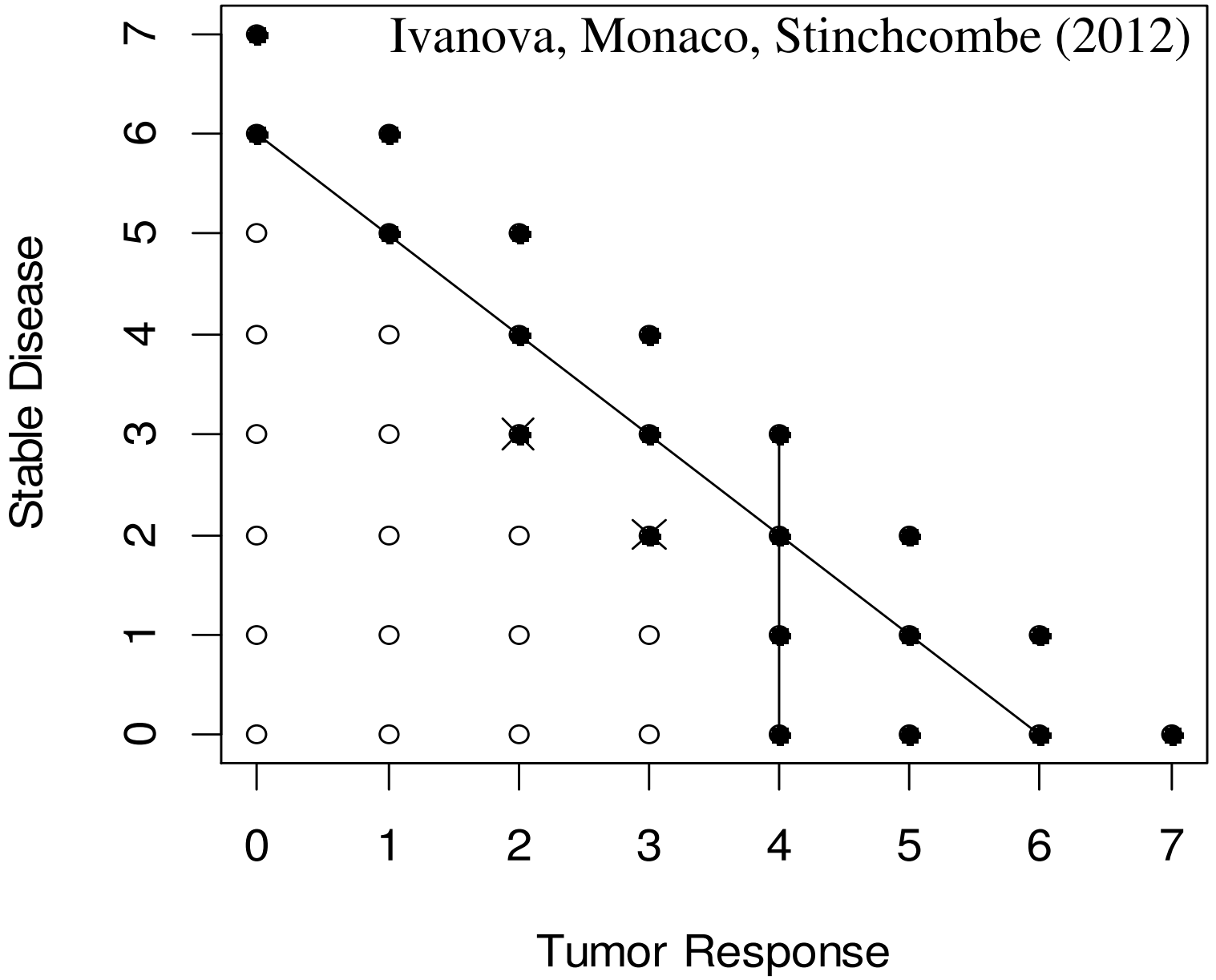
$$H_{0D}: \; p_D < p_{0D} \quad \text{versus} \quad H_{1D}: \; p_D > p_{0D}$$

We would like to test $H_{0T} \cap H_{0D}$

The drug is accepted if TR rate is good OR DC rate is good, that is, if one of the individual hypothesis is rejected.

The drug is rejected if BOTH TR rate and DC rate are poor.

Example. $p_{0T} = 0.15$, $p_{1T} = 0.55$, $p_{0D} = 0.35$, $p_{1D} = 0.75$, $\alpha = 0.05$



Example: Phase II trial in NSCLC, testing TR and DC simultaneously

Design used in the trial

$p_{0T} = 0.05$, $p_{1T} = 0.25$,

$p_{0D} = 0.25$, $p_{1D} = 0.50$, $\alpha = 0.05$

$n = 18$, $n_1 = 12$

Power for $H_{0T} \cap H_{0D}$ under various alternatives

BOTH	TR only	DC only
$p_{1T} = 0.25$	$p_{1T} = 0.25$	$p_{1T} = 0$
$p_{1D} = 0.50$	$p_{1D} = 0.25$	$p_{1D} = 0.50$
0.80	**0.76**	**0.70**

Example: A Phase II trial with relaxed stopping for futility

A Phase II trial of a chemotherapy and a monoclonal antibody in older patients with previously untreated diffuse large B-cell lymphoma

