



存活分析 (1)

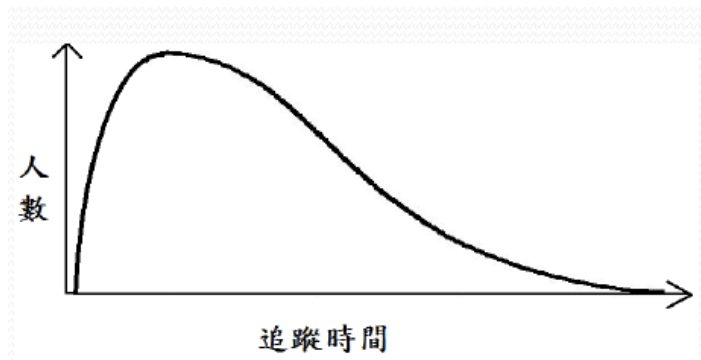
醫學研究部生統小組

陳俊朋

2023/12/6

Survival Analysis

- 將『時間』變項列入分析的統計方法
- 從一時間點至事件(event)發生的時間(time to event) , 稱為survival time
 - Start time- 研究起始時間/確診癌症時間
 - End time- 疾病發生時間/死亡時間
- 資料特性：資料通常不是常態分配

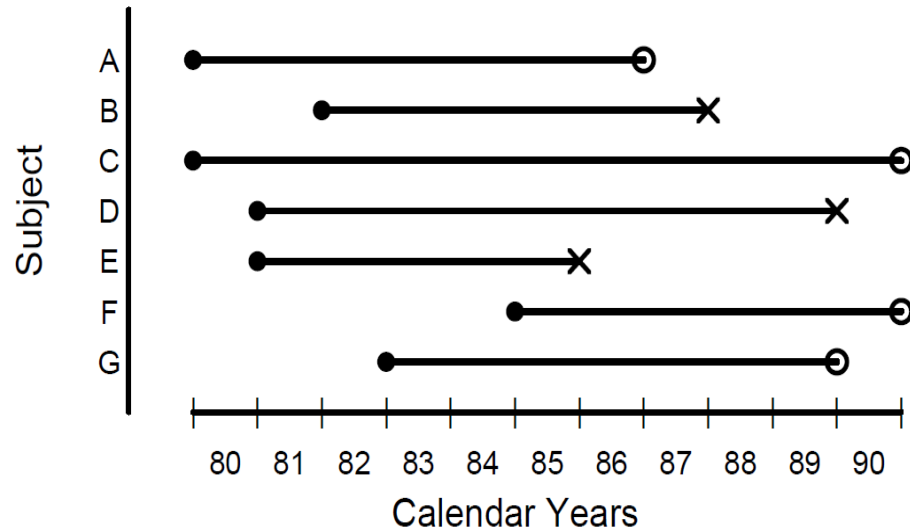


資料分類

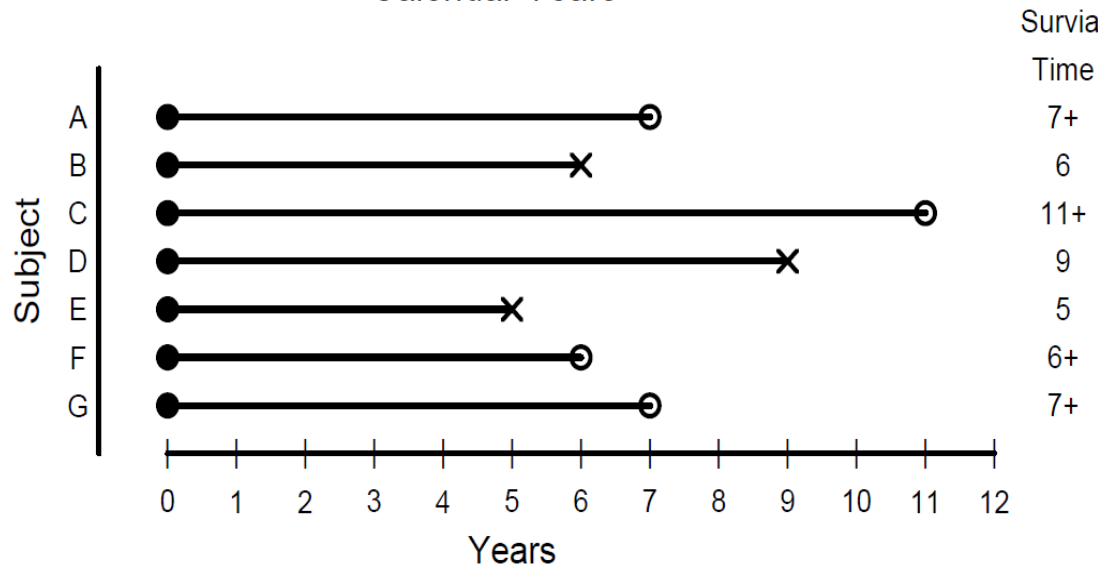
- 完整資料(complete data)
 - 觀察期間提供事件發生的時間點
- 設限資料(censored data)
 - 觀察期間失去聯絡或觀察結束時仍未發生事件
- 競爭死因資料(Competing cause of deaths)
 - 觀察期間死於其他死因之時間

Survival time

(● initial follow-up, × death, ○ alive.)



- Survival time= 事件發生時間 - 追蹤的起始時間 (需大於0)
- 個案BDE是完整資料；個案ACFG為設限資料



存活資料建檔

基本資料

存活狀態

起訖時間

分析整理

| ID | Sex | Age | Tx | event (death) | Start_date | Death_date | follow_date | End_date | follow_year |
|--------|-----|-----|----|---------------|------------|------------|-------------|------------|-------------|
| Case1 | 0 | 45 | 0 | 0 | 2012/5/8 | | 2016/7/15 | 2016/7/15 | 4.19 |
| Case2 | 1 | 55 | 0 | 1 | 2013/12/9 | 2015/1/16 | | 2015/1/16 | 1.10 |
| Case3 | 0 | 56 | 0 | 0 | 2012/5/10 | | 2014/3/12 | 2014/3/12 | 1.84 |
| Case4 | 1 | 51 | 1 | 1 | 2014/9/11 | 2016/7/18 | | 2016/7/18 | 1.85 |
| Case5 | 0 | 62 | 1 | 1 | 2012/8/12 | 2016/12/19 | | 2016/12/19 | 4.35 |
| Case6 | 1 | 70 | 0 | 1 | 2018/7/13 | 2019/7/20 | | 2019/7/20 | 1.02 |
| Case7 | 1 | 58 | 0 | 0 | 2015/5/14 | | 2016/8/21 | 2016/8/21 | 1.27 |
| Case8 | 0 | 66 | 0 | 0 | 2018/8/15 | | 2019/7/22 | 2019/7/22 | 0.93 |
| Case9 | 0 | 60 | 1 | 0 | 2019/5/16 | | 2019/7/23 | 2019/7/23 | 0.19 |
| Case10 | 0 | 63 | 1 | 0 | 2017/10/17 | | 2018/7/24 | 2018/7/24 | 0.77 |
| Case11 | 0 | 66 | 1 | 0 | 2016/3/20 | | 2017/7/25 | 2017/7/25 | 1.35 |
| Case12 | 0 | 69 | 1 | 0 | 2014/8/22 | | 2016/7/26 | 2016/7/26 | 1.93 |

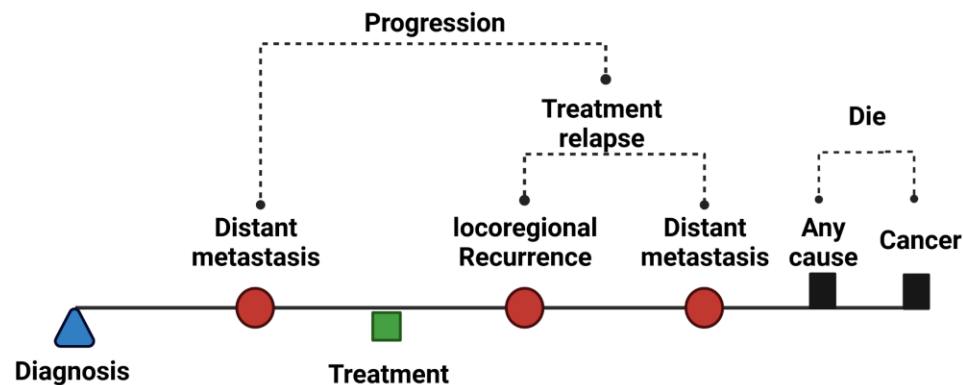
- 日期格式統一西元年
- 可分組比較存活曲線差異，如Sex (F vs M), Age (<60 vs >=60)
- 事件(event): 0=存活; 1=死亡
- 未死亡個案End_date，研究者需設定結束時間(建議可用最後看診日期)。

Definition of Time to Event Variables



| 指標 | 事件 | 設限 |
|--|-----------------------------|-------------------------|
| 1. 無疾病進展存活期 (progression free survival, PFS) | 疾病進展或死亡 | 無疾病進展 |
| 2. 無疾病存活期 (disease free survival, DFS) | 疾病復發或死亡 | 無疾病復發 |
| 3. 治療至疾病進展時間 (time to progression, TTP) | 疾病進展 | 無疾病進展、無疾病狀態死亡(如車禍或不明原因) |
| 4. 治療至治療失敗時間 (time to treatment failure, TTF) | 疾病進展、疾病復發、藥物副作用、死亡 | 無疾病進展或復發 |
| 5. 總存活時間 (overall survival, OS) | 死亡 | 存活 |
| 6. 疾病特定存活期 (disease specific survival, DSS) | 由疾病本身所導致的死亡(車禍或其他原因要歸為設限資料) | 存活、因其他原因死亡 |

Definition for the Assessment of Time-to-event Endpoints in Cancer



Time to progress

The time from diagnosis/treatment to the progression of tumor (in any aspect).

Progression-free survival

The time from diagnosis/treatment to the progression of tumor (in any aspect) or death (for any cause).

Progression-free interval

The time from treatment to locoregional recurrence.

Relapse-free survival

The time from treatment to the relapse (local, regional, distant).

Recurrence-free survival

The time from treatment to locoregional recurrence or death (for any cause)

Disease-free survival

The time from treatment to locoregional recurrence, metastasis or death (for any reason)

Overall survival

The time from diagnosis/treatment to death (for any reason).

Disease-specific survival

The time from diagnosis/treatment to death (cancer-specific).

