Clinical Trials Statistics for Non-Statisticians

Constance T. Cirrincione, MS
CALGB Statistical Center

November 2009

Components of Research

1. Formulate hypotheses
2. Define sample
3. Design the study
4. Analyze the data
5. Report the results
   – Aim of our research is to evaluate cancer treatment

Planning a Study: How It Starts

Investigator → Idea !

New treatment = Rx E(xperimental)

Standard treatment = Rx S(tandard)

Planning a Study: Quantifying the Idea

- Aim of the study
  - Compare Rx E with Rx S
  - Rx S is basis of comparison for Rx E
- Study hypothesis
  - Rx E will be better than Rx S
- Define study endpoint
  - How will 'better' be measured?

Planning a Study: Defining Endpoints

What do we mean ‘is better’ ?

Efficacy = effective = it works !

- How long patient lives (survival)
- How tumor responds (tumor response)
- If/when disease returns (disease recurrence)

Planning a Study: Identifying a Target Population

Want to make conclusions about population

- Usually cannot assess population
- Take sample of population
  - Random or at least representative
- Generalize to population
- Probability of being correct/incorrect

Planning a Study: Identifying a Target Population

- Cannot randomly select patients
- Can randomize treatment
- Can stratify

Designing the Study

Target pop: Patients with metastatic breast cancer

Endpoint: Tumor response [TR] (Response/non-response)

Given: TR of Rx S is 20%

Want to test: TR of Rx E is 40%

Number of patients

Error probabilities

Designing the Study

- Hypotheses
- Error probabilities – margin of error
- Power probability
- Number of patients (N)

- All components are interrelated

Designing the Study: Setting Error Probabilities

Decision	Truth	
	No difference	**Difference**
Difference	**alpha**	**power (1-beta)**
No difference	**confidence (1-alpha)**	**Beta**

Designing the Study

- Probability of being incorrect
  - Decide there IS a difference when there really IS NOT
  - Alpha → small
- Probability of being correct
  - Decide there IS a difference when there really IS
  - Power → large

Designing the Study: Example 1

- Testing new drug *in vitro*
- If ‘significant’, continue development
- If not, abandon drug

Designing the Study: Example 1

Decision	Truth	
	Drug not effective	**Drug is effective**
Is effective	**alpha = do more experiments**	**power (1-beta)**
Not effective	**confidence (1-alpha)**	**Beta = abandon effective drug**

Strategy = Set larger alpha, smaller beta
(→ larger power)

Designing the Study: Example 2

- Testing new drug in patients
- Several existing alternative Rxs
- If ‘significant’, market drug
- If not, abandon drug

Designing the Study: Example 2

Decision	Truth	
	Drug not effective	**Drug is effective**
Is effective	**alpha = use useless drug, deprive good standard**	**power (1-beta)**
Not effective	**confidence (1-alpha)**	**Beta = abandon effective drug (but others available)**
Strategy = Set smaller alpha		

Designing the Study

Aim is to make a definitive decision
Therefore, set cutoff that will define significance'

Set
- Alpha small (incorrect)
- Power large (correct)

Convention
- Alpha 0.05 – 0.10
- Power 0.80 +

Designing the Study

Given 20%

To test 40%
with alpha = 0.05
and power = 0.80

How many patients do we need?

Designing the Study

Given 20%

To test 40%
with alpha = 0.05
and power = 0.80

How many patients do we need?
Requires 80 pts/Rx

Designing the Study: Power & N

Given 20%

To test	40%	40%
alpha =	0.05	0.05
power =	0.80	0.90
Pts/Rx	80	107

Designing the Study: Power & N

Given 20%

To test	40%	40%	40%
alpha =	0.05	0.05	0.05
power =	0.80	0.90	0.95
Pts/Rx	80	107	133

Designing the Study: Alpha & N

Given 20%

To test	40%	40%
alpha =	0.05	0.10
power =	0.80	0.80
Pts/Rx	80	63

Designing the Study: Alpha & N

Given 20%

To test	40%	40%	40%
alpha =	0.05	0.10	0.15
power =	0.80	0.80	0.80
Pts/Rx	80	63	53

