

SAS統計軟體操作(1)

醫學研究部 基礎醫學科 生統小組:陳韻伃 博士 授課日期:113年5月22日

SAS: The Role of Statistic in Clinical Trials

- Sample size estimation
- Analyze results (hypothesis testing and effect estimation)
- Control for confounding

SAS 9.4 起始畫面

585	SAS - [編	<u> </u> 露 - 未命名1]		_	
	檔案(F)	編輯(E) 檢視(V)	工具(T)	執行(R) 解決方案(S) 視窗(W) 說明(H)	_ 8 ×
	/		`	2 🗅 🛩 🖬 🖨 🖻 % 🖻 🛍 🗠 🐌 💁 * 🗙 🛈 🛷	
	SAS 環境」(SAS 環境」) 資料館 武的最変 文	<u>的内容</u>	X	libname data 'C:\Users\YYChen\Downloads\ 統計\data';	TSH
	副結果	🔍 檔案總領		副 輸出 - (未命名) 日誌 - (未命名) 武 編輯器 - 未命名1	
				C:\Users\YYChen Ln 1	, Col 1
				Program Editor	

設定SAS永久檔位置

	■ SAS - [編輯器 - 未命名1]	x c
	■ 檔案(F) 編輯(E) 檢視(V) 工具(T) 執行(R) 解決方案(S) 視窗(W) 說明(H)	_ 8 ×
	V D 🚔 🔲 / 🎒 🔃 / ½ 🖻 🛍 / ½ 🗐 🖳 / ½ 🔇 🏈	
	Ibname data 'D:\副研究員\中榮醫研部-生統小約 Ibname data 'D:\副研究員\中榮醫研部-生統小約 \全院教育課程規劃-2022oct\113年生統課程規劃 \20240522-SAS統計軟體操作-1-Amelia\課程講義 及教材\SAS dataset';	j]
/	檔案總管 「Data」的內容 名稱 大小 類型 IIID Drug_compliance 128.0KB Table IIIID Event 128.0KB Table IIIILab 128.0KB Table IIIILab 128.0KB Table IIIIILAB 128.0KB Table IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	~
	□ □<	>
	E C:\Users\YYChen Ln 1, Co	ol 1 🔡
	Results Questioner Dutput - (Untitled) Dutput - (Untitled) Dutput - (Untitled) Program Editor	

Clinical Trial Data Management using SAS

Clinical Trial Data Management using SAS mostly for the following purposes:

> Designing a protocol and beginning a study : Sample size

Health Data Management : Data Management using SAS

Analyzing Data

Designing a Protocol and Beginning a Study : Sample Size

▶ **範例**:想驗證新藥AAA的療效

6

■ 執行一個第三期、隨機、雙盲、多中心試驗,比較新藥AAA與安慰劑兩組在降血壓的療效上有無差異?

- ▶ /主要療效:8週後血壓 v.s. 基礎血壓質的差異
- ▶ 統計方法: 連續型變數的T-test
- ▶ **虛無假說**:新藥AAA與安慰劑兩組在降血壓的效果**無差異**
- ▶ 對立假說:新藥AAA與安慰劑兩組在降血壓的效果有差異

Designing a Protocol and Beginning a Study : Sample Size

虛無假說:新藥AAA與安慰劑兩組在降血壓的效果**無差異** 對立假說:新藥AAA與安慰劑兩組在降血壓的效果**有差異**

7

		實際	狀況	
		新藥 AAA 沒有降血 壓的效果	新藥 AAA 是有降血 壓的效果(8 週後至 少降 10 mmHg)	
t-test 檢 定	新藥 AAA 沒有降血 壓的效果	正確決定	型 II 誤差 β (type II error)	
後的決定	新藥 AAA 是有降血 壓的效果(8 週後至 少降 10 mmHg)	型 I 誤差 α (type I error)	正確決定 (1-β) (power)	

在樞紐性第三期臨床試驗相關法規規定, 整體型一誤差應控制在雙尾5%或以下; 整體型二誤差應控制在20%或以下; 檢定力 > 80%

Designing a Protocol and Beginning a Study : Sample Size

8

計算 連續型變數 要考慮的因素

1. 試驗目標	以安慰劑為對照之較優性試驗
2. 主要評估指標	8 週後的血壓值與基礎值血壓值之差異
3. 虛無假說 (H0)	$\mu_{Placebo} = \mu_{Drug AAA}$
4. 對立假說 (Ha)	$\mu_{Placebo} eq \mu_{Drug AAA}$
5. 統計檢定量	t-test
6. 型一誤差機率	<i>α</i> = 0.05 (雙尾)
7. 型二誤差機率 (檢定力)	$\beta = 0.20 (1 - \beta = 0.80)$
8. 預期療效大小 (d)	3 mmHg
9. 標準差 (σ)	8 mmHg
10. 受試者隨機分派比	1 : 1 (m : 1)
(AAA :placebo)	
11. 預期中途退出率(p)	p = 10%



人工樣本數計算

$$\frac{(1+1)^*(1.96+0.845)^2*8^2}{3^2} \approx 112, \quad \frac{112}{(1-0.1)^2} = 138$$

