

# Outline

---

- Davis et al., Clinical Cancer Research
- Surrogate markers
- Defining outcomes
- Issues relating to measurement

# Motivating Example:

## PD Analysis of Effects of SU5416 or SU6668

---

- Clinical outcomes of interest are biomarkers:
  - receptor phosphorylation (VEGFR and PDGFR)
  - Biomarkers of angiogenesis
- Biopsies were performed at baseline (day 0) and at approximately day 28.
- Data acquisition included 3 to 5 regions in two to three sequential biopsy sections from each tumor at each of baseline and follow-up.

# Motivating Example:

---

- How is this study different than other studies?
  - Primary outcomes are 'efficacy-related', but not clinical.
  - Outcomes are 'surrogate' outcomes or 'correlative' outcomes.
  - Measuring these outcomes is more invasive and more costly than standard safety or efficacy trials.
  - Measurement of these outcomes can be complicated.

# Surrogate Markers

---

- If receptors change as we expect, what can we conclude?
- Is it true, then, that SU6668 and SU5416 are efficacious?
- Not necessarily.....how closely tied are these **markers** and **clinical response**?
- Surrogate outcomes: outcomes in the causal pathway of true outcome.

# Surrogate Markers

---

- Replace a distal endpoint (response) by proxy endpoint (tumor metabolism and/or blood flow).
- Benefits of using surrogate markers
  - Reduction in sample size
  - Reduction in trial duration
  - Reduction in cost
  - Reduction in time to evaluate new therapies
- Their use is NOT AS EASY AS IT SOUNDS...
- Use of a marker as surrogate for outcome requires that you first identify one...

# What is a surrogate marker?

---

## □ ***Defining Characteristic:***

- a marker must predict clinical outcome, in addition to predicting the effect of treatment on clinical outcome

## □ ***Operational Definition***

- establish an association between marker & clinical outcome
- establish an association between marker, treatment & clinical outcome, in which marker mediates relationship between clinical outcome and treatment


# Surrogate Markers

---

**[ Clinical Outcome | Treatment & Marker ]**

=

**[ Clinical Outcome | Marker ]**

 Effect of the treatment on the clinical outcome is completely explained by the effect of the marker

Knowing the status of the marker, the treatment adds nothing more to explaining the clinical outcome.

# Surrogate Markers

---

1) establish an association between marker & clinical outcome.

