

住院病患於急性病情變化前之 Modified Early Warning Score (MEWS) 指數是否具有預測價值

第一組：陸貳伍壹肆

組長：統計所 洪若縈

組員：統計所 莊誠奉

統計所 廖毅桓

統計所 簡柏翔

指導教授：鄭順林 教授

June 12, 2020

目錄

- 1 研究簡介
- 2 資料介紹
- 3 資料分析
- 4 結論與討論

問題

- 非預期病況變差的病人，包括
 - 轉入 ICU
 - 院內死亡
 - 病危出院
- 諮詢者認為在發生非預期事件的前 48 小時，病人的病情應該可以分成下列三種情況：
 - 病情由好轉壞
 - 病情由壞轉好再轉壞
 - 病情狀況維持低下
- 諮詢者的問題：病人在發生事件前的病情變化呈現何種情形。

目標

1. 針對發生非預期事件的病人，了解他們在發生非預期事件前生命徵象的變化情形，試圖找出病人病況變差的判斷依據。
 - 資料僅包含發生非預期事件的病人 (case)。
2. 探討 MEWS 分數能否推斷病人會不會發生非預期事件。
 - 資料包含有發生事件 (case) 及沒有發生事件 (control) 的病人。

MEWS 簡介

● Modified Early Warning Score (MEWS) :

Score	3	2	1	0	1	2	3
血壓(mmHg)	< 70	71 – 80	81 – 100	101 – 199		≥ 200	
心跳 (bpm)		< 40	41 – 50	51 – 100	101 – 110	111 – 129	≥ 130
呼吸速率(bpm)		< 9		9 – 14	15 – 20	21 – 29	≥ 30
溫度(C)		< 35		35 – 38.4		≥ 38.5	
意識狀況				14 – 15	10 – 13	4 – 9	3

資料介紹

	第一筆資料	第二筆資料
資料筆數	20807	711522
變數個數	29	21
病患人數	3552	23756
來源	急診室	護理師的電腦
時間	2018/6 ~ 2019/6	2018/6 ~ 2018/12
對象	case	all
取得時間	3/17	4/2
備註	每 8 小時進行量測， 且該次住院至多 6 筆	從入院到出院的所有紀錄

變數介紹(第一筆資料)

- 病歷號，床號 (病房 + 床號)，性別，年齡
- 出院時間，進 ICU 時間，手術時間，紀錄時間，紀錄時辰
- 呼吸速率，脈搏，體溫，血壓，意識 (亂碼)，尿量 (大量遺失值)
- 呼吸速率分數，脈搏分數，體溫分數，血壓分數，意識分數，尿量分數
- RSS Level，總分 (呼吸 + 脈搏 + 體溫 + 血壓)，新總分 (呼吸 + 脈搏 + 體溫 + 血壓 + 意識 + 尿量)
- 國際疾病分類標準 (ICD9, ICD10)，非預期狀況，主治醫師代碼

變數介紹(第二筆資料)

- 床號 (病房 + 床號)，病例號，年齡，住院號
- 入院時間，紀錄時間
- 呼吸速率，脈搏，體溫，血壓，意識，尿量 (大量遺失值)
- 呼吸速率分數，脈搏分數，體溫分數，血壓分數，意識分數，尿量分數
- 總分 (呼吸 + 脈搏 + 體溫 + 血壓)，新總分 (呼吸 + 脈搏 + 體溫 + 血壓 + 意識 + 尿量)

綜合第一筆與第二筆資料，未來分析上我們所能夠考慮的生命徵象有呼吸速率、脈搏、體溫及血壓。

資料前處理

1. 異常值處理：改為 NA

- 呼吸速率 > 80
- 體溫 > 45 或體溫 < 30
- 血壓 < 60
- 脈搏 > 300

2. 若一筆紀錄中有太多個生命徵象是遺失值，則進行插補並沒有意義，我們選擇將這種資料刪除。

病歷號	呼吸	脈搏	體溫	血壓	總分
1	20	111	36.4	134	3
1	19	95	36.3	NA	-1
1	19	NA	NA	NA	-1
1	NA	NA	NA	NA	-1

