

# CT 41: 第二期臨床試驗

## Phase II Design

林建甫

台北大學 統計系

2019-2020

# Phase II (SE, SA)

- Safety and Efficacy Studies (SE)
- Safety-Activity (SA)
- Larger Size (20~100)
- Carry forward recommended dose from phase I
- Seeking to establish evidence of activity
- Not comparative (II, IIa) vs. comparative (IIb)
- Sometimes controlled
- Several months to 2 years
- Practice for phase 3

# 目錄

- 1 Phase II Overview
- 2 Two-stage Phase II Design
- 3 Simon's Two-Stage Design
- 4 Admissible Two-Stage Designs

# Phase II Design: Introduction

- Looking for evidence of activity,  
with respect to what is already available to subjects.
- “Phase II” is convenient word but loosely defined.
- Piantadosi uses term middle development.

# Phase II Design Overview

- feasibility of treatment
- side effects and toxicity
- logistics of administration and cost
- dose finding (lowest dose level with good efficacy)
- Major issue: Is there enough evidence of efficacy of the new drug to move to phase III?
- Surrogate markers are often used.
- Usually, one-arm (no comparison)

表1: When is phase 2 step most critical?

Factors in favor of Phase 2	Factors in favor of skipping Phase 2
Many competing experimental therapies in pipeline	Lack of available therapies for disease (either experimental or approved)
Pessimistic about likelihood of success	Optimistic about likelihood of success
Availability of short-term, easily measured surrogate efficacy outcomes	Efficacy usefully measured by hard clinical outcomes
Highly prevalent disease	Rare disease
Opportunity cost for failed phase 3 trial would be high	Opportunity cost for time required to do phase 2 is high

## 測驗 1

A disease has a population prevalence of 2%. You have a diagnostic test for the disease that has sensitivity and specificity equal to 0.9. What is the post-test probability that an individual who tests positive has the disease?

- 1 0.155
- 2 0.431
- 3 0.783
- 4 0.900

## 測驗 2

Repeat the previous question with a true disease prevalence of 20%.

- 1 0.420
- 2 0.531
- 3 0.692
- 4 0.990



# Drug Development Pipeline I

- Let  $W$  denote drugs that are worthwhile
- Let  $W^C$  denote drugs that are not worthwhile
- Illustration of phase 2 goal:

# Drug Development Pipeline II

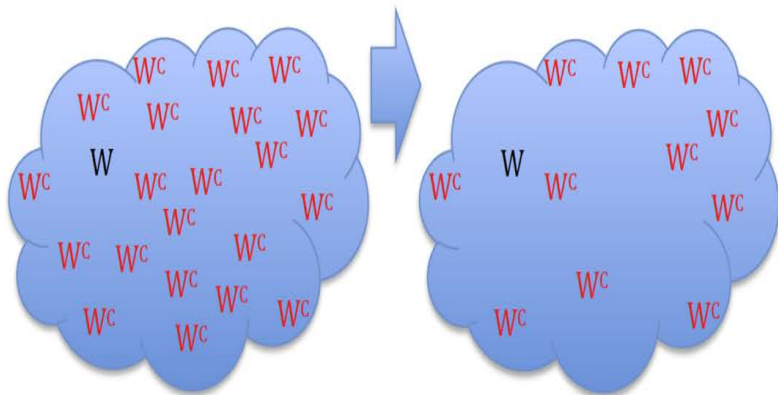


圖 1: 第二期臨床試驗目標

# Drug Development Pipeline III

- $P_{\text{ph2}}(W)$ : % of truly “worthwhile” drug entering phase 2 stage
- $\alpha, \beta$ : type I and type II error rates in phase 2
- $S_2$  indicate that a drug is selected in phase 2 pipeline
- From Bayes Rule, true positive finding from phase 2 is:

$$P(W | S_2) = \frac{P(S_2 | W)P_{\text{ph2}}(W)}{P(S_2 | W)P_{\text{ph2}} + P(S_2 | W^c)P_{\text{ph2}}(W^c)} \quad (1.1)$$

$$= \frac{(1 - \beta)P_{\text{ph2}}(W)}{(1 - \beta)P_{\text{ph2}}(W) + \alpha(1 - P_{\text{ph2}}(W))}. \quad (1.2)$$

# Drug Development Pipeline IV

- If 2% of drugs entering phase 2 study are truly worthwhile.
- $\alpha = 0.05$ , and  $\beta = 0.20$ , then

$$\frac{(0.80 \times 0.02)}{(0.80 \times 0.02 + 0.05 \times 0.95)} \approx 25\%. \quad (1.3)$$

- Approximate 25% of drugs leaving phase 2 study and entering phase 3 study, are truly worthwhile.

