

以傾向分數配對法評估 糖尿病論質計酬方案之成效

林文德^{1,*} 謝其政^{2,3} 邱尚志⁴
吳慧俞¹ 黃一展⁵

目標：健保局於2001年11月推行糖尿病論質計酬方案，過去研究雖呈正向的成效，但未考慮加入方案者選擇偏差的問題，因此本研究目的旨在以糖尿病可避免住院為品質結果指標，並以傾向分數法配對產生可比較之配對組，以釐清該方案之成效。**方法：**擷取2002-2003年承保抽樣歸入檔中因糖尿病至門診就診並申報檢查者前後一年之就醫資料，藉由傾向分數(PS, propensity score)配對法，先依病患各項特性建構加入方案之機率(即PS)模型，再自未加入方案者(即對照組，共6,855人)中，以1:1的方式配對產生與加入方案者(即介入組，共647人)加入機率相當之配對組樣本(647人)，之後再以廣義估計方程式比較組間可避免住院勝算之比值。**結果：**對照組在前一年之病患特性、就醫情形、檢查申報比例及住院等變項上與介入組有顯著差異，但配對組與介入組各變項之差異均不顯著。方案實施後，介入組之各項檢查申報比例均顯著高於對照組與配對組，但其糖尿病相關可避免住院率增加幅度高於對照組(勝算比1.01, 95%信賴區間為0.98-1.04)，卻低於配對組(勝算比0.98, 95%信賴區間為0.94-1.02)，惟均未達顯著意義。**結論：**參加論質計酬方案者與未參加者之特性及過去檢查申報比例並不相同。經傾向分數配對法產生與介入組條件相當之配對組，並據此評估論質計酬方案之成效，與傳統上以未參加者為對照組的比較方式有不同的結果。以傾向分數法做為政策評估的工具，或可提供健康政策成效評估的另一種觀點。(台灣衛誌 2010；29(1)：54-63)

關鍵詞：糖尿病、論質計酬、傾向分數配對法、可避免住院

前 言

健保局於2001年11月推行醫療給付改善方案(以下簡稱論質計酬方案)，並將糖尿病

列為其中之一[1]。該方案在推行之初，主要著重於過程面的品質，如定期空腹血糖、糖化血紅素等多項檢查的執行，卻缺乏反應該計畫成效之結果面指標。學者指出，衛生政策之品質改善成效評估宜以結果(outcome)為導向[2]。因此，以結果面品質做為評估糖尿病論質計酬方案應有其必要。

可避免住院(avoidable hospitalization)是門診照護品質的重要結果指標之一，其主要的邏輯為適當且即時的門診照護可以避免不必要之住院。該指標最初由Weissman等人於1992年提出[3]，後經美國健康照護研究暨品質局(Agency for Healthcare Research and Quality, AHRQ)定義出門診照護敏感

¹長榮大學醫務管理學系暨碩士班

²亞東技術學院醫務管理學系

³國立台灣大學醫療機構管理研究所

⁴亞洲大學醫務管理學系

⁵美國佛羅里達大學醫學院流行病學及健康政策研究學系

*通訊作者：林文德

聯絡地址：台南縣歸仁鄉長榮路一段396號

E-mail: vincelin@mail.cjcu.edu.tw

投稿日期：98年9月3日

接受日期：99年1月15日



疾病(ambulatory care sensitive conditions, ACSCs), 或預防性品質指標(prevention quality indicator, PQI)[4], 其中糖尿病短期併發症住院、糖尿病長期併發症住院、糖尿病控制不良住院及糖尿病患下肢截肢四項, 均被列為可避免的住院之一。研究指出, 糖尿病的住院有7%是可避免的[5]。故以可避免住院率為指標, 應可評估糖尿病論值計酬方案的結果面品質。

回顧我國糖尿病論質計酬成效評估之相關文獻, 多數均得到正面的結果[6-9], 惟過去雖有研究以可避免住院做為成效評估的指標[6], 但多以未加入論質計酬者為對照組, 即或以差異中的差異法(difference in difference)克服兩組可能的基礎點(baseline)差異所造成的影響, 仍無法解決對照組與介入組不可比較(non-comparable)的問題。亦即, 加入計畫者可能是檢查等醫囑遵從性較佳, 或對病情控制動機較強之病患, 故加入計畫後可避免住院較少, 可能僅是反映加入者與對照組在影響糖尿病控制之相關特性有所不同, 而非計畫本身所帶來的成效, 如此可能產生選擇偏誤(selection bias)的問題。晚近已有國內學者對於病患選擇偏誤的問題進行研究, 結果發現參加論質計酬方案的醫師有篩選病患的情形[10], 因此該方案的成效實有必要再予以檢驗。