Designing a Protocol and Beginning a Study : Sample Size

■ 利用 SAS 計算 連續型變數 T-test 差異的方法

9

dm "log;clear;output;clear;program;recall"; title 'Clinical trials - two independent sample: contiuous variables'; PROC POWER; TWOSAMPLEMEANS TEST=DIFF MEANDIFF=3 STDDEV=8 APLHA=0.05 POWER=0.8 NPERGROUP=.; run; Clinical trials - two independent sample: contiuous variables

POWER 程序 平均差的雙樣本 t 檢定

固定案例元素			
分布	常態		
方法	精準		
Alpha	0.05		
平均差	3		
標準差	8		
名目次方	0.8		
側邊敷目	2		
Null 差異	0		

計算 每個群組的 N				
實際次方 每個群組的 N				
0.801 113				

Designing a Protocol and Beginning a Study : Sample Size

10

計算 類別型變數 要考慮的因素

1.	試驗目標	以安慰劑為對照之較優性試驗	
2.	主要評估指標	8 週後的血壓值<90 mmHg 的比率	
3.	虛無假說 (H0)	$p_{1(placebo)} = p_{2(AAA)}$	
4.	對立假說 (Ha)	$p_{1(placebo)} \neq p_{2(AAA)}$	
5.	統計檢定量	Z-test	
6.	型一誤差機率	<i>α</i> = 0.05 (雙尾)	-
7.	型二誤差機率 (檢定力)	$\beta = 0.20 (1 - \beta = 0.80)$	
8.	預期之療效大小(d)	10%	$P_2 - P_1 = 20\% - 10\% = 10\%$
9.	p ₁ 各組 <90 m	mHg 的比例 p ₁ = 10%	P ₂ = 10% + 10% = 20%
10.	受試者隨機分派比	1 : 1 (m : 1)	-
	(AAA :placebo)	(/	
11.	預期中途退出率 (p)	p = 5%	



人工樣本數計算



Designing a Protocol and Beginning a Study : Sample Size

▶ 利用 SAS 計算 類別型變數 比例差異的方法

11

title 'Cclinical trials - two independent sample: proportion difference '; PROC POWER; TWOSAMPLEFREQ TEST=PCHI GROUPPROPORTIONS = (0.1 0.2) NULLPROPORTIONDIFF = 0 ALPHA=0.05 POWER=0.8 NPERGROUP=.; run; Cclinical trials - two independent sample: proportion difference

POWER 程序 兩個比例的 Pearson 卡方棱定

固定案例元素				
分布	漸近常態			
方法	常態近似			
Null 比例差異	0			
Alpha	0.05			
群組1比例	0.1			
群組 2 比例	0.2			
名目次方	0.8			
側邊敷目	2			

計算 每個群組的 N				
每個群組的 N	實際次方			
199	0.800			

Designing a Protocol and Beginning a Study : 12 Sample Size

利用 SAS 計算 兩組相對風險 的方法 title 'Clinical trials - two independent sample: relative risk'; PROC POWER; TWOSAMPLEFREQ TEST=PCHI RELATIVERISK=0.6667 REFPROPORTION=0.6 ALPHA=0.05 POWER=0.8 NPERGROUP=.; run;

REFPROPORTION: 參考組 事件數 的比例

Clinical trials - two independent sample: relative risk

POWER 程序 兩個比例的 Pearson 卡方被定

固定案例元素				
分布	漸近常態			
方法	常態近似			
Alpha	0.05			
参考 (群組 1) 比例	0.6			
相對風險	0.6667			
名目次方	0.8			
側邊數目	2			
Null 相對風險	1			

計算 每個群組的 N					
實際次方 每個群組的 N					
0.800	97				

Clinical Trial Data Management using SAS

Clinical Trial Data Management using SAS mostly for the following purposes:

- > Designing a protocol and beginning a study : Sample size
- Health Data Management : Data Management using SAS
- Analyzing Data

Data Management using SAS: 保留重要分析變數

14

```
/*清理資料*/
                                      只有data.subject檔案有Group
data data.subject; set data.subject;
keep Center_no Screening_no Ran_no Group Visit1-Visit6 Gender Age; *保留必要的變數;
      刪除不必要的變數;
*drop
if Center no=. then delete;
if Group=. then delete; run;
data data.Drug compliance; set data.Drug compliance;
     Center no Screening no Ran no Visit Visit date taken no returned no lost no period day
keep
compliance r;
if Center no=. then delete; run;
data data.event; set data.event;
keep Center no Screening no Ran no Visit start visit start visit date end visit end visit date
total no record AF total;
if Center no=. then delete; run;
data data.vital sign; set data.vital sign;
keep Center no Screening no Ran no Visit sbp dbp;
if Center no=. then delete; run;
```

data data.lab; set data.lab; keep Center_no Screening_no Ran_no Visit hemoglobin WBC RBC AST ALT creatinine; if Center_no=. then delete; run;

