Diagnostic test

醫學研究部　生物統計小組
徐倩儀
2023/5/23

內容大綱

- Clinical trial & Diagnostic test
  - Bland–Altman plot
- Hosmer and Lemeshow test (goodness of fit)
- Receiver Operating Characteristic (ROC) Curve Analysis

Clinical trial & Diagnostic test

- 體外診斷醫療器材
  - 取人體之檢體，診斷疾病或評估健康狀態
- 臨床診斷效能試驗
  - 量化測定方式
  - 評估每一個受試者之檢體與真實情況，或與其他已核准之對照方法比對臨床效能的研究

Clinical trial & Diagnostic test

- 評估指標
  - **目標族群**中抽樣採集足夠檢體
  - 新器材判讀結果與**黃金標準方法**做比較
- 黃金標準方法
  - 目前公認最精確的診斷方式
- 非黃金標準(non-gold standard)
  - 相同預期檢測標的且經核准之產品

Clinical trial & Diagnostic test

- 臨床效能
  - Sensitivity (敏感度)
    - 醫療器材偵測受試者**特定條件 (如有病)**的能力
  - Specificity (特異度)
    - 醫療器材偵測受試者**不屬於特定條件 (如沒有病)**的能力
- 檢測判讀
  - 目標族群中受試者屬於或不屬於某特定條件
  - 體外診斷醫療器材與所選擇的對照器材之比較
    - 2x2列聯表

Clinical trial & Diagnostic test

- 2×2列聯表(與黃金標準方法比較)
  - Sensitivity (敏感度)、Specificity (特異度)
    - 最常被使用的評估指標
    - 不受盛行率(prevalence)影響

表二、 檢體經新體外診斷醫療器材與黃金標準方法分別判讀的結果

		黃金標準方法		總數
		陽性(+)	陰性(−)	
新體外診斷醫療器材	陽性(+)	44	1	45
	陰性(−)	7	168	175
總數		51	169	220

    - 新體外診斷醫療器材
      - Sensitivity 86.3% (44/51)
      - Specificity 99.4% (168/169)

Clinical trial & Diagnostic test

- 常用術語
  - Type I error (α)
    - 將實驗處理組誤判為有治療效果的機率
    - 偽陽性率(False-positive rate)
  - Type II error (β)
    - 將原本具有治療效果的試驗組誤判為沒有效果的機率
    - 偽陰性率(False-negative rate)
    - 統計檢定力
      - Power = 1 - β

Clinical trial & Diagnostic test

- The agreement between **two quantitative methods** of measurement
  - correlation coefficient (r)
    - the relationship between one variable and another
    - not the differences (not recommended)
  - Bland–Altman plot
    - In 1983 Altman and Bland (B&A) proposed
    - **Mean difference** and **constructing limits of agreement**

Clinical trial & Diagnostic test

- Bland–Altman plot
  - scatter plot
    - Y axis – the difference between the two paired measurements (A-B)
      - Normally distributed (Gaussian) - 95% of differences d-1.96s and d +1.96s
        - not normally distributed - logarithmic transformation
    - X axis – the average of these measures ((A+B)/2)

Clinical trial & Diagnostic test

- Bland–Altman plot

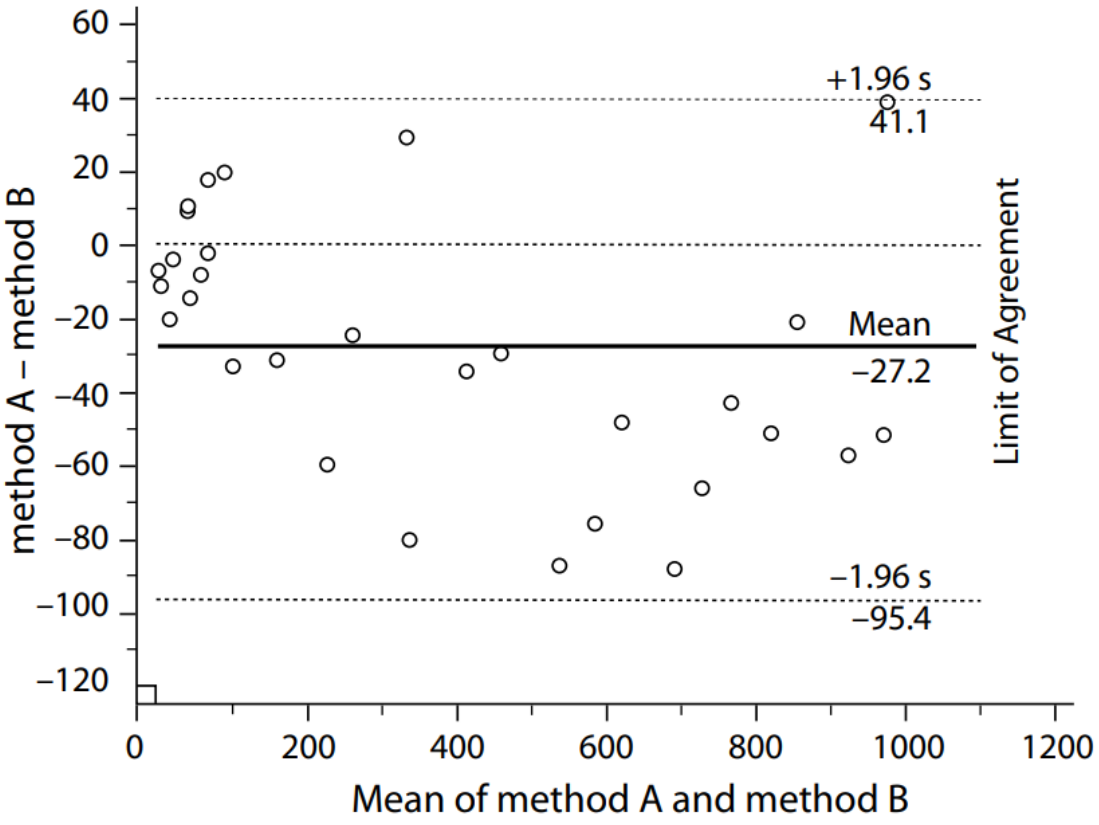
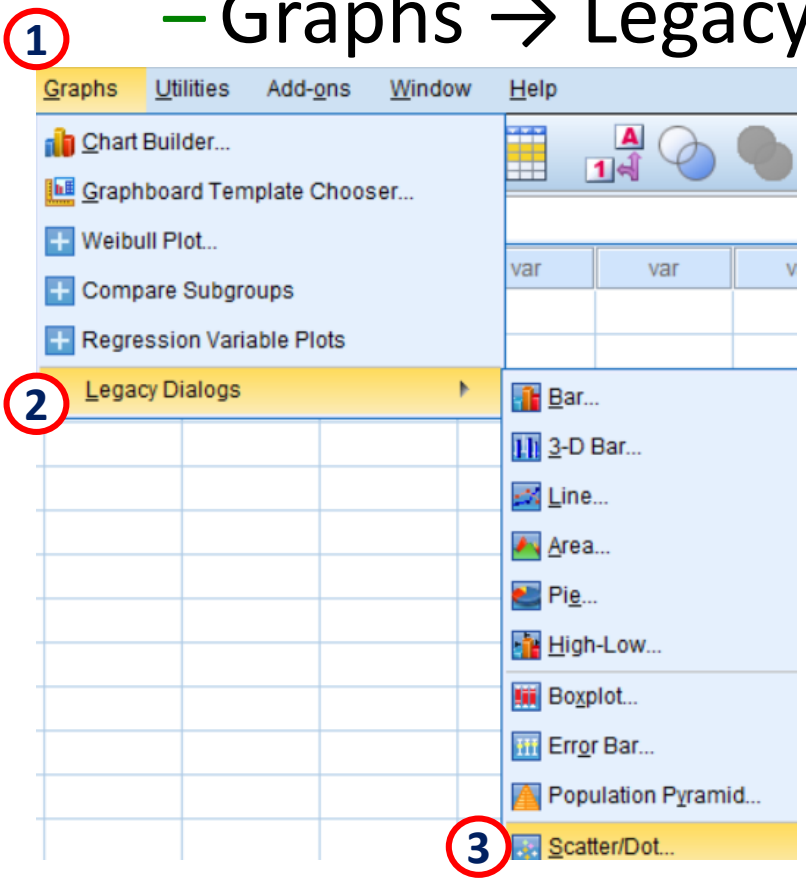


Figure 5. Bland and Altman plot for data from the table 1, with the representation of the limits of agreement (doted line), from -1.96*s* to +1.96*s*.

Bland–Altman plot

- SPSS
  - Graphs → Legacy Dialogs → Scatter/Dot



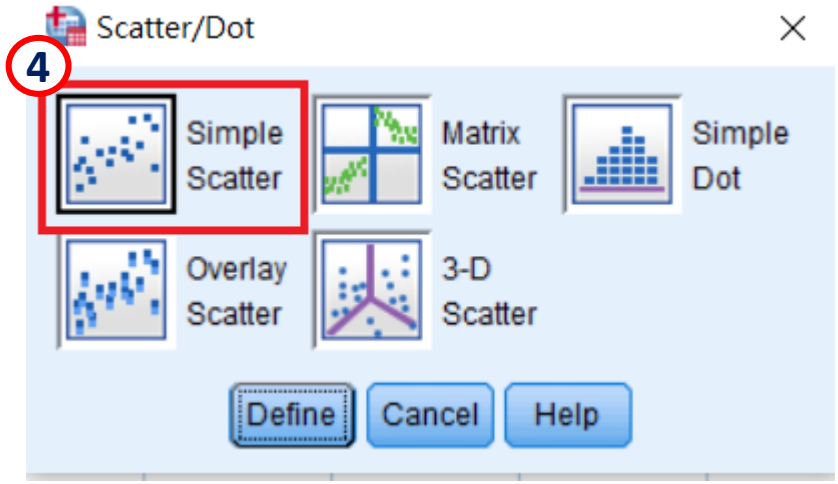


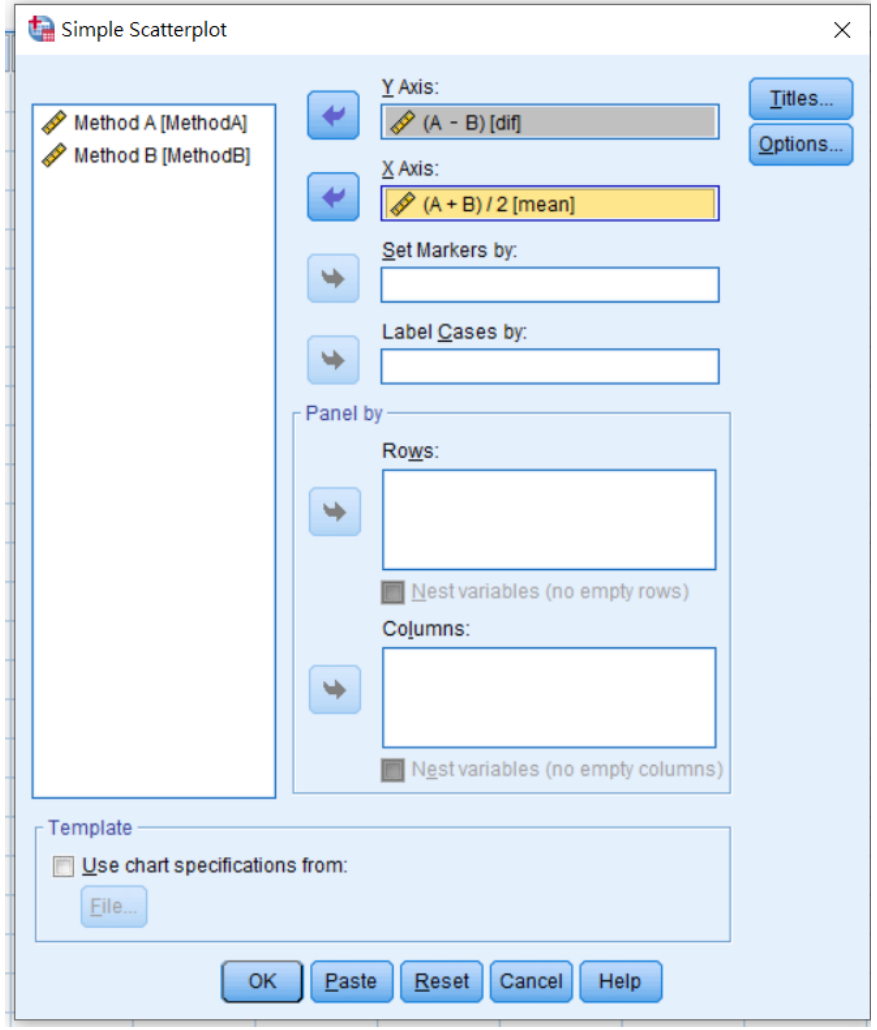
Table 1. Hypothetical data of an agreement between two methods (Method A and B).

Method A (units)	Method B (units)	Mean (A+B)/2 (units)	(A – B) (units)	(A – B)/ Mean (%)
1.0	8.0	4.5	-7.0	-155.6%
5.0	16.0	10.5	-11.0	-104.8%
10.0	30.0	20.0	-20.0	-100.0%
20.0	24.0	22.0	-4.0	-18.2%
50.0	39.0	44.5	11.0	24.7%
40.0	54.0	47.0	-14.0	-29.8%
50.0	40.0	45.0	10.0	22.2%
60.0	68.0	64.0	-8.0	-12.5%
70.0	72.0	71.0	-2.0	-2.8%
80.0	62.0	71.0	18.0	25.4%
90.0	122.0	106.0	-32.0	-30.2%
100.0	80.0	90.0	20.0	22.2%

	MethodA	MethodB	mean	dif
1	1	8	5	-7
2	5	16	11	-11
3	10	30	20	-20
4	20	24	22	-4
5	50	39	45	11
6	40	54	47	-14
7	50	40	45	10
8	60	68	64	-8
9	70	72	71	-2
10	80	62	71	18