近年來, 愈來愈多的衛生政策評估採用傾向分數法(propensity score, PS)來克服觀察性研究的選擇偏誤問題[11-13], 甚至在隨機控制試驗不易進行的情況下, 以PS將觀察性研究模擬為隨機控制的情境, 而取得可比較的介入組與控制組[14]。糖尿病相關的介入計畫, 亦有採用此方法者[15]。因此, 在論質計酬方案成效難以採隨機控制試驗之方式加以驗證下, 本研究旨在透過傾向分數法, 由未加入論質計酬之糖尿病患者中配對選出與加入者(介入組)加入機率相當的配對組, 以模擬隨機控制的情形, 並以差異中的差異法評估論質計酬方案能否減少糖尿病相關之可避免住院。

材料與方法

研究設計及研究對象

本研究採用類實驗之雙組前後測設計, 自國家衛生研究院全民健康保險研究資料庫(<http://w3.nhri.org.tw/nhird//index.php>) 20萬的承保抽樣歸人檔中擷取2002年至2003年因糖尿病至門診就醫之病患做為研究對象。介入組個案係依糖尿病論質計酬方案之規定[1], 將費用清單之診察項目代碼或醫令清單之醫令代碼申報完整性門診初診照護費(代碼為P1401C)認定為介入組病患(共680人), 並扣除前一年就醫申報資料未有糖尿病診斷(即國際疾病分類臨床修正版第九版(ICD-9-CM)前三碼為250)或進行相關之檢查檢驗者, 實際共得647位介入組病患。介入組病患第一次申報完整性門診初診照護費之就醫日期, 被視為加入論質計酬方案之日期, 並以之計算其加入計畫前、後年的病患特性及就醫相關資料。

對照組的產生, 係先排除2004年前加入糖尿病論質計酬方案者, 並以2002-2003年為研究時間窗, 選取該段時間窗內門診申報資料中至少有四次主、次診斷為糖尿病, 且2002年及2003年至少各有一次就醫主、次診斷為糖尿病者, 做為本研究之對照組(如圖1所示)。為了在對照組中取得類似介入組病患之加入計畫時間, 以做為計算加入前、後年的病患特性及就醫相關資料, 本研究以對照組門診就醫資料中費用最高者做為指標門診就醫(index ambulatory visit)。因此, 如圖1中指標門診就醫日期為2002年11月1日, 該時間點的病患特性及其前一年的就醫情形, 即可加以計算。經過此處理後, 介入組的加入日期與對照組的指標門診就醫日期在2002-2003年的分布即趨於一致。

配對組的產生, 係以加入計畫與否為依變項, 並以加入計畫日期或指標門診就醫日期前一年之各項病患特性及就醫相關情形做為自變項(自變項的內容詳如下節所述), 以羅吉斯迴歸模型估算糖尿病患者加入論質計酬方案之機率(即傾向分數)。隨後, 藉由Parsons提出之傾向分數配對巨集程式[16]以

一比一配對為原則，先依介入組之傾向分數選取對照組中傾向分數與之差距在0.00001之內者做為配對組，再逐步放寬條件至傾向分數差距在0.1之內者做為配對組，如此可以避免自對照組中選取出配對組時，因傾向分數的差距限制而減少了介入組與配對組的對數。

變項之測量

本研究估計樣本加入計畫之機率(傾向分數)所投入的自變項包含以下幾種：

- (一) 病患特性：包括性別、年齡、病情分階、新舊個案與否，其中病情分階採 Gonnella 等人於1984年提出之糖尿病併發症發生的進程，將糖尿病由輕至重分為四個階段，第一階段為單純糖尿病，第二階段為合併一個或以上的身體系統疾病，第三階段為合併糖尿病昏迷或酮酸中毒，第四階段為死亡[17]。由於本研究排除死亡的病患，故僅使用一至三階的分階。另外新舊個案係以病患加入計畫日期或指標門診就醫日期前二年是否有糖尿病主、次診斷之就醫紀錄為判斷依據，有則為舊個案，反之為新個案；
- (二) 就醫情形：包括是主要就醫率(usual provider continuity)、加入計畫前或指標門診就醫日期前的糖尿病門診就醫次數、主要就醫地區的都市化程度、主要就醫院所之權屬別及評鑑等級，其中主要就醫率係以糖尿病就醫次數最多之院所就醫次數占所有就醫次數之比例，做為衡量就醫持續性的指標[18]；先前的糖尿病門診就醫次數係反應其先前就醫的密集度；就醫地區之都市化程度係以病人主要就醫院所做為地區別判定，並以劉介宇等人的台灣地區都市化分層方法加以分類[19]。該分類方法將台灣地區369個鄉鎮市區依人口密度、大專教育以上人口比例、農業人口比例、65歲以上人口比例及每十萬人醫師數五個因素分為七個集群，本研究為求精簡，再

將其高度及中度都市化地區合併為「都市地區」、新興及一般地區合併為「城鎮地區」、高齡化及農業化地區合併為「鄉村地區」、偏遠地區仍保留為「偏遠地區」；主要就醫院所之權屬別及評鑑等級則以糖尿病患者糖尿病就醫次數最多之院所權屬別及評鑑等級為認定的依據，權屬別分公立醫院、法人醫院及私立醫院，評鑑等級則分為醫學中心、區域醫院、地區醫院及基層院所四級。

