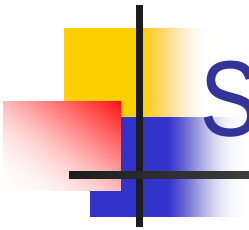
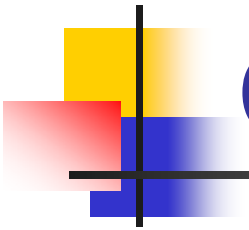


Statistical Designs of Clinical Trials

Instructor: Jen-pei Liu, Ph.D.
Division of Biometry
Department of Agronomy
National Taiwan University
&
Division of Biostatistics and Bioinformatics
National Health Research Institutes

at TCOG Nurse Training Program
August 26, 2006
Taichung, Taiwan



Outlines

- Introduction
- Designs for Exploratory Trials
  - Phase I Trials
  - Phase II Trials
- Designs for Confirmatory Trials
  - Phase III Trials
- Types of Trials
- Discussion and Summary

INTRODUCTION

Clinical Trials

- FDA (21 CFR 312.3, April 1994)
  A clinical trial is the clinical investigation of a drug which is administrated or dispensed or used involving one or more human subjects.

- Chow and Liu (July 1998)
  A clinical trial is a clinical investigation in which treatments are administrated, dispensed or used involving one or more human subjects for evaluation of the treatments

Three Key Components

- **Experiment Unit**
  - A subject from a targeted population under study.
- **Treatment**
  - It could be a placebo or any combinations of a new pharmaceutical entity, a new diet, a surgical procedure, a diagnostic test, a medical device, or no treatment .

Three Key Components

- **Evaluation**
  - Efficacy analysis
    - Clinical endpoint
  - Safety assessment
    - Adverse events
    - Laboratory test results
  - Quality of life assessment
  - Pharmacoeconomic analysis
  - Outcomes research

INTRODUCTION

Bias and Precision

Goal : To make a **unbiased** inference with the possibly **best precision** to scientifically answer clinical questions with respect to a **targeted patient population**.

(1) To minimize bias.

(2) To maximize precision.

INTRODUCTION

- Phases of Clinical Development
  - Phase I: Safety: < 100 subjects
  - Phase II: Efficacy: hundreds patients
  - Phase III: Confirmation of safety and efficacy: hundreds to tens of thousands
  - Phase IV: After approval

    Additional evaluation of long-term safety and efficacy in other patient population

Designs for Exploratory Trials - Cancer Phase Ⅰ Studies

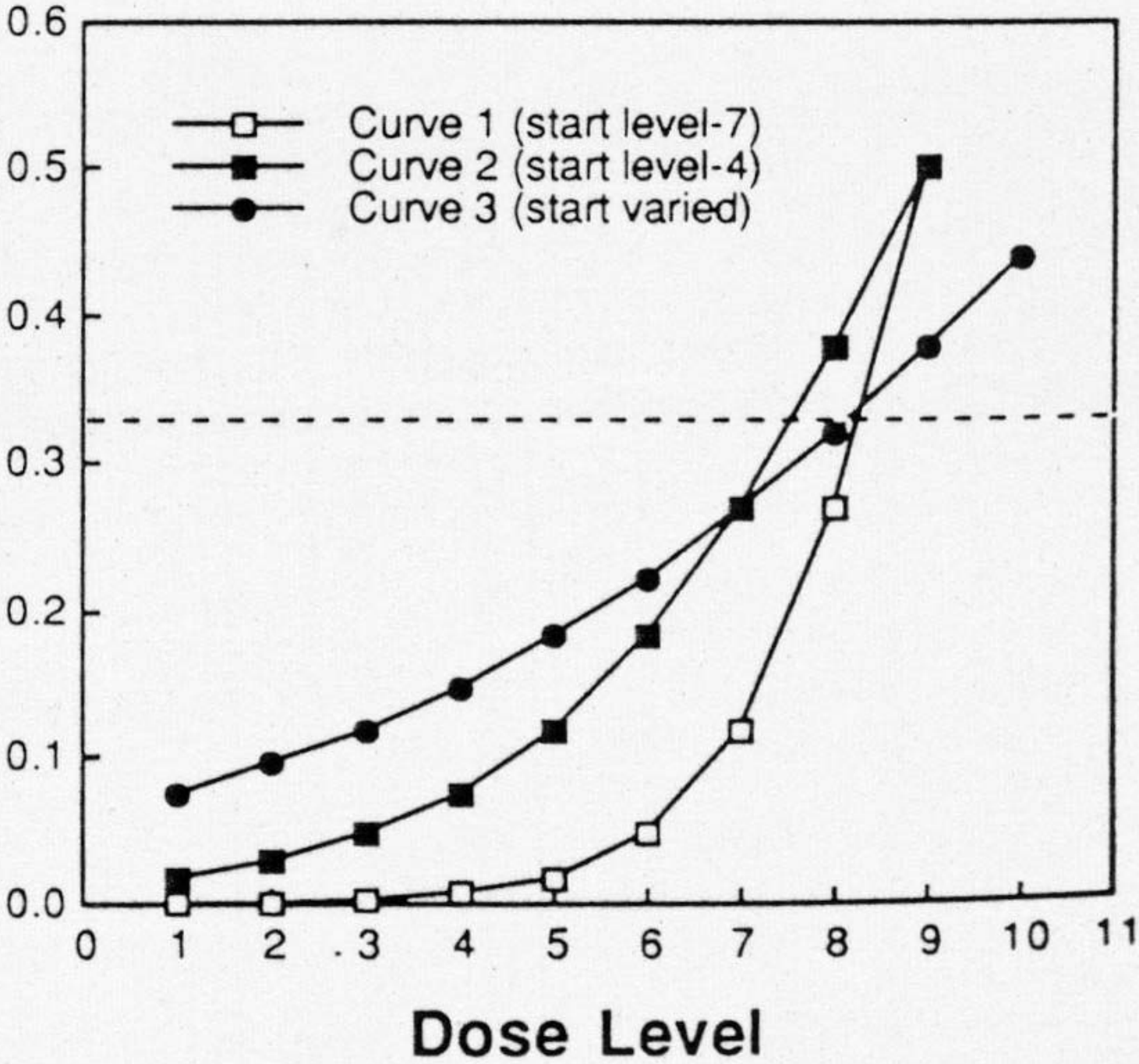
- Designs for determination of maximum tolerable dose
  - For Phase Ⅰ cancer chemotherapy
  - Pre-selected fixed dose levels
- Maximum Tolerable Dose (MTD)
  - Quantitative Definition

    Some percentile of a tolerance distribution w.r. to some definitive dose-limiting clinical toxicity

    Storer (1989), Korn, et al (1994)

$$\log it[P(x,\theta)] = \alpha + \beta x, \quad where\ \theta = (\alpha, \beta)'$$

$$MTD = X_m = (k_q - \alpha)/\beta, \quad where\ k_q = \log it(q)$$


Curve 1 (start level-7)
Curve 2 (start level-4)
Curve 3 (start varied)
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
1
2
.3
4
5
6
7
8
9
10
11
Dose Level

Standard Dose Escalation Design

- Step 1

  A group of three patients are treated with the initial dose

- Step 2

  If no toxicity is observed in all three patients, then the dose for the next group of three patients is escalated to the next higher dose level. Otherwise, the next group of three patients is treated at the same dose.

