

評估健保制度中肺結核的 論質計酬專案*

謝鈺婷

高雄市立小港醫院總務室辦事員

管中閔

國立台灣大學財務金融學系「台大講座」教授

蔡蕙安

中山大學經濟學研究所教授

謝鈺偉

紐約大學經濟學系博士候選人

「論質計酬制」是世界各國正逐漸推行的健保給付制度。本文針對台灣健保制度中的結核病「論質計酬專案」進行制度評量，也就是估計並檢定醫療院所加入此計畫後，結核病的完治率（處置效果）是否顯著提高。實證作法上，我們採用非參數的最近鄰域配對法來估計平均處置效果。實證結果發現，「論質計酬專案」對結核病的完治率的確產生顯著效果，醫療院所加入此計畫可使整體完治率提高約 13%。但若考慮不同的層級，我們發現此計畫對診所與地區醫院的效果較為顯著（診所的完治率可提高約 46%，地區醫院完治率提高約 10%），但是對層級較高的區域醫院與醫學中心則無顯著效果。若考慮不同的權屬，我們發現在診所和地區醫院中，私立院所的處置效果都大於公立者，但財團法人醫院不管在那一層級處置效果皆不顯著。我們亦發現若未適當控制各種變因，將會大幅高估此計畫的處置效果。

關鍵字：平均處置效果、配對估計、論質計酬制、制度評量、肺結核、
完治率

* 作者們特別感謝兩位匿名評審，許績天、連賢明、蔡偉德、劉錦添、羅紀瓊、林常青、陳宜廷諸位教授，以及 WETA 與台灣經濟學會 2007 年年會與會者所提供的寶貴建議。本研究部分資料來自行政院衛生署疾病管制局，但文中任何闡釋或結論並不代表該單位之立場；我們亦感謝羅紀瓊教授協助取得部分資料。文中任何錯誤，當然由作者們自行負責。

收稿日期：98 年 8 月 27 日；接受刊登日期：99 年 3 月 5 日

壹、前言

健康保險的支付制度是健康經濟學中的重要研究課題。支付制度對醫事服務提供者產生不同的誘因，所以會影響醫療效果、醫病關係及病人權益。傳統的支付制度大致可分為：論量計酬制 (fee for service)、論病例計酬 (case payment)、論人計酬 (capitation payment)、總額支付制度 (global budget)、論日計酬制 (per diem)，與薪水制 (salary) 六種，而這些制度都有其優缺點。如何建立更好的支付制度，一直是學術界與實務界所關心的重點。近年來，許多地方（包括台灣）正在試驗推行論質計酬制 (pay for quality)，也就是由保險提供者訂定醫治各種疾病的「品質」標準，再根據醫事服務提供者對病人醫療品質的高低（而非數量多寡）來支付醫療費用。論質計酬制以醫療品質做為計價基礎，避免了傳統支付方法中以「量」或「總額」為標準的缺點。這個制度藉由提供直接有效的財務誘因，使醫療機構有意願去改進醫療品質，也引導醫療院所提供完整且持續的醫療照護。有關論質計酬制的討論，參見 Millenson (2004: 323-324)、Rosenthal et al. (2004: 127-141)、Smith and York (2004: 112-118)，與 Bodenheimer et al. (2005: 1-4)。

然而制度的優劣不能只依據直觀來評斷，應該透過嚴謹的數量方式加以評量。近年來，許多經濟學與社會學的實證研究都利用計量方法來進行「制度評量」(program evaluation)，以衡量社會上各種制度的優劣。例如，Dehejia and Wahba (1999: 1053-1062) 估計勞工受訓計畫的效果；Angrist and Lavy (2001: 343-369) 探討小學教師參與在職訓練計畫對學生分數的影響；Chou et al. (2003: 1873-1894) 評估台灣全民健保制度是否對家庭預防性儲蓄產生影響；Winkelmann (2006: 131-145) 則分析德國健保制度改革如何改變病患就診次數等。台灣的全民健保實施多年後，近年已對某些疾病（乳癌、結核病、糖尿病、氣喘、高血壓）的醫療採取論質計酬制，但國內對此一制度的嚴謹計量評估卻仍付闕如。

本文的目的即是採用「制度評量」文獻中的新計量方法，估計並檢定結

核病論質計酬制的平均處置效果 (average treatment effect)，也就是醫療院所加入此計畫後，是否能顯著提高對結核病的完治率（處置效果）。我們也分別評估不同權屬和層級醫院的處置效果，藉以了解哪種類型的醫院，其品質會得到較大的改善。參加論質計酬制的醫院，雖然理論上應有較高的處置效果，但實際上是否確實如此則是一個實證問題。因此，我們的實證假說為：論質計酬計畫的處置效果是否顯著異於零。這些實證結果，應可作為未來健保支付制度興革的參考。

「制度評量」通常採用最小平方法來估計模型中的平均處置效果；Wooldridge (2002) 對此有詳細的討論。為了避免模型誤設，近年的研究則常採用非參數方法來估計處置效果，包括傾向分數 (propensity score) 估計法，如 Dehejia and Wahba (1999: 1053-1062)；核迴歸 (kernel regression) 估計法，如 Heckman et al. (1997: 605-654)；series estimator，如 Hahn (1998: 315-331)，以及最近鄰域配對法 (nearest neighbor matching)，如 Abadie and Imbens (2006: 235-267)。這些非參數方法主要是透過相近或類似的資料加以比較，所以資料如何配對是此一方法的關鍵。當影響制度的變因 (covariate) 很多時，找到非常相近資料來配對就變成很困難的工作。傾向分數估計法將各種變因合為一個傾向分數，並以此作為配對的基礎；這個方法雖然相對方便，但也會造成估計效率的損失 (Hahn, 1998: 315-331)。本研究採用最近鄰域配對法來估計論質計酬制的平均處置效果；目前採用此方法的實證研究在國外尚不多見，本文更是國內相關研究的首創。

台灣的結核病論質計酬計畫提供了制度評量的難得機會。實務上由於病人鮮少知道究竟有哪些醫院有加入計畫，醫院也不會以加入計畫來招攬病患，因此我們可以排除許多樣本選擇偏誤的問題。本文的實證結果發現，「論質計酬專案」對提高結核病的完治率的確有顯著效果，醫療院所加入此計畫可使整體完治率提高 13%。若考慮不同的層級，我們發現此計畫對診所與地區醫院的效果較為顯著（診所的完治率可提高約 46%，地區醫院完治率提高約 10%），但是對層級較高的區域醫院與醫學中心則無顯著效果，亦即無法再提高結核病完治率。針對不同醫院權屬比較的結果顯示，在診所和地區醫院中，私立院所的處置效果都大於公立者，而財團法人醫院在任何層級的處置

效果皆不顯著。此外，我們亦發現若未適當控制各種變因，將會大幅高估「論質計酬專案」的效果。

本文的結構安排如下。第貳節介紹結核病相關的健保支付制度。第參節介紹平均處置效果及其估計方法。第肆節則說明本文的資料，變數，與配對估計。第伍節為實證結果的討論，以及成本有效性分析。第陸節為結論。

貳、結核病的背景與健保制度

一、結核病與相關健保專案

結核病是一種有三千年歷史的古老疾病。目前全球約有三分之一的人口受結核菌感染，每年約有 800 萬人新發生結核病，300 萬人死於結核病。台灣地區每年約有 15,000 個肺結核新病例，其中 4,000 多例具有傳染性，每年約 900 人死於肺結核。結核病是歷年來法定傳染病發生的第一位，也是歷年來法定傳染病死亡的第一位。依據衛生署疾病管制局資料顯示，我國結核病罹患率是美國的 10 倍，日本的 2 倍，且多重抗藥性問題更在這 15 年間增加了 10 倍以上。這些都顯示結核病的防治工作是刻不容緩的事情。

結核病雖然是重大傳染病，但隨著抗結核藥物不斷的演進，至今有 95% 的結核病患是可以完治的，而完治的治療期間約為六到九個月。結核病能否完治，服藥遵從性是個很關鍵的因素。可是在台灣，病人往往因藥物副作用的問題，或各種個人與經濟因素，而自行停藥，以致無法完成治療；請見胡曉雲等（2005: 348-359）的討論。

針對結核病的治療，政府於 2001 年 10 月實施「全民健康保險肺結核病患醫療改善方案試辦計畫」；此專案為針對肺結核病患制定，並不包括所有的結核病患。該專案將整個療程分成不同階段，實行包裹式給付。而在 2004 年 1 月 1 日起，政府改採新的結核病專案：「全民健康保險結核病醫療給付改善方案試辦計畫」。此計畫之特點在於改採「論質計酬制」，並且收案對象擴大

到其他結核病患。¹ 中央健保局推動論質計酬制，其目的在於鼓勵醫療院所提供整體性照護與個案管理，以提升結核病患的照護品質及完整治療。此計畫將結核病患的治療療程分成多個階段，每三個月為一階段。論質計酬的支付標準為依階段給付，支付的重點為：醫療機構若能對病人連續疾病管理照護三個月以上，則可得到疾病管理照護費每月 500 點，此外還有其他額外的支付點數（如醫師確診診察費、完成治療診察費等）。中央健保局希望這些額外的給付能形成對醫療院所與醫師的財務誘因；有關此一專案的詳細規定請參見附錄一、二，以及中央健康保險局（2004）。

