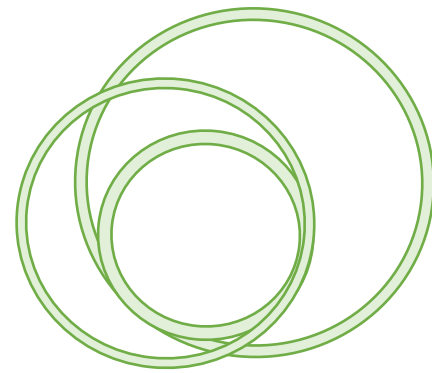




# 如何運用SPSS進行臨床資料分析 ~以Appendicitis為例~

2016.02.24



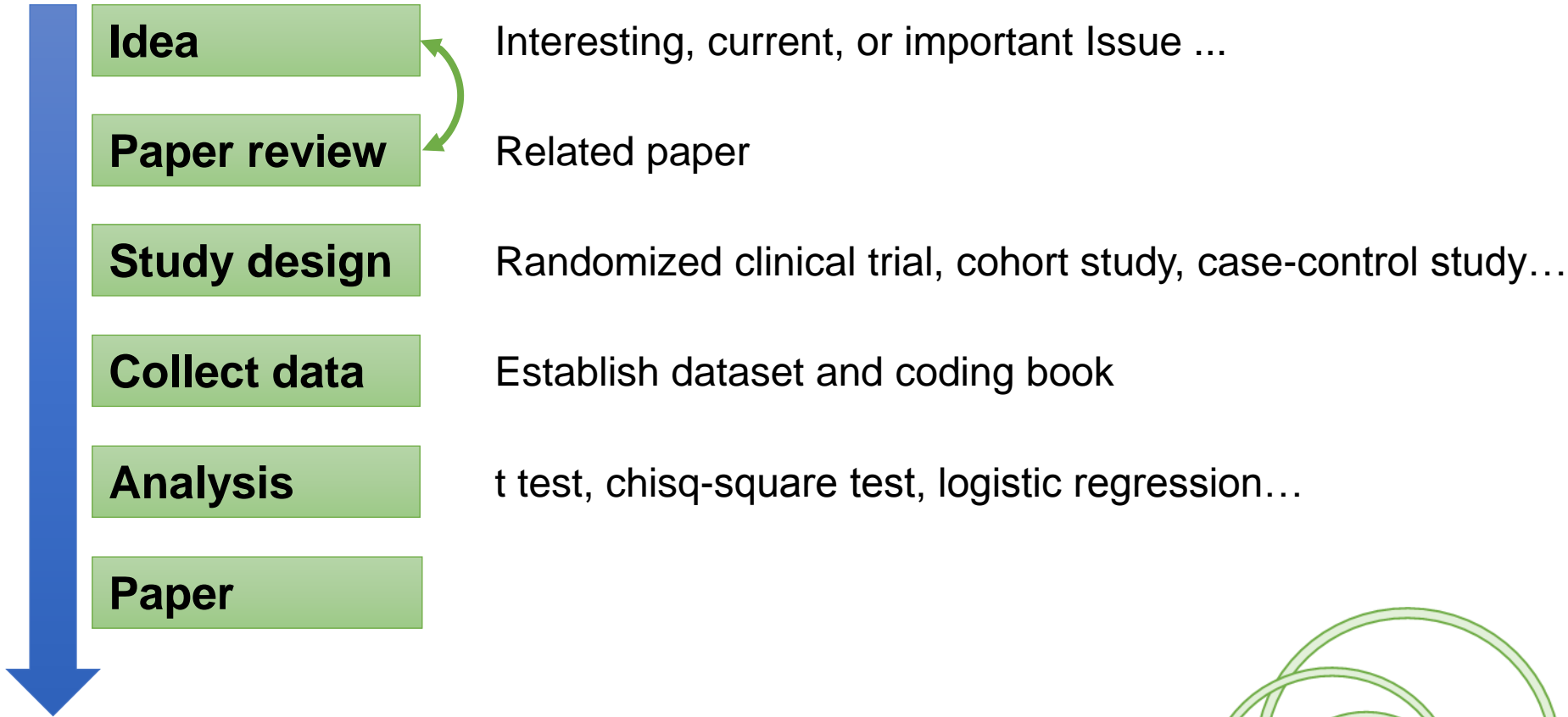
---

先將檔案下載並存放在桌面。  
路徑：\公用資料夾\醫學研究部\臨床醫學研究中心  
20160224\_SPSS課程



# Introduction

---



# Idea and paper review

我該以什麼臨床資料做為本次課程的範例???

外科

手術

Appendicitis and Appendectomy

## Review paper

Original Investigation

### Antibiotic Therapy vs Appendectomy for Treatment of Uncomplicated Acute Appendicitis The APPAC Randomized Clinical Trial

Paulina Salminen, MD, PhD; Hannu Paajanen, MD, PhD; Tero Rautio, MD, PhD; Pia Nordström, MD, PhD; Markku Aarnio, MD, PhD; Tuomo Rantanen, MD, PhD; Risto Tuominen, MPH, PhD; Saija Hurme, MSc; Johanna Virtanen, MD; Jukka-Pekka Mecklin, MD, PhD; Juhani Sand, MD, PhD; Airi Jartti, MD; Irina Rinta-Kiikka, MD, PhD; Juha M. Grönroos, MD, PhD

(JAMA 2015)

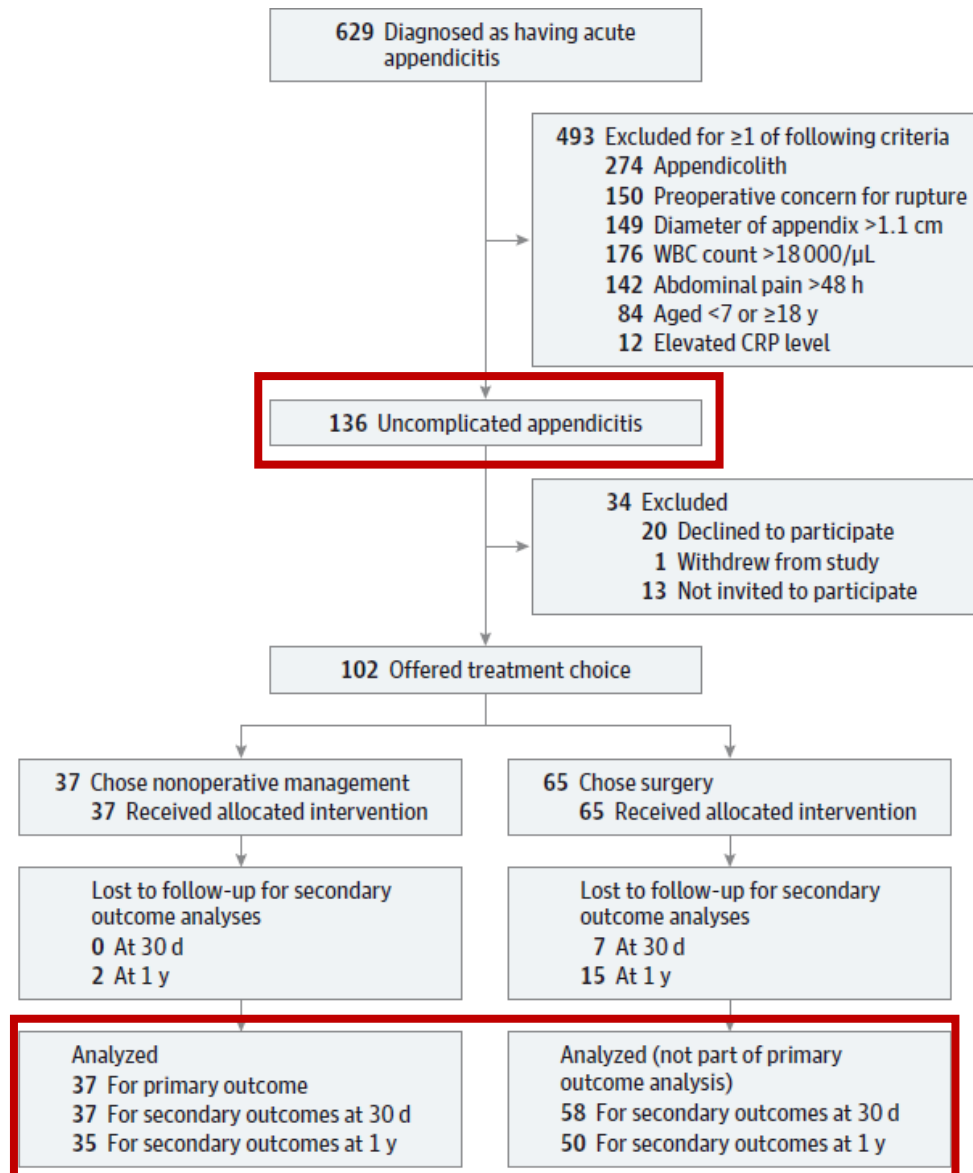
Original Investigation

### Effectiveness of Patient Choice in Nonoperative vs Surgical Management of Pediatric Uncomplicated Acute Appendicitis

Peter C. Minneci, MD, MHSc; Justin B. Mahida, MD, MBA; Daniel L. Lodwick, MD, MS; Jason P. Sulkowski, MD; Kristine M. Nacion, MPH; Jennifer N. Cooper, PhD, MS; Erica J. Ambeba, PhD, MPH; R. Lawrence Moss, MD; Katherine J. Deans, MD, MHSc

(JAMA Surg. 2015)

## Study design: prospective patient choice cohort study



**Table 1. Baseline Demographic and Clinical Characteristics for All 102 Enrolled Patients By Treatment Decision**

Characteristic	Nonoperative Management (n = 37)	Surgery (n = 65)	P Value
Age, median (IQR), y	11 (10-14)	12 (9-13)	.87
Male, No. (%)	24 (64.9)	45 (69.2)	.65
White, No. (%)	25 (67.6)	53 (81.5)	.09
Hispanic, No. (%)	4 (10.8)	3 (4.6)	.16
Transferred from another institution, No. (%)	5 (13.5)	28 (43.1)	.002
Primary language other than English in home, No. (%)	8 (21.6)	2 (3.1)	.004
Insurance status, No. (%)			
Private	25 (67.6)	41 (63.1)	.93
Medicaid	11 (29.7)	21 (32.3)	
None	1 (2.7)	3 (4.6)	
Symptoms on initial presentation			
Duration of abdominal pain, median (IQR), h	13 (8-23)	16 (10-24)	.20
Fever, No. (%)	8 (21.6)	10 (15.4)	.43
Vomiting, No. (%)	18 (48.6)	21 (32.3)	.10
Diarrhea, No. (%)	4 (10.8)	6 (9.2)	.80
Ultrasonography, No. (%)	33 (89.2)	51 (78.5)	.28
CT scan, No. (%)	11 (29.7)	18 (27.7)	.83
WBC count, median (IQR), / $\mu$ L	12 900 (9000-14 800)	12 900 (10 000-14 600)	.78

**Mann-Whitney U test****Chisq-square test****Fisher exact test**

Table 2. Comparison of Outcomes Between Nonoperative Management and Surgery

Characteristic	Nonoperative Management	Surgery	P Value
Appendicitis-related medical care within 30 d, No. (%)	(n = 37)	(n = 58)	
Outpatient visit or urgent care <sup>a</sup>	4 (10.8)	12 (20.7)	.21
Emergency department	3 (8.1)	2 (3.4)	.37
Hospital admission	2 (5.4)	0	.15
Surgery or other invasive procedure	2 (5.4)	0	.15
1 y after enrollment	(n = 35)	(n = 50)	
Disability days, median (IQR) <sup>b</sup>	8 (5-18)	21 (15-25)	<.001
Complicated appendicitis, No. (%) <sup>c</sup>	1 (2.7)	8 (12.3)	.15
Health care costs, median (IQR), \$			
Initial hospital stay	3641 (2474-4227)	4960 (4588-5373)	<.001
All appendicitis-related care	4219 (2514-7795)	5029 (4596-5482)	.01
HRQOL questionnaire score, median (IQR)			
Child <sup>d</sup>	95.7 (89.1-98.9)	91.3 (87.0-98.9)	.31
Parent proxy <sup>e</sup>	91.9 (87.0-98.9)	93.0 (87.0-97.8)	.76

Abbreviations: HRQOL, health-related quality of life; IQR, interquartile range.

