

Tsao et al. (2007) JCO  
Prognostic and Predictive Importance of p53  
and RAS for Adjuvant Chemotherapy in Non-  
Small Cell Lung Cancer

Statistical Discussion

12/14/07

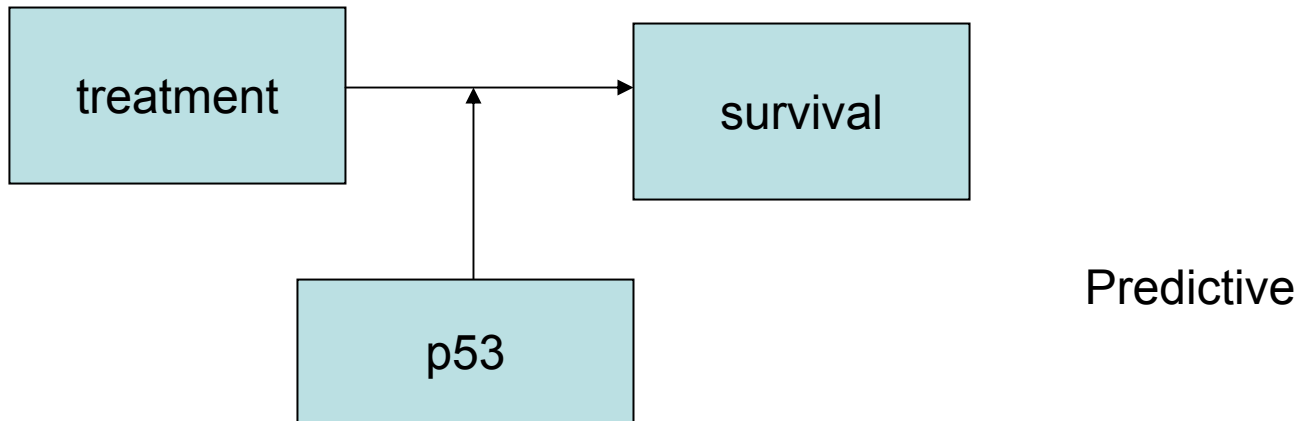
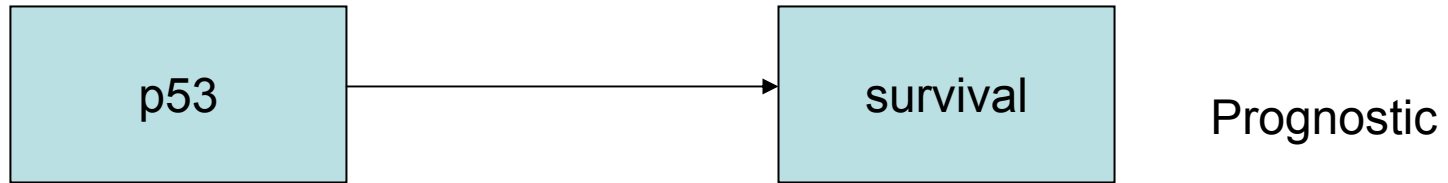
# www.cancer.gov

- prognostic factor (prog-NOS-tik FAK-ter)  
A situation or condition, or a characteristic of a patient, that can be used to estimate the chance of recovery from a disease or the chance of the disease recurring (coming back).
- predictive factor (prih-DIK-tiv FAK-ter)  
A situation or condition that may increase a person's risk of developing a certain disease or disorder.

# Tsao et al.

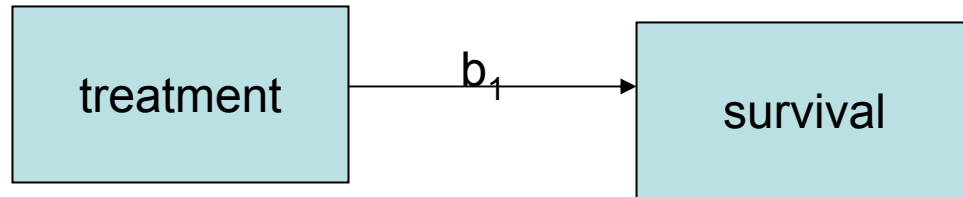
- Prognostic markers: patient or tumor factors that, independent of treatment, predict survival outcome
- Predictive markers: factors that may influence and predict the outcome of treatment in terms of either response or survival benefit

