

## 醫療法律案例之實證醫學分析--以胃癌案件為例

Evidence Based Analysis of Legal Issues - Applied on a Gastric Cancer Case

陳杰峰、吳俊穎<sup>a</sup>、徐圭璋<sup>b</sup>、郭耿南<sup>c</sup>、陳榮基<sup>c,d</sup>

台北醫學大學 市立萬芳醫院、台中榮民總醫院<sup>a</sup>、

高雄榮民總醫院<sup>b</sup>、國家衛生研究院<sup>c</sup>、恩主公醫院<sup>d</sup>

### 摘要

本文藉著一個實際胃癌醫療法律案例，以實證醫學的角度來探討。評估文獻證據等級或證據力有其困難度，不同的學者、不同的研究機構，可能會有不同證據等級的評鑑結果。在法律上，重視的是證據的證據能力，而在醫學上強調醫學研究文獻的證據等級。傳統上，社會尊重權威和經驗，會依據專家意見來決定正確的醫療行為，而實證醫學所強調的是以醫學文獻為依據。

本案例之特色是一個有關胃鏡檢查之診斷及預後之爭議，相對於治療型爭議，這部份較少被討論到。因此本文以實證醫學之五步驟，分別演述學理之觀點，及面對臨床不確定性時，演算機率之邏輯工具，如敏感度(sensitivity, Sn)，特異度(specificity, Sp)，陽性預測值(positive predictive value)，陰性預測值(negative predictive value)，陽性相似比(positive likelihood ratio, LR+)，陰性相似比(negative likelihood ratio, LR-)。希望藉由這些實證醫學的方法，讓醫法雙方形成一個理性討論的平台。

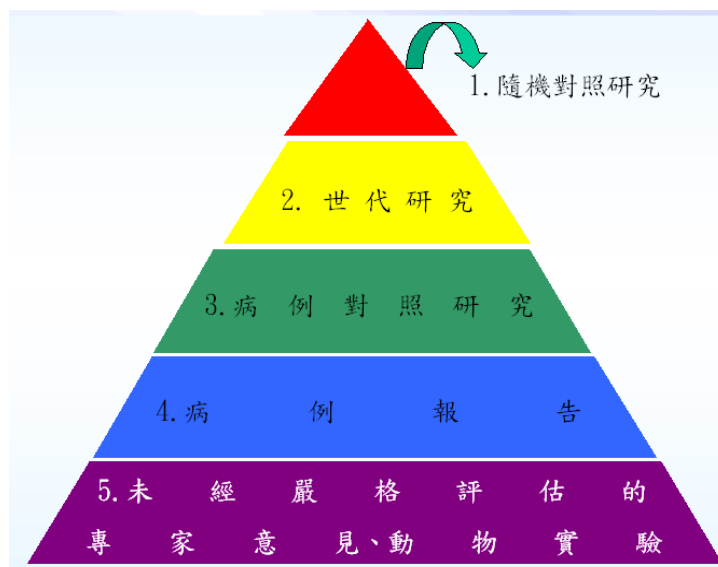
臨床醫療工作有其不確定性，常存在被法律定罪的風險。本文僅就文獻證據，以實證醫學的角度來做分析，真正的法律定論，仍有待社會中各法學領域專家討論後，形成實務操作所依據之共識。

## 前言

在行政院衛生署指導及國立台灣大學醫學院協辦之下，台灣實證醫學學會以及財團法人國家衛生研究院於 2008 年 6 月，共同主辦了國內第一次“實證醫學與法律工作坊”，活動的特色是與會人員法界醫界各佔一半，法界包含法官、檢察官、律師以及法學專家學者等，醫界則以醫師、各類臨床工作者、實證醫學專家及其他相關學者為主。在這次為期半天的工作坊，這二大專業團體針對醫療法律案件，經由講課及混合編組之分組討論，做了初步的意見交換。

醫學文獻在法律上是有證據能力的。根據我國大法官釋字第五八二號解釋，定義“證據能力”為：「係指證據得提出於法庭調查，以供作認定犯罪事實之用，所應具備之資格；此項資格必須證據與待證事實具有自然關聯性，符合法定程式，且未

受法律之禁止或排除，始能具備。如證人須依法具結，其證言始具證據能力；被告之自白，須排出於不正之方法，始具證據資格」[1]。美國聯邦法院明確要求聯邦法院，有義務確認專家、證人所提供的證詞及與案件相關之證據是可靠的。Daubert 認為只要具有下述三項特點，就可以形成法律上面，具有證據能力的文獻：(1)證據是合理的，(2)在科學上是有根據而正確的，(3)該證據可以應用在該特定案件上[2]。一旦擁有證據能力後，法院可以本於自由心證，在不違背經驗法則及論理法則的情況下，決定該證據的證明力。於法，法律上重視文件的證據能力及證據力；於理，醫學上強調醫學研究文獻的證據等級。評估文獻證據等級有其困難度，不同的學者、不同的研究機構，可能會有不同證據等級的評鑑結果。例如，醫學裡面的治療型文獻，其證據等級是依照研究品質而分類，如圖一。



圖一：證據等級(改編自 Oxford Center, 2001) [3]

實證醫學中，治療文獻的 Level 1 證據是隨機對照研究，而 Level 2 是世代研究，Level 3 是病例對照研究，Level 4 是病例報告，而其他未經嚴格評估的專家意見及動物實驗則為 Level 5[3]。