Data Management using SAS ■ 利用 SAS 進行 資料管理

短資料格式 適合治療前後 (配對) 差異比較							
ID	Group	SBP_Visit_1	SBP_Visit_2	SBP_Visit_3	DBP_Visit_1	DBP_Visit_2	DBP_Visit_3
\$101	1	98	120	99	62	68	63
S102	2	141	130	168	91	78	82
長資料格式		適合重複測量分析					
ID	Group	Visit	SBP	DBP	AST	Creatinine	AF_event
\$101	1	1	98	62	43	70	0
S101	1	2	120	68	65	89	1
\$101	1	3	99	63	51	67	0
S102	2	1	141	91	55	111	1
S102	2	2	130	78	90	99	1
\$102	2	3	168	82	41	83	1

Data Management using SAS ■ 利用 sas 進行 資料管理

短資料格式:Data.subject_A

短資料格式:Data.subject_B

	Center_no	Screening_no	Ran_no	Group	Visit1	Visit2	Visit3	Visit4	Visit5	Visit6	Gender	Age
1	1	S101	F-R101	2	02SEP2019	02SEP2019	09SEP2019	23SEP2019	28NOV2019	10FEB2020	1	70.024657534
2	1	S102	R-R101	1	21FEB2020	21FEB2020	****	*****	****	06AUG2020	1	64.084931507
3	1	S103	R-R102	1	09JUL2020	09JUL2020	16JUL2020	27JUL2020	090CT2020	28DEC2020	1	64.44109589
4	2	S201	F-R201	2	10JUL2019	10JUL2019	18JUL2019	06AUG2019	010CT2019	26DEC2019	1	37.260273973
5	2	S202	F-R202	1	28AUG2019	28AUG2019	10SEP2019	24SEP2019		26NOV2019	1	66.728767123
6	2	S203	F-R203	1	29NOV2019	29NOV2019	10DEC2019	24DEC2019	***	09JUN2020	1	44.391780822
7	2	S204	R-R201	1	03JAN2020	03JAN2020	14JAN2020	04FEB2020	21APR2020	14JUL2020	0	68.726027397
8	3	S301	R-R301	2	010CT2019	010CT2019	080CT2019	220CT2019	31DEC2019	***	0	67.260273973
9	3	\$303	F-R301	2	150CT2019	150CT2019	220CT2019	05NOV2019		14JAN2020	1	71.671232877
10	3	S304	R-R302	1	01NOV2019	01NOV2019	08NOV2019	22NOV2019	31JAN2020	24APR2020	1	68.205479452
		· · · · · · · · · · · · · · · · · · ·										

先sort 照相同id排序,再直向 合併不同檔案:set 短資料格式:Data.subject

												4			
	Center_no Screening_no	Ran_no	Group	Visit1	Visit2	Visit3	Visit4	Visit5	Visit6	Gender	Age				
12	3 \$306	F-R303		2 22NOV2019	22NOV2019	29NOV2019	13DEC2019	21FEB2020	****	0	70.205479452				
13	3 \$307	R-R303		2 10DEC2019	10DEC2019	17DEC2019	31DEC2019	****	02JUN2020	1	48.315068493				
14	3 \$308	R-R304		1 13DEC2019	13DEC2019	20DEC2019	03JAN2020		05JUN2020	1	75.638356164				
15	3 \$309	R-R305		2 14JAN2020	14JAN2020	21JAN2020	04FEB2020	14A	a aant d	lata-data		7.			
16	3 \$310	R-R306		2 04FEB2020	04FEB2020	12FEB2020	26FEB2020	*** PLO		iata-uata	. subject	_^,			
17	3 \$311	R-R307		1 20FEB2020	20FEB2020	27FEB2020	***	*** рЛ	Center_	_no Scre	ening_no	;	run;	* 龙貞科俗工: \	'isit
18	3 \$312	F-R304		1 27FEB2020	27FEB2020	***	****	*** 1-V	isit 6;						
19	3 \$313	R-R308		1 *****	*****	****	01APR2020	10Л							
20	3 \$314	R-R309		2 03APR2020	03APR2020	10APR2020	30APR2020	03J pro	a sort d	lata=data	a subject	в:			
								P10							
								by	Center	no Scre	ening no	;	<pre>run;</pre>	*短資料格式: 、	/isit
								1-V	isit 6;	_	—				
								dat	a data.s	subject;					

- set data.subject_A (in=a) data.subject_B (in=b) ;
- by Center_no Screening_no ; run;

Data Management using SAS ■ 利用 sas 進行 資料管理

短資料格式:Data.subject

	Center_no	Screening_no	Ran_no	Group	Visit1	Visit2	Visit3	Visit4	Visit5	Visit6	Gender	Age
1	1	S101	F-R101	2	02SEP2019	02SEP2019	09SEP2019	23SEP2019	28NOV2019	10FEB2020	1	70.024657534
2	1	S102	R-R101	1	21FEB2020	21FEB2020	****	****	****	06AUG2020	1	64.084931507
3	1	\$103	R-R102	1	09JUL2020	09JUL2020	16JUL2020	27JUL2020	090CT2020	28DEC2020	1	64.44109589
4	2	S201	F-R201	2	10JUL2019	10JUL2019	18JUL2019	06AUG2019	010CT2019	26DEC2019	1	37.260273973
5	2	S202	F-R202	1	28AUG2019	28AUG2019	10SEP2019	24SEP2019		26NOV2019	1	66.728767123
6	2	S203	F-R203	1	29NOV2019	29NOV2019	10DEC2019	24DEC2019	***	09JUN2020	1	44.391780822
7	2	S204	R-R201	1	03JAN2020	03JAN2020	14JAN2020	04FEB2020	21APR2020	14JUL2020	0	68.726027397
8	3	\$ \$301	R-R301	2	010CT2019	010CT2019	080CT2019	220CT2019	31DEC2019	***	0	67.260273973
9	3	\$ \$303	F-R301	2	150CT2019	150CT2019	220CT2019	05NOV2019		14JAN2020	1	71.671232877
10	3	\$ \$304	R-R302	1	01NOV2019	01NOV2019	08NOV2019	22NOV2019	31JAN2020	24APR2020	1	68.205479452