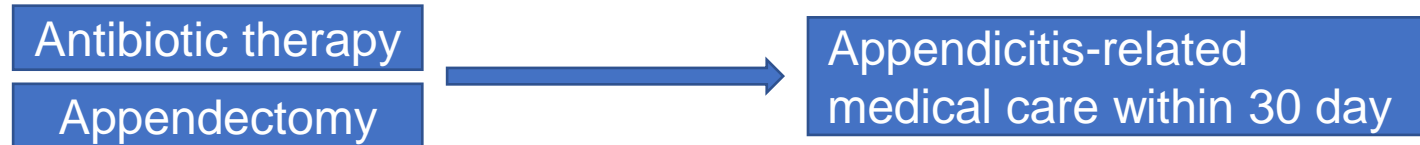
parent or guardian; and office, emergency department, and inpatient visits.

**Mann-Whitney U test****Fisher exact test**

# Study design and data collection

以PC Minneci *et al*, JAMA Surg, 2015為參考範例

- Aim:



- **Study design:** Cohort study
- **Study subject:** Patients with appendicitis
- **Study outcome:** appendicitis-related medical care within 30 day
- **Study variable:** age, sex, **treatment**, symptoms on initial presentation, ultrasonography, CT scan, WBC count



# Study design and data collection

- 在建立資料庫前，一定要先建好資料格式!

## 建立資料格式與譯碼簿

- 變數名稱、變數格式與編碼、建立譯碼簿

A1 : X ✓ fx ID													
	A	B	C	D	變項	說明	編碼	K	L	M	N	O	P
1	ID	sex	birthday_d	index_d	sex	性別	0=女, 1=男	CT	WBC	opd_30d	emg_30d	adm_30d	op_30d
2					birthday_d	出生日期							
3					index_d	闌尾炎發生日期							
					treatment	治療	0=抗生素治療, 1=闌尾切除術						
					duration_ap	腹部疼痛天數(單位:小時)							
					fever	發燒	0=無, 1=有						
					vomiting	嘔吐	0=無, 1=有						
					diarrhea	腹瀉	0=無, 1=有						
					ultra	超音波檢查	0=無, 1=有						
					CT	CT檢查	0=無, 1=有						
					WBC	WBC計量(單位:ul)							
					opd_30d	30天內有無門診	0=無, 1=有						
					emg_30d	30天內有無急診	0=無, 1=有						
					adm_30d	30天內有無住院	0=無, 1=有						
					op_30d	30天內有無手術或重大處置	0=無, 1=有						



## 輸入收案資料

- 依照格式將資料輸入，並確定資料正確性與格式

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
ID	sex	birthday_d	index_d	treatment	duration_ap	fever	vomiting	diarrhea	ultra	CT	WBC	opd_30d	emg_30d	adm_30d	op_30d
0082f8c2e	0	1976/3/17	2006/1/23	0	19	0	0	0	0	1	0	16217	0	0	0
00f7e6f0c	0	1993/9/20	2010/12/14	0	29	1	0	0	1	1	23219	0	0	0	0
01de9e3f	0	1992/4/19	2005/10/20	0	16	1	0	0	1	0	13420	0	0	0	0

# 建立資料格式與編碼之注意事項

- **如果資料庫已建好，要先確定資料格式與編碼沒問題!**

- 若須修改，切記要保留最原始的資料(未修改前)，以被查證用。

- **變數命名建議**

- 考慮會使用到的作業系統語言

- 英文或繁體中文為主(建議使用英文)

- 注意事項:

- 不可重複命名

- 不得有空白、特殊符號 (可包含 " @ # \$ \_ . " )、避免使用全形符號

- "數字、#、\$、\_、."不可用在變數名稱的第一字元

- "\_、."不可用在變數名稱最後一個字元

- 不可使用SPSS保留字，如 ALL, AND, BY, EQ(=), GE( $\geq$ ), GT(>), LE(<), LT( $\leq$ ), NE( $\neq$ ), NOT, OR, TO, WITH等等

# SPSS基本介紹

---

- 如何使用院內**SPSS**

- 向資訊室申請**anspss**主機之使用權限。(填寫OA電子表單)
- 遠端桌面連線至**anspss**
- 院內**SPSS**為第21版

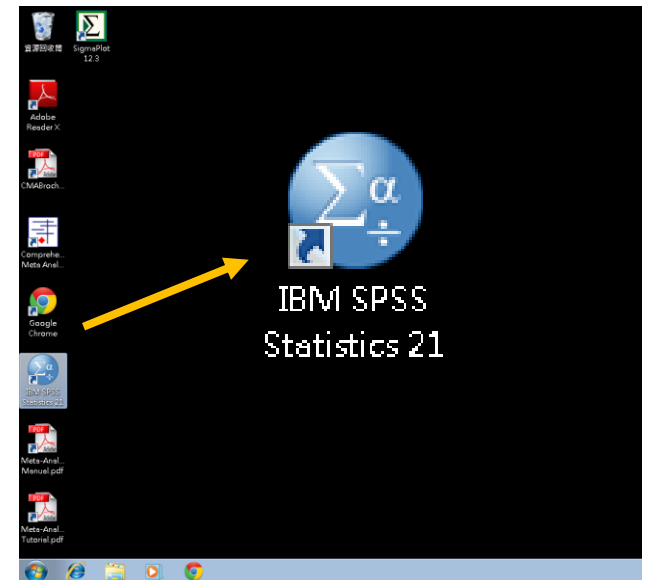
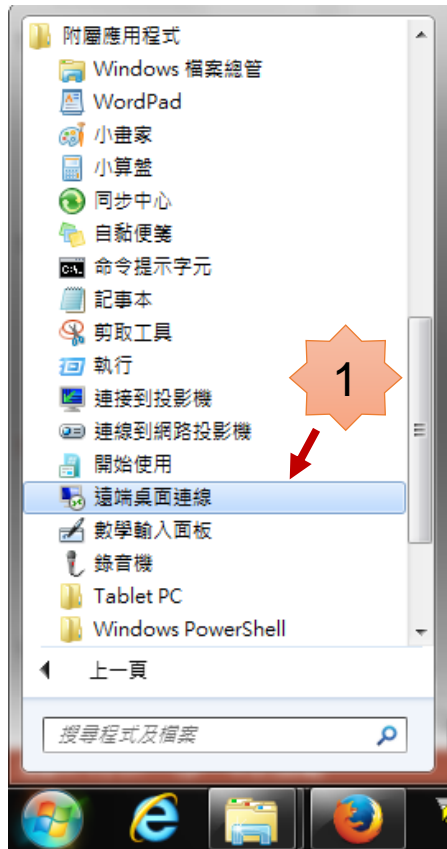
- **SPSS**檔案類型與基本視窗

- 資料視窗→資料檔 SPSS data files (\*.sav files)
- 語法視窗→語法檔 SPSS syntax command files (\*.sps files)
- 輸出視窗→輸出檔 SPSS output files (\*.spo files)
- 資料編輯視窗

# 如何使用院內SPSS - ANSPSS電腦主機

先填電子表單申請使用ANSPSS權限，通過後即可透過遠端連線操作SPSS

**STPE:**  
開始→附屬應用程式→遠端桌面連線→電腦 (anspss)→連線→輸入員編與密碼→確定



# SPSS檔案類型與基本視窗

## 資料視窗→資料檔 data files (\*.sav files)

**資料檢視** 可檢視與編輯觀測值。

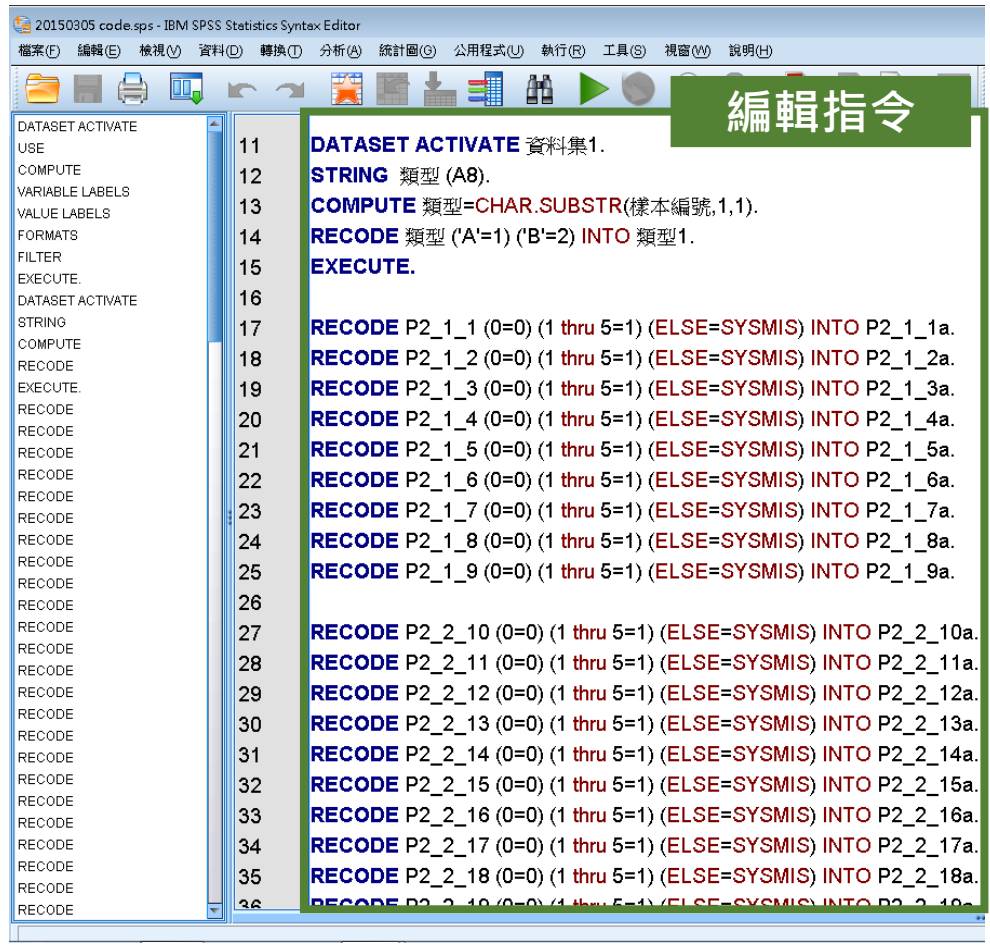
	ID	sex	birthday_d	index_d
1	0082f8c2e4c7a5dd5bc...	0	17-Mar-1976	23-Jan-2006
2	00f7e8f0dfe8c36311e3...	0	20-Sep-1993	14-Dec-2010
3	01de9e3fa520c1d28a6...	0	19-Apr-1992	20-Oct-2005
4	45d4f6f0...	0	16-Dec-1977	10-Nov-2002
5	fa80725...	0	28-Nov-1985	07-Jun-2007
6	021b97c1d3c79fa9cba...	0	13-Dec-1966	13-Jul-2005
7	0363eb02a2dff4e915f2...	0	03-Aug-1932	06-Nov-2002
8	04f20efd13f0d5a6b2d2...	0	08-Apr-1959	19-Apr-2003
9	06d004c46cf823134db...	1	03-Jun-1978	03-Sep-2005
10	079c327f0621fec1c9be...	1	05-Dec-1958	25-Jul-2003
11	088bc662d0bc15f5b3e...	1	01-Mar-1966	19-Apr-2003
12	095b2af5da28225dd64...	0	18-Aug-1981	22-Jan-2001
13	0bf72e8e703874f90990...	0	25-Oct-1983	22-Feb-2007
14	0d1a87e3e709f3acd77...	1	03-Oct-1989	24-Oct-2007

**變數檢視** 可檢視與編輯變項的特徵。

	名稱	類型	寬度	小數	標記
1	ID	字串	32	0	
2	sex	數字的	11	0	
3	birthday_d	日期	11	0	
4	index_d	日期	11	0	
5	treatment	數字的	11	0	treatment
6	duration_ap	數字的	11	0	
7	fever	數字的	11	0	
8	vomiting	數字的	11	0	
9	diarrhea	數字的	11	0	
10	ultra	數字的	11	0	
11	CT	數字的	11	0	
12	WBC	數字的	11	0	
13	opd_30d	數字的	11	0	outpatient visit ...
14	emrg_30d	數字的	11	0	emergency dep...
15	adm_30d	數字的	11	0	hospital admiss..