Principal Investigator Steven I. Park

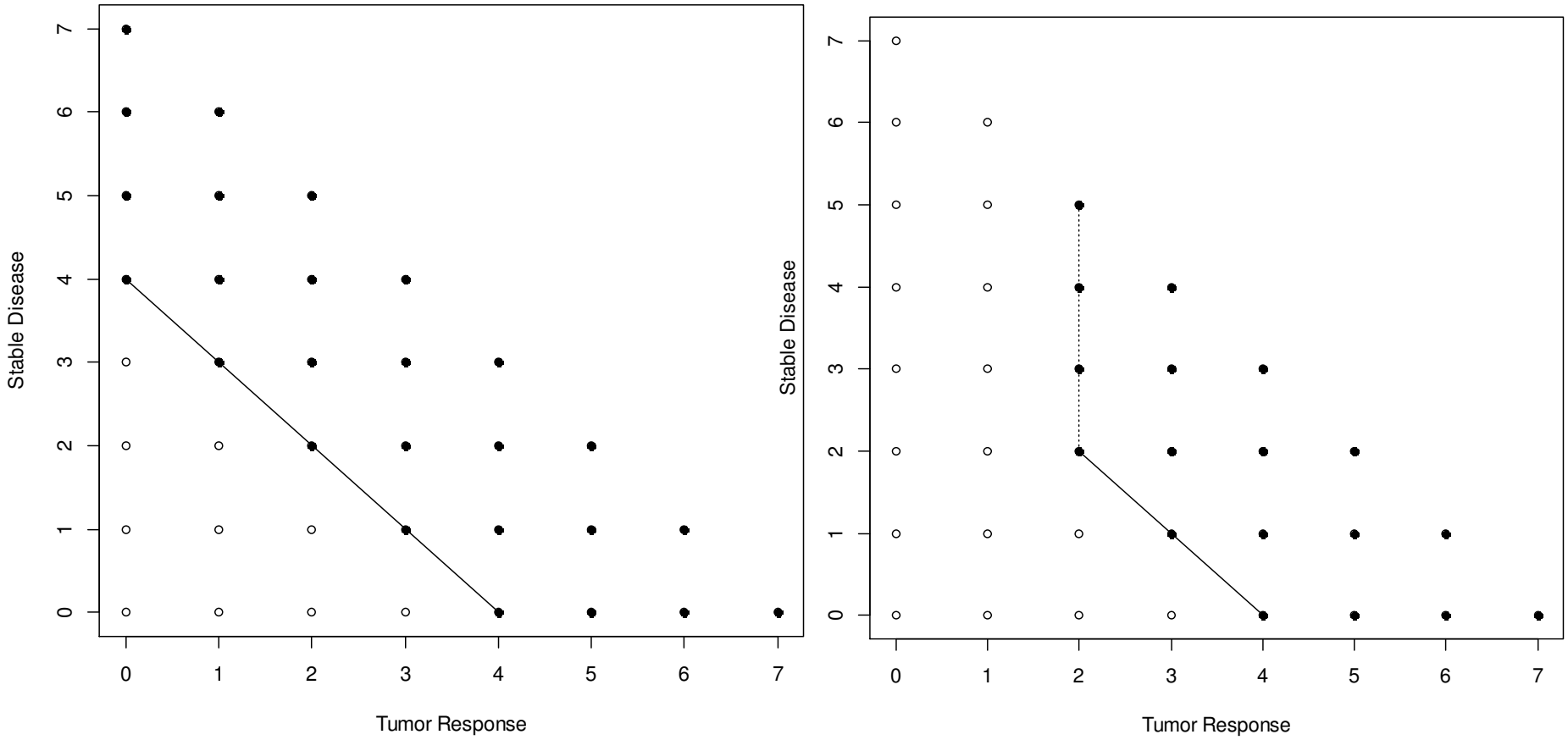
- **Tumor Response (TR) is the primary outcome**

  Would like to use Simon's two-stage design with TR as the primary endpoint.

- **Would like to base futility stopping on the rate of Disease Control**

  DC = TR + Stable Disease, at the end of the trial

Interim analysis for futility



Phase II oncology trials

Simon (1989)	Extensions
Single arm	Two-arm (Kepner, 2011)
Two-stage	Three-stage (Cheng, 2007)
Early stopping for futility only	Early stopping for futility and efficacy (Fleming, 1982)
min(N) or min(E_0N)	Admissible (Jung et al., 2004) min[$wN + (1-w)E_0N$]
Binary endpoint	Ordinal outcome Lu et al. (2005), Ivanova et al. (2012)

Software for Phase II oncology trials

Easy to use web-based software

http://cancer.unc.edu/biostatistics/program/ivanova/

- Simon's admissible designs
- Fleming's admissible designs
- Two-stage designs with relaxed stopping for futility
- Two-stage designs with ordinal outcome
- Continuous toxicity monitoring

- Rapid Enrollment Design (RED) for dose finding (coming soon)

References

Cheng, T (2007). Optimal three-stage designs for phase II cancer clinical trials. *Statistics in Medicine* 16: 2701-2711

Fleming, TR (1982). One-Sample multiple testing procedure for Phase II clinical trials, *Biometrics* 38: 143-151

Ivanova, A, Monaco, J and Stinchcombe, T. (2012). Efficient designs for phase II oncology trials with ordinal outcome. *Statistics and its interface* 5: 463-469.

Jung, S, Lee, T, Kim, KM, George, SL (2004). Admissible two-stage designs for phase II cancer clinical trials. *Statistics in Medicine* 23: 561-569.

Lu, Y, Jin, H, Lamborn, KR. (2005). A design of phase II cancer trials using total and complete response endpoints. *Statistics in Medicine* 24: 3155–3170.

Simon, R. (1989). Optimal two-stage designs for phase II clinical trials. *Controlled Clinical Trials* 10: 1-10.