Bland–Altman plot

- SPSS



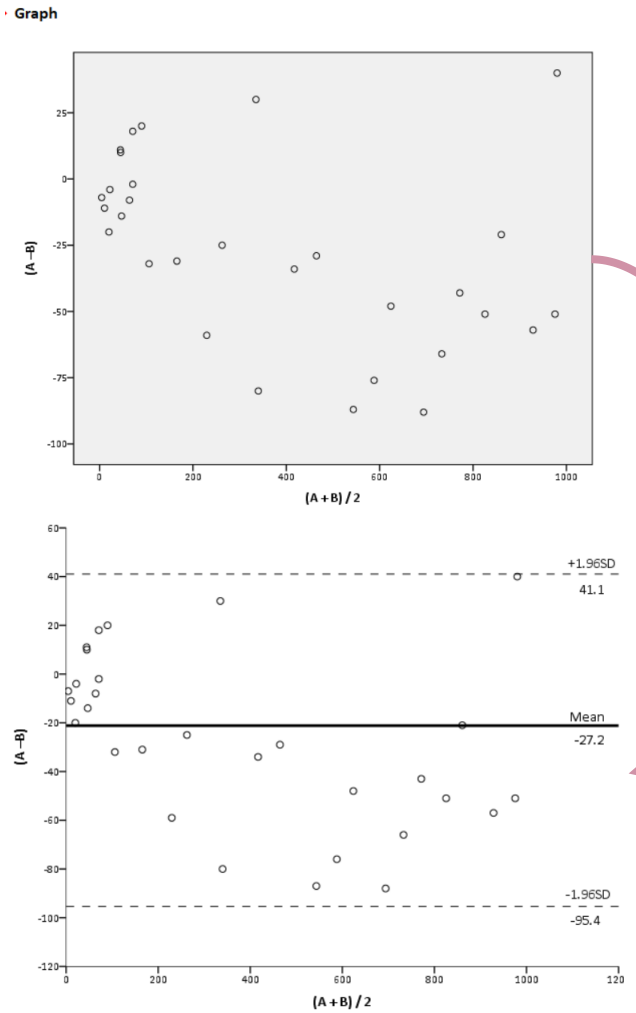
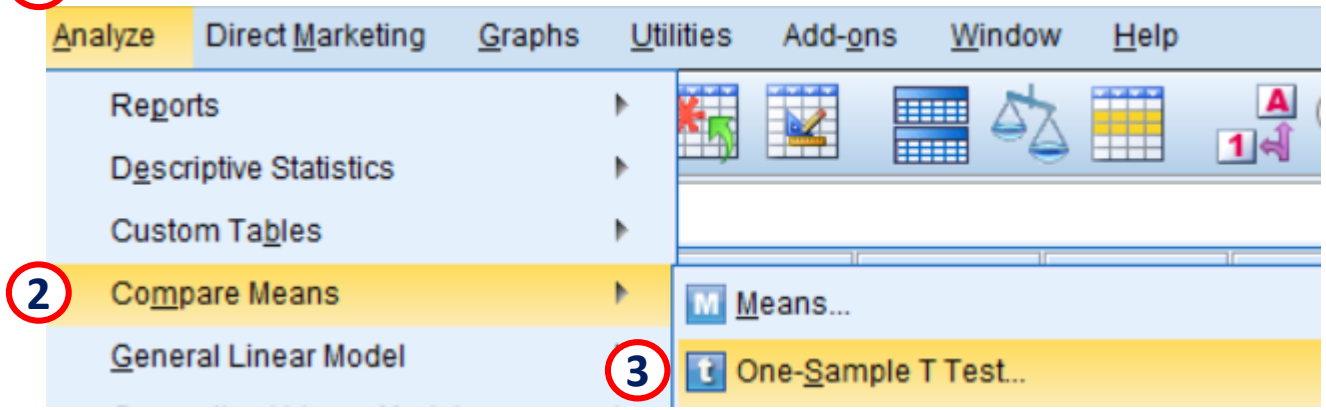


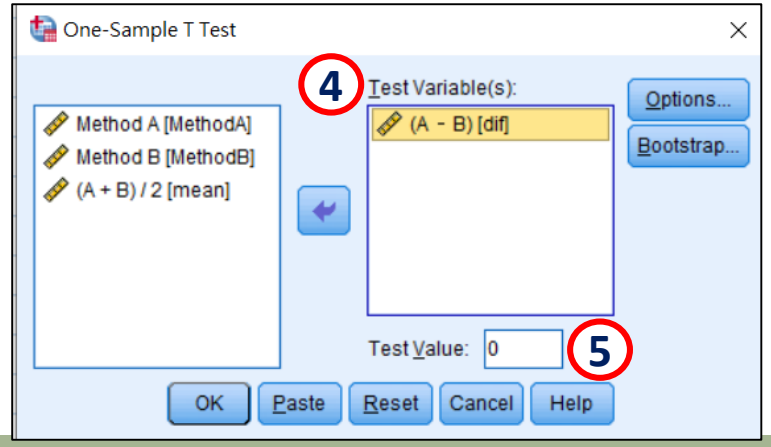
Table 1. Hypothetical data of an agreement between two methods (Method A and B).

Method A (units)	Method B (units)	Mean (A+B)/2 (units)	(A – B) (units)	(A – B)/ Mean (%)
1.0	8.0	4.5	-7.0	-155.6%
5.0	16.0	10.5	-11.0	-104.8%
10.0	30.0	20.0	-20.0	-100.0%
20.0	24.0	22.0	-4.0	-18.2%
50.0	39.0	44.5	11.0	24.7%
40.0	54.0	47.0	-14.0	-29.8%
50.0	40.0	45.0	10.0	22.2%
60.0	68.0	64.0	-8.0	-12.5%
70.0	72.0	71.0	-2.0	-2.8%
80.0	62.0	71.0	18.0	25.4%
90.0	122.0	106.0	-32.0	-30.2%
100.0	80.0	90.0	20.0	22.2%
150.0	181.0	165.5	-31.0	-18.7%
200.0	259.0	229.5	-59.0	-25.7%
250.0	275.0	262.5	-25.0	-9.5%
300.0	380.0	340.0	-80.0	-23.5%
350.0	320.0	335.0	30.0	9.0%
400.0	434.0	417.0	-34.0	-8.2%
450.0	479.0	464.5	-29.0	-6.2%
500.0	587.0	543.5	-87.0	-16.0%
550.0	626.0	588.0	-76.0	-12.9%
600.0	648.0	624.0	-48.0	-7.7%
650.0	738.0	694.0	-88.0	-12.7%
700.0	766.0	733.0	-66.0	-9.0%
750.0	793.0	771.5	-43.0	-5.6%
800.0	851.0	825.5	-51.0	-6.2%
850.0	871.0	860.5	-21.0	-2.4%
900.0	957.0	928.5	-57.0	-6.1%
950.0	1001.0	975.5	-51.0	-5.2%
1000.0	960.0	980.0	40.0	4.1%

Bland–Altman plot

- SPSS
  - Analyze → Compare Means → One-Sample T test





T-Test

One-Sample Statistics

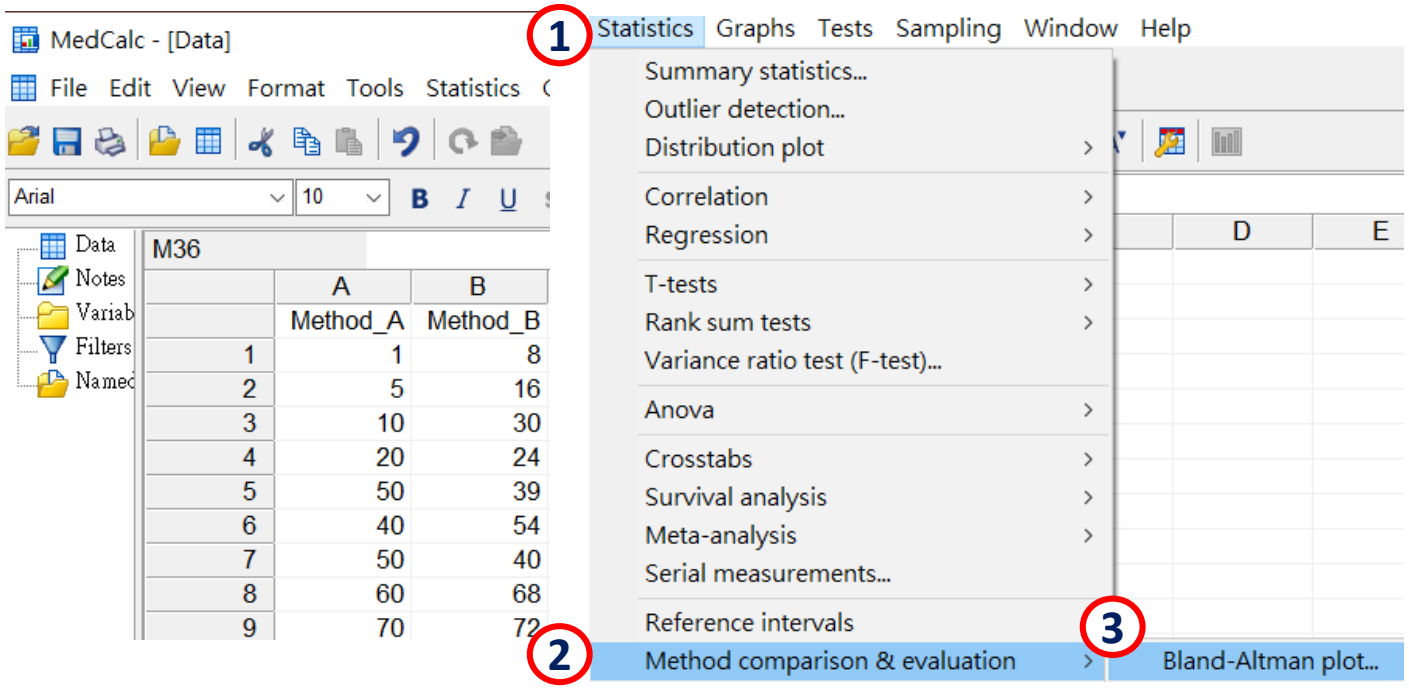
	N	Mean	Std. Deviation	Std. Error Mean
(A - B)	30	-27.17	34.806	6.355

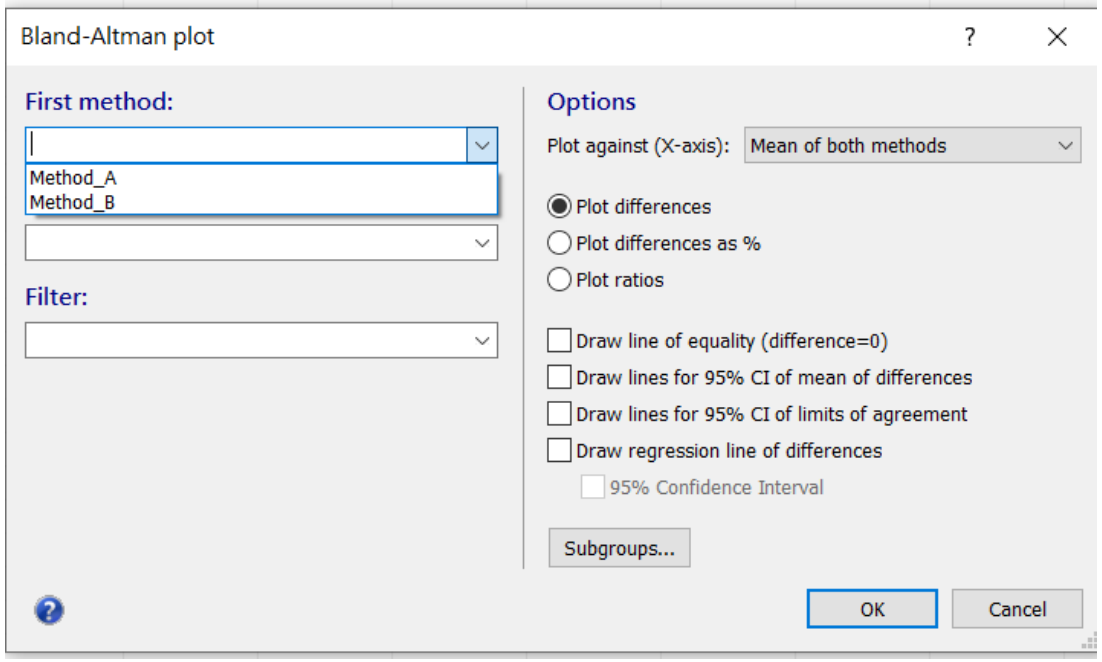
One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
(A - B)	-4.275	29	.000	-27.167	-40.16	-14.17

Bland–Altman plot

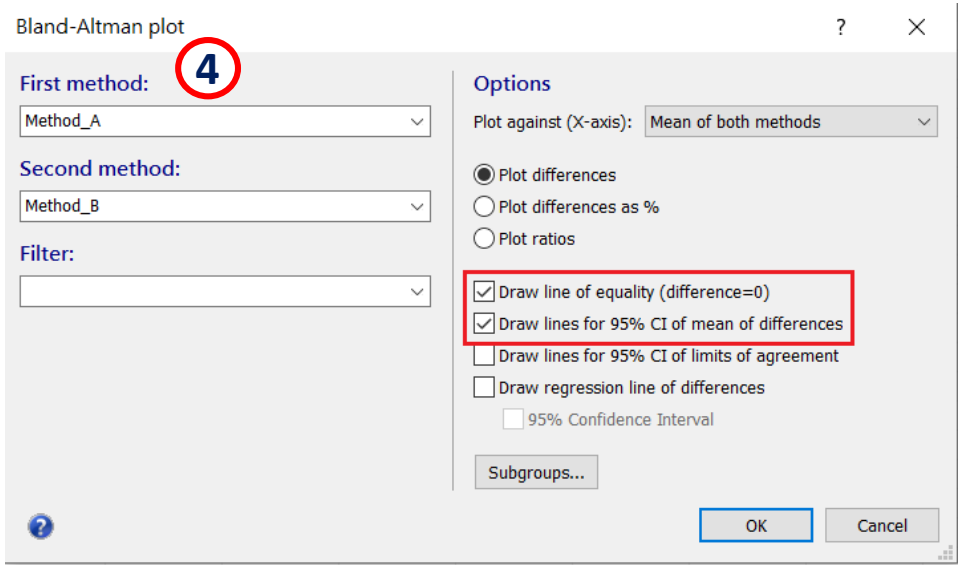
- Medcalc
  - Statistics → Method comparison & evaluation→ Bland–Altman plot

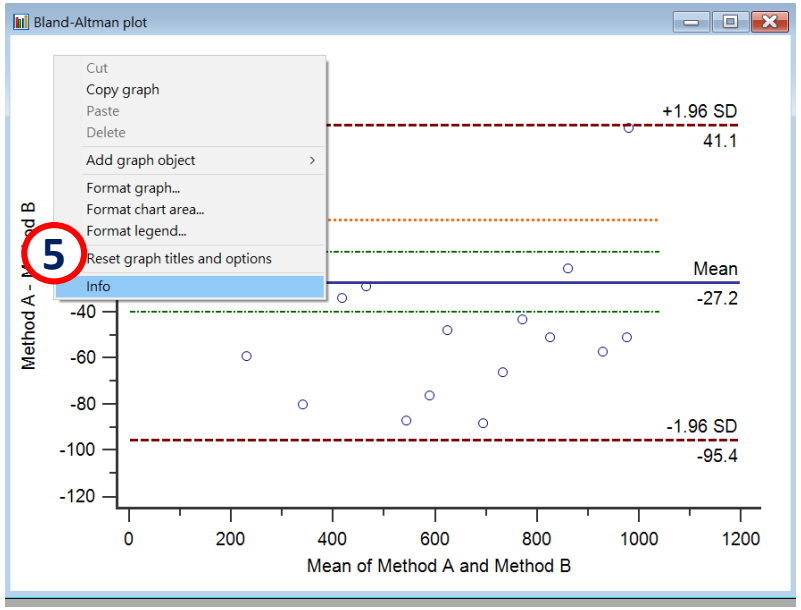




Bland–Altman plot

- Medcalc





Bland-Altman plot

Method A	Method_A Method A
Method B	Method_B Method B

Differences

Sample size	30
Arithmetic mean	-27.1667
95% CI	-40.1634 to -14.1699
P (H_0: Mean=0)	0.0002
Standard deviation	34.8059
Lower limit	-95.3863
95% CI	-117.8487 to -72.9240
Upper limit	41.0530
95% CI	18.5906 to 63.5153

Reference (Clinical trial)