橫向 合併不同檔案:merge

長資料格式:Data.vital_sign

	0.1	a :	D	TT: 11	aron	DDD
	Center_no	Screening_no	Ran_no	V isit	2Bb	DBP
1	1	S101	F-R101	1	138	79
2	1	S101	F-R101	2	138	79
3	1	S101	F-R101	3	144	74
4	1	S101	F-R101	4	140	68
5	1	S101	F-R101	5	121	67
6	1	S101	F-R101	6	126	74
7	1	S102	R-R101	1	133	88
8	1	S102	R-R101	2	133	88
9	1	S102	R-R101	3	121	69
10	1	S102	R-R101	4	133	89
11	1	S102	R-R101	5	125	80
12	1	S102	R-R101	6	127	81

長資料格式:Data.lab

	Center_no	Screening_no	Ran_no	Visit	Hemoglobin	RBC	WBC	Creatinine
1	1	S101	F-R101	1	15.8	5.31	5440	1.15
2	1	S101	F-R101	3	15.6	5.29	5510	1.06
3	1	S101	F-R101	6	14.7	5.06	5050	1.24
4	1	S102	R-R101	1	13.3	4.54	5260	1.21
5	1	S102	R-R101	3	12.7	4.4	6070	1.14
6	1	S102	R-R101	6	13.6	4.54	6.34	1.1
7	1	S103	R-R102	1	16.8	5.41	9.09	1.33
8	1	S103	R-R102	3	16.3	5.01	8.93	1.2
9	1	S103	R-R102	6	12.8	3.94	8.36	1.23
10	2	S201	F-R201	1	16.2	4.98	5200	1.22
11	2	S201	F-R201	3	15.5	4.72	9100	1.1
12	2	S201	F-R201	6	15.6	4.6	7600	1.2

Data Management using SAS ■ 利用 SAS 進行 資料管理

先sort 照相同id排序,再横向 合併不同檔案:merge

```
proc sort data=data.subject;
              Screening no ; run; *短資料格式: Visit 1-Visit 6;
by Center no
proc sort data=data.Drug compliance;
by Center no Screening no Visit; run; *長資料格式→有2次資料: Visit 3, Visit 5;
proc sort data=data.event;
by Center no Screening no Visit; run; *長資料格式→有4次資料: Visit 1, Visit 3, Visit 4, Visit
6;
proc sort data=data.Vital sign; *長資料格式→有六次資料: Visit 1-Visit 6;
by Center no Screening no Visit; run;
proc sort data=data.lab;
by Center no Screening no Visit; run; *長資料格式→有3次資料: Visit 1, Visit 3, Visit 6;
data data.all;
merge data.Vital sign (in=a) data.Drug compliance (in=b) data.event (in=c) data.lab
(in=d)
by Screening no visit ; run;
data data.final;
merge data.subject (in=a) data.all (in=b) ;
     Screening no ; run;
by
```