### 案例介紹

在這次的工作坊，藉著醫療法律案例而進行分組討論，案例事實摘錄如下[1]：

A 病患自費至 X 醫院進行健康檢查，經 B 醫師以胃鏡檢查病患胃部，發現螢幕上有黑色狀。B 醫師稱係胃潰瘍並無大礙，吃藥即可。四個月後，A 病患發現體重有不正常下降，經病理檢查發現異常，後經外科醫師將病患胃脾全數切除，其後接續做六次化學治療，終因病情嚴重不治死亡。

原告 A 病患家屬主張：A 病患因 B 醫師診斷及治療之錯誤而錯失治療胃癌之先機，終致病逝。按依 Y 醫院消化系惡性腫瘤報告，「若局限在胃黏膜層有 95% 的治癒率，若至黏膜下層仍有 92%，僅肌肉內層還有 76%」，依 A 病患之體檢報告，「食道、十二指腸、乙狀結腸」均無異常，僅「胃前壁潰瘍」，若非被告 B 醫師之過失延誤，顯有 95% 之治癒機率，故被告 B 醫師所辯 A 病患之死亡與其無關云云，應非事實。

被告 B 醫師主張：專家證人 Z 醫院外科 C 醫師，於被告 X 醫院檢視 A 病患之病歷原本、前後兩次間隔四個月的胃鏡照片及隨後的手術紀錄之後，說明為：A 病患所罹患的胃腺癌，是一種分化不良十分惡性之癌症，在手術時，癌細胞已侵蝕穿透

胃壁，在醫理上已是第三期的胃腺癌。就時間推算，在 A 病患體檢時，當時應有胃腺癌之存在，「在當時應該若不是胃腺癌第三期就是第二期，不可能是第一期」，「從開刀時之病理報告顯示，是第三期之胃腺癌，很難說早四個月發現有何影響」等語，也就是說 A 病患若是在接受首次胃鏡後不久，即發現胃腺癌而接受開刀切除，其結果與 A 病患在四個月後再做切除手術，其「預後存活率之結果沒差異」，A 病患因胃腺癌死亡之結果，並不因被告未及時予以「切片檢驗」而有所差異。

根據法院見解，依被告 B 醫師所提出之日本 D 醫師所發表之「Some Factors which Influence Prognosis after Surgery for Advanced Gastric Cancer」之文章[4]，胃癌二期之五年存活率為 39.4%，三期之五年存活率是 32.4%，而澳洲 E 醫師所發表之文章「Gastric Remnant Cancer Has a Better Prognosis Than Primary Gastric Cancer」[5]，則認為胃癌二期之五年存活率為 25%，胃癌三期之五年存活率是 20%，而依據原告 A 病患家屬所提出之台灣某醫院關於胃癌之說明，認為第二期之五年存活率約為 50%，而第三期之五年存活率則僅為 15% 或更少，有該等兩造所不爭執各文獻附卷可參。足見第二期及第三期之胃癌，其五年之存活率確有差異，被告主張無差異云云，尚無可採。至於被告 B 醫師另所提之其他文獻，所為說明或是指二年之存活率、或是指三年之存活率，均非在說明五年之存活率，自難據而認為二期或三期胃癌之五年存活率無差異。A 病患在接受體檢當時之胃癌屬於

第二期之機率極大，而被告 B 醫師復未舉證說明 A 病患之胃癌在體檢當時，確已屬第三期，則當時 A 病患之胃癌情形，應係第二期較接近實情，而參酌前開說明，胃癌二期及三期之五年存活率不同，則本件如能早期（即在體檢時）即發現，關於 A 病患之五年存活率自屬較高，茲因被告 B 醫師之過失，未能及早發現 A 病患身上之胃癌，而錯失治療胃癌之先機，致 A 病患之五年存活率及治癒率均降低，並致 A 病患因治療無效而死亡，是該 A 病患之死亡之結果，自應認為與被告 B 醫師之上開過失行為有相當因果關係。

### 實證醫學分析

在操作實證醫學分析時，可以依下列 5 個步驟執行[6]：

- I. 整理出一個可以回答的問題( Ask)
- II. 尋找文獻的最佳證據 (Acquire)
- III. 嚴格評核相關證據 (Appraise)
- IV. 將評核結果整合至臨床專業知識及病人意願作決定 (Apply)
- V. 對上述四個步驟進行稽核 (Assess)

本文將依這些步驟分析此案例。

#### Step 1 Ask：形成問題

國內法律案例的裁判書，可由司法院法學資料檢索系統[1]搜尋到(前述大法官釋字第 582 號解釋文，亦可自此資料庫獲取)，只要在資料庫的裁判書查詢鍵入”胃癌”，即可找到判決書之全文。此檢索系統內之裁判書，類似本案的”胃癌病患胃

鏡偽陰性(false negative)結果”之訴訟案件不只一例，此外，在實證醫學與法律工作坊之綜合討論中，亦有與會醫師分享，其所服務之醫院也有類似的訴訟案件，因此我們可以推論，在真實醫療環境中類似的案件一定不少。本工作坊之討論，主要是下列二點：(1)本土的醫學文獻與國外之文獻，在證據等級上是否有所不同?在訴訟上的證據能力及證明力是否有所不同?(2)原告引用相關醫學文獻證明存活率的差異，專家證人則就具體個案判定，未即時切片對於死亡結果之發生無影響，法院最後認定存活率有差異，被告醫師是否須負損害賠償責任?緣此，本文嘗試著從實證醫學的角度來做本醫療法律案件之分析。本案為有關診斷及預後之問題，客觀數據如為胃鏡檢查診斷的敏感度(sensitivity)以及特異度(specificity)及正確率(accuracy)是多少?以及漏診[miss diagnosis)，亦即為”偽陰性”，有別於誤診(misdiagnosis)]機率多大?手術後可以存活多久?