Standard Dose Escalation Design

- Step 3

  The dose of the next group of three patients is escalated to the next higher dose level, if the pre-specified clinical toxicity is observed at most one patient of the six patients from both step 1 and 2, otherwise the trial stops.

- Step 4

  Repeat step 2 and 3 two consecutive groups of three patients until the trial stops.

- Traditionally from the standard design, the MTD is defined to be either the dose at which the trial stops or the next lower dose.

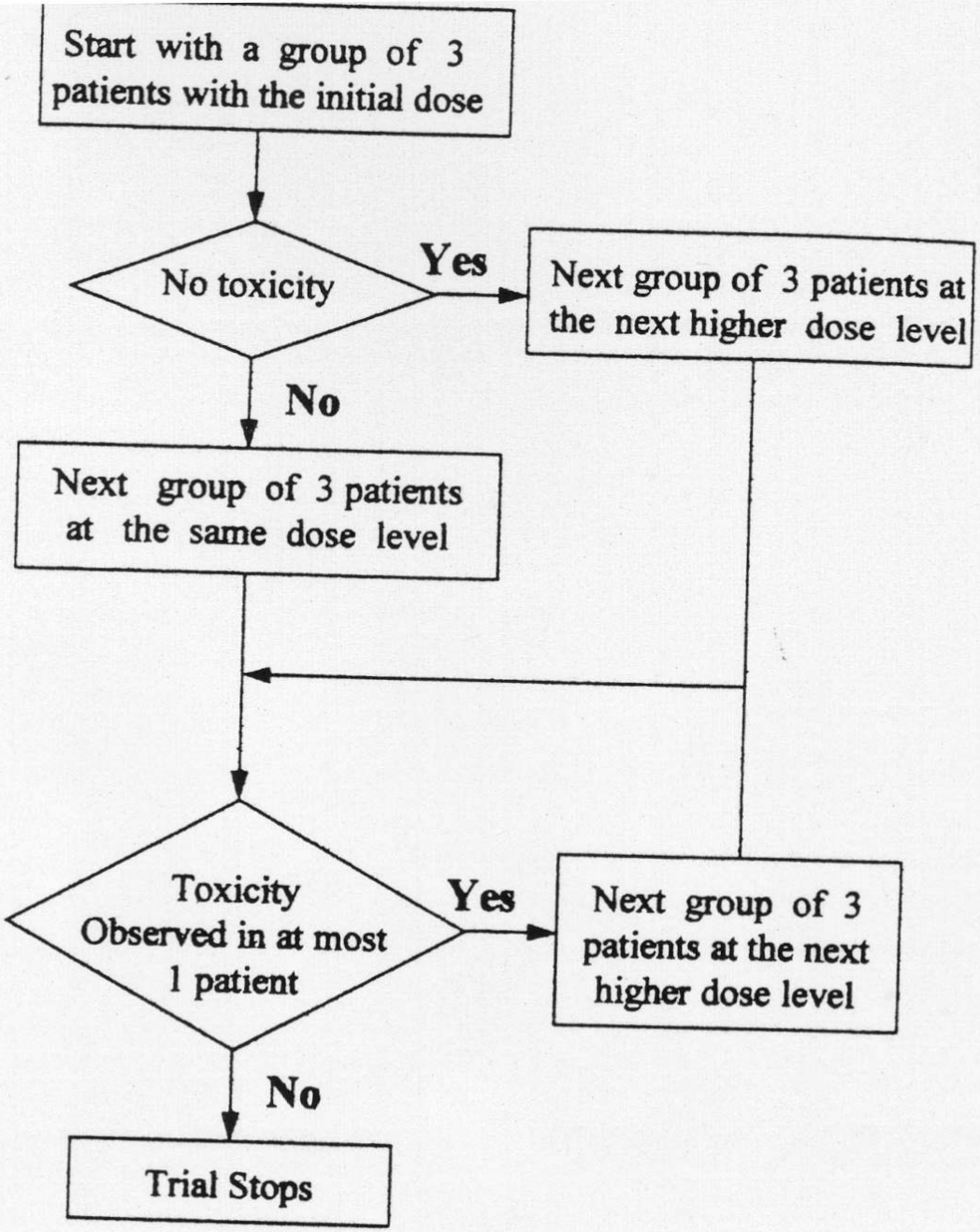


Figure 5.4.1 Flowchart for design *A*.

Drawbacks of the Current Practice for Standard Design

- No room for de-escalation
- No further analysis of data
- No objective estimation of MTD with statistical models
- No sampling error and no confidence interval
- Slow escalation

Accelerated Titration Designs Richard Simon (1997)

- Rationale
  - Address the flaws of traditional designs
  - Attempt to obtain information about inter-patient variability and cumulative toxicity - stay for 3 courses to allow for intra-patient dose modifications
  - Distinguish between moderate and dose-limiting toxicities

Accelerated Titration Designs Richard Simon (1997)

Scheme

- The first stage

  1 patient per level until 1 DLT or 2 moderate toxicities

- The second stage

  Traditional design, i.e. add 2 patients to the current dose that triggered the switch.

Accelerated Titration Designs Richard Simon (1997)

- MTD

  Estimated as the highest dose where at most 1/6 patients developed DLT

Compared to traditional designs

- Go through the lower doses quickly, and thus reduces under-treated patients in absolute sense and speed up the completion
- Obtain similar estimate of MTD
- Provide more information. Upon completion, a model can be fitted to estimate inter- and intra-patient variability
- Require careful patient management to track the toxicity over multiple course

Issues of Phase Ⅰ Designs for MTD

- Complete the trials with
  - Minimum amount of patients, and
  - Minimum amount of time
- Recognize differential dosing-limiting clinical toxicity
- Include a stopping rule to allow flexibility to extend to higher or lower dose levels
- Investigators and regulatory agencies dictate the dose level for the first patient

Cancer Phase II Trials

- Allows early termination for inactivity or high activity.
- Define
  - P_0: undesirable level (lower bound)
  - P_1: target level

The rationale is based on the hypothesis testing

$$H_0 : p \le p_0 \ against \ H_1 : p \ge p_1$$

and the error limits : $Type\ I \le \alpha, Type\ II \le \beta$

Simon's design

- Procedure
  - Stage 1: If $X_1 > r_1 \Rightarrow$ go to stage 2
    $\leqq r_1 \Rightarrow$ stop and reject the drug
  - Stage 2: If X_1+X_2 is $\leqq r \Rightarrow$ reject the drug
    $> r \Rightarrow$ accept the drug
- Given p_0, p_1, α, β, then (n_1, n_2, r_1, r) are optimized to minimize either
  - The expected sample size under p_0, or
  - The maximal sample size $n_1 + n_2$
- Not readily evaluable, but tables of designs under different values of parameter are available from the paper.