資料前處理

3. 篩選出年齡 ≥ 18

- 由於年齡 < 18 者生命徵象的標準不同，因此只對年齡 ≥ 18 者做分析。

4. 將同一人的所有量測紀錄依紀錄時間做排序

資料前處理

5. 遺失值處理

- 對於遺失值的處理，直觀上有兩種處理方式：刪除或是插補。
- 在實際資料中有一種常見狀況是，

病歷號	呼吸	脈搏	體溫	血壓	總分
1	20	111	36.4	134	3
1	19	95	36.3	NA	-1
1	19	96	35.8	107	2

- "When a predictor value was missing for a time interval, the previous value was carried forward. If no previous value was available, the median value for that variable was imputed."¹

¹Matthew M Churpek et al. "Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards". In: *Critical care medicine* 44.2 (2016), p. 368. ◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ≡ ↺ ↻

目標一、病人在發生事件前的病情變化呈現何種情形

- 針對發生非預期事件的病人，運用機器學習的非監督式學習方法將每一筆紀錄進行分群。
- 方法：k-means
- 變數：呼吸、脈搏、血壓、體溫、MEWS score

分析結果

- 利用所有事件組資料，並初步分為 2 群。



第一主成份和第二主成份可解釋 67.4% 的變異

MEWS		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
predict	1	41	861	1206	1368	1253	1176	839	580	315	177	86	37	14	7	2 (偏高)
	2	75	5233	3324	1107	604	367	165	88	33	12	10	4	0	0	0 (偏低)

分析結果

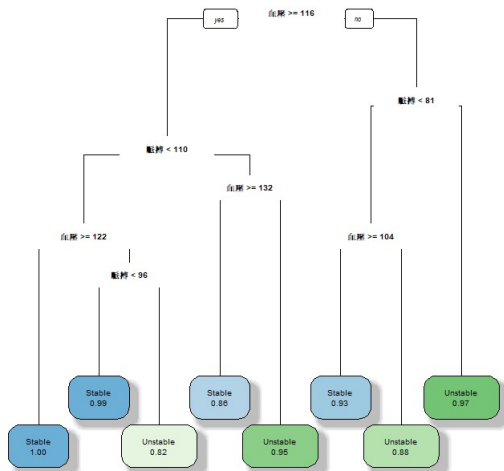
- 從分群結果和 MEWS 的 table 觀察可得知，發生非預期事件前 48 小時資料仍可稍微分出病況穩定、模糊地帶與狀況危險三種程度，不過中間的模糊地帶很大。
- 依據 K-means 的分群結果來定義病況：

	MEWS	Defined Status
1	偏高	Unstable
2	偏低	Stable

- 在使用 K-means 分群後，以決策樹的方式來了解分群的依據。

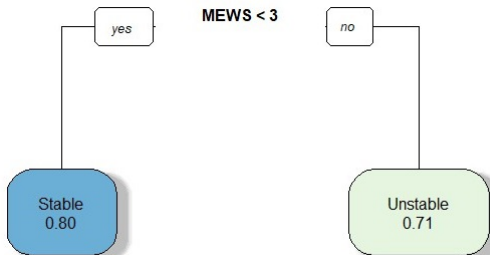
分析結果

- 由下圖可知，**血壓**與**脈搏**會是決策樹分類的重要依據。



分析結果

- 由下圖可知，若 MEWS score < 3 就代表是病況穩定的。



目標二、MEWS能否推斷病人會不會發生非預期事件

- 配對出對照組：

由於 case 及 control 組病人數量差距懸殊，且特徵有所差異，因此使用配對方式找到與事件組特徵相似的病人作為對照組。

1. 方法：傾向評分匹配 (PSM)

- 選定想控制的變數
- 使這些變數在 case, control 兩組有相似的分布

配對出對照組

2. 配對結果

資料中僅有年齡及病房兩個病人基本資料能夠用來進行匹配，我們使用病人第一筆資料的年齡及所在病房進行配對。由於病房總共有 38 類，不宜直接進行配對，因此將病房先進行分類，分類方式以樓層為主。