二、論質計酬支付制度

Rosenthal et al. (2004: 127-141) 對論質計酬制有很深入的討論。他們指出，論量計酬制無法使醫師提供對病人最適當的醫療處置，甚至可能造成醫師採取對病人較不利的醫療處置；但論質計酬制卻能獎勵提供高品質的醫事服務提供者。然而論質計酬制最困難之處在於如何測量醫療品質。一般的測量指標包括：(一)醫院結構（如醫療設備、醫院的人力水準等），(二)療程（醫事服務提供者是否完成某項特定工作或醫療處置），(三)結果（病人治療後的健康狀況或對治療的滿意度）。美國醫療品質的衡量指標大都是針對療程；若只衡量結果，醫生會承擔過多的財務風險，以致於拒收重症病患。我國健保制度的結核病專案中，則是以衡量療程為主，結果為輔；也就是評斷醫院是否達成結核病連續疾病管理照護（療程），而且結核病患是否確實完治（結果）。

論質計酬制下的支付方法有競爭式與非競爭式兩種。在競爭式下，醫事服務提供者會根據其醫療品質被打分數和排名，排名前 10% 的醫院可以拿到醫療支出中一定比例（如 2%）的紅利；排名越後則紅利拿到的越少，甚至當排名墊底時，還可能被罰款。非競爭式的給付法則沒有醫院間相互比較的問題；只要醫療處置超過某一固定的醫療衡量指標，或病人達到一定的醫療指

1 病患包括主診斷碼為結核病 010-018 之病患，其中 010 原發性結核感染，011 肺結核病，012 其他呼吸道結核病，013-018 肺外結核。

標，醫院就可以得到紅利。Rosenthal et al. (2004: 127-141) 指出，競爭式的給付法會使醫事服務提供者有更強的誘因去提高醫療品質，但在非競爭式的支付法之下，醫院提高醫療品質的誘因會相對較弱。

論質計酬制也有其缺點。Smith and York (2004: 112-118) 指出，論質計酬制雖然對慢性病的疾病管理有重大影響，但是論質計酬多給的報酬卻往往追不上醫事服務提供者增加的工作量與成本；Bodenheimer et al. (2005: 1-4) 也有類似的看法。從執行層面來看，醫療指標以及如何衡量指標是論質計酬制的關鍵，錯誤的衡量方式和指標會侵蝕保險業者與醫事服務提供者之間的信任，也容易形成廣泛的舞弊行為；參見 Millenson (2004: 323-324)。Trude et al. (2006: 537-542) 另外指出，保險業者的權屬和利益也會影響論質計酬制的推行；例如，很難說服私立或財團法人的保險業者給醫事服務提供者額外的績效報酬。

綜觀國內外文獻，目前尚缺乏採用嚴謹計量方法來評量論質計酬的實證論文。國內現有的研究多直接衡量「計畫下治療」與「計畫外治療」兩組病患之完治率差異，這種方法的缺點為無法控制其他可能影響完治率的因素。蔡文正等 (2005) 以卡方雙變項分析，發現加入計畫與否與完治率沒有相關性。然而此方法僅為相關性分析，既非估計與檢定平均處置效果，也無法控制其他可能影響完治率的因素。因此本文針對健保局於 2004 年起推動的「論質計酬專案」(亦即「全民健康保險結核病醫療給付改善方案試辦計畫」)，嘗試以新的計量方法加以評量，以確認該制度是否真能有效改善結核病患的醫療狀況。由於本文利用的是全台所有結核病患的實際(匿名)資料，而非抽樣問卷，因此資料兼具客觀性與普遍性，這使得本文的評量結果具有重要的參考價值。

參、平均處置效果與估計方法

本節中我們將討論如何評量特定政策或制度的效果。將任何政策或制度視做一種處置，我們要根據接受與未接受處置者的反應 (response)，設法評估該政策或制度的處置效果。以本文將探討的「論質計酬專案」為例，該方

案希望透過獎勵方式給予醫生適當的財務誘因，促使其妥善追蹤照護病人，進而提高完治率。在此，論質計酬專案是處置，病人完治與否則是反應。

一、平均處置效果

令 Y_{1i} 為虛擬變數，表示病人 i 接受處置後的反應。而 Y_{0i} 為病人 i 未接受處置的反應虛擬變數。例如， Y_{1i} 為病人在論質計酬制下接受治療的結果， Y_{0i} 為病人在舊制度下接受治療的結果。文獻上稱 Y_{1i} 與 Y_{0i} 為可能反應 (potential response)， $Y_{1i} - Y_{0i}$ 代表該處置對該病人 i 的個別效果 (individual effect)，而 $Y_{1i} - Y_{0i} > 0$ ，代表該處置有效。

對個別 i ，我們只能觀察到兩種可能反應中的一種，所以個別效果是無法被認定的 (not identified)。文獻上稱無法觀察的反應為擬真 (counterfactual)：如果觀察到 Y_{1i} ， Y_{0i} 為擬真；反之 Y_{1i} 為擬真。令 D 表示處置分派 (treatment assignment)， $D_i = 1$ 代表病人 i 接受處置，否則 $D_i = 0$ 。則可觀察的反應表示為 $Y_i = D_i Y_{1i} + (1 - D_i) Y_{0i}$ ；所以當 $D_i = 1$ 時， $Y_i = Y_{1i}$ ，而 $D_i = 0$ 時， $Y_i = Y_{0i}$ 。我們雖無法認定個別的處置效果，但對於接受處置的處置組 (treatment group，或稱實驗組) 與未接受處置的控制組 (controlled group，或稱對照組)，我們可設法估計兩組之間的平均處置效果： $E(Y_1 - Y_0)$

我們首先考慮以下的「外生性」條件。

假設 1. $(Y_1, Y_0) \perp D$ ，其中 $\perp$ 代表獨立。

當處置分派 D 為隨機時， D 和反應無關，故假設 1 在隨機實驗時成立。但是當某些群體一律加入特定制度而沒有選擇的餘地時， D 與 (Y_1, Y_0) 之間也是獨立的；這種實驗被稱為準實驗 (quasi-experiment)，或者自然實驗 (natural experiment)。假設 1 保證處置組和控制組相當類似，因而可以互相比較；這是適當估計平均處置效果的基本要求。令 $E(Y|D=1)$ 代表處置組均數， $E(Y|D=0)$ 代表控制組均數。在假設 1 之下，這兩組均數的差是：

$$E(Y|D=1) - E(Y|D=0) = E(Y_1|D=1) - E(Y_0|D=0) = E(Y_1) - E(Y_0),$$

也就是平均處置效果。所以，估計平均處置效果就是估計兩組均數，並加以比較。

大多數的社會科學研究中，處置分派通常是非隨機的。例如，醫院可以自行選擇是否加入特定計畫，此時 D 不是隨機分派，假設 1 也不一定成立。實際情況裡，處置分派 D 可能會受某一變數 X （如醫療院所層級）的影響；譬如健保局可能會要求層級較高的醫療院所（如醫學中心）配合新的試辦計畫，而通常大醫院成本較低，比較可能吸收因較複雜的行政程序與作法所產生的額外成本。當然， X （醫療院所層級）也可能影響到反應 Y （完治率）。因此， Y 不但受 D 影響，也受 X 的影響。既然 X 為 D 和 Y 的共同因素，假設 1 就不可能成立。文獻上稱呼 X 這一類的變數為變因，或混淆因子（confounder）。

當假設 1 不成立時，我們就不能用處置組與控制組的均數差來估計平均處置效果。舉例來說，高層級的醫療院所通常有較高的機率參加計畫，但也有較高的完治率。若處置組裡面有較多的高層級醫療院所，則不論計畫有效與否，我們都可能觀察到較高的完治率。所以若不能適當的控制變因（如醫療院所層級），處置組和控制組之間的直接比較是沒有意義的。蔡文正等（2005）直接用兩組樣本的平均數之差來判斷論質計酬制對完治率的影響，就是忽略了非隨機樣本可能造成的偏誤。這種因遺漏可觀察的變因所產生的偏誤，文獻上稱之為顯性偏誤（overt bias）。

我們現在考慮另外兩個條件；令 X 代表所有變因所成的向量。

假設 2.（無混淆，unconfoundedness） $(Y_1, Y_0) \perp D | X$ 。

假設 3.（有重疊，overlap） $0 < \text{prob}(D=1|X) < 1$ 。

假設 2 為假設 1 的延伸，也就是控制了所有可觀察的變因之後， D 與 (Y_1, Y_0) 之間獨立。假設 3 保證了處置組與控制組之間是可以比較的。例如，若 $X=1$ 代表男性，則 $\text{prob}(D=1|X=1)=1$ 代表男性受試者均為處置組，而控制組完全沒有男性。此時因性別因素無法得到控制，兩組之間就無從比較。