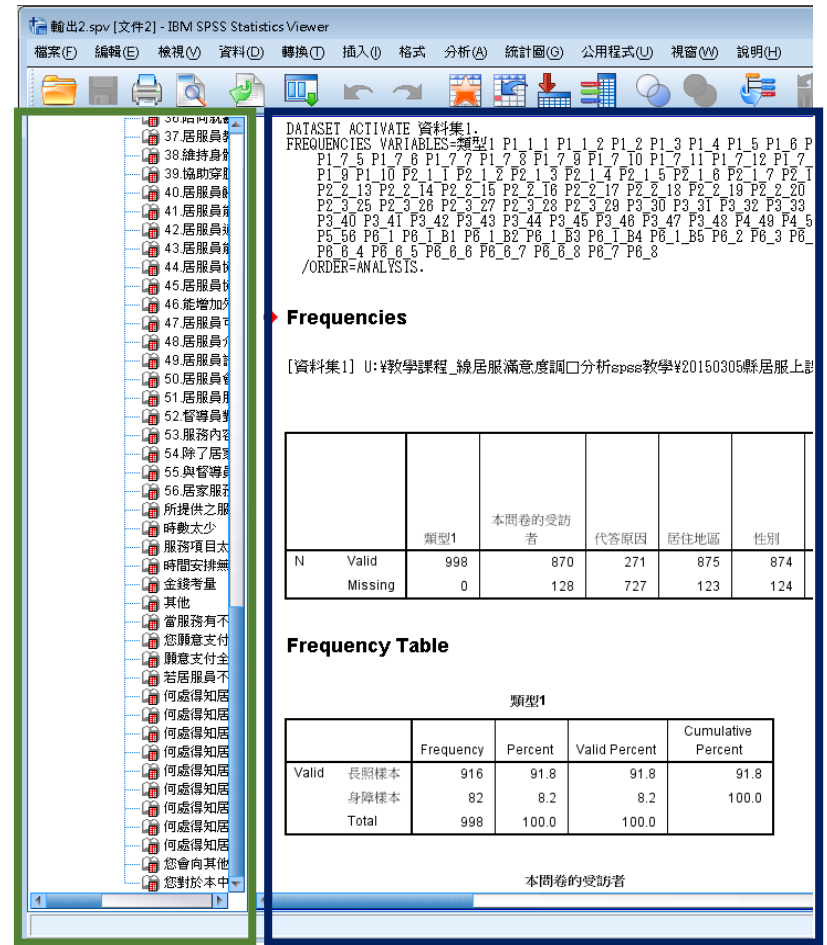
# SPSS檔案類型與基本視窗

語法視窗→  
語法檔 syntax command files (\*.sps files)



可由對話視窗匯出或自行編輯。執行例行性作業時可節省時間，且記錄資料分析的歷程

輸出視窗→  
輸出檔 output files (\*.spo files)



樹狀輸出列表

輸出內容區  
→ 可以直接複製表格

# SPSS實際操作

---

- 資料讀取與儲存
- 資料處理
  - 變數格式定義
  - 檢查數值合理性與遺漏值處理
  - 資料轉換
    - 計算變數
    - 重新編碼成不同變數
  - 選擇觀察值
- 統計分析

- ✓ Sample size: 1000 appendicitis patients
- ✓ Outcome: opd\_30d, emg\_30d, op\_30d
- ✓ Variables: sex, treatment, duration\_ap, fever, vomiting, diarrhea, ultra, CT, WBC

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	ID	sex	birthday_d	index_d	treatment	duration_ap	fever	vomiting	diarrhea	ultra	CT	WBC	opd_30d	emg_30d	adm_30d	op_30d
2	0082f8c2e	0	1976/3/17	2006/1/23	0	19	0	0	0	1	0	16217	0	0	0	0
3	00f7e6f0c	0	1993/9/20	2010/12/14	0	29	1	0	0	1	1	23219	0	0	0	0
4	01de9e3f	0	1992/4/19	2005/10/20	0	16	1	0	0	1	0	13420	0	0	0	0
5	01f8ba65f	0	1977/12/16	2002/11/10	0	8	1	0	0	1	1	21195	0	0	0	0
6	0203364b	0	1985/11/28	2007/6/7	0	22	1	0	0	1	1	13016	0	0	0	0
7	021b97c1	0	1966/12/13	2005/7/13	0	15	1	0	0	1	1	12344	0	0	0	1
8	0363eb02	0	1932/8/3	2002/11/6	0	23	1	1	1	1	1	11791	1	0	0	0
9	04f20efd1	0	1959/4/8	2003/4/19	0	17	1	0	0	0	1	13228	1	0	0	0
10	06d004c4	1	1978/6/3	2005/9/3	0	8	1	0	0	0	1	23808	1	0	1	0
11	079c327f0	1	1958/12/5	2003/7/25	0	28	1	0	0	0	1	11223	1	0	0	0
12	086bc662	1	1966/3/1	2003/4/19	0	5	1	0	0	0	1	17930	1	0	0	0
13	095b2af5	0	1981/8/18	2001/1/22	0	4	1	0	0	0	1	15935	0	0	0	0
14	0bf72e8e	0	1983/10/25	2007/2/22	0	27	1	1	0	0	1	13389	0	0	0	0
15	0d1a87e3	1	1969/10/3	2007/10/24	0	19	1	0	0	0	1	11538	0	0	0	0
16	0e35b003	1	1981/10/21	2002/3/22	0	6	1	0	0	0	1	14352	0	0	0	0
17	0e51f590	1	1926/5/5	2008/7/12	0	5	1	0	0	0	1	22318	1	1	0	0
18	0ea72131	0	1991/10/20	2010/9/28	0	4	1	0	0	1	1	16280	1	1	0	0
19	0ef0b262	0	1977/2/23	2003/1/11	0	27	1	0	0	1	1	23008	1	1	0	0
20	123beb0e	1	1980/1/31	2001/5/15	0	21	1	1	0	1	1	19959	1	1	0	0
21	125a21dd	0	1977/7/18	2005/4/19	0	6	1	1	0	1	1	17390	1	1	0	0
22	12b6d81f	0	1966/5/26	2009/4/15	0	12	1	1	0	1	1	16270	0	0	0	0
23	1398212b	1	1967/12/25	2006/5/4	0	27	1	1	0	1	1	24797	0	0	0	0
24	13e26513	1	1991/2/19	2001/2/21	0	16	1	1	0	0	1	12743	0	0	0	0
25	13fb622f8	1	1966/4/22	2002/1/22	0	12	1	1	0	0	1	15836	0	0	0	0
26	155e79d1	0	1985/2/18	2000/8/7	0	12	1	1	0	0	1	21064	0	0	0	0
27	15b9a2f5	1	1987/4/30	2010/4/6	0	14	1	1	0	0	0	14898	0	0	0	0
28	17b3090f	0	1980/3/15	2008/1/11	0	13	1	1	0	0	0	15158	0	0	0	0
29	185fe7f0c	0	1988/10/16	2003/10/3	0	19	1	0	0	0	0	19947	0	0	0	0



- ✓ Sample size: 1000 appendicitis patients
- ✓ Outcome: opd\_30d, emg\_30d, adm\_30d, op\_30d
- ✓ Variables: sex, birthday\_d, index\_d, treatment, duration\_ap, fever, vomiting, diarrhea, ultra, CT, WBC

變項	說明	編碼
ID	病歷號	
sex	性別	0=女, 1=男
birthday_d	出生日期	
index_d	闌尾炎發生日期	
treatment	治療	0=抗生素治療, 1=闌尾切除術
duration_ap	腹部疼痛天數(單位:小時)	
fever	發燒	0=無, 1=有
vomiting	嘔吐	0=無, 1=有
diarrhea	腹瀉	0=無, 1=有
ultra	超音波檢查	0=無, 1=有
CT	CT檢查	0=無, 1=有
WBC	WBC計量(單位:ul)	
opd_30d	30天內有無門診	0=無, 1=有
emg_30d	30天內有無急診	0=無, 1=有
adm_30d	30天內有無住院	0=無, 1=有
op_30d	30天內有無手術或重大處置	0=無, 1=有