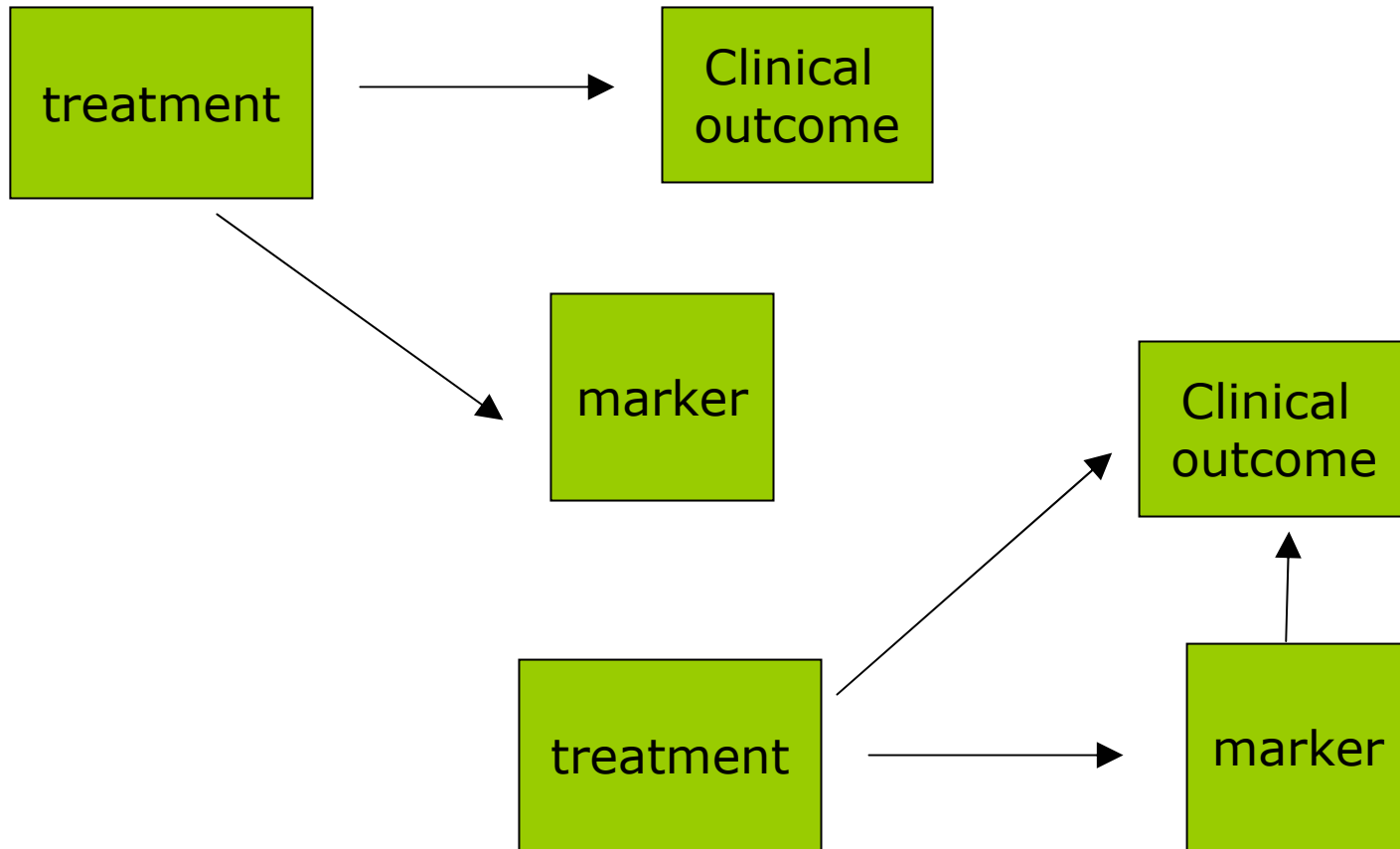


2) establish an association between marker, treatment & clinical outcome, in which marker completely mediates relationship between clinical outcome and treatment.



# NOT Surrogate Markers

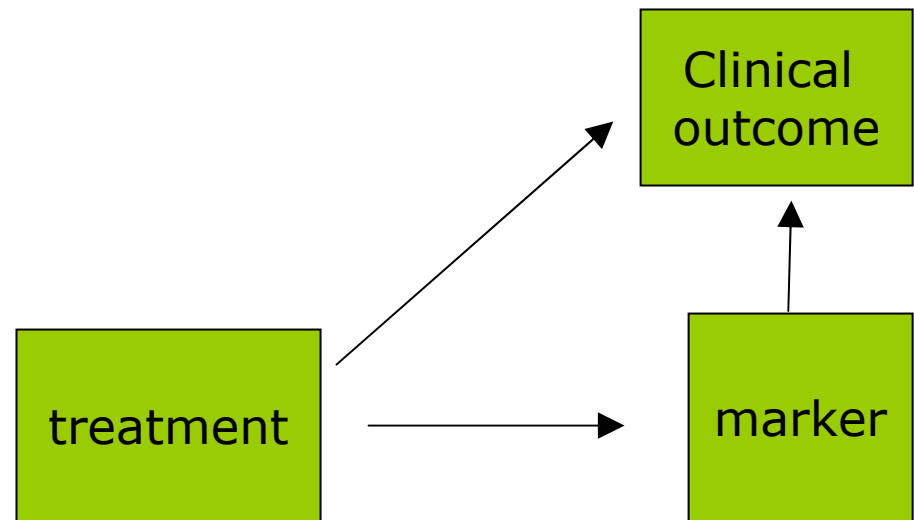
---



# Correlative Studies

---

- ❑ **It is still valid to look at markers that you would expect to be correlated with the clinical outcome.**
- ❑ But, we do not want to be overconfident by saying that they are true 'surrogates.'
- ❑ Correlative studies might include:
  - Pharmacokinetics
  - Pharmacodynamics
  - Other biologic markers that can be measured in serum, biopsy samples, etc.



# Choosing the timing

---

## □ In laboratory studies

- Usually a pre-post setting: how does the measure compare after treatment to before treatment.
- Baseline (before treatment) measure is needed.
- Post-treatment measures:
  - When is treatment at its most potent?
  - How often should response be measured?
  - Expensive? Invasive?
  - Is it sufficient to look at clinic visit times?