Designs for $p_1 - p_0 = 0.20$

		Optimal Design				Minimax Design			
		Reject Drug if Response Rate				Reject Drug if Response Rate			
P_1	p_0	≦ r1/n1	≦ r/n	EN(p0)	PET(p0)	≦ r1/n1	≦ r/n	EN(p0)	PET(p0)
0.05	0.25	0/9	2/24	14.5	0.63	013	2/20	1634	0.51
		0/9	2/17	12.0	0.63	012	2/16	13.8	0.54
		0/9	3/30	16.8	0.63	015	3/25	20.4	0.46
0.10	0.30	1/12	5/35	19.8	0.65	1/16	4/25	20.4	0.51
		1/10	5/29	15.0	0.74	1/15	5/25	19.5	0.55
		2/18	6/35	22.5	0.71	2/22	6/33	26.2	0.62
0.20	0.40	3/17	10/37	26.0	0.55	3/19	10/36	28.3	0.46
		3/13	12/43	20.6	0.75	4/18	10/33	22.3	0.50
		4/19	15/54	30.4	0.67	5/24	13/45	31.2/	0.66
0.30	0.50	7/22	17/46	29.9	0.67	7/28	15/39	35.0	0.36
		5/15	18/46	23.6	0.72	6/19	16/39	25.7	0.48
		8/24	24/63	34.7	0.73	7/24	21/53	36.6	0.56
0.40	0.60	7/18	22/46	30.2	/0.56	11/28	20/41	33.8	0.55
		7/16	23/46	24.5	0.72	17/34	20/39	34.4	0.91
		11/25	32/66	36.0	0.73	12/28	27/54	38.1	0.64
0.50	0.70	11/21	26/45	29.0	0.67	11/23	23/39	31.0	0.50
		8/15	26/43	23.5	0.70	12/23	23/37	27.7	0.66
		13/24	36/61	34.0	0.73	14/27	32/53	36.1	0.65
0.60	0.80	6/11	26/38	25.4	0.47	18/27	24/35	28.5	0.82
		7/11	30/43	20.5	0.70	8/13	25/35	20.8	0.65
		12/19	37/53	29.5	0.69	15/26	32/45	35.9	0.48
0.70	0.90	6/9	22/28	17.8	0.54	11/16	20/25	20.1	0.55
		4/6	22/27	14.8	0.58	19/23	21/26	23.2	0.95
		11/15	29/36	24.2	0.70	13/18	26/32	22.7	0.67

Designs for Confirmatory Trials – Phase III Trials

- **Parallel Group Designs**

  The patients are randomized to one of two or more arms, each arm being allocated to a different treatment.

  - Advantages：
    - Simple and easy to implement.
    - Less complicated analysis and interpretation.
  - Drawbacks：
    - Relative large variability
    - Inter-patient + Intra-patient

Crossover Design

Each subject is randomized to a sequence of two or more treatments.

- Advantages
  - Subjects act their own control for treatment comparison
  - Reduction of variability
  - Fewer patients required
  - Standard design for bioequivalence trials of generic drug products

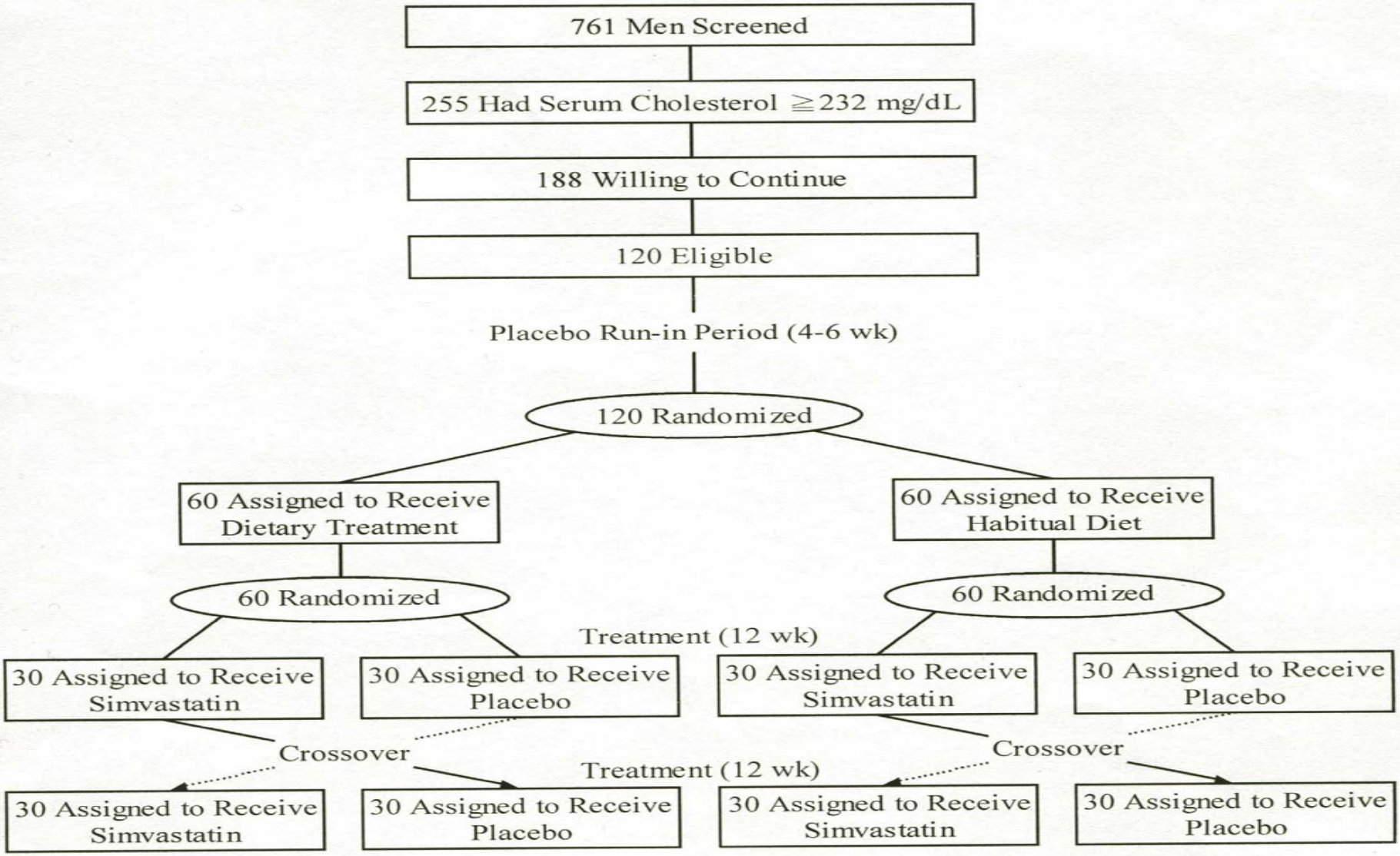


Figure 5.4.3 Joint Application of Parallel-group and Crossover Designs.