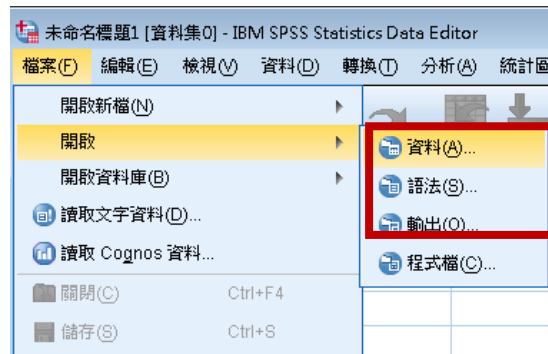
## 1. 打開SPSS



## 2. 將跳出的視窗按取消

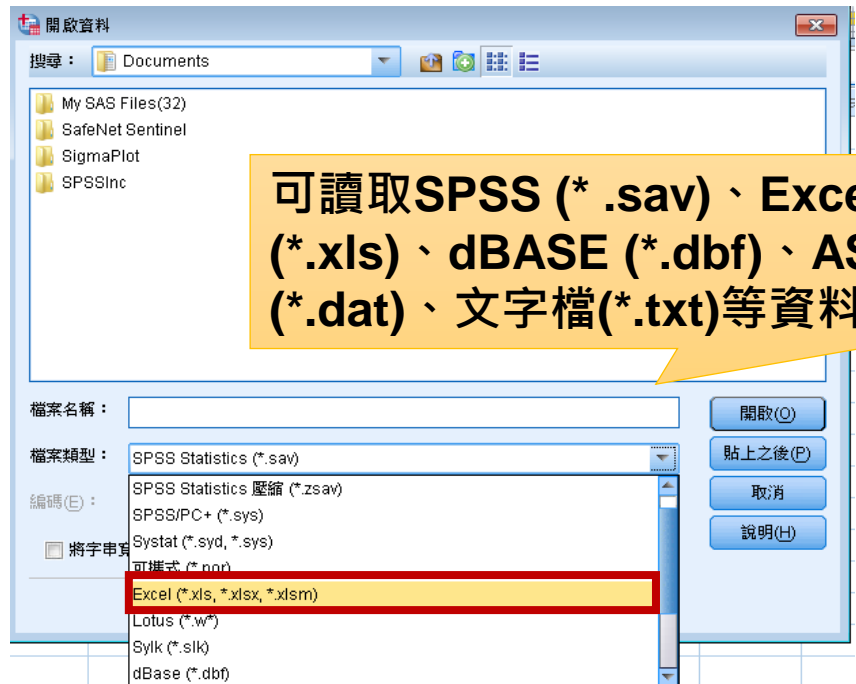


## 3. 檔案→開啟→資料

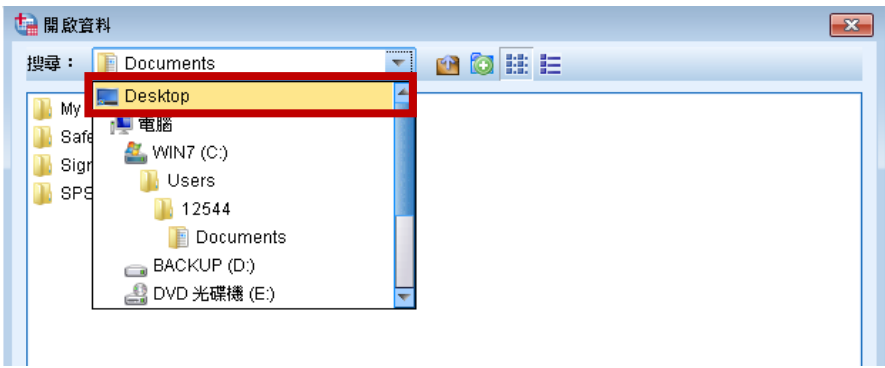


除了開啟資料檔外，也可開啟語法檔及輸出檔

## 4. 選取檔案類型 (選取Excel檔)



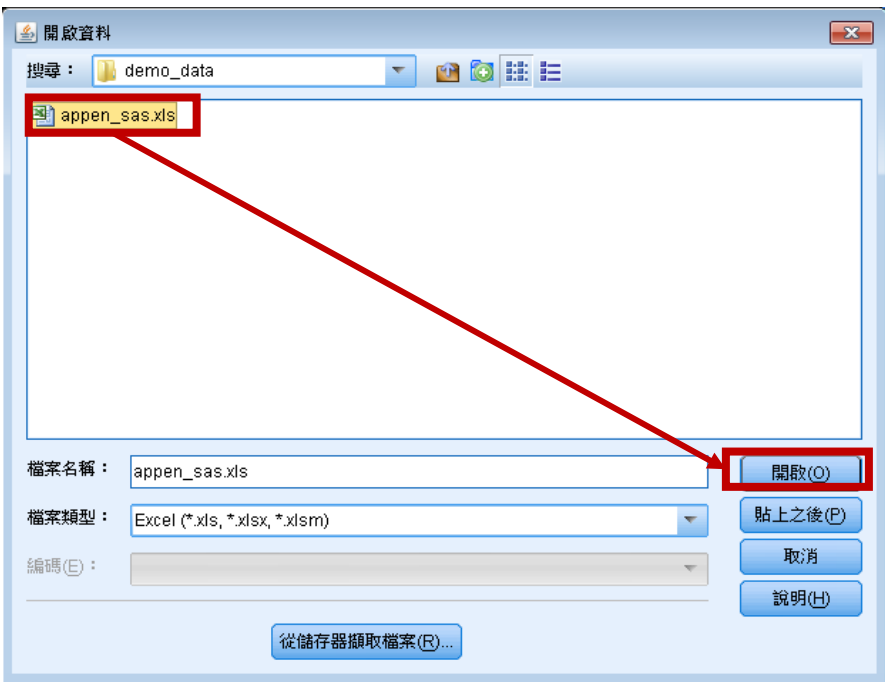
## 5. 選取檔案路徑



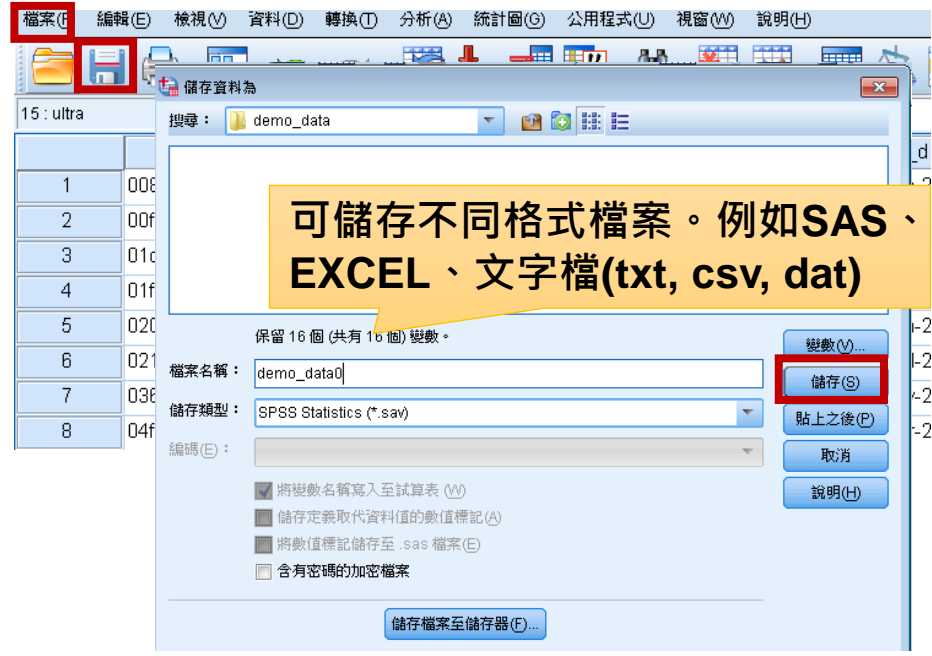
## 7. 工作單請選取你的資料表



## 6. 選取要開啟的資料檔後，按開啟



## 8. 成功開啟檔案後，請存檔： 按 或 檔案 → 另存新檔 → 儲存



- 變數格式定義
- 檢查數值合理性與遺漏值處理
  - 了解原始資料分布狀態
  - 處理不合理數值、意義不明的編碼
    - ✓ 遺漏值設定或回去查原始資料(病歷資料)修改
- 新增變數 (以年齡變項為例)
  - 計算變數、重新編碼成不同變數

## 從變數檢視去設定

demo0.sav [資料集2] - IBM SPSS Statistics Data Editor

檔案(F) 編輯(E) 檢視(V) 資料(D) 轉換(T) 分析(A) 統計圖(G) 公用程式(U) 視窗(W) 說明(H)

\*demo1.sav [資料集1] - IBM SPSS Statistics Data Editor

檔案(F) 編輯(E) 檢視(V) 資料(D) 轉換(T) 分析(A) 統計圖(G) 公用程式(U)

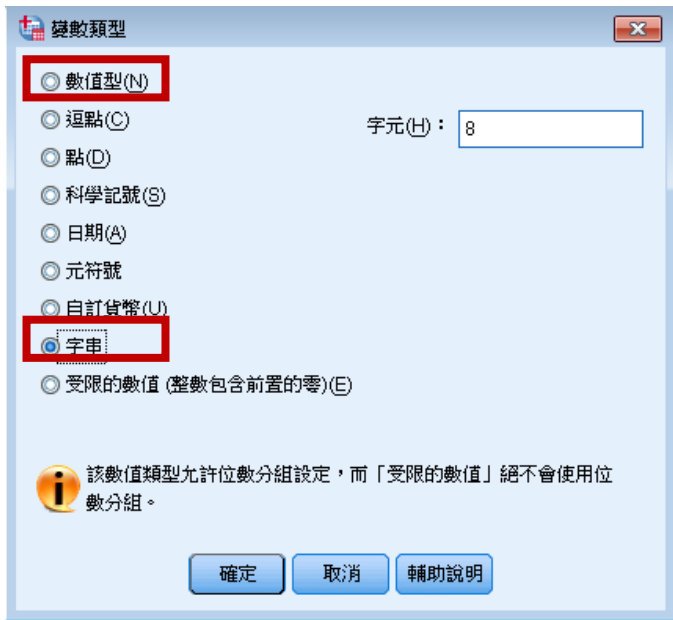
**替變項作註解**

名稱	名稱	類型	寬度	小數	標記	值	遺漏	欄	對齊	測量	角色
1	ID	字串	32	0		無	無	18	≡ 靠左	名義(N)	↘ 輸入
2	sex	數字的	11	0		{0, female}...	無	4	≡ 靠右	名義(N)	↘ 輸入
3	birthday_d	日期	11	0		無	無	11	≡ 靠右	尺度(S)	↘ 輸入
4	index_d	日期	11	0	appendicitis date	無	無	11	≡ 靠右	尺度(S)	↘ 輸入
5	treatment	數字的	11	0	treatment	{0, antibiotic t...	無	8	≡ 靠右	名義(N)	↘ 輸入
6	duration_ap	數字的	11	0	duration of abd...	無	無	11	≡ 靠右	尺度(S)	↘ 輸入
7	fever	數字的	11	0		{0, no}...	無	5	≡ 靠右	名義(N)	↘ 輸入
8	vomiting	數字的	11	0		{0, no}...	無	7	≡ 靠右	名義(N)	↘ 輸入
9	diarrhea	數字的	11	0		{0, no}...	無	7	≡ 靠右	名義(N)	↘ 輸入
10	ultra	數字的	11	0		{0, no}...	無	5	≡ 靠右	名義(N)	↘ 輸入
11	CT	數字的	11	0		{0, no}...	無	4	≡ 靠右	名義(N)	↘ 輸入
12	WBC	數字的	11	0		無	無	7	≡ 靠右	尺度(S)	↘ 輸入
13	opd_30d	數字的	11	0	outpatient visit ...	{0, no}...	無	8	≡ 靠右	名義(N)	↘ 輸入
14	emg_30d	數字的	11	0	emergency dep...	{0, no}...	無	8	≡ 靠右	名義(N)	↘ 輸入
15	adm_30d	數字的	11	0	hospital admiss...	{0, no}...	無	8	≡ 靠右	名義(N)	↘ 輸入
16	op_30d	數字的	11	0	surgery or other...	{0, no}...	無	8	≡ 靠右	名義(N)	↘ 輸入

資料檢視 變數檢視

資料檢視 變數檢視

## 變數類型 (Type)



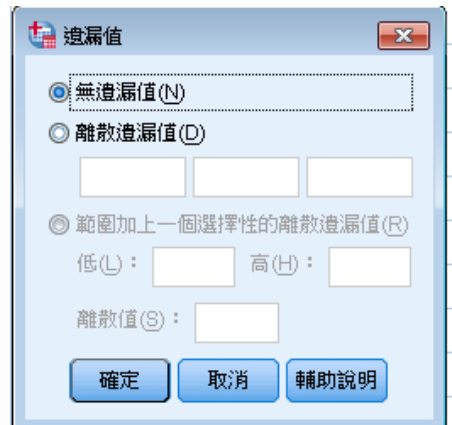
數值型(Numeric) & 字串 (String) 較常使用。由於部份統計無法使用字串的變項，故建議將類別變項(如性別)以數值型方式輸入。