TTP

PFS

PFI

RpFS

RFS

DFS

OS

DSS

存活分析－統計方法

- 無母數分析
 - Kaplan-Meier
 - Log-rank test
 - Cox Proportional Hazards Model

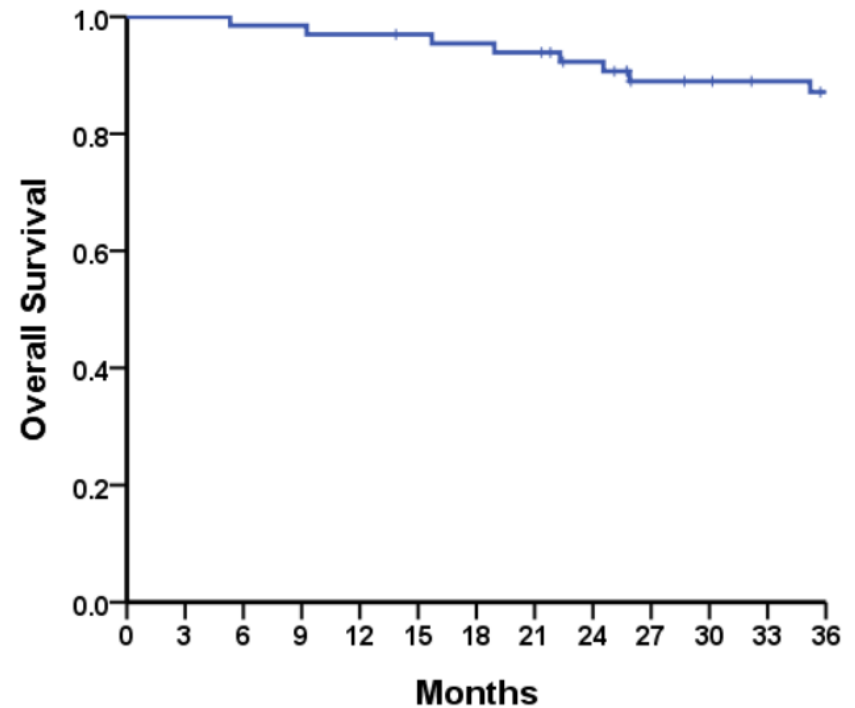
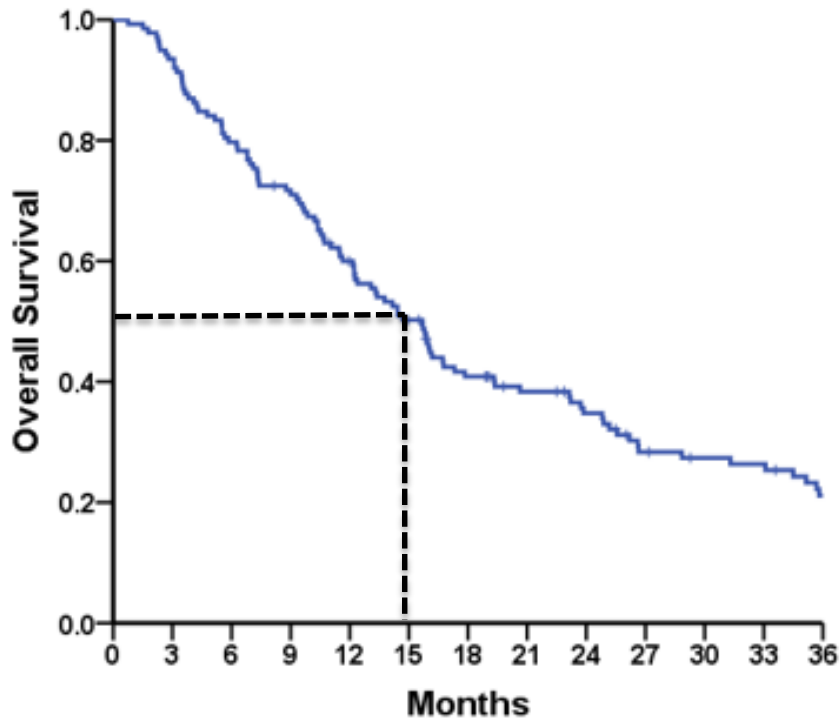
2.2. Statistical Analysis

The most important censored point of this study was the first episode of UBR. UBR-free survival was defined as the interval between the date of NUx and the date of the first episode of UBR. Survival data was analyzed using both the Kaplan–Meier method and the logrank test. Univariate and multivariate analyses by Cox’s proportional hazards model were used to determine the relevance between each of the clinicopathological factors and UBR. P-values less than 0.05 were defined as statistically significant. All statistical analyses were performed with SPSS (Statistical Package for the Social Sciences, version 22.0, IBM, NY, USA).

Chen, Chuan-Shu, et al. *Diagnostics* 10.4 (2020): 201.

Kaplan-Meier

- 常用來估計存活曲線的方法，此方法用每一個事件發生時間點及設限點來設定區間
- 可估計存活中位數及不同時間點的存活率



Kaplan-Meier Estimates



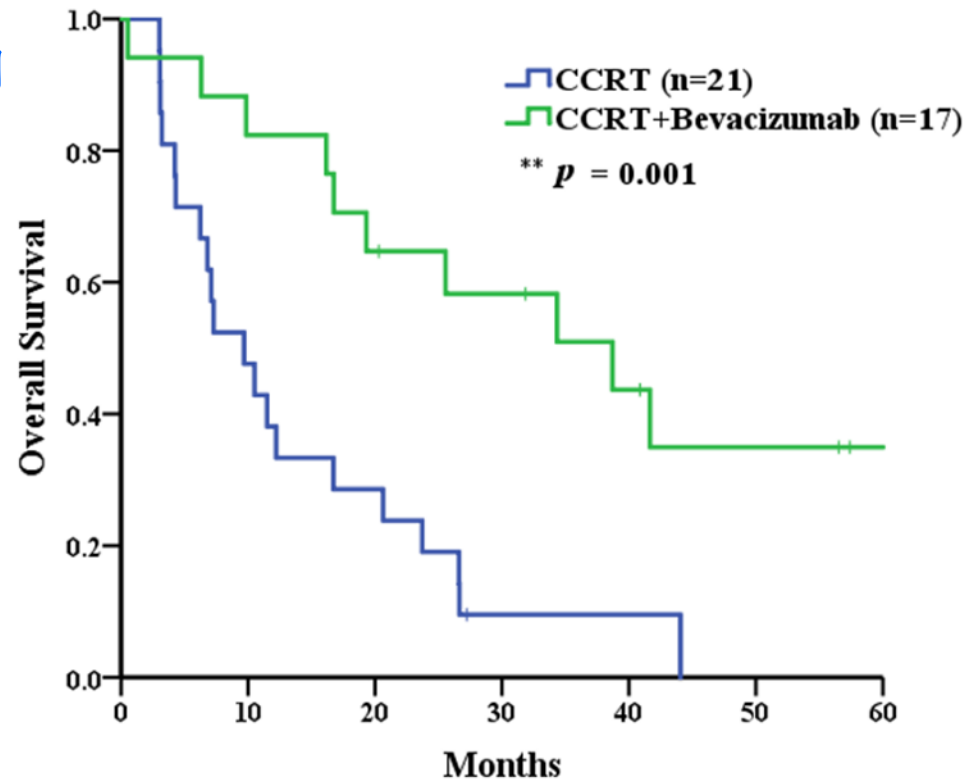
| ID | fu_time (month) | status |
|----|-----------------|--------|
| 1 | 2 | 0 |
| 2 | 2 | 0 |
| 3 | 2 | 1 |
| 4 | 3 | 0 |
| 5 | 3 | 0 |
| 6 | 4 | 1 |
| 7 | 4 | 0 |
| 8 | 5 | 0 |
| 9 | 6 | 1 |
| 10 | 6 | 0 |

| month | Number at risk | Number of events | Number of censored | Conditional Probability | Survival Function |
|-------|----------------|------------------|--------------------|-------------------------|----------------------|
| 1 | 10 | 0 | 0 | $10/10 = 1.00$ | 1.00 |
| 2 | 10 | 1 | 2 | $9/10 = 0.90$ | $0.90 * 1.00 = 0.90$ |
| 3 | 7 | 0 | 2 | $7/7 = 1.00$ | $1.00 * 0.90 = 0.90$ |
| 4 | 5 | 1 | 1 | $4/5 = 0.80$ | $0.80 * 0.90 = 0.72$ |
| 5 | 3 | 0 | 1 | $3/3 = 1.00$ | $1.00 * 0.72 = 0.72$ |
| 6 | 2 | 1 | 1 | $1/2 = 0.50$ | $0.50 * 0.72 = 0.36$ |

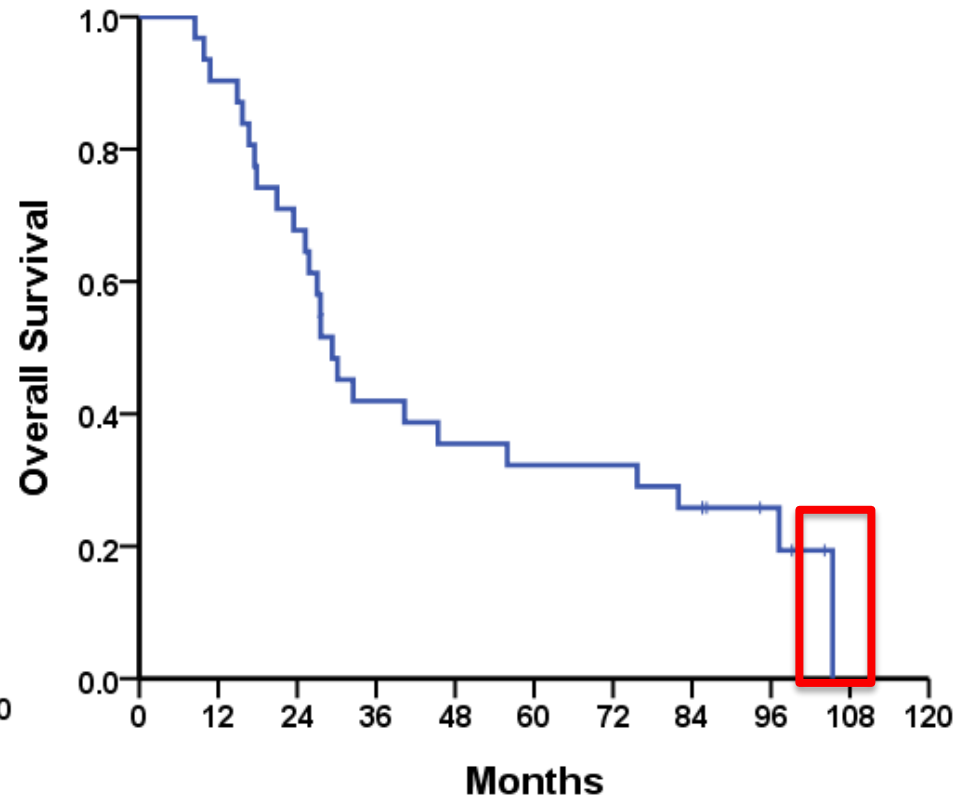
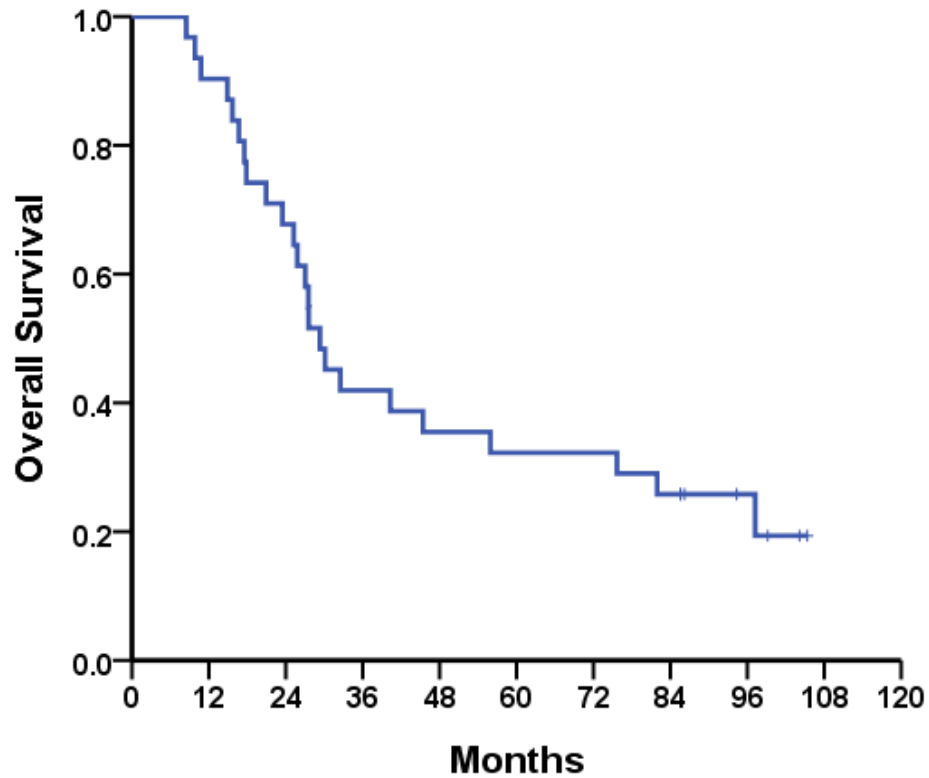
Log-rank Test



