

Phase II Clinical Trial

Phase I では安全性を評価したのに対して、phase II では薬効を評価します。

薬剤に関する phase II trials

- (1) 初期用量は phase I study(ies)の結果より設定します。
- (2) 用量は通常増量しません (デザインによる); むしろ 1 人でも患者さんに毒性を認めたら誰も毒性を示さないところまで減量します。
- (3) まず他に良い治療の存在しない病気のステージの患者さんの場合 (不治の病) まず 14-20 人を対象に薬を投与します。例: 14 人に投与し全く効果を認めなければその時点で中止します。なぜなら真の反応率が 20% だとした場合、14 人中誰も反応しない確率は 0.05 以下であり統計学的に起こりにくいことです。もしも、少なくとも 1 人が薬剤に反応したら、新たに 10-25 人の患者さんに薬剤を投与し、反応率と信頼区間を算出します。よって 20-80 人位の患者さんを対象とします。通常は 100-200 人を超えることはありません。
- (4) 他にも治療薬が存在する場合、短期間に phase II を終えます。
- (5) 薬剤に対する反応性、その他の効果が phase II study における主な目的ですが、毒性や副作用についての情報収集も重要です。
- (6) phase I とはすこし異なった患者層を対象とします。健丈人を対象とせず、薬剤の対象疾患を持つ患者さんを対象とします。しかしながら希な病気に対しては健丈人を対象に行なわれることもあります。薬剤の反応性を測定可能な病気を対象とします。患者さんの余命が薬剤の反応をみるのに十分な時間あることを必要とします。
- (7) 通常はその薬剤試験の対象疾患は 1 つに絞られます。
- (8) 経費と時間が限られます。しかも他社との競合等も存在し、複雑な世界です。また phase III trial は高額のコストがかかるため、phase II で十分な自信を得て次のステップにすすむ必要があります。しかし疾患および薬剤の性格によっては、観察期間が短かったり、用量設定がうまくいかなかったり、健康すぎる、あるいは病気が重過ぎて本当の効果をうまく測定できないことも十分ありえます。

(Design 1) Rising Single dose studies

個々の患者さんに対して、頭痛薬の phase II trial を行なっています。最低の用量から投薬を始め、大きな問題がなければ徐々に用量を増やしていきます。通常は placebo をランダムに挿入します。いくつか例をあげましたが、薬物動態によってデザインは当然異なってきます。投薬後すぐに効果がみられるものもあれば 1 週間かかるものもあるでしょうし、washout に 1 週間かかるものもあれば数時間でよいものもあるでしょう。また抗生剤や解熱剤のように 1 度で効いてしまうこともあるかもしれません。

対象	100 mg	200 mg	300 mg	400 mg
1	A	A	P	A
2	A	A	P	A
3	A	P	A	A
4	P	A	A	A
5	A	A	A	P
6	P	A	A	A
7	A	P	A	A
8	A	A	A	P

A: active drug administration, P; placebo

(Design 2) Alternating-panel rising single-dose study

dose mg	Subject numbers															
	Panel I								Panel II							
100	14	2	7	10	16	3	6	12	1	5	15	13	4	11	9	8
200	a	a	a	a	a	p	p	p	a	a	a	a	a	p	p	p
300	p	p	p	a	a	a	a	a	p	p	p	a	a	a	a	a
400	a	p	a	p	a	p	a	a	a	p	a	p	a	p	a	a
500	p	a	p	a	p	a	a	a	p	a	p	a	p	a	a	a
600																
700																
800																

A: active drug, p; placebo, panel I は Panel II を開始する少なくとも 24 時間前に修了する。各個人が薬の量を増やすときは少なくとも 48 時間あける。