二、估計方法

以下我們介紹估計平均處置效果的迴歸與配對方法。令

$$\begin{aligned}\mu(x, d) &= E(Y|X=x, D=d), \quad \mu_d(x) = E(Y_d|X=x), \\ \sigma^2(x, d) &= V(Y|X=x, D=d), \quad \sigma_d^2(x) = V(Y_d|X=x). \text{ 若假設 2 成立, 則} \\ \mu(x, d) &= \mu_d(x), \quad \sigma^2(x, d) = \sigma_d^2(x).\end{aligned}$$

(一) 迴歸估計法

假設處置的個別效果是不隨個人改變的固定效果： $Y_{1i} - Y_{0i} = \tau$, for all i , 則平均處置效果亦為 τ 。設 $Y_{0i} = \alpha + X_i' \beta + \varepsilon_i$, 我們可以得到一個包含了虛擬變數 D_i 的迴歸模型：

$$Y_i = Y_{0i} + D_i(Y_{1i} - Y_{0i}) = \alpha + \tau D_i + \beta' X_i + \varepsilon_i. \quad (1)$$

此迴歸模型控制了變因 X , 估計 τ 所得到的 $\hat{\tau}$ 即為平均處置效果的估計式。從另一個角度來看, 在假設 2 之下,

$$\begin{aligned}E(Y|D=1, X) - E(Y|D=0, X) &= E(Y_1|D=1, X) - E(Y_0|D=0, X) \\ &= E(Y_1|X) - E(Y_0|X),\end{aligned}$$

這是 conditional 平均處置效果 (表示為 $\tau(X)$)。平均處置效果即為 $\tau(X)$ 的期望值：

$$\tau = E[E(Y_1 - Y_0|X)] = E[\tau(X)],$$

其對應的估計式為

$$\hat{\tau} = \frac{1}{N} \sum_{i=1}^N [\hat{\mu}_1(X_i) - \hat{\mu}_0(X_i)], \quad (2)$$

其中 $\hat{\mu}_1(X_i)$ 為 $E(Y|D=1, X)$ 的估計值, $\hat{\mu}_0(X_i)$ 為 $E(Y|D=0, X)$ 之估計值。所以可用 $\hat{\mu}_1(X_i) - \hat{\mu}_0(X_i)$ 估計 $\tau(X)$, 而其平均值則是 $E[\tau(X)] = \tau$ 的估計值。

由上述討論可知，估計平均處置效果的重點是估計條件均數 $E(Y|D, X)$ 。若用最小平方方法估計包含虛擬變數 D_i 的迴歸模型(1)， $D_i=1$ 之下配適值為 $\hat{\mu}_1(X_i) = \hat{\alpha} + \hat{\tau} + \hat{\beta}'X_i$ ； $D_i=0$ 之下配適值為 $\hat{\mu}_0(X_i) = \hat{\alpha} + \hat{\beta}'X_i$ 。兩者的差為估計的平均處置效果 $\hat{\tau}$ 。

(二) 配對估計法

線性迴歸並非唯一可以估計 $E(Y|D, X)$ 或者設算擬真的方法；近期相關文獻多採用非參數方法來估計平均處置效果，譬如本文採用的最近鄰域配對估計法，Hahn (1998: 315-331) 的 series estimator，和 Heckman et al. (1997: 605-654) 的核迴歸方法等。無母數估計法相較於傳統的迴歸分析（如 OLS 和 logit 模型等），在研究健康經濟議題上有許多優點。由於 logit 模型仰賴分配上的假設，也需要假設應變數與自變數之間的函數形式，這些均容易造成模型誤設的問題。在某些經濟議題（例如生產函數）的研究中，經濟理論多少提供了這一類模型設定的基礎（如 CES 生產函數）。然而健康狀況與其他變因之間的函數形式，文獻上並無確定的理論。以無母數方法出發的最近鄰域配對估計法，正好可以避免模型誤設的問題，提供較為可靠的實證估計結果。

顧名思義，最近鄰域配對估計式（簡稱配對估計式）就是把處置組和控制組中類似的（變因值相近的）受試者配對後再做比較。² 配對估計式簡單且容易實作，因此在文獻中已有很長的歷史，尤其一對一配對（pair matching）更常見於生物與醫學統計的文獻（Rosenbaum, 2002）。Abadie and Imbens (2006: 235-267)，底下簡稱 AI，建立了本文的配對估計式的極限性質，使我們不僅能估計平均處置效果，更能檢定平均處置效果是不是顯著異於零（也就是檢定政策是否有效）。

我們先從資料裡找出和受試者 i 在相對的組別裡最靠近的 M 個人以進

2 文獻中亦常根據「傾向分數」加以配對。由於本文並未採用此法，我們在此不予討論；有關該方法的討論參見 Dehejia and Wahba (1999: 1053-1062)、Smith and Todd (2001: 112-118)、Imbens (2004: 4-29)，與 Dehejia (2005: 355-364)。

行配對。令 $j_m(i)$ 為一指標 (index) 同時滿足 $D_j=1-D_i$ ，且

$$\sum_{k:D_k=1-D_i} 1\{\|X_k - X_i\| \leq \|X_j - X_i\|\} = m,$$

其中 $1\{\cdot\}$ 為指標函數 (indicator function)， $\|X\|=(x'x)^{1/2}$ 。 $j_m(i)$ 標出在相對 i 的組別裡，第 m 個靠近 i 的樣本。例如，若 i 屬於處置組，我們將找出在控制組中第 m 個性狀最靠近 i 的樣本，然後用 M 個最靠近的樣本設算 i 的擬真。這些樣本所屬的指標所成的集合可寫作：

$$J_M(i)=\{j_1(i),\dots,j_M(i)\}。$$

給定使用 M 個配對，令 $K_M(i)$ 代表樣本 i 被用來設算擬真的次數：

$$K_M(i)=\sum_{k=1}^N 1\{i \in J_M(k)\}。$$

AI 分析此種形式的配對估計式：

$$\hat{Y}_{0i}=\begin{cases} Y_i & \text{if } D_i=0 \\ \frac{1}{M} \sum_{j \in J_M(i)} Y_j & \text{if } D_i=1 \end{cases}$$

與

$$\hat{Y}_{1i}=\begin{cases} \frac{1}{M} \sum_{j \in J_M(i)} Y_j & \text{if } D_i=0 \\ Y_i & \text{if } D_i=1 \end{cases}$$

我們以最靠近 i 的 M 個樣本取其平均值來設算擬真，而平均處置效果的估計式為：

$$\hat{\tau}_{N,M}=\frac{1}{N} \sum_{i=1}^N (\hat{Y}_{1i} - \hat{Y}_{0i}),$$

此稱為簡單配對估計式。

AI 證明，當配對的數目 M 固定不變，樣本數 N 趨近無窮大時，此估計式收斂速度為 $O_p(N^{-1/K})$ ，其中 $k \geq 2$ 為連續型變因 X 的維度。換言之，不論 M

的大小， $\hat{\tau}_{N,M}$ 均為一致性估計式。此外，變因 X 的維度越高，收斂速度就越慢。當受試者 i 的變因為 X_i ，若所有用來配對的樣本均和受試者 i 有完全一樣的變因值：

$$X_{jm(i)} = X_i, \quad \forall j_m(i) \in J_M(i),$$

這種配對稱為完全配對。從直觀上想，當維度越高時越難找到完全配對，因此收斂速度就會變的越慢。

AI 並證明在一定條件下且 $k=1$ 時， $\hat{\tau}_{N,M}$ 的極限分配為：

$$V^{-1/2} \sqrt{N} (\hat{\tau}_{N,M} - \tau) \Rightarrow N(0,1)$$

其中 V 為 $\hat{\tau}_{N,M}$ 的極限變異數。估計 V 涉及到估計條件變異數 $\sigma_{D_i}^2(X_i)$ 。一種方法是使用無母數的方式來估計；無母數方法雖可確保一致性，但有選擇核函數以及帶寬值 (bandwidth) 的問題，且 X 維度越高估計就越不準，這就是所謂的 curse of dimensionality 問題。AI 則根據配對的概念來估計，但此時不是從相對的組別裡去找性狀接近的樣本，而是在同一個組內尋找性狀接近的樣本。舉例來說，若我們找到兩個樣本變因值相等且屬於同一組，則可得到下列關係：

$$E[(Y_i - Y_j)^2 | X_i = X_j = x, D_i = D_j = d] = 2\sigma_d^2(x)。$$

因此， $(Y_i - Y_j)^2/2$ 即為 $\sigma^2(X_i, D_i)$ 之不偏估計式（但非一致估計式）。

我們可進一步拓展此一概念，令 $l_j(i)$ 為第 j 個靠近 i 的樣本，且兩者屬於同一組。利用 J 個樣本所估計 $\sigma_{D_i}^2(X_i)$ 的不偏估計式為：

$$\hat{\sigma}_{D_i}^2(X_i) = \frac{J}{J+1} \left(Y_i - \frac{1}{J} \sum_{j=1}^J Y_{l_j(i)} \right)^2。$$