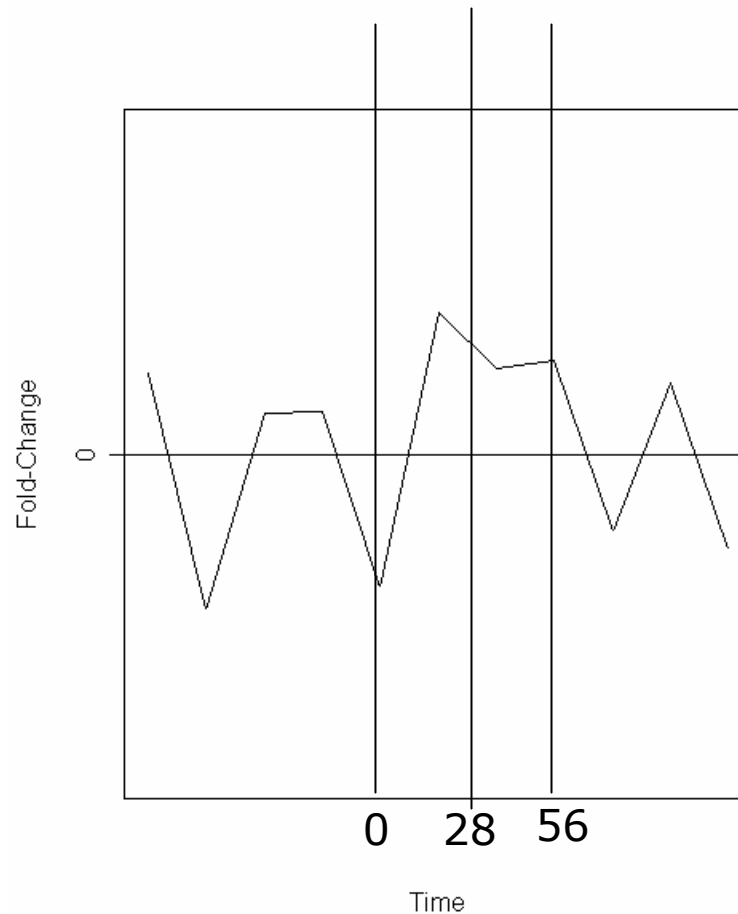
# Defining outcomes in laboratory studies

---

- ❑ They are usually messy
- ❑ More common binary outcomes have nice properties:
  - “Looking for 40% response vs. 20% response”
- ❑ Laboratory outcomes are not so nice:
  - Often skewed.
  - Often have ‘undetectable’ range.
  - Often do not know what to expect.
  - ➔ This makes it hard to plan (i.e. sample size, power).
- ❑ Novel assays: Not obvious what expected changes would be without treatment.
  - How much fluctuation would we expect to see?

# Expected Fluctuations

---



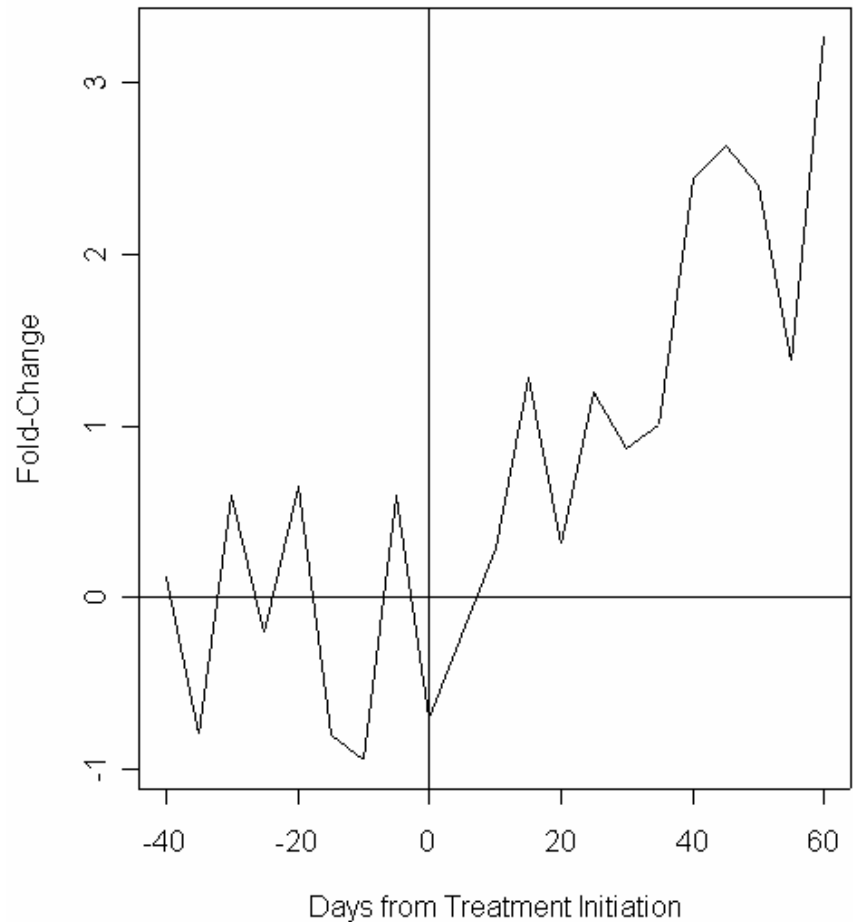
But, with multiple patients, these changes would tend to average to zero.