首先，實證醫學中建議先將問題形成 PICO，即問題四分法(4-part question)。所謂 PICO，是 P (patient；population)－病人或病人族群、I (intervention)－醫療介入方式、C (comparison)－比較可用的療法及 O (outcome)－可衡量的結果。若將本案依問題四分法分析，將形成 PICO 如下：

P (patient，病人)－自費健康檢查者，經胃鏡檢查發現螢幕上有黑色狀，診斷為胃潰瘍。

I (intervention，介入方式)－胃鏡檢查

C (comparison, 比較) – 胃鏡檢查並切片

O (outcome, 結果) – 診斷正確率(診斷), 死亡率或存活率(預後)

PICO 問題的形成後, 第 2 步驟搜尋所需的關鍵字, 也呼之欲出了。

## Step 2 Acquire : 文獻搜尋

進行文獻搜尋的時候, 比較省力的方式, 是從衍生性期刊 (secondary journals) 或是已經有整理好結論的資料庫下手, 臨床應用時, 亦包含診療指引 (clinical practice guideline, CPG)。方便記憶之口訣為 "Secondary first, primary second", 意即 "先找衍生性文獻, 再搜尋原始著作"。本案件有可參考的診療指引嗎? 搜尋診療指引時, 第一個想到的就是美國的國家指引資料庫 (<http://www.guideline.gov/>), 我們鍵入 Key word: gastric cancer, 可以找到 Management of oesophageal and gastric cancer. A national clinical guideline [7], 但是在指引內容描述到有關於胃鏡診斷胃癌的部份, 並沒有找到和本案件有關的資料。另外又鍵入 "biopsy protocol", 找到一篇 Tissue sampling and analysis [8], 但沒有可以解釋本案之診斷流程。

接著進入 PubMed 資料庫 (<http://www.ncbi.nlm.gov/sites/entrez>)。首先, 以 MeSH database 校正搜尋關鍵字。鍵入 "gastric cancer", 出現 "Stomach Neoplasms" [MeSH]; 接著在 PubMed 中鍵入 "Stomach Neoplasms" [Mesh] AND "Diagnostic Errors" [Mesh] AND "Endoscopy" [Mesh], 搜尋到 183 篇文獻, 其中有 7 篇文獻, 第一

篇即為 2008 年日本之胃癌篩檢指引 [9]。

其次, 可以善用各資料庫之文獻品質篩選器 (filter), 如 PubMed 的 clinical query, 可以有效率地找到關鍵文獻。本案例即可在 clinical query 點選 diagnosis 和高特異度 (narrow, specific search) 之選項, 鍵入 "Stomach Neoplasms" [MeSH] AND "Biopsy" [MeSH], 找到 56 篇文獻, 再聚焦於探討胃癌病患在胃鏡沒有檢驗出之文獻, 茲分析如下。

## Step 3 Appraise : 嚴格評讀

EBM 教科書 [6] 指出, 在文獻評讀的部份和預後有關的證據是否有效時, 必須檢測下列幾點:

1. 某特定的、具代表性的病患樣本, 是否聚集在疾病過程的同一點?
2. 研究中的病患是否追蹤得夠久? 夠完整?
3. 研究結果是否有客觀的標準, 並以 "盲法" 表現?
4. 假如分成不同預後的次族群, 是否是針對重要預後因子作調整?

首先參照 2008 年日本之胃癌篩檢指引 [9]。本文獻根據一篇中國之研究報告 [10], 主張使用胃鏡篩檢胃癌, 其胃癌死亡率和標準死亡率之風險比值為 1.01 (95% CI: 0.32 - 1.37), 並無統計學之顯著差異。

原始文獻的第一篇為 "評估以胃鏡檢查胃部腫瘤的診斷正確性" [11], 該文研究胃癌病患在胃鏡沒有檢驗出者, 其臨床病理特色及預後。在這篇芬蘭的文章裡, 收集了 1996-2001 年的病患, 納入 3.5 年內, 國家癌症登記系統裡面, 總共 284 個胃癌

病患，其中 13 位為做過胃鏡而結果為偽陰性者。在第一次做胃鏡的時候，其平均年齡是 72.4 歲，這 13 位經過胃鏡檢查後，平均延遲發現胃癌的時間是 11.5 個月，他們的細胞學檢查，全部都是胃癌 (gastric carcinoma)，這個研究的敏感度和特異度分別是 0.93 和 1.00。容我們在此再深入說明一下這幾個有關診斷的流行病學術語。一般檢驗結果分為陽性 (positive)、陰性 (negative)，用以判斷疾病的有、無，有病的人檢驗結果是陽性就稱為真陽性 (true positive)，沒有病的人檢驗結果是陰性就稱為真陰性 (true negative)，當然也是會有偽陰、偽陽 (false positive) (如表一)。下列這個 2 X 2 表 (two by two table) 是一種列聯表 (contingency table) 的表達方式，有利於對診斷機率的說明。