事實上若有 $\sigma_{D_i}^2(X_i)$ 的不偏估計式即足以一致估計 V 。AI 提供了 $\hat{\tau}_{N,M}$ 的極限變異數的一致估計式如下：

$$\hat{V} = \frac{1}{N} \sum_{i=1}^N (\hat{Y}_{1i} - \hat{Y}_{0i} - \hat{\tau}_M)^2 + \frac{1}{N} \sum_{i=1}^N \left[\left(\frac{K_M(i)}{M} \right)^2 + \left(\frac{2M-1}{M} \right) \left(\frac{K_M(i)}{M} \right) \right] \hat{\sigma}_{D_i}^2(X_i)$$

三、比較迴歸方法與配對估計法

我們這節討論迴歸與配對方法的異同。首先，線性迴歸為參數模型，可能產生模型誤設；配對為非參數方法，故可避免此一問題，但此方法仍有因為選取 M 而產生的任意性的問題。其次，從上文的討論中，我們知道以 OLS 估計平均處置效果，其實已經假設了每個人的處置效果都一樣。使用配對法則可以允許每個人可以有不一樣的處置效果，因此避免了這個很強的假設。再者，線性迴歸和配對是用不同方式處理樣本選擇偏誤。在迴歸方法中，我們估計線性模型 $Y_i = \alpha + \tau D_i + \beta' X_i + \varepsilon_i$ ， $\hat{\tau}$ 即為估計的平均處置效果。由 Frisch-Waugh-Lovell 定理可知， $\hat{\tau}$ 是已經去除了 X 對 Y 以及對 D 的線性影響所得到的估計值。配對則直接把控制組和處置組當中，變因值相近的受試者予以比較。因此不管 X 對 Y 以及對 D 的影響是線性或非線性，配對法均能夠處理。

最後，這兩種方法對那些樣本該納入計算的處理方式不同。令 X_1 與 X_0 分別代表處置組受試者與控制組受試者變因的分配， $s(X_1)$ 與 $s(X_0)$ 為其對應的支撐集 (support)。配對僅取用兩組支撐集重疊處 $s(X_1) \cap s(X_0)$ 附近的樣本進行估計。如果重疊範圍很大，以變因的角度來看兩組是非常相近的，因此可以做比較。若重疊範圍很小，代表兩組特性差異很大，其實可能不適合比較。配對法只分析重疊的部分，所以只有一部分的樣本會被納入計算；文獻上稱此現象為支撐集問題 (support problem)。極端的例子是 $s(X_1) \cap s(X_0)$ 為空集合，兩組變因既無重疊，也就無法配對。舉例來說，若政府推行一項補助計畫，規定年收入低於 20 萬者全部參加。如此所得變數就無法用來配對，因為處置組全部都是低所得，而控制組全部都是高所得。然而迴歸方法不管支撐集重疊範圍大小，均使用全部樣本來進行估計。參見 Lee (2005) 與 Imbens (2004: 4-29) 的討論。

肆、資料與配對估計

一、資料與變數

我們的資料來源分成三部分。第一部分為行政院衛生署疾病管制局提供的 2005 年結核病資料庫 (2005)，內含 22,666 個觀察值。第二部分為中央健康保險局提供的 2005 年加入結核病醫療給付改善試辦計畫的醫療院所名單，及其層級等資料 (2005)，共 172 個觀察值。第三部分則是由中央健康保險局網頁查知的 2007 年醫事機構名單以及層級等資料 (2007)。結核病資料庫內有 22,666 個觀察值，但為了得到所需變數，我們必須藉由「銷案單位代碼」來判斷病患所在的醫療院所層級、醫療院所權屬、以及醫療院所是否加入該計畫。因此，我們剔除 2,688 個銷案單位代碼有問題的觀察值 (包括 742 個沒填銷案醫事機構代碼者，1,946 個銷案單位代碼輸入錯誤者)，以及 22 個沒有銷案原因的觀察值，最後得到 19,956 個觀察值。

由結核病資料庫的銷案單位代碼中，我們得知當時全國共有 803 間醫療院所收治結核病患。對於加入計畫的醫療院所，我們根據健保局提供的 2005 年資料判斷其醫療層級。但對未加入計畫的醫療院所，我們卻不知道其當時的真正醫療層級，所以我們改用 2007 年的醫療院所層級資料作為判斷的依據。³ 由於醫療院所層級一般而言變動不大，此一方式應不致產生太大誤差。

我們得到的變數包含：銷案單位代碼、通報單位代碼、X 光診斷結果、痰塗片結果、痰培養結果、銷案原因、出生日期、性別、醫療院所是否位於城市、醫療院所層級、醫療院所權屬、醫療院所是否加入計畫等。⁴ 變數中的「銷案原因」可用來判斷病人是否完治。對於「醫療院所是否位於城市」這個

3 此資料可從中央健保局網站下載，檔案名稱為「健保特約醫療院所名冊壓縮檔」，下載日期 2007 年 5 月 5 日，網址 <http://www.nhi.gov.tw/> (中央健康保險局，2007)。

4 資料庫中一些社會經濟變數，如婚姻狀態和職業，並不完整，所以無法採用。例如，婚姻狀態變數的遺漏值高達 68.45%，而職業變數當中亦有 89.47% 為「其他」，也就是不知道其職業為何。

變數，我們以台北縣的人口密度為城市的劃分標準。⁵ 台北縣已晉升為直轄市，也是全國人口最多之處；超過台北縣人口密度的包括：高雄市、台北市、台中市、新竹市、與基隆市，其人口密度每平方公里皆大於 1,800 人，這些地方與台北縣均列為城市。「醫療院所層級」共有四種：醫學中心，區域醫院，地區醫院，與診所。「醫療院所權屬」代表醫院所有權的歸屬，共有公立，私立，財團法人三種。⁶ 為了簡化分析，我們以病患的銷案單位代碼（銷案醫院）為主，只要病患的銷案醫院在 2005 年度加入計畫，（不管其加入計畫時間多長、何時加入、何時退出），我們都假設該個案在結核病論質計酬計畫下治療。

我們實證分析的依變數為完治率。由於痰培養檢驗為陽性者，從細菌學的角度上來看代表結核病的「確診」病例，我們於是依據痰培養檢驗為陽性的病患來計算完治率。我們從整理過後的通報病患 19,956 個觀察值中，剔除未做痰培養者 (3,252 個)、痰培養已驗未出者 (364 個)、痰培養陰性者 (7,263 個)，最後剩下 9,077 個痰培養陽性的觀察值作為本文分析的樣本。

此一樣本中，男性占約 69.5%，為女性的兩倍多；從年齡上來看，65 歲以上的老年人約占五成六，45 至 65 歲者占約四分之一。X 光片的檢驗結果可以區分為正常、異常但無空洞、異常且有空洞。據此可作為病情嚴重度的判斷，其中以異常且有空洞最為嚴重。病患中，約七成的胸部 X 光為異常但無空洞，約兩成的胸部 X 光為異常且有空洞。根據痰培養陽性病患所計算的完治率為 69.17%。而在 9,077 個觀察值中，6920 人 (76.2%) 是在加入結核病論質計酬試辦計畫的醫療院所中接受治療（以下簡稱「計畫下治療」），其完治率為 77.59%；但病患若在未加入試辦計畫的醫療院所中治療（以下簡稱「計畫外治療」），其完治率僅為 42.19%，遠低於前者的完治率。

我們以下亦將依照醫院的層級 (level) 與權屬 (ownership) 劃分，藉此控制這些特性的影響，然後分別估計不同層級與不同權屬之下的處置效果。

5 我們劃分的重點是為了估計時能控制不同地區所可能產生的影響。我們也曾區分北中南東四區來估計，結果大體和現在差不多。

6 社團法人與私立醫院屬性及其本質其實相同（衛生署當初為了解決私立醫院經營上的一些問題，所以鼓勵私立醫院改登記為社團法人），加上社團法人僅有 291 個病患，我們於是將社團法人醫院與私立醫院合併。

不同層級的醫院代表著不同的服務功能、任務與發展目標，也反映了病床數（醫院規模）、醫院評鑑結果（即醫療品質）等資訊；一般來說，層級越高的醫院代表該醫院的醫療資源和品質越好。不同權屬的醫療機構由於所有權的歸屬不同，一方面有不同的賦稅負擔（例如公立醫療機構可免徵營業稅、所得稅、房屋稅等），另一方面則可能有不同的營運目標與管理方式（例如公私立醫院有不同的財務分配方式）。在健康經濟或醫務管理的實證研究中，醫院層級與權屬都是重要的考量因素，例如許績天、連賢明（2007: 415-450）。蔡文正等（2002; 2005）在評估與結核病相關的計畫時，亦曾針對醫院層級加以分析。