(Design 3) Alternating-panel rising single-dose crossover study

#	dose day	10mg 1	10mg 2	20mg 1	20mg 2	50mg 1	50mg 2	100mg 1	100mg 2	250mg 1	250mg 2	500mg 1	500mg 2
1	1	a	p			p	a			p	a		
2	1	p	a			p	a			a	p		
3	1	p	a			a	p			p	a		
4	1	a	p			a	p			a	p		
5	2			a	p			p	a			a	p
6	2			a	p			a	p			p	a
7	2			p	a			p	a			a	p
8	2			p	a			a	p			p	a

A: active, p; placebo

(Design 4) Rising single-dose crossover study

	Period					
subject	1	2	3	4	5	6
1	2.5	5	10	20	p	40
2	2.5	5	p	10	20	40
3	2.5	5	10	20	40	p
4	p	2.5	5	10	20	40
5	2.5	5	10	p	20	40
6	2.5	p	5	10	20	40
7	2.5	5	10	20	p	40
8	2.5	5	10	p	20	40
9	2.5	p	5	10	20	40
10	2.5	5	p	10	20	40
11	2.5	5	10	20	40	p
12	p	2.5	5	10	20	40

(Design 5) Parallel panel rising single-dose study

	Dose level (mg)				
	subject	10	20	40	80
I	1	a	a	a	a
	2	a	a	a	a
	4	a	a	a	a
	6	a	a	a	a
	7	a	a	a	a
	9	a	a	a	a
	10	a	a	a	a
II	3	p	p	p	p
	5	p	p	p	p
	8	p	p	p	p

(Design 6) Parallel-panel multiple –dose time-lagged study

		Dose level (mg)		
		250	500	1000
		Period 1	Period 2	Period 3
I	1	p		
I	2	a		
I	3	a		
I	4	a		
II	5		p	
II	6		a	
II	7		a	
II	8		a	
III	9			a
III	10			a
III	11			a
III	12			p

(Design 7) Parallel-panel rising multiple-dose design

	Dose (mg)	Duration (days)	Study days
Group A	100	7	1 - 7
	250	7	8 - 14
	500	7	15 - 21
	placebo	7	22 - 28
Group B	placebo	28	1 - 28

(Design 8) Altering-panel multiple-dose rinsing dosage study

		Study day						
panel	#	1	2	3	4	5	6	7
I	1	A	A	A	C	C	C	
	2	A	A	A	C	C	C	
	3	A	A	A	C	C	C	
	4	P	P	P	P	P	P	
II	5		B	B	B	D	D	D
	6		P	P	P	P	P	P
	7		B	B	B	D	D	D
	8		B	B	B	D	D	D

P: placebo, A: 100mg, B: 250mg, C: 500mg, D: 1000mg,
 用量を上げる前に washout period 挿入する。

(Design 9) Parallel multiple-dose design

	Treatment		
	250 mg b.i.d	500 mg b.i.d	Placebo b.i.d.
subject	1, 5, 8, 9	3, 6, 7, 12	2, 4, 10, 11

治療は1週間ずつ

用量依存性パラレルテスト

新しい抗うつ薬を検討しようと思います。既に販売されている薬剤 A のスコアは用量を増すと改善（スコアの減少）をみえています。新薬 B も似た傾向を示しています。はたしてどちらの方が効果があるといえるでしょうか？ A に対する B の potency という形で表し、95% CI も計算してください。

スコアでみた薬効

	標準的治療薬 A の用量(mg/kg)				新薬 B の用量 (mg/kg)				
	10	30	100	300	1	3	10	30	100
	48	50	26	20	44	35	33	10	6
	49	37	20	14	48	39	32	19	5
	52	36	25	12	48	42	33	19	20
	53	39	26	16	56	52	48	27	17
	34	34	27	15	47	41	33	21	15
	50	36	24	11	56	44	28	16	9
	58	41	28	18					
	48	40	25	16					
	46	30	22	14					
	56	40	23	13					
合計	494	383	246	149	299	253	197	112	72
平均	49.4	38.3	24.6	14.9	49.8	42.2	32.8	18.7	12.0