## 值 (Values)



為輸入的數值做定義

## 遺漏值 (Missing)



指定特定不納入分析的值

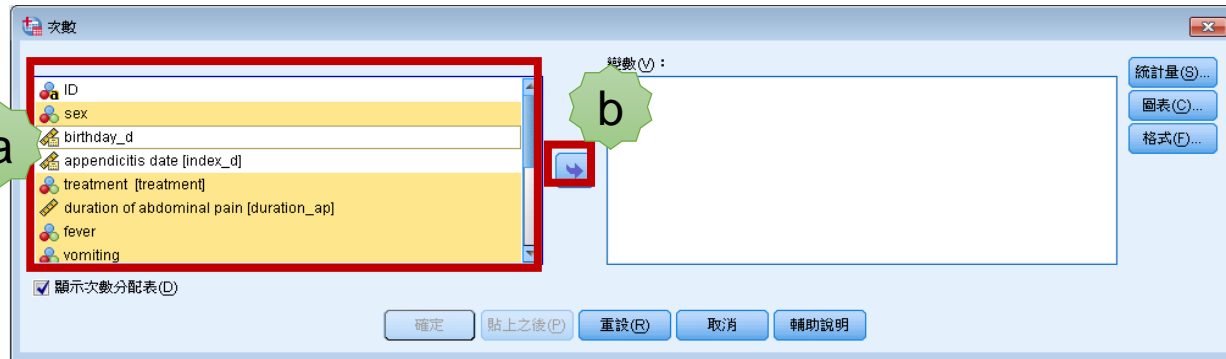
# SPSS資料處理:檢查數值合理性與遺漏值處理

檢查每個變數的型態與編碼 → 利用次數分配表來了解資料分布狀態

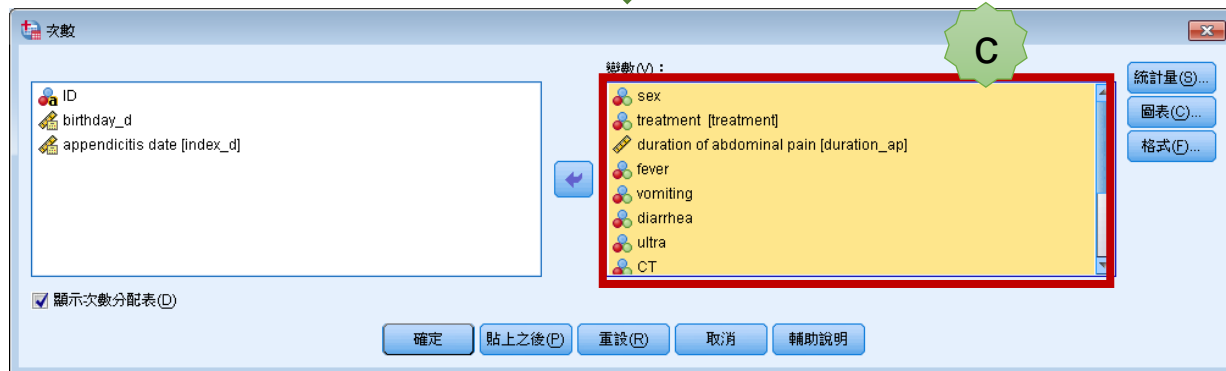
## 1. 分析 → 敘述統計 → 次數分配表



## 2. 將左邊框框內的預檢視變數選取 → 放入變數



變數會跑進變數框框中



# SPSS資料處理：檢查數值合理性與遺漏值處理

檢查每個變數的型態與編碼→利用次數分配表來了解資料分布狀態

## 3.設定統計量、圖表、格式

次數分配表：統計量

**百分位數值**

四分位數(Q)

切割觀察組為(U)： 10 相同值組別

百分位數(P)：

新增(A) 變更(C) 移除(R)

**集中趨勢**

平均數(M)

中位數(D)

眾數(O)

總和(S)

觀察值為組別中點(L)

**分散情形**

標準差(T)  最小值(O)

變異數(V)  最大值(X)

範圍(A)  平均數的標準誤(E)

**分配**

偏態(W)

峰度(K)

繼續 取消 輔助說明

次數分配表：圖表

**圖表類型**

無(O)

長條圖(B)

圓餅圖(P)

直方圖(H)：

在直方圖顯示常態分佈曲線(S)

**圖表值**

次數分配表(F)  百分比(C)

繼續 取消 輔助說明

次數分配表：格式

**順序依據**

依觀察值遞增排序(A)

依觀察值遞減排序(D)

依個數遞增排序(E)

依個數遞減排序(N)

**多重變數**

比較變數(C)

依變數組織輸出(O)

不列出具有許多類別的表格

最大類別數(M)： 10

繼續 取消 輔助說明



# SPSS資料處理：檢查數值合理性與遺漏值處理

檢查每個變數的型態與編碼→利用次數分配表來了解資料分布狀態

## 4.設定完成後按 確定/貼上之後(執行語法)



會跳出語法視窗。接著選取這段程式→按綠色箭頭

請見範例輸出1



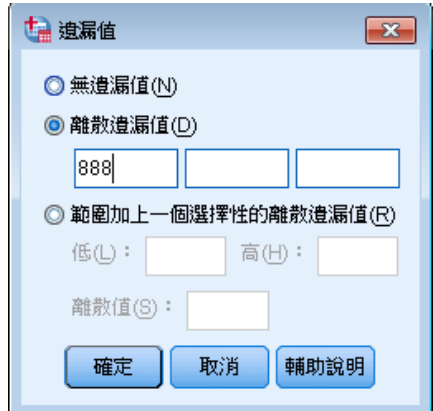
Statistics

		sex	treatment	duration of abdominal pain	fever	vomiting	diarrhea	ultra	CT	WBC	outpatient visit or urgent care within 30 days	emergency department within 30 days	hospital admission within 30 days	surgery or other invasive procedure with 30 days
N	Valid	997	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	Missing	3	0	0	0	0	0	0	0	0	0	0	0	0
Mean		.49	.61	16.30	.18	1.39	.10	.92	.30	17400.66	.17	1.04	.03	.02
Median		.00	1.00	15.00	.00	.00	.00	1.00	.00	17465.50	.00	.00	.00	.00
Std. Deviation		.500	.489	28.911	.380	31.582	.304	.273	.460	4478.955	.371	31.591	.165	.143
Minimum		0	0	1	0	0	0	0	0	0	0	0	0	0
Maximum		1	1	888	1	999	1	1	1	24976	1	999	2	1
Percentiles	25	.00	.00	8.00	.00	.00	.00	1.00	.00	13407.25	.00	.00	.00	.00
	50	.00	1.00	15.00	.00	.00	.00	1.00	.00	17465.50	.00	.00	.00	.00
	75	1.00	1.00	23.00	.00	1.00	.00	1.00	1.00	21323.25	.00	.00	.00	.00



處理不合理數值與編碼  
→設定為遺漏值或查原始資料修改

若要設定遺漏值→請回變數檢視



# SPSS資料轉換:新增變數-計算變數

## 新增age變數

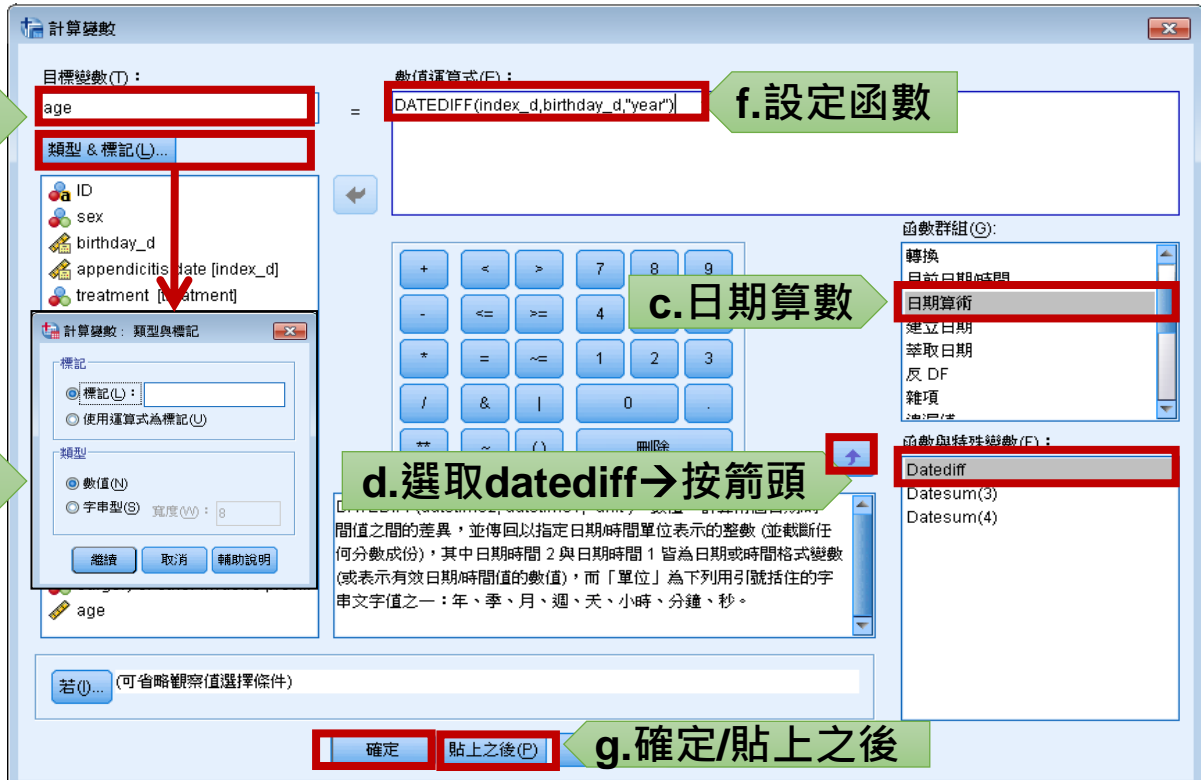
### 1.轉換→計算變數



2. 目標變數(新變項名稱)→類型&標記(類型選數值)→函數名稱(日期算數)→函數與特殊變數(datediff)→朝上箭頭→數值運算式(DATEDIFF(index\_d,birthday\_d,"year"))→確定/貼上之後

a.取新變項名稱

b.類型選擇字串



兩個日期相減的其他計算方式:

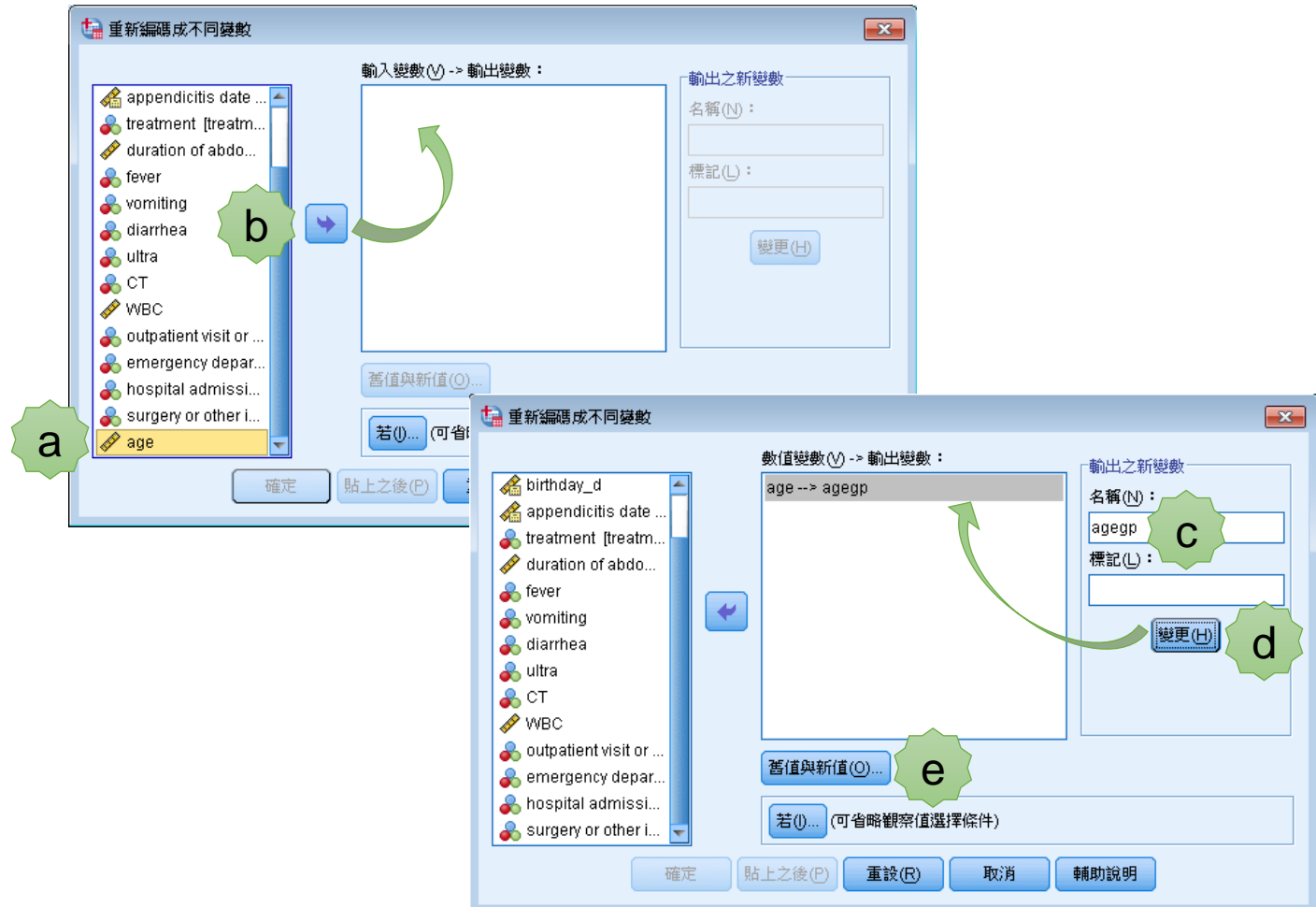
- 1.計算變數→數值運算式:  $(index\_d - birthday\_d) / (60 * 60 * 24 * 365.25)$
- 2.轉換→日期和時間精靈