檢驗結果	疾病		
	有	無	
陽性 (positive)	真陽 (a)	偽陽 (b)	檢驗陽性 (a+b)
陰性 (negative)	偽陰 (c)	真陰 (d)	檢驗陰性 (c+d)
	有病 (a+c)	無病 (b+d)	

敏感度 (sensitivity, Sn) =  $a/(a+c)$

特異度 (specificity, Sp) =  $d/(b+d)$

陽性預測值 (positive predictive value) =  $a/(a+b)$

陰性預測值 (negative predictive value) =  $d/(c+d)$

陽性相似比 (真陽相似比) (positive likelihood ratio, LR+) =  $Sn/(1-Sp)$

陰性相似比 (偽陰相似比) (negative likelihood ratio, LR-) =  $(1-Sn)/Sp$

測前機率 (pre-test probabilities)  $(a+c)/(a+b+c+d)$

(註：當測前機率在描述整個 population 時，可稱作盛行率 (prevalence))

表一 檢驗結果及相關參考公式 [6]

這個列聯表 (contingency table) 有個綽號為 confusion table，因為其中眾多公式常讓人困惑。但是在評估檢測之有效性時，這個表格是有用的工具。流行病學家有時被定義為“一個透過 2 X 2 表格來看世界的人” [12]。

本研究敏感度和特異度分別是 0.93 和 1.00，可套用到陽性相似比 (positive likelihood ratio, LR+) 的公式 =  $Sn/(1-Sp)$ 。依照公式算出，陽性相似比為無限大，表示如果被胃鏡診斷為有胃癌，幾乎就是可以確定診斷，然而偽陰性的相似比 (negative likelihood ratio, LR-) 公式為  $(1-Sn)/Sp$ ，計算出來，偽陰性的相似比為 0.07，表示陰性的檢查並不排除有胃癌的可能性，然而在這些病患當中，死亡的病患其平均存活的時間，沒有漏診的 (n=191) 為 9.4 個月，而漏診的 (n=10) 為平均 7.3 個月，其死亡的病患其平均存活的時間，並沒有統計學上的差異 (p=0.15)，所以由這一個全國性的登錄資料結論顯示，的確有少部份的胃癌，沒有辦法經由胃鏡而檢查出來，所以造成診斷上的延遲，但是延遲診斷以及沒有延遲診斷的胃癌病患，他們的預後是同樣的不好。文獻中診斷特異度為 1.00 的情形，可以用“SpPin” (高特異度納入法) 的觀念來解釋。SpPin 即是當一個徵兆/檢查/症狀有高特異度 (Sp) 時，一個陽性結果 (P) 可以幫忙納入 (in) 診斷 [6]。也就是說，被胃鏡診斷為有胃癌，幾乎就是可以確定診斷；但是並不表示未被胃鏡診斷為有胃癌者沒有胃癌。由於這一篇是以全國性的登錄資料來做分析，我們在參照牛津大學實證醫學中

心所擬定的有關診斷的文獻等級，如表二，將這篇文獻歸類為 1c 等級：

Level	Diagnosis
1a	SR (with homogeneity*) of Level 1 diagnostic studies; CDR† with 1b studies from different clinical centres
1b	Validating** cohort study with good††† reference standards; or CDR† tested within one clinical centre
1c	Absolute SpPins and SnNouts††
2a	SR (with homogeneity*) of Level >2 diagnostic studies
2b	Exploratory** cohort study with good††† reference standards; CDR† after derivation, or validated only on split-sample§§§ or databases
2c	
3a	SR (with homogeneity*) of 3b and better studies
3b	Non-consecutive study; or without consistently applied reference standards
4	Case-control study, poor or non-independent reference standard
5	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"

表二 診斷類文獻之證據等級。

Oxford Centre for Evidence-based Medicine Levels of Evidence (May 2001), <http://www.cebm.net/index.aspx?o=1047> (Produced by Bob Phillips, Chris Ball, Dave Sackett, Doug Badenoch, Sharon Straus, Brian Haynes, Martin Dawes since November 1998)[3]

接著，利用上一篇文章在 PubMed 的相關文獻(related articles)裡，發現最相關(related)的文獻為”上腸胃道癌症的漏診病患”(Missed diagnoses in patients with upper gastrointestinal cancers)[13]，在這英國

一系列的病患中，14 位胃癌病人在確診之前 3 年內有做過胃鏡，其中有 71.4%在剛開始時，胃鏡檢查的診斷為胃發炎、胃潰瘍或是有可能的病灶，情形和本案類似，其結論為經過胃鏡的診斷之後，而沒有檢查出的胃癌是常見的(frequent)。由這篇英國的文獻，可以知道在注意癌症早期的臨床症狀非常重要。如本案以胃鏡檢查病患胃部，發現螢幕上有黑色狀之可疑病灶時，可以考慮多處切片。這篇文獻證據等級為 2b。