Total score in A = 1272, total score in B = 933, total score in A and B = 2205

Total subjects in A = 40, total subjects in B = 30, total subjects in A and B = 70

C = correction term = $(2205)^2/70 = 69457.5$

Total SS = $(48)^2 + (49)^2 + \dots + (9)^2 - C =$

Doses SS = $=(494)^2/10 + (383)^2/10 + \dots + (72)^2/6 - C =$

Preparation SS = $(1272)^2/40 + (933)^2/30 - C =$

SS: sum of squares

Drug A

$$X = 10 \times \log 10 + 10 \times \log 30 + 10 \times \log 100 + 10 \times \log 300 = 69.542$$

$$X_{\text{mean-A}} = 69.542/40 = 1.7386$$

$$[X^2]_A = [10 \times (\log 10) + 10 \times (\log 30) + 246 \times (\log 100) + 149 \times (\log 300)] - (69.5420)^2/40 = 12.2762$$

$$[XY]_A = 494 \times \log 10 + 383 \times \log 30 + 10 \times \log 100 + 10 \times \log 300 - (69.5420)(1272)/40 = -290.6184$$

Drug B

$$X = 6 \times \log 1 + 6 \times \log 3 + 6 \times \log 10 + 6 \times \log 30 + 6 \times \log 100 = 29.7252$$

$$X_{\text{mean}} = 29.7252/30 = 0.9908$$

$$[X^2]_B = [10 \times (\log 1)^2 + 10 \times (\log 3)^2 + 10 \times (\log 10)^2 + 10 \times (\log 30)^2 + 10 \times (\log 100)^2] - (29.7252)^2/30 = 15.0038$$

$$[XY]_B = 299 \times \log 1 + 253 \times \log 3 + 197 \times \log 10 + 112 \times \log 30 + 112 \times \log 100 - (15.0038)(933)/30 = -297.3122$$

$$[X^2] = [X^2]_A + [X^2]_B = 27.2800$$

$$[XY] = [XY]_A + [XY]_B = -587.9306$$

$$\text{Regression SS} = ([XY])^2 / [X^2] = (587.9306)^2 / 27.2800 = 12670.9087$$

$$\text{Parallelism SS} = [XY]_A^2 / [X^2]_A + [XY]_B^2 / [X^2]_B - ([XY])^2 / [X^2] = (-290.6184)^2 / 12.2762 + (297.3122)^2 / 15.0038 - (-587.9306)^2 / 27.2800 = 100.4705$$

$$\text{Curvature SS} = \text{Doses SS} - \text{PREP SS} - \text{Regression SS} - \text{Parallelism SS} = 98.0875$$

$$\text{Within doses SS} = \text{Total SS} - \text{Doses SS} = 1747.6333$$

$$M = \log_{10}(\text{relative potency}), R = 10^M, \quad b_c = \text{common slope とします。}$$

$$b_c = [XY] / [X^2] = (-587.9306) / 27.2800 = -21.5517$$

$$M = X_{A-\text{MEAN}} - X_{B-\text{MEAN}} - [(Y_{A-\text{MEAN}} - Y_{B-\text{MEAN}}) / b_c] = (1.7386 - 0.9908) - [(1272/40 - 933/30) / (-21.5517)] = 0.7803$$

$$R = 10^{0.7803} = 6.03$$

$$S^2 = (\text{within doses SS}) / (\text{within degree of freedom}) = 28.6497$$

$$t_{0.05, 61} = 1.9969$$

$$g = (t^2 s^2) / b_c^2 [X^2] = (1.9969)^2 (28.6497) / (-21.5517)^2 (27.2800) = 0.0090$$