Data Management using SAS

▶ 利用 SAS 進行 資料管理

長資料格式:Data.final

先sort 照相同id排序,再橫向 合併不同檔案:merge

	Center_no Screening_no Ran_no	Group	Visit1	Visit2	Visit3	Visit4	Visit5	Visit6	Gender	Age	Visit	SBP	DBP	Visit_date	Taken_no	Returned_no	Lost_no	Period_day
1	1 S101 F-R101		2 02SEP2019	02SEP2019	09SEP2019	23SEP2019	28NOV2019	10FEB2020	1	70.024657534	1	138	79					
2	1 S101 F-R101		2 02SEP2019	02SEP2019	09SEP2019	23SEP2019	28NOV2019	10FEB2020	1	70.024657534	2	138	79					
3	1 S101 F-R101		2 02SEP2019	02SEP2019	09SEP2019	23SEP2019	28NOV2019	10FEB2020	1	70.024657534	3	144	74	09SEP2019	200	84	10	J 80
4	1 \$101 F-R101		2 02SEP2019	02SEP2019	09SEP2019	23SEP2019	28NOV2019	10FEB2020	1	70.024657534	4	140	68					
5	1 \$101 F-R101		2 02SEP2019	02SEP2019	09SEP2019	23SEP2019	28NOV2019	10FEB2020	1	70.024657534	5	121	67	28NOV2019	212	82	C	J 74
6	1 \$101 F-R101		2 02SEP2019	02SEP2019	09SEP2019	23SEP2019	28NOV2019	10FEB2020	1	70.024657534	6	126	74					
7	1 \$102 R-R101		1 21FEB2020	21FEB2020	****	****	****	06AUG2020	1	64.084931507	1	133	88					
8	1 S102 R-R101		1 21FEB2020	21FEB2020	****	****	***	06AUG2020	1	64.084931507	2	133	88					
9	1 \$102 R-R101		1 21FEB2020	21FEB2020	****	****	***	06AUG2020	1	64.084931507	3	121	69	***	203	77	14	4 72
10	1 S102 R-R101		1 21FEB2020	21FEB2020	****	****	***	06AUG2020	1	64.084931507	4	133	89					
11	1 S102 R-R101		1 21FEB2020	21FEB2020	****	****	****	06AUG2020	1	64.084931507	5	125	80	***	252	42		J 84
12	1 \$102 R-R101		1 21FEB2020	21FEB2020	****	****	****	06AUG2020	1	64.084931507	6	127	81					
13	1 \$103 R-R102		1 09JUL2020	09JUL2020	16JUL2020	27JUL2020	090CT2020	28DEC2020	1	64.44109589	1	138	78					
14	1 \$103 R-R102		1 09JUL2020	09JUL2020	16JUL2020	27JUL2020	090CT2020	28DEC2020	1	64.44109589	2	138	78					
15	1 \$103 R-R102		1 09JUL2020	09JUL2020	16JUL2020	27JUL2020	090CT2020	28DEC2020	1	64.44109589	3	118	69	16JUL2020	252	42		J 85
16	1 \$103 R-R102		1 09JUL2020	09JUL2020	16JUL2020	27JUL2020	090CT2020	28DEC2020	1	64.44109589	4	122	68					
17	1 \$103 R-R102		1 09JUL2020	09JUL2020	16JUL2020	27JUL2020	090CT2020	28DEC2020	1	64.44109589	5	140	75	090CT2020	229	65		J 80
18	1 \$103 R-R102		1 09JUL2020	09JUL2020	16JUL2020	27JUL2020	09OCT2020	28DEC2020	1	64.44109589	6	158	86					

Data Management using SAS

長

▶ 利用 SAS 進行 資料管理

資料格式:Vital_sign	/*長資料改	女成短資料格式*	/			
	data vi	.tal_sign; <mark>se</mark>	t data	.vital_sign;		
	SBP1=.;	if visit=1	then	SBP1=SBP;		
	SBP2=.;	if visit=2	then	SBP2=SBP;		
	SBP3=.;	if visit=3	then	SBP3=SBP;		
	SBP4=.;	if visit=4	then	SBP4=SBP;		
	SBP5=.;	if visit=5	then	SBP5=SBP;		
	SBP6=.;	if visit=6	then	SBP6=SBP;	run;	

檔室總管	x		Center_no Scree	ning_no Ran_no	Visit	SBP	DBP	SBP1	SBP2	SBP3	SBP4	SBP5	SBP6
[W-1] (bbb x		1	1 \$101	F-R101	1	138	79	138					
Work」的内容		2	1 \$101	F-R101	2	138	79		138				
名稱	大小 類型	3	1 \$101	F-R101	3	144	74			144			
Final	256.0KB Tak	4	1 \$101	F-R101	4	140	68				140		
Vital sign	128 OKB Tak	5	1 \$101	F-R101	5	121	67					121	
	120.000 140	6	1 \$101	F-R101	6	126	74						126
Vital_sign1	128.0KB Tab	7	1 \$102	R-R101	1	133	88	133					
Vital_sign2	128.0KB Tab	8	1 \$102	R-R101	2	133	88		133				
Vital_sign3	128.0KB Tak	9	1 \$102	R-R101	3	121	69			121			
Vital sign4	128.0KB Tak	10	1 \$102	R-R101	4	133	89				133		
Vital sign5	128.0KB Tak	11	1 \$102	R-R101	5	125	80					125	
	120.000 Tac	12	1 \$102	R-R101	6	127	81						127
Vital_signb	128.0KB Tab	13	1 \$103	R-R102	1	138	78	138					
		14	1 \$103	R-R102	2	138	78		138				
		15	1 \$103	R-R102	3	118	69			118			
		16	1 \$103	R-R102	4	122	68				122		
		17	1 \$103	R-R102	5	140	75					140	
		18	1 \$103	R-R102	6	158	86						158
		10 40	0 0004	7 P.004									

Data Management using SAS

▶ 利用 SAS 進行 資料管理

短資料格式: Vital_sign1 - Vital_sign6

1						
<u></u> 宏總管		x		Center_no	Screening_no	SBP1
			1	1	S101	138
work」的內容			2	1	S102	133
名稱	大小	類型	3	1	S103	138
Final	256.0KB	Tab	4	2	S201	148
Vital sign	128 OKB	Tab	5	2	S202	138
Vital sign	120.010	T-1	6	2	S203	121
Vital_sign i	128.UKB	Tab	7	2	S204	120
Vital_sign2	128.0KB	Tab	8	3	\$301	180
Vital_sign3	128.0KB	Tab	9	3	\$302	131
Vital_sign4	128.0KB	Tab	10	3	\$303	123
Vital sign5	128 OKB	Tab	11	3	\$304	135
	120.010	Tab	12	3	\$305	122
e vital_signo	128.UKB	Tab	13	3	\$306	119
			14	3	\$307	124
			15	3	\$308	128
			16	3	\$309	153
			17	3	\$310	127
			18	3	S311	125