相關文獻裡的第二篇是 Hosokawa 1998 年的研究報告[14]。這是一篇非常大規模的回溯性世代研究(retrospective cohort study)，證據等級為 1b，總共有 37094 個對有腸胃症狀者之胃鏡檢查(有別於健康者之篩檢)，交叉比對了日本福井縣(Fukui Prefecture, Japan,人口約 80 萬人)的癌症登錄系統，其中真陽，偽陽，以及偽陰分別是 659 人、6 人及 155 人，計算出敏感度(sensitivity)、特異度(specificity)、陽性預測值及陰性預測值分別是 81.0%、接近 100.0%、99.1%及 99.6%，胃鏡診斷正確率(overall diagnostic accuracy)整體而言是 99.6%。此時可以下載牛津大學實證醫學中心之公用軟體 CATmaker (<http://www.cebm.net/index.aspx?o=1216>)，並以此分析文獻之數值，可以估算其信賴區間(如圖二)，並推算出陽性相似比(LR+)為 4895(95% CI: 2198 - 10903)，偽陰性的相似比(LR-)為 0.19(95% CI: 0.17 - 0.22)。數值之估算，最好加上信賴區間(confidence

interval)。通常使用“95%信賴區間(95%CI)”，這表示：我們有95%的信心，相信母群體的估計值將落在這個範圍內。信賴區間越窄，我們對這個結果越有信心。信賴區間包含1，表示沒有統計上的顯著差異。

Your Question		the Study Patients	the Study Evidence	the Bottom Line	the Other Stuff
Analysis 1 of 1		TARGET DISORDER cancer			
		Present	Absent		
TEST	Positive	659	6	a	b
scope	Negative	155	36274	c	d
				95% Confidence Intervals	
SENSITIVITY		a / (a+c)		81 %	78 to 84
SPECIFICITY		d / (b+d)		100 %	100 to 100
Pre-test Probability ("Prevalence")		(a+c)/(a+b+c+d)		2 %	2 to 2
Positive Predictive Value:		a / (a+b)		99 %	98 to 100
Negative Predictive Value:		d / (c+d)		100 %	100 to 100
LIKELIHOOD RATIO +		sens / (1 - spec)		4895.27	2197.84 to 10903.29
LIKELIHOOD RATIO -		(1 - sens) / spec		0.19	0.17 to 0.22

圖二 以CATmaker分析數值，可以自動運算其信賴區間。

因此從這篇文獻裡，可以了解胃鏡對於胃癌的檢查效果是很好的，但是仍然有不可忽視的偽陰性結果，因此反覆持續的胃鏡檢查在胃癌的篩檢上非常重要。反覆持續的檢查，可以降低偽陰結果之可能性。“相似比(likelihood ratio)”在這裡表現出另一個特性：我們可以“串連”各個獨立檢查的相似比。我們將第一個診斷性檢查所得到的測後勝算比(post-test odds)，作為第二個檢查的測前勝算比(pre-test odds)。因此，我們可以不斷地將

此數值乘以下一個檢查所得的相似比。例如，本文獻中胃鏡篩檢的胃癌測前機率是2.2%。

$$(659 + 155) / 37094 = 2.2\%$$

轉換成測前勝算比亦接近2.2%

$$2.2\% / (1 - 2.2\%) = 2.2\%$$

本文獻中胃鏡篩檢的胃癌偽陰性的相似比(LR-)為0.19。則其胃癌偽陰性的測後機率就是他的測前勝算比，乘以其陰性相似比(LR-)(0.19)之乘積，將所得的測後勝算比(0.004)，再換算成機率%(除以其值+1)，即為測後偽陰性之機率，亦即：

$$(0.022 / 0.978) \times 0.19 = 0.004 ; 0.004 / (0.004 + 1) = 0.4\%$$

其實，這裡0.4%也就是偽陰155人除以總檢查數37094，計算式如下：

$$155 / 37094 = 0.4\%$$

當重覆檢測時，以相似比估算測後機率就很好用。例如套用本文獻之數字，做兩次胃鏡偽陰的機率會大幅下降到0.08%，計算式如下：

$$(0.022 / 0.978) \times 0.19 \times 0.19 = 0.08\% ; 0.08\% / (0.08\% + 1) = 0.08\%$$

做兩次胃鏡仍為偽陰的機率0.08%算是很低了。但是乘以文獻中37094個胃鏡檢查，仍然有30個偽陰性結果。也就是說，以胃鏡兩次當篩檢，福井縣仍然有30個醫師可能因偽陰性結果而被告。

有趣的是，那做3次胃鏡呢?答案是仍然有6個醫師可能因偽陰性結果而被告。反覆持續的胃鏡檢查在胃癌的篩檢上非常重要。問題是，要反覆檢查幾次，才可以



讓醫師免於被告的恐懼呢？我們可以模仿 number needed to treat (NNT, 治益比)的整數表達方式，將測後偽陰性之機率的倒數，命名為” number needed to sue (NNS, 治告比)”，即醫師每執行檢驗多少次，就有可能被告 1 次的風險，本文獻中胃鏡檢測胃癌之測後偽陰性之機率為 0.4%，其倒數為 250，亦即醫師每操作 250 次胃鏡，就有 1 次可能因為偽陰性結果被告，其治告比(NNS)為 250。若一位腸胃內科醫師一輩子做 1 萬件胃鏡，要讓偽陰性的機率小於 1，即治告比大於 1 萬，經過計算，每一可疑病灶必須反覆檢查 4 次，才能免於被告的恐懼，那時醫療環境會如 Andrew Grove 所言，”只有偏執狂才能生存”(Only the paranoid survive.)[15]。