表 1 中首先列出痰培養陽性病患在各醫療院所層級的比例與完治率，以及在計畫內外治療的比例和各自的完治率。表中顯示醫療院所層級越高，其完治率越高；例如，診所完治率僅約 39%，醫學中心完治率則為其兩倍，高達 79%。除了診所以外，其餘層級的病患多數在計畫下治療。除了區域醫院這個層級，其他層級在計畫下治療的病患完治率皆高於計畫外治療的病患。例如在醫學中心，計畫下治療的完治率高於計畫外治療者約 5%，但在診所層級兩者相差卻 56%。表 2 分別列出不同層級下，痰培養陽性病患在各權屬的比例與完治率，以及在計畫內外治療的比例和各自的完治率。我們可以看到，財團法人院所的完治率通常較高（除了在地區醫院層級中，財團法人完治率略低於公立者），而診所以外的公立醫院完治率均高於私立醫院。不論公私立院所，我們都可看到計畫下治療的完治率高於計畫外治療者的完治率；但是在財團法人的地區醫院與區域醫院中，反而是計畫外治療者的完治率略高。

表 1：按醫療院所層級區分的痰培養陽性病患比例與完治率

層級別	診所		地區醫院		區域醫院		醫學中心	
觀察值	1,688		2,015		2,719		2,655	
比例%	18.6		22.2		29.95		29.25	
完治率%	38.74		72.41		76.17		78.91	
是否在計畫下治療	是	否	是	否	是	否	是	否
比例%	14.87	85.13	73.25	26.75	94.19	5.81	99.13	0.87
完治率%	86.85	30.34	76.22	61.97	76.06	77.85	78.95	73.91

表 2-1：診所層級下按權屬區分的痰培養陽性病患比例與完治率

權屬別	公立		私立		財團法人	
比例%	86.55		13.33		0.12	
完治率%	34.91		63.11		100	
是否在計畫下治療	是	否	是	否	是	否
比例%	7.8	92.2	60.89	39.11	0	100
完治率%	78.95	31.18	93.43	15.91	NA	100

表 2-2：地區醫院層級下按權屬區分的痰培養陽性病患比例與完治率

權屬別	公立		私立		財團法人	
比例%	42.53		34.69		21.99	
完治率%	76.43		66.38		75.17	
是否在計畫下治療	是	否	是	否	是	否
比例%	80.05	19.95	59.37	40.63	84.65	15.35
完治率%	78.86	66.67	73.49	55.99	74.40	79.41

表 2-3：區域醫院層級下按權屬區分的痰培養陽性病患比例與完治率

權屬別	公立		私立		財團法人	
比例%	33.76		24.02		42.22	
完治率%	75.49		71.21		79.53	
是否在計畫下治療	是	否	是	否	是	否
比例%	97.49	2.51	100	0	88.24	11.76
完治率%	76.09	52.17	71.21	NA	79.17	82.22

表 2-4：醫學中心層級下按權屬區分的痰培養陽性病患比例與完治率

權屬別	公立		私立		財團法人	
比例%	36.38		0.45		63.16	
完治率%	78.47		33.33		79.49	
是否在計畫下治療	是	否	是	否	是	否
比例%	100	0	100	0	98.63	1.37
完治率%	78.47	NA	33.33	NA	79.56	73.91

二、模型配對估計

我們的實證包括三種模型：全體模型為針對全體樣本所設定的模型；層級模型為針對四個層級（診所、地區醫院、區域醫院、醫學中心）的子樣本所設定的模型；權屬模型為針對各層級下不同權屬的子樣本（例如公立診所、私立地區醫院、財團法人區域醫院等）所設定的模型。在全體模型中，我們估計全體樣本的平均處置效果： $E(Y_1 - Y_0)$ 。在層級模型中，我們估計不同層級下的平均處置效果： $E(Y_1 - Y_0 | level)$ 。而在權屬模型中，我們估計不同醫療院所層級與權屬下的平均處置效果為： $E(Y_1 - Y_0 | level, ownership)$ 。在這裡，平均處置效果就是參加計畫與否所造成完治率的差別。

我們採用配對估計式估計平均處置效果，即第參・二・(二)節的 $\hat{\tau}_{N,M}$ ；程式為 Abadie et al. (2001: 1-18) 所撰寫，我們也評估此一效果是否具有統計上的顯著性。在決定配對個數的 M ，我們考慮 $M=1, 5, 10$ 。不同的 M 會帶來不同的結果。 M 小時，偏誤較小而變異數較大； M 大時，偏誤較大但變異數較小。由於各種估計結果相差不多，我們僅列出 $M=5$ 的結果。估計條件變異數 $\sigma_{bi}^2(X_i)$ 時，我們將樣本數 J 設為 5。決定變因值是否夠靠近的準則： $\|X\|_V = (x' V x)^{1/2}$ 中， V 為取 x 的共變數矩陣的逆矩陣主對角線的部份。在 Abadie 等人的程式中，對於某些變因，我們也要求執行完全配對。所謂完全配對指的是，把處置組和控制組變因完全相同的病人予以配對。若加入此指令會給予需要完全配對的變數在相對應 V 的位置乘上 1,000，也就是給較大的懲罰 (penalty)。因此在配對時，會優先以需要完全配對的變數為篩選基礎挑出配對樣本。而做到完全配對的比例，可作為配對品質好壞的初判。

我們針對病患年齡、轉院與否、性別、胸部 X 光、醫療院所權屬、醫療院所層級、醫療院所是否位於城市等變因加以配對。我們對性別、胸部 X 光、層級別皆要求完全配對，但對年齡、轉院與否、醫院所在地在城市與否、權屬別則不做此要求。因為年齡相差幾歲對病情影響不會太大，故不需完全配對。在轉院方面，我們並沒有考慮病患為何轉院、是在什麼層級中轉換等因素，所以我們用較寬鬆的角度來看待此一變因。至於院所是否位於城市，一般而言，在城市的院所（相對於非位於城市者）擁有較多的醫療資源，也讓

民眾就醫較為便利。但這些並非絕對，還是有些擁有充沛醫療資源的非城市院所（例如花蓮慈濟醫院、高雄縣義大醫院）；而在非城市的醫院就醫，交通也不必然就不便利。在權屬別方面，由於某些子樣本（例如私人醫學中心、財團法人診所）的樣本數過少，若要求完全配對會導致配對困難，所以對這些變因我們亦不要求完全配對。

伍、實證結果

一、配對估計結果

本節中我們分別討論不同模型的估計結果。由於資料中的變因可能有缺漏值（missing value），使得配對估計式無法配對，所以我們扣除了 627 個缺乏胸部 X 光片結果的觀察值後，得到 8,450 個痰培養陽性病患的觀察值。

估計結果顯示，全體樣本的處置效果為完治率提高 12.98%（標準差為 0.0321），其中完全配對為 95.11%。這些效果均在 1% 的水準下顯著，而且完全配對的情況很高，這表示論質計酬方案能使完治率提高約 13%。蔡文正等（2005）的分析顯示加入計畫與否和完治率沒有相關性，此和我們的結果完全不同。由於完治率介於 0 和 1 之間，我們也利用 logit 模型估計平均處置結果，結果顯示完治率提高 14.96%（標準差 0.0379），較配對估計結果超出約 15%。

表 3 列出不同層級醫療院所的處置效果，其中，診所的處置效果約為 46%，地區醫院的處置效果則為 10%，且均在 1% 水準下顯著。但區域醫院與醫學中心的效果在 5% 水準下卻都不顯著。所以我們推論處置效果隨層級提高而降低；但層級提高到區域醫院或更高時，此一專案似乎已無效果。⁷

7 我們也曾刪去曾經轉院的觀察值後重新估計全體與層級模型，估計的結果多數和現有之結果差不多，但診所的處置效果卻降低很多。這是因為轉院的觀察值幾乎都是轉入沒有加入計畫的診所，刪除這些轉院者的資料，等於刪除沒有加入計畫的診所。而這類診所本就偏低，刪除這些觀察值後，就不容易顯現出處置效果。作者們在此感謝台灣經濟學會年會評審所提的建議。

表 3：層級模型配對結果

層級別	診所	地區醫院	區域醫院	醫學中心
觀察值	1,559	1,893	2,563	2,435
完治率提升%	45.75***	10.4***	0.15	10.79
標準差	(0.0466)	(0.0274)	(0.0396)	(0.111)
完全配對	98.18	100	98.76	84.03

*, **, *** 分別代表 10%, 5%, 1% 水準下顯著。

我們可由兩方面來解釋上述的結果。第一，從生產面來看，品質生產力較低時，邊際報酬會較高，反之則邊際報酬較低。由於層級越高的醫療院所的完治率平均比較高（參見第肆・一節），所以診所品質的邊際報酬會較高，也就是完治率可以提高的較多，而區域醫院與醫學中心品質的邊際報酬較低，故完治率的增加也相對有限。第二，從誘因面來看，健保局將參加計畫的額外報酬先給付給醫療院所，再由醫療院所分配給其醫護人員。大醫院可能拿額外報酬的一部份去投資醫院其他建設，只將其中一部份分給醫師，而醫師所分到的比例亦隨每間醫院分配報酬的方式而不同。但在較小的醫療院所中，這種額外報酬較易與醫師收入直接掛鉤，形成較強的財務誘因。