SBP1, SBP2.....SBP5, SBP6

檔案總管		x			Center_no	Screening_no	SBP6
			1	1	1	S101	126
			Ш	2	1	S102	127
名稱	大小	類型		3	1	S103	158
Final	256.0KB	Tab	Ш	4	2	S201	125
Vital sign	128 OKB	Tab		5	2	S202	119
	420.0/0	- L	Ш	6	2	S203	116
Vital_sign1	128.0KB	Тар	Ш	7	2	S204	130
Vital_sign2	128.0KB	Tab	Ш	8	3	\$301	135
Vital_sign3	128.0KB	Tab		9	3	\$303	123
Vital_sign4	128.0KB	Tab		10	3	\$304	126
Vital sign5	128.0KB	Tab	Ш	11	3	\$305	133
Withol sing 6	100.0//0	Tab	Ш	12	3	\$306	135
Vital_signo	126.UKB	Tab	Ш	13	3	\$307	129
			Ш	14	3	\$308	111
			Ш	15	3	\$309	132
				16	3	\$310	129
				17	3	\$311	125
				18	3	\$312	116

data vital_sign1; set vital_sign; if visit=1; keep Center_no
data vital_sign2; set vital_sign; if visit=2; keep Center_no
data vital_sign3; set vital_sign; if visit=3; keep Center_no
data vital_sign4; set vital_sign; if visit=4; keep Center_no
data vital_sign5; set vital_sign; if visit=5; keep Center_no
data vital_sign6; set vital_sign; if visit=6; keep Center_no
Screening_no sbp6; run;

data data.vital_sign_m; merge vital_sign1 vital_sign2 vital_sign3 vital_sign4 vital_sign5
vital_sign6;

Data Management using SAS ■ 利用 sas 進行 資料管理

短資料格式:data.Vital_sign_m

data	<pre>vital_sign1;</pre>	set	vital_sign;	if	visit= 1;	keep	Center_no	Screening_no	sbp1;	<pre>run;</pre>
data	<pre>vital_sign2;</pre>	set	vital_sign;	if	visit= 2;	keep	Center_no	Screening_no	sbp2;	<pre>run;</pre>
data	vital_sign3;	set	vital_sign;	if	visit= 3;	keep	Center_no	Screening_no	sbp3;	<pre>run;</pre>
data	vital_sign4;	set	vital_sign;	if	visit= 4;	keep	Center_no	Screening_no	sbp4;	<pre>run;</pre>
data	vital_sign5;	set	vital_sign;	if	visit= 5;	keep	Center_no	Screening_no	sbp5;	<pre>run;</pre>
data	vital sign6;	set	vital sign;	if	visit= 6;	keep	Center no	Screening no	sbp6;	run;

data data.vital sign m; merge vital sign1 vital sign2 vital sign3 vital sign4 vital sign5 vital sign6;

Screening_no; run; by

88	B CAC	IV/IE\A/TABLE	Data Vital	sign ml	
	9 SAS -	UNLIVIADLE.	Data.vital_	agu_mj	

	檔案(F)	編輯(E)	檢視(V)	工具(T)	資料(D)	解決方案(S)	視窗(W)	說明(H)
--	-------	-------	-------	-------	-------	---------	-------	-------

✓	~	Ľ	🖻 🔳 🖨	🗿 🖪 🖻 🖻 🗠 🗙	↓ª ↓z 🏢	= 6 6	🛐 🔯 🧇			
宝宝绚智		x		Center_no Screening_no	SBP1	SBP2	SBP3	SBP4	SBP5	SBP6
		_	1	1 \$101	138	138	144	140	121	126
Data] 的内容			2	1 \$102	133	133	121	133	125	127
名稱	大小	類型	3	1 \$103	138	138	118	122	140	158
All	192.0KB	Tab	4	2 \$201	148	148	139	131	133	125
Drug compliance	128 OKB	Tab	5	2 \$202	138	138	134	128		119
E compliance	120.000		6	2 \$203	121	121	127	114	122	116
Event	128.0KB	Tab	7	2 \$204	120	120	118	133	143	130
📑 Final	192.0KB	Tab	8	3 \$301	180	180	137	138	142	135
📑 Lab	128.0KB	Tab	9	3 \$302	131	131				
Subject	128.0KB	Tab	10	3 \$303	123	123	131	143		123
Vital sign	128.0KB	Tab	11	3 \$304	135	135	129	131	154	126
	120.0//0	Tab	12	3 \$305	122	122	102	106	132	133
vital_sign_m	120.UKB	Tab	13	3 \$306	119	119	119	113	108	135
			14	3 \$307	124	124	140	126	127	129
			15	3 \$308	128	128	127	124	134	111
			16	3 \$309	153	153	129	106	133	132
			17	3 \$310	127	127	138	145	140	129
			18	3 \$311	125	125	136	128	135	125