# Drug Development Pipeline V

- Phase 2 goal is to enrich phase 3 population with worthwhile drugs  
so that positive findings in phase 3 are highly likely to be true positives:
- If 25% of drugs entering phase 3 study are truly worthwhile.
- $\alpha = 0.05$ , and  $\beta = 0.20$ , then

$$\frac{(0.80 \times 0.25)}{(0.80 \times 0.25 + 0.05 \times 0.75)} \approx 84\%. \quad (1.4)$$

- 84% of drugs leaving phase 3 study,  
i.e. submitted for regulatory approval, are truly worthwhile.

# General Dichotomy of Phase II Designs

- Phase II or Phase IIA: any evidence of activity?  
Lower threshold, single arm, fewer subjects
- Phase IIB: evidence of greater efficacy?  
Higher threshold, randomized (multiple arms), more subjects.
- In reality, blurry distinction between these two.

# Statistical Setup for Single Arm Phase II I

- All subjects enrolled to new therapy.
- Followed for outcome (“response”)
- Response often (forced to be) binary,  $Y \in \{0, 1\}$ .
- May take time to occur, e.g.
  - tumor shrinkage of X% by 3 months.
  - clinical improvement of symptoms by 6 months.
- Should not take too long.

# Statistical Setup for Single Arm Phase II II

- Two motivating questions:
  - What is current response rate to best available therapy?  
( $\pi_0$ ; historical)
  - What response rate would suggest activity? ( $\pi_1$ )



# Phase II with Single Arm: Hypothesis Testing I

- $\pi = P(Y = 1) = P(\text{response})$

$$H_0 : \pi = \pi_0 \quad \text{versus} \quad H_1 : \pi = \pi_1. \quad (1.5)$$

- Reject  $H_0$  = conclude further study warranted
- Do not reject  $H_0$  = conclude no further study warranted

# Phase II with Single Arm: Hypothesis Testing II

表2: 假設檢定: 型一誤差與型二誤差

真實的狀況	可能的決策	
	接受 $H_0$	拒絕 $H_0$
$H_0$ 為真	(4) 正確決策 特異性 (Specificity)	(1) 型一錯誤 (FP) $\alpha$ 誤差
$H_A$ 為真	(2) 型二錯誤 (FN) $\beta$ 誤差	(3) 正確決策 敏感度 (Sensitivity)

# Phase II with Single Arm: Hypothesis Testing III

- One-sided hypothesis test.
- $R$ : total # of responses
- Estimate  $\pi$  with  $\hat{\pi} = R/n$ .
- Reject  $H_0$  if  $R/n > r/n$  for some constant  $r$
- What sample size  $n$  is required?

# Phase II with Single Arm:

## Normal-based Sample Size

- $H_0 : \pi = \pi_0$  versus  $H_A : \pi = \pi_1$ .
- Effect size:  $\delta = \pi_1 - \pi_0$ .
- $n$ : sample size.

$$n = \frac{\left( Z_{1-\alpha} \sqrt{\pi_0(1-\pi_0)} + Z_{1-\beta} \sqrt{\pi_A(1-\pi_1)} \right)^2}{(\pi_1 - \pi_0)^2}. \quad (1.6)$$

# Comments on Formula

- Large-sample approximation based on normality assumption.
- Actually easy to calculate “exact” sample size directly using binomial distribution (hopefully get similar results).
- Phase II trials have limited budgets–negotiations are inevitable.
- One of your jobs is to communicate ramifications of different choices.
- Various formulas exist in softwares.