# Prognostic vs. Predictive

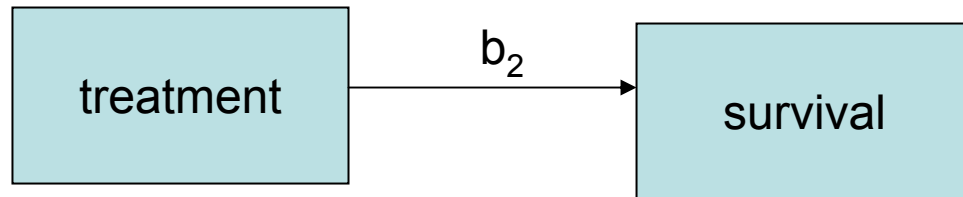


# Another conception of predictive

For p53 IHC +



For p53 IHC -

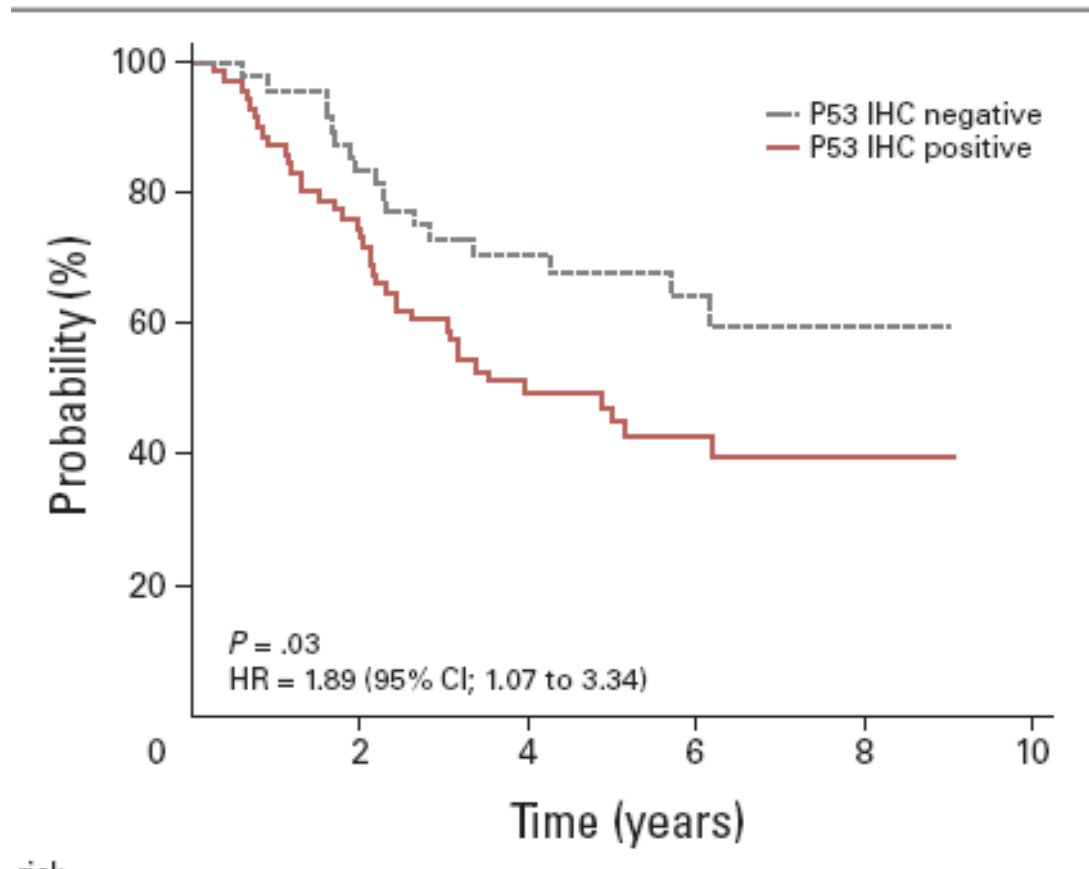


# The problem

- People throw these terms around ALL THE TIME
- There is little consistency:
  - Prediction
  - Prognosis
  - Association

# Common mistake

- Example:

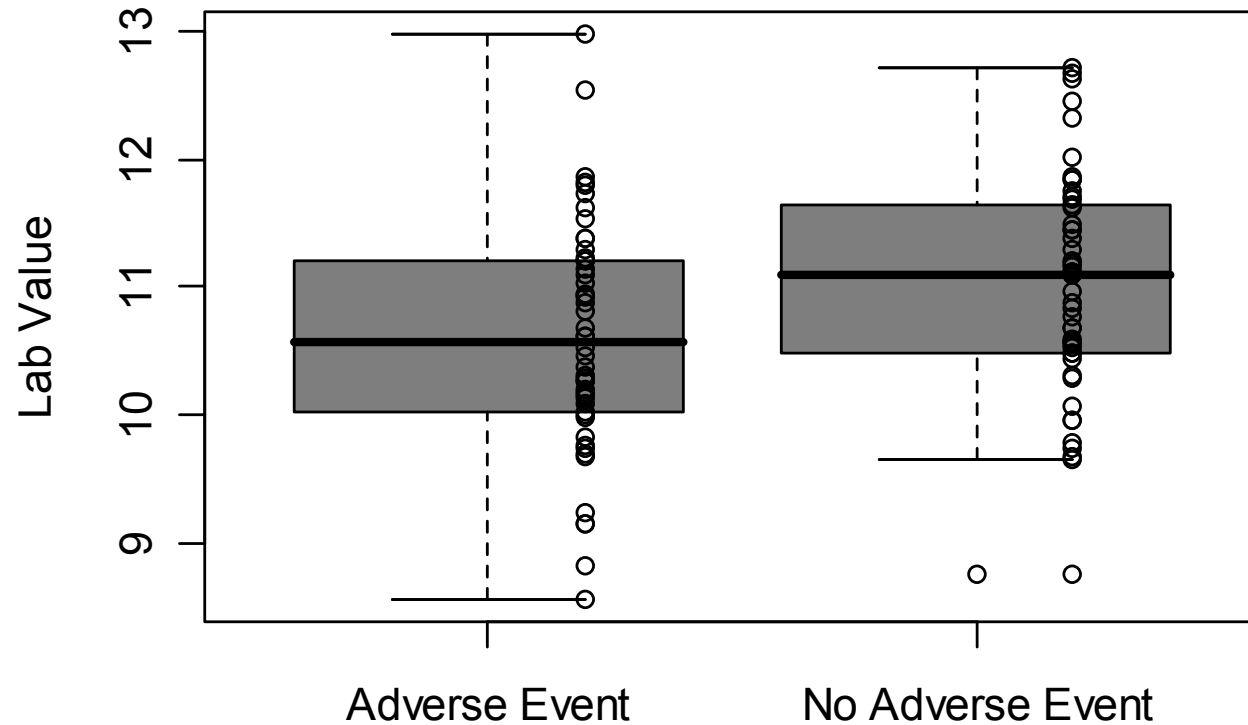


# Common mistake

- Example:
- $P=0.03$
- “Therefore, p53 predicts survival”
- A better interpretation:
  - p53 is associated with survival