# SPSS資料轉換:重新編碼成不同變數

新增年齡組別變數:將年齡分成 0-6, 7-17, 18-60, 60+

1.轉換→重新編碼成不同變數

2.選取變項→按向右箭頭丟到輸入變數中→輸出之新變數(命名新變數名稱)→變更→舊值與新值



# SPSS資料轉換:重新編碼成不同變數

新增年齡組別變數:將年齡分成 0-6, 7-17, 18-60, 60+

3.舊值(範圍:0-6)→新值(數值=1)→新增→【其他範圍設定】→繼續→確定/貼上之後



4.去變數檢視設定該變數特徵。(小數、標記、值...)

	名稱	類型	寬度	小數	標記	值	遺漏	欄	對齊
17	age	數字的	8	0		無	無	8	≡ 靠右
18	agegp	數字的	8	0		{1, 0-6}...	無	10	≡ 靠右

	ultra	CT	WBC	opd_30d	emg_30d	adm_30d	op_30d	age	agegp
1	1	0	16217	0	0	0	0	29	3
2	1	1	23219	0	0	0	0	17	2
3	1	0	13420	0	0	0	0	13	2

完成

# 貼上之後~

為什麼會強調一定要按'貼上之後'呢?

1. 減少重複作業時間
2. 紀錄處理過程

- ✓ 舉例來說，你有47個變數需要重新編碼，且編碼方式相同。
  - ✓ 你可以想像處理完第1個變數後，還要處理剩下的47個變數嗎?
  - ✓ 萬一每年都要跑一樣的資料，還要重複這些步驟...

因此，我們需要'貼上之後'，幫我們輸出這段程式的語法，之後只要修改語法的變項名稱即可。往後的資料只要執行這一大串語法，就可以節省你處理資料的時間!

```
17 RECODE P2_1_1 (0=0) (1 thru 5=1) (ELSE=SYSMIS) INTO P2_1_1a.
18
19
20
21 RECODE P2_1_2 (0=0) (1 thru 5=1) (ELSE=SYSMIS) INTO P2_1_2a.
22 RECODE P2_1_3 (0=0) (1 thru 5=1) (ELSE=SYSMIS) INTO P2_1_3a.
23 RECODE P2_1_4 (0=0) (1 thru 5=1) (ELSE=SYSMIS) INTO P2_1_4a.
24 RECODE P2_1_5 (0=0) (1 thru 5=1) (ELSE=SYSMIS) INTO P2_1_5a.
25 RECODE P2_1_6 (0=0) (1 thru 5=1) (ELSE=SYSMIS) INTO P2_1_6a.
26 RECODE P2_1_7 (0=0) (1 thru 5=1) (ELSE=SYSMIS) INTO P2_1_7a.
27 RECODE P2_1_8 (0=0) (1 thru 5=1) (ELSE=SYSMIS) INTO P2_1_8a.
28 RECODE P2_1_9 (0=0) (1 thru 5=1) (ELSE=SYSMIS) INTO P2_1_9a.
29
30 RECODE P2_2_10 (0=0) (1 thru 5=1) (ELSE=SYSMIS) INTO P2_2_10a.
31 RECODE P2_2_11 (0=0) (1 thru 5=1) (ELSE=SYSMIS) INTO P2_2_11a.
32 RECODE P2_2_12 (0=0) (1 thru 5=1) (ELSE=SYSMIS) INTO P2_2_12a.
33 RECODE P2_2_13 (0=0) (1 thru 5=1) (ELSE=SYSMIS) INTO P2_2_13a.
34 RECODE P2_2_14 (0=0) (1 thru 5=1) (ELSE=SYSMIS) INTO P2_2_14a.
```

只要修改變項名稱就可以了!  
雖然第一次改程式會比較久，  
但之後可以一勞永逸!

# SPSS資料處理：選擇觀察值

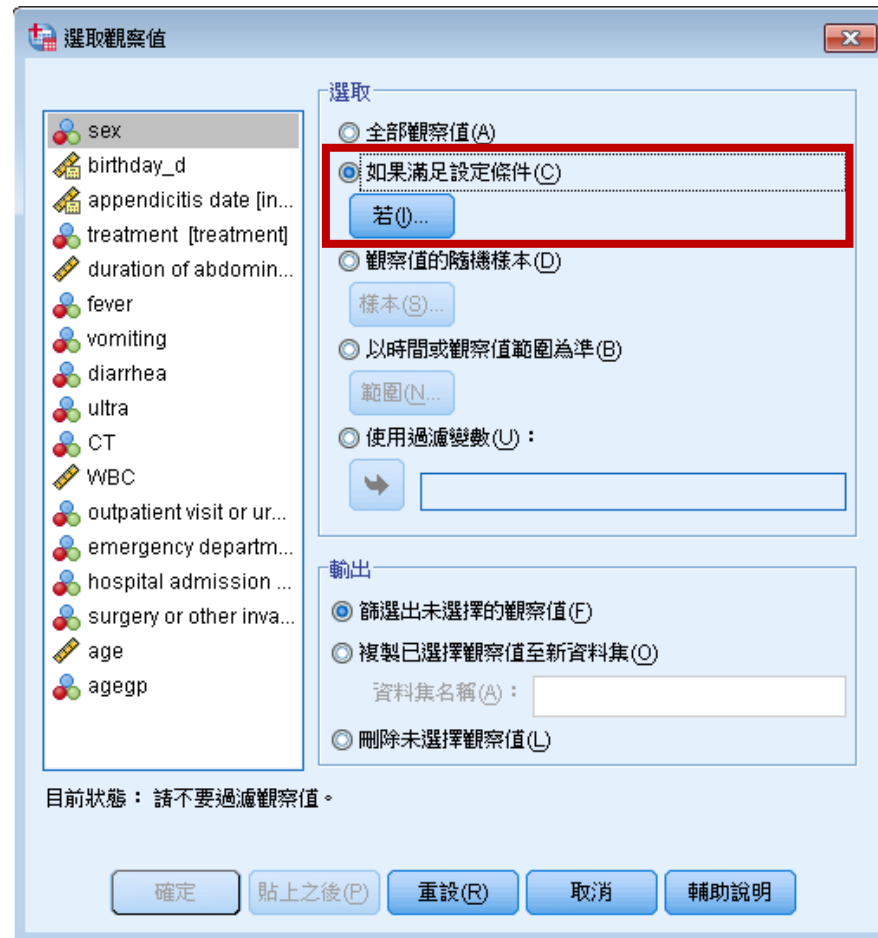
篩選出我們要的觀察值

- 以PC Minneci *et al*, JAMA Surg, 2015為例，只挑選7-17歲病人 (age group=2)

## 1. 資料 → 選擇觀察值



## 2. 選'如果滿足設定條件' → 若



# SPSS資料處理：選擇觀察值

## 3. 函數群組(全部) → 函數與特殊變數(Any) → 按朝上箭頭

## 4. 設定Any函數 → 繼續



函數: Any(?,?) → 滿足後會回傳'1'這個數值

變數      變數的值



# SPSS資料處理：選擇觀察值

## 5. 確定/貼上之後

目前狀態：該不要過濾觀察值。

篩選後的資料，可另存新檔

完成篩選後，不要的觀察值會用斜線表示  
(filter\_\$值為missing)

	CT	WBC	opd_30d	emg_30d	adm_30d	op_30d	age	agegp	filter_\$
<del>1</del>	1	0	16217	0	0	0	29	3	0
<del>2</del>	1	1	23219	0	0	0	17	2	1
<del>3</del>	1	0	13420	0	0	0	13	2	1
<del>4</del>	1	1	21195	0	0	0	24	3	0
<del>5</del>	1	1	13016	0	0	0	21	3	0
<del>6</del>	1	1	12344	0	0	1	38	3	0
<del>7</del>	1	1	11791	1	0	0	70	4	0
<del>8</del>	0	1	13228	1	0	2	44	3	0

## 6. 要取消篩選，就按‘全部觀察值’

目前狀態：藉由下列的值過濾觀察值： filter\_\$

Sample size: 1000 → 199

# 資料型態與適用統計方法

依變數		自變數				
		連續變數	類別變數			
			兩組		三組或以上	
			獨立樣本	相依樣本	獨立樣本	相依樣本
連續變數	常態分布	Pearson's correlation/ Linear regression	t-test	Paired t-test	ANOVA	Repeated measures ANOVA
	非常態分布	Spearman's correlation	Wilcoxon rank-sum test	Wilcoxon signed-rank test	Kruskal-Wallis test	Friedman test
類別變數	兩組	Logistic regression	Chi-square test	McNemar's test	Chi-square test	Cochren's Q test
	三組	Multinomial/ Ordinal logistic regression	Chi-square test	Cochren's Q test	Chi-square test	Cochren's Q test

# SPSS統計分析:獨立樣本 T 檢定

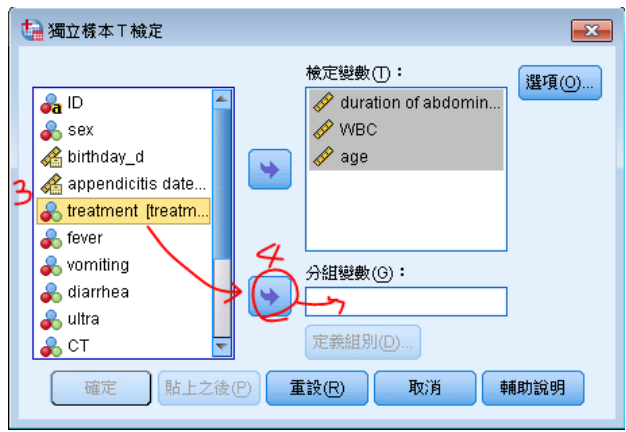
目的:了解age, duration of abdominal pain, WBC count在不同治療的分布狀況。

假定: age, duration of abdominal pain, WBC count為常態分佈。

## 執行步驟

1.分析→敘述統計  
→獨立樣本T檢定

2.將依變數放入檢定變數中→將欲分組變數放入分組變數中→定義組別→分別輸入分組變數的編碼→繼續→選項(無須改變設定)→繼續→貼上之後/確定



# SPSS統計分析:獨立樣本 T 檢定

目的:了解age, duration of abdominal pain, WBC count在不同治療的分布狀況。

假定: age, duration of abdominal pain, WBC count為常態分佈。

## 輸出內容

### 描述性統計量

Group Statistics

	treatment	N	Mean	Std. Deviation	Std. Error Mean
duration of abdominal pain	antibiotic therapy	392	15.50	8.382	.422
	appendectomy	607	15.37	8.798	.339
WBC	antibiotic therapy	391	17510.52	4415.257	223.200
	appendectomy	607	17387.23	4417.884	179.300
age	antibiotic therapy	393	32.31	18.531	.940
	appendectomy	607	33.94	18.958	.737