- Kaplan-Meier法僅能了解不同組別的存活曲線分佈
- 比較組別間是否差異
 - H0: 兩條存活曲線相同
 - H1: 兩條存活曲線不相同



Survival Curve



Survival Curve



Survival Table

| | Time | Status | Cumulative Proportion Surviving at the Time | | N of Cumulative Events | N of Remaining Cases |
|----|---------|----------|---|------------|------------------------|----------------------|
| | | | Estimate | Std. Error | | |
| 1 | 8.467 | Death | .968 | .032 | 1 | 30 |
| 2 | 9.833 | Death | .935 | .044 | 2 | 29 |
| 3 | 10.767 | Death | .903 | .053 | 3 | 28 |
| 4 | 14.900 | Death | .871 | .060 | 4 | 27 |
| 5 | 15.667 | Death | .839 | .066 | 5 | 26 |
| 6 | 16.667 | Death | .806 | .071 | 6 | 25 |
| 7 | 17.533 | Death | .774 | .075 | 7 | 24 |
| 8 | 17.833 | Death | .742 | .079 | 8 | 23 |
| 9 | 20.933 | Death | .710 | .082 | 9 | 22 |
| 10 | 23.467 | Death | .677 | .084 | 10 | 21 |
| 11 | 25.233 | Death | .645 | .086 | 11 | 20 |
| 12 | 25.800 | Death | .613 | .087 | 12 | 19 |
| 13 | 27.033 | Death | .581 | .089 | 13 | 18 |
| 14 | 27.533 | Death | .548 | .089 | 14 | 17 |
| 15 | 27.567 | Death | .516 | .090 | 15 | 16 |
| 16 | 29.333 | Death | .484 | .090 | 16 | 15 |
| 17 | 30.133 | Death | .452 | .089 | 17 | 14 |
| 18 | 32.500 | Death | .419 | .089 | 18 | 13 |
| 19 | 40.300 | Death | .387 | .087 | 19 | 12 |
| 20 | 45.400 | Death | .355 | .086 | 20 | 11 |
| 21 | 55.933 | Death | .323 | .084 | 21 | 10 |
| 22 | 75.667 | Death | .290 | .082 | 22 | 9 |
| 23 | 81.933 | Death | .258 | .079 | 23 | 8 |
| 24 | 85.567 | Survival | . | . | 23 | 7 |
| 25 | 85.600 | Survival | . | . | 23 | 6 |
| 26 | 86.233 | Survival | . | . | 23 | 5 |
| 27 | 94.333 | Survival | . | . | 23 | 4 |
| 28 | 97.233 | Death | .194 | .081 | 24 | 3 |
| 29 | 99.167 | Survival | . | . | 24 | 2 |
| 30 | 104.167 | Survival | . | . | 24 | 1 |
| 31 | 105.367 | Survival | . | . | 24 | 0 |

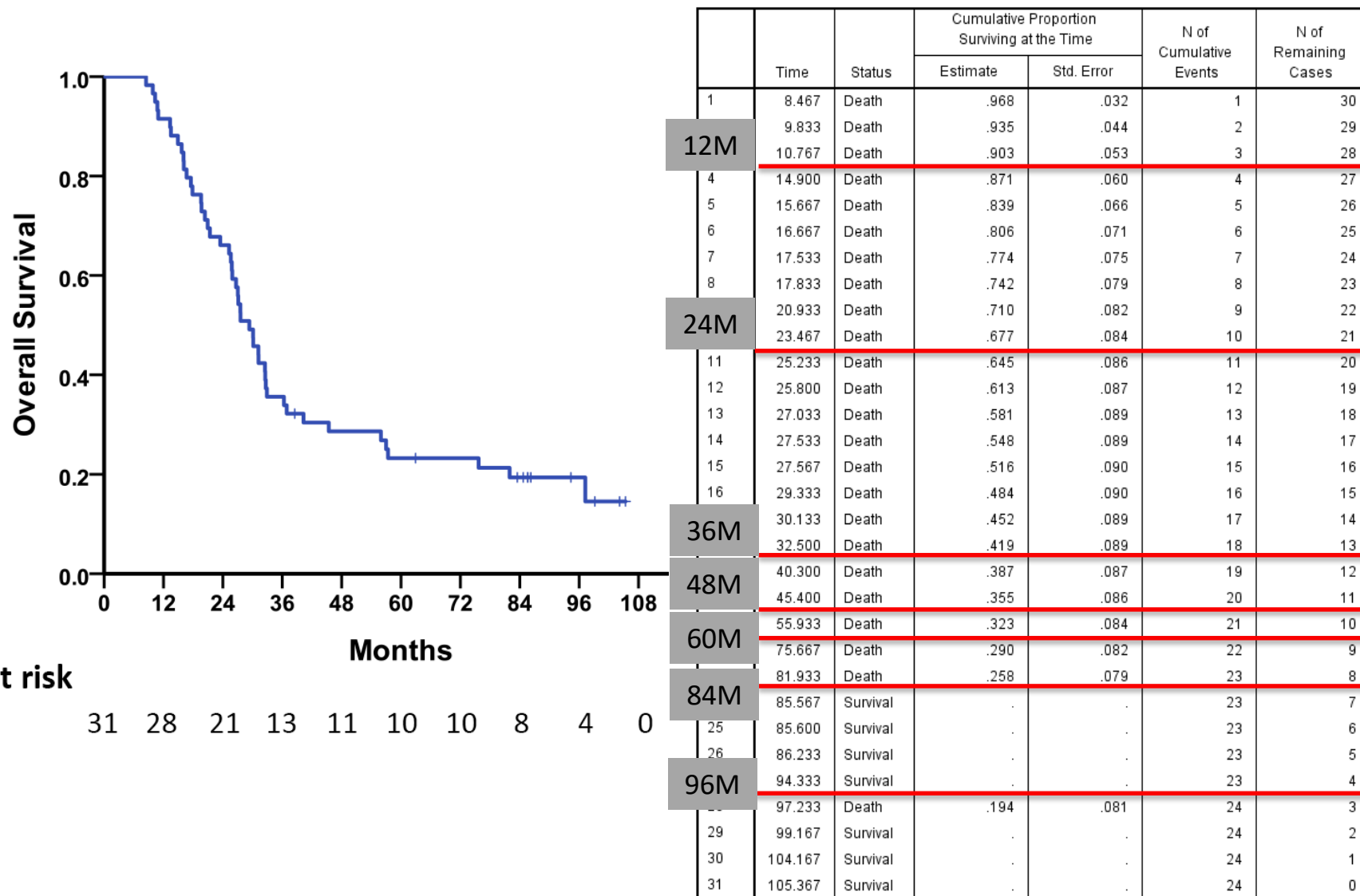
Survival Table

| | Time | Status | Cumulative Proportion Surviving at the Time | | N of Cumulative Events | N of Remaining Cases |
|----|---------|----------|---|------------|------------------------|----------------------|
| | | | Estimate | Std. Error | | |
| 1 | 8.467 | Death | .968 | .032 | 1 | 30 |
| 2 | 9.833 | Death | .935 | .044 | 2 | 29 |
| 3 | 10.767 | Death | .903 | .053 | 3 | 28 |
| 4 | 14.900 | Death | .871 | .060 | 4 | 27 |
| 5 | 15.667 | Death | .839 | .066 | 5 | 26 |
| 6 | 16.667 | Death | .806 | .071 | 6 | 25 |
| 7 | 17.533 | Death | .774 | .075 | 7 | 24 |
| 8 | 17.833 | Death | .742 | .079 | 8 | 23 |
| 9 | 20.933 | Death | .710 | .082 | 9 | 22 |
| 10 | 23.467 | Death | .677 | .084 | 10 | 21 |
| 11 | 25.233 | Death | .645 | .086 | 11 | 20 |
| 12 | 25.800 | Death | .613 | .087 | 12 | 19 |
| 13 | 27.033 | Death | .581 | .089 | 13 | 18 |
| 14 | 27.533 | Death | .548 | .089 | 14 | 17 |
| 15 | 27.567 | Death | .516 | .090 | 15 | 16 |
| 16 | 29.333 | Death | .484 | .090 | 16 | 15 |
| 17 | 30.133 | Death | .452 | .089 | 17 | 14 |
| 18 | 32.500 | Death | .419 | .089 | 18 | 13 |
| 19 | 40.300 | Death | .387 | .087 | 19 | 12 |
| 20 | 45.400 | Death | .355 | .086 | 20 | 11 |
| 21 | 55.933 | Death | .323 | .084 | 21 | 10 |
| 22 | 75.667 | Death | .290 | .082 | 22 | 9 |
| 23 | 81.933 | Death | .258 | .079 | 23 | 8 |
| 24 | 85.567 | Survival | . | . | 23 | 7 |
| 25 | 85.600 | Survival | . | . | 23 | 6 |
| 26 | 86.233 | Survival | . | . | 23 | 5 |
| 27 | 94.333 | Survival | . | . | 23 | 4 |
| 28 | 97.233 | Death | .194 | .081 | 24 | 3 |
| 29 | 99.167 | Survival | . | . | 24 | 2 |
| 30 | 104.167 | Survival | . | . | 24 | 1 |
| 31 | 105.367 | Death | .000 | .000 | 25 | 0 |