95% CI

$$M_L, M_U = M + g / (1-g) (M - X_{A-\text{MEAN}} + X_{B-\text{MEAN}}) \pm [t_{0.05, 61} / b_c (1 - g)] \times \{S^2 [(1-g) / (1/N_A + 1/N_B)] + (M - X_{A-\text{MEAN}} + X_{B-\text{MEAN}}) / [X^2]\} = 0.7806 \pm (-0.1205) = 0.6601, 0.9011$$

$$R_L, R_U = 4.57, 7.96$$

解答

新薬 B の 1mg/kg は標準的治療薬 A の 6.03 mg/kg に相当します。その 95% CI は 4.57 - 7.96 の範囲です。また 95% CI は 1 を含まないことから、A と B の間には有意差が存在します。

新しい抗高血圧薬を検討しようと思います。既に販売されている抗高血圧薬 A は用量依存的に収縮期血圧を下降させます。新薬 B も似た傾向を示しています。はたしてどちらの方が高い薬効を持つでしょうか？ A に対する B の potency という形で表し、95% CI も計算してください。

	抗高血圧薬 A (mg/kg)				新薬 B (mg/kg)				
	15	30	60	120	25	50	100	200	
	4	10	14	20	2	4	5	8	
	6	11	15	21	3	3	6	9	
	6	9	13	20	4	5	7	7	
	4	10	12	22	2	2	4	10	
	5	9	16	19	1	6	6	6	
total	25	49	70	102	12	20	28	40	346
means	4.6	9.8	14	20.8	2.4	4	5.6	8	

Total score in A = 246, total score in B = 100, total score in A and B = 346

Total subjects in A = 20, total subjects in B = 20, total subjects in A and B = 40

C = correction term = $(346)^2/40 = 2992.9$

Total SS = $(4)^2 + (5)^2 + \dots + (6)^2 - C = 1311.1$

Doses SS = $= (25)^2/5 + (49)^2/5 + \dots + (40)^2/5 - C = 1258.9$

Preparation SS = $(246)^2/20 + (100)^2/20 - C = 532.9$

SS: sum of squares

Drug A

$$X = 5 \times \log 15 + 5 \times \log 30 + 5 \times \log 60 + 5 \times \log 120 = 32.5524$$

$$X_{\text{mean-A}} = 32.5524/20 = 1.62762$$

$$[X^2]_A = [5 \times (\log 15) + 5 \times (\log 30) + 5 \times (\log 60) + 5 \times (\log 120)] - (32.5524)^2/20 = 3.70925$$

$$[XY]_A = 25 \times \log 15 + 49 \times \log 30 + 70 \times \log 60 + 102 \times \log 120 - (32.5524)(246)/20 = 37.9329$$

Drug B

$$X = 5 \times \log 25 + 5 \times \log 50 + 5 \times \log 100 + 5 \times \log 200 = 36.9887$$

$$X_{\text{mean}} = 36.9887/20 = 1.849485$$

$$[X^2]_B = [5 \times (\log 25)^2 + 5 \times (\log 50)^2 + 5 \times (\log 100)^2 + 5 \times (\log 200)^2] - (36.9887)^2/20 = 2.2653$$

$$[XY]_B = 12 \times \log 25 + 20 \times \log 50 + 28 \times \log 100 + 40 \times \log 200 - (36.9887)(100)/20 = -13.8474$$

$$[X^2] = [X^2]_A + [X^2]_B = 5.97455$$

$$[XY] = [XY]_A + [XY]_B = 51.7803$$

$$\text{Regression SS} = \frac{(\sum XY)^2}{\sum X^2} = (51.7803)^2/5.97455 = 448.770$$

$$\text{Parallelism SS} = \frac{(\sum XY)_A^2}{\sum X^2_A} + \frac{(\sum XY)_B^2}{\sum X^2_B} - \frac{(\sum XY)^2}{\sum X^2} = (37.9329)^2/3.70925 + (13.8474)^2/2.2653 - (51.7803)^2/5.97455 = 23.8$$

$$\text{Curvature SS} = \text{Doses SS} - \text{PREP SS} - \text{Regression SS} - \text{Parallelism SS} = 253.43$$