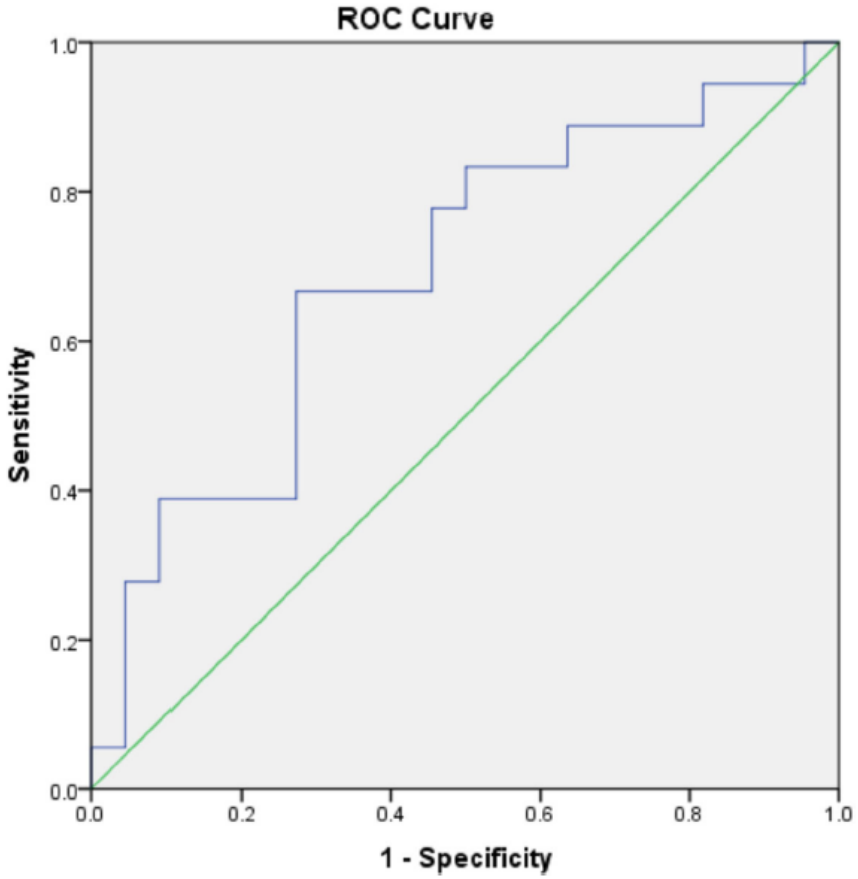


Fig. 1 Receiver Operating Characteristics (ROC) Curve representing the balance between sensitivity (true positive rate) and one minus the specificity (false positive rate) for patients reporting perceived improvements (GRC ≥2) given various cut off points for change in the largest curve

Reference (Clinical trial)

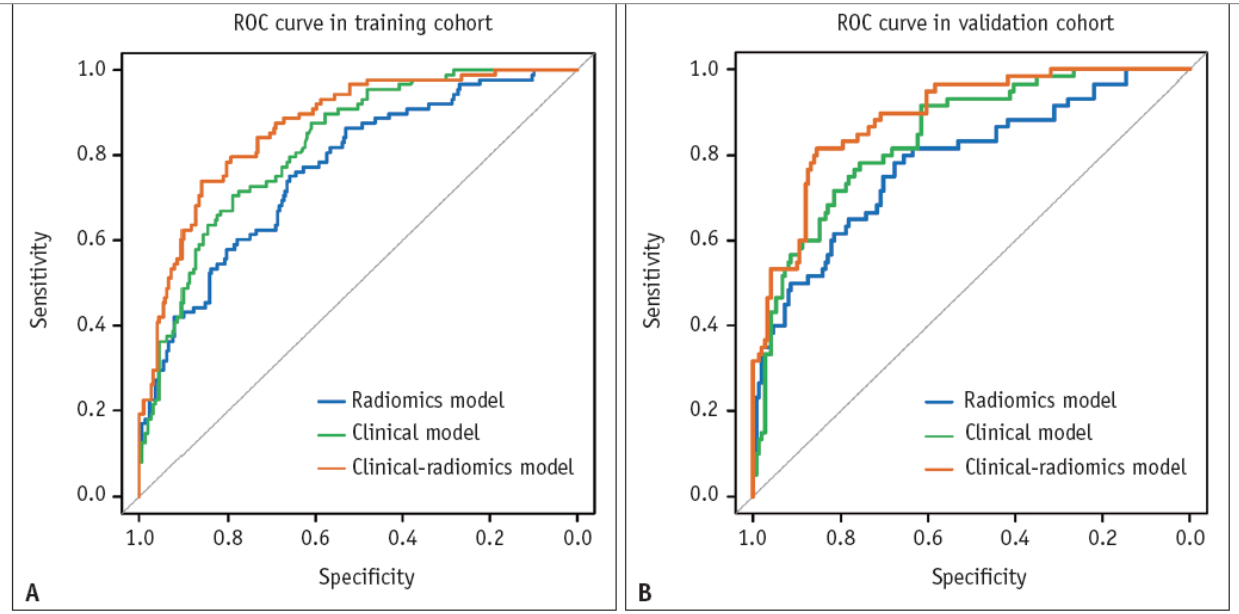


Fig. 2. ROC curves of the radiomics model, clinical model, and clinical-radiomics model in the training (A) and validation (B) cohorts. ROC = receiver operating characteristic

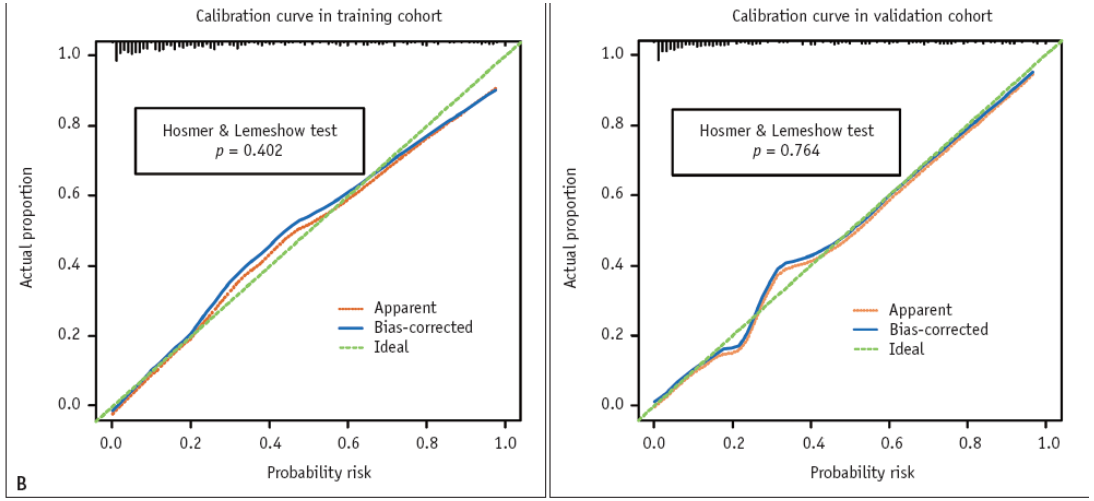
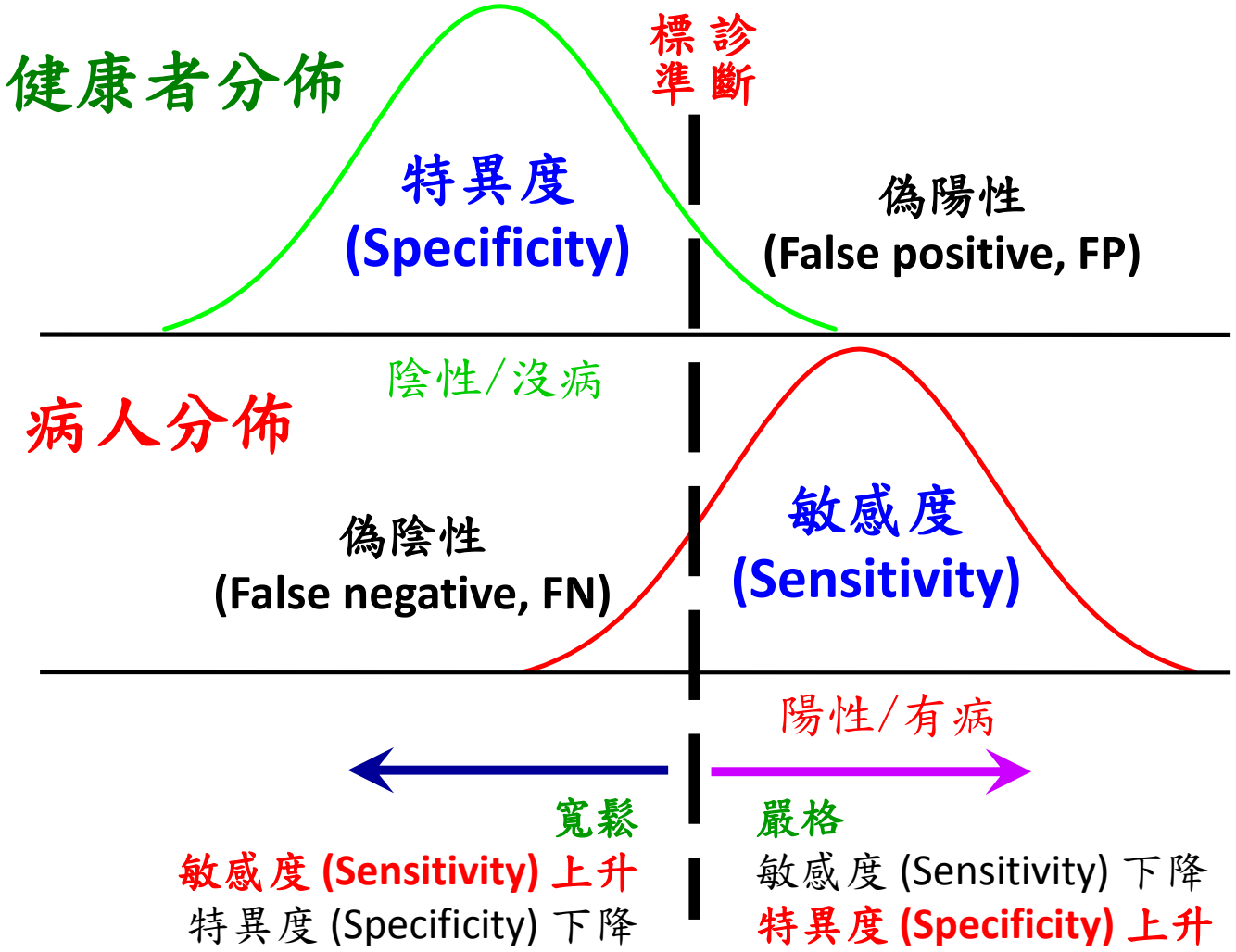


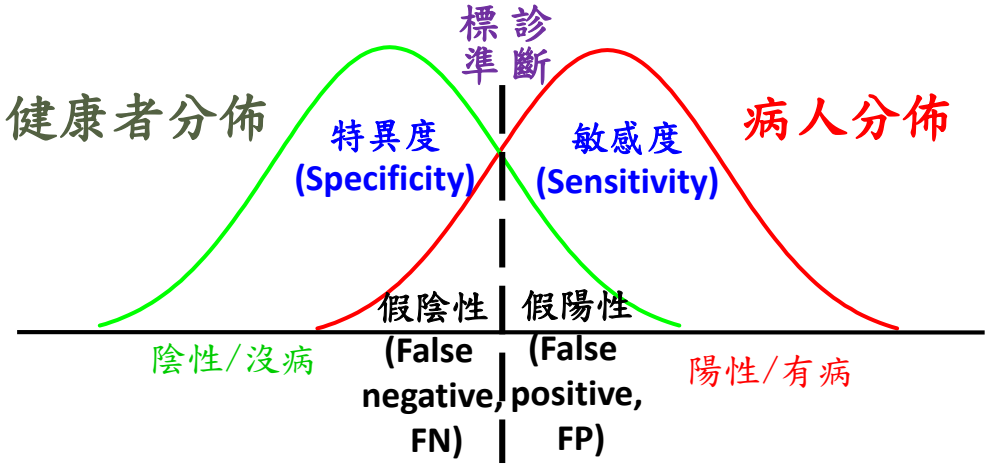
Fig. 3. The clinical-radiomics nomogram for predicting acute ischemic stroke outcomes.
A. The developed nomogram based on the clinical-radiomics prediction model to predict the risk of poor stroke outcome. Diabetes: 0, no diabetes; 1, diabetes. Sex: 0, female; 1, male. Stroke history: 0, no stroke history; 1, stroke history; $mRS_{baseline}$: 0, ≤ 2; 1, > 2. **B.** Calibration curves for the nomogram in the training and validation cohorts. The green dashed line represents the ideal prediction and the red dashed line represents the predictive ability of the nomogram. The closer the dashed red line fit to the dashed green line, the greater the prediction accuracy of the nomogram. **C.** Decision curve analysis for the nomogram. The black line represents the net benefit of assuming no stroke patients have

The predictive performance of the radiomics, clinical, and clinical-radiomics models was evaluated using **receiver operating characteristic (ROC) curves**. The area under the ROC curve **(AUC)** and balanced **sensitivity** and **specificity** at the cutoff yielding the largest **Youden index** value were calculated. The performance of the three models was tested in the training and validation cohorts. The Delong test was used to compare the AUC between the models.
The calibration curve and **Hosmer–Lemeshow test** were used to assess the calibration performance of the clinicalradiomics nomogram.

Korean J Radiol. 2022 Aug;23(8):811-820. Feasibility of a Clinical-Radiomics Model to Predict the Outcomes of Acute Ischemic Stroke. Yiran Zhou, Di Wu, Su Yan, Yan Xie, Shun Zhang , Wenzhi Lv, Yuanyuan Qin, Yufei Liu, Chengxia Liu, Jun Lu, Jia Li, Hongquan Zhu, Weiyin Vivian Liu, Huan Liu, Guiling Zhang, Wenzhen Zhu

Diagnostic test





Diagnostic test

		Gold Standard Test (True class)		
		Disease (Positive)	Non-disease (Negative)	
Screen Test (Hypothesized Class)	Positive (Yes)	a (True Positive)	b (False Positive)	a + b
	Negative (No)	c (False Negative)	d (True Negative)	c + d
		a + c	b + d	a + b + c + d

- Sensitivity (敏感度) = a / (a + c)
- Specificity (特異度) = d / (b + d)
- Positive predictive value, PPV = a / (a + b)
- Negative predictive value, NPV = d / (c + d)
- Accuracy (精確度) = (a + d) / (a + b + c + d)
- False positive, FP (偽陽性) = b / (a + b)
- False negative, FN (偽陰性) = c / (c + d)

Diagnostic test

- Likelihood ratio positive, LR+ = Sensitivity / (1 – Specificity)
- Likelihood ratio negative, LR- = (1 – Sensitivity) / Specificity

概似比的臨床意義	
LR	**Interpretation**
>10	Strong evidence to rule in disease
5-10	Moderate evidence to rule in disease
2-5	Weak evidence to rule in disease
0.5-2.0	No signification change in the likelihood of disease
0.2-0.5	Weak evidence to rule out disease
0.1-0.2	Moderate evidence to rule out disease
< 0.1	Strong evidence to rule out disease

Source: Sackett, Richardson, Rosenberg, Haynes. Evidence-Based Medicine: How to Practice and Teach EBM. Churchill Livingstone, London, 1997.