# Another example:

$p = 0.02$

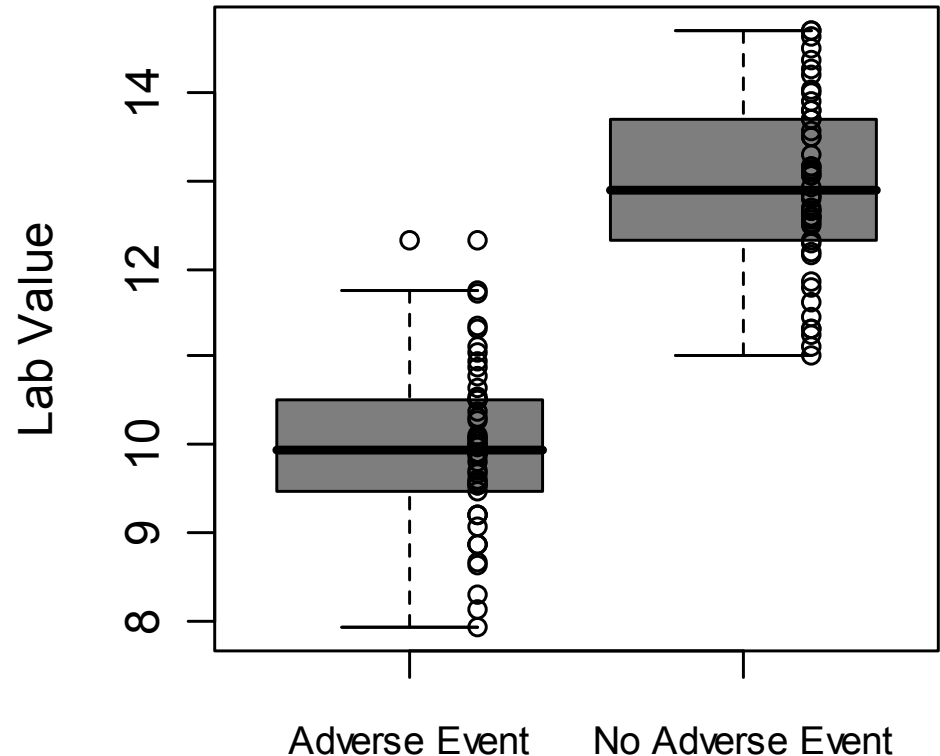


“A low X level is an independent predictor of adverse events. We recommend that all patients have X levels measured preoperatively to identify patients at risk for these events.”

# When could that claim be made?

Two situations

1. the intervention is safe and simple (e.g. vitamin supplementation) and could be given to everyone without risk
2. the distributions are 'well-separated'



# Association $\neq$ Prediction

- Common leap
- Need caution
  - Interpreting the literature!
  - Using the terminology yourself
- Need to ask
  - What would i do differently if i had the information?
  - Should treatment be modified based on this knowledge?
- *Removing treatment is generally a worse error than adding a potentially ineffective (safe) treatment*

# How to distinguish?

- Ignore p-values!
  - Well, not completely
  - They measure **association**
  - Yes, you need a small p-value for ‘prediction’ **but it doesn’t ensure it**
  - **There are many situations with strong association but weak ‘prediction’**
- Other measures
  - Sensitivity and specificity type measures
  - AUC (another time...)

# Sensitivity

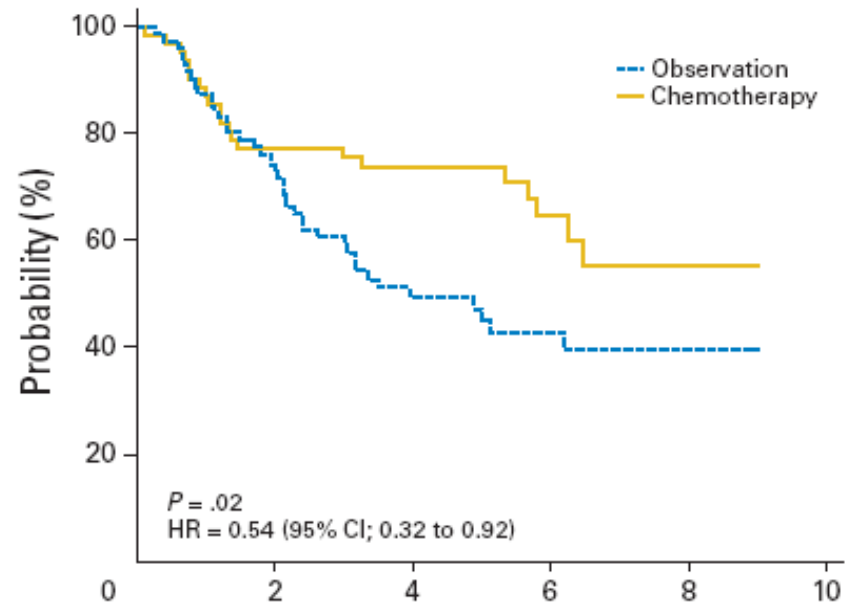
- Classic interpretation:
  - The probability of positive test given you have the disease
- Modified to our context:
  - The probability of p53+ given death < 4 years

# Specificity

- Classic interpretation:
  - The probability of negative test given you do not have the disease
- Modified to our context:
  - The probability of p53- given death > 4 years

