113年度衛生福利資料科學中心統計軟體推廣課程

2024.09.30 (一) 09:00-12:00 @國立成功大學 75201電腦教室

健康資料研究與R軟體

《生物統計分析方法》
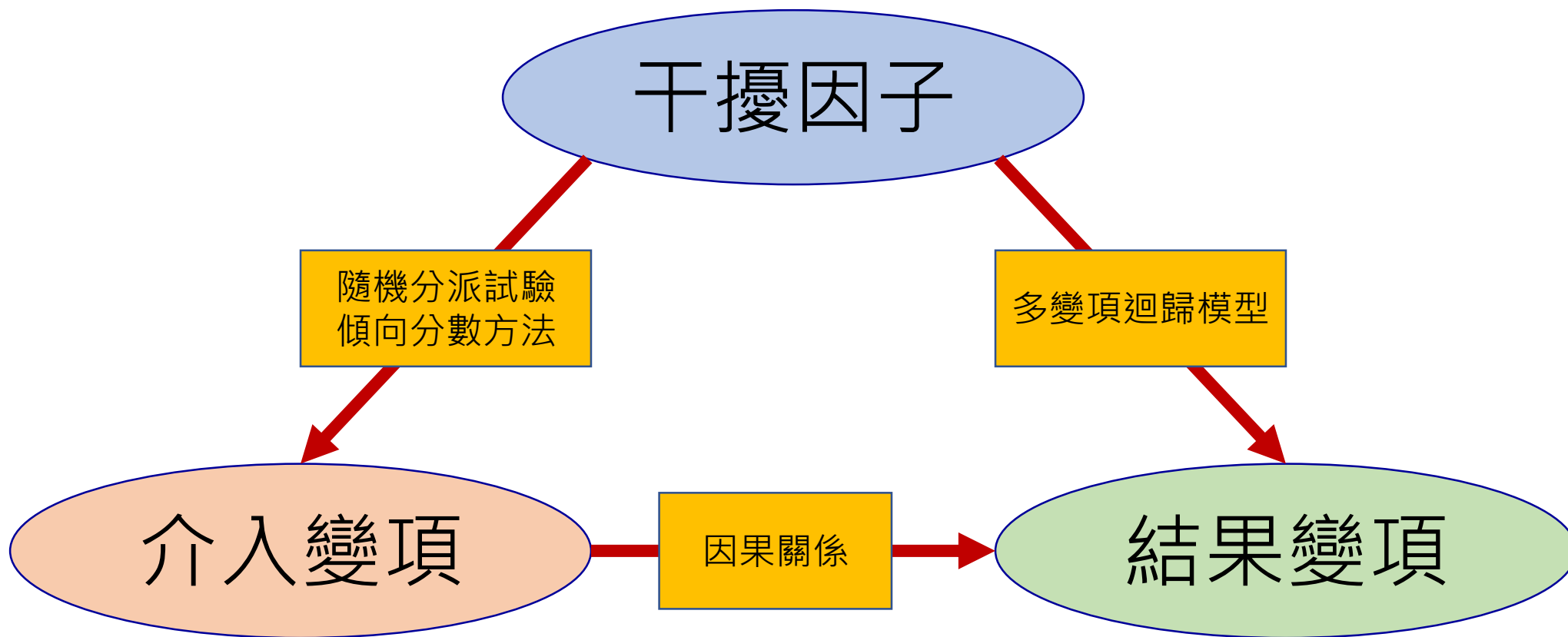
劉品崧　統計諮詢暨分析師

花蓮慈濟醫院健康長壽中心

課程大綱

- 干擾因子與統計分析

- 基礎統計分析方法
  - 降血壓藥物隨機分派試驗
  - 統計分析計畫

- 模擬試驗存活分析
  - 抗凝血藥物選擇與未來糖尿病併發症風險
  - 統計分析計畫

統計分析與干擾因子

干擾因子

介入變項

結果變項

隨機分派試驗
傾向分數方法

多變項迴歸模型

因果關係

基礎統計分析方法

- 降血壓藥物隨機分派試驗

- 統計分析計畫

降血壓藥物隨機分派試驗

- 對象 Population

- 介入 Intervention

- 對照 Control

- 結果 Outcome

- 設計 Study

- 高血壓病人

- 新開發藥物

- 安慰劑

- 血壓下降

- RCT

依據你最感興趣的結果變數（Y）分為

- 數值型態

  - Mean、SD、Person's r、Box-plot

  - t-test、ANOVA

  - Linear regression

- 類別型態

  - N、Percent

  - 交叉表、χ^2test、Fisher exact test

  - Logistic regression

模擬試驗存活分析

- AF & DM病人口服抗凝血劑選擇對併發症影響

- 統計分析計畫

AF & DM病人口服抗凝血劑選擇對併發症影響

Diabetes-Related Complications and Mortality in Patients With Atrial Fibrillation Receiving Different Oral Anticoagulants : A Nationwide Analysis

Huei-Kai Huang [1], Peter Pin-Sung Liu [2], Shu-Man Lin [3], Jin-Yi Hsu [4], Jih-I Yeh [5], Edward Chia-Cheng Lai [6], Carol Chiung-Hui Peng [7], Kashif M Munir [8], Ching-Hui Loh [4], Yu-Kang Tu [9]