Designing the Study: Effect Size & N

Given 20%

To test	40%	50%
alpha =	0.05	0.05
power =	0.80	0.80
Pts/Rx	80	38

Designing the Study: Effect Size & N

Given 20%

To test	40%	50%	60%
alpha =	0.05	0.05	0.05
power =	0.80	0.80	0.80
Pts/Rx	80	38	22

Analyzing the Data: P-value

- Calculated from data
- Probability of getting a result as big or bigger than what we've gotten *if truly no difference*
- Convention for 'significance'

    0.05 or smaller

Analyzing the Data: P-value

We found: TR of Rx S = 21% P=0.02

TR of Rx E = 39%

- *If truly no difference,* repeating the trial many times would give a result as or more extreme than our trial (<21% vs >39%) 2% of the times.
- There is a 2% probability that results are caused by random chance.

Analyzing the Data: Confidence Interval (CI)

- Calculated from data
- Measure of precision
- Convention is 95% CI
- If we were to repeat the trial many times, then the true value would be included in the interval 95% of times.

Analyzing the Data: Confidence Interval

- We found:

  TR of Rx E was 39%

  with 95% CI = 28% – 50%

- If we were to repeat the same trial many times, 95% of times the true TR would lie somewhere between 28% and 50%.
- The true TR would be <28% or >50%.

Analyzing the Data: Confidence Interval

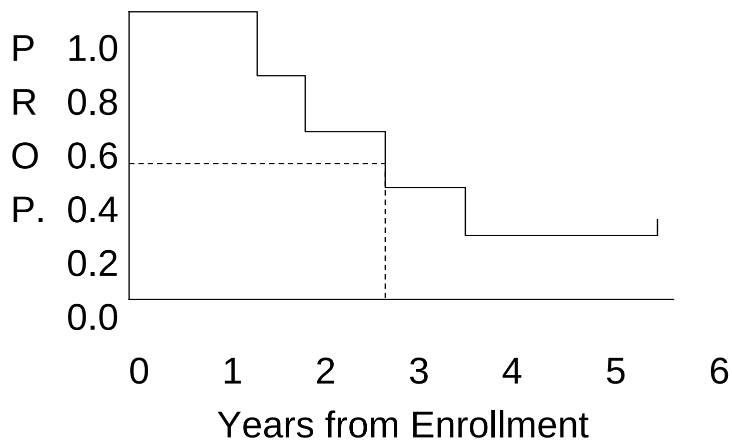
- Role of sample size
- Study reports TR of 65%

# Pts	95% CI
60	52 – 77%
250	59 – 71%
500	61 – 69%
3000	63 – 67%

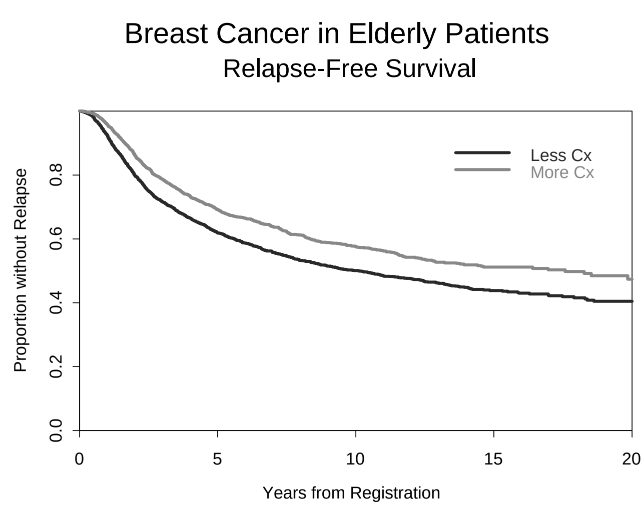
Analyzing the Data: Confidence Interval

- All things remaining equal [eg, hypotheses, alpha] increasing the sample size will result in a narrower interval.