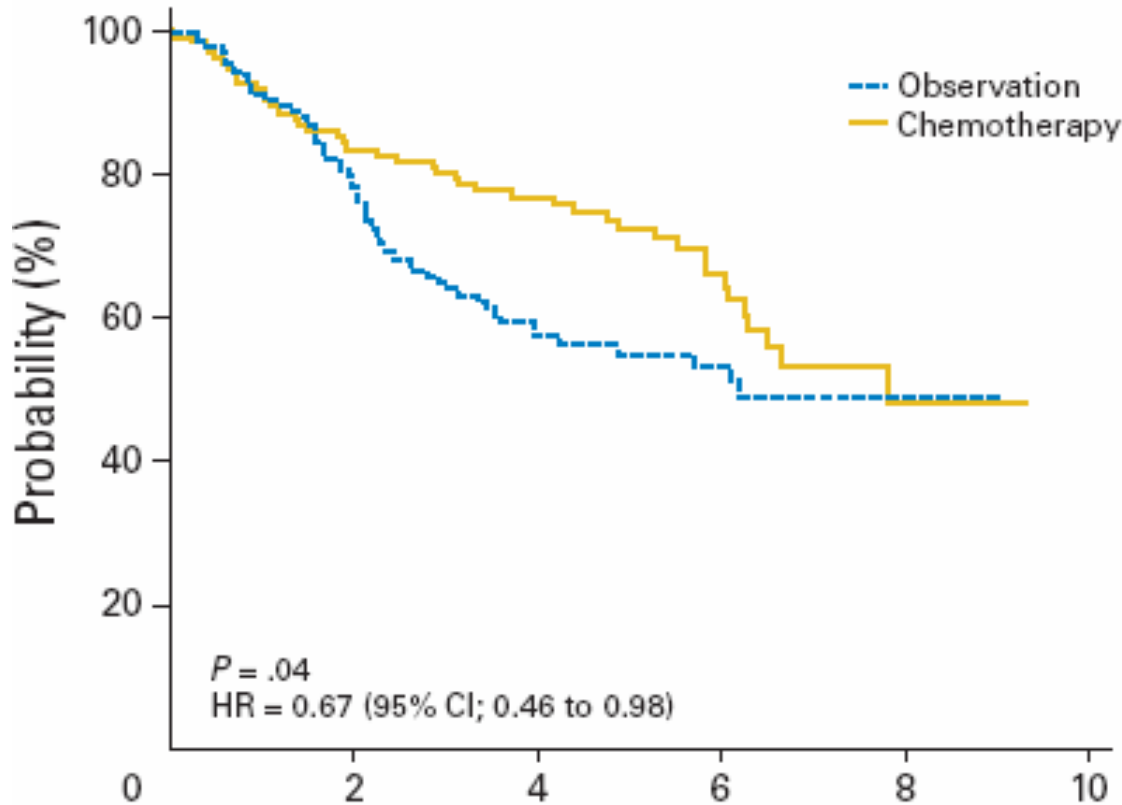
# Predicti

- Our example:
  - PPV = probability of d
  - NPV = probability of a



- More interesting PVs: within p53+ patients
  - probability of alive given chemo
    - **0.75** at 4 years
  - probability of alive given observation
    - **0.50** at 4 years

# What about p53 wildtype?



At 4 years?

At 7 years?

Do you believe  
7 year data?

No. at risk

Observation	136	106	55	27	10	0
Chemotherapy	137	114	78	36	9	0

# Prediction Models in Survival Analysis

- Commonly seen these days
- “Nomograms”
- Example: Prostate Cancer\*

---

## Variable

---

Chemotherapy type (vs q3w docetaxel)

Weekly docetaxel

Mitoxantrone

Liver metastases

No. metastatic sites (>2 vs ≤2)

Pain at baseline

Performance status (≤70 vs ≥80)

Progression type

Measurable disease

Bone scan progression

Baseline PSADT (<55 vs ≥55 d)

Baseline log PSA [for every unit rise in log(PSA) in ng/dL]