- (三) 檢查情形：即糖尿病論質計酬方案中，列入必要檢查、檢驗的項目，包括：1.空腹血漿葡萄糖、2.糖化血紅素、3.血清總膽固醇、中性脂肪、高密度及低密度脂蛋白膽固醇、4.血清肌酸酐、5.血清麩胺酸丙酮酸轉胺基酶、6.尿液分析及7.眼睛等七項檢查或檢驗之有無。本研究係以研究對象的健保資料中判斷其是否接受上述檢查，故反映的是上述檢查是否有加以申報，不一定是真實的檢查情形；
- (四) 加入方案前一年糖尿病相關可避免住院之有無：糖尿病相關可避免住院依據美國AHRQ的定義，包括：1.短期併發症住院(ICD-9-CM診斷碼為250.1-250.3)、2.糖尿病長期併發症住院(ICD-9-CM診斷碼為250.4-250.9)、3.糖尿病控制不良住院(ICD-9-CM診斷碼為250.0)、及4.糖尿病患下肢截肢(ICD-9-CM手術碼為84.1)，並排除國際疾病分類為兒科(主要診斷類別為15)、產科(主要診斷類別為14)或外傷就醫者[4]。

另在論質計酬成效評估部分，則是以介入組個案的加入方案日期，或配對組及對照組之指標門診就醫日期後一年之糖尿病相關住院有無做為標的，糖尿病相關可避免住院的定義同前段(四)所述。

分析方法

本研究以SAS 9.1.3統計軟體進行資料處理與分析。介入組、配對組與對照組之病患特性、就醫情形及各項檢查等，採次數分配(百分比)、平均數(標準差)呈現分布

結 果

情形，並以t檢定(非類別變項)以及卡方檢定(類別變項)進行實驗與對照、實驗與配對的兩組間差異比較，並採雙尾方式加以檢定，統計顯著水準 α 設為0.05。計算傾向分數採用的羅吉斯模型以c統計值呈現其鑑別度(discrimination)，而以Hosmer-Lemeshow的配適度檢定其正確度(calibration)[20]。

此外，針對組別間可避免糖尿病相關住院之前後差異比較，由於研究對象在指標門診日期前、後測得之住院並非獨立，故以廣義估計方程式(GEE, generalized estimating equation)建構模式，並以羅吉函數(logit function)做為連結函數，進行兩組間前、後住院率之差異中差異分析。羅吉函數如下：

$$\log\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 * I + \beta_2 * T + \beta_3 * (I * T) + e$$

其中，P代表發生可避免住院的機率，I代表介入組(I=1)或對照組及配對組(I=0)，T代表後測(T=1)或前測(T=0)，由於

$$\beta_2 + \beta_3 = \log\left(\frac{P_{(I=1, T=1)}}{1 - P_{(I=1, T=1)}}\right) - \log\left(\frac{P_{(I=1, T=0)}}{1 - P_{(I=1, T=0)}}\right)$$

表示介入組在指標門診就醫日期前、後的可避免住院對數勝算比(log odds ratio)，而

$$\beta_2 = \log\left(\frac{P_{(I=0, T=1)}}{1 - P_{(I=0, T=1)}}\right) - \log\left(\frac{P_{(I=0, T=0)}}{1 - P_{(I=0, T=0)}}\right)$$

表示對照組(或配對組)在指標門診就醫日期前、後的可避免住院對數勝算比，因此，兩者之差 β_3 即為差異中之差異，如以對數的運算邏輯觀之，實際上即為兩組的勝算比之比值。

為呈現傾向分數配對後在分析上與傳統迴歸模型的差異，本研究將上述的迴歸模型分為：一、「傳統模型」，包括介入組與對照組的研究對象，除了以組別(T，參考組為對照組)、前、後測(I，參考組為前測)、及此兩者的交互作用項(T*I，即差異中的差異變項)為模型的自變項外，並納入其他可能影響糖尿病相關住院的控制變項，包括病患特性及就醫情形；二、「傾向分數模型」，即包括介入組與配對組的研究對象。

表一顯示，未加入論質計酬方案的對照組在年齡上顯著較介入組高(+2.0歲， $p < 0.001$)、病情較嚴重及較輕微的比例顯著較高、新個案比例顯著較高(+3.0%， $p = 0.022$)、指標門診就醫日期前的糖尿病就醫次數顯著較低(-6.8次， $p < 0.001$)、主要就醫院所以醫學中心及基層院所為主(25.6%及27.1%， $p < 0.001$)、申報空腹血漿葡萄糖、糖化血紅素、空腹血脂及眼睛檢查的申報比例顯著較低($p < 0.001$)、前一年的糖尿病相關住院顯著較高(+11.1%， $p < 0.001$)。