而另外一篇英國 2002 年[16]的報告數據顯示，129 個做過胃鏡的胃癌病患，其中有 18 位(14%)在初次胃鏡中，未能診斷為胃癌，其診斷敏感度為 0.86。這篇文獻證據等級為 2b。

本訴訟案件所提出的文獻，一篇為 1968 年日本 Maruta 所發表的文獻[4]，另外一篇為 1994 年的澳洲報告[5]，這是二篇都是有關”手術治療後”追蹤病患存活率(預後)的文獻，因此我們在參照牛津大學實證醫學中心所擬定的有關預後的文獻等級，如表三：

Level	Prognosis
1a	SR (with homogeneity*) of inception cohort studies; CDR† validated in different populations
1b	Individual inception cohort study with $\geq 80\%$

	follow-up; CDR† validated in a single population
1c	All or none case-series
2a	SR (with homogeneity*) of either retrospective cohort studies or untreated control groups in RCTs
2b	Retrospective cohort study or follow-up of untreated control patients in an RCT; Derivation of CDR† or validated on split-sample§§§ only
2c	"Outcomes" Research
3a	
3b	
4	Case-series (and poor quality prognostic cohort studies***)
5	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"

表三 預後類文獻之證據等級。

Oxford Centre for Evidence-based Medicine Levels of Evidence (May 2001), <http://www.cebm.net/index.aspx?o=1047>Produced by Bob Phillips, Chris Ball, Dave Sackett, Doug Badenoch, Sharon Straus, Brian Haynes, Martin Dawes since November 1998.[3]

在 Maruta 的這篇文獻中[4]，收錄了 1953-1965 年 12 年間，總共 541 位因胃癌而做外科手術的所有病人，其中 382 位做胃除切手術，而 130 位做姑息性治療，但是本研究排除了手術後死亡的病例 9 位，以及因為其他病因死亡的 3 位，所以總共收錄 225 位可切除的病患，都追蹤了五年以上，所以這是一個很好的世代研究，這 245 位病患中，144 位接受胃全切除手術，而 101 位接受姑息性切除手術，而所有的 245 位病患，全部都完整的追蹤，因此其

追蹤完成比例是 100%，這是一個 Level 2b 的證據，然而在本文獻中，已排除手術後死亡的 9 位病患，關於這一點我們也應該讓病人知道，還是有在手術後一個月之內死亡的風險，而這風險並沒有在存活分析裡面呈現。本文中胃癌二期五年存活率 39.4%，三期之五年存活率 32.4%。這篇文獻有幾點需要進一步討論，第一個是 1968 年的文獻，距離現在為 40 年前，另外日本和台灣的醫療情形以及癌症的臨床表現，亦有可能有差異，這是值得進一步討論。

另外一篇為 1994 年的澳洲報告[5]，在這篇澳洲的報告中胃癌的存活率，胃癌二期五年存活率為 25%，三期五年存活率為 20%，是一家澳洲醫院所做的世代研究，總共追蹤了五年，這篇文獻等級是 2b。

如何找到本土資料呢？目前本土的資料其主要資料庫是 CEPS (<http://www.gcscholar.com/>)，在鍵入“胃癌”以後，我們亦發現了有本土的資料，是 1994 年的醫學報告，文獻顯示[17]，在手術後一個月內死亡的手術死亡率 (operative mortality) 是 3.4%，整體而言，胃癌病患接受手術後的五年整體存活率平均是 35.7%，其中 265 例(56.1%)屬於絕對治癒性切除，150 例(31.8%)為相對治癒性切除，57 例(12.1%)為非治癒性切除；其 5 年存活率分別為 54.4%，13.7%，及 5.3%。根據牛津大學的分類，這篇本土資料為 Level 2b。

## 討論

### Step 4 Apply : 實務應用

世代研究(cohort study)的定義為，研究者追蹤一組或多組罹患目標疾病的個體一段時間，並觀察是否出現所關心的結果，是回答診斷及預後問題最好的研究設計。儘管隨機試驗研究之參與者，無法完全代表罹患某疾病的母群體，隨機試驗也可以是預後資料的來源(尤其是這類研究通常包含詳細的基本資料文件)。所以經過這完整的實證醫學搜尋，無法確立切片和死亡之間的因果關係，但是告知切片之必要，為醫師之義務之一。

本案例之“診斷”爭點為：“…未能及早發現 A 病患身上之胃癌，而錯失治療胃癌之先機，致 A 病患之五年存活率及治癒率均降低，並致 A 病患因治療無效而死亡…” [1]，而本案例初次胃鏡為 1999/5/8，到切片診斷日 1999/9/3，約延遲 4 個月。若以延遲診斷的方向搜尋，所發現之文獻，較接近本案例病患之實際情形，芬蘭 2005 年的文獻[11]，認為胃鏡偽陰性結果到確診之平均延遲診斷期間為 11.5 個月，而本案例約延遲 4 個月，優於芬蘭這篇文獻的平均值。另一篇英國 2002 年[16]的報告數據顯示，平均延遲診斷期間，為 13 個星期(介於 3-102 星期之間)，其平均值和本案例類似。

另一問題為“4 個月對於胃癌的自然病程之意義是什麼？”參照一篇 2000 年的文獻[18]，觀察 56 位經切片確定診斷為早期胃癌後，6-137 個月未治療之患者，其中