由於資料的限制，在醫學中心這個層級，控制組（23 人）與處置組（2,632 人）的樣本數相差懸殊。在這種情形下，並不容易在控制組裡找到適當的病人來設算處置組的擬真。從統計方法來說，極限變異數的公式將會反映出處置組與控制組樣本數懸殊的問題：兩組樣本數差異越大，其變異數也越大。換言之，我們很難準確估計醫學中心的處置效果，檢定時越不容易拒絕虛無假設 $ATE=0$ 。這也說明了儘管估計醫學中心的處置效果將近 11%，但在統計推論上卻無法拒絕虛無假設。⁸

表 4 為不同院所權屬的各種處置效果。綜合而言，財團法人院所以及公私立區域醫院與醫學中心，其處置效果在 5% 的水準下均不顯著；公私立診所和地區醫院則在 1% 水準下有顯著的處置效果，其中又以私立者較大。例

8 我們感謝一位評審對此問題的指正。

表 4：權屬模型配對結果

院所層級與權屬	公立診所	公立地區醫院	公立區域醫院	公立醫學中心
觀察值	1,341	798	798	887
完治率提升%	46.14***	7.92*	24.42*	缺乏控制組
標準差	(0.0532)	(0.0431)	(0.1266)	
完全配對%	97.42	98.44	72.7	
院所層級與權屬	私人診所	私人地區醫院	私人區域醫院	私人醫學中心
觀察值	216	654	625	11
完治率提升%	69.74***	14.73***	缺乏控制組	缺乏控制組
標準差	(0.0607)	(0.0387)		
完全配對%	96.49	99.78		
院所層級與權屬	財團法人診所	財團法人地區醫院	財團法人區域醫院	財團法人醫學中心
觀察值	2	425	1,100	1,537
完治率提升%	缺乏處置組	0.02	-3.98	9.99
標準差		(0.0771)	(0.0361)	0.107
完全配對%		96.03	98.79	84.7991

* , ** , *** 分別代表 10%, 5%, 1% 水準下顯著。

如，私立診所的處置效果為 70%。而公立診所的處置效果則為 46%。私立地區醫院的處置效果約 15%，而公立地區醫院的處置效果僅約 8%。

綜合以上結果，我們發現若不考慮醫療院所權屬，「論質計酬專案」主要對診所和地區醫院這兩層級較有影響。考慮權屬之後，此一計畫對私立診所和私立地區醫院的效果都比公立者為大。私立院所之所以會有較高的處置效果，可能是因為私立院所以營利為目的，醫師的薪資給付方式也有別於公立院所，故可能有較強的財務誘因去提高醫療品質，以獲得額外的報酬。⁹ 醫療文獻也指出，私立院所因較強的財務誘因，會積極進入較有利潤的醫療服務項目 (Sloan, 2003)。

若未控制變因，我們可以將「計畫下治療的平均完治率」與「計畫外治療的平均完治率」相減，作為處置效果的估計。在這種方法下，痰培養陽性

9 財團法人地區醫院的效果亦不顯著，其原因可能是財團法人醫院與公立醫院在組織與經營上有其相似之處，故醫師的財務誘因可能相對較低。

病患完治率提高 35.3%（標準差 0.64），遠高於控制了變因的配對估計結果（全體病患完治率提高 13%），所以顯然放大了處置效果。表 5 列出這種方法在不同層級下的估計結果；與表 3 相比，未控制變因的估計方法亦高估診所與地區醫院的處置效果，而且高估的程度都相當大。例如，在診所層級，未控制變因的估計結果為 58%（配對估計的結果為約 46%），而在地區醫院層級，未控制變因的估計結果為 14%（配對估計的結果約 10%），均高估約 40%。此外，未適當控制變因亦使我們高估在權屬部分的處置效果；但為了節省篇幅，我們不列出此部分的結果。

表 5：不同層級下傳統估計結果

層級別	診所	地區醫院	區域醫院	醫學中心
觀察值	1,559	1,893	2,563	2,435
完治率提升%	58.43***	14.07***	-1.89***	6.02
標準差	(0.5565)	(0.6414)	(0.5929)	(0.6017)

*, **, *** 分別代表 10%, 5%, 1% 水準下顯著。

二、成本有效性分析

結核病一直被各國視為重要的公共衛生指標之一；1996 年 WHO 估計全世界一年要花費 240 億美元來防治結核病。在國內研究報告中，王怡婷等（1999: 263-270）估計 1999 年時，台灣因結核病死亡造成國民生產力損失約 1 億 2,500 多萬美元。國內文獻在結核病醫療支出方面的研究很少；Brewer et al.（1998: 232-238）的研究發現，平均一位結核病患所耗用的醫療費用從 1,113 美元到 1,500 美元不等。Josephine（1994: 107-147）指出，多重抗藥性結核病患的總醫療費用是對藥物有感受性結核菌病患的 10 倍，其藥品費為 21 倍。由此可知結核病醫療費用的支出成本相當驚人。以下我們對實施「論質計酬專案」計畫作一簡單的成本有效性分析。

根據蔡文正等（2002）的研究報告，在計畫外治療的結核病患，若在九個月內完治，總醫療費用平均約為新台幣 20,158 元。我們又根據中央健康保險局（2004），平均每位在計畫下治療的病患，健保局要多支出新台幣 6,000

元的額外報酬給醫事服務提供者（假設在九個月內完治）。最後我們根據本篇研究用最近鄰域配對估計式所估出來的痰培養陽性病患，若在結核病論質計酬計畫下治療可提高完治率約 13%。在此計畫下，健保局要對 100 位結核病患多支付 60 萬元，但可使額外 13 位結核病患得到完治。若這 13 個人因在計畫外治療而導致未完治，進而變成多重抗藥性病患（假設並未死亡），其醫療成本將成為 10 倍，此時健保局必須多支出約 260 萬元（ $20,158 \times 10 \times 13$ ）來醫治這些病患。簡言之，平均每 100 個病患，多支出 60 萬元時就可能減少約 260 萬元的支出，而全國每年若有 2 萬個病患，健保局就可能省下約 4 億元的支出。所以，即使「論質計酬專案」計畫下完治率僅提高約 13%，但仍是值得實行的政策。

三、樣本選擇性的問題

本文資料有兩種樣本選擇性的問題：醫院加入計畫與否的選擇與病人就醫的決策行為。這兩者都可能產生內生性，從而使估計結果產生偏誤。然而本文受限於資料，無法在方法或模型上針對選擇性加以調整。本節的說明，目的在使讀者瞭解此一問題，以及本文實證結果可能有的侷限性。

在就醫的決策上，一種內生性的可能性是：病情嚴重的病人都集中到層級較高的醫學中心就診，因此造成醫學中心的完治率會偏低。然而在我們的樣本中，病情（X 光片檢驗結果）的分配在各層級醫院都非常相似（請參見表 6），因此我們可以初步排除此種可能性。

另一種內生性的可能情況是，層級越高的醫院越傾向加入計畫（權屬別的影響則不顯著），而層級又與醫療品質高度相關。如果病人握有充分的醫療訊息（知道結核病的治療方式與後果），或有較高的自我察覺意識（知道自己

表 6：各層級醫院病患的胸部 X 光片分配

層級別	診所	地區醫院	區域醫院	醫學中心
胸部 X 光正常%	3.98	3.96	5.27	3.70
胸部 X 光異常無空洞%	75.18	75.22	71.40	75.36
胸部 X 光異常有空洞%	20.85	20.81	23.33	20.94

病情的嚴重性)，他們會傾向去高醫療品質的醫院就醫。¹⁰ 通常這種病人的配合度也較高，此時不論該醫院是否加入計畫，對病人的治療其實影響不大。相對而言，選擇在診所醫治結核病的病患，本身可能比較缺乏醫療訊息和自我察覺意識，此時診所若能實行連續的疾病管理照護（也就是論質計酬專案的重點），將易使完治率提高。所以，此計畫對診所的效果可能會相對較大。這些因樣本選擇所產生的問題，可能使高層級院所的處置效果遭到低估，也可能使低層級院所的處置效果被高估。但本文樣本中兩類觀察值數目接近，平均後的處置效果雖仍可能有偏誤，但應不至於對整體處置效果的估計產生太大影響。

對於樣本選擇所產生的內生性，文獻上有三種處理方法：工具變數，追蹤資料，以及 Heckman 樣本選擇模型。受限於現有的資料，我們並沒有適當的工具變數。追蹤資料是控制不可觀測因素的有效工具；然而即使有追蹤資料，實證上仍然很難估計處置效果。舉例來說，在醫學中心就醫的病患每年都有數百人，因此我們有充分的院內樣本來進行配對。但診所收治的病患每年可能只有個位數，造成配對上的困難。為了估計，我們勢必仍得併用橫斷面資料（即其他診所的病患資料）來配對，但這也違背了用追蹤資料控制不可觀測因素的原意。另一種可能性是直接對醫院與病人的決策行為建立模型，再用 Heckman 兩階段估計法來修正內生性與不可觀測因素的影響。我們遭遇最大的難題在於病人決策方面的建模。一般的作法，如 Kessler and McClellan (2000: 577-615) 是以病人居住地為圓心，選定一定範圍內的醫院作為病人選擇就醫的替選方案集合，再以病人與醫院的特性為外生變數估計其選擇行為。但本文資料中缺乏地址資訊，故也無法建立適當的選擇模型。