以羅吉斯迴歸建立之論質計酬方案加入機率估計模型，其c統計值為0.719，在可接受的範圍，而Hosmer與Lemeshow配適度檢定則顯示其配適度佳(卡方值為9.33， $p = 0.32$)。依此模型估計之傾向分數配對產生之配對組共647人，在各自變項的分布上均與介入組沒有顯著的差異(詳如表一)。

方案實施後，介入組與對照組申報檢查比例均有所提高，但配對組提高的情形並不明顯。比較三組的檢查申報比例差異，發現介入組的病患在七種檢查的申報比例上均顯著高於另兩組，顯示本方案在提高介入組的檢查申報比例上，確實頗為顯著。此外，三組的糖尿病相關住院比較，則是對照組顯著高於介入組(12.9% VS 9.0%， $p < 0.001$ ，詳如表二)。

以GEE模式分析介入組與對照組在指標門診前、後可避免住院對數勝算比的差異(即傳統模型，詳如表三)，結果顯示將可能的影響因素加以控制後，介入組在指標門診前的可避免住院勝算低於對照組(勝算比為0.94，95%信賴區間為0.92至0.96)，對照組在指標門診後的可避免住院勝算上升(勝算比為1.02，95%信賴區間為1.01至1.03)，雖然介入組的可避免住院勝算增加幅度大於對照組，但未達顯著意義(勝算比之比值為1.01，95%信賴區間為0.98至1.04)。同樣的分析運用在介入組與配對組的比較(即傾向分數模型)，結果則有所不同，經過傾向分數配對，兩組的指標門診前住院勝算已無差

表一 論質計酬實施前一年糖尿病患者之特性、就醫情形、檢查檢驗及可避免住院分布情形表

變項	介入組(n=647)		配對組(n=647)		p值	對照組(n=6,855)		p值
	人數	百分比	人數	百分比		人數	百分比	
病患特性								
年齡(歲) [#]	58.2	(12.7)	58.1	(12.9)	0.908	60.2	(13.2)	<0.001
性別(女性)	332	51.3	357	55.2	0.164	3,500	51.1	0.910
病情					0.848			0.021
第1階	366	56.6	375	58.0		4,089	59.7	
第2階	249	38.5	239	36.9		2,307	33.7	
第3階	32	5.0	33	5.1		459	6.7	
糖尿病新舊個案(新個案)	57	8.8	58	9.0	0.922	811	11.8	0.022
就醫情形								
主要就醫率 [#]	78.6	(20.5)	76.5	(21.5)	0.077	77.6	(21.3)	0.249
糖尿病就醫次數 [#]	36.8	(23.1)	37.5	(27.6)	0.623	30.0	(24.0)	<0.001
就醫地區					0.245			0.698
都市地區	447	69.9	450	69.6		4,868	71.0	
城鎮地區	149	23.0	163	25.2		1,475	21.5	
鄉村地區	22	3.4	16	2.5		243	3.5	
偏遠地區	29	4.5	18	2.8		269	3.9	
權屬別					0.491			0.241
公立醫療機構	218	33.7	198	30.6		2,442	35.6	
法人醫療機構	174	26.9	181	28.0		1,940	28.3	
私立醫療機構	255	39.4	268	41.4		2,473	36.1	
評鑑等級					0.746			<0.001
醫學中心	108	16.7	112	17.3		1,752	25.6	
區域醫院	234	36.2	216	33.4		1,745	25.5	
地區醫院	162	25.0	165	25.5		1,500	21.9	
基層診所	143	22.1	154	23.8		1,858	27.1	
檢查情形								
空腹血漿葡萄糖	630	97.4	627	96.9	0.617	6,136	89.5	<0.001
糖化血紅素	439	67.9	439	67.9	1.000	2,712	39.6	<0.001
空腹血脂	446	68.9	453	70.0	0.673	3,944	57.5	<0.001
血清肌酸酐	408	63.1	430	66.5	0.201	4,062	59.3	0.059
血清麩胺酸丙酮酸轉胺基酶	350	54.1	343	53.0	0.696	3,734	54.5	0.855
尿液分析	63	9.7	68	10.5	0.645	541	7.9	0.099
眼睛檢查	188	29.1	189	29.2	0.951	1,167	17.0	<0.001
住院	39	6.0	39	6.0	1.000	758	11.1	<.001

註：介入組和配對組、介入組和對照組分別進行卡方檢定(類別變項)或t檢定(非類別變項)，[#]為平均值及(標準差)。

異，但介入組的糖尿病相關住院勝算增加幅度則小於配對組(勝算比之比值為0.98，95%信賴區間為0.94至1.02)，惟亦未達顯著。

討 論

本研究發現，參加論質計酬方案的介入組患者較對照組年輕，在未參加方案前的多項檢查申報比例顯著高於未參加者，而其糖尿病可避免住院率則低於對照組。如以傳統迴歸模型進行差異中的差異分析，雖然介入