# Determining $r$

- Reject  $H_0$  if  $R/n > r/n$  at the end.
- If sufficiently many subjects respond.
- $\pi_0 = 0.20$  and  $\pi_1 = 0.35$
- $\alpha = 0.05$  and  $\beta = 0.20$
- one-sided test
- $n = 50$  and  $r = 15$  (reject  $H_0$ )

$$Pr(R > r \mid \pi = \pi_0 = 0.20) = 0.031 \quad (\text{reject } H_0) \quad (1.7)$$

$$Pr(R > r \mid \pi = \pi_1 = 0.35) = 0.720. \quad (1.8)$$

# Determining $r$ II

- So if 15 or more responses (out of  $n = 50$ ), move to phase 3
- Reality check: 95% score-based CI for  $\pi$  is (0.19, 0.44).
- CI lower limit =  $0.19 < \pi_0 = 0.20$
- Evidence not overwhelming.

# 目錄

- 1 Phase II Overview
- 2 Two-stage Phase II Design**
- 3 Simon's Two-Stage Design
- 4 Admissible Two-Stage Designs



# Stopping for Futility I

- In one-stage design, all subjects enrolled before decision
- What if no responses after 5, 10, 15 subjects?
- Worthwhile, ethical to continue?
- Motivation for interim futility analyses
- futile = no use in trying
- lack of evidence for activity

# Stopping for Futility II

- Simple futility analysis
- First stage:  
assesses likelihood of no responses under  $H_1$ .

表3: Stopping for Futility When  
 $R = 0$

subjects enrolled	$P(R = 0 \mid \pi = 0.35)$
1	0.650
2	0.423
3	0.275
4	0.179
5	0.116
6	0.075
7	0.049
8	0.032

# Two-stage design (Gehan, 1961)

- Stage One:
  - Enroll initial cohort of  $n_1$  subjects.
  - Stop for futility if no responders
- Stage Two:
  - Otherwise, enroll remaining  $n - n_1$  subjects.
  - Conduct standard hypothesis test at end.
- Gehan originally proposed  $n_1 = 14$  based upon  $H_1 : \pi = 20\%$ .
- Circa 1960 (early chemotherapeutic era),  
20% response rate was impressive.
- We can use  $n_1 = 7$  based upon larger 35% response rate.

### 測驗 3

---

Relative to a one-stage design, adding a futility analysis will generally

- ① increase the total type I error rate
- ② not change the total type I error rate
- ③ decrease the total type I error rate

## 測驗 4

How many unique values of  $n$ , i.e. the actual sample size, can a two-stage design actually have, assuming all rules are followed precisely?

- 1 the actual sample size is pre-specified in advance
- 2 there are two possible sample size
- 3  $n$ : enrollment could stop at any point in the trial

# Modified Gehan's Designs

- Modified Gehan's Designs
- Continue if  $R_1 > 0$  (first stage),  $R > r$  (second stage)

$$\text{Type I Error} \quad (2.1)$$

$$= P(R > r, R_1 > 0 \mid \pi = \pi_0) \quad (2.2)$$

$$= \sum_{x=1}^{n_1} P(R > r, R_1 = x \mid \pi = \pi_0) \quad (2.3)$$

$$= \sum_{x=1}^{n_1} P(R - R_1 > r - x, R_1 = x \mid \pi = \pi_0) \quad (2.4)$$

$$= \sum_{x=1}^{n_1} P(R - R_1 > r - x \mid \pi = \pi_0) \times P(R_1 = x \mid \pi = \pi_0). \quad (2.5)$$

表4: 比較第二期試驗設計: One Stage 與 Modified Gehan

	Type I Error	Power	P (Early Termination)	$\mathcal{E}$ (Enrollment)
One Stage	0.061	0.812	0.00	50
Modified Gehan	0.057	0.785	0.21	41

# Comments

- Tradeoff: enroll 9 fewer subjects (on average) for slight loss of power.
- However, futility analysis is conservative:  
 $P(R_1 = 0 \mid \pi = \pi_0) = 0.210$ ,  
i.e. almost 80% chance of enrolling max possible subjects under  $H_0$
- Possible refinement: new stage 1 stopping rule:  $R_1 > r_1$ .  
Tune  $\{r_1, n_1\}$  to stop when  $H_0 : \pi = \pi_0$  appears likely