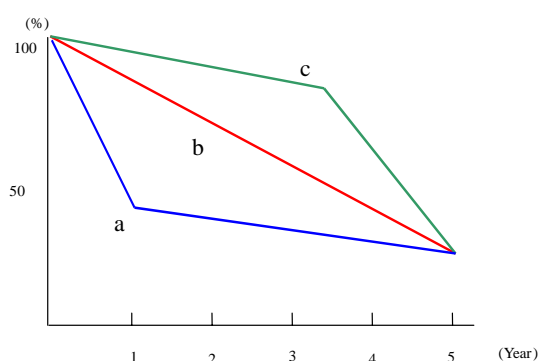
20 位仍為早期(early)胃癌，36 位演變為進展期(advanced)胃癌，維持在早期(early)胃癌的平均時間為 44 個月。4 個月對於胃癌的自然病程而言，不易讓早期胃癌演變為進展期胃癌；反推之，若確立診斷時為進展期胃癌，4 個月之前不大可能是早期胃癌。這和本法律案例之判決” … A 病患在接受體檢當時之胃癌屬於第二期之機率極大…” ，有所差異。

另一診斷爭點為” 作胃鏡如發現異常，是否有應作切片的標準流程(SOP)的 guideline?” 在第 2 步驟之搜尋時，相關指引並無明確描述 SOP[7] [8] [9]。前述之英國研究[13]建議，胃鏡檢查時，宜降低切片之閾值(threshold)，並考慮多處切片採樣，以提高早期胃癌篩檢率。偽陰性結果被視為胃鏡檢查伴隨的害處(harm)之一[9]，緣此，如果作了切片而呈偽陰性，醫師是否可據以免責?如果此論點可被法界接受，將來是否對遵照臨床診療指引(CPG)行事的醫師(即作胃鏡發現異常即作切片)，較有保障?當病人拒作切片時，應請其切結書面聲明，並登錄於病歷。如此，好的病歷記載，也可保障醫師。

本案例之” 預後” 爭點為：” …足見第二期及第三期之胃癌，其五年存活率確有差異…” [1]。如果胃癌由二期演變到三期，如文獻所描述，其五年存活率由 39.4% 降到 32.4%[4]，或由 25%降到 20% [5]，當病患最後死亡的時候，可以把這個結果全部歸結到因為延誤所增加的那 5-7%的風險嗎？還是它比較可能是原有的 75%死亡

率所造成的呢？換句話說，如果某個過失將風險從 75%升高到 80%，而風險真的發生了的時候，責任要算在這個過失者身上嗎?如果要，那麼只由 75%提高到 76%又如何？如果這種情形責任不必算在過失者身上，那麼當風險提高到 90%又如何呢？這比例的拿捏分寸到底在哪裡呢？實證醫學專家也許可以提出文獻數據以供參考，但是分寸尺度的拿捏上，仍然有待法界定奪。遺憾的是，在本案中僅以五年存活率的降低，就一概認定其過失與「死亡」有因果關係，似乎仍有討論空間。而如果認定其過失與「存活機率的降低」有因果關係，似為較公允的看法，但此時損害之客體則變為” 存活機率” 而非” 生命”，損害賠償之計算基礎亦隨之改變，賠償的金額當有不同之結果。

以五年存活率來評估或解釋案例時亦容易產生迷思。統計的描述，會影響閱聽者之感受。例如本案例中，若胃癌患者五年存活率 30%，會讓人覺得到五年後還有 30%的存活機率；若以存活月數衡量，本案例中，患者 1999 年 9 月手術，隔一年 10 月死亡，存活了 13 個月，其術後存活期，甚至超過芬蘭未漏診胃癌存活之平均值(7.3 個月)[11]。同樣五年存活率 30%，其實仍包含 3 種可能的過程：a、早期預後不佳，死亡率慢慢減緩；b、穩定的預後；c、早期預後良好，之後變壞。如圖三五年存活率 30%，有可能第一年已死亡 50%以上，如曲線 a。



圖三 5年存活率30%的3種可能進程

存活率會隨時間而改變，而且，早期追蹤通常比後期追蹤包含更多的病人及結果，因此，存活率曲線在追蹤的早期會比較準確(也就是，有較窄的信賴區間)。一個好的預後研究，其正文、表格或圖表中，應包括預後估計值的信賴區間。研究的病人數越少，信賴區間就越廣。但是，即使信賴區間很廣，它還是可以給我們一些提示，我們應該注意信賴區間的上下限。

手術本身存在的風險，在本案例中較少被提及，甚至所提出的文獻中[4]，有9位病人是手術後一個月內死亡，都被當做手術併發症死亡，不納入存活分析。這一點會讓我們在和病人解釋病情時，忽略了手術仍有存在的風險。

最後我們討論誤差的種類。本案例可歸類為診斷的第二型誤差(type II error( $\beta$ ))。流行病學中的誤差有二種，第一型誤差(type I error( $\alpha$ ))的定義是當二種實驗組結果實際上相同時，偵測兩組間具統計上顯著差異的或然率(也就是偽陽性結果的機會)。例如 Hosokawa 1998年報告指出[14]，事實上沒有胃癌的36280個胃鏡