$$\text{Within doses SS} = \text{Total SS} - \text{Doses SS} = 52.2$$

$M = \log_{10}(\text{relative potency}), R = 10^M, b_c = \text{common slope}$ とします。

$$b_c = \frac{\sum XY}{\sum X^2} = (51.7803)/5.97455 = 8.6668$$

$$M = X_{A-MEAN} - X_{B-MEAN} - \frac{(Y_{A-MEAN} - Y_{B-MEAN})}{b_c} = (1.62762 - 1.849485) - [12.3 - 5.0]/(8.6668) = -1.06416$$

$$R = 10^{-1.06416} = 0.086266$$

$$S^2 = (\text{within doses SS})/(\text{within degree of freedom}) = 52.2/32 = 1.63125$$

$$t_{0.05, 32} = 2.042$$

$$g = (t^2 s^2)/b_c^2 \sum X^2 = (2.042)^2(1.63125)/(8.6668)^2(5.97455) = 0.01515$$

95% CI

$$M_L, M_U = M + \frac{g}{(1-g)} (M - X_{A-MEAN} + X_{B-MEAN}) \pm [t_{0.05, 61} / b_c (1 - g)] \times \{S^2[(1-g)(1/N_A+1/N_B) + (M - X_{A-MEAN} + X_{B-MEAN})^2 / \sum X^2]\} = -1.182227, -0.972007$$

$$R_L, R_U = 0.06573, 0.16657893$$

解答

新薬 B の 1mg/kg は標準的治療薬 A の 0.09 mg/kg に相当します。その 95%CI は 0.07 - 0.17 の範囲です。また 95%CI は 1 を含まないことから、A と B の間には有意差が存在します。

Trend test

傾向を調べる統計学的手法です。特に用量依存性に反応が上がることを検証するとき便利です。

$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$

$H_A: \mu_1 \neq \mu_2 \neq \mu_3 \neq \dots \neq \mu_k$

どれか1つでも違えば H_0 を棄却できます。例えば新薬の用量を増やしていったとき、下のように反応率も増えていったとします。反応率の低い方から順番に番号をふり(score)、logistic regression を用いてtrend test を行ないます。

```
. list
```

	response	non	score	N
1.	0	25	1	25
2.	10	15	2	25
3.	14	11	3	25
4.	21	4	4	25

```
. blogit response N score
```

Logit estimates

```
Number of obs = 100  
LR chi2(1) = 41.14  
Prob > chi2 = 0.0000  
Pseudo R2 = 0.2989
```

Log likelihood = -48.246151

```
-----+-----  
_outcome |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]  
-----+-----  
score |  1.426121   .2753138     5.180  0.000   .8865161   1.965726  
_cons | -3.883305   .7721328    -5.029  0.000  -5.396658  -2.369953  
-----+-----
```

用量依存性が証明されました。

Cross-Over Design / Carry Over Effect

cross over design を用いてテニス選手の痛みが新薬 Z により軽減するかどうかを検討したいと思います。痛みの軽減程度は 1 悪化、2 無変化、3 軽度改善、4 中等度改善、5 ほとんど改善、6 完全に改善、で評価します。A グループの 42 人は最初に新薬 Z を 2 週間使用し washout period なく placebo を 2 週間使用します。逆に B グループは placebo を 2 週間使用し washout period なく新薬 Z を 2 週間使用します。結果は下表の通りでした。

	A				B			
	新薬 Z	placebo	差	平均	新薬 Z	placebo	差	平均
mean	3.833	3.762	0.071	3.798	4.214	2.857	1.357	3.536
SD	1.188	1.574	1.813	1.060	1.353	1.160	1.376	1.056
n	42	42	42	42	42	42	42	42