Diagnostic test

- Positive predictive value, PPV = a / (a + b)
- Negative predictive value, NPV = d / (c + d)

		Gold Standard Test		
		Disease	Non-disease	
Screen Test	+	a	b	a + b
	-	c	d	c + d
		a + c	b + d	a + b + c + d

$$\text{PPV} = \frac{\text{true positive}}{\text{all positive}} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}}$$

$$\text{NPV} = \frac{\text{true negative}}{\text{all negative}} = \frac{\text{true negative}}{\text{true negative} + \text{false negative}}$$

True positive = all cases (a+c) * sensitivity

False positive = all health (b + d) * (1 – specificity)

All cases = total (a + b + c + d)* prevalence

All health = total (a + b + c + d)* (1 – prevalence)

$$\text{PPV} = \frac{(\text{total} * \text{prevalence}) * \text{sensitivity}}{(\text{total} * \text{prevalence}) * \text{sensitivity} + (\text{total} * (1 - \text{prevalence}) * (1 - \text{specificity}))} = \frac{\textbf{prevalence} * \text{sensitivity}}{(\textbf{prevalence} * \text{sensitivity}) + (1 - \textbf{prevalence}) * (1 - \text{specificity})}$$

$$\text{NPV} = \frac{(1 - \textbf{prevalence}) * \text{specificity}}{(1 - \textbf{prevalence}) * \text{specificity} + \textbf{prevalence} * (1 - \text{sensitivity})}$$

- Prevalence ⬆
  - PPV⬆ NPV⬇
- Prevalence ⬇
  - PPV⬇ NPV⬆

Hosmer and Lemeshow test

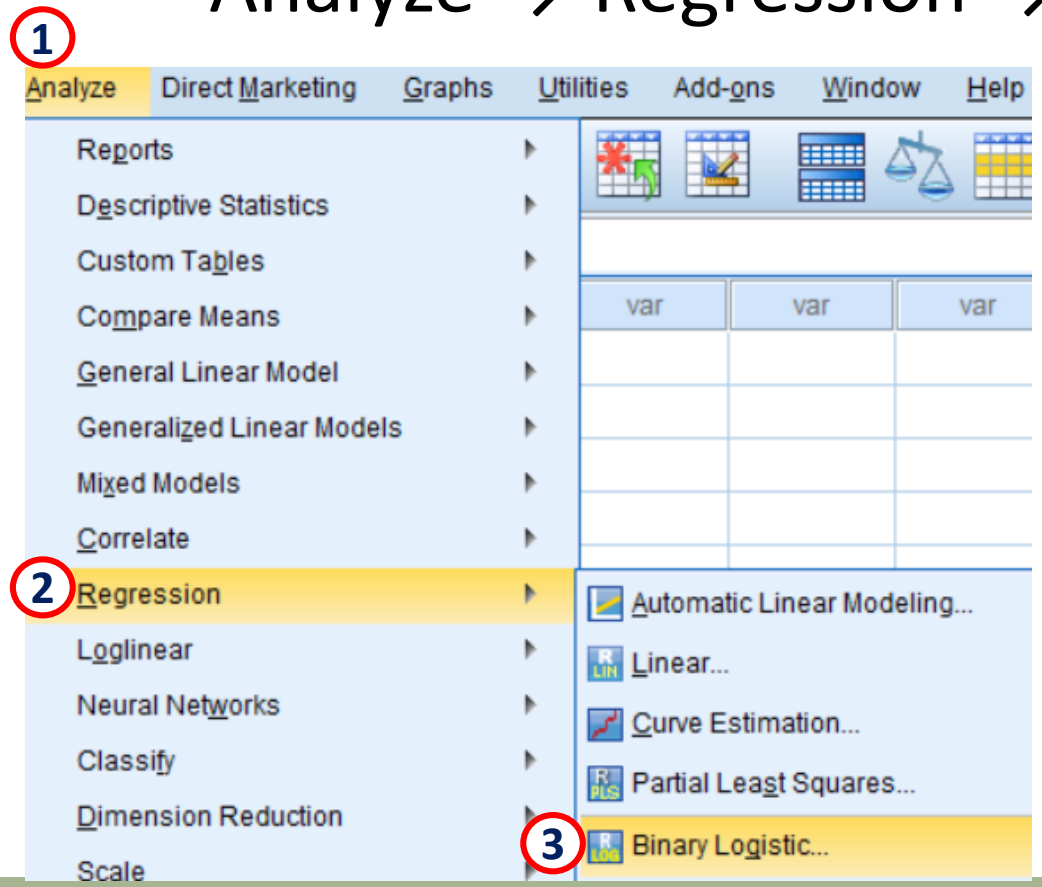
- 對logistic迴歸模型擬合優度的檢定方法
- 根據預測機率值將資料分成大致相同規模的10個組
  - 不管模型中有多少共變類型
- 將觀測資料按其預測機率做升序排列
  - 第一組機率最小
  - 最後一組估計機率最大
- 實際上
  - 很多觀測案例有同樣的預測機率
  - 相同預測機率之觀測案例放在同一組
  - 每組的規模不可能完全相同

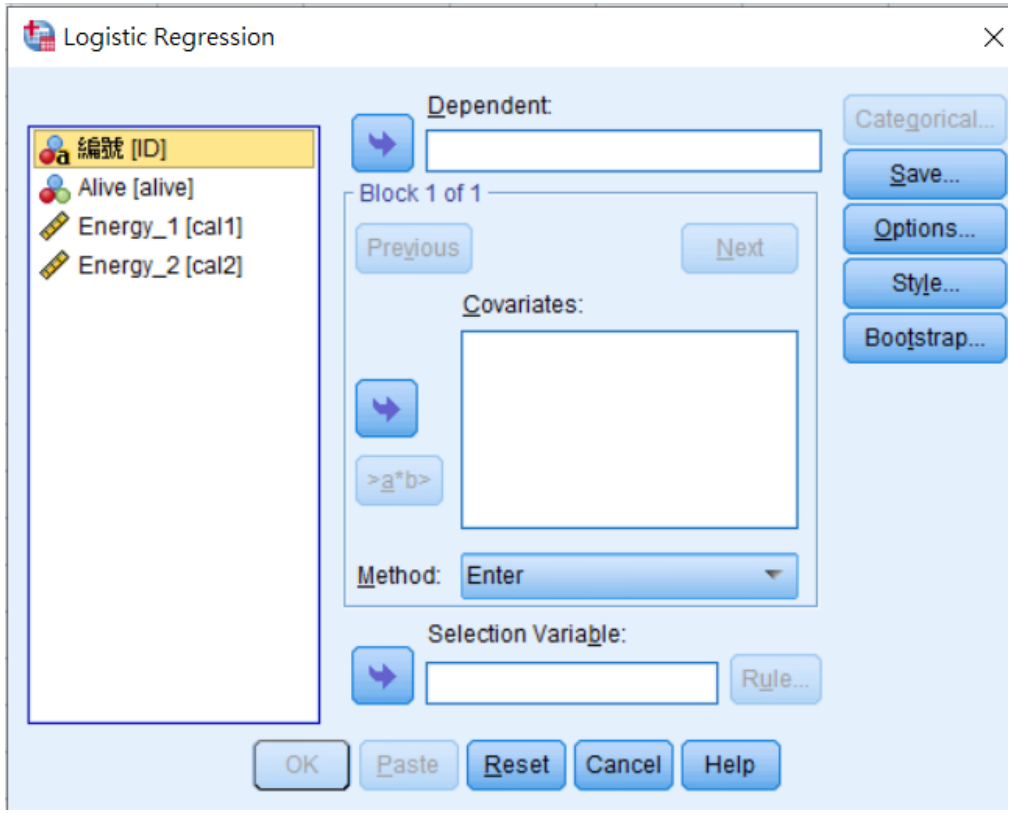
Hosmer and Lemeshow test

- 類似於皮爾遜X^2統計量的指標
- 可從觀測頻數和預測頻數構成的2*G交互表中求得
- HL指標與X^2分佈相比較
  - $p > 0.05$ (Acceptable Calibration)
    - 不能拒絕關於模型擬合資料好的假設
    - 模型很好地擬合資料

Hosmer and Lemeshow test

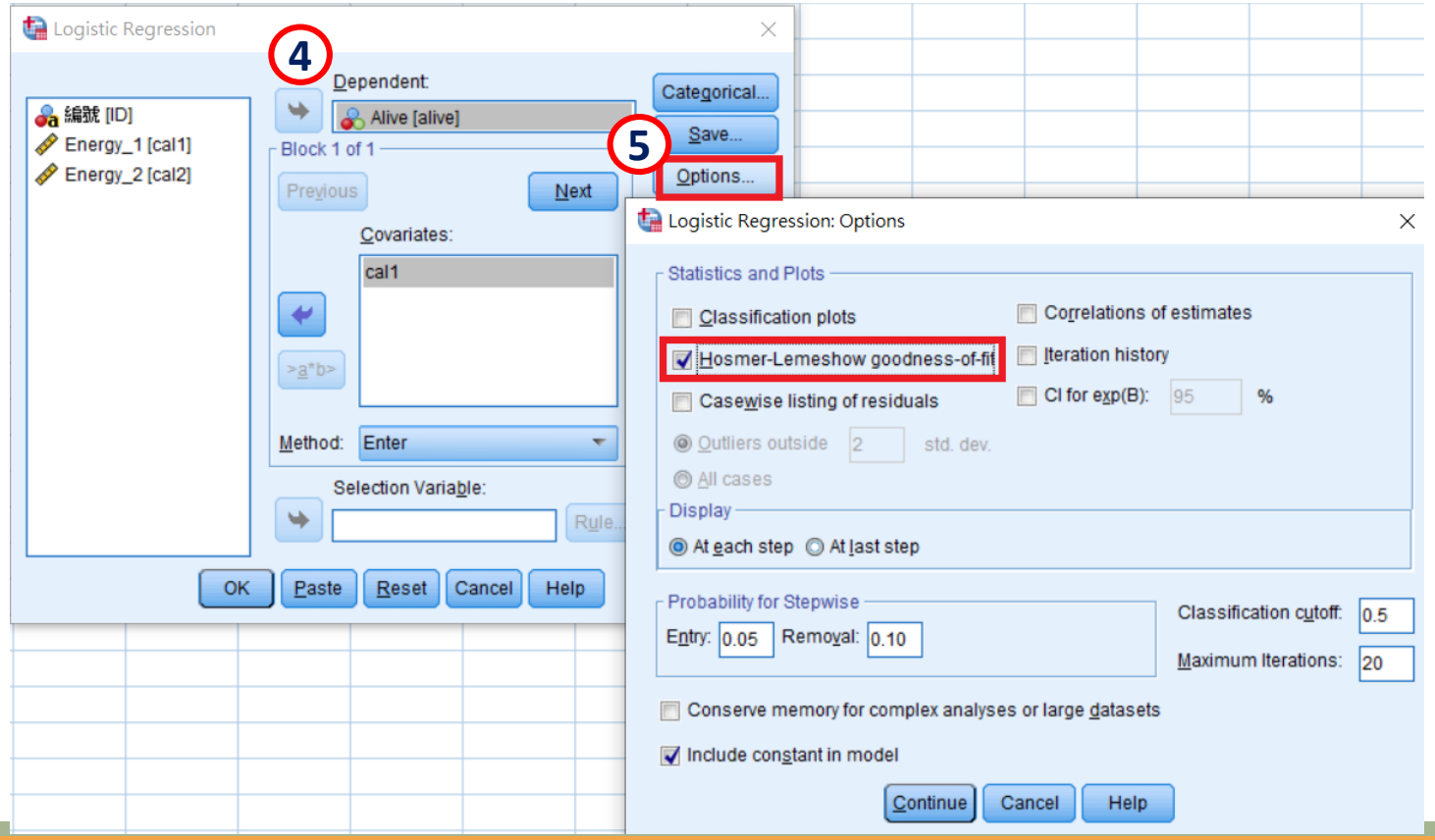
- SPSS
  - Analyze → Regression → Binary Logistic





Hosmer and Lemeshow test

- SPSS
  - Analyze → Regression → Binary Logistic



Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	7.332	7	.395

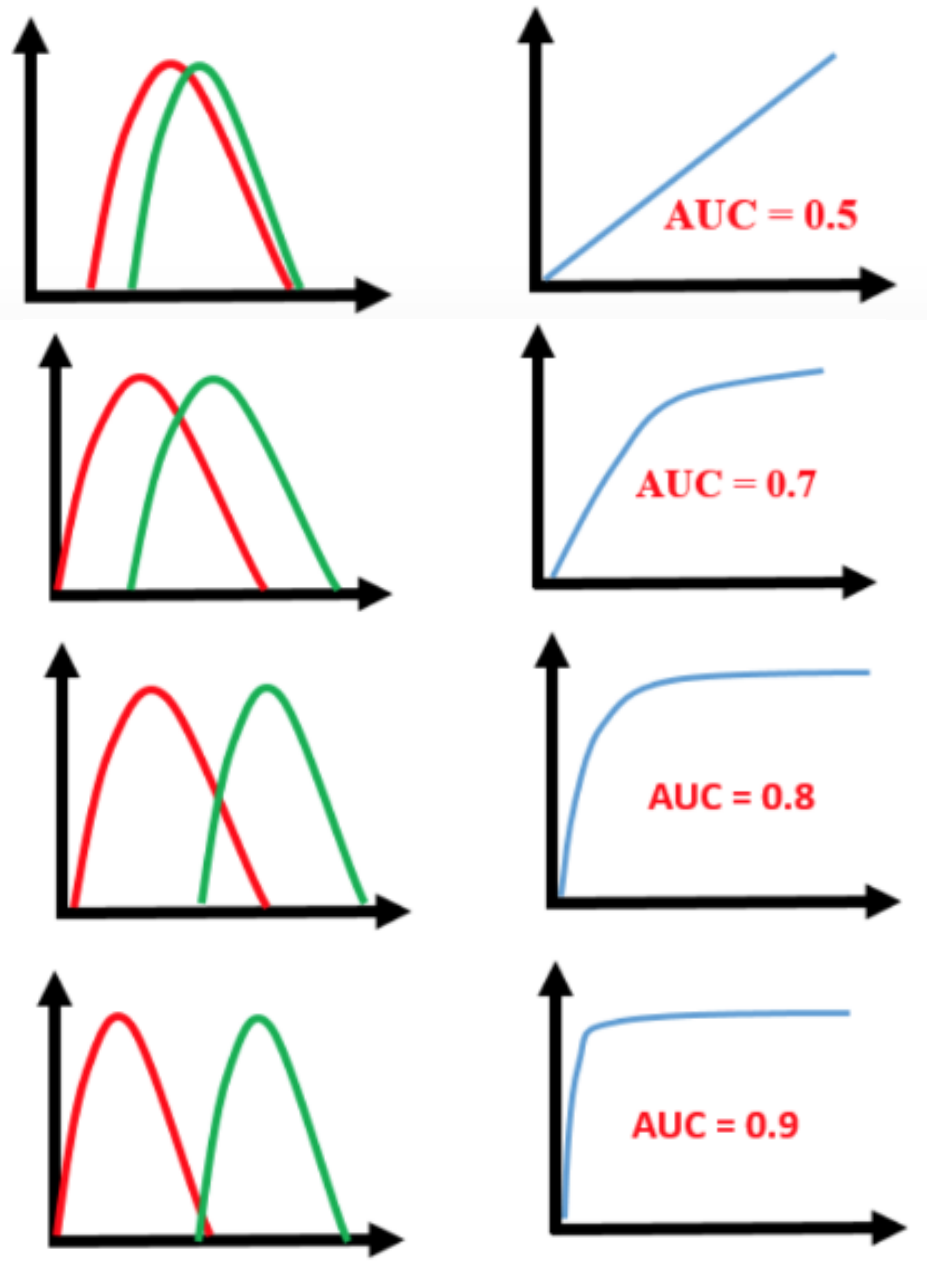
***p* > 0.05**
Acceptable calibration

ROC Curve Analysis

- 使用時機
  - 當開發新的檢驗方法，無法決定臨界值 (Cut-off value)
  - 利用連續數值預測結果 (二元分類)
- 目的
  - 將連續數值決定臨界值
  - 用來比較不同工具的好壞