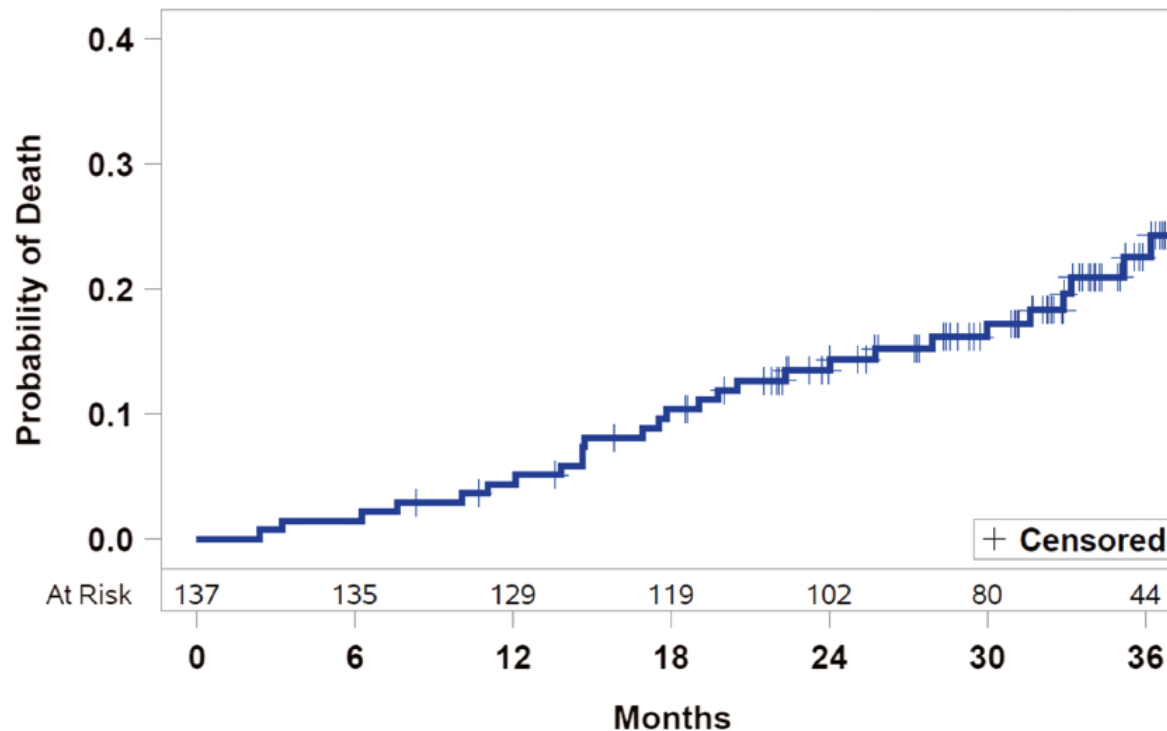
Presentation of Survival Curve



Survival Table



Presentation of Survival Curve



Months represent months from the date of therapy activation.

Figure 5. Kaplan–Meier curve of mortality. Kaplan–Meier curve showing estimated mortality through 36 months of active therapy using the pooled treatment and former control groups. Patients who did not die were censored at last contact if they did not reach 36 months of active therapy.

Cox Proportional Hazards Model

- 評估多個變數對存活(Time to event)的影響
- 可使用類別/連續變數呈現危險因子，並估算出這些危險因子對outcome的影響

$$\log \frac{h(t)}{h_0(t)} = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

| Hazard ratio (HR) | Explanation |
|-------------------|--|
| HR=1 | Two groups have the same survival experience |
| HR>1 | Survival is better in the control group |
| HR<1 | Survival is better in the treatment group |

Kaplan-Meier Example



- 比較在第三期子宮內膜癌術後治療，Sandwich組和CT alone兩組治療預後的差異

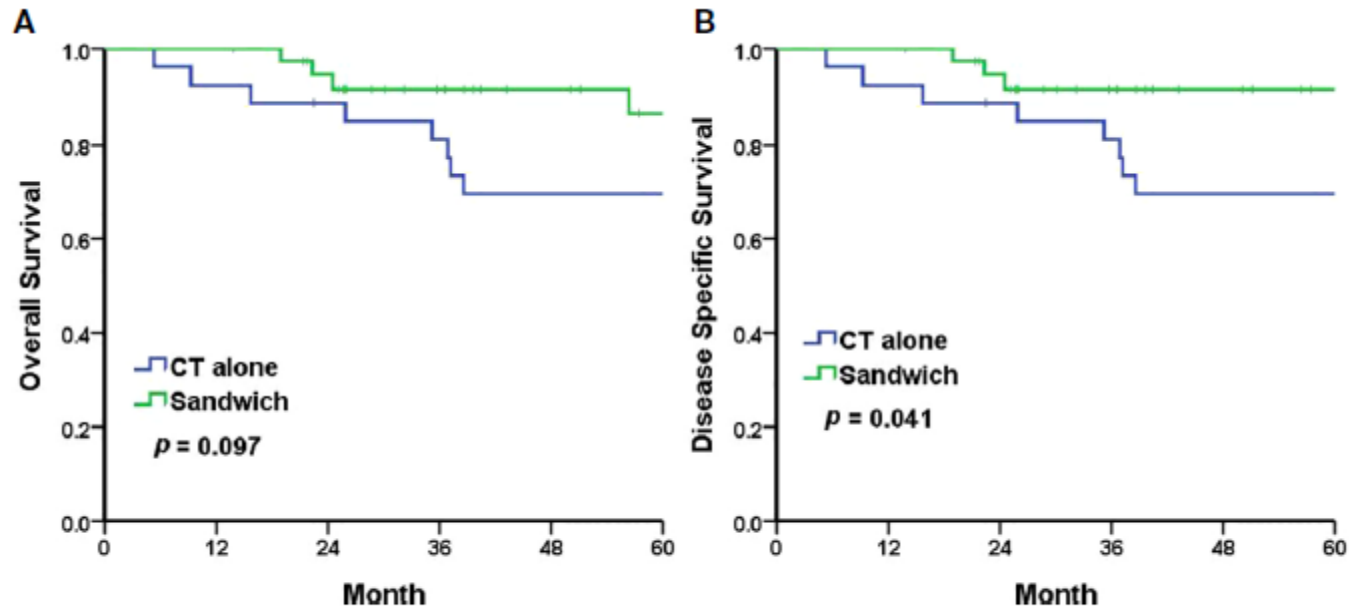


FIGURE 2

The Kaplan–Meier survival curves for 5-year overall survival (A) and 5-year disease-specific survival (B). CT chemotherapy.

Outcomes of " sandwich " chemoradiotherapy compared with chemotherapy alone for the adjuvant treatment of FIGO stage III endometrial cancer. *Frontiers in oncology*, 12 (2022), 946113-946113.

Cox Regression Example



5-year disease-specific survival

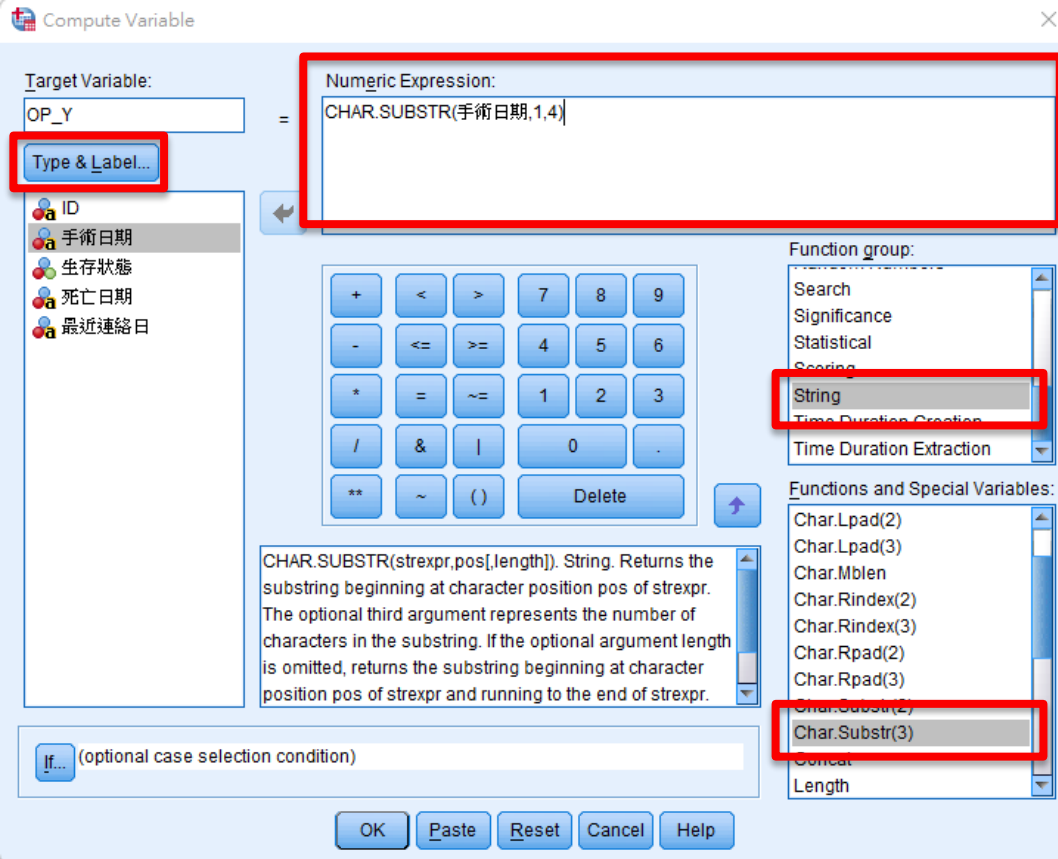
| | Univariate | | | Multivariable | | |
|--------------------------|------------|--------------|---------|---------------|--------------|---------|
| | HR | 95% CI | p-value | HR | 95% CI | p-value |
| Age group | | | | | | |
| <60 | Reference | | | | | |
| ≥60 | 1.34 | (0.36-5.07) | 0.665 | | | |
| FIGO stage | | | | | | |
| IIIA and IIIB and IIIC1 | Reference | | | | | |
| IIIC2 | 1.45 | (0.44-4.75) | 0.541 | | | |
| Histology grading | | | | | | |
| Grades 1 and 2 | Reference | | | Reference | | |
| Grade 3 | 8.70 | (1.11-68.01) | 0.039* | 9.16 | (1.17-71.70) | 0.035* |
| Treatment | | | | | | |
| CT alone | Reference | | | Reference | | |
| Sandwich | 0.27 | (0.07-1.04) | 0.056 | 0.23 | (0.06-0.87) | 0.030* |
| LVSI | | | | | | |
| Absent | Reference | | | | | |
| Present | 3.82 | (0.49-29.89) | 0.201 | | | |
| Deep myometrial invasion | | | | | | |
| Absent | Reference | | | Reference | | |
| Present | 7.85 | (1.00-61.34) | 0.050 | 9.44 | (1.20-74.15) | 0.033* |

Outcomes of "sandwich" chemoradiotherapy compared with chemotherapy alone for the adjuvant treatment of FIGO stage III endometrial cancer. *Frontiers in oncology*, 12 (2022), 946113-946113.