Clinical Trial Data Management using SAS

Clinical Trial Data Management using SAS mostly for the following purposes:

- > Designing a protocol and beginning a study : Sample size
- Health Data Management : Data Management using SAS
- Analyzing Data

▶ 利用 SAS 進行 資料計算

```
data data.vital sign m; set data.vital sign m;
mean SBP = mean (SBP1, SBP2, SBP3, SBP4, SBP5, SBP6);
Avg SBP=(SBP1+SBP2+SBP3+SBP4+SBP5+SBP6)/6;
min SBP = min (SBP1, SBP2, SBP3, SBP4, SBP5, SBP6);
max SBP = max (SBP1, SBP2, SBP3, SBP4, SBP5, SBP6);
SBP sqrt=sqrt (mean SBP) ;
SBP square=Avg SBP**2 ;
SBP exp=exp (mean SBP) ;
run;
```

短資料	格式:	data.V	ital_sig	n_m								
/ /												
SBP1	SBP2	SBP3	SBP4	SBP5	SBP6	mean_SBP	Avg_SBP	min_SBP	max_SBP	SBP_sqrt	SBP_square	SBP_exp
138	138	144	140	121	126	134.5	134.5	121	144	11.597413505	18090.25	2.5858767E58
133	133	121	133	125	127	128.66666667	128.66666667	121	133	11.343133018	16555.111111	7.572222E55
138	138	118	122	140	158	135.66666667	135.66666667	118	158	11.647603473	18405.444444	8.3039497E58
148	148	139	131	133	125	137.33333333	137.33333333	125	148	11.718930554	18860.444444	4.3965179E59
138	138	134	128		119	131.4		119	138	11.462983905		1.1649168E57
121	121	127	114	122	116	120.16666667	120.16666667	114	127	10.962055768	14440.027778	1.5407077E52
120	120	118	133	143	130	127.33333333	127.33333333	118	143	11.284207253	16213.777778	1.996016E55
180	180	137	138	142	135	152	152	135	180	12.328828006	23104	1.0298198E66
131	131					131		131	131	11.445523142		7.8086711E56
123	123	131	143		123	128.6		123	143	11.340194002		7.0838666E55

利用SAS進行常態性檢定資料分佈:

/*Normality Test*/
data lab; set data.final;
if visit=1; keep Group age SBP AST Creatinine; run;

proc univariate data=lab normal ;
var sbp;

class group; run;

	UNIVARIATE 程序 變數: SBP (SBP) Group = 2							
V	lean (s	std) 🔹	的差					
	N	34	總和權重	34				
	平均值	130.558824	總和觀測	4439				
	標準差	18.6664359	變異勲	348.435829				
	偏態	0.42712288	峰度	0.42294941				
	未校正平方和	591049	校正平方和	11498.3824				
	變異係數	14.2973377	標準誤差平均值	3.20126734				

SAS 系統

	基本統計量值								
位置		變異性							
平均值	130.5588	標準差	18.66644						
中位熱	129.5000	變異數	348.43583						
眾數	119.0000	全距	86.00000						
		內四分位距	21.00000						

分位數 (定	義 5)
層級	分位數
100% 最大值	180.0
99%	180.0
95%	164.0
90%	154.0
75% Q3	140.0
50% 中位熱	129.5
25% Q1	119.0
10%	107.0
5%	101.0
1%	94.0
0% 最小值	94.0

常態性檢定							
檢定	新	計值	p 值				
Shapiro-Wilk	w	0.981536	Pr < W	0.8198			
Kolmogorov-Smirnov	D	0.100622	Pr > D	>0.1500			
Cramer-von Mises	W-Sq	0.046868	Pr > W-Sq	>0.2500			
Anderson-Darling	A-Sq	0.26081	Pr > A-Sq	>0.2500			

常態性檢定→ 符合常態分佈 P >0.05 Kolmogorov-Smirnov (K-S) 檢定: 樣本數50個以上

Shapiro-Wilk (S-W) 檢定: 樣本數50個以下

26

Diabetes Care 2019;42:1067–1074 | https://doi.org/10.2337/dc19-0093



Figure 1—Median (IQR) change from baseline to week 24 in HbA_{1c} (P = 0.018) (A), FPG (P = 0.27) (B), and 2-h glucose after an OGTT (P = 1) (C) in the two intervention groups.

利用SAS進行常態性檢定資料分佈:

proc sort data=lab; by group; proc univariate data=lab normal plot; var sbp; by group; run;



利用SAS進行常態性檢定資料分佈:

```
proc univariate data=lab ;
```

title histogram; histogram sbp / midpoint = 60 to 200 by 5 Normal; class group; run;

Histogram



長資料格式:data.final

▶ 利用 SAS 進行 描述性統計分析:連續型變數

PROC MEANS DATA=data.final
MAXDEC=2 N MEAN STD MEDIAN Q1 Q3
/*MEDIAN Q1 Q3 MIN MAX N VAR SKEWNESS KURTOSIS T PROBT CLM /*CLM 樣本信賴區間*/;
VAR SBP DBP Creatinine AST ALT ;
CLASS visit ;
/*BY gender ; */ RUN;

