



# 內科PGY OSCE測驗題組品質分析

## —— 累計三年十二梯次學員測驗

王竹賢 ◆ 內科部重症醫學科／臨床技能訓練中心 主任

### 背景與目的

彰基自2011年起針對畢業後一般醫學訓練計畫醫師（PGY）輪訓內科三個月結束時舉行一次內科PGY OSCE，並自2013年至2014年新編寫內科PGY OSCE題組兩套（各十題），針對2014年起新訓PGY輪訓內科三個月期間進行內科PGY OSCE前後測。前後測藍圖相對應題組主題與核心能力類似但非重複，主旨在評核PGY在輪訓內科三個月的學習成效。每站結構為讀題兩分鐘，病史詢問身體檢查或技能操作十分鐘，回饋兩分鐘。每梯次學員約9-12人。我們嘗試分析小規模但多次舉行的OSCE，該如何進行OSCE測驗題組品質分析。

### 方法

自2014年8月起至2017年5月止，共針對12梯次學員舉行內科PGY OSCE前後測。因為每梯次學員數很少，及格標準採Modified Angoff & Global rating combined score由編寫教案專家群訂定。我們嘗試綜合12梯次學員成績做OSCE題組品質分析：每題信度、困難度與鑑別度分析。另外，我們也比較同一學員前後測成績的差異，探討內科PGY訓練的教學成效。

表一、前測藍圖

內科PGY OSCE前測藍圖						
	心	肺	感染	腎臟	其他	合計
病史詢問/身體檢查		意識變化 (林OO醫師)	發燒 (吳OO醫師)		體重下降 (李OO醫師)	3
醫病溝通/衛教	DVT (賴OO醫師)	肺炎 (邱OO醫師)	院內感染 (吳OO醫師)			3
判讀				血尿 (謝OO醫師)		1
醫療倫理					告知壞消息 (謝OO醫師)	1
臨床技能	氣道處理(王OO醫師)					1
EBM	EBM(廖OO醫師)					1
	合計					10

表二、後測藍圖

內科PGY OSCE後測藍圖							
	心	肺	感染	腎臟	神經	其他	合計
病史詢問/身體檢查		體重下降 (李OO醫師)	發燒 (吳OO醫師)		意識變化 (林OO醫師)		3
醫病溝通/衛教	DVT (賴OO醫師)	肺炎 (邱OO醫師)	院內感染 (吳OO醫師)				3
判讀				血尿 (謝OO醫師)			1
醫療倫理						告知壞消息 (林OO醫師)	1
臨床技能	氣道處理(王OO醫師)						1
EBM	EBM(廖OO醫師)						1
	合計						10

### ■ 及格標準設定

Modified Angoff & Global rating combined score: 學員分數採Modified Angoff method score與 Global rating score各占50%比重。

[Modified Angoff method score (100%) + Global rating score (60/100)] / 2

Global rating score: 由考官依學員表現憑直覺五等級打分，但考官說明會有特別說明心目中及格群考生下限為60分。請考官給予實際分數（0-100分）；差（0-54分）、待加強（55-64分）、普通（65-74分）、良好（75-84分）、優秀（85-100分）。



### ■ 難易度／鑑別度分析方法說明

根據題組總分的高低依序排列（註：題組總分：以題組滿分100分來轉換計算）

從最高分部分向下取總人數的27%為高分組，再從最低分部分向上取總人數的27%為低分組。分別計算高分組與低分組在此一試題的及格人數與百分比。

難易度（ $0 \leq \text{難易度} \leq 1$ ）=（高分組及格率+低分組及格率）/2

極難：0.34以下；難：0.35 ~0.59；中：0.6~.84；易：0.85以上。

鑑別度（ $-1 \leq \text{鑑別度} \leq 1$ ）=（高分組及格率-低分組及格率）

極佳：0.4以上；佳：0.3~0.39；尚可：0.2~0.29；差：0.19以下

### ■ 題組各站信度分析說明

考題內部之間一致性整體信度以 Cronbach's  $\alpha$  係數表示，其值愈高，表示考題間有相當高的一致性。計算 $\alpha$ 係數常見的公式是根據「平均問項間相關係數」（average inter-item correlation）的，其公式為

$$\alpha = \frac{k \bar{r}}{1 + (k - 1) \bar{r}}$$

其中K是問項的數目

$\alpha$ 係數值	可信程度的參考範圍
$\alpha \leq 0.3$ ：	不可信
$0.3 < \alpha \leq 0.4$ ：	初步研究，勉強可信
$0.4 < \alpha \leq 0.5$ ：	稍微可信
$0.5 < \alpha \leq 0.7$ ：	可信
$0.7 < \alpha \leq 0.9$	很可信
$0.9 < \alpha$ ：	十分可信

### 結果

我們共累計三年12梯次內科PGY OSCE前測127人與後測118人。前後測題組每題信度、困難度與鑑別度品質分析結果見（表3-5）：前後測對應題組，學員前後測成績比較見（圖1），學員成績後測進步率比較見（圖2）。

表三、前後測題組信度分析：前後測對應題組信度分析，除一站（體重減輕）較不理想，其餘信度都達到可信到很可信程度。

前後測題組信度分析

站名	前測信度	後測信度
意識不清	0.667	0.699
血尿	0.606	0.539
困難氣道	0.637	0.615
DVT	0.715	0.744
發燒	0.663	0.676
院內感染	0.671	0.644
體重減輕	0.506	0.299
告之壞消息	0.671	0.683
EBM	0.684	0.660
肺炎	0.619	0.712