# How to Determine $n_1$ and $n$ ? I

- Determine  $n_1$ : denote  $X = \#$  responses out of  $n_1$  subjects.

$$P(X = 0) = (1 - \pi)^{n_1}. \quad (2.6)$$

- $\pi_0 =$  minimal efficacy
- If  $\pi > \pi_0 \Rightarrow$  investigate the new drug in phase III.
- Control the probability of discarding the new drug early if in fact it is promising.
- $n_1$  has to satisfy:

$$P(X = 0) = (1 - \pi)^{n_1} \leq \alpha_0, \quad \forall \pi \geq \pi_0. \quad (2.7)$$

- where  $\alpha_0$  is our tolerance.

# How to Determine $n_1$ and $n$ ? II

- $P(X = 0) = (1 - \pi)^{n_1}$  is a decreasing function of  $\pi$ ,
- Only need  $n_1$  for

$$(1 - \pi_0)^{n_1} \leq \alpha_0. \quad (2.8)$$

$$\Rightarrow n_1 \log(1 - \pi_0) \leq \log(\alpha_0) \quad (2.9)$$

$$\Rightarrow \text{need } n_1 \geq \frac{\log(\alpha_0)}{\log(1 - \pi_0)} \quad (2.10)$$

- $\pi_0 = 0.20, \alpha_0 = 0.05$

$$n_1 \geq \frac{\log(0.05)}{\log(1 - 0.20)} \approx 14. \quad (\text{round up}) \quad (2.11)$$

# How to Determine $n_1$ and $n$ ? III

- Interpretation:
- If see 0/14 responses,  
upper limit of 95% CI on  $|pi$  is less than 0.2
- 95% sure response rate with new therapy is less than  $\pi_0$ .

# How to Determine $n_1$ and $n$ ? IV

- $n_2 = n - n_1$  chosen to give sufficient precision for estimating  $|p_i$  after all patients observed
- Gehan chooses  $n_2$  after seeing  $X_1$ , but most applications fix  $n_1$  and  $n_2$  (or  $n$ ) in advance.

# How to Determine $n_1$ and $n$ ? V

- Gehan determine  $n$ : based on precision of 95% CI.
- Want to be 95% sure that the estimate is within  $\pm 15\%$  of the minimum  $\pi_0 = 0.2$ :

$$1.96 \times \left( \frac{0.2 \times 0.8}{2} \right)^{\frac{1}{2}} = 0.15. \quad (2.12)$$

- Solve  $n \approx 28$ .

# How to Determine $n_1$ and $n$ ? VI

- Expected sample size under Gehan design

$$n = \begin{cases} n_1, & \text{with probability } (1 - \pi)^{n_1}; \\ n, & \text{with probability } 1 - (1 - \pi)^{n_1}. \end{cases} \quad (2.13)$$

$$\text{mean}(n) = n_1(1 - \pi)^{n_1} + (n_1 + n_2) \left( 1 - (1 - \pi)^{n_1} \right) \quad (2.14)$$

$$= n_1 + n_2 \left( 1 - (1 - \pi)^{n_1} \right). \quad (2.15)$$

# 目錄

- 1 Phase II Overview
- 2 Two-stage Phase II Design
- 3 Simon's Two-Stage Design**
- 4 Admissible Two-Stage Designs

# Simon's Two-Stage Design (Simon, 1989)

- Goal: identify design set  $\{r_1, n_1, r, n\}$  that satisfies type I error  $\alpha$  and power  $1 - \beta$  constraints.
- Lots of such sets exist
- Simon proposed two designs that are best by some definition (Simon, Control Clinical Trials, 1989)
- More than 2800 citations currently



# Simon's Two-Stage Design: Statistical Considerations I

- Suppose two values  $\pi_0 < \pi_1$  are pre-specified
- Divide  $\pi \in (0, 1)$  into 3 Regions: Drug is ineffective | Indifference region | effective
- If  $\pi \leq \pi_0$  then declare the drug ineffective with high probability,  $1 - \alpha$ , where  $\alpha$  is taken to be small.
- If  $\pi \geq \pi_1$ , then consider the drug for further investigation with high probability,  $1 - \beta$ , where  $\beta$  is taken to be small.
- $\alpha$  and  $\beta$  are generally 0.05 and 0.20.