Affiliations + expand
PMID: 35157495   DOI: 10.7326/M21-3498

AF & DM病人口服抗凝血劑選擇對併發症影響

- 對象 Population

- 介入 Intervention

- 對照 Control

- 結果 Outcome

- 設計 Study

- Patients with AF & DM

- NOAC

- Warfarin

- DM complications

- Cohort study

依據你最感興趣的結果變數（Y）分為

- 數值型態
  - Mean、SD、t-test、ANOVA

- 類別型態
  - N、Percent、交叉表、χ^2test、Fisher exact test

- 處理干擾因子
  - Propensity score matching

- Time-to-event analysis
  - Kaplan-Meier plot
  - Cox proportional hazard regression

原始樣本背景特質比較

Table 1. Baseline characteristics of patients (original cohort)

Variable	NOAC N = 6,916		Warfarin N = 3,335		SMD
	N	(%)	N	(%)	
Age*	73.61	9.73	68.56	11.36	0.477
Male	3,602	52.1	1,848	55.4	0.067
Index year group					0.846
2012-2013	849	12.3	1,504	45.1	
2014-2015	2,444	35.3	1,063	31.9	
2016-2017	3,623	52.4	768	23.0	
CHA_2DS_2-VASc Score	2.02	1.42	2.02	1.41	0.001
Hyperlipidemia	2,435	35.2	1,164	34.9	0.006
CKD	815	11.8	492	14.8	0.088
Cancer	365	5.3	176	5.3	<0.001

* Expressed as mean and SD.