先看Equality of variance那欄，

1. 若P值(Sig)<0.05→ t-test結果參考 Equal variances not assumed
2. 若P值(Sig)≥0.05→ t-test結果參考 Equal variances assumed

### 獨立樣本 T 檢定

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
duration of abdominal pain	Equal variances assumed	2.950	.086	.233	997	.816	.130	.560	-.968	1.228
	Equal variances not assumed			.235	863.357	.814	.130	.554	-.957	1.217
WBC	Equal variances assumed	.026	.873	.430	996	.667	123.290	286.415	-438.756	685.336
	Equal variances not assumed			.431	832.428	.667	123.290	286.378	-438.818	685.398
age	Equal variances assumed	.541	.462	-1.336	998	.182	-1.625	1.217	-4.013	.762
	Equal variances not assumed			-1.342	850.613	.180	-1.625	1.211	-4.002	.751

a

b

# SPSS統計分析:Mann-Whitney U test

目的:了解age, duration of abdominal pain, WBC count在不同治療的分布狀況。

假定: age, duration of abdominal pain, WBC count為非常態分佈。

## 執行步驟

### 1.分析→無母數檢定→獨立樣本

分析(A) 統計圖(G) 公用程式(U) 視窗(W) 說明(H)

- 報表(P)
- 敘述統計(E)
- 比較平均數法(M)
- 一般線性模式(G)
- 概化線性模式(Z)
- 混合模式(X)
- 相關(C)
- 迴歸(R)
- 對數線性(O)
- 分類(Y)
- 維度縮減(D)
- 尺度(A)
- 無母數檢定(N)**
  - ▲ 單一樣本(O)...
  - ▲ 獨立樣本(I)...**
  - ▲ 相關樣本(R)...
  - 歷史對話記錄(L) ▶
- 預測(T)
- 存活分析(S)
- 複選題(U)
- 模擬...
- 品質控制(Q)
- ROC 曲線(V)...
- IBM SPSS Amos...

Variable	Value
dex_d	treatment
-Jan-2006	
Dec-2010	
-Oct-2005	
-Nov-2002	
-Jun-2007	
3-Jul-2005	
-Nov-2002	
-Apr-2003	
-Feb-2007	
-Oct-2007	
-Mar-2002	
2-Jul-2008	

### 2.設定目標→自動比較群組間的分配

無母數檢定：兩個以上的獨立樣本

目標 欄位 設定

使用無母數檢定識別兩個以上群組之間的差異。無母數檢定不會假設您的資料以常態分配。

您的目標是什麼？

各個目標會對應到「設定」索引標籤中的不同預設組態，希望的話可進一步自訂。

- 自動比較群組間的分配(U)
- 比較群組間的中位數(S)
- 自訂分析(C)

說明

自動使用 Mann-Whitney U 檢定比較 2 個樣本時群組間的分配，或使用 Kruskal-Wallis 單因子變異數分析比較 k 個樣本時群組間的分配。選擇的檢定會根據您的資料有所不同。

▶ 執行(R) 貼上之後(P) 重設(R) 取消 輔助說明

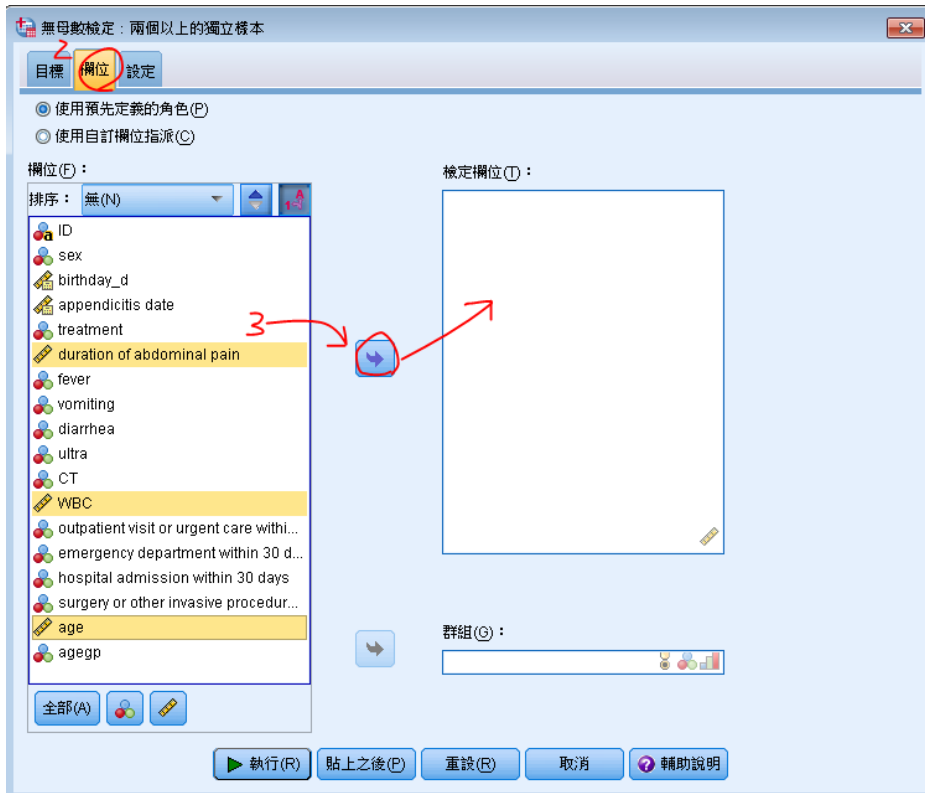
# SPSS統計分析:Mann-Whitney U test

目的:了解age, duration of abdominal pain, WBC count在不同治療的分布狀況。

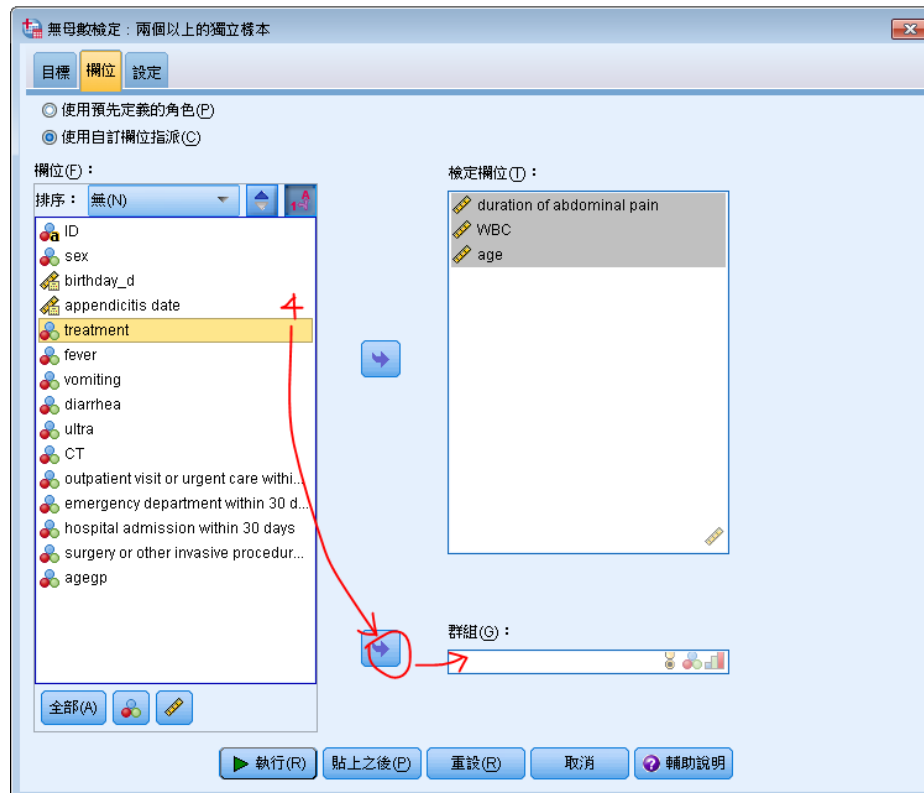
假定: age, duration of abdominal pain, WBC count為非常態分佈。

## 執行步驟

### 3.設定欄位→將依變數放入檢定欄位



### 4.將欲分組變數放入群組



# SPSS統計分析:Mann-Whitney U test

目的:了解age, duration of abdominal pain, WBC count在不同治療的分布狀況。

假定: age, duration of abdominal pain, WBC count為非常態分佈。

## 執行步驟

5.設定→勾選自訂檢定→勾選Mann-Whitney U →檢定選項與使用者遺漏值無需設定

無母數檢定：兩個以上的獨立樣本

目標 欄位 設定 5

選取項目(Z)：

選擇檢定(S)

檢定選項

使用者遺漏值

自動根據資料選擇檢定(U)

自訂檢定(C) 6

(Wilcoxon rank-sum test)

比較群組間的分配

Mann-Whitney U (2 個樣本)(M) 7

Kruskal-Wallis 單因子變異數分析(k 個樣本)(K)

Kolmogorov-Smirnov (2 個樣本)(O)

隨機性檢定順序 (Wald-Wolfowitz 適用於 2 個樣本)(Q)

多重比較(M)：所有成對

檢定有序對立 (Jonckheere-Terpstra 適用於 k 個樣本)(J)

假設排序(H)：最小至最大

多重比較(O)：所有成對

比較群組間的範圍

Moses 極端反應 (2 個樣本)(X)

計算樣本中的偏離值(Y)

自訂偏離值數目(U)

偏離值(O)：1

比較群組間的中位數

中位數檢定(k 個樣本)(K)

合併樣本中位數(C)

自訂(O)

中位數(E)：0

多重比較(M)：所有成對

估計群組間的信賴區間

Hodges-Lehman 估計 (2 個樣本)(H)

8

執行(R) 貼上之後(P) 重設(R) 取消 輔助說明

無母數檢定：兩個以上的獨立樣本

目標 欄位 設定

選取項目(Z)：

選擇檢定(S)

顯著水準(I)：0.05

檢定選項

信賴區間(%) (O)：95.0

使用者遺漏值

排除的觀察值

依檢定排除觀察值(X)

完全排除觀察值(U)

目標 欄位 設定

選取項目(Z)：

選擇檢定(S)

檢定選項

使用者遺漏值

類別欄位的使用者遺漏值

排除(X)

包含(U)

連續欄位中含有使用者遺漏值的觀察值永遠會被排除。

# SPSS統計分析:Mann-Whitney U test

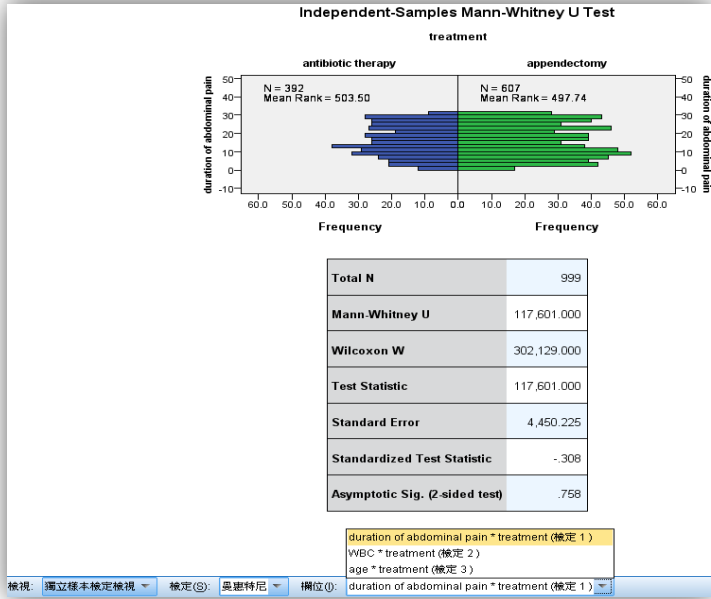
目的:了解age, duration of abdominal pain, WBC count在不同治療的分布狀況。

假定: age, duration of abdominal pain, WBC count為非常態分佈。

## 輸出內容

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of duration of abdominal pain is the same across categories of treatment .	Independent-Samples Mann-Whitney U Test	.758	Retain the null hypothesis.
2	The distribution of WBC is the same across categories of treatment .	Independent-Samples Mann-Whitney U Test	.672	Retain the null hypothesis.
3	The distribution of age is the same across categories of treatment .	Independent-Samples Mann-Whitney U Test	.188	Retain the null hypothesis.