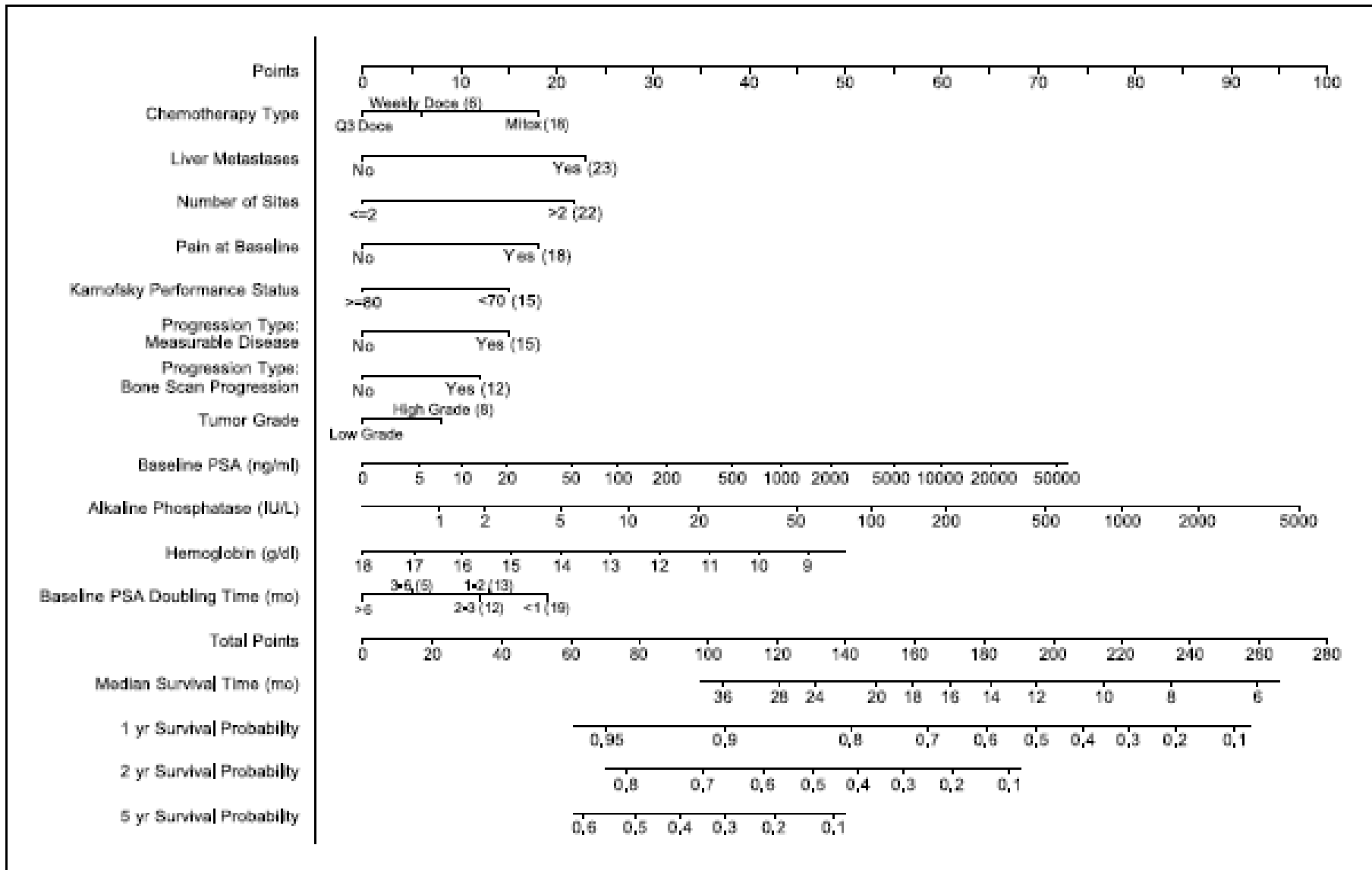
Tumor grade (Gleason ≥8 or WHO 3-4 vs Gleason ≤7 or WHO 2-3)

Alkaline phosphatase, log scale (per log unit rise, IU/L)

Hemoglobin (per unit decline, g/dL)

---

\* Armstrong et al. Clinical Cancer Research, Nov 2007.



# Is that all?

- You need to show that your model has strong predictive ability!
- Binary outcomes: AUC
- Survival endpoints: C-index
  - Ranges from 0 to 1
  - $C = 0.5$  implies prediction is essentially random
  - Generally, around 0.70 is minimally acceptable for a useful prediction tool
- Tsao et al.? Didn't approach predictive power
- Armstrong et al.:  $C = 0.69$