Abbreviations: n, number; SD, standard deviation; CKD, chronic kidney disease

計算與評估傾向分數（Propensity score, PS）

- Logistic regression model 干擾因子如何影響分組結果？

  - <mark>使用藥物</mark> ~ 性別 + 年齡（暫時忽略病史及指標年分）

  - $ln(\frac{P(NOAC=1)}{P(NOAC=0)}) = -3.87 + (-0.10 * male) + (0.04 * age)$
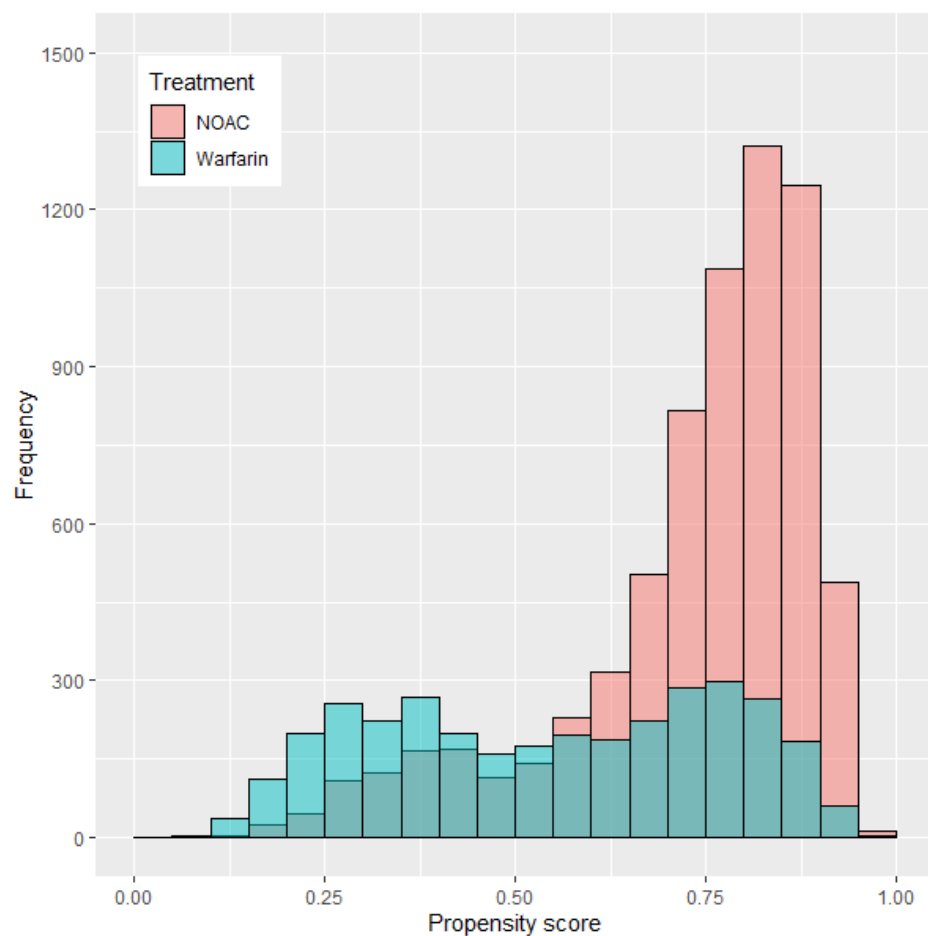
- Propensity score 這個樣本的條件被分到治療組的機率？

  - 65歲男性

  - $\hat{p} = \frac{1}{1+e^{(-(-3.87+\beta_{male}X_{male}+\beta_{age}X_{age}))}}$
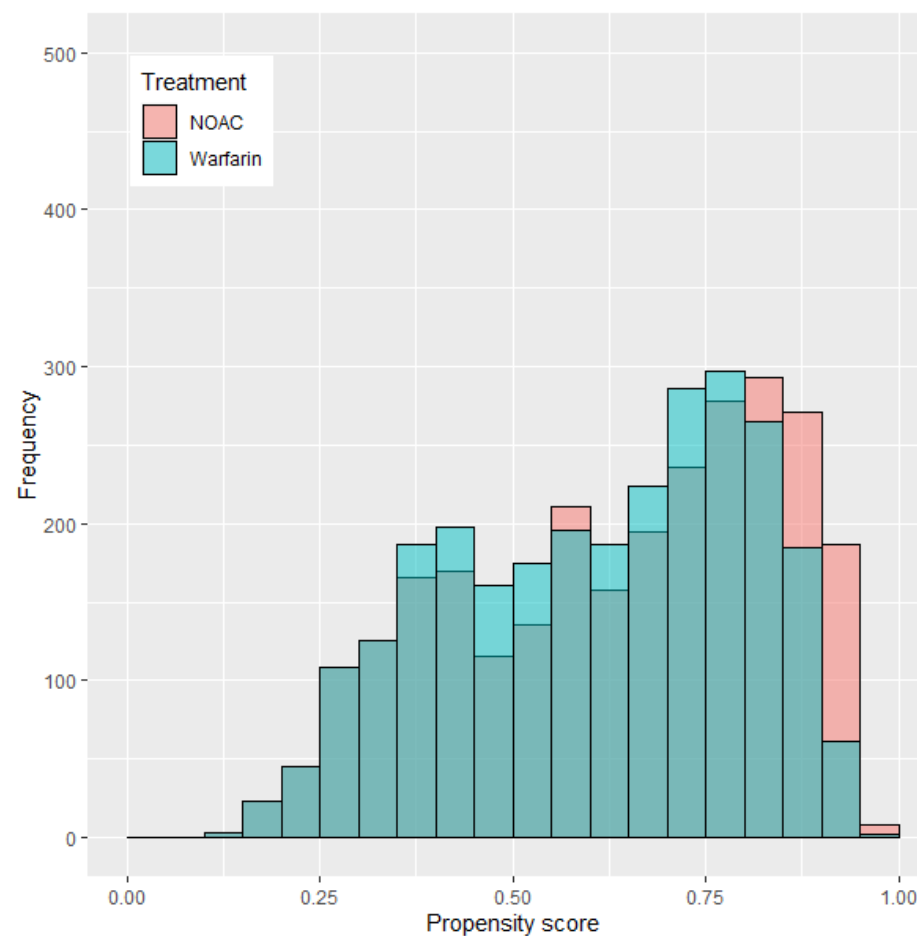
  - $\hat{p} = \frac{1}{1+e^{(-(-3.87+(-0.10*1)+(0.04*65)))}} = 0.2026$

