



(19)中華民國智慧財產局

(12)發明說明書公開本

(11)公開編號：TW 202238515 A

(43)公開日：中華民國 111 (2022) 年 10 月 01 日

(21)申請案號：110144557

(22)申請日：中華民國 110 (2021) 年 11 月 30 日

(51)Int. Cl. : **G06T7/00 (2017.01)** **A61B5/00 (2006.01)**

(30)優先權：2020/12/01 美國 63/120,102

(71)申請人：美商壯生和壯生企業創新公司 (美國) JOHNSON & JOHNSON ENTERPRISE INNOVATION INC. (US)

美國

(72)發明人：華席科 喬治 WASHKO, GEORGE R., JR. (US) ; 艾斯堤帕 勞爾 ESTEPAR, RAUL SAN JOSE (US) ; 金希 查爾斯 KINSEY, CHARLES MATTHEW (US) ; 史帝文森 克里斯多福 STEVENSON, CHRISTOPHER SCOTT (US)

(74)代理人：何愛文；王仁君

申請實體審查：無 申請專利範圍項數：118 項 圖式數：9 共 103 頁

(54)名稱

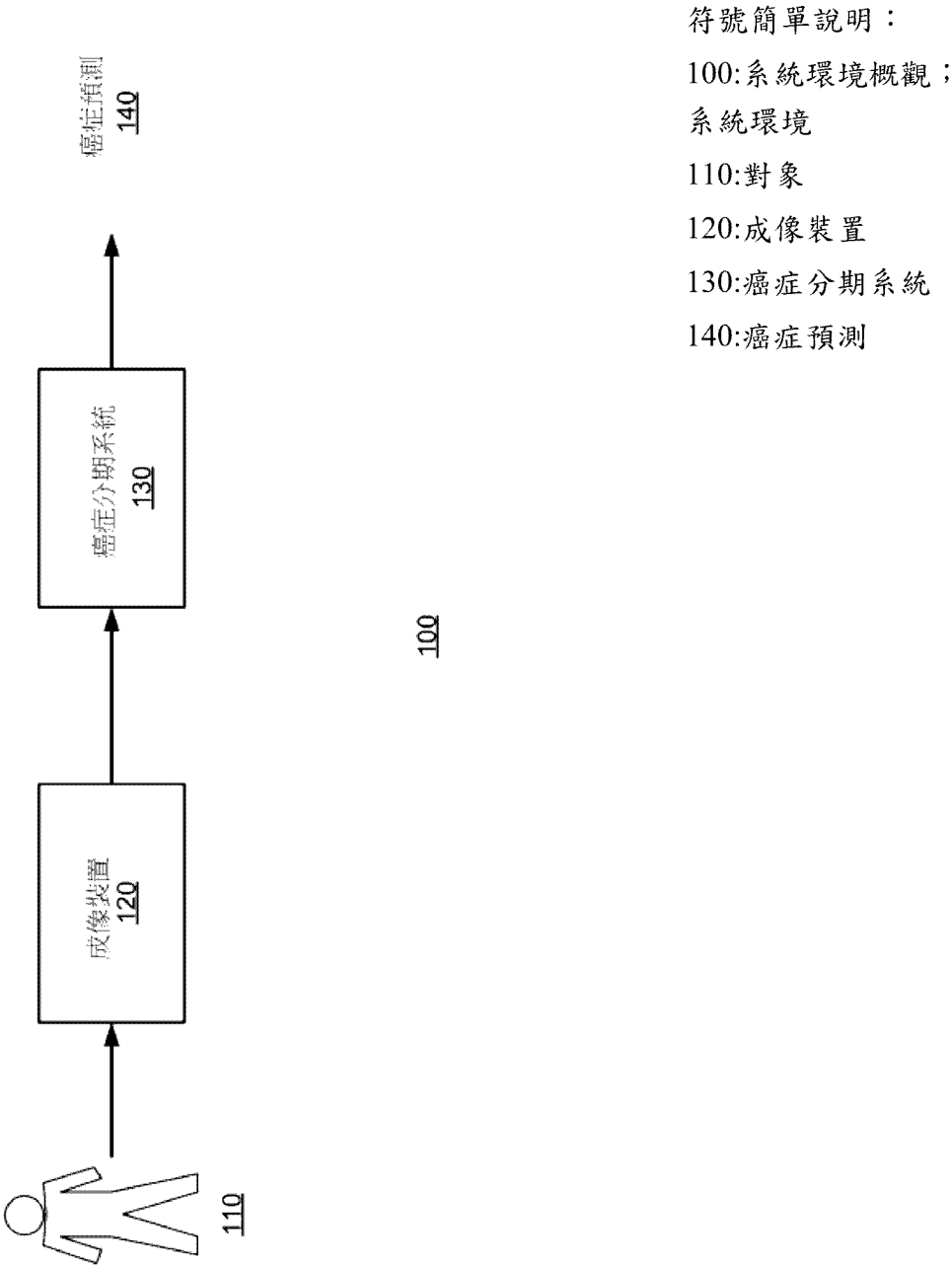
機器學習預測侵犯淋巴結的癌症

(57)摘要

本文揭示用於判定轉移性癌症之一對象水準風險的方法，其涉及訓練及/或部署模型以判定：1) 個體淋巴結侵犯之淋巴結水準風險及/或 2) 淋巴結侵犯之對象水準風險。因此，該等方法可識別出具有淋巴結疾病之高或低風險的患者，並可選地例如經由治療來實現癌症患者之導引介入。