表四、前測題組鑑別度難易度分析

前測題組鑑別度難易度分析

站名	及格分數	高分組及格率(%)	低分組及格率(%)	鑑別度	判定	難易度	判定
意識不清	60	94	47	0.47	極佳	0.71	中
血尿	55	94	74	0.20	尚可	0.84	中
困難氣道	52.5	79	41	0.38	佳	0.60	中
DVT	60	91	41	0.50	極佳	0.66	中
發燒	55	71	53	0.18	差	0.62	中
院內感染	55	50	39	0.11	差	0.45	難
體重減輕	67.5	97	56	0.41	極佳	0.71	中
告之壞消息	60	100	62	0.38	佳	0.81	中
EBM	60	79	41	0.38	佳	0.60	中
肺炎	64.4	22	13	0.09	差	0.18	難
平均				0.31		0.618	

及格分數: Angoff & global rating combined score



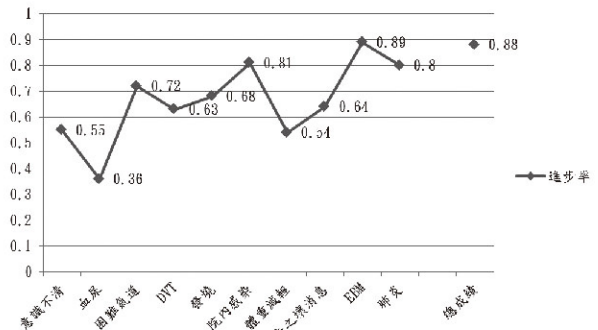


表五、後測題組鑑別度難易度分析

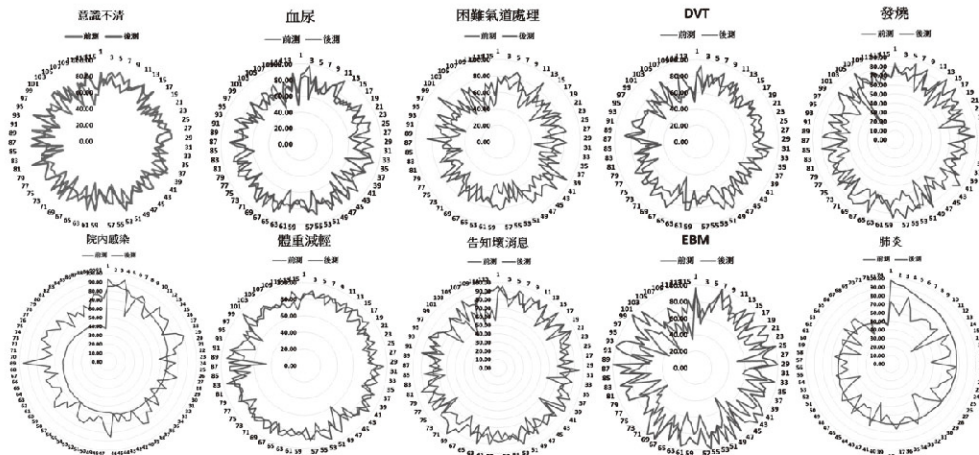
後測題組鑑別度難易度分析							
站名	及格分數	高分組及格率(%)	低分組及格率(%)	鑑別度	判定	難易度	判定
意識不清	60	97	69	0.28	尚可	0.83	中
血尿	55	97	69	0.28	尚可	0.83	中
困難氣道	52.5	97	59	0.38	佳	0.78	中
DVT	60	90	50	0.40	佳	0.70	中
發燒	55	97	69	0.28	尚可	0.83	中
院內感染	55	66	31	0.35	佳	0.49	難
體重減輕	55	100	100	0.00	差	1.00	易
告之環消息	55	100	100	0.00	差	1.00	易
EBM	60	100	97	0.03	差	0.99	易
肺炎	64	75	15	0.60	極佳	0.45	難
平均				0.26		0.79	

及格分數: Angoff & global rating combined score

後測進步率



表二、前後測對應題組，學員成績後測進步率(進步：後測成績/前測成績 > 1)



圖一、前後測對應題組，學員前後測成績比較雷達圖。

### 討論

及格成績採用Modified Angoff & Global rating combined score，Modified Angoff method score與 Global rating score各占50%比重。會這樣設計主要是考慮每站評分表細項也許無法反映學員測驗主題全面向能力，而由專家考官評分Global rating納入總分計算，也許能補足此部分顧慮，但客觀性可能因此降低，因為PGY OSCE本質為教學型OSCE，教學與評量並重，客觀性是希望有的品質但非全部，如果評分表設計不夠完美時，本次分析採用的combined score也許有互補之處。

前後測對應題組鑑別度難易度比較，發現後測題組平均鑑別度下降及平均難易度上升（趨易），學員總成績後測進步率達88%。這些數據變化應是反映PGY醫師在內科三個月訓練成效，而非反映後測難易度。

前後測對應題組，比較學員成績後測進步率，發現部分題組進步率偏低。

例如意識不清病史詢問站、血尿判讀站、體重減輕病史詢問站。這部分訊息可供一般內科示範中心安排核心教學時參考加強。