However, if half of patients have increase and half have decrease, we might conclude that treatment is 50% effective!

# Expected Fluctuations

---

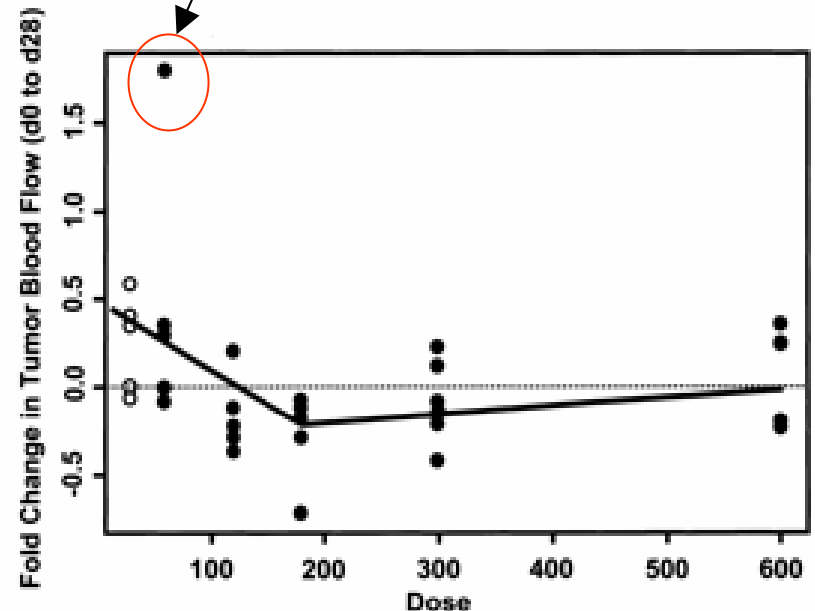
- Follow-up times are compared to baseline
  - That puts a lot of stock in baseline measure
  - Why not consider a 'burn-in' period?
- If baseline is inaccurately measured, all comparisons will be incorrect.



# Measurement Issues

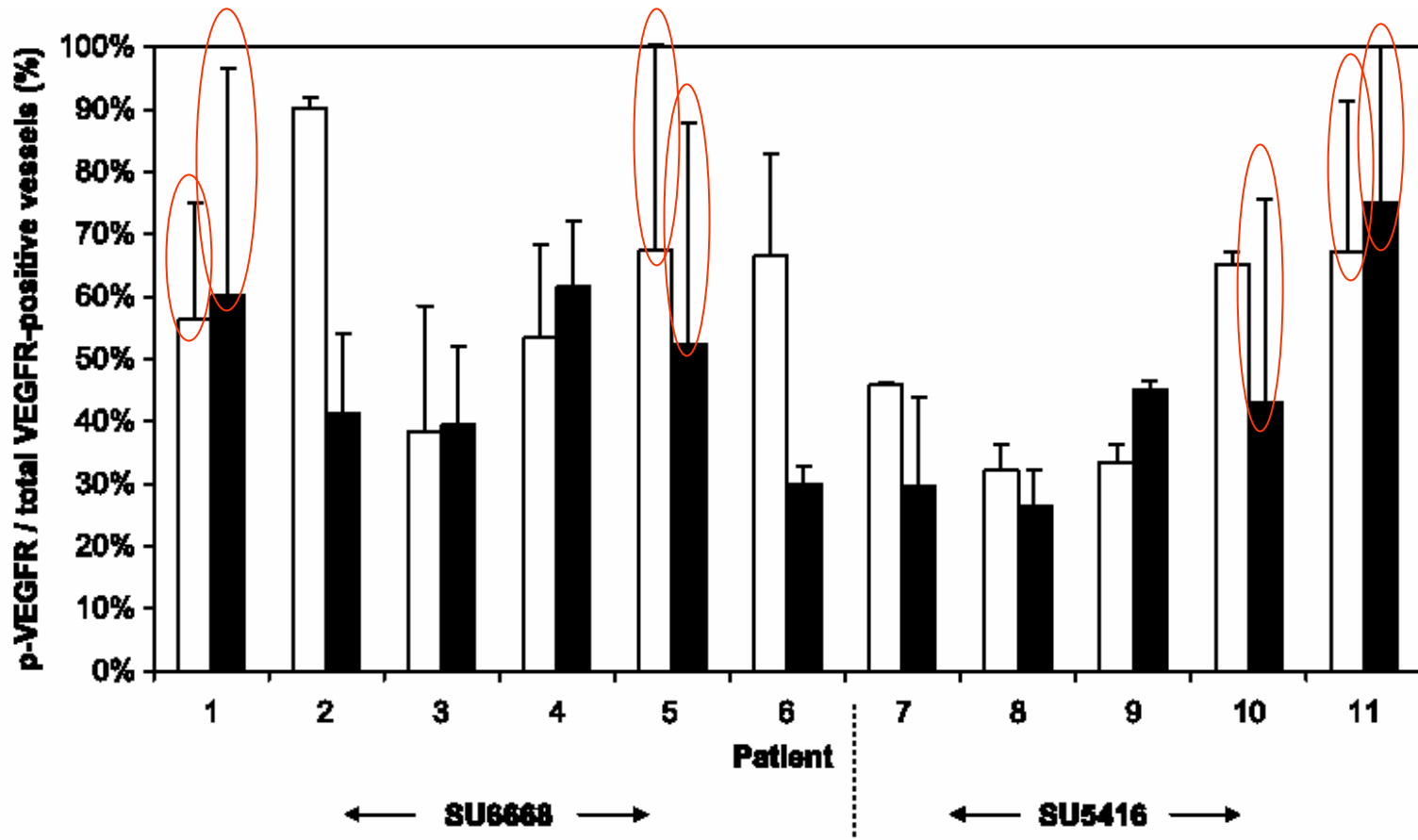
- ❑ Regardless of natural fluctuations...
- ❑ How accurately are we measuring our outcome?
- ❑ How accurately can we measure the target of interest?
- ❑ Are we on target?