Source: Jula, et.al. (2002)

Factorial Design (Combination Trials)

- Two or more treatments are evaluated simultaneously in the same sets of patients via various of combinations of two treatments.

- ***Example*** :
  FACET International Study Group
  (NEJM 1997;337:1405-11)

Factorial Design (Combination Trials)

- ***Example :***
  - Effect of inhaled formoterol and budesonide on exacerbation of asthma
  - Double-blind, randomized, parallel group
  - 4-week run-in period with budesonide with 800 μg bid followed by 12-month DB, randomized treatments
- Fixed dose-combination: A+B

Factorial Design (Combination Trials)

- **Example :**

Treatment	Budesonide	Formoterol
I	100 μg bid	Placebo
II	100 μg bid	12 μg bid
III	400 μg bid	Placebo
IV	400 μg bid	12 μg bid

Types of Trials – Multicenter Trials

- A multicenter study is a single study conducted under a common protocol, involving several centers (e.g., clinics, practices, hospitals) where the data collected are intended to be analyzed as whole (as opposed to a post-hoc decision to combine data or results from separate studies)

Types of Trials – Multicenter Trials

- **Goals**：
  - To accrue patients efficiently (all stages)
  - To provide a basis for generalization of its findings (later phases)
  - Generalizability：
    The extent to which the findings of a clinical trial can be reliable extrapolated from the subjects who participated in the trial to a broader patient population and a broader range of clinical settings.

- ***Examples :***

Tacrine in Alzheimer's disease

(Farlow, et al. JAMA 1992;268:2523-2529)

A multinational and multicenter trial

- Targeted population :
  - 468 randomized patients
  - 50-89 years old
  - Criteria by NINCDS-ADRDA
  - Min-Mental State Examination
  - (MMSE) score: 1-26
- No. Centers: 23 in 2 countries

- ***Issues :***
  - Variations in implementing protocol
    - Common protocol
    - Standardization of procedures
    - Pre-study investigator's meeting
    - Training of personnel
    - Careful monitoring
  - Variation in the number of patients
    - Few small centers vs. lots of large centers
    - Few large centers vs. lots of small centers
    - All small centers

- ***Issues :***
  - Heterogeneity of treatment effects
    Treatment-by-center interaction
    - Interaction
      The situation in which a treatment contrast (difference between investigational product and control) is dependent on another factor (center)
    - Quantitative
      The magnitude of contrast differs at the different levels of the factor

      Estimates of treatment effect in same direction
    - Qualitative
      The direction of the contrasts differs for at least one level of the factor
      May require additional clinical trials

Types of Trials – Targeted Clinical Trials

- Examples
- *HER2/neu* gene in metastatic breast cancer - Herceptin® - requirement of screening the patients with over-expressed *HER2* level (Slamon, 2001).
- Estrogen receptor ploymorphism - Estrogen Replacement Atherosclerosis trial (ERA, Herrington, et al, 2002): a total of 9 SNPs were identified and interaction between treatment of HRT and some of SNPs in elevation of lipid levels is suggested
- Sample size determination: Fijal, et al. (2000) and Maitournam and Simon (2005).

Types of Trials – Targeted Clinical Trials

- Other examples
- The epidermal growth factor receptor (EGFR) inhibitor for the non-small cell lung cancer.
- Iressa (gefitnib) and Tarceva (Erlotinib) are targted at the EGFR pathway.
- Efficacy is correlated to