類別	病房
4F	4A, 4C, 4D
5F	5C
6F	6A, 6B, 6C
7F	7A, 7B, 7C
8F	8A, 8B, 8C
9F	9A, 9B, 9C
10F	10A, 10B, 10C
11F	11A, 11B
12F	12A, 12B, 12C, 12D
ICU	CCU, MI2, MIC, RCC, RIC, BMT
other	3B1, 3S1, 3S2, 3S3, 4N1, 4P1, 4P2

配對出對照組

配對後之病人的年齡與病房種類有更相似的分佈，配對的結果如下表：

Table 1: 配對前

	事件組 (n = 3367)	對照組 (n = 16803)	SMD
年齡 (mean (SD))	66.48 (15.19)	57.58 (17.71)	0.539
病房類別 (%)			0.783
4F	59 (1.8)	1757 (10.5)	
5F	96 (2.9)	909 (5.4)	
6F	534 (15.9)	2921 (17.4)	
7F	299 (8.9)	3744 (22.3)	
8F	630 (18.7)	2500 (14.9)	
9F	495 (14.7)	1607 (9.6)	
10F	820 (24.4)	994 (5.9)	
11F	151 (4.5)	1188 (7.1)	
12F	218 (6.5)	1036 (6.2)	
ICU	45 (1.3)	102 (0.6)	
other	20 (0.6)	45 (0.3)	

Table 2: 配對後

	事件組 (n = 3367)	對照組 (n = 3367)	SMD
年齡 (mean (SD))	66.48 (15.19)	66.45 (15.04)	0.002
病房類別 (%)			0.076
4F	59 (1.8)	44 (1.3)	
5F	96 (2.9)	78 (2.3)	
6F	534 (15.9)	544 (16.2)	
7F	299 (8.9)	291 (8.6)	
8F	630 (18.7)	637 (18.9)	
9F	495 (14.7)	518 (15.4)	
10F	820 (24.4)	799 (23.7)	
11F	151 (4.5)	174 (5.2)	
12F	218 (6.5)	229 (6.8)	
ICU	45 (1.3)	42 (1.2)	
other	20 (0.6)	11 (0.3)	

建立模型

- 考慮 3 種代表生命徵象的指標，並使用 3 種模型進行預測。
- 考慮的生命徵象：呼吸、脈搏、血壓、體溫。

生命徵象指標	模型
Original Vital Sign	Logistic regression
MEWS score	Decision tree
Centile-based score	Random forest

- 定義反應變數：
 - 對於有發生非預期事件的病人，其所有量測紀錄都定義為 1
 - 對於沒有發生非預期事件的病人，其所有量測紀錄都定義為 0

Centile-based score

Centile-based score² 的給分方式為：

- 3 分：1% 以下或 99% 以上。
- 2 分：介於 1% 至 5% 之間或介於 95% 至 99%。
- 1 分：介於 5% 至 10% 之間或介於 90% 至 95%。
- 0 分：介於 10% 至 90% 之間。

下表為根據我們的資料所建的 centile-based score：

Score	3	2	1	0	1	2	3
呼吸	≤ 12	12-16		17-22	23-24	25-32	≥ 33
脈搏	≤ 51	52-59	60-63	64-112	113-121	122-140	≥ 141
體溫	≤ 35	35.1-35.6	35.7-35.8	35.9-37.4	37.5-37.9	38.0-38.9	≥ 39
血壓	≤ 66	67-89	90-98	99-155	156-163	164-181	≥ 182

²Lionel Tarassenko et al. "Centile-based early warning scores derived from statistical distributions of vital signs". In: *Resuscitation* 82.8 (2011), pp. 1013–1018.

分析結果整理

- Logistic regression 分析結果如下表：

Model:	N	AUC	threshold	Sensitivity	Specificity	Cutpoint
Logistic regression						
Original vital sign		0.72	0.35	0.70	0.58	-
MEWS score	45398	0.73	0.34	0.67	0.66	1.53
Centile-based score		0.70	0.33	0.65	0.64	0.52