Analyzing the Data: Time to Clinical Event Survival



Analyzing the Data



Analyzing the Data: Subset Analyses

- Examples of subset analyses
  - Patients with ER-positive tumors
  - Patients < 50 years
- Advantages
  - Explore patterns
  - Use results as basis for future studies
- Concerns
  - Study often not powered for subset analyses

Designing the Study

Given 20%

To test 40%
with alpha = 0.05
and power = 0.80

How many patients do we need?
Requires 80 pts/Rx = 160

Analyzing the Data

Given 20%

To test 40%
with alpha = 0.05
and power = 0.80

Requires 80 pts/Rx = 160

Subset analysis on age < 50
40 pts/Rx = 80

Designing the Study

Given 20%

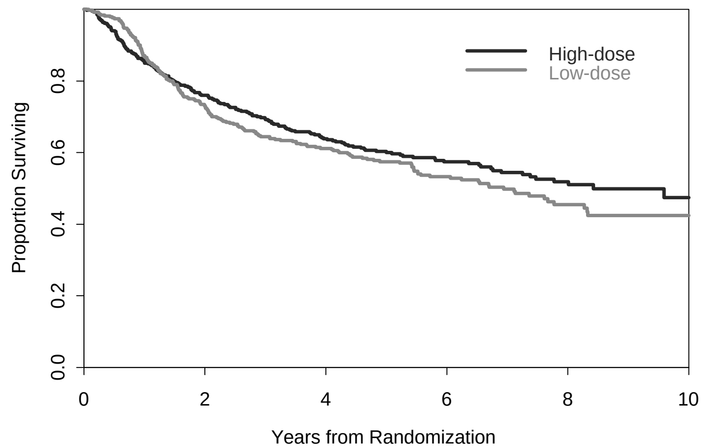
To test 40%
with alpha = 0.05
and power = 0.80 → 0.51

Subset analysis
40 pts/Rx = 80

Just as good as flipping a coin !