傾向分數分布比較

- 配對前

- 配對後

樣本背景特質比較（配對前 vs 配對後）

Table 1. Baseline characteristics of patients (original cohort)

Variable	NOAC N = 6,916		Warfarin N = 3,335		SMD
	N	(%)	N	(%)	
Age*	73.61	9.73	68.56	11.36	0.477
Male	3,602	52.1	1,848	55.4	0.067
Index year group					0.846
2012-2013	849	12.3	1,504	45.1	
2014-2015	2,444	35.3	1,063	31.9	
2016-2017	3,623	52.4	768	23.0	
CHA_2DS_2-VASc Score	2.02	1.42	2.02	1.41	0.001
Hyperlipidemia	2,435	35.2	1,164	34.9	0.006
CKD	815	11.8	492	14.8	0.088
Cancer	365	5.3	176	5.3	<0.001

* Expressed as mean and SD.

Abbreviations: n, number; SD, standard deviation; CKD, chronic kidney disease

Table 1. Baseline characteristics of patients (matched cohort)

Variable	NOAC N = 2,728		Warfarin N = 2,728		SMD
	N	(%)	N	(%)	
Age*	72.41	11.65	70.29	11.17	0.185
Male	1,297	47.5	1,448	53.1	0.111
Index year group					0.160
2012-2013	842	30.9	907	33.2	
2014-2015	918	33.7	1,053	38.6	
2016-2017	968	35.5	768	28.2	
CHA_2DS_2-VASc Score	2.34	1.51	2.01	1.41	0.222
Hyperlipidemia	1,162	42.6	953	34.9	0.158
CKD	540	19.8	356	13.0	0.183
Cancer	241	8.8	142	5.2	0.142

* Expressed as mean and SD.