我們在實證作法上既然未能對樣本選擇性加以調整，故本文估計上的偏誤仍可能存在，這是本文最大的侷限，也是解讀實證結果時必須特別注意的。

10 關於醫療品質，我們可採用醫策會的醫院評鑑結果做為衡量醫療品質的客觀標準。然而本文未針對此項因素加以控制的原因有二。第一，診所層級並無評鑑。第二，在估計整體醫院的平均處置效果時，我們也已經控制了醫院層級。由於醫院層級和醫療品質兩者高度相關。所以控制了醫院層級，也多少控制了醫療品質的影響。

陸、結論

本文針對台灣的結核病論質計酬計畫進行制度評量。我們採用 Abadie and Imbens (2006: 235-267) 的最近鄰域配對估計式，估計此計畫的平均處置效果（結核病人完治率提高的程度）。我們發現此一制度使整體樣本的完治率提高 13%，此數字遠低於未控制變因之下所估計的處置效果。若改用 logit 估計處置效果、刪除轉院樣本、及分析樣本選擇偏誤的影響下，此結論均大體不變。

我們進一步發現在不同醫療院所層級下，此計畫只對診所與地區醫院有顯著效果，而且對診所的效果較大，但對區域醫院和醫學中心則無效果。若進一步考慮醫療院所權屬，我們發現實施此一計畫對私立院所的平均處置效果高於公立院所，而對財團法人院所則無顯著效果。我們提出兩種假說來解釋不同醫療層級處置效果的差異：從生產面及誘因面來解釋其差異性。然而受限於資料，我們難以處理樣本選擇的問題，也沒有額外的資訊可以檢定我們提出的假說。如何發展實證方法來解決這些困難，並適當解釋不同醫療層級處置效果的差異，使之成為健保制度興革的重要參考資料，將是我們未來研究的方向。

參考資料

A. 中文部分

中央健康保險局

- 2004 〈全民健康保險結核病醫療給付改善方案試辦計畫〉。2007年4月3日，取自 <http://www.nhi.gov.tw/>
- 2005 「加入結核病醫療給付改善試辦計畫的醫療院所名單」，Excel 資料檔，中央健康保險局 2005 年提供。
- 2007 〈健保特約醫療院所名冊壓縮檔〉。2007 年 5 月 5 日，取自 <http://www.nhi.gov.tw/>

王怡婷、莊乃蓉、黃曼媛、許清曉

1999 〈台灣地區結核病防治績效不彰的原因探討及建議〉，《院內感染控制》9: 263-270。
行政院衛生署疾病管制局

2005 〈結核病資料庫〉。Excel 資料庫，購自行政院衛生署疾病管制局。

2006 《結核病診治指引》。台北：行政院衛生署疾病管制局。

胡曉雲、蔡文正、龔佩珍

2005 〈肺結核病患未完成治療原因探討〉，《台灣衛誌》24: 348-359。

許績天、連賢明

2007 〈賺得越少，洗得越多？——台灣血液透析治療的誘發性需求探討〉，《經濟論文叢刊》35: 415-450。

蔡文正、龔佩珍、江自得

2002 〈全民健康保險肺結核病患醫療改善方案試辦計畫之評估〉。行政院衛生署九十一年度委託研究計畫。

蔡文正、龔佩珍、沈光漢、施純明

2005 〈健保結核病專案對於結核病防治工作的影響評估〉。行政院衛生署疾病管制局九十四年度科技研究發展計畫。

B. 英文部分

Abadie, A., D. Drukker, J. L. Herr, and G. W. Imbens

2001 "Implementing Matching Estimators for Average Treatment Effects in Stata," *The Stata Journal* 1: 1-18.

Abadie, A. and G. W. Imbens

2006 "Large Sample Properties of Matching Estimators for Average Treatment Effects," *Econometrica* 74: 235-267.

Angrist, J. D. and V. Lavy

2001 "Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools," *Journal of Labor Economics* 19: 343-369.

Bodenheimer, T., H. May, R. A. Berenson, and J. Coughlan

2005 "Can Money Buy Quality? Physician Response to Pay for Performance," *Center for Studying Health System Change* 102: 1-4.

Brewer, T. F., S. J. Heymann, and M. Ettling

1998 "An Effectiveness and Cost Analysis of Presumptive Treatment for Mycobacterium Tuberculosis," *American Journal of Infection Control* 26: 232-238.

Chou, S. Y., J. T. Liu, and J. K. Hammitt

2003 "National Health Insurance and Precautionary Saving: Evidence from Taiwan," *Journal of Public Economics* 87: 1873-1894.

Dehejia, R. H.

2005 "Practical Propensity Score Matching: A Reply to Smith and Todd," *Journal of Econometrics* 125: 355-364.

Dehejia, R. H. and S. Wahba

1999 "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of

- Training Programs," *Journal of the American Statistical Association* 94: 1053-1062.
- Hahn, J.
1998 "On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects," *Econometrica* 66: 315-331.
- Heckman, J. J., H. Ichimura, and P. E. Todd
1997 "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Econometrica* 64: 605-654.
- Imbens, G. W.
2004 "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review," *Review of Economics and Statistics* 86: 4-29.
- Josephine, G.
1994 "Controlling Resurgent Tuberculosis: Public Health Agencies, Public Policy, and Law," *Journal of Health Politic Policy, and Law* 19: 107-147.
- Kessler, D. P. and M. B. McClellan
2000 "Is Hospital Competition Socially Wasteful?" *The Quarterly Journal of Economics* 115: 577-615.
- Lee, M. J.
2005 *Micro-Econometrics for Policy, Program, and Treatment Effects*. New York: Oxford University Press.
- Millenson, M. L.
2004 "Pay for Performance: The Best Worst Choice," *Qual. Saf. Health Care* 13: 323-324.
- Rosenbaum, P. R.
2002 *Observational Studies*. New York: Springer-Verlag.
- Rosenthal, M. B., R. Fernandopulle, H. S. R. Song, and B. Landon
2004 "Pay for Quality: Providers' Incentives for Quality Improvement," *Health Affairs* 23: 127-141.
- Sloan, F.
2003 "Not-for-Profit Ownership and Hospital Behavior," pp.1142-1174 in Frank A. Sloan (ed.), *Handbook of Health Economics*. Amsterdam: Elsevier Science.
- Smith, J. A. and P. E. Todd
2001 "Reconciling Conflicting Evidence on the Performance of Propensity Score Matching Methods," *American Economic Review* 91: 112-118.
- Smith, P. C. and N. York
2004 "Quality Incentives: The Case of U.K. General Practitioners," *Health Affairs* 23: 112-118.
- Trude, S., M. Au, and J. B. Christianson
2006 "Health Plan Pay-for-Performance Strategies," *The American Journal of Managed Care* 12: 537-542.
- Winkelmann, R.
2006 "Reforming Health Care: Evidence from Quantile Regressions for Counts,"

Journal of Health Economics 25: 131-145.

Wooldridge, J. M.

2002 *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.

附錄一、結核病相關醫療名詞及定義

根據（行政院衛生署疾病管制局，2006）所發佈之結核病診治指引，我們將一些重要名詞解釋整理於此附錄。

個案定義：

1. 疑似個案：任何有疑似結核症狀之病人，尤其是咳嗽三週以上者。
2. 結核個案：A. 細菌學確診之病人：結核菌培養陽性之病人。B. 其他確診之病人：非細菌學確診之病人，但醫師診斷為結核病，並決定施予一完整療程之抗結核治療之結核病人。

細菌學（顯微鏡檢痰塗片檢查結果）：

1. 痰塗片陽性肺結核：
 - (1)至少兩次顯微鏡檢痰塗片陽性；或
 - (2)至少一次顯微鏡檢痰塗片陽性、且經醫師判定胸部X光之病灶符合肺結核之變化、決定施予一完整療程之抗結核治療；或
 - (3)至少一次顯微鏡檢痰塗片陽性且該檢體結核菌培養陽性。
2. 痰塗片陰性肺結核：

定義一：

- (1)至少三套痰檢體顯微鏡檢痰塗片檢查皆為陰性；且
- (2)胸部X光之病灶符合活動性肺結核之變化，且
- (3)臨床上對一週之廣效抗生素治療無反應；且
醫師決定施予一完整療程之抗結核治療。