# Simon's Two-Stage Design: Statistical Considerations II

- Find integers  $n_1, n, r_1, r$ , with  $n_1 < n, r_1 < n_1$ , and  $r < n$
- Enroll  $n_1$  subjects in the first stage.
  - If  $r_1$  subjects or less respond,  
then declare the treatment a failure and stop.
  - If more than  $r_1$  respond,  
then add  $(n - n_1)$  subjects for a total of  $n$  subjects.

# Simon's Two-Stage Design: Statistical Considerations III

- At the second stage:
  - if the total # that respond among all  $n$  subjects is greater than  $r$ , then declare the treatment a success.
  - otherwise, declare it a failure.
- stop the trial at stage 1
  - if the # of responses among  $n_1$  is greater than  $r$ , then declare the treatment a success.

# Simon's Two-Stage Design: Statistical Considerations IV

- $X_1$  = the # of responses in stage 1 (out of  $n_1$  subjects)
- $X_2$  = the # of responses in stage 2 (out of  $n_2 = n - n_1$  subjects)

$$X_1 \sim \text{Bin}(n_1, \pi), X_2 \sim \text{Bin}(n_2, \pi) \quad (X_1 \text{ and } X_2 \text{ are ind}). \quad (3.1)$$

# Simon's Two-Stage Design: Statistical Considerations V

- Declare the new drug a failure if

$$(X_1 \leq r_1) \text{ OR } \{(X_1 > r_1) \text{ AND } (X_1 + X_2 \leq r)\}. \quad (3.2)$$

- The new drug is declared a success if

$$(X_1 > r_1) \text{ AND } (X_1 + X_2 > r). \quad (3.3)$$

# Simon's Two-Stage Design: Statistical Considerations VI

- Design constraints

$$P\left((X_1 > r_1) \text{ AND } (X_1 + X_2 > r)\right) \leq \alpha, \forall \pi \leq \pi_0; \quad (3.4)$$

$$P\left((X_1 > r_1) \text{ AND } (X_1 + X_2 > r)\right) \geq \beta, \forall \pi \geq \pi_1. \quad (3.5)$$

# Simon's Two-Stage Design: Statistical Considerations VII

- Denote the power function by

$$\Psi(\pi) = P\left((X_1 > r_1) \text{ AND } (X_1 + X_2 > r) \mid \pi\right). \quad (3.6)$$

- $\Psi(\pi)$  is an increasing function of  $\pi$  for any  $n_1, r_1, n, r$ .

# Simon's Two-Stage Design: Statistical Considerations VIII

- Therefore, criteria (3.4) and (3.4) are equivalent to

$$P\left((X_1 > r_1) \text{ AND } (X_1 + X_2 > r) \mid \pi = \pi_0\right) = \alpha, \quad (3.7)$$

$$P\left((X_1 > r_1) \text{ AND } (X_1 + X_2 > r) \mid \pi = \pi_1\right) = \beta. \quad (3.8)$$



# Simon's Two-Stage Design:

## Statistical Considerations IX

- How to calculate  $P\left((X_1 > r_1) \text{ AND } (X_1 + X_2 > r) \mid \pi\right)$  for each  $\pi$ ?
- By independence of  $X_1$  and  $X_2$ :

$$P\left((X_1 > r_1) \text{ AND } (X_1 + X_2 > r) \mid \pi\right) \quad (3.9)$$

$$= \sum_{m_1 > r_1, m_1 + m_2 > r} P(X_1 = m_1, X_2 = m_2) \quad (3.10)$$

$$= \sum_{m_1 > r_1, m_1 + m_2 > r} P(X_1 = m_1) P(X_2 = m_2). \quad (3.11)$$

# Simon's Two-Stage Optimal Design I

- Many combinations of  $n_1, n, r_1, r$  satisfy (3.4) and (3.4).
- Optimal Design:  
has smallest expected sample size when  $\pi = \pi_0$   
(when the new drug is ineffective)