A



\* Herbst, Abbruzzese et al. Development of Biologic Markers of Response and Assessment of Antiangiogenic Activity in Clinical Trial of Human Recombinant Endostatin. JCO v. 20, 2002.

# How accurate?



Data acquisition included three to five regions in two to three sequential biopsy sections from each tumor that contained viable tumor cells. Data are represented as a percentage of total tumor microvessels positive for phosphorylated VEGFR-2 total VEGFR-2. Bars, SD.

# Measurement Issues

---

- Why would it be measured inaccurately?
  - Often LONG protocol to get 'final' measure.
  - Lots of room for errors!
- Although the classic efficacy outcome of 'response' is relatively soft, it has been objectively defined.
- Laboratory endpoints require assays and other measurements, and also assumptions.
- Often, assay is being developed along with the trial.

# More Measurement Issues

---

## □ Reliability of procedure

- This includes both
  - Potential heterogeneity within subjects
  - Reproducibility of the assay
- How reproducible are the results?
  - Two samples taken from the same patient on the same day from different lesions?
  - Two samples taken from the same patient on the same day from the same lesion?
  - One sample analyzed twice using the same method?
- Subjectivity
- Inter-rater agreement
- Intra-rater agreement
- In what ways can 'error' come into the procedure?

Great study + bad assay = bad study

# Measurement Issues

---

- ❑ But, how can we know the reliability of the measures?
- ❑ Preliminary studies (pre-clinical).
- ❑ Build it into the study design!
  - Reliability substudy:
    - ❑ Inter-rater agreement
    - ❑ Intra-rater agreement
  - Incorporate burn-in period
  - Take multiple measures
  - Run the assay (test) more than once

# Practical Issue: Biopsies in clinical trials


---

- Including biopsies is great for research
  - Allows investigation of correlative endpoints
  - Helps understand mechanism of action
- But, practically:
  - Often hurts accrual
  - Mandatory versus optional?
  - Potentially large proportion are 'unevaluable'
  - Might only be useful if **both** of paired samples are evaluable
  - Is it worth the effort if you only end up with useable information on a subset of the patients?

# Summary Comments

---

- ❑ Correlative studies often need creative/novel thinking about study design
- ❑ Understanding the properties (i.e. reliability, sensitivity) of your assay/procedure are crucial.
- ❑ Try to build reliability testing into study design
- ❑ Including biopsy in clinical trial has feasibility issues to consider
- ❑ Think carefully about:
  - Are your markers truly surrogates?
  - How often you should measure the outcome?
  - How can we quantify/improve reliability of measurements?



Statisticians usually can provide **guidance**, but not necessarily **answers** to these questions!

Answers need to be based on experience with the novel agent and the underlying biology.