ROC Curve Analysis

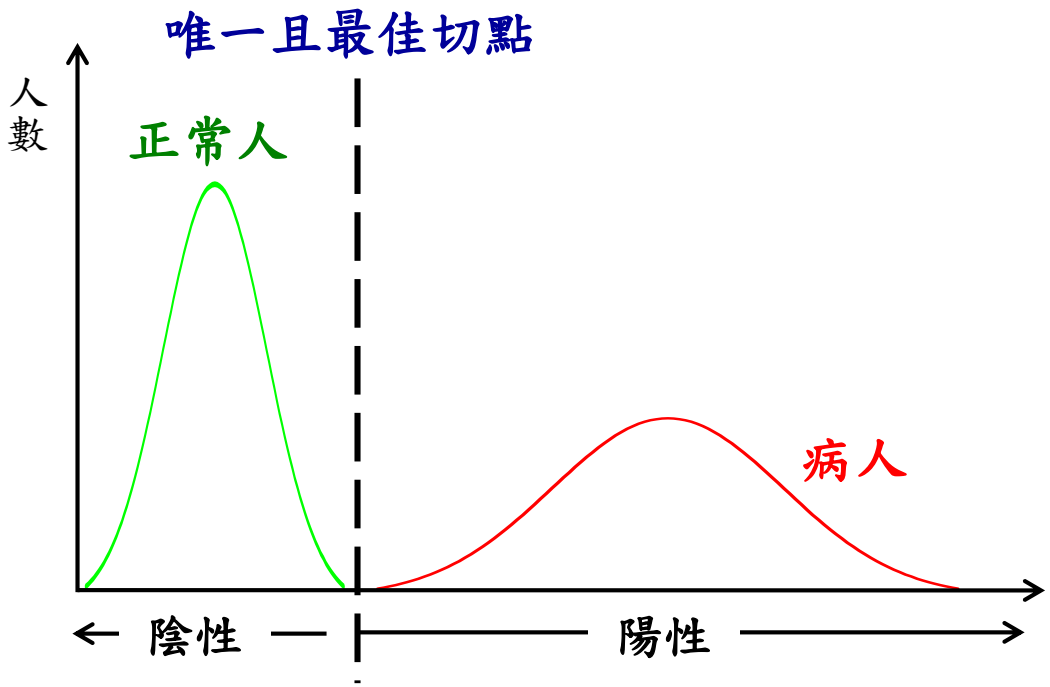
AUC	Discrimination
0.5	No discrimination
0.7-0.8	Acceptable discrimination
0.8-0.9	Excellent discrimination
0.9-1.0	Outstanding discrimination

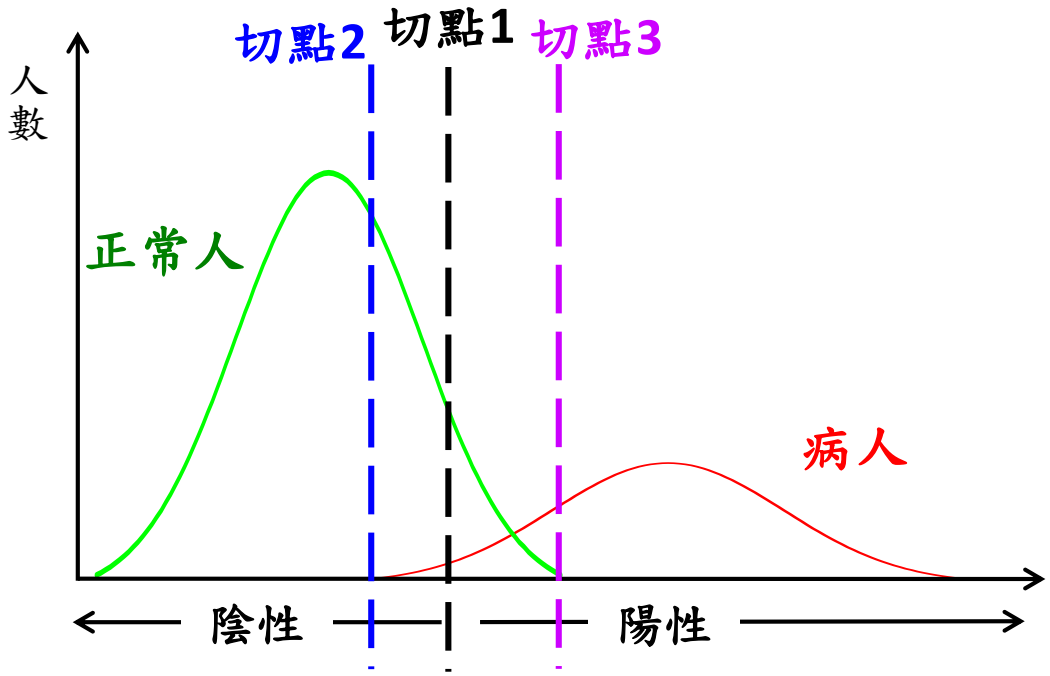


Ref: https://reurl.cc/288Xr9

ROC Curve Analysis

- Cut-point



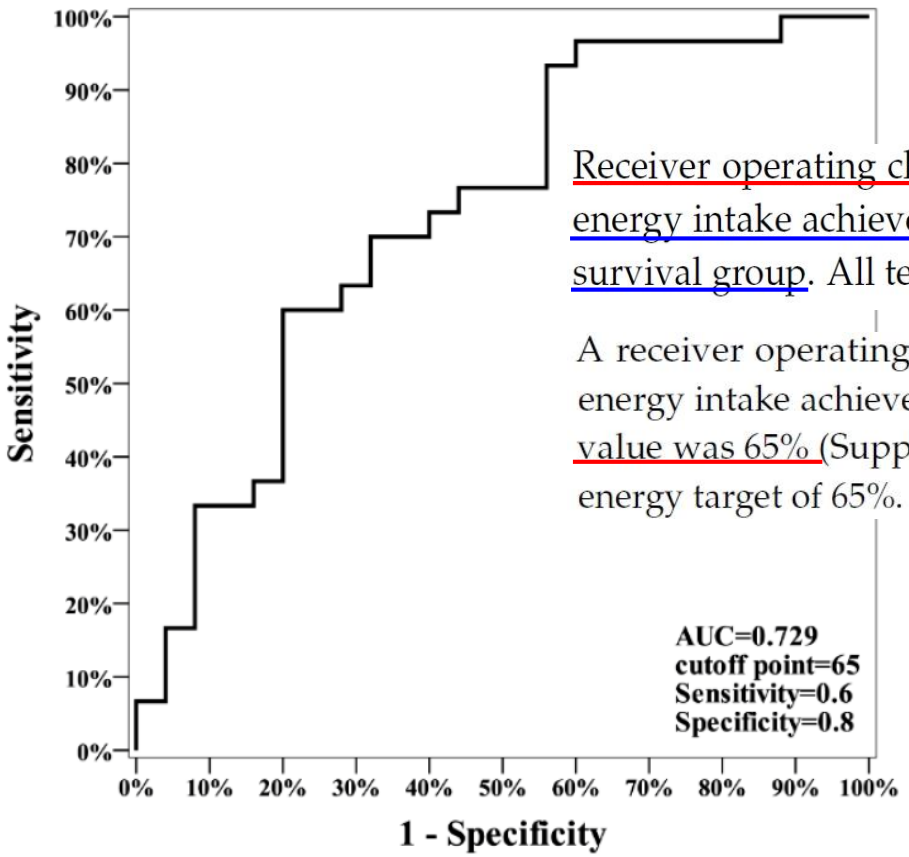


ROC Curve Analysis

- Youden's index
  - 反映在有疾病和沒有疾病的陽性結果可能機率
  - 公式
    - Sensitivity + Specificity - 1
    - Sensitivity - False positive
  - Range 0-1
  - 不受盛行率影響
  - Maximum
    - Cut-point

Reference (ROC curve)

Supplementary Materials



Receiver operating characteristic (ROC) curves were used to evaluate the discriminative ability of energy intake achievement rates on day 3 after the initiation of small bowel feeding to identify the survival group. All tests were two-sided, with $p < 0.05$ considered significant.

A receiver operating characteristic (ROC) curve was used to evaluate the discriminative ability of energy intake achievement rates 3 days after SBEN initiation to identify the survival group; the cutoff value was 65% (Supplement Figure S1). In the survival group, two-thirds of the patients achieved the energy target of 65%.

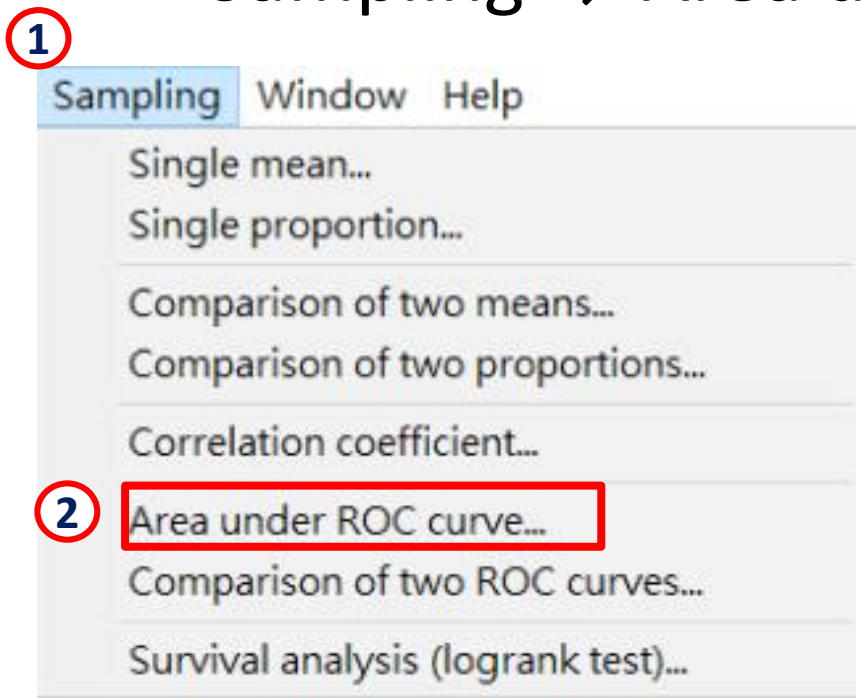
Figure 1. Receiver operating characteristic (ROC) curve to determine the cutoff point for the feeding target between surviving and non-surviving malnourished patients administered SBEN.

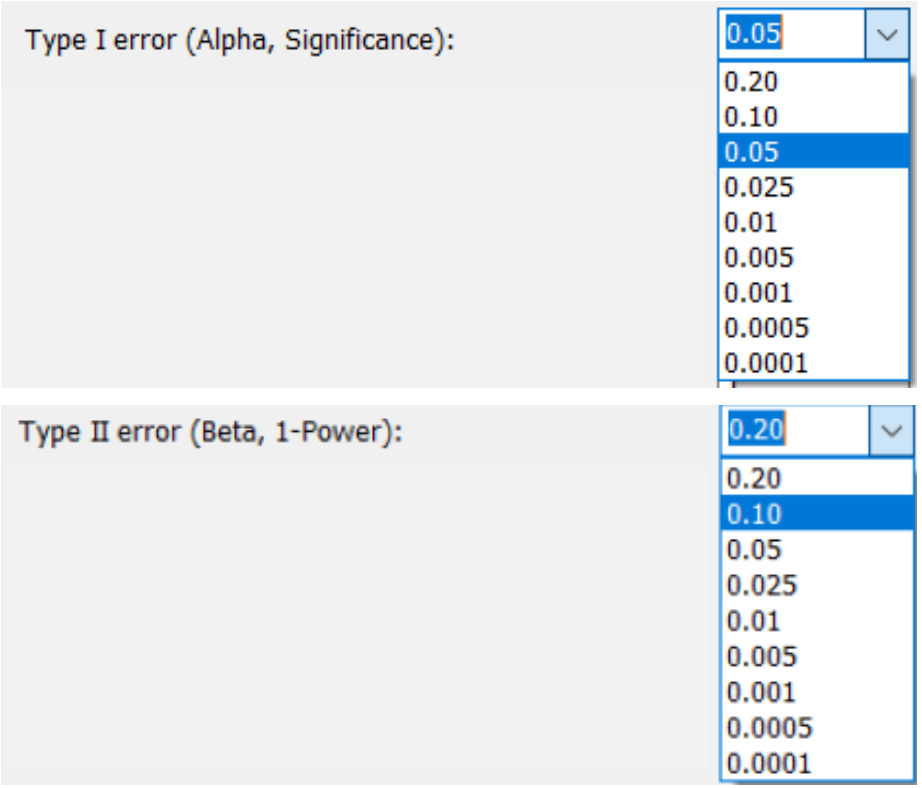
ROC Curve Analysis

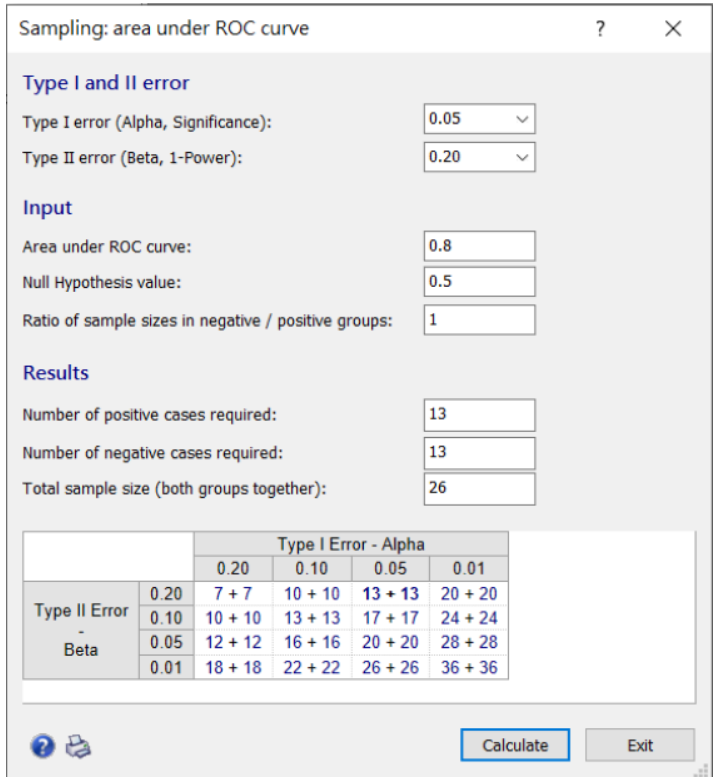
- 樣本量的計算直接影響著臨床試驗的陰性或者陽性
  - 假設：AUC = 0.8 、 α = 0.05、 β = 0.2
  - Medcalc計算樣本數
    - Sampling → Area under ROC curve
      - Type I error (Alpha, Significance) = 0.05
      - Type II error (Beta, 1-Power) = 0.20
      - Area under ROC curve = 0.80
      - Ratio of sample sizes in negative/positive groups = 1