檢查中，有6個被誤判為有胃癌，此為第一型誤差，機率小於0.0002，這個 $\alpha$ 值是相當低的；第二型誤差(type II error( $\beta$ ))的定義是當二種實驗組結果實際上是很不相同時，偵測兩組間不具統計上顯著差異的或然率(也就是偽陰性結果的機會)。這種偽陰性結果的機會就是第二型誤差，如 Hosokawa 的報告[14]，事實上應有胃癌的814個胃鏡檢查中，155個檢查無法偵測來，此為第二型誤差，機率約為0.19，這個 $\beta$ 值基本上算是有相當的顯著性。在這兩種難以避免的誤差下，醫療工作實際上冒著被刑法定罪的風險。這個流行病學的觀念，也可以套用在法律上：第一型誤差就是清白者被定罪，例如冤獄；第二型誤差為逍遙法外，無法讓有罪者定罪。現實生活中，法律工作同樣存在為這兩種難以避免的誤差(冤獄或縱放)。

### Step 5 Assess：對上述四個步驟進行稽核

- 1.我的問題是否已回答了？案例可以形成其他問題嗎？
- 2.有其他適用的關鍵字嗎？可以再用其他資料庫搜尋？
- 3.文獻中可能的偏差(bias)，是否已被充分評讀？
- 4.文獻應用在案例中，可能讓案情有不同結果嗎？

實證醫學的第五步驟，讓我們重新省思研究的過程，並開啓未來改善之可能。

### 結論

本案例為診斷之第二型誤差，是臨床診斷中難免會發生的，而法律中“應注

意，並能注意，而不注意者”也可以說是第二型誤差的一種，但是法律若以 0 誤差之標準(standards)，要求診斷，將使診斷之潛在社會成本提升，不見得符合群體之最大利益。

本文僅就文獻證據，以實證醫學的角度來做分析，真正的法律定論，仍有待社會中各法學領域專家討論後，形成實務操作所依據之共識。

### 誌謝

感謝國家衛生研究院研究計劃”醫學文獻與醫療行為準則做為醫療糾紛訴訟之證據的實證研究”(HD-097-PP-22)之支持，台灣實證醫學學會、財團法人國家衛生研究院工作同仁之協助，以及法界人士之熱心參與，使本實證醫學與法律工作坊圓滿順利。本文之完成承曾珮娟小姐彙整文章內容，謹致謝忱。

### 參考文獻

1. 司法院法學資料檢索系統. [cited; Available from: <http://jirs.judicial.gov.tw/Index.htm>.]
2. Daubert, v., *Merrell Dow Pharmaceuticals, Inc*, 1993. 133 S.CT.2786.
3. Bob Phillips, et al. *Oxford Centre for Evidence-based Medicine Levels of Evidence*. 2001 [cited; Available from: <http://www.cebm.net/index.aspx?o=1047>.]
4. Maruta, K. and Shida, H., *Some factors which influence prognosis after surgery for advanced gastric cancer*. *Ann Surg*, 1968. **167**(3): p. 313-8.
5. Pointner, R., et al., *Gastric remnant cancer has a better prognosis than primary gastric cancer*. *Arch Surg*, 1994. **129**(6): p. 615-9.
6. Sharon E. Straus , et al., *Evidence Based Medicine:How to Practice and Teach EBM*. 2005, Churchill Livingstone; 3 edition
7. *Management of oesophageal and gastric cancer. A national clinical guideline*. 2006 [cited; Available from: <http://www.guideline.gov/>.]
8. Faigel, D.O., et al., *Tissue sampling and analysis*. *Gastrointest Endosc*, 2003. **57**(7): p. 811-6.
9. Hamashima, C., et al., *The Japanese guidelines for gastric cancer screening*. *Jpn J Clin Oncol*, 2008. **38**(4): p. 259-67.
10. Riecken, B., et al., *No impact of repeated endoscopic screens on gastric cancer mortality in a prospectively followed Chinese population at high risk*. *Prev Med*, 2002. **34**(1): p. 22-8.
11. Voutilainen, M.E. and Juhola, M.T., *Evaluation of the diagnostic accuracy of gastroscopy to detect gastric tumours: clinicopathological features and prognosis of patients with gastric cancer missed on endoscopy*. *Eur J Gastroenterol Hepatol*, 2005. **17**(12): p. 1345-9.
12. Schulz, K.F. and Grimes, D.A., *Handbook of Essential Concepts in Clinical Research*. 2006: Elsevier Science Health Science div

13. Yalamarathi, S., et al., *Missed diagnoses in patients with upper gastrointestinal cancers*. *Endoscopy*, 2004. **36**(10): p. 874-9.
14. Hosokawa, O., et al., *Diagnosis of gastric cancer up to three years after negative upper gastrointestinal endoscopy*. *Endoscopy*, 1998. **30**(8): p. 669-74.
15. Grove, A.S., *Only The Paranoid Survive* 1996: Currency Doubleday Publishing
16. Amin, A., et al., *Gastric adenocarcinoma missed at endoscopy*. *J R Coll Surg Edinb*, 2002. **47**(5): p. 681-4.
17. 李威傑, et al., *Postresection Prognosis of Gastric Cancer Patients: Comparison of Staging Systems* 慈濟醫學雜誌 1994. **6**(1): p. 7-13.
18. Tsukuma, H., et al., *Natural history of early gastric cancer: a non-concurrent, long term, follow up study*. *Gut*, 2000. **47**(5): p. 618-21.