  race

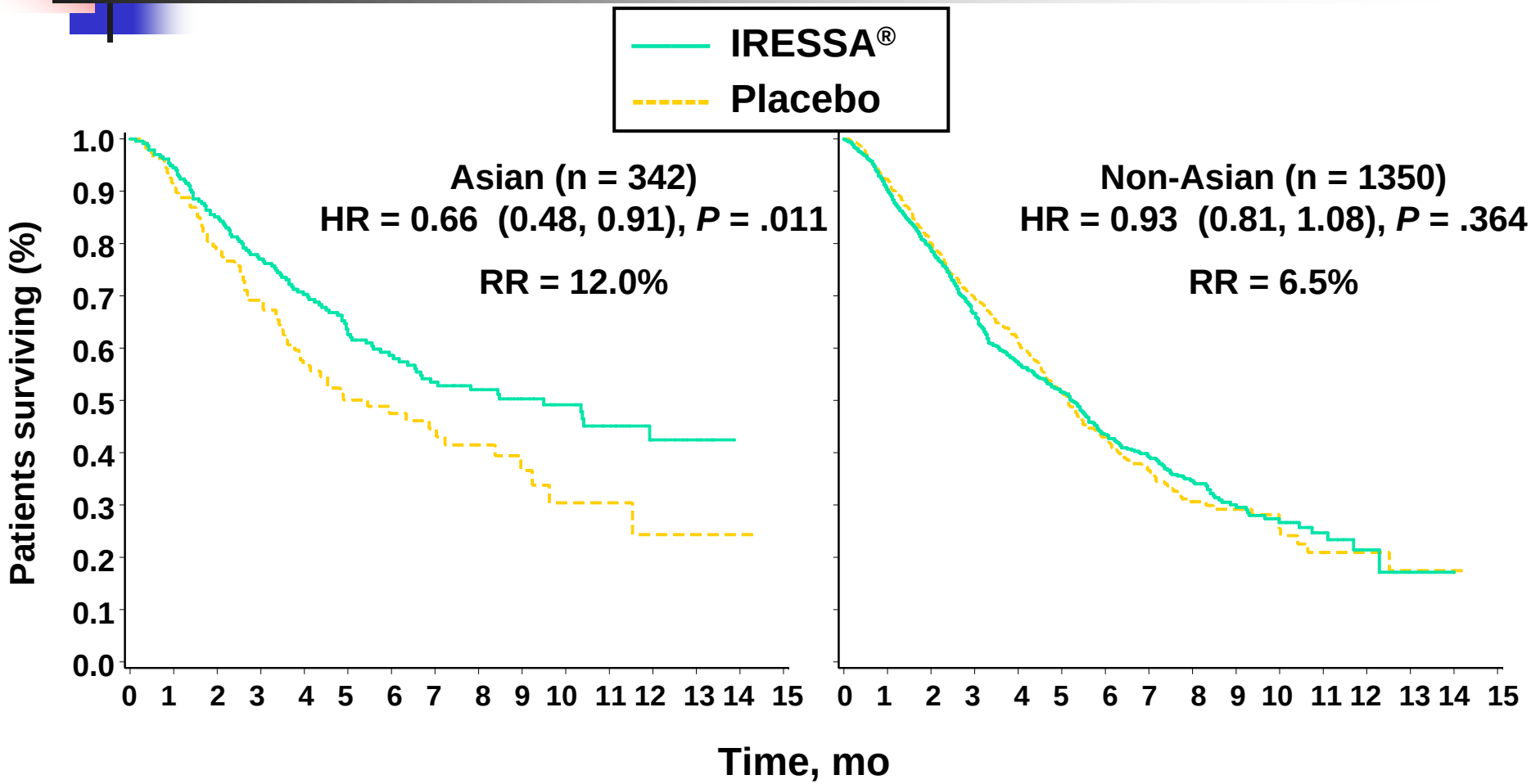
  number of gene copies

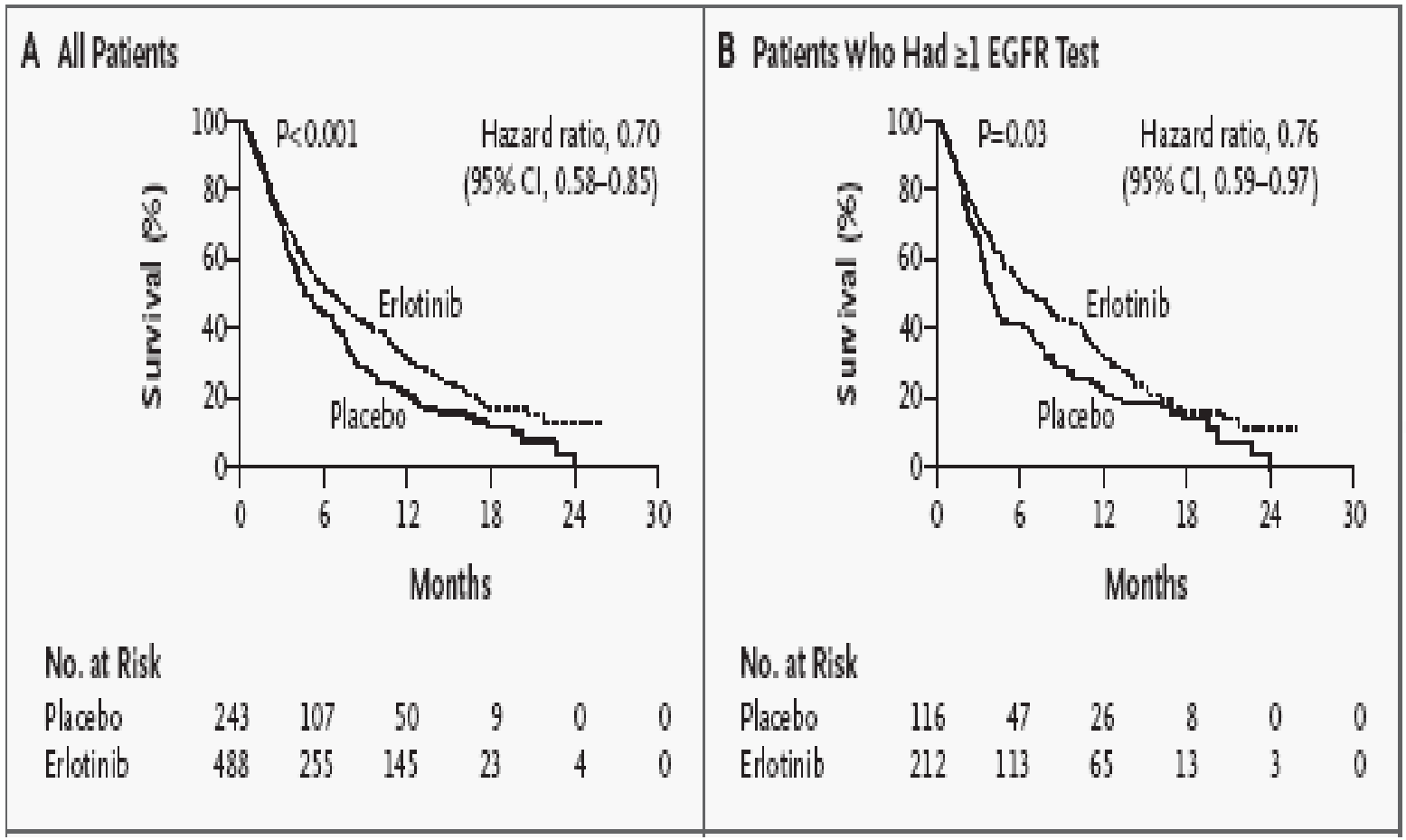
  protein expression

  EGFR mutation

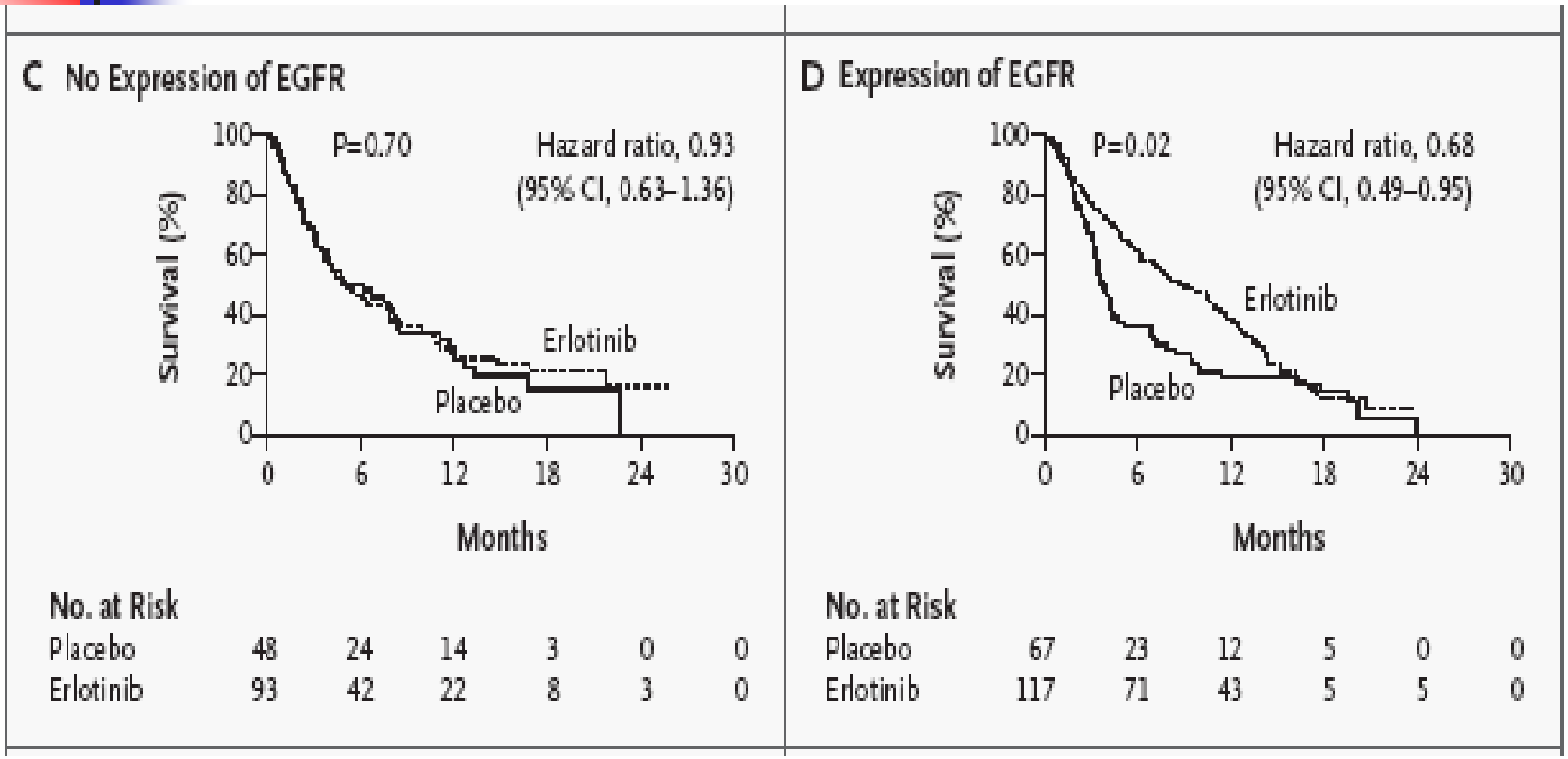
Gappuzzo et al. (JNCI, 2005), Tsao, et al (NEJM, 2005)

Survival by Ethnic Origin

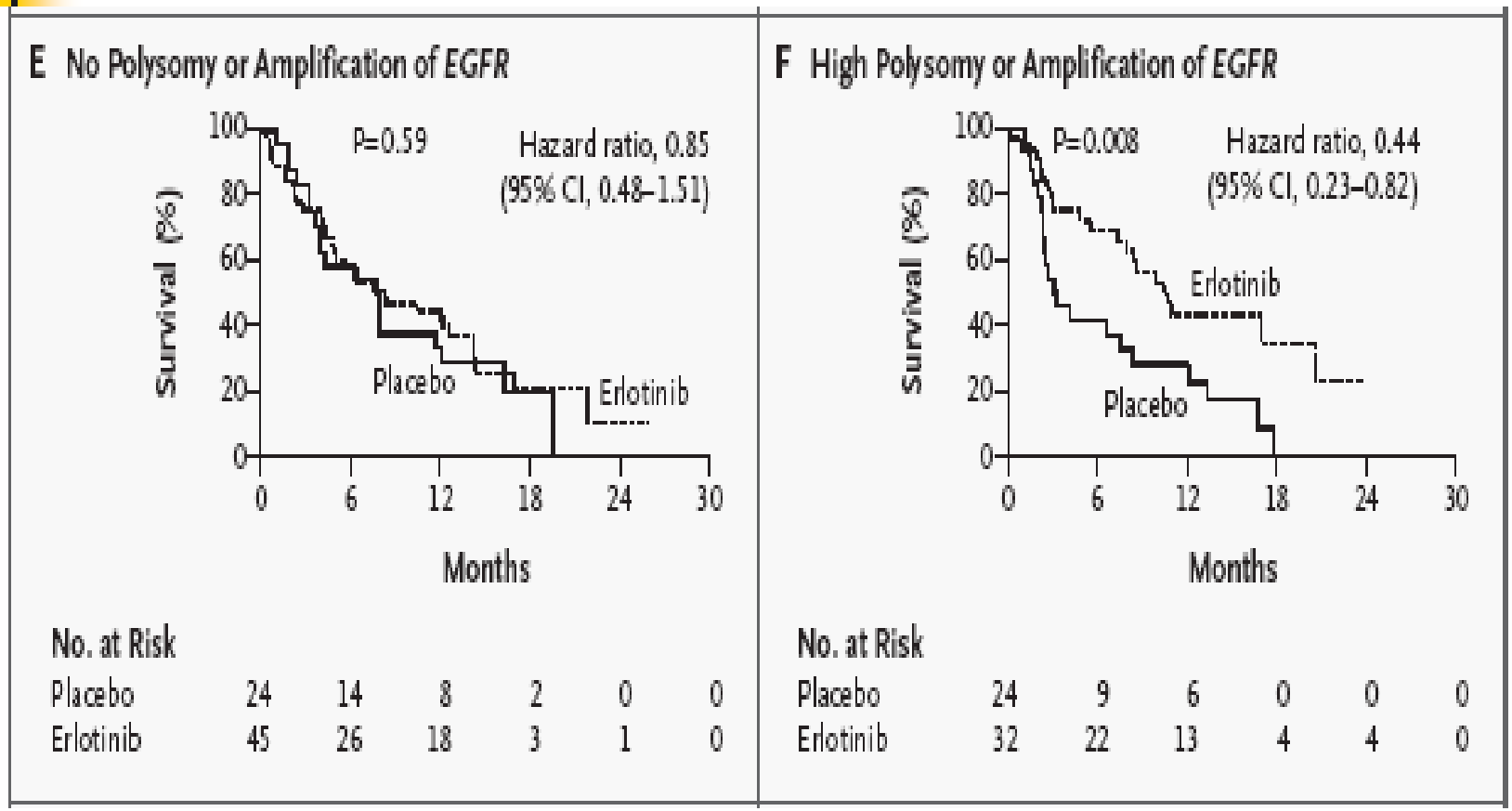




From: Tsao, et al (2005, NEJM)



From: Tsao, et al (2005, NEJM)



From: Tsao, et al (2005, NEJM)

Types of Trials – Superiority Trials

The objective of the trial is to establish the efficacy by demonstrating that the test treatment is superior to

- a concurrent placebo control
- a concurrent active treatment control

or

- a dose-response relationship

Equivalence or Non-inferiority Trials

The objective of the trial is to show that the efficacy of the test treatment is either

- similar (or equivalent) to or
- no worse than the concurrent active treatment control.