ROC Curve Analysis

- Medcalc計算樣本數
  - Sampling → Area under ROC curve

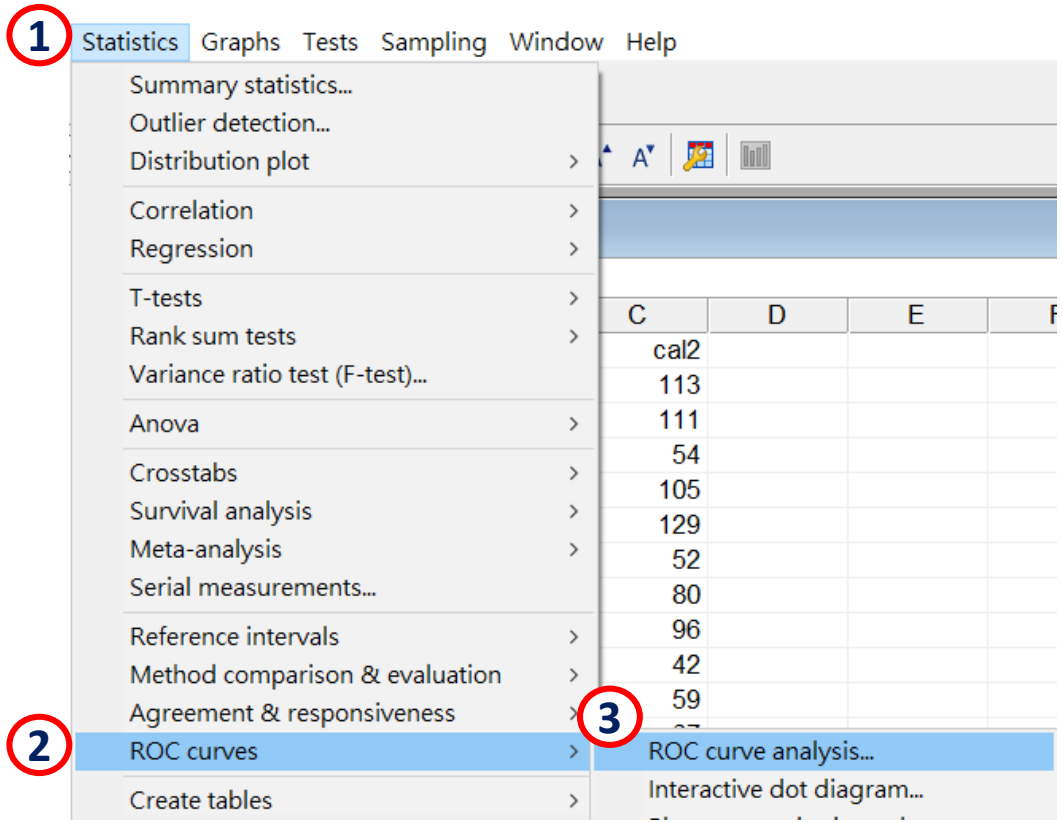


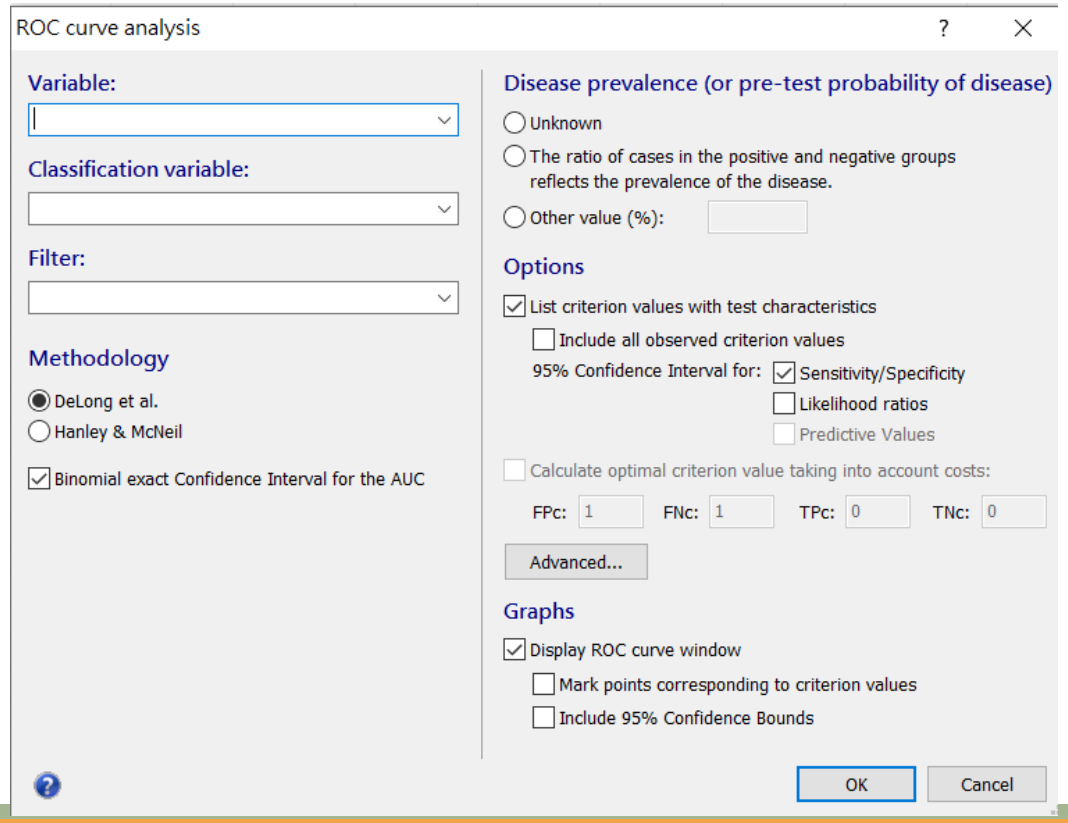




ROC Curve Analysis

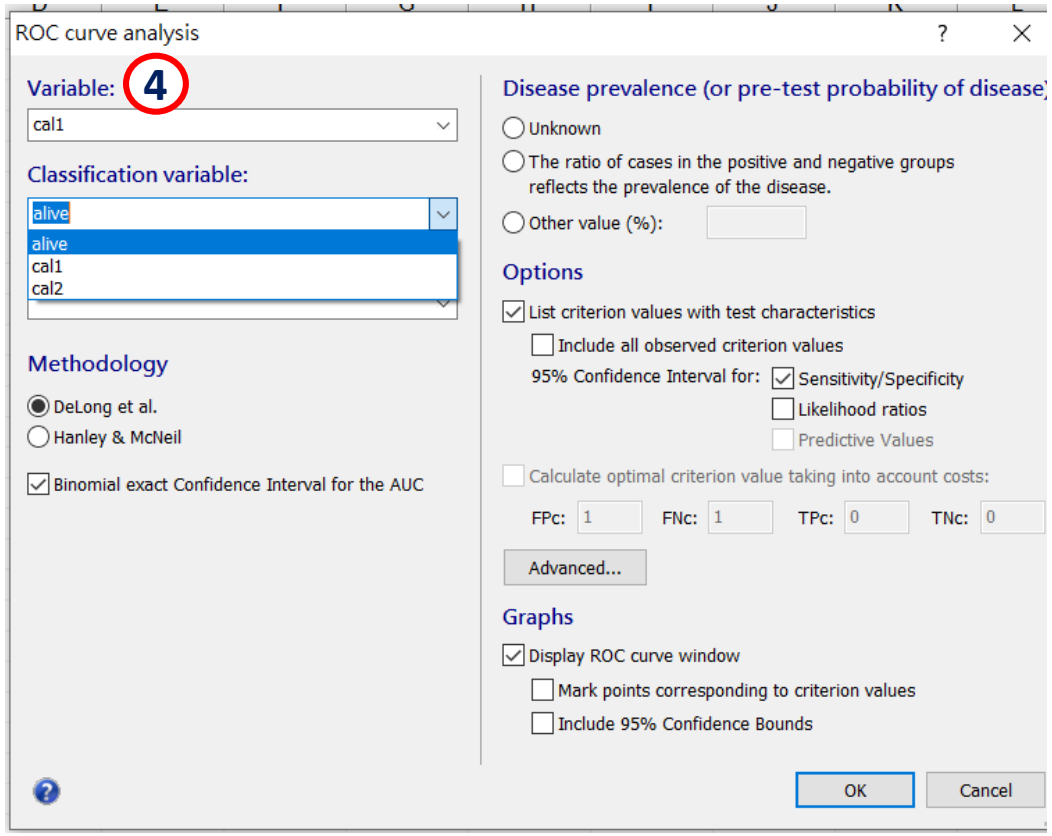
- Medcalc
  - Statistics → ROC curves → ROC Curve Analysis

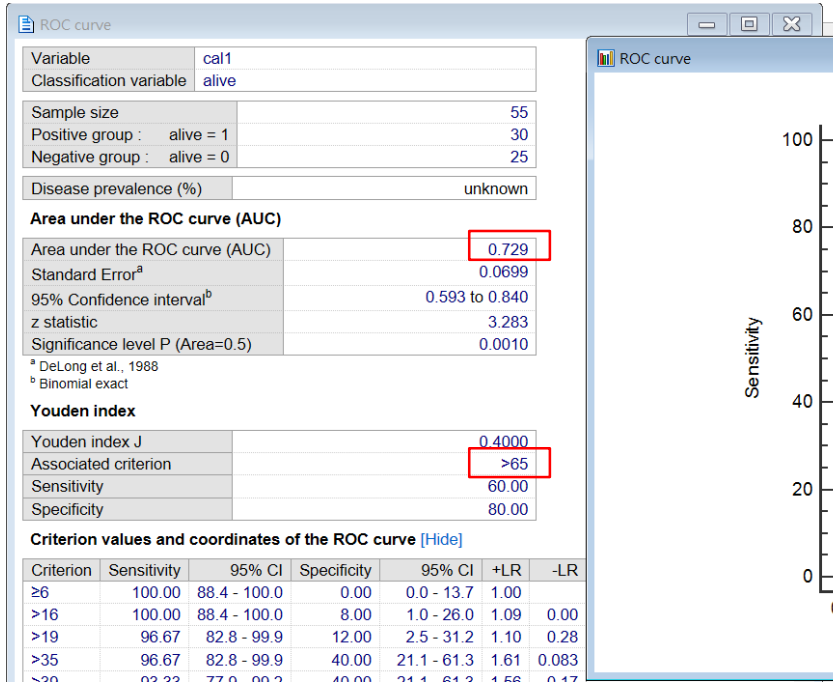




ROC Curve Analysis

- Medcalc
  - Statistics → ROC curves → ROC Curve Analysis





ROC curve

Variable	cal1
Classification variable	alive

Sample size		55
Positive group :	alive = 1	30
Negative group :	alive = 0	25

Disease prevalence (%)	unknown

Area under the ROC curve (AUC)

Area under the ROC curve (AUC)	0.729
Standard Error[a]	0.0699
95% Confidence interval[b]	0.593 to 0.840
z statistic	3.283
Significance level P (Area=0.5)	0.0010

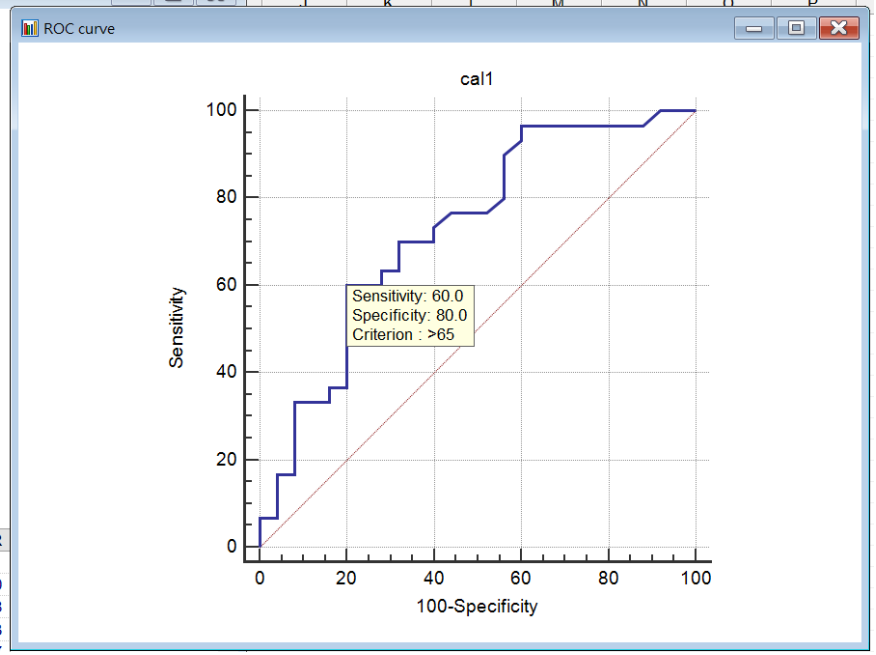
[a] DeLong et al., 1988
[b] Binomial exact

Youden index

Youden index J	0.4000
Associated criterion	>65
Sensitivity	60.00
Specificity	80.00

Criterion values and coordinates of the ROC curve [Hide]

Criterion	Sensitivity	95% CI	Specificity	95% CI	+LR	-LR
≥6	100.00	88.4 - 100.0	0.00	0.0 - 13.7	1.00	
>16	100.00	88.4 - 100.0	8.00	1.0 - 26.0	1.09	0.00
>19	96.67	82.8 - 99.9	12.00	2.5 - 31.2	1.10	0.28
>35	96.67	82.8 - 99.9	40.00	21.1 - 61.3	1.61	0.083



ROC Curve Analysis

- Medcalc
  - Statistics → ROC curves → ROC Curve Analysis

ROC curve

Variable	test3
Classification variable	Dis

Sample size	100
Positive group : Dis = 1	20
Negative group : Dis = 0	80

Disease prevalence (%)	unknown

Area under the ROC curve (AUC)

Area under the ROC curve (AUC)	0.739
Standard Error[a]	0.0512
95% Confidence interval[b]	0.641 to 0.822
z statistic	4.660
Significance level P (Area=0.5)	<0.0001

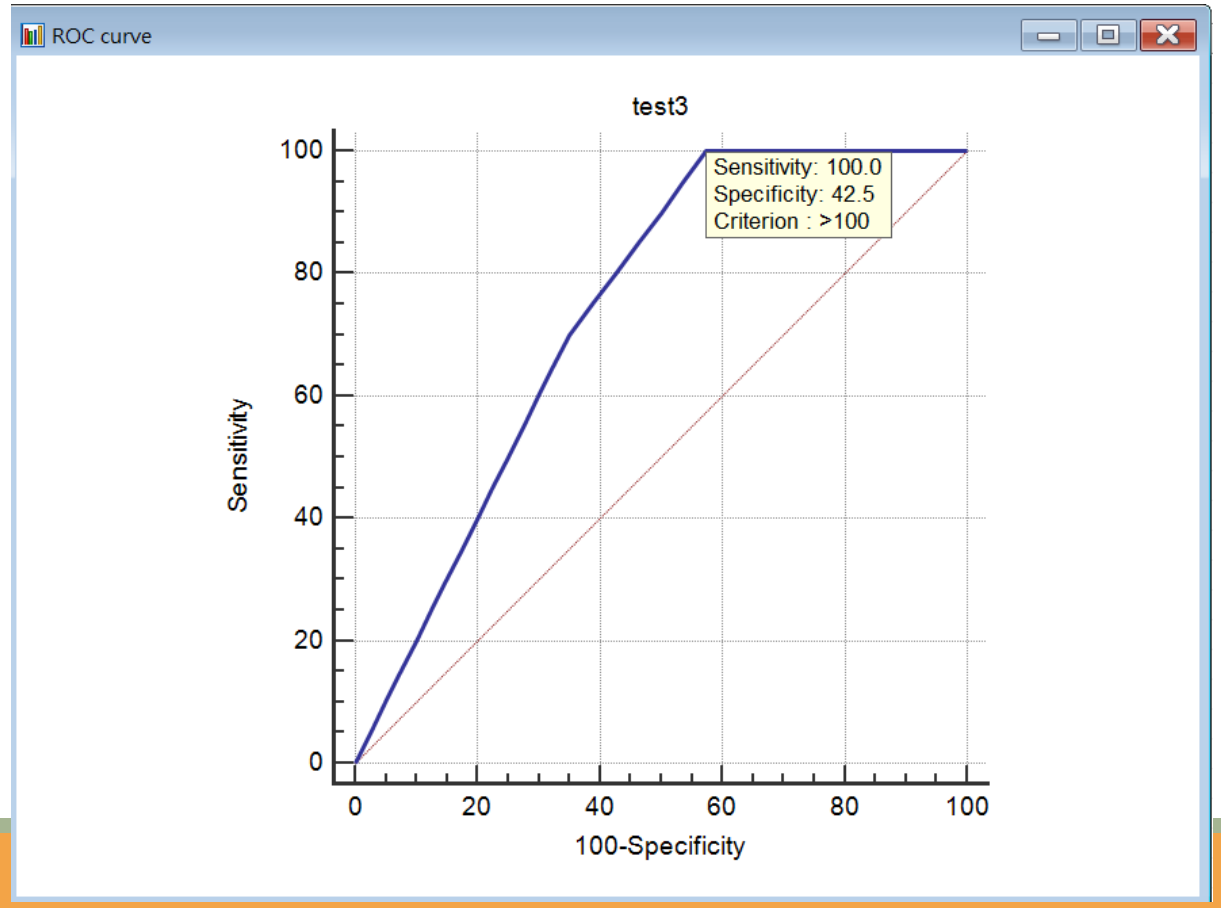
[a] DeLong et al., 1988
[b] Binomial exact