Abbreviations: n, number; SD, standard deviation; CKD, chronic kidney disease

Kaplan-Meier plot（配對後）
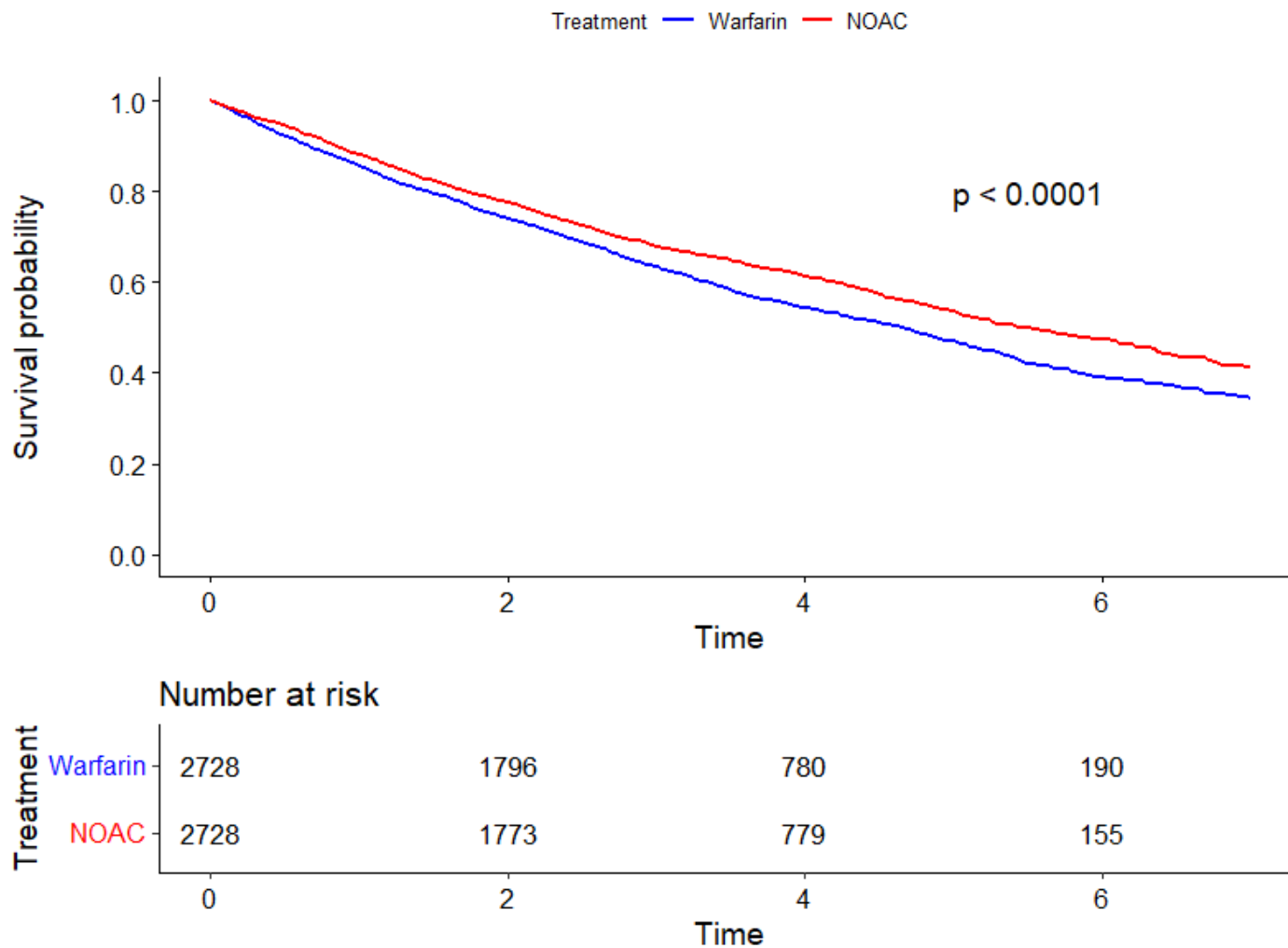
Cox proportional hazard regression

Table 2. Risk of macrovascular complication

Original cohort	N	Events	FU	IR	aHR (95% CI)	p value
Warfarin	3335	1692	11476	147.44	1.00 (reference)	
NOAC	6916	2480	19582	126.65	0.79 (0.74-0.85)	<.0001
Matched cohort	N	Events	FU	IR	HR (95% CI)	p value
Warfarin	2728	1342	8786	152.75	1.00 (reference)	
NOAC	2728	1125	8929	125.99	0.83 (0.75-0.90)	<.0001

Abbreviations: n, number; FU, follow-up time (years); IR, incidence rate per 1,000 person years; aHR, adjusted hazard ratio; CI, confidence intervals.

Summary

- 研究目的

- 統計分析計畫

- 干擾因子

- 傾向分數配對

- 開放提問時間

- FB：劉品崧

- Peter Pin-Sung Liu

- psliu520@gmail.com

- https://github.com/PSLiu/

▶ 109年度R基礎課程-劉品崧老師 🔍