Visit	觀測值數目	變數	標籤	N	平均值	標準差	中位數	下四分位數	上四分位熱	Visit
1	76	SBP	SBP	76	132.80	17.30	130.00	121.50	140.00	4
		DBP	DBP	76	80.20	11.26	79.00	73.50	89.00	
		Creatinine	Creatinine	75	1.05	0.32	0.99	0.86	1.15	
		AST	AST	76	22.71	10.44	20.00	17.00	23.00	
		ALT	ALT	76	24.86	20.70	20.00	15.00	25.50	
2	73	SBP	SBP	73	132.77	17.84	130.00	121.00	140.00	5
		DBP	DBP	73	79.85	11.58	79.00	73.00	89.00	
		Creatinine	Creatinine	0						
		AST	AST	0						
		ALT	ALT	0						
3	71	SBP	SBP	71	128.68	15.69	127.00	120.00	138.00	6
		DBP	DBP	71	77.55	10.45	78.00	70.00	86.00	
		Creatinine	Creatinine	71	1.03	0.29	1.00	0.87	1.18	
		AST	AST	71	22.15	9.33	20.00	17.00	23.00	
		ALT	ALT	71	23.28	16.33	20.00	15.00	26.00	

Visit	觀測值敷目	變數	標籤	N	平均值	標準差	中位熱	下四分位熱	上四分位熱
4	71	SBP	SBP	69	125.58	16.29	125.00	114.00	135.00
		DBP	DBP	69	78.03	14.18	76.00	70.00	82.00
		Creatinine	Creatinine	0					
		AST	AST	0					
		ALT	ALT	0					
5	71	SBP	SBP	64	128.08	14.53	130.00	119.50	138.50
		DBP	DBP	64	77.83	11.09	79.00	68.50	85.50
		Creatinine	Creatinine	0					
		AST	AST	0					
		ALT	ALT	0					
6	71	SBP	SBP	71	128.90	15.30	128.00	119.00	138.00
		DBP	DBP	71	77.32	9.99	79.00	70.00	85.00
		Creatinine	Creatinine	68	1.03	0.28	1.00	0.89	1.12
		AST	AST	68	21.10	7.72	20.00	17.00	23.00
		ALT	ALT	68	21.54	10.77	19.00	15.00	27.00

/*類別型變數 + Chi-square test*/
data lab; set data.final;
if visit=1; keep Group SBP gender; run;

PROC FREQ data=lab;

table gender*group / chisq expected exact ; run;

▶ 利用 SAS 進行 描述性統計分析:

類別型變數 + Chi-square test

SAS 系統

FREQ 程序

の動	Table of G	ender b	y Group			
朝		Gro	oup(Group)			
」 百分比	Gender(Gender)	1	2	總計		
百分比	0	7	8	15		
		7.8169	7.1831			
		9.86	11.27	21.13		
		46.67	53.33			
		18.92	23.53			
	1	30	26	56		
		29.183	26.817			
		42.25	36.62	78.87		
		53.57	46.43			
		81.08	76.47			
	總計	37	34	71		
		52.11	47.89	100.00		
	次郵)遺漏 = !	5			

表格 Group-Gender*s 的統計值

統計值	自由度	值	機率
卡方	1	0.2260	0.6345
概度比卡方	1	0.2259	0.6346
連續性調整卡方	1	0.0340	0.8537
Mantel-Haenszel 卡方	1	0.2228	0.6369
Phi 係熱		-0.0564	
列聯係數		0.0563	
Cramer V		-0.0564	

Fisher 精準檢定						
儲存格 (1,1) 次敷 (F)	7					
<u>左</u> 邊 Pr <= F	0.4262					
右邊 Pr >= F	0.7782					
表格機率 (P)	0.2044					
雙邊 Pr <= P	0.7729					

● 利用 SAS 進行 描述性統計分析: Student's t test

```
/*Student's t test*/
data lab; set data.final;
if visit=1; keep Group SBP gender; run;
```

PROC TTEST DATA=lab ;
CLASS group; VAR SBP ; RUN;

TTEST 程序

變勳:SBP (SBP)

Group	N	平均值	標準差	標準誤差	最小值	最大值
1	37	134.9	17.1188	2.8143	107.0	174.0
2	34	130.6	18.6664	3.2013	94.0000	180.0
Diff (1-2)		4.3871	17.8757	4.2467		

Group	方法	平均值	平均值 95% CL 平均值		標準差	95% CL 標準差	
1		134.9	129.2	140.7	17.1188	13.9212	22.2366
2		130.6	124.0	137.1	18.6664	15.0559	24.5702
Diff (1-2)	集區	4.3871	-4.0848	12.8591	17.8757	15.3269	21.4491
Diff (1-2)	Satterthwaite	4.3871	-4.1207	12.8949			

方法	變異數	自由度	t值	Pr > [t]
集區	均等	69	1.03	0.3052
Satterthwaite	不均等	67.023	1.03	0.3071

變異熱相等性					
方法	法 分子自由度 分母自由度		F值	Pr > F	
Folded F	33	36	1.19	0.6104	





Thank you for listening

東中資料庫上手論文寫