日期變數處理-SPSS

- 日期/時間計算 (函數說明)
 - 月或日為99時，利用計算變數轉換數值。月=99轉為7、日=99轉為15。
 - 保留特定區段的字串(CHAR.SUBSTR)
 - CHAR.SUBSTR(變數,起始位置,保留長度)。
 - EX: CHAR.SUBSTR(20090731,1,4) = 2009

日期變數處理



Target Variable: OP_Y =

Numeric Expression: CHAR.SUBSTR(手術日期,1,4)

Type & Label...

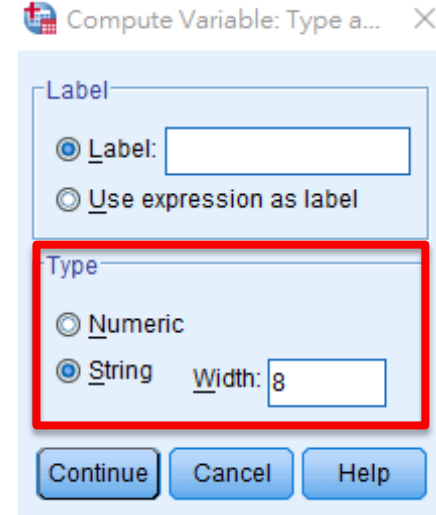
Function group: String

Functions and Special Variables: Char.Substr(3)

CHAR.SUBSTR(strexp, pos[, length]). String. Returns the substring beginning at character position pos of strexp. The optional third argument represents the number of characters in the substring. If the optional argument length is omitted, returns the substring beginning at character position pos of strexp and running to the end of strexp.

If... (optional case selection condition)

OK Paste Reset Cancel Help



Label

Label:

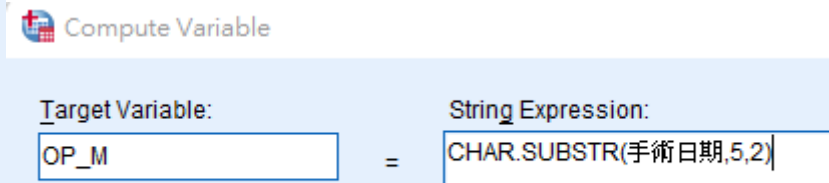
Use expression as label

Type

Numeric

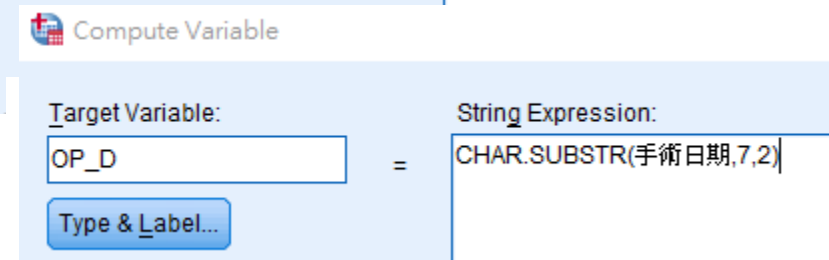
String Width: 8

Continue Cancel Help



Target Variable: OP_M =

String Expression: CHAR.SUBSTR(手術日期,5,2)



Target Variable: OP_D =

String Expression: CHAR.SUBSTR(手術日期,7,2)

Type & Label...

日期變數處理

| ID | 手術日期 | 生存狀態 | 死亡日期 | 最近連絡日 | OP_Y | OP_M | OP_D |
|--------|----------|------|----------|----------|------|------|------|
| case1 | 20060303 | .0 | 20080525 | | 2006 | 03 | 03 |
| case2 | 20120517 | .0 | 20121026 | | 2012 | 05 | 17 |
| case3 | 20060503 | .0 | 20090617 | | 2006 | 05 | 03 |
| case4 | 20101112 | .0 | 20160402 | | 2010 | 11 | 12 |
| case5 | 20120803 | .0 | 20130512 | | 2012 | 08 | 03 |
| case6 | 20110224 | .0 | 20140514 | | 2011 | 02 | 24 |
| case7 | 20121199 | .0 | 20141016 | | 2012 | 11 | 99 |
| case8 | 20130821 | .0 | 20141212 | | 2013 | 08 | 21 |
| case9 | 20140523 | .0 | 20160608 | | 2014 | 05 | 23 |
| case10 | 20160113 | .0 | 20170811 | | 2016 | 01 | 13 |
| case11 | 20150303 | .0 | 20180329 | | 2015 | 03 | 03 |
| case12 | 20160905 | .0 | 20190812 | | 2016 | 09 | 05 |
| case13 | 20141219 | .0 | 20200125 | | 2014 | 12 | 19 |
| case14 | 20120420 | 1.0 | | 20150710 | 2012 | 04 | 20 |
| case15 | 20151229 | 1.0 | | 20211004 | 2015 | 12 | 29 |
| case16 | 20160420 | 1.0 | | 20210909 | 2016 | 04 | 20 |
| case17 | 20190827 | 1.0 | | 20211018 | 2019 | 08 | 27 |
| case18 | 20050820 | 1.0 | | 20170831 | 2005 | 08 | 20 |
| case19 | 20051016 | 1.0 | | 20180816 | 2005 | 10 | 16 |
| case20 | 20111224 | 1.0 | | 20151018 | 2011 | 12 | 24 |

日期變數處理

| OP_Y | OP_M | OP_D | OP_D2 |
|------|------|------|-------|
| 2006 | 03 | 03 | 03 |
| 2012 | 05 | 17 | 17 |
| 2006 | 05 | 03 | 03 |
| 2010 | 11 | 12 | 12 |
| 2012 | 08 | 03 | 03 |
| 2011 | 02 | 24 | 24 |
| 2012 | 11 | 99 | 15 |
| 2013 | 08 | 21 | 21 |
| 2014 | 05 | 23 | 23 |
| 2016 | 01 | 13 | 13 |
| 2015 | 03 | 03 | 03 |
| 2016 | 09 | 05 | 05 |
| 2014 | 12 | 19 | 19 |
| 2012 | 04 | 20 | 20 |
| 2015 | 12 | 29 | 29 |
| 2016 | 04 | 20 | 20 |
| 2019 | 08 | 27 | 27 |
| 2005 | 08 | 20 | 20 |
| 2005 | 10 | 16 | 16 |
| 2011 | 12 | 24 | 24 |

Compute Variable

Target Variable: OP_D2 = String Expression: OP_D

Type & Label...

Compute Variable

Target Variable: OP_D2 = String Expression: "15"

Type & Label...

ID
 手術日期
 生存狀態
 死亡日期
 最近連絡日
 OP_Y
 OP_M
 OP_D
 OP_D2

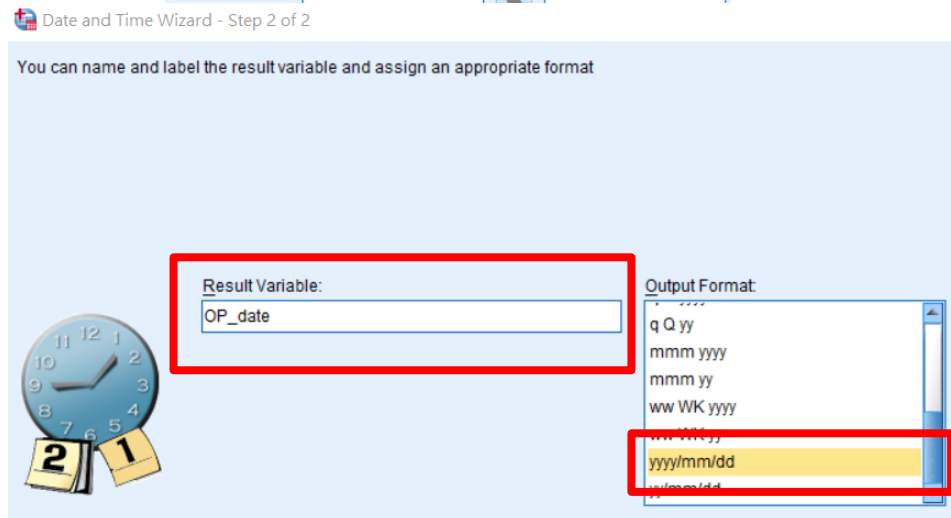
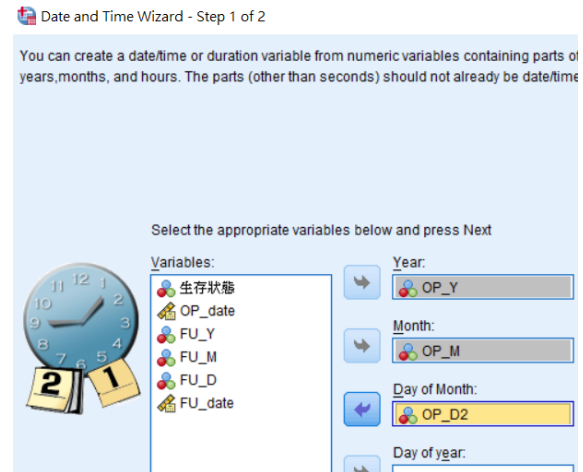
If... OP_D = "99"

OK Paste Reset Cancel Help

日期變數處理

- OP_Y、OP_M、OP_D2 → 轉為數值型資料
- 轉換>日期和時間精靈>從保有部分日期或時間的變數建立日期/時間變數

| OP_Y | OP_M | OP_D | OP_D2 |
|------|------|------|-------|
| 2006 | 3 | 03 | 3 |
| 2012 | 5 | 17 | 17 |
| 2006 | 5 | 03 | 3 |
| 2010 | 11 | 12 | 12 |
| 2012 | 8 | 03 | 3 |
| 2011 | 2 | 24 | 24 |
| 2012 | 11 | 99 | 15 |
| 2013 | 8 | 21 | 21 |
| 2014 | 5 | 23 | 23 |
| 2016 | 1 | 13 | 13 |
| 2015 | 3 | 03 | 3 |
| 2016 | 9 | 05 | 5 |
| 2014 | 12 | 19 | 19 |
| 2012 | 4 | 20 | 20 |
| 2015 | 12 | 29 | 29 |
| 2016 | 4 | 20 | 20 |
| 2019 | 8 | 27 | 27 |
| 2005 | 8 | 20 | 20 |
| 2005 | 10 | 16 | 16 |
| 2011 | 12 | 24 | 24 |



日期變數處理

| OP_Y | OP_M | OP_D | OP_D2 | OP_date |
|------|------|------|-------|------------|
| 2006 | 3 | 03 | 3 | 2006/03/03 |
| 2012 | 5 | 17 | 17 | 2012/05/17 |
| 2006 | 5 | 03 | 3 | 2006/05/03 |
| 2010 | 11 | 12 | 12 | 2010/11/12 |
| 2012 | 8 | 03 | 3 | 2012/08/03 |
| 2011 | 2 | 24 | 24 | 2011/02/24 |
| 2012 | 11 | 99 | 15 | 2012/11/15 |
| 2013 | 8 | 21 | 21 | 2013/08/21 |
| 2014 | 5 | 23 | 23 | 2014/05/23 |
| 2016 | 1 | 13 | 13 | 2016/01/13 |
| 2015 | 3 | 03 | 3 | 2015/03/03 |
| 2016 | 9 | 05 | 5 | 2016/09/05 |
| 2014 | 12 | 19 | 19 | 2014/12/19 |
| 2012 | 4 | 20 | 20 | 2012/04/20 |
| 2015 | 12 | 29 | 29 | 2015/12/29 |
| 2016 | 4 | 20 | 20 | 2016/04/20 |
| 2019 | 8 | 27 | 27 | 2019/08/27 |
| 2005 | 8 | 20 | 20 | 2005/08/20 |
| 2005 | 10 | 16 | 16 | 2005/10/16 |
| 2011 | 12 | 24 | 24 | 2011/12/24 |