- 綜合考量 AUC、Sensitivity 及 Specificity，MEWS score 的模型表現最佳。

分析結果整理

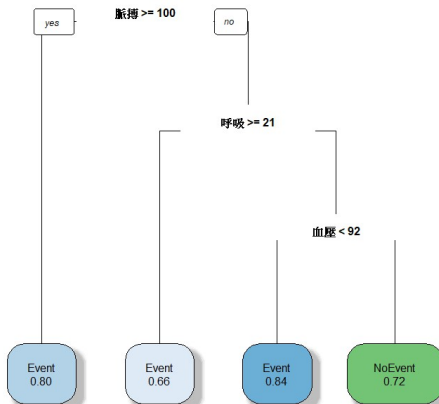
- Decision tree 分析結果如下表：

Model:	N	Accuracy	Sensitivity	Specificity	5-Fold CV (Accuracy)
Decision tree					
Original vital sign		0.73	0.51	0.89	0.73
MEWS score	45398	0.73	0.43	0.94	0.73
Centile-based score		0.68	0.33	0.94	0.68

- 綜合考量 Accuracy、Sensitivity 及 Specificity，Original vital sign 及 MEWS score 的模型表現相當。

分析結果

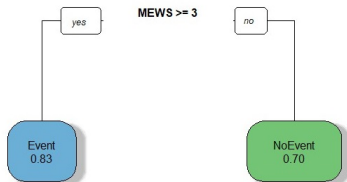
- Decision tree
- 解釋變數：呼吸、脈搏、體溫、血壓



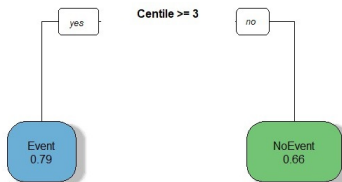
分析結果

- Decision tree

- 解釋變數：MEWS score



- 解釋變數：Centile-based score



分析結果整理

- Random forest 分析結果如下表：

Model:	N	Mtry	Accuracy	Sensitivity	Specificity	5-Fold CV (Accuracy)
Random forest						
Original vital sign		2	0.79	0.65	0.88	0.77
MEWS score	45398	1	0.73	0.43	0.93	0.73
Centile-based score		1	0.68	0.35	0.91	0.68

備註：3 個 Random forest 模型的參數皆為 Number of trees：500，
Target node size：1，Variable important mode：impurity，Splitrule：gini。

- 綜合考量 Accuracy、Sensitivity 及 Specificity，Original vital sign 的模型表現最佳。
- 對於直接使用生命徵象數值做預測的模型，重要的生命徵象為脈搏、血壓。

結論與討論 — 目標一

針對發生非預期事件的病人，

1. 病人在發生事件前 48 小時的病況可以分為穩定及不穩定兩種狀態。根據分群的結果，我們可以訂出 MEWS 超過 3 的紀錄代表病況不穩定，但依據現有資料，並無法評估這樣的標準是否正確。

結論與討論 — 目標二

對於有發生非事件及沒有發生事件的病人，

1. 原始生命徵象、MEWS 分數及 Centile-based 分數之間，由於 MEWS 分數比原始生命徵象數值更符合實務操作的方便性，且其預測能力優於 Centile-based 分數，因此我們推斷 MEWS 為較好的早期預警系統。
2. 根據這筆資料，綜合各種模型的分析結果，可以訂出 MEWS 分數 2 是一個好的分界點。
3. 模型的建立需要定義出反應變數，我們將有發生非預期事件的病人，其所有量測紀錄都定為 event，雖然無法反應病人的真實病況變化情形，但根據文獻的分析方式以及現有資料所能提供的資訊，目前並沒有合理的反應變數界定方式。

結論與討論 — 目標二

4. 僅用呼吸、脈搏、血壓、體溫 4 個生命徵象預測病人是否會發生非預期事件，整體的預測表現約有 7 成的準確率。
5. 現有資料僅使用呼吸、脈搏、血壓、體溫 4 個生命徵象，但 Kyriacos 等人 (2011)³ 所提出的文獻回顧指出，現行許多早期預警系統有使用包括意識及血氧飽和度的生命徵象。未來的分析若能再加入此二者，應能使我們對病患狀況的掌握度更高，提升病況預測的準確度。

³Una Kyriacos, J Jelsma, and S Jordan. "Monitoring vital signs using early warning scoring systems: a review of the literature". In: *Journal of nursing management* 19.3 (2011), pp. 311-330.

Thank you for listening!