表二 論質計酬實施後一年糖尿病患者之檢查及可避免住院分布情形

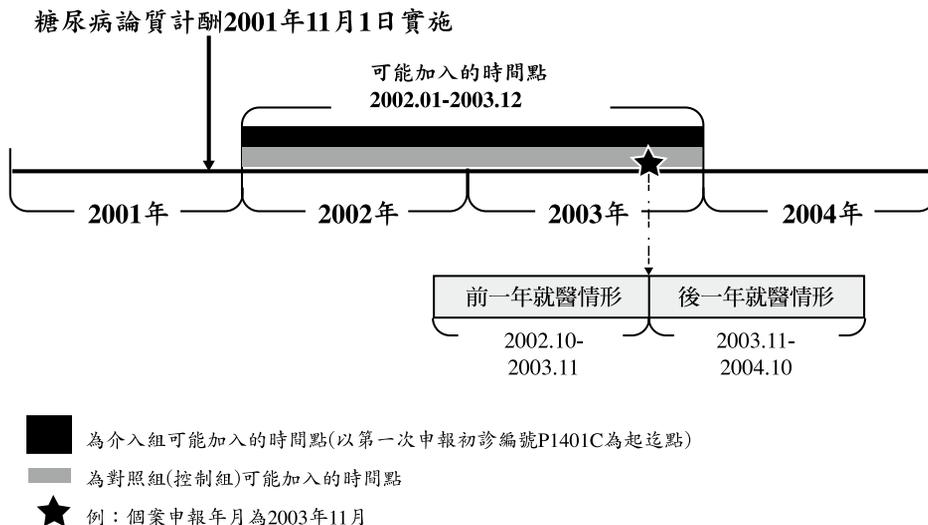
變項	介入組(n=647)		配對組(n=647)		p值	對照組(n=6,855)		p值
	人數	百分比	人數	百分比		人數	百分比	
檢查情形								
空腹血漿葡萄糖	646	99.9	609	94.1	<0.001	6,283	91.7	<0.001
糖化血紅素	622	96.1	388	60.0	<0.001	3,234	47.2	<0.001
空腹血脂	627	96.9	438	67.7	<0.001	4,207	61.4	<0.001
血清肌酸酐	620	95.8	433	66.9	<0.001	4,399	64.2	<0.001
血清麩胺酸丙酮酸轉胺基酶	619	95.7	396	61.2	<0.001	4,068	59.3	<0.001
尿液分析	535	82.7	58	9.0	<0.001	569	8.3	<0.001
眼睛檢查	559	86.4	136	21.0	<0.001	1,179	17.2	<0.001
住院	58	9.0	71	11.0	0.228	882	12.9	<0.001

註：介入組和配對組、介入組和對照組分別進行卡方檢定

表三 傳統模型及傾向分數模型分析論質計酬實施前後可避免住院之勝算比

變項	傳統模型(N=7,502)			傾向分數模型(N=1,294)				
	勝算比	95%信賴區間	p值	勝算比	95%信賴區間	p值		
差異中的差異								
介入組(vs.對照(或配對)組)	0.94	0.92	0.96	<.0001	0.98	0.95	1.01	0.123
後測(vs.前測)	1.02	1.01	1.03	<.0001	1.06	1.02	1.09	<.0001
介入組*後測(vs.對照(或配對)組 前、後測住院勝算比)	1.01	0.98	1.04	0.432	0.98	0.94	1.02	0.222
病患特性								
年齡(歲)	1.00	1.00	1.00	<.0001	1.00	1.00	1.00	0.001
性別(女性)	1.03	1.02	1.04	<.0001	1.03	1.00	1.05	0.029
病情第二階(vs.第一階)	1.12	1.10	1.13	<.0001	1.10	1.08	1.13	<.0001
病情第三階(vs.第一階)	1.17	1.14	1.21	<.0001	1.10	1.03	1.16	0.003
新個案(vs.舊個案)	1.05	1.03	1.07	<.0001	1.06	1.02	1.11	0.006
就醫特性								
主要就診率	0.89	0.87	0.92	<.0001	0.87	0.82	0.93	<.0001
前一年就醫次數	1.00	1.00	1.00	<.0001	1.00	1.00	1.00	<.0001
城鎮地區(vs.都市地區)	1.00	0.99	1.02	0.574	1.01	0.98	1.04	0.489
鄉村地區(vs.都市地區)	1.02	0.99	1.05	0.180	0.99	0.93	1.05	0.733
偏遠地區(vs.都市地區)	1.01	0.98	1.03	0.679	0.93	0.89	0.98	0.006
法人機構(vs.公立機構)	1.01	1.00	1.02	0.038	1.00	0.98	1.03	0.019
私立機構(vs.公立機構)	1.01	1.00	1.02	0.144	1.00	0.98	1.03	0.751
區域醫院(vs.醫學中心)	1.03	1.01	1.05	0.001	1.06	1.02	1.09	0.001
地區醫院(vs.醫學中心)	0.99	0.97	1.01	0.184	1.03	0.99	1.07	0.149
基層診所(vs.醫學中心)	0.97	0.95	0.99	0.001	1.02	0.98	1.06	0.305