新薬 Z は有効であると結論できますか？

$$D_{mean} = (0.071 + 1.357)/2 = 0.714$$

$$SD^2 = [(n_A - 1)SD_{DA}^2 + (n_B - 1)SD_{DB}^2]/(n_A + n_B - 2) = [(42 - 1)1.813^2 + (42 - 1)1.376^2]/(42 + 42 - 2) = 2.590$$

$$\text{Standard error of } d_{mean} = se(d_{mean}) = [SD^2/4(1/n_A + 1/n_B)] = [2.590^2/4(1/42 + 1/42)] = 0.176$$

$$t = d_{mean}/se(d_{mean}) = 0.714/0.176 = 4.07, \quad p < 0.001$$

$$95\% \text{ CI} : d_{mean} \pm t_{(n_A + n_B - 2)} \times se(d_{mean}) = 0.714 \pm 1.989 \times 0.176 = (0.365, 1.063)$$

よって新薬 Z は有効であると結論できます。しかし、もう一度表をよく見てください。A ではあまり差がないにもかかわらず、B においては大きな差があります。A においては最初に使用した新薬 Z が placebo 期間にも影響して、差を認めなかった可能性が高いと思われます。このような現象を、carry-over effects と呼びます。十分な washout period をおかないと carry over effects を発生します。それでは carry over effects が存在するかどうかを統計学的にいうにはどうしたらよいのでしょうか？

$$\underline{x_{Amean} = 3.798, SD_A = 1.060, n_A = 42}$$

$$x_{Bmean} = 3.536, SD_B = 1.056, n_B = 42$$

$$SD^2 = [(n_A - 1)SD_A^2 + (n_B - 1)SD_B^2]/(n_A + n_B - 2) = [(41)1.060^2 + (41)1.056^2]/(42 + 42 - 2) = 1.119$$

$$t = (x_{Amean} - x_{Bmean}) / [SD^2(1/n_A + 1/n_B)] = (3.798 - 3.536) / [1.119(1/42 + 1/42)] = 0.262/0.231 = 1.135, \quad p > 0.05,$$

よって本研究において carry over effect がないという仮説を棄却できませんでした。でも表をみた印象ではかなり強く cross over effect を疑うことができます。一般的に

carry over effect を検出する上の方法は power が弱く、一部の統計学者は $\alpha = 0.05$ ではなく $\alpha=0.10$ を推奨しています。それでも上のデータは有意差を検出できません。もしも有意差があったときはどうするべきでしょうか？薬の使用にかかわらず後半行なったデータは使用できません。よって A の最初を薬剤投与群として、B の最初を placebo として unpaired t test を行ないます。

Sample size calculation

Cross-over design では carry-over effects が存在しない限り通常の randomized clinical trial よりも少ない数で済みます。

$$n = S_d^2 (z_\alpha + z_\beta)^2 / 2\delta^2$$

$SD_d^2 = \text{variance of response on treatment A} - \text{variance of response on treatment 2}$

$\delta = \text{treatment benefit}$

閉経後女性の拡張期血圧に対する女性ホルモンの影響を調べようと思います。Aグループに対しては最初の 4 週間ホルモンを投与し、2 週間の washout period の後、4 週間 placebo を投与します。逆に B では最初の 4 週間 placebo を投与し、2 週間の washout period の後、4 週間ホルモンを投与します。血圧はそれぞれの患者さんで 4 週間の治療期間の終了時に測定し前半と後半で比較します。過去に行なったパイロット研究では 2 つの期間、拡張期血圧の差は 2mmhg で $S_d^2=31$ でした。それぞれのアームに何人の女性が必要ですか？80%の power で検討するものとします。

$$n = S_d^2 (z_\alpha + z_\beta)^2 / 2\delta^2 = 31 (1.96 + 0.84)^2 / 2 \times 2^2 = 30.4$$

よって各アーム 31 人、合計 62 人必要です。もしも通常の randomized clinical trial で行なったとすると、

$$n = 2S_d^2 (z_\alpha + z_\beta)^2 / \delta^2 = 4 \times 30.4 = 122 / \text{arm, total 244 subjects}$$

両者は同じ結果を出すのに必要とする患者さんの数において 4 倍も違います。