日期變數處理

| 生存狀態 | 死亡日期 | 最近連絡日 | FU |
|------|----------|----------|----------|
| .0 | 20080525 | | 20080525 |
| .0 | 20121026 | | 20121026 |
| .0 | 20090617 | | 20090617 |
| .0 | 20160402 | | 20160402 |
| .0 | 20130512 | | 20130512 |
| .0 | 20140514 | | 20140514 |
| .0 | 20141016 | | 20141016 |
| .0 | 20141212 | | 20141212 |
| .0 | 20160608 | | 20160608 |
| .0 | 20170811 | | 20170811 |
| .0 | 20180329 | | 20180329 |
| .0 | 20190812 | | 20190812 |
| .0 | 20200125 | | 20200125 |
| 1.0 | | 20150710 | 20150710 |
| 1.0 | | 20211004 | 20211004 |
| 1.0 | | 20210909 | 20210909 |
| 1.0 | | 20211018 | 20211018 |
| 1.0 | | 20170831 | 20170831 |
| 1.0 | | 20180816 | 20180816 |
| 1.0 | | 20151018 | 20151018 |

Compute Variable

Target Variable: = Numeric Expression:

Compute Variable

Target Variable: = Numeric Expression:

- ID
- 手術日期
- 生存狀態
- 死亡日期
- 最近連絡日
- FU
- OP_Y
- OP_M
- OP_D
- OP_D2
- OP_date

計算追蹤時間

轉換>日期和時間精靈>以日期和時間計算>計算介於兩個日期之間的時間單位數


Transform Analyze Direct Marketing Graphs

- Compute Variable...
 - Programmability Transformation...
- Count Values within Cases...
 - Shift Values...
- Recode into Same Variables...
- Recode into Different Variables...
- Automatic Recode...
 - Create Dummy Variables
- Visual Binning...
- Optimal Binning...
 - Prepare Data for Modeling
- Rank Cases...
- Date and Time Wizard...**
- Create Time Series...
- Replace Missing Values...
- Random Number Generators...

Date and Time Wizard - Step 2 of 3

Calculate the number of time units between two date or date/time variables.

The result will be an integer variable. Any fractional part of a unit will be discarded. The result will be a duration variable. Only duration variables are shown in the variables list below.



Variables:

- Current date and ti...

Date1:

- FU_date

minus Date2:

- OP_date

Unit:

- Years

Result Treatment:

- Truncate to integer
- Round to integer
- Retain fractional part

For month and year units, the result is based on average unit length unless truncation is used.

STIME is the current date and time.

Date and Time Wizard - Step 3 of 3


Calculation: FU_date - OP_date

Result Variable: OS_year Units: Years

Variable Label:

Execution:

- Create the variable now
- Paste the syntax into the syntax window



計算追蹤時間

- 日期格式靠右
- $OS_year = FU_date - OP_date$

| OP_date | FU_date | OS_year |
|------------|------------|---------|
| 2006/03/03 | 2008/05/25 | 2.23 |
| 2012/05/17 | 2012/10/26 | .44 |
| 2006/05/03 | 2009/06/17 | 3.12 |
| 2010/11/12 | 2016/04/02 | 5.39 |
| 2012/08/03 | 2013/05/12 | .77 |
| 2011/02/24 | 2014/05/14 | 3.22 |
| 2012/11/15 | 2014/10/16 | 1.92 |
| 2013/08/21 | 2014/12/12 | 1.31 |
| 2014/05/23 | 2016/06/08 | 2.05 |
| 2016/01/13 | 2017/08/11 | 1.58 |
| 2015/03/03 | 2018/03/29 | 3.07 |
| 2016/09/05 | 2019/08/12 | 2.93 |
| 2014/12/19 | 2020/01/25 | 5.10 |
| 2012/04/20 | 2015/07/10 | 3.22 |
| 2015/12/29 | 2021/10/04 | 5.77 |
| 2016/04/20 | 2021/09/09 | 5.39 |
| 2019/08/27 | 2021/10/18 | 2.14 |
| 2005/08/20 | 2017/08/31 | 12.03 |
| 2005/10/16 | 2018/08/16 | 12.83 |
| 2011/12/24 | 2015/10/18 | 3.82 |

KM and Cox Regression SPSS dataset



| ID | group | OP_date | End_Date | STATUS_5y | STATUS_DSS_5y | Month_OS | Age_gp | Hist_grading |
|----|-------|------------|------------|-----------|---------------|----------|--------|--------------|
| 1 | 1 | 2016/04/20 | 2021/09/09 | 1 | 0 | 64.66 | 0 | 2 |
| 2 | 1 | 2012/04/27 | 2020/02/20 | 1 | 0 | 93.80 | 1 | 1 |
| 3 | 1 | 2015/12/30 | 2021/01/19 | 1 | 0 | 60.68 | 0 | 1 |
| 4 | 0 | 2010/11/12 | 2016/04/02 | 1 | 0 | 64.66 | 0 | 2 |
| 5 | 1 | 2012/11/21 | 2014/10/01 | 0 | 1 | 22.31 | 0 | 2 |
| 6 | 1 | 2013/10/08 | 2020/02/12 | 1 | 0 | 76.16 | 0 | 1 |
| 7 | 0 | 2006/03/29 | 2008/05/25 | 0 | 1 | 25.89 | 0 | 1 |
| 8 | 1 | 2012/06/19 | 2020/12/23 | 1 | 0 | 102.14 | 1 | 2 |
| 9 | 1 | 2018/10/29 | 2021/10/20 | 1 | 0 | 35.71 | 1 | 1 |
| 10 | 1 | 2019/08/27 | 2021/10/18 | 1 | 0 | 25.72 | 0 | 2 |
| 11 | 0 | 2010/02/05 | 2017/11/24 | 1 | 0 | 93.60 | 0 | 2 |
| 12 | 0 | 2011/02/24 | 2014/05/14 | 0 | 1 | 38.60 | 0 | 2 |
| 13 | 1 | 2017/08/25 | 2021/01/05 | 1 | 0 | 40.38 | 1 | 1 |
| 14 | 1 | 2012/04/20 | 2015/07/10 | 1 | 0 | 38.64 | 0 | 2 |
| 15 | 0 | 2005/10/06 | 2015/02/18 | 1 | 0 | 112.43 | 0 | 2 |
| 16 | 1 | 2018/06/25 | 2021/10/14 | 1 | 0 | 39.66 | 0 | 2 |
| 17 | 1 | 2015/04/08 | 2021/09/28 | 1 | 0 | 77.70 | 1 | 1 |
| 18 | 1 | 2014/10/17 | 2021/10/27 | 1 | 0 | 84.34 | 0 | 1 |
| 19 | 0 | 2012/08/03 | 2013/05/12 | 0 | 1 | 9.26 | 0 | 2 |
| 20 | 1 | 2013/01/14 | 2017/09/26 | 0 | 0 | 56.38 | 0 | 2 |

Kaplan-Meier 分析操作



分析 > 存活分析
> Kaplan-Meier 統計

The image shows the SPSS Analyze menu with the following items:

- Reports
- Descriptive Statistics
- Tables
- Compare Means
- General Linear Model
- Generalized Linear Models
- Mixed Models
- Correlate
- Regression
- Loglinear
- Neural Networks
- Classify
- Dimension Reduction
- Scale
- Nonparametric Tests
- Forecasting
- Survival** (highlighted)
 - Life Tables...
 - Kaplan-Meier...** (highlighted)
 - Cox Regression...
 - Cox w/ Time-Dep Cov...
- Multiple Response
- PS Matching
- Missing Value Analysis...
- Multiple Imputation

The background shows a data view with columns STATUS_5y and STATUS_DSS_5y, and a value of 1 in the STATUS_5y column.

Kaplan-Meier 分析操作



Kaplan-Meier

- ID
- OP_date
- End_Date
- STATUS_5y
- Age_gp
- Hist_grading
- FIGO_stage
- LVSI
- Deep myometrium invasion [...]
- BMI25_gp
- Cx involvement [Cxinvolveme...]
- 病歷號碼

Time: Month_OS

Status: STATUS_DSS_5y(1)

Define Event...

Compare Factor...

Save...

Options...

Factor: group

Strata:

Label Cases by:

OK Paste Reset Cancel

Kaplan-Meier: Compare Factor Levels

Test Statistics

Log rank Breslow Tarone-Ware

Linear trend for factor levels

Pooled over strata

Pairwise over strata

For each stratum

Pairwise for each stratum

Continue Cancel Help

Kaplan-Meier: Define Event For Status ...

Value(s) indicating event has occurred

Single value: 1

Range of values: through

List of values:

Add

Change

Remove

Continue Cancel Help

Kaplan-Meier: Options

Statistics

Survival table(s)

Mean and median survival

Quartiles

Plots

Survival

One minus survival

Hazard

Log Survival

Continue Cancel Help

Kaplan-Meier Output



Case Processing Summary

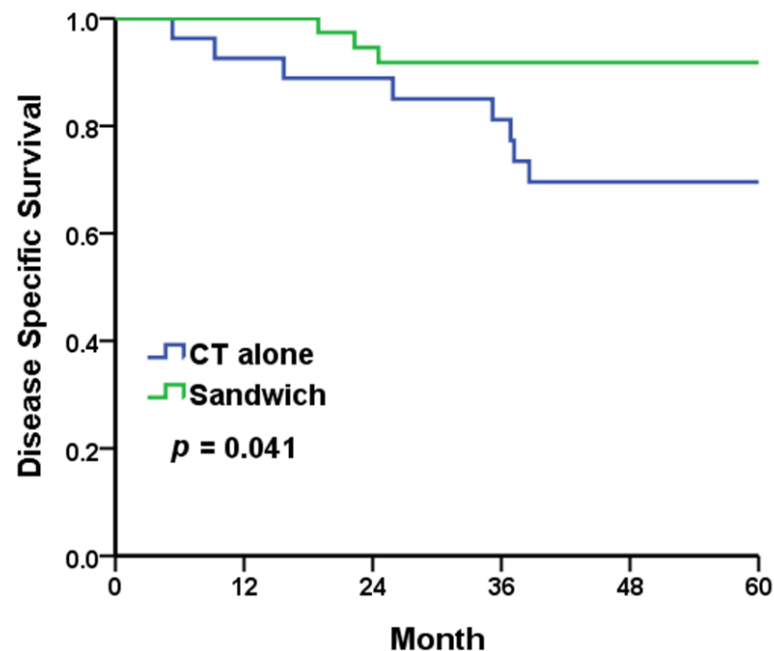
| group | Total N | N of Events | Censored | |
|----------|---------|-------------|----------|---------|
| | | | N | Percent |
| CT alone | 27 | 8 | 19 | 70.4% |
| Sandwich | 39 | 3 | 36 | 92.3% |
| Overall | 66 | 11 | 55 | 83.3% |