定義二：

- (1)顯微鏡檢痰塗片陰性但痰培養陽性之病人
- (2)醫師決定施予一完整療程之抗結核治療。

通報：

1. 細菌學確診之病人及其他確診之病人皆應通報。
2. 細菌學確診之病人即使確診前死亡而未接受抗結核藥物治療亦應通報。

3. 所有接受抗結核藥物治療之結核病人皆應通報。
4. 實驗室應通報所有耐酸性染色陽性及結核菌培養陽性之病人。

治療結果：

1. 治癒 (Sputum negative, cured)：病人於治療過程中至少一次痰陰性且最後一個月之治療時痰陰性。
2. 完治 (Sputum examination not done, treatment completed)：病人已完成治療但其痰檢驗無法歸類為治癒或治療失敗。
3. 失敗 (Sputum positive, failure)：新病人於治療第五個月依然痰陽性、復治病人於治療最後依然痰陽性，或者治療前痰陰性、治療二個月後變成痰塗片或培養陽性的病人。
4. 死亡 (Died)：病人於治療過程中因任何死因死亡。
5. 失落 (Defaulted)：連續中斷治療兩個月以上、或在評估治療結果時仍在治療中（登記季 (Quarter) 後 12 個月）。
6. 轉出 (Transferred)：病人轉到其他單位治療且其治療結果不詳。

附錄二、結核病專案健保支付標準

「全民健康保險結核病醫療給付改善方案試辦計畫」是將整個治療過程分成幾個階段，依每個階段治療結果給予費用加成。此費用包含疾病管理照護費、確診診察費、完成治療診察費、特殊個案管理獎勵費、特殊個案醫師照護診察費、結核病個案DOTS (Directly Observed Treatment Short-Course) 執行服務費以及部份檢驗檢查費用。未列入計畫內的醫療費用支出標準，則按現行「全民健康保險醫療費用支付標準」核實申報。若個案於任一階段皆未符合連續管理，疾病管理照護費將折半給付，且該個案後續排除試辦計畫外申報（中央健康保險局，2004）。此計畫的支付標準，民國 93 年修定適用，以下為通則。

- 一、本試辦計畫支付標準按表訂項目及點數辦理，未訂項目按現行全民健康保險醫療費用支付標準辦理。
- 二、符合下列條件者，各階段結核病管理照護費按本標準所定點數支付。

(一)每一階段必須符合連續管理。

(二)每一階段至少需鍵入衛生署疾病管制局資料庫：

1. 一次胸部 X 光檢查結果；肺外結核者（肋膜積水除外）第二階段後免輸入胸部 X 光檢查結果。
2. 一次驗痰結果（痰培養陽性者需含藥物敏感試驗結果）；肺外結核者（肋膜積水除外）第二階段後免輸入驗痰檢查結果。
3. 一次生化檢查結果。
4. 肺外結核者第一階段需輸入病理檢查或結核菌培養結果。

(三)需附個案管理紀錄（每月乙次）、就診紀錄。

三、不符合上述條件者，但附有管理紀錄、就診紀錄（如為院所轉診，請加附轉診紀錄）者，各階段疾病管理照護費折半給付（即按 P1317C、P1327C、P1337C、P1347C 申報）。

支付標準 編 號	診療項目	支付點數
P1316C P1310C P1317C	第一階段（確診且連續服藥三個月）： 1. 期間：個案經檢驗檢查結果確認診斷、通報，且連續投藥滿三個月後病患回診當次申報此費用。 2. 內含第一階段疾病管理照護費（P1316C）1500 點及結核病例醫師確診診察費（P1310C）500 點。 3. 若未連續管理 3 個月，則每案每月個案管理費（P1317C）折半給付 250 點（每月）。 註：原支付標準 A1001C 結核病例發現診察費依現行申報方式申請費用。	1500 500 250/月
P1326C P1320C P1327C	第二階段（連續服藥第四至第六個月）： 1. 期間：自個案投藥第四個月起，實施相關醫療檢驗，並連續投藥滿三個月。 2. 內含第二階段疾病管理照護費（P1326C）1500 點，如於六個月內完治，並得申報結核病例醫師完成治療診察費（P1320C）1000 點。 3. 若未連續管理 3 個月，則每案每月個案管理費（P1327C）折半給付 250 點（每月）。 註：原支付標準 A1002C 結核病完成治療費依現行申報方式申請費用。	1500 1000 250/月

P1336C	第三階段（連續服藥第七至第九個月）： 1. 期間：自個案投藥第七個月起，實施相關醫療檢驗，持續投藥至病例經檢驗檢查結果確認完成治療。	500
P1320C	2. 於完治前每月給付疾病管理照護費（P1336C）500點，如於九個月內完治，並得申報結核病例醫師完成治療診察費（P1320C）1000點，於完治時一併給付。	1000
P1337C	3. 若未連續管理，則每案每月個案管理費（P1337C）折半給付 250 點（每月）。 註：原支付標準 A1002C 結核病完成治療費依現行申報方式申請費用。	250/月
P1346C	第四階段（連續服藥第十個月至第十二個月） 1. 本階段需符合以下條件之一：①抗藥性個案（含多重抗藥性個案）②糖尿病個案③前六個月用藥治療時有嚴重副作用需停藥試藥之個案④肺外結核（肋膜積水者除外）⑤其他，經「結核病診療諮詢小組」審核通過者。 2. 期間：自個案投藥第十個月起，實施相關醫療檢驗，持續投藥至病例經檢驗檢查結果確認完成治療。	500
P1320C	3. 於每月給付疾病管理照護費（P1346C）500點，完治時並得申報結核病例醫師完成治療診察費（P1320C）1000點；另如為多重抗藥性個案，於完治時得再申報特殊個案管理獎勵費（P1330C）2000點及特殊個案醫師照護診察費（P1340C）1000點。	1000
P1330C		2000
P1340C		1000
P1347C	4. 若未連續管理，則每案每月個案管理費（P1347C）折半給付 250 點（每月）。 5. 連續服藥超過第十三個月起者，不再繼續給付疾病管理照護費，惟如為多重抗藥性個案，於完治時得申報特殊個案管理獎勵費（P1330C）2000點及特殊個案醫師照護診察費（P1340C）1000點。 註：原支付標準 A1002C 結核病完成治療費依現行申報方式申請費用。	250/月
P1330C		2000
P1340C		1000

P1301C	痰液抹片耐酸菌顯微鏡檢 P1301C 註：不得與全民健康保險醫療費用支付標準 13006C 重複申報。	75
P1302C	分枝桿菌培養 P1302C 註：不得與全民健康保險醫療費用支付標準 13012C 重複申報。	240
P1303C	分枝桿菌藥物敏感性試驗 P1303C 註：不得與全民健康保險醫療費用支付標準 13015C 重複申報。	520
P1304C	抗酸菌鑑定 P1304C 註：不得與全民健康保險醫療費用支付標準 13013C 重複申報。	440
P1341C P1342C	結核病個案 DOTS (Directly Observed Treatment Short-course) 執行服務費 由特約藥局或設有藥師之特約院所提供 (每月)。 由護理人員居家提供 (每月)。 註：1.配合衛生署疾病管制局 DOTS 推廣計畫，且遵照其 DOTS 執行作業規範執行 DOTS 服務之特約醫事服務機構得申報本項費用。 2.申報 DOTS 執行服務之個案以占所照護結核病個案之 30%為目標，山地離島地區則應全力推動。 3.住院個案不得申報。 4.P1341C 及 P1342C 不得同時申報 5.如於病患開始治療或結束治療當月，未滿三十天者，按治療天數比例申報費用。	600 1200

Evaluating a Quality Incentive Program for TB Treatment in Taiwan*

Yu-ting Hsieh

Clerk, General Affairs Office
Kaohsiung Municipal Hsiao-Kang Hospital

Chung-ming Kuan

Professor
Department of Finance, National Taiwan University

Diana HweiAn Tsai

Professor
Department of Economics, National Sun Yat-Sen University

Yu-wei Hsieh

Doctor Candidate
Department of Economics, New York University

ABSTRACT

Pay-for-quality is a new reimbursement scheme around the world. In this paper, we evaluate the pay-for-quality program of the National Health Insurance in Taiwan. We estimate the average treatment effect (ATE) of this program on the success rate of tuberculosis treatment. In particular, the nonparametric nearest neighbor matching estimator is employed. We find the treatment success rate is improved up to 13% on average for the hospitals that participate in the program. Conditional on the hospital level, this program is most effective in clinics and local

* Author for correspondence: Chung-Ming Kuan, Department of Finance, National Taiwan University, 1, Sec. 4, Roosevelt Road, Taipei 106, TAIWAN; ckuan@ntu.edu.tw

hospitals. The ATEs are 46% and 10%, respectively. For regional hospitals and medical centers, the estimated ATEs are insignificant. Conditional on the hospital ownership, private hospitals have higher ATEs than public hospitals at the level of clinic and local hospital. The ATE of non-profit proprietary hospitals is insignificant at all levels. Moreover, we show that the ATE of this program would be largely inflated if relevant confounders are not controlled properly.

Key Words: average treatment effect, matching estimator,
pay-for-quality, program evaluation, tuberculosis,
treatment success rate