# Simon's Two-Stage Optimal Design II

- The expected sample size when  $\pi = \pi_o$

$$\mathcal{E}(n \mid \pi = \pi_o) \quad (3.12)$$

$$= n_1 P(\text{stop at stage 1}) + n P(\text{did not stop at stage 1}) \quad (3.13)$$

$$= n_1 \left( P(X_1 \leq r_1 \mid \pi = \pi_o) + P(X_1 > r \mid \pi = \pi_o) \right) \\ + n P(r_1 + 1 \leq X_1 < r \mid \pi = \pi_o). \quad (3.14)$$

# Simon's Two-Stage Optimal Design III

- Through computer search, the optimal design can be identified.
- Among all sets of  $n_1, n, r_1, r$  satisfy (3.4) and (3.4) constraints, identify set that minimizes  $\mathcal{E}(n \mid \pi = \pi_0)$
- Minimize number of subjects enrolled when trial should not have been run.
- Simon (1989), Optimal two-stage designs for Phase II clinical trials. Controlled Clinical Trials. 10: 1-10.

# Simon's Two-Stage Minimax Design

- Through computer search, the optimal design can be identified.
- Among all sets of  $n_1, n, r_1, r$  satisfy (3.4) and (3.4) constraints, identify set that minimizes  $n$ .
- Aimed at identifying effective agents using as few patients as possible.

## 測驗 5

---

For studying a rare disease in which patient accrual is slow, which of the Simon's design would likely make more sense

- 1 optimal
- 2 minimax

# Simon's two-Stage Design

- Implemented in `clinfun` package (Seshan, 2015)

# 目錄

- 1 Phase II Overview
- 2 Two-stage Phase II Design
- 3 Simon's Two-Stage Design
- 4 Admissible Two-Stage Designs



# Admissible Two-Stage Designs

- Consider expected cost function for any possible design set satisfying  $\alpha$  and  $\beta$  constraints given by

$$C(q, \{r_1, n_1, r, n\}) = qn + (1 - q)\mathcal{E}(n \mid \pi = \pi_0) \quad (4.1)$$

- It is weighted average of maximum sample size and expected sample size under  $H_0$
- Design set  $n_1, n, r_1, r$  is called admissible if it achieves smallest possible cost across all possible design sets for at least one  $q \in [0, 1]$ .

# Simon's Design and Admissible Design

- Simon's optimal design is admissible because it minimizes risk when  $q = (?)$
- Simon's minimax design is admissible because it minimizes risk when  $q = (?)$
- There are usually other admissible designs that minimize risk for  $q \in (0, 1)$ .
- Jung et al., 2004

# Discussion I

- Phase 2 trials are most often futility trials.
- Goal is to prune inefficacious treatments
- Could trial stop for efficacy at stage 1? Should it?
- Different implications from stopping early because  $R_1 \leq r_1$  and failing to reject  $H_0$  because  $R \leq r$
- Nothing special about two stages.
- More stages possible.

# Discussion II

- Strict sample sizes and adherence to interim analyses may be difficult to adhere to.
- Causes trial conduct to deviate from trial design and potentially invalidates type I error and power properties.
- Inference is conditional on following design as it is laid out

# Inference depends on design: Ex 1 I

- Suppose that we conduct minimax design with  $\pi_0 = 0.20$ ,  $\pi_1 = 0.35$ ,  $\alpha = 0.05$ , and  $1 - \beta = 0.80$ :  
 $\{r_1, n_1, r, n\} = \{6, 31, 15, 53\}$
- We successfully conclude the trial with  $R = 16 > r = 15$  responders.