註：傳統模型係以介入組與對照組為對象進行分析，傾向分數模型則以介入組與配對組為對象進行分析。



圖一 研究時間窗示意圖

組個案在加入方案後一年的糖尿病可避免住院率仍低於對照組，但其增加幅度卻高於對照組，如此反顯示計畫的成效不彰。本研究經由傾向分數產生之配對組，在計畫前一年，不論人口特性、就醫情形、檢查申報比例及糖尿病可避免住院率，均與介入組無顯著差異，透過差異中的差異分析，則顯示介入組的糖尿病可避免住院率增加幅度低於配對組，雖未達顯著意義，但計畫的成效仍屬正面。

過去的研究皆顯示介入組參加方案後，相關的檢查申報比例均顯著提升[6-9]，但申報比例的提升僅代表過程面品質的提升，再加上申報資料與實際作為是否一致並無從求證，故需有結果面的品質加以驗證。惟過去研究即使以可避免住院做為品質結果的指標，但多以未加入計畫者做為對照組，比較的結果尚有疑慮。由本研究可以發現，未加入論質計酬者原有的糖尿病可避免住院即顯著較高，如未配對而僅比較介入組與對照組，反而出現對照組糖尿病可避免住院率增加幅度小於介入組的情形(雖未達顯著)。因此未適當選擇配對組以為比較，可能使方案的成效評估產生相反的結果。本研究以傾向

分數配對後的配對組做為比較的對象，並經差異中的差異法控制兩組指標門診前的基礎線狀態，結果顯示論質計酬仍具有正向的成效，雖亦未達顯著，但應可做為未來研究的參考。

本研究所採用傾向分數方式，在加入方案者較少或事件(例如住院)發生數較少的情形下，可減低多變項模型參數估計不具效率的問題[21]。惟本研究所採用健保資料庫無法取得其他影響住院的因子，如教育程度、收入、病患的動機、遵從性等變項，故本研究實驗與配對兩組樣本是否能完全達到類似隨機分配而排除相關干擾因素，仍有待研究。未來研究可以問卷調查方式探知是否有健保資料庫無法含括但可能影響住院的因素，應能進一步證實本研究的結果。此外，本研究僅比較加入方案一年內糖尿病可避免住院的有無，對於需要較長時間才能顯現的成效，適當性或有疑慮。目前糖尿病論質計酬方案以糖化血色素做為糖尿病控制良好與否的指標，該指標確較能反映短期的血糖控制好壞之影響，惟該指標之數值並非由客觀的第三者蒐集，且亦非健保局定期對外公開的資料，故以其做為政策評估的指標有一定

的困難。本研究以糖尿病可避免住院為指標，雖然有短期內成效不顯的缺點，惟未來應可拉長評估期間而顯現計畫的成效。

綜上，本研究發現加入論質計酬方案的糖尿病患者雖在各項檢查的申報比例達到較高的水準，並因此而得到論質計酬的「酬勞」，但這些患者在未加入前，即可能有較高的檢查遵從性，因此在加入後要達到「論質」的品質標準可能較為容易。國外亦有論質計酬的研究發現，得到較大獎勵者可能是治療的病患先前健康狀況即較好，而非因病患的健康狀況進步最多[22]。因此國外研究者即建議，應針對目前糖尿病控制情形較差者設計較強的報酬誘因，而非只以達到品質目標為誘因[23]。未來如能考量品質進步的幅度而提高給付，或可使論質計酬發揮更大的成效。

致 謝

本研究部分資料來源為衛生署中央健康保險局提供、財團法人國家衛生研究院管理之『全民健康保險研究資料庫』，特此誌謝。文中任何闡釋或結論並不代表衛生署中央健康保險局、或財團法人國家衛生研究院之立場。

參考文獻

1. 中央健康保險局：全民健康保險糖尿病醫療服務改善方案試辦計畫支付標準。第一版。台北：中央健康保險局，2001。
2. Brook RH, McGlynn EA, Shekelle PG. Defining and measuring quality of care: a perspective from US researchers. *Int J Qual Health Care* 2000;**12**:281-95.
3. Weissman JS, Gatsonis C, Epstein AM. Rates of avoidable hospitalization by insurance status in Massachusetts and Maryland. *JAMA* 1992;**268**:2388-94.
4. Agency for Healthcare Research and Quality. AHRQ quality indicators-prevention quality indicators overview, 2004. Available at: http://www.qualityindicators.ahrq.gov/pqi_overview.htm. Accessed January 3, 2008.
5. Niefeld NR, Braunstein JB, Wu AW, Saudek CD, Weller WE, Anderson GF. Preventable hospitalization among elderly Medicare beneficiaries with type 2 dia-

betes. *Diabetes Care* 2003;**26**:1344-9.