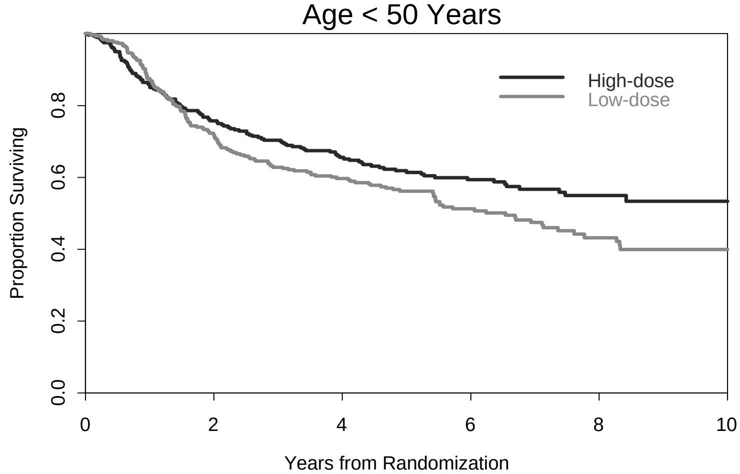
Reporting the Results

Chemotherapy in Breast Cancer
Overall Survival

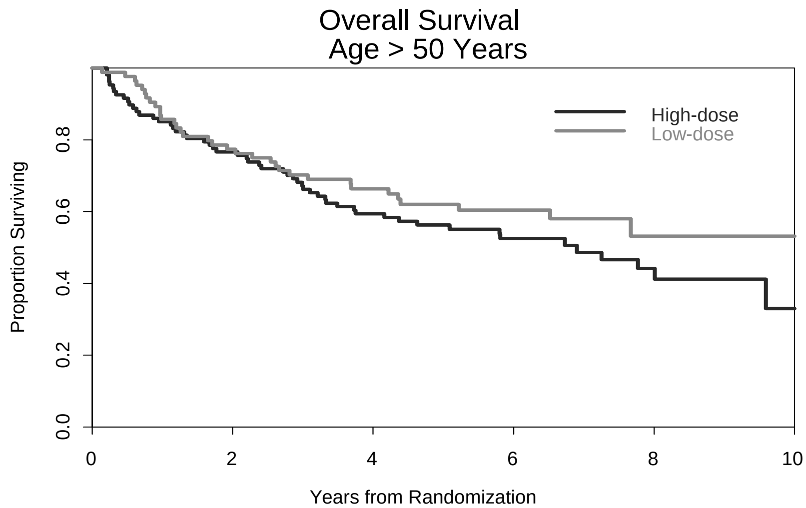


Reporting the Results

Overall Survival
Age < 50 Years



Reporting the Results



Reporting the Results

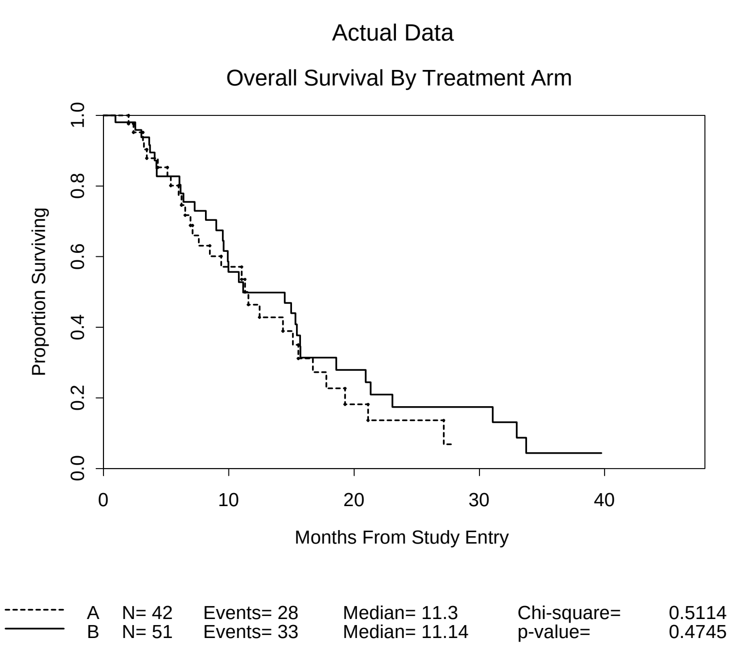
Results: Overall survival was not significantly different between the two treatment groups (P=0.24).

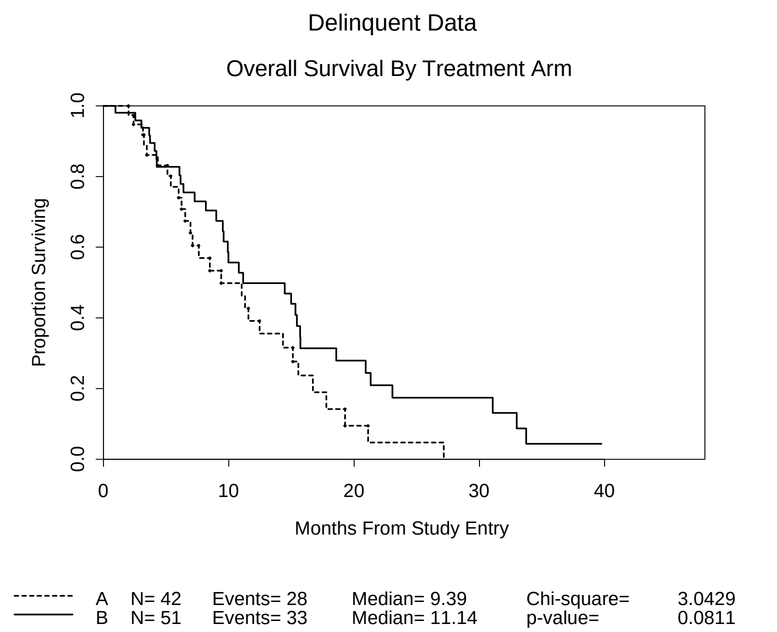
There appears to be a significant interaction on survival between treatment and age. Women younger than 50 randomized to high-dose experienced a longer time to disease recurrence compared to women in the same age category randomized to low-dose (P=0.02).

Summary: High-dose was not superior to std low-dose for overall survival in women with high-risk breast cancer. However, among women under 50 years of age, high-dose had improved survival.

Data Maturity/Currency: Tumor Response

	Rx A N=40	Rx B N=40
RESPONSE	50%	50% ?
		63% ?
		38% ?





Reference

Motulsky, H. (1995). *Intuitive Biostatistics*, Oxford University Press, New York, NY.