點兩下



# SPSS統計分析: 卡方檢定

目的: 了解sex, agegp, fever, vomiting, diarrhea, ultra, CT, outcome (opd\_30d, emg\_30d, adm\_30d, op\_30d)在不同治療的分布狀況。

## 執行步驟

### 1. 分析 → 敘述統計 → 交叉表

分析(A) 統計圖(G) 公用程式(U) 視窗(W) 說明(H)

- 報表(P)
- 敘述統計(E)
  - 次數分配表(F)...
  - 描述性統計量(D)...
  - 預檢資料(E)...
  - 交叉表(C)...**
  - 比率(R)...
  - P-P 圖(P)...
  - Q-Q 圖(Q)...
- 比較平均數法(M)
- 一般線性模式(G)
- 概化線性模式(Z)
- 混合模式(X)
- 相關(C)
- 迴歸(R)
- 對數線性(O)
- 分類(Y)
- 維度縮減(D)
- 尺度(A)
- 無母數檢定(N)
- 預測(T)
- 存活分析(S)
- 複選題(U)
- 模擬...
- 品質控制(Q)
- ROC 曲線(M)...
- IBM SPSS Amos...

### 2. 將分組變數放入‘列’中 → 將依變數放入‘欄’中 → 設定統計量、儲存格、格式 → 確定/貼上之後

交叉表

列(C):

欄(R):

圖層 1 來自 1

顯示表格層中的階層變數(A)

顯示集群長條圖(B)

隱藏表格(T)

確定 貼上之後(P) 重設(R) 取消 輔助說明

交叉表

列(C):

欄(R):

圖層 1 來自 1

顯示表格層中的階層變數(A)

顯示集群長條圖(B)

隱藏表格(T)

確定 貼上之後(P) 重設(R) 取消 輔助說明

交叉表

列(C):

欄(R):

圖層 1 來自 1

顯示表格層中的階層變數(A)

顯示集群長條圖(B)

隱藏表格(T)

確定 貼上之後(P) 重設(R) 取消 輔助說明

# SPSS統計分析: 卡方檢定

目的: 了解sex, agegp, fever, vomiting, diarrhea, ultra, CT, outcome (opd\_30d, emg\_30d, adm\_30d, op\_30d)在不同治療的分布狀況。

## 執行步驟

### 統計量



交叉表: 統計量

卡方分配(H)  相關(R)

名義

- 列聯係數(O)
- Phi 與 Cramer's V
- Lambda 值(L)
- 不確定係數(U)

次序的

- Gamma 參數(G)
- Somers' d 值(S)
- Kendall's tau-b 相關係數(K)
- Kendall's tau-c 統計量數(C)

名義變數對等距變數

- Eta 值(E)

Kappa 統計量數(K)

風險(I)

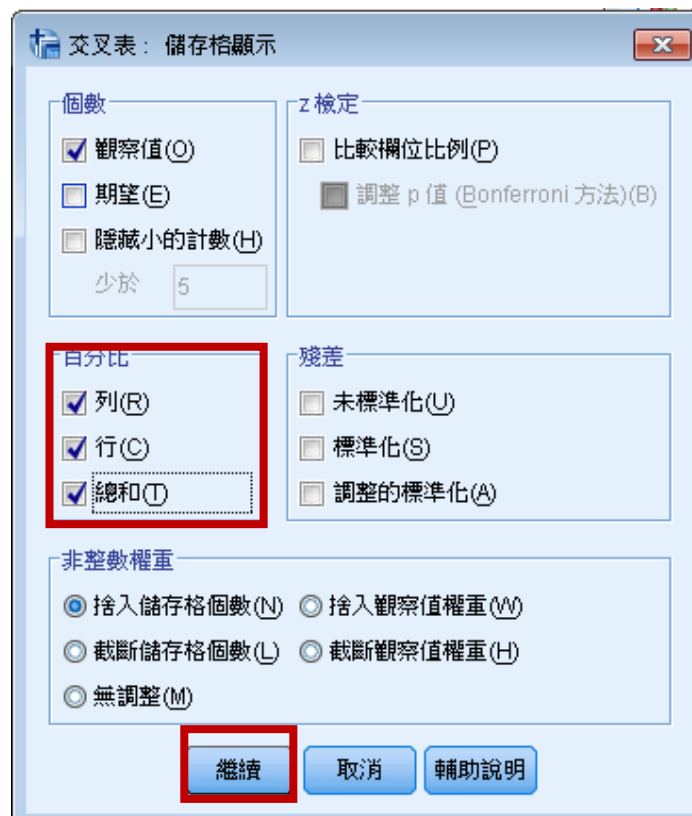
McNemar 檢定(M)

Cochran's 與 Mantel-Haenszel 統計量(A)

檢定共同 odds 比率等於: 1

繼續  取消  輔助說明

### 儲存格



交叉表: 儲存格顯示

個數

- 觀察值(O)
- 期望(E)
- 隱藏小的計數(H)

少於 5

Z 檢定

- 比較欄位比例(P)
- 調整 p 值 (Bonferroni 方法)(B)

百分比

- 列(R)
- 行(C)
- 總和(T)

殘差

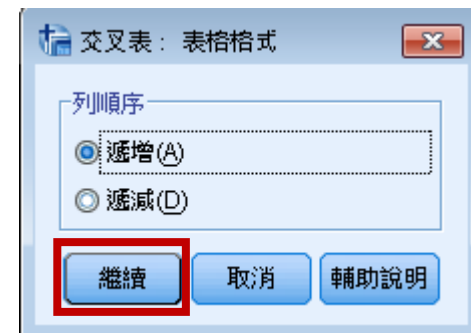
- 未標準化(U)
- 標準化(S)
- 調整的標準化(A)

非整數權重

- 捨入儲存格個數(N)
- 捨入觀察值權重(W)
- 截斷儲存格個數(L)
- 截斷觀察值權重(H)
- 無調整(M)

繼續  取消  輔助說明

### 格式



交叉表: 表格格式

列順序

- 遞增(A)
- 遞減(D)

繼續  取消  輔助說明

# SPSS統計分析: 卡方檢定

目的: 了解sex, agegp, fever, vomiting, diarrhea, ultra, CT, outcome (opd\_30d, emg\_30d, adm\_30d, op\_30d)在不同治療的分布狀況。

## 輸出內容

### 資料表摘要

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
treatment * sex	997	99.7%	3	0.3%	1000	100.0%
treatment * fever	1000	100.0%	0	0.0%	1000	100.0%
treatment * vomiting	999	99.9%	1	0.1%	1000	100.0%
treatment * diarrhea	1000	100.0%	0	0.0%	1000	100.0%
treatment * ultra	1000	100.0%	0	0.0%	1000	100.0%
treatment * CT	1000	100.0%	0	0.0%	1000	100.0%
treatment * outpatient visit or urgent care within 30 days	1000	100.0%	0	0.0%	1000	100.0%
treatment * emergency department within 30 days	999	99.9%	1	0.1%	1000	100.0%
treatment * hospital admission within 30 days	999	99.9%	1	0.1%	1000	100.0%
treatment * surgery or other invasive procedure with 30 days	1000	100.0%	0	0.0%	1000	100.0%
treatment * agegp	1000	100.0%	0	0.0%	1000	100.0%

### 分析結果

treatment \* sex

Crosstab

頻率分布

			sex		
			female	male	Total
treatment	antibiotic therapy	Count	235	156	391
		% within sex	46.6%	31.6%	39.2%
	appendectomy	Count	269	337	606
		% within sex	53.4%	68.4%	60.8%
Total		Count	504	493	997
		% within sex	100.0%	100.0%	100.0%

### 卡方檢定

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	23.473 <sup>a</sup>	1	.000		
Continuity Correction <sup>b</sup>	22.849	1	.000		
Likelihood Ratio	23.597	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	23.450	1	.000		
N of Valid Cases	997				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 193.34.

b. Computed only for a 2x2 table

- ✓ Asymp.Sig. 就是P值 (若P 值<0.05 , 則代表兩組分部有差異存在)
- ✓ 若細格數<5 , 則要看Exact.Sig. (Fisher's exact test)

# SPSS統計分析: 結果整理

Table1. Baseline Demographic and Clinical Characteristics for Patients by Treatment

	antibiotic therapy (N=393)	appendectomy (N=607)	P value
Characteristics , No. (%)			
Age, median (IQR), y	15 (9-23)	15 (8-23)	0.188
0-6	17 (4.3)	19 (3.1)	0.538
7-17	78 (19.8)	121 (19.9)	
18-60	264 (67.2)	401 (66.1)	
60+	34 (8.7)	66 (10.9)	
Sex			
Female	235 (60.1)	269 (44.4)	<0.001
Male	156 (39.9)	337 (55.6)	
Symptoms on initial presentation			
Duration of abdominal pain, median (IQR),h	29 (18-44)	31 (18-48)	0.758
Fever	86 (21.9)	89 (14.7)	0.003
Vomiting	194 (49.5)	199 (32.8)	<0.001
Diarrhea	43 (10.9)	60 (9.9)	0.591
Ultrasonography	366 (93.1)	553 (91.1)	0.251
CT scan	106 (27.0)	198 (32.6)	0.058
WBC count, median (IQR), /ul	17667 (13485-21460)	17327 (13403-21300)	0.672

Table2 Comparison of outcomes between treatment

	antibiotic therapy (N=393)	appendectomy (N=607)	P value
Appendicitis- related medical care within 30 days ,No.(%)			
outpatient visit or urgent care within 30 days	45 (11.5)	120 (19.8)	0.001
emergency department within 30 days	18 (4.6)	18 (3.0)	0.178
hospital admission within 30 days	14 (3.6)	10 (1.6)	0.052
surgery or other invasive procedure with 30 days	15 (3.8)	6 (1.0)	0.002

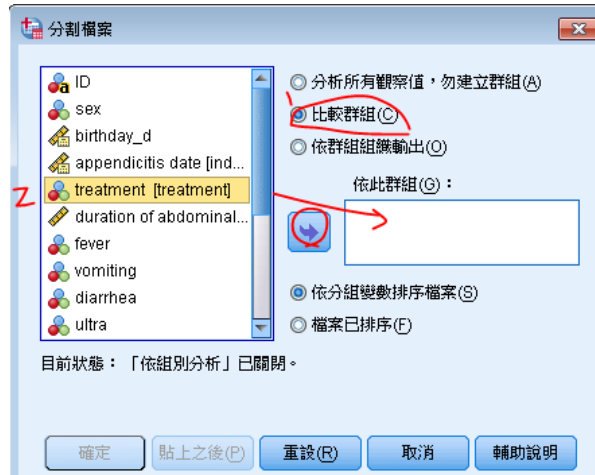
# SPSS資料處理:分割檔案

## 執行步驟 為取得描述性統計量-Median and IQR

### 1.資料→分割檔案



### 2.勾選比較群組→將欲分組變數放入依此群組→確定/貼上之後之後



### 3.將age, duration of abdominal pain, WBC count進行次數分配表分析

### 4.要取消比較群組，就按‘分析所有觀察值，勿建立群組’



**Discussion~**