Disease specific survival

| | Total | DSS | Censored | | Survival rate (%) | | | p for log rank |
|-----------|-------|-----|----------|-------|-------------------|-------|-------|----------------|
| | | | n | % | 1y | 3y | 5y | |
| Treatment | | | | | | | | 0.041 |
| CT alone | 27 | 8 | 19 | 70.4% | 92.6% | 81.2% | 69.6% | |
| Sandwich | 39 | 3 | 36 | 92.3% | 100.0% | 97.4% | 91.8% | |

Survival Table

| group | Time | Status | Cumulative Proportion Surviving at the Time | | N of Cumulative Events | N of Remaining Cases | |
|----------|------|--------|---|------------|------------------------|----------------------|----|
| | | | Estimate | Std. Error | | | |
| CT alone | 1 | 5.322 | DSS | .963 | .036 | 1 | 26 |
| | 2 | 9.265 | DSS | .926 | .050 | 2 | 25 |
| | 3 | 15.704 | DSS | .889 | .060 | 3 | 24 |
| | 4 | 22.472 | no-DSS | . | . | 3 | 23 |
| | 5 | 25.889 | DSS | .850 | .069 | 4 | 22 |
| | 6 | 35.187 | DSS | .812 | .076 | 5 | 21 |
| | 7 | 36.862 | DSS | .773 | .082 | 6 | 20 |
| | 8 | 37.191 | DSS | .734 | .086 | 7 | 19 |
| | 9 | 38.604 | DSS | .696 | .090 | 8 | 18 |
| | 10 | 63.639 | no-DSS | . | . | 8 | 17 |
| | 11 | 64.526 | no-DSS | . | . | 8 | 16 |
| | 12 | 64.657 | no-DSS | . | . | 8 | 15 |
| | 13 | 72.279 | no-DSS | . | . | 8 | 14 |
| | 14 | 85.290 | no-DSS | . | . | 8 | 13 |
| | 15 | 89.823 | no-DSS | . | . | 8 | 12 |



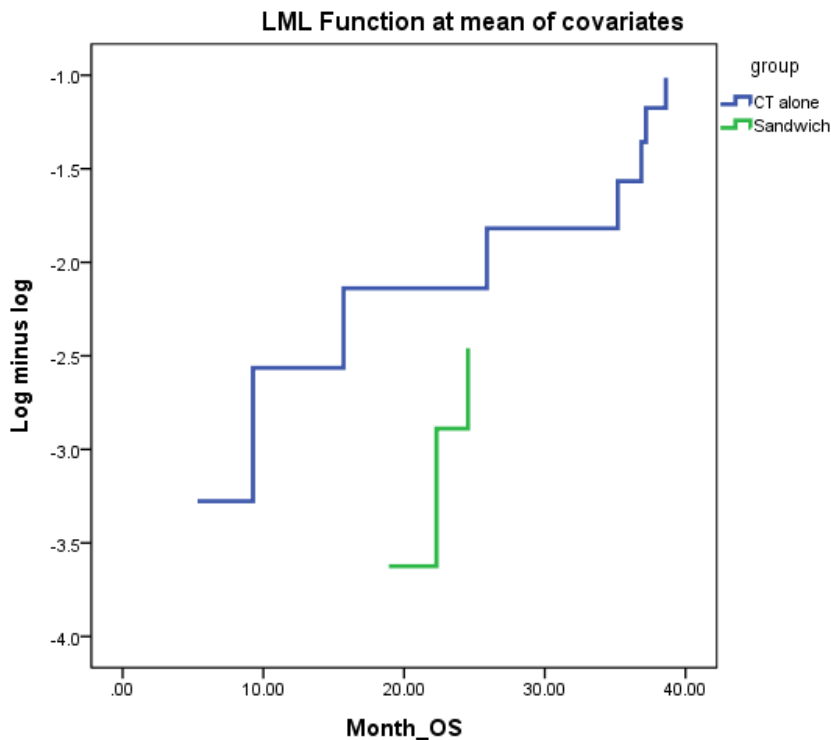
Overall Comparisons

| | Chi-Square | df | Sig. |
|-----------------------|------------|----|------|
| Log Rank (Mantel-Cox) | 4.169 | 1 | .041 |

Test of equality of survival distributions for the different levels of group.

Cox PH assumption

- 透過Cox model LML圖形判定是否符合假設
 - 交叉代表未符合
 - 平行代表符合



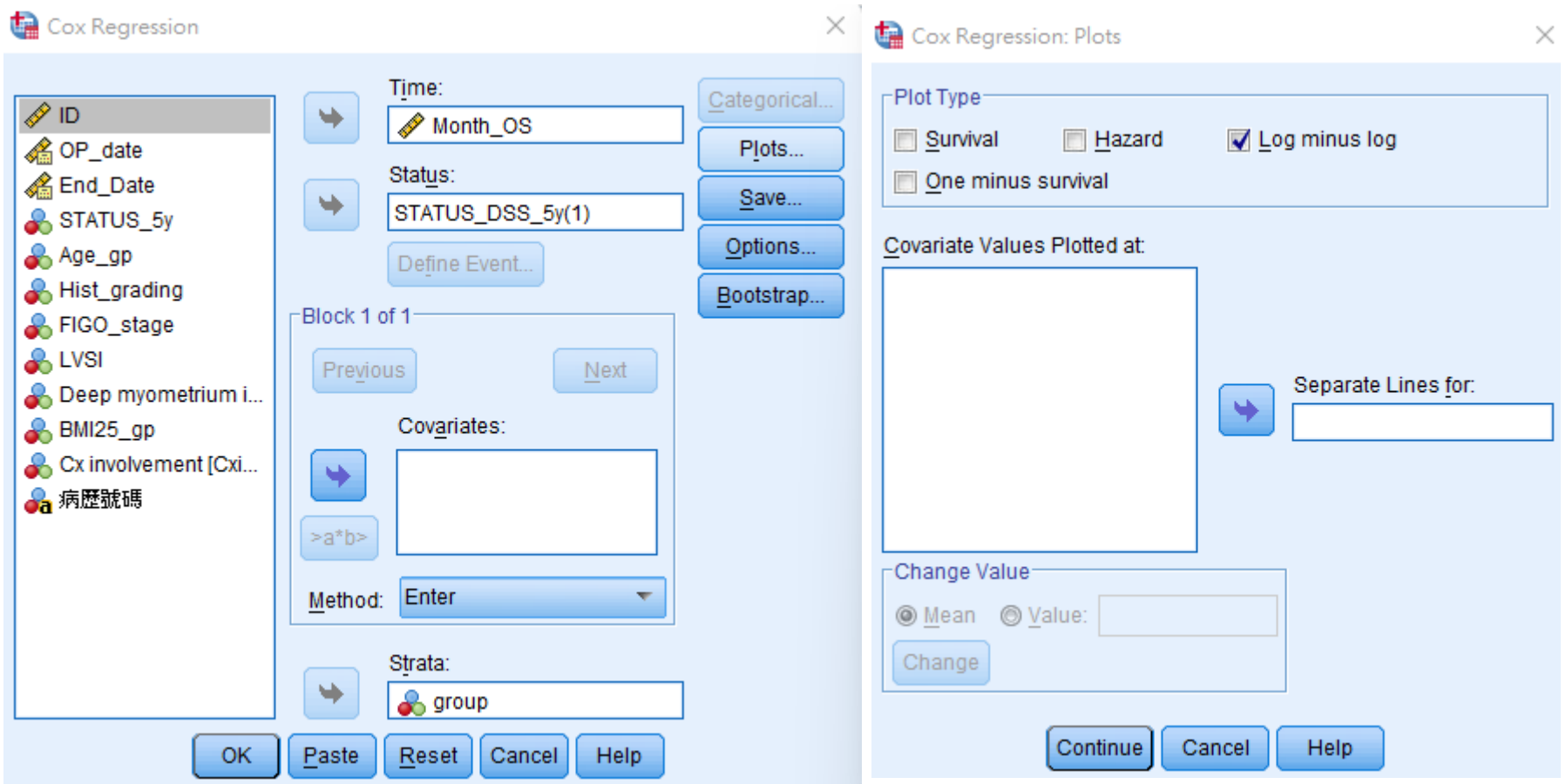
- 在Cox model加入time-dependent變數的交互作用項判定是否符合假設
 - 顯著代表未符合
 - 未顯著代表符合

Variables in the Equation

| | B | SE | Wald | df | Sig. | Exp(B) |
|--------|-------|-------|------|----|------|--------|
| T_COV_ | -.022 | .062 | .120 | 1 | .729 | .979 |
| group | -.786 | 1.581 | .247 | 1 | .619 | .456 |

Cox PH assumption

- Log minus log (LML) 圖形



The image shows two overlapping dialog boxes from the SPSS software interface. The left dialog is the main 'Cox Regression' dialog, and the right dialog is the 'Cox Regression: Plots' sub-dialog.

Cox Regression Dialog (Left):

- Time:** Month_OS
- Status:** STATUS_DSS_5y(1)
- Block 1 of 1:** Includes 'Previous' and 'Next' buttons.
- Covariates:** An empty text box for entering covariate names.
- Method:** Enter
- Strata:** group
- Buttons:** OK, Paste, Reset, Cancel, Help

Cox Regression: Plots Dialog (Right):

- Plot Type:** Survival, Hazard, Log minus log, One minus survival
- Covariate Values Plotted at:** An empty text box.
- Separate Lines for:** A button with a right-pointing arrow and an empty text box.
- Change Value:** Mean, Value: [empty text box], Change button
- Buttons:** Continue, Cancel, Help

Cox PH assumption



- Cox time-dependent變數的交互作用項

The image shows the SPSS Analyze menu with the Survival sub-menu open. The Survival sub-menu includes options like Life Tables..., Kaplan-Meier..., Cox Regression..., and Cox w/ Time-Dep Cov... The Cox w/ Time-Dep Cov... option is highlighted.

| Menu Item | STATUS_5y | STATUS_DSS |
|---------------------------|-----------|------------|
| Reports | | |
| Descriptive Statistics | | |
| Tables | | |
| Compare Means | | |
| General Linear Model | | |
| Generalized Linear Models | 1 | |
| Mixed Models | 1 | |
| Correlate | 1 | |
| Regression | 1 | |
| Loglinear | 0 | |
| Neural Networks | 1 | |
| Classify | 0 | |
| Dimension Reduction | 1 | |
| Scale | 1 | |
| Nonparametric Tests | 1 | |
| Forecasting | 1 | |
| Survival | | |
| Multiple Response | | |
| PS Matching | | |
| Missing Value Analysis... | | |
| Multiple Imputation | | |