# Inference depends on design: Ex 1 II

- Could report p-value:

$$p = P(R > 15, R_1 > 6 \mid \pi = \pi_0) \quad (4.2)$$

$$= \sum_{x=7}^{31} P(R > 15, R_1 = x \mid \pi = \pi_0) \quad (4.3)$$

$$= \sum_{x=7}^{31} P(R - R_1 > 15 - x, R_1 = x \mid \pi = \pi_0) \quad (4.4)$$

$$= \sum_{x=7}^{31} P(R - R_1 > 15 - x \mid \pi = \pi_0) \\ \times P(R_1 = x \mid \pi = \pi_0) \quad (4.5)$$

$$= 0.0498. \quad (4.6)$$

# Inference depends on design: Ex 1 III

```
1 # Inference depends on design: Ex 1
2 # Design 1
3 r1 = 6; n1 = 31; R = 15; n = 53;
4 p0 = 0.20
5 sum(pbinom(R - ((r1+1):n1), n - n1, p0, lower.tail = F) *
6     dbinom((r1+1):n1, n1, p0));
7 [1] 0.04979161
```

# Inference depends on design: Ex 2 I

- Adapted from (Lindley and Phillips, 1976).
- Two designs for testing  $H_0 : \pi = 0.2$ .



# Inference depends on design: Ex 2 II

- Design 1:
- Enroll  $n = 25$  patients.
- $R$  is binomial.
- Observe  $R = 8$  responses.
- p-value is

$$P(R \geq 8 \mid \pi = 0.2) = \sum_{x=8}^{25} \binom{25}{8} (0.2)^x (0.8)^{25-x} \quad (4.7)$$

$$= 0.109. \quad (4.8)$$

```
1 # Inference depends on design: Ex 2
2 pbinom(7 , 25, 0.2, lower.tail = F) # 1-binomial CDF
3 [1] 0.1091228
```

# Inference depends on design: Ex 2 III

- Design 2:
- Enroll patients until  $f = 17$  non-responders.
- $R$  is negative binomial. Observe  $R = 8$  responses.
- p-value is (negative-binomial)

$$P(R \geq 8 \mid \pi = 0.2) = \sum_{x=8}^{\infty} \binom{x+16}{x} (0.2)^x (0.8)^{17} \quad (4.9)$$

$$= 0.0892. \quad (4.10)$$

```
1 # Inference depends on design: Ex 2
2 # Design 2
3 pnbinom(7, 17, 1-0.2, lower.tail = F) # 1 - neg-binom CDF
4 [1] 0.08917126
```

# Inference depends on design: Ex 2 IV

- Design 1:  $p = 0.109$ .
- Design 2:  $p = 0.089$ .
- If  $\alpha = 0.10$  were significance threshold, then same data (but different designs) yield different conclusions.

# Likelihood Principle

- In both examples,  
likelihood of final data was equal for both designs.
- Frequentist inference based on what could have occurred  
and not just what did occur.
- Violates likelihood principle:  
inference should be based only upon likelihood function  
and not design.

# K-Stage Design (Fleming, 1982)

- Gehan design allows for early termination only if new therapy seems ineffective
- Fleming (1982, Biometrics)
- K-stage designs allow for early stopping if
  - treatment appears ineffective
  - or treatment appears overwhelmingly effective
- preserve (approximately) size and power of a single stage design

# Practical Uses of Fleming's Designs

- Rarely happens that want more than 2 stages
- Rarely terminate phase II trial early because treatment is effective.
- Phase III takes time to develop.
- Phase III is not often ready in time to take advantage of early stopping
- Would like more information before starting phase III.

# More Phase II Designs

- Bayesian Single Stage Design  
Sylvester, Biometrics, 1988.
- Bayesian Two-Stage Design  
Herson, Biometrics, 1979.
- Bayesian Two-Stage Design  
Simon and Thall, Controlled Clinical Trials, 1989.  
Simon and Thall, Biometrics, 1994.
- Bivariate Designs (Early Stopping for Response and Toxicity)  
Bryant and Day Biometrics, 1995. Conaway and Petroni,  
Biometrics, 1995.

# R Packages: Phase II Design

- `clinfun`
- `ph2simon`
- `ph2mult`
- `DoseFinding`
- `MCPMod`
- `OneArmPhaseTwoStudy`
- `ph2mult`



# 目錄

- 1 Phase II Overview
- 2 Two-stage Phase II Design
- 3 Simon's Two-Stage Design
- 4 Admissible Two-Stage Designs

# Randomized Controlled Phase II Designs

- Two Arms Design and Multiple Arms Design
- Single Stage and Multiple Stages
- Non-adaptive Design Adaptive Design

Thanks!