Equivalence or Non-inferiority Trials

- ***Equivalence Trial***
  A trial with the primary objective of showing that the response to two or more treatments differs by an amount which is clinically unimportant
- ***Non-inferiority Trial***
  A trial with the primary objective of showing that the response to the investigational product is not clinically inferior to a comparative agent (active or placebo control)

Joint Applications of Superiority and Equivalence Trials

- Moseley, et al. NEJM, 2002: 347:81-8
- Patient: Osteoarthritis of the knee
- **Design**: Randomized, parallel-group, Placebeo-controlled, evaluator-blind
- Treatments:
- Arthroscopic debridement (n=59)
- Arthroscopic lavage (n=61)
- Placebo surgery (n=60)

Joint Applications of Superiority and Equivalence Trials

- Objectives:
- (1) Superiority of the arthroscopic procedures over placebo surgery
  - **Results: NO**
- (2) If lack of evidence of superiority, equivalence of arthroscopic procedures to placebo surgery
  - **Results: YES**

Group Sequential Trials

A group sequential trial allows to evaluate the efficacy and safety of test treatment by means of interim analyses during the study for possible early termination based on convincing evidence of either benefit or harm before its scheduled completion

- Description of statistical methods and pre-planned interim analyses in the protocol with adjustment of p-values
- Documentation of everything
- Independent data monitoring committee

Women Health Initiatives (WHI)

- US Physician Health Study: An all men trial
- Too few women in clinical trials
- Women Health Initiatives (WHI) by NIH director B. Healy (a cardiologist) in 1991
- Designed to run for 15 years to recruit 160,000 women with a price tag of US $ 727 millions by 2007
- One of the components is the WHI trial

Women Health Initiatives (WHI)

- A randomized, DB placebo-controlled primary prevention trial
- Investigate the benefits and risks of hormone replacement therapy (HRT)
- Estrogen + progestin vs. Placebo
- 16608 healthy postmenopausal women aged 50-79 years with intact uterus recruited between 1993 and 1998 with expectation of final analysis in 2005 after an average of approximately 8.5 years of follow-up

Women Health Initiatives (WHI)

- A primary prevention trial
- Healthy subjects
- Low incidence, mortality and morbidity rates
- Large sample size and infeasible to repeat due to cost and long-term nature
- More comprehensive approach to monitoring the primary prevention trials
- Considerations of overall health benefit vs. risk into formal termination procedure

Women Health Initiatives (WHI)

- Primary efficacy endpoint

  incidence of CHD
- Primary safety endpoint

  incidence of invasive breast cancer
- Competing benefits and risks

  colorectal cancer, hip fracture

  stroke, pulmonary embolism, endometrial cancer, death due to other causes

Women Health Initiatives (WHI)

- A weight global (Freedman et al, 1996)

  $W = w_1 d_1 + \ldots + w_8 d_8$

  d_i: difference in proportions between the two groups for outcome i, i =1,…8

  w_i: weights for outcome i – expected proportion of diagnosed patients who will die of that disease within a specific years of diagnosis

Benefits and risks are not symmetric

Women Health Initiatives (WHI)

A mixed approach to early termination

1. O-B boundaries for each of eight outcomes and for global index
2. Asymmetric upper and lower boundaries:

   one-sided α = 0.025 for benefit

   one-sided α = 0.05 for adverse effects

   adverse-effect boundaries were adjusted using Bonferroni method
3. Trial stops if the upper or lower boundaries were crossed and the result from global index was supportive at 0.20 level

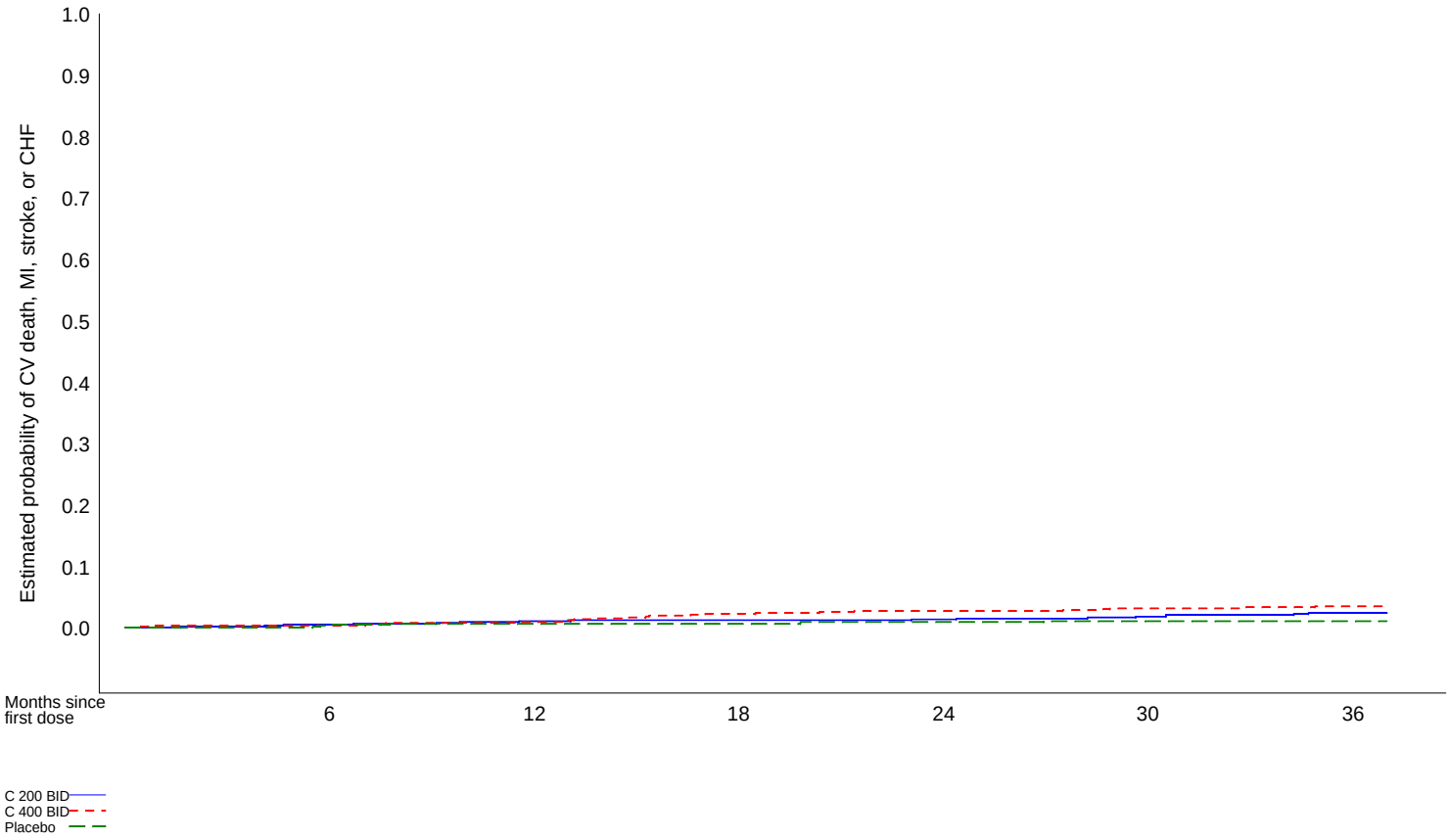
Women Health Initiatives (WHI)