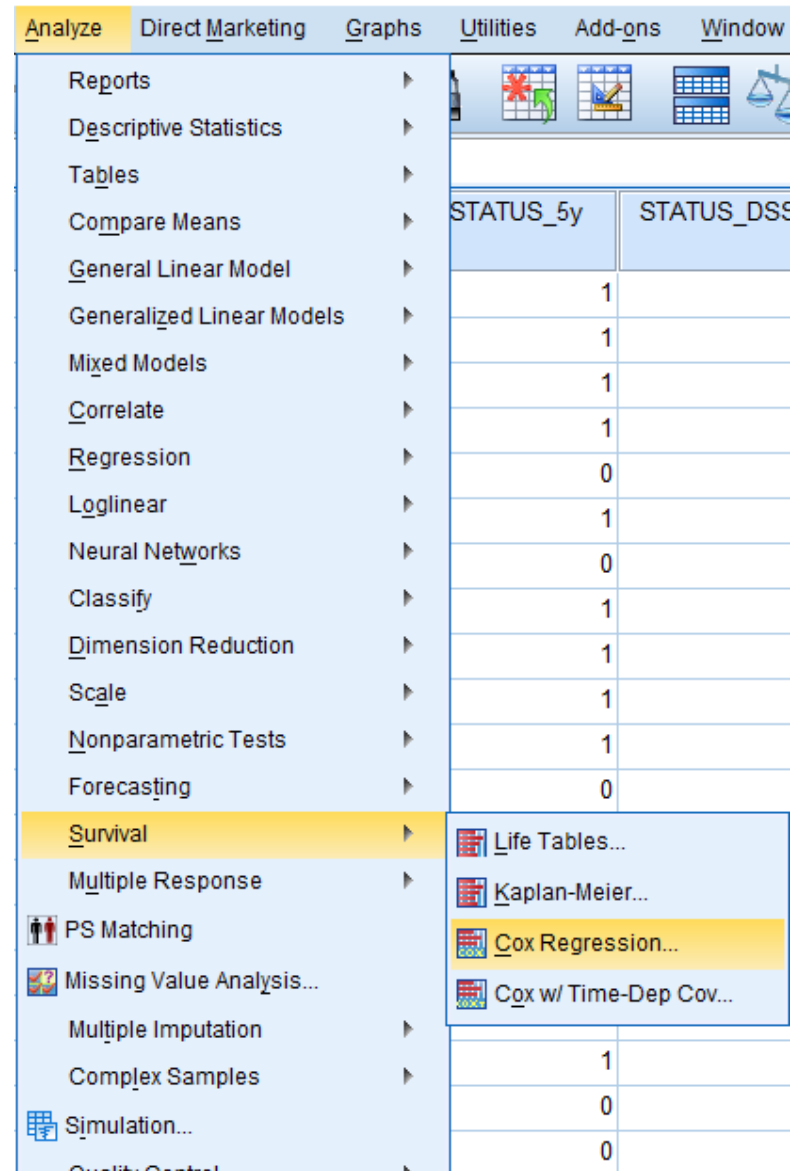
The Compute Time-Dependent Covariate dialog box shows the expression for T_COV_ as T_* group. The Model... button is highlighted with a red box.

The Cox Regression dialog box shows the Time variable as Month_OS and the Status variable as STATUS_DSS_5y(1). The Categorical... button is highlighted with a red box.

The Cox Regression: Define Categorical Covariates dialog box shows the Covariates list containing T_COV_[T_COV_] and the Categorical Covariates list containing group(Indicator(first)). The Contrast is set to Indicator and the Reference Category is set to Last.

Cox model-分析操作

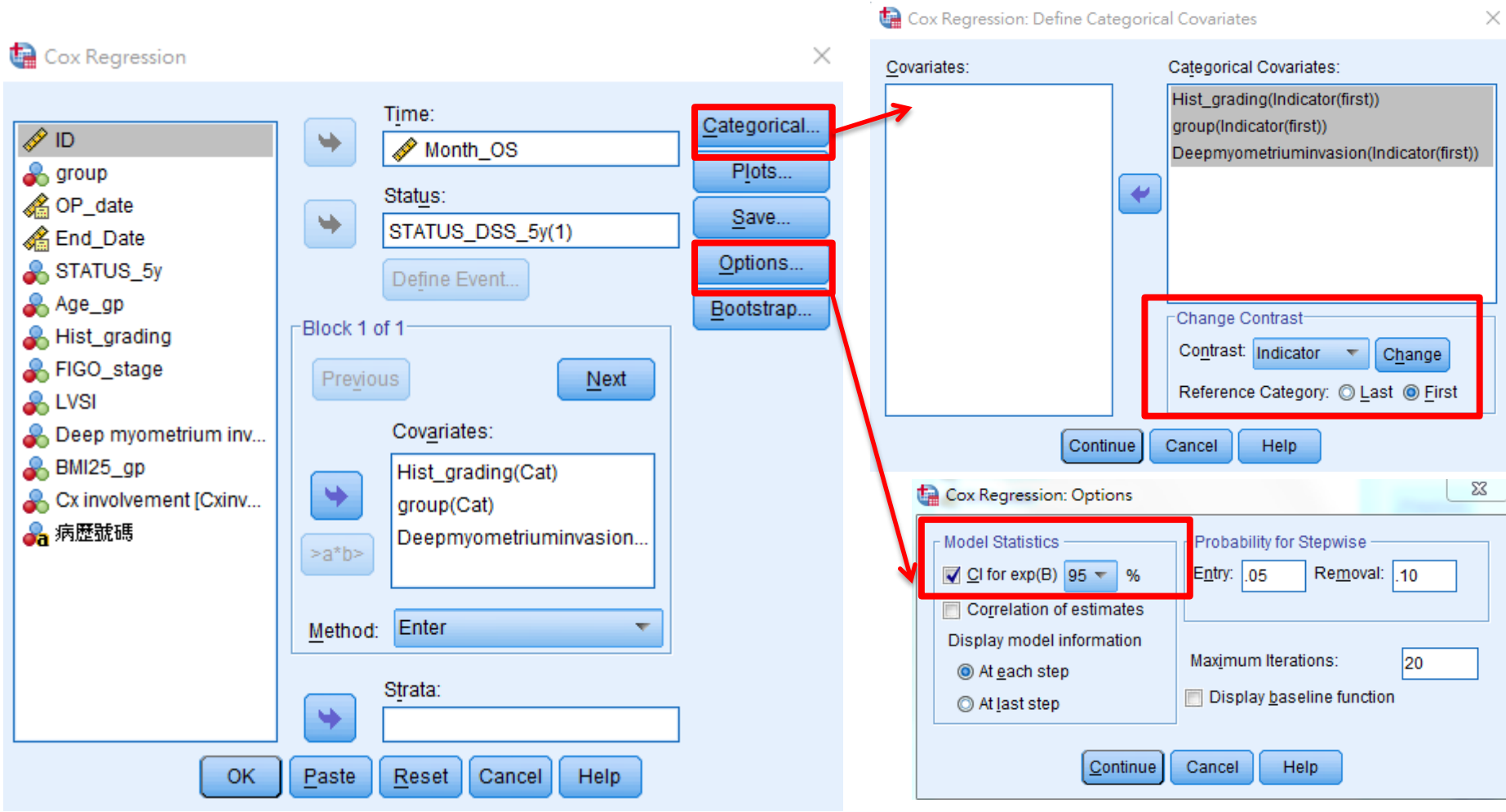
分析>存活分析>Cox迴歸



The screenshot shows the SPSS software interface with the 'Analyze' menu open. The 'Survival' option is selected, and its sub-menu is displayed, with 'Cox Regression...' highlighted. The background shows a data table with columns 'STATUS_5y' and 'STATUS_DSS'.

| | STATUS_5y | STATUS_DSS |
|--|-----------|------------|
| | 1 | |
| | 1 | |
| | 1 | |
| | 1 | |
| | 0 | |
| | 1 | |
| | 0 | |
| | 1 | |
| | 1 | |
| | 1 | |
| | 1 | |
| | 0 | |
| | 1 | |
| | 1 | |
| | 0 | |
| | 1 | |
| | 0 | |
| | 0 | |

Cox model-分析操作



The image shows three overlapping SPSS dialog boxes for Cox Regression analysis, with red boxes and arrows highlighting specific options:

- Cox Regression (Main Dialog):**
 - Time:** Month_OS
 - Status:** STATUS_DSS_5y(1)
 - Block 1 of 1:** Previous, Next
 - Covariates:** Hist_grading(Cat), group(Cat), Deepmyometriuminv...
 - Method:** Enter
 - Strata:** (empty)
 - Buttons:** OK, Paste, Reset, Cancel, Help
- Cox Regression: Define Categorical Covariates:**
 - Covariates:** (empty)
 - Categorical Covariates:** Hist_grading(Indicator(first)), group(Indicator(first)), Deepmyometriuminv...
 - Change Contrast:** Contrast: Indicator, Reference Category: Last, First
 - Buttons:** Continue, Cancel, Help
- Cox Regression: Options:**
 - Model Statistics:** CI for exp(B) 95 %
 - Probability for Stepwise:** Entry: .05, Removal: .10
 - Display model information:** At each step, At last step
 - Maximum Iterations:** 20
 - Display baseline function:**
 - Buttons:** Continue, Cancel, Help

Red boxes highlight the **Categorical...** button in the main dialog, the **Options...** button in the main dialog, the **Change Contrast** section in the Define Categorical Covariates dialog, and the **Model Statistics** section in the Options dialog. Red arrows point from the **Categorical...** button to the Define Categorical Covariates dialog and from the **Options...** button to the Options dialog.

Cox model-Output

Categorical Variable Codings^{a,c,d}

| | | Frequency | (1) |
|---|-----------------|-----------|-----|
| group ^b | 0=CT alone | 27 | 0 |
| | 1=Sandwich | 39 | 1 |
| Hist_grading ^b | 1=Grade 1 and 2 | 29 | 0 |
| | 2=Grade 3 | 37 | 1 |
| Deepmyometriuminvasio n ^b | 0=Absent | 27 | 0 |
| | 1=Present | 39 | 1 |

a. Category variable: group

b. Indicator Parameter Coding

c. Category variable: Hist_grading

d. Category variable: Deepmyometriuminvasion (Deep myometrium invasion)

經多變項調整後, Sandwich相較於CT alone降低0.23倍的風險死於子宮內膜癌且有統計差異(p= 0.037)

Variables in the Equation

| | B | SE | Wald | df | Sig. | Exp(B) | 95.0% CI for Exp(B) | |
|----------------------------|--------|-------|-------|----|------|--------|---------------------|--------|
| | | | | | | | Lower | Upper |
| Hist_grading | 2.215 | 1.050 | 4.452 | 1 | .035 | 9.161 | 1.170 | 71.701 |
| group | -1.474 | .679 | 4.711 | 1 | .030 | .229 | .060 | .867 |
| Deepmyometriuminvasio n | 2.245 | 1.051 | 4.559 | 1 | .033 | 9.442 | 1.202 | 74.150 |

Cox Model v.s Logistic Model

Cox Regression

- 依變項為時間×事件
- 加入追蹤時間及設限資料
- 使用Hazard Ratio (HR)估計

Logistic Regression

- 依變項為類別變數(0 or 1)
- 沒有時間變項
- 使用Odds Ratio (OR)估計



臺中榮民總醫院
Taichung Veterans General Hospital



中榮官網



中榮FB



感謝您的聆聽！

Thank you !