Youden index

Youden index J	0.4250
Associated criterion	>100
Sensitivity	100.00
Specificity	42.50

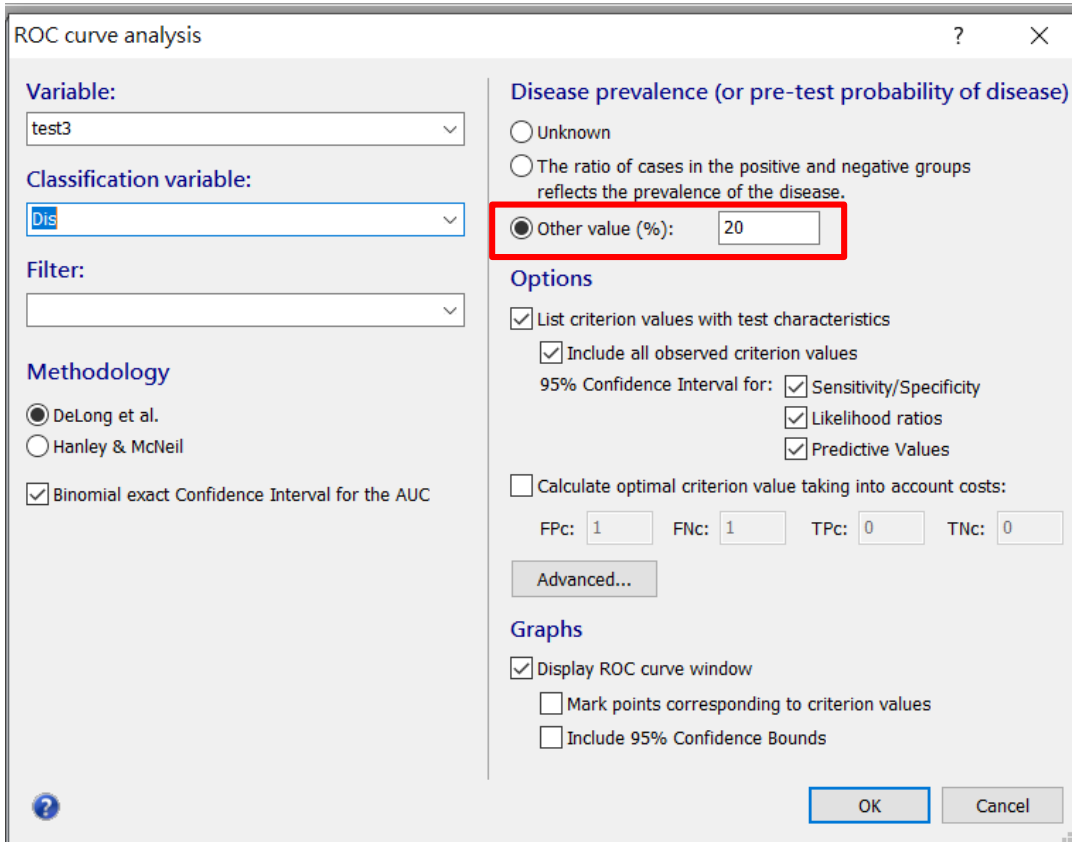
Criterion values and coordinates of the ROC curve [Hide]

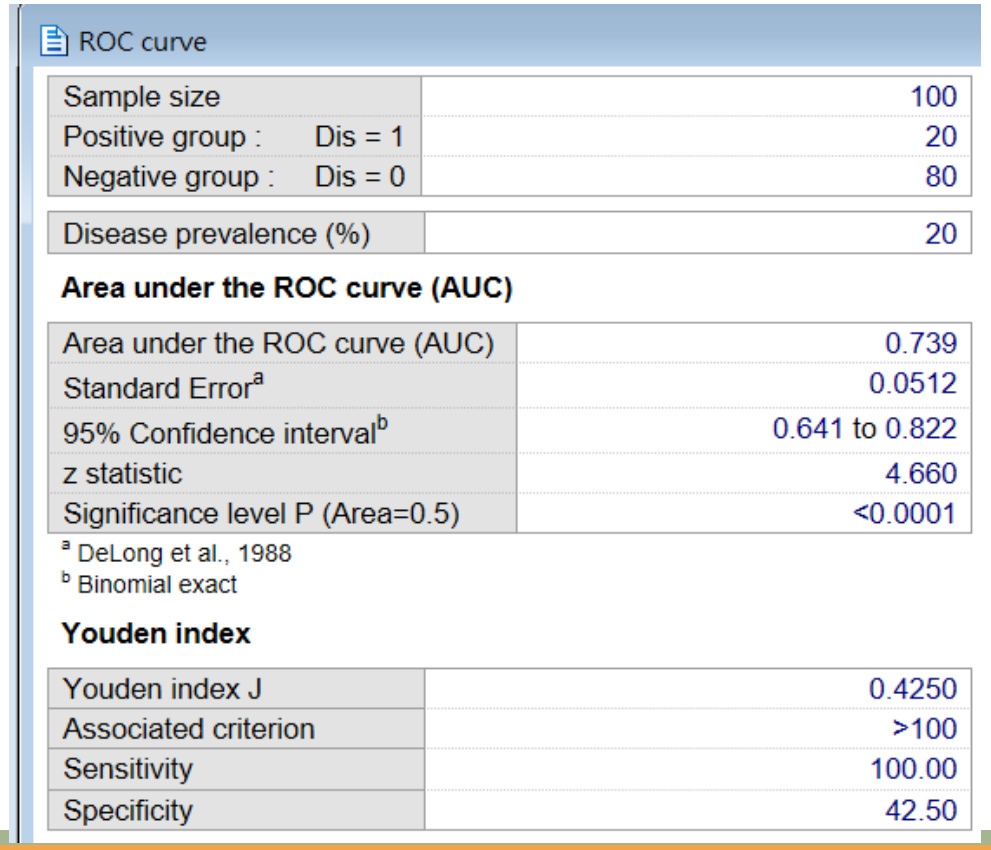
Criterion	Sensitivity	95% CI	Specificity	95% CI	+LR	-LR
≥80	100.00	83.2 - 100.0	0.00	0.0 - 4.5	1.00	
>100	100.00	83.2 - 100.0	42.50	31.5 - 54.1	1.74	0.00
>110	95.00	75.1 - 99.9	46.25	35.0 - 57.8	1.77	0.11



ROC Curve Analysis

- Medcalc
  - Statistics → ROC curves → ROC Curve Analysis





ROC curve

Sample size		100
Positive group :	Dis = 1	20
Negative group :	Dis = 0	80

Disease prevalence (%)	20

Area under the ROC curve (AUC)

Area under the ROC curve (AUC)	0.739
Standard Error[a]	0.0512
95% Confidence interval[b]	0.641 to 0.822
z statistic	4.660
Significance level P (Area=0.5)	<0.0001

[a] DeLong et al., 1988
[b] Binomial exact

Youden index

Youden index J	0.4250
Associated criterion	>100
Sensitivity	100.00
Specificity	42.50

ROC Curve Analysis

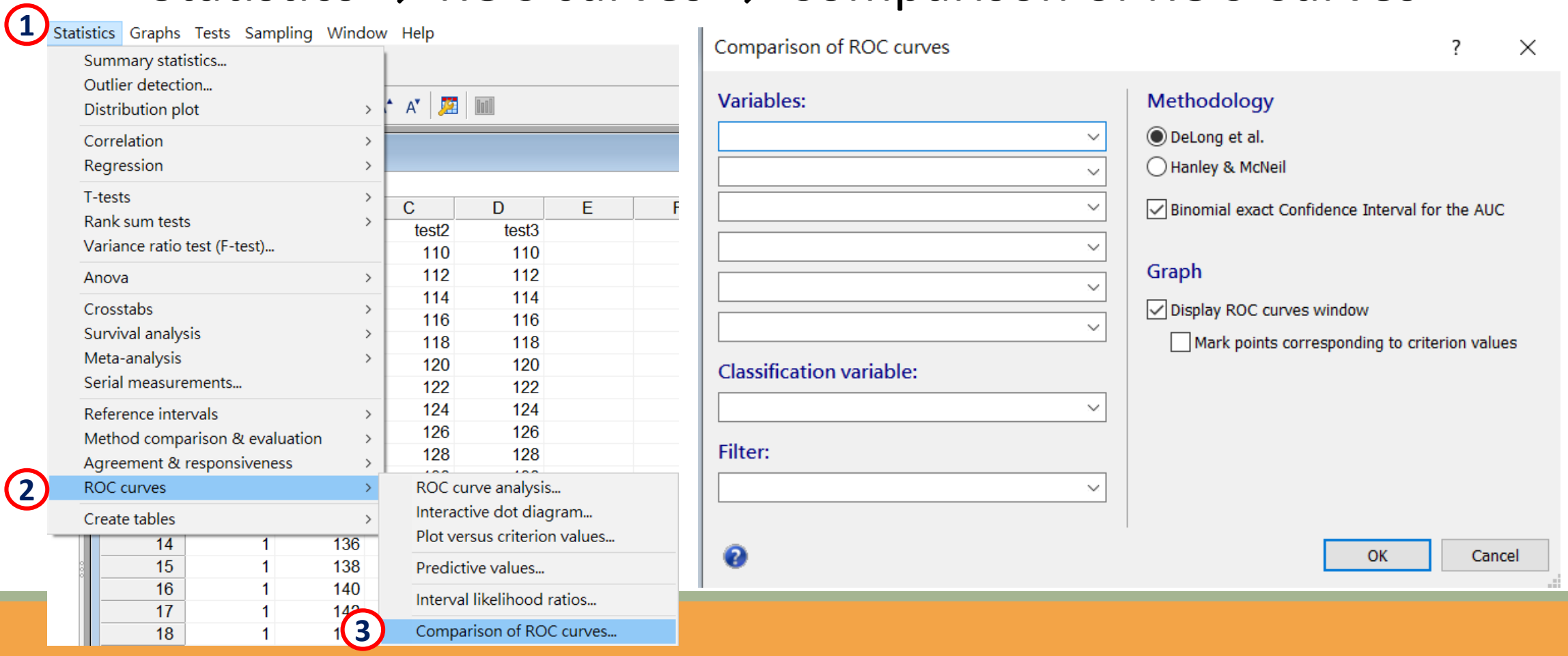
- Medcalc
  - Statistics → ROC curves → ROC Curve Analysis

Criterion values and coordinates of the ROC curve [Hide]

Criterion	Sensitivity	95% CI	Specificity	95% CI	+LR	95% CI	-LR	95% CI	+PV	95% CI	-PV	95% CI
≥80	100.00	83.2 - 100.0	0.00	0.0 - 4.5	1.00	1.0 - 1.0			20.0	12.7 - 29.2		
>80	100.00	83.2 - 100.0	2.50	0.3 - 8.7	1.03	1.0 - 1.1	0.00		20.4	12.9 - 29.7	100.0	15.8 - 100.0
>82	100.00	83.2 - 100.0	5.00	1.4 - 12.3	1.05	1.0 - 1.1	0.00		20.8	13.2 - 30.3	100.0	39.8 - 100.0
>84	100.00	83.2 - 100.0	7.50	2.8 - 15.6	1.08	1.0 - 1.2	0.00		21.3	13.5 - 30.9	100.0	54.1 - 100.0
>86	100.00	83.2 - 100.0	10.00	4.4 - 18.8	1.11	1.0 - 1.2	0.00		21.7	13.8 - 31.6	100.0	63.1 - 100.0
>88	100.00	83.2 - 100.0	12.50	6.2 - 21.8	1.14	1.1 - 1.2	0.00		22.2	14.1 - 32.2	100.0	69.2 - 100.0
>90	100.00	83.2 - 100.0	17.50	9.9 - 27.6	1.21	1.1 - 1.3	0.00		23.3	14.8 - 33.6	100.0	76.8 - 100.0
>92	100.00	83.2 - 100.0	22.50	13.9 - 33.2	1.29	1.1 - 1.5	0.00		24.4	15.6 - 35.1	100.0	81.5 - 100.0
>94	100.00	83.2 - 100.0	27.50	18.1 - 38.6	1.38	1.2 - 1.6	0.00		25.6	16.4 - 36.8	100.0	84.6 - 100.0
>96	100.00	83.2 - 100.0	32.50	22.4 - 43.9	1.48	1.3 - 1.7	0.00		27.0	17.4 - 38.6	100.0	86.8 - 100.0
>98	100.00	83.2 - 100.0	37.50	26.9 - 49.0	1.60	1.4 - 1.9	0.00		28.6	18.4 - 40.6	100.0	88.4 - 100.0
>100	100.00	83.2 - 100.0	42.50	31.5 - 54.1	1.74	1.4 - 2.1	0.00		30.3	19.6 - 42.9	100.0	89.7 - 100.0
>110	95.00	75.1 - 99.9	46.25	35.0 - 57.8	1.77	1.4 - 2.2	0.11	0.02 - 0.7	30.6	19.6 - 43.7	97.4	86.2 - 99.9
>112	90.00	68.3 - 98.8	50.00	38.6 - 61.4	1.80	1.4 - 2.3	0.20	0.05 - 0.8	31.0	19.5 - 44.5	95.2	83.8 - 99.4
>114	85.00	62.1 - 96.8	53.75	42.2 - 65.0	1.84	1.4 - 2.5	0.28	0.10 - 0.8	31.5	19.5 - 45.6	93.5	82.1 - 98.6
>116	80.00	56.3 - 94.3	57.50	45.9 - 68.5	1.88	1.3 - 2.6	0.35	0.1 - 0.9	32.0	19.5 - 46.7	92.0	80.8 - 97.8
>118	75.00	50.9 - 91.3	61.25	49.7 - 71.9	1.94	1.3 - 2.8	0.41	0.2 - 0.9	32.6	19.5 - 48.0	90.7	79.7 - 96.9
>120	70.00	45.7 - 88.1	65.00	53.5 - 75.3	2.00	1.3 - 3.0	0.46	0.2 - 0.9	33.3	19.6 - 49.5	89.7	78.8 - 96.1
>122	65.00	40.8 - 84.6	67.50	56.1 - 77.6	2.00	1.3 - 3.1	0.52	0.3 - 1.0	33.3	19.1 - 50.2	88.5	77.8 - 95.3
>124	60.00	36.1 - 80.9	70.00	58.7 - 79.7	2.00	1.2 - 3.3	0.57	0.3 - 1.0	33.3	18.6 - 51.0	87.5	76.8 - 94.4
>126	55.00	31.5 - 76.9	72.50	61.4 - 81.9	2.00	1.2 - 3.4	0.62	0.4 - 1.0	33.3	18.0 - 51.8	86.6	76.0 - 93.7
>128	50.00	27.2 - 72.8	75.00	64.1 - 84.0	2.00	1.1 - 3.6	0.67	0.4 - 1.1	33.3	17.3 - 52.8	85.7	75.3 - 92.9
>130	45.00	23.1 - 68.5	77.50	66.8 - 86.1	2.00	1.1 - 3.8	0.71	0.5 - 1.1	33.3	16.5 - 54.0	84.9	74.6 - 92.2
>132	40.00	19.1 - 63.9	80.00	69.6 - 88.1	2.00	1.0 - 4.0	0.75	0.5 - 1.1	33.3	15.6 - 55.3	84.2	74.0 - 91.6
>134	35.00	15.4 - 59.2	82.50	72.4 - 90.1	2.00	0.9 - 4.3	0.79	0.6 - 1.1	33.3	14.6 - 57.0	83.5	73.5 - 90.9
>136	30.00	11.9 - 54.3	85.00	75.3 - 92.0	2.00	0.9 - 4.7	0.82	0.6 - 1.1	33.3	13.3 - 59.0	82.9	73.0 - 90.3
>138	25.00	8.7 - 49.1	87.50	78.2 - 93.8	2.00	0.8 - 5.2	0.86	0.7 - 1.1	33.3	11.8 - 61.6	82.4	72.6 - 89.8
>140	20.00	5.7 - 43.7	90.00	81.2 - 95.6	2.00	0.7 - 6.0	0.89	0.7 - 1.1	33.3	9.9 - 65.1	81.8	72.2 - 89.2
>142	15.00	3.2 - 37.9	92.50	84.4 - 97.2	2.00	0.5 - 7.3	0.92	0.8 - 1.1	33.3	7.5 - 70.1	81.3	71.8 - 88.7
>144	10.00	1.2 - 31.7	95.00	87.7 - 98.6	2.00	0.4 - 10.2	0.95	0.8 - 1.1	33.3	4.3 - 77.7	80.9	71.4 - 88.2
>146	5.00	0.1 - 24.9	97.50	91.3 - 99.7	2.00	0.2 - 21.0	0.97	0.9 - 1.1	33.3	0.8 - 90.6	80.4	71.1 - 87.8
>148	0.00	0.0 - 16.8	100.00	95.5 - 100.0			1.00	1.0 - 1.0			80.0	70.8 - 87.3