Semiannual DSMB meeting since the fall of 1997

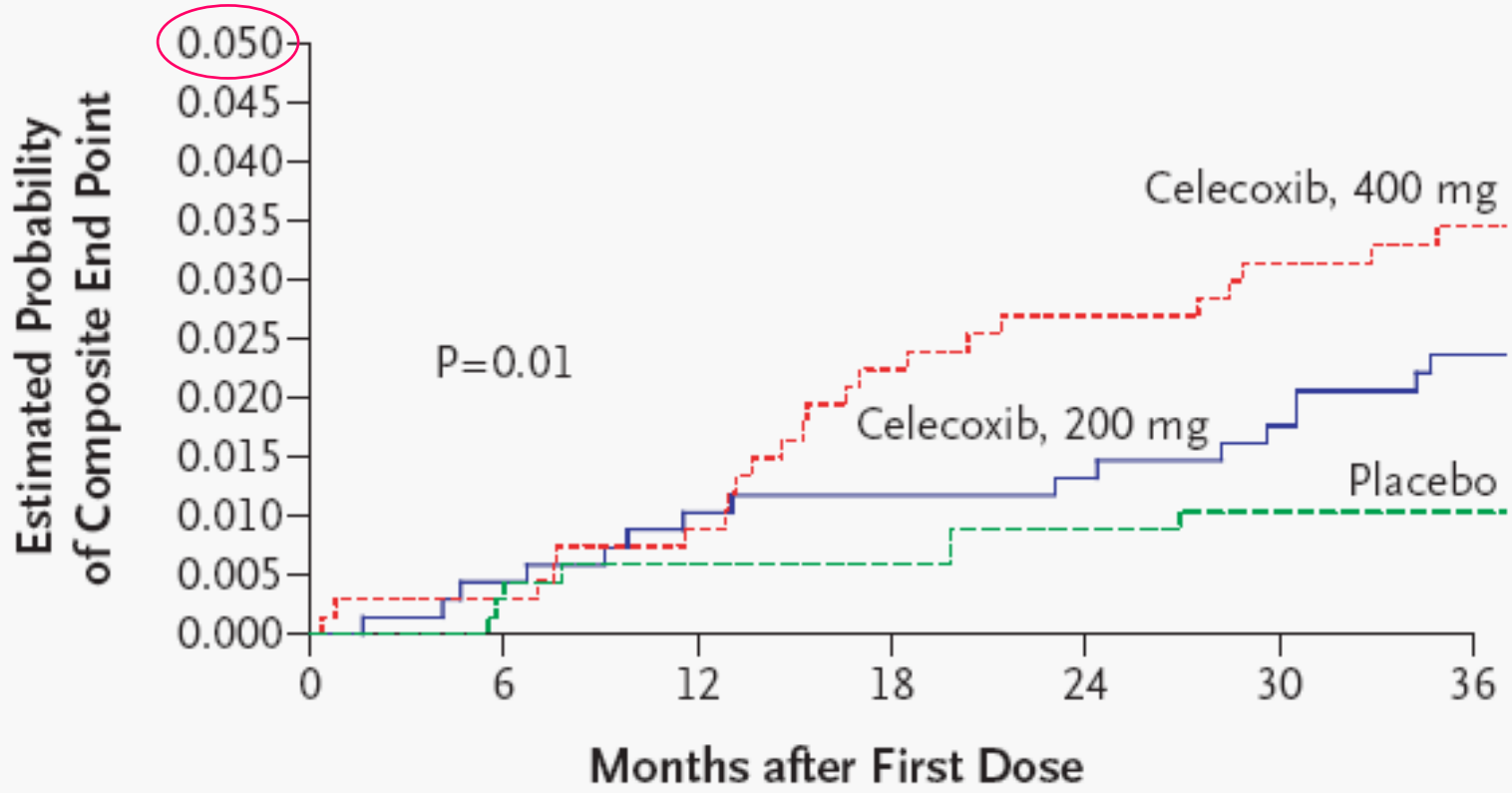
The tenth interim analysis on May 31, 2002

1. The weighted log-rank test statistic z = -3.19 crossed the lower boundary for adverse event z= -2.32
2. The global index is supportive (z = -1.62)
3. Additional evidence of risks on CHD, stroke, pulmonary embolism outweighed the evidence of benefit on hip fracture and colon cancer
4. DSMB recommended early termination of the estrogen plus progestin component of WHI

Kaplan-Meier Estimates of the Risk of Serious CV Events in the APC Trial by Treatment Arm*



Kaplan-Meier Estimates of the Risk of Serious CV Events in the APC Trial by Treatment Arm*



No. at Risk

Celecoxib, 400 mg	671	669	665	655	651	648	576
Celecoxib, 200 mg	685	681	676	675	673	670	595
Placebo	679	677	675	672	668	667	585

VI. DISCUSSION AND SUMMARY

- ***Bias***

  Controls, Blinding, randomization

- ***Statistical Designs***
  - Phase I for MTD
  - Phase II for Activity
  - Parallel
  - Crossover
  - Factorial
  - Group Sequential

VI. DISCUSSION AND SUMMARY

- ***Types of Trials***
  - Multicenter
  - Superiority
  - Non-inferiority and Equivalence
  - Targeted Clinical Trials

References - Books :

Chow, S.C., and Liu, J.P. (2004) *Design and Analysis of Clinical Trials*, 2nd Ed., John Wiley and sons, New York, New York.

Pocock, S.J. (1996) *Clinical Trials: A Practical Approach*,2nd edition, John Wiley and sons, New York, new York.

Meinert, C.L. (1986) *Clinical Trials: Design, Conduct, and Analysis*, Oxford University Press, Oxford, UK.

Spilker, B. (1991) *Guide to Clinical Trials*, Raven, New York, New York.

Piantadosi, S. (1997) *Clinical Trials, A Methodology Perspectives*, John Wiley and sons, New York, New York.

References **-** *Guidelines or Guidances :*

The US FDA *Guideline for the Format and Content of the Clinical and Statistical Sections of New Drug Applications* (1988)

International Conference on Harmonisation: Guideline E8 *General Considerations for Clinical Trials* (1997).

International Conference on Harmonisation: Guideline E6 *Good Clinical Practice: Consolidated Guideline* (1996).

International Conference on Harmonisation: Guideline E3 *Structure and Content of Clinical Study Reports* (1996).

International Conference on Harmonisation: Guideline E9 *Statistical Principles for Clinical Trials* (1998).