6. 李玉春：全民健保支付與管理制度對病人就醫品質之影響：以可避免之住院為例。行政院衛生署九十五年度科技研究計畫，計畫編號DOH95-HI-1002。台北：行政院衛生署，2006。
7. 李玉春：建立全民健保以共同照護模式為基礎的糖尿病人疾病管理計畫先導性研究(III)。行政院衛生署九十一委託計畫，計畫編號DOH91-NH-1002。台北：行政院衛生署，2002。
8. 侯佳雯：「糖尿病醫療服務改善方案」試辦計畫對第二型糖尿病人醫療資源利用之影響。台北：台灣大學醫療機構管理研究所碩士論文，2004。
9. 李待弟：全民健康保險糖尿病醫療給付改善方案初步影響評估。台北：台灣大學衛生政策與管理研究所碩士論文，2006。
10. Chang RE, Lin SP, Swei SC. Patient selection under a performance-based care improvement program. In: *Proceedings of the First Annual Meeting on Center of Health Insurance Research*, 2008. Taipei, Taiwan: Center of Health Insurance Research, 2008.
11. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;**70**:41-55.
12. D'Agostino RB Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med* 1998;**17**:2265-81.
13. Fu AZ, Dow WH, Liu GG. Propensity score and difference-in-difference methods: a study of second-generation antidepressant use in patients with bipolar disorder. *Health Serv Outcomes Res Method* 2007;**7**:23-38.
14. Tarakji KG, Brunken R, McCarthy PM, et al. Myocardial viability testing and the effect of early intervention in patients with advanced left ventricular systolic dysfunction. *Circulation* 2006;**113**:230-7.
15. Barnett TE, Chumbler NR, Vogel WB, Beyth RJ, Qin H, Kobb R. The effectiveness of a care coordination home telehealth program for veterans with diabetes mellitus: a 2-year follow-up. *Am J Manag Care* 2001;**12**:467-74.
16. Parsons LS. Reducing bias in a propensity score matched-pair sample using greedy matching techniques. In: *Proceedings of the Twenty-Sixth Annual SAS® Users Group International Conference*. Cary, NC: SAS Institute Inc, 2001.
17. Gonnella JS, Hornbrook MC, Louis DZ. Staging of disease. A case-mix measurement. *JAMA* 1984;**251**:637-44.
18. Saultz JW, Lochner J. Interpersonal continuity of care and care outcomes: a critical review. *Ann Fam Med*



- 2005;**3**:159-66.
19. 劉介宇、洪永泰、莊義利等：台灣地區鄉鎮市區發展類型應用於大型健康調查抽樣設計之研究。健康管理學刊 2006；**4**：1-22。
 20. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. Clin Chem 2008;**54**:17-23.
 21. Stümer T, Joshi M, Glynn RJ, Avorn J, Rothman KJ, Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. J Clin Epidemiol 2006;**59**:437-47.
 22. Rosenthal MB, Frank RG, Li Z, Epstein AM. Early experience with pay-for-performance: from concept to practice. JAMA 2005;**294**:1788-93.
 23. Rosenthal MB, Dudley RA. Pay-for-performance: will the latest payment trend improve care? JAMA 2007;**297**:740-4.

An effectiveness evaluation of a pay-for-performance program for diabetes based on the propensity score matching method

WENDER LIN^{1,*}, CHI-JENG HSIEH^{2,3}, SHANG-JYH CHIOU⁴
HUI-YU WU¹, I-CHAN HUANG⁵

Objectives: This study aimed to evaluate the pay-for-performance (P4P) program for diabetes by using the propensity score method to construct a comparable matching group, and by using the frequency of diabetes-related avoidable hospitalizations to assess outcome. **Methods:** In order to construct the intervention group, we selected diabetic cases (N=647) from the NHI beneficiaries claim data from 2002 to 2003. Using the propensity score, we selected the matched group (N=647) from the control group (N=6,855) by matching their characteristics and covariates relating to the probability of their participation in the program. Subsequently, we took a generalized equation estimate (GEE) approach with logit link to compare the likelihood of avoidable hospitalization among groups. **Results:** The distributions of predictors such as patient characteristics, prior ambulatory care utilization, and the rates of laboratory tests and hospitalizations were significantly different between the intervention group and the control group while there were no significant differences between the intervention group and the matched group. Although the proportion of laboratory tests was higher in the intervention group, after the P4P program this group had a higher likelihood of avoidable hospitalization than did the control group with an odds ratio (OR) of 1.01 (95% C.I.: 0.98-1.04) while having a lower likelihood than the matched group with an OR of 0.98 (95% C.I.: 0.94-1.02). Neither difference was significant. **Conclusions:** The characteristics of the diabetics who participated in the P4P were different from those who did not. After using propensity scores to create a matched group with the same probability as those who participated in the P4P, we found that the results were different from comparisons made with traditional methods. The propensity score method may provide another insight into understanding the effectiveness of a particular health policy. (*Taiwan J Public Health. 2010;29(1):54-63*)

Key Words: diabetes, pay-for-performance, propensity score matching, avoidable hospitalization

¹ Department of Health Care Administration, Chang Jung Christian University, No.396, Sec.1, Chang Jung Rd., Kway Jen, Tainan, Taiwan, R.O.C.