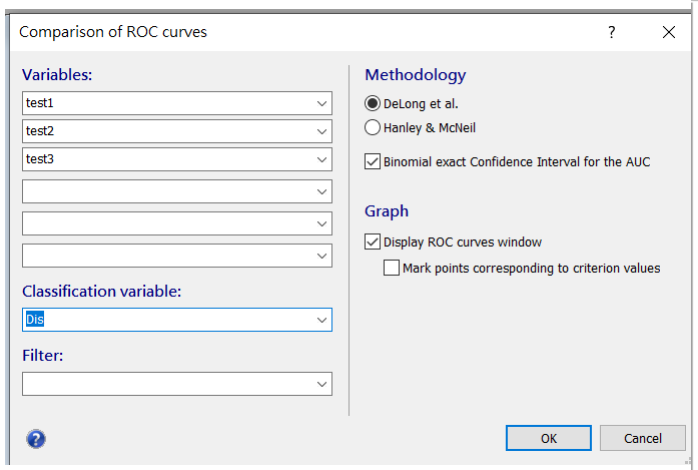
ROC Curve Analysis

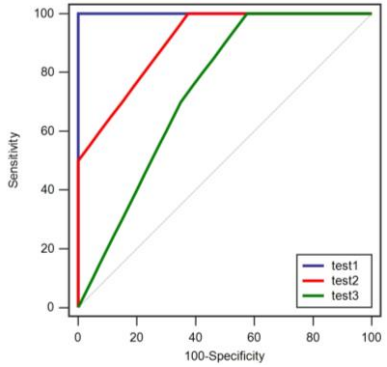
- Medcalc
  - Statistics → ROC curves → Comparison of ROC Curves



ROC Curve Analysis

- Medcalc
  - Statistics → ROC curves → Comparison of ROC Curves





Variable 1	test1
Variable 2	test2
Variable 3	test3
Classification variable	Dis

Sample size	100
Positive group : Dis = 1	20
Negative group : Dis = 0	80

	AUC	SE [a]	95% CI [b]
test1	1.000	0.000	0.964 to 1.000
test2	0.906	0.0324	0.831 to 0.955
test3	0.739	0.0512	0.641 to 0.822

[a] DeLong et al., 1988
[b] Binomial exact

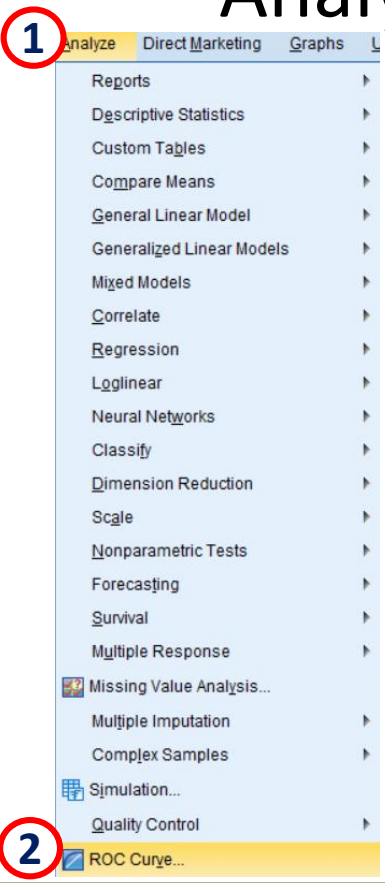
Pairwise comparison of ROC curves

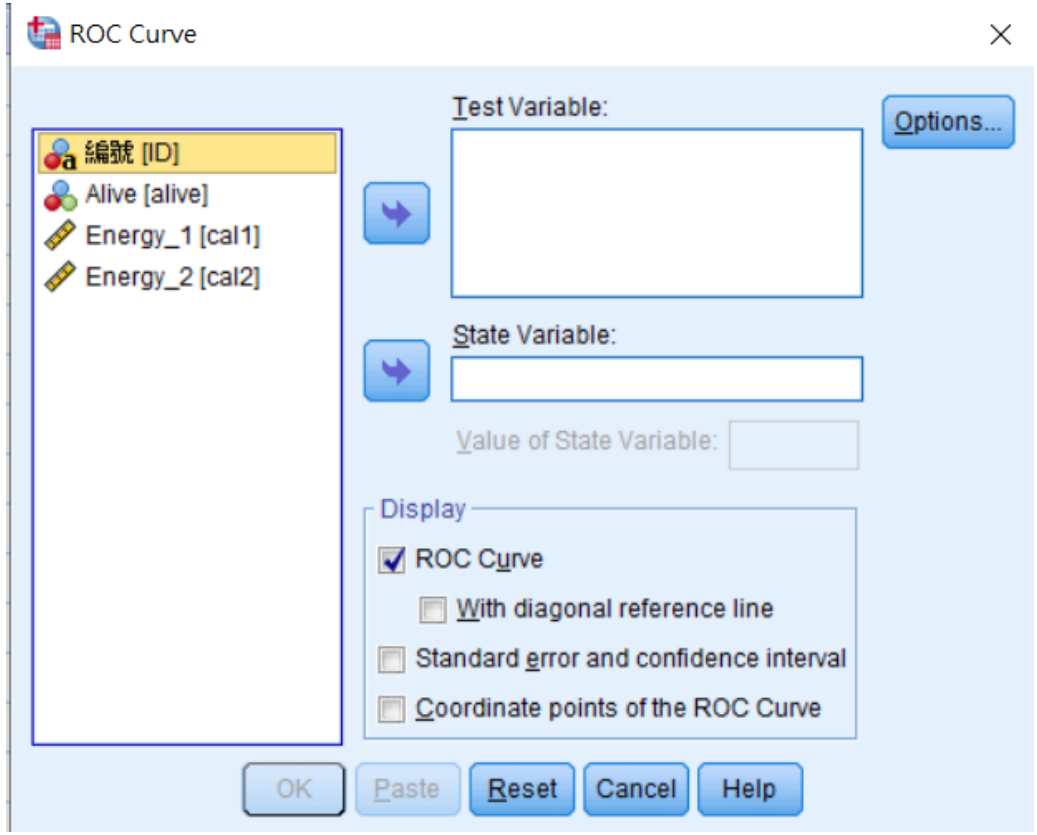
test1 ~ test2	
Difference between areas	0.0937
Standard Error[c]	0.0324
95% Confidence Interval	0.0302 to 0.157
z statistic	2.892
Significance level	P = 0.0038
test1 ~ test3	
Difference between areas	0.261
Standard Error[c]	0.0512
95% Confidence Interval	0.161 to 0.362
z statistic	5.099
Significance level	P < 0.0001
test2 ~ test3	
Difference between areas	0.168
Standard Error[c]	0.0349
95% Confidence Interval	0.0991 to 0.236
z statistic	4.797
Significance level	P < 0.0001

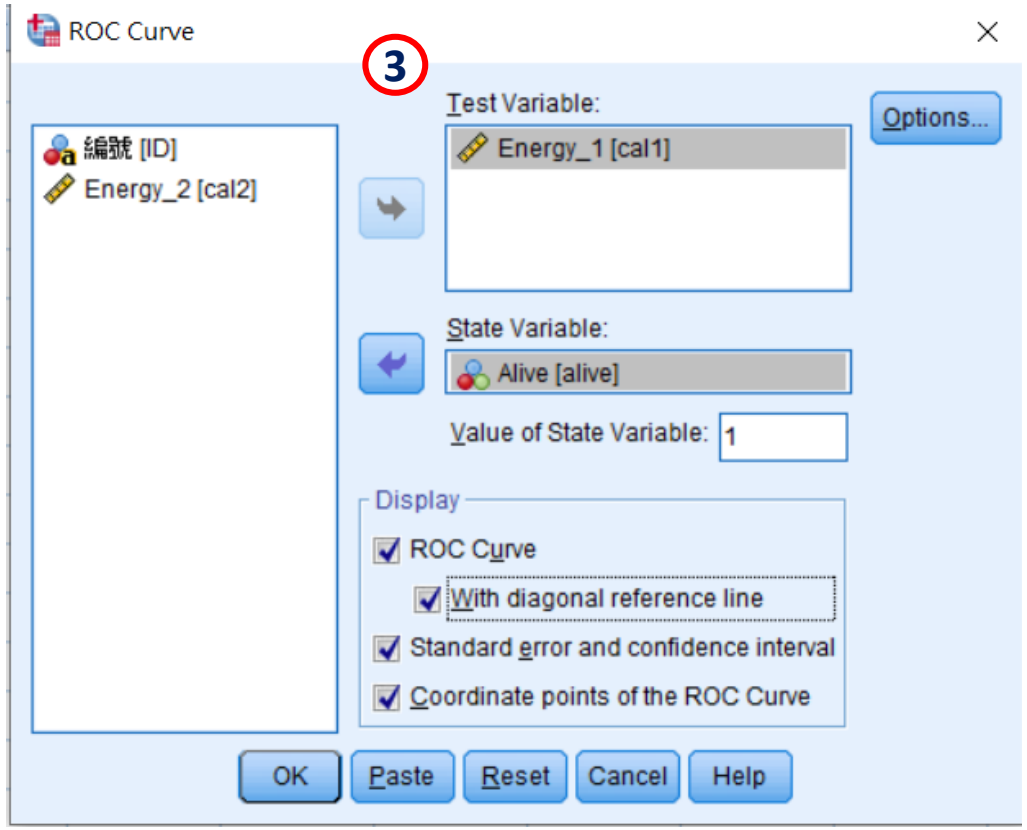
[c] DeLong et al., 1988

ROC Curve Analysis

- SPSS
  - Analyze → ROC curve







ROC Curve Analysis

- SPSS

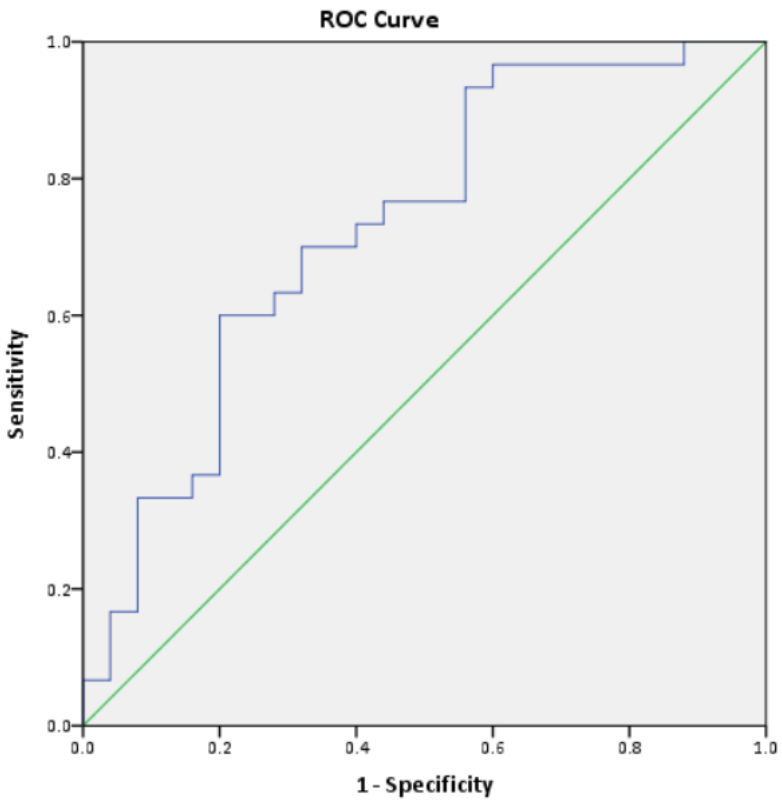
ROC Curve

Case Processing Summary

Alive	Valid N (listwise)
Positive[a]	30
Negative	25

Larger values of the test result variable(s) indicate stronger evidence for a positive actual state.

a. The positive actual state is Alive.



Area Under the Curve

Test Result Variable(s): Energy_1

Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.729	.069	.004	.594	.865

a. Under the

b. Null hypothesis

Coor

Test Result Varia

Positive if Greater Than or Equal To[a]	Sensitivity	1 - Specificity
4.82	1.000	1.000
10.68	1.000	.960
17.07	1.000	.920
18.76	1.000	.880
21.61	.967	.880
24.43	.967	.840
25.66	.967	.800

AUC = 0.73
(0.59-0.87)
Acceptable discrimination

ROC Curve Analysis

- Cut-point
  - Youden's index (Maximum)
    - Sensitivity + Specificity - 1

IF =B4+D4-1

	A	B	C	D	E
1	Coordinates of the Curve				
2	Test Result Variable(s):				
3	Positive if Greater Than or Equal To[a]	Sensitivity	1 - Specificity	Specificity	Youden's index
4	65.22	.600	.200	0.800	=B4+D4-1
5	54.90	.700	.320	0.680	0.380
6	40.99	.933	.560	0.440	0.373
7	36.66	.967	.600	0.400	0.367
8	66.28	.567	.200	0.800	0.367
9	64.15	.600	.240	0.760	0.360
10	60.46	.633	.280	0.720	0.353
11	55.91	.667	.320	0.680	0.347
12	41.64	.900	.560	0.440	0.340
13	53.00	.700	.360	0.640	0.340
14	39.65	.933	.600	0.400	0.333
15	66.95	.533	.200	0.800	0.333

生物統計小組
統計方法教育訓練

統計方法教育訓練

Thank you

課程問卷

Diagnostic test

For your attention!!