² Department of Health Care Administration, Oriental Institute of Technology, Taipei, Taiwan, R.O.C.

³ Graduate institute of Health Care Organization Administration, National Taiwan University, Taipei, Taiwan, R.O.C.

⁴ Department of Health Care Administration, Asia University, Taichung, Taiwan, R.O.C.

⁵ Department of Epidemiology and Health Policy Research, College of Medicine, University of Florida, U.S.A.

*Correspondence author. E-mail: vincelin@mail.cjcu.edu.tw

Received: Sep 3, 2009 Accepted: Jan 15, 2010

評論：以傾向分數配對法評估糖尿病論質計酬方案之成效

政策評估通常需要較長的觀察成效期間，但在實際操作時，常需於政策實施後，即時評估政策，以瞭解政策介入的效應，並提供政策修正。差異中差異法(difference-in-difference method, DID)可免除長期觀察時間，為方便常用的評估分析工具，但此方法最大的限制，可能為未考慮加入政策方案者選擇偏差(selection bias)的問題。之前評估我國糖尿病論質計酬方案成效之研究，即多以DID為研究設計方法，且結果多為正向的政策成效[1-5]。

本研究以傾向分數法做為政策評估的工具，研究設計主要是為解決上述對照組與介入組不可比較(non-comparable)的問題，再增加配對產生配對組，與原介入組及對照組比較，有異於傳統上僅以未參加者為對照組，更有助釐清政策方案之成效，於方法學上有其創新之處，值得肯定。

然研究者仍需多方嚴謹考量研究設計，否則從對照組中，加入若干配對條件，僅是研究方法異於傳統方法，從具有某些特質的對照組，配對出仍具對照組特質之配對組，不盡然真可比較，恐反陷於另一selection bias之研究設計迷思。另如本研究以對照組門診就醫費用最高者作為指標門診就醫(index ambulatory visit)，為配對組介入時間之條件，然就醫費用可能因該次門診再確認所有檢查值、醫院換藥或其他因素之影響，似乎非為最適當配對介入時間之條件。可否考量其他條件為指標，如以介入組個案介入時間為配對條件基礎，或先觀察各次門診費用之分佈，說明最高費用指標時間點具有鑑別度，或可使本研究設計更臻完美。

1990年代之後，隨著可避免住院(avoidable hospitalization)發展成熟，逐漸被使用測量許多研究的照護結果[6]。本研究亦以可避免住院為評估成效指標之一，然僅

測量有無，並未更進一步量化可避免住院者之嚴重程度，如次數、費用、天數等，後續研究或可加入此研究設計，當可更加釐清評估政策之成效。

次級資料分析(secondary analysis)最大的限制，自是資料庫內容及效度。但本研究至少兩次強調研究結果僅反映檢查是否申報，不一定是真實檢查情形。此部分除非是申報造假，但即使如此，對於所有研究對象，不論在任何組別，所受影響均為相同，似無需過慮；反而是受限資料庫內容變項，如前段略述，無法完全隨機配對，排除相關干擾因素，致生selection bias，恐非本研究原意，可能是強調配對方法的本研究，最重要的研究限制。

本研究於傳統政策評估方法領域另闢蹊徑，提供不同方法觀點評估政策成效，殊為難得。雖仍有諸多限制，待後續研究克服，然實非戰之罪，俟資料庫更加完整蒐集，即可竟其功。

參考文獻

1. Lee TT, Cheng SH, Chen CC, Lai MS. A pay-for-performance program for diabetes care in Taiwan: a preliminary assessment. *Am J Manag Care* 2008;16:65-9.
2. 李玉春：全民健保支付與管理制度對病人就醫品質之影響：以可避免之住院為例。行政院衛生署九十五年度科技研究計畫，計畫編號DOH95-HI-1002。台北：行政院衛生署，2006。
3. 李玉春：建立全民健保以共同照護模式為基礎的糖尿病人疾病管理計畫先導性研究(III)。行政院衛生署九十一委託計畫，計畫編號DOH91-NH-1002。台北：行政院衛生署，2002。
4. 侯佳雯：「糖尿病醫療服務改善方案」試辦計畫對第二型糖尿病人醫療資源利用之影響。台北：台灣大學醫療機構管理研究所碩士論文，2004。
5. 李待弟：全民健康保險糖尿病醫療給付改善方案初步影響評估。台北：台灣大學衛生政策與管理研究所碩士論文，2006。
6. Lin W, Huang IC, Wang SL, Yang MC, Yaung CL. Continuity of diabetes care is associated with avoidable hospitalizations: evidence from Taiwan's National Health Insurance scheme. *Int J Qual Health Care* 2010;22:3-8.

中國醫藥大學醫務管理系暨碩士班
黃光華
聯絡地址：台中市學士路91號
E-mail: khhuang@mail.cmu.edu.tw