Disclosed herein are methods for determining a subject level risk of metastatic cancer involving the training and/or deployment of models to determine 1) a lymph node level risk of individual lymph node involvement and/or 2) a subject level risk of lymph node involvement. Thus, the methods can identify patients who are high or low risk for having nodal disease and optionally enable the guided intervention of cancer patients, for example, via treatment.

指定代表圖：



【圖1A】

【發明摘要】

【中文發明名稱】機器學習預測侵犯淋巴結的癌症

【英文發明名稱】MACHINE LEARNING TO PREDICT CANCER

INVOLVEMENT OF LYMPH NODES

【中文】

本文揭示用於判定轉移性癌症之一對象水準風險的方法，其涉及訓練及/或部署模型以判定：1)個體淋巴結侵犯之淋巴結水準風險及/或2)淋巴結侵犯之對象水準風險。因此，該等方法可識別出具有淋巴結疾病之高或低風險的患者，並可選地例如經由治療來實現癌症患者之導引介入。

【英文】

Disclosed herein are methods for determining a subject level risk of metastatic cancer involving the training and/or deployment of models to determine 1) a lymph node level risk of individual lymph node involvement and/or 2) a subject level risk of lymph node involvement. Thus, the methods can identify patients who are high or low risk for having nodal disease and optionally enable the guided intervention of cancer patients, for example, via treatment.

【指定代表圖】圖1A

【代表圖之符號簡單說明】

100:系統環境概觀；系統環境

110:對象

120:成像裝置

130:癌症分期系統

140:癌症預測

【特徵化學式】

無

【發明說明書】

【中文發明名稱】機器學習預測侵犯淋巴結的癌症

【英文發明名稱】MACHINE LEARNING TO PREDICT CANCER

INVOLVEMENT OF LYMPH NODES

【技術領域】

相關申請案之交互參照

【0001】 本申請案主張2020年12月1日申請之美國臨時專利申請案第63/120,102號之優先權及權益，該案之整體揭露出於所有目的以全文以引用方式併入本文中。

【先前技術】

【0002】 胸部的電腦斷層攝影(computed tomographic, CT)成像已成為用於評估處於肺癌風險者的臨床標準。這些努力著重於偵測肺結節，該等肺結節可表示在其最早及最容易治癒期的疾病。此一肺結節的偵測觸發更廣泛的臨床評估，以判定其是否為癌性的，並且若是，則判定其是否已擴散至胸部或甚至身體的遠端區域中的其他結構。此係用以判定肺癌之期別，其繼而決定治療介入之計劃。

【0003】 肺癌分期的一個態樣係判定其在胸腔淋巴結中存在或不存在（淋巴結狀態）。目前標準照護涉及分期方案，其中首先將結節判定為良性或惡性，接著完成額外檢測，諸如針對淋巴結的正子發射斷層攝影(PET)掃描及手術、內視鏡、或基於支氣管鏡之採樣。組織採樣程序增加住院診斷工作(work up)的風險、費用、及時間。需要新的非侵入性方法以減輕這些風險並縮短介於

診斷與治療之間的時間。如本文進一步詳細說明者，應用於用於肺癌篩檢之初始胸腔CT掃描的機器學習及基於深度學習之技術可解決這些需求。

【0004】 肺癌偵測、分期及治療規劃係可費時數週或數月的多步驟程序。除了此一臨床住院診斷工作的風險及費用外，此程序所需的額外時間亦會導致顯著患者焦慮及提高間隔轉移(interval metastasis)。事實上，較長的診斷至治療時間與更糟的整體存活期相關聯。減輕這些疑慮的例示性方法是允許幾乎同時的癌症診斷、分期、及（在支氣管鏡療法的情況下）治療。

【發明內容】

【0005】 本文中所揭露之本發明的實施例係基於用於篩檢高風險族群中肺癌的胸腔CT掃描之分析。在一些態樣中，識別處於癌症已轉移至胸部中之至少一個淋巴結(LN)之高風險且因此需要額外檢測（諸如PET掃描及淋巴結活體組織切片）的患者。此允許臨床提供者為最需要檢測者排定優先順序。相反地，處於具有淋巴結轉移之低風險的疑似患有肺癌的患者可直接進行至立即治療，諸如手術切除或合併的支氣管鏡診斷及治療。此將消除疾病診斷及療法之間的時間。具體而言，本文中所揭露之本發明實施例實現識別出具有淋巴結疾病之高或低風險的患者，從而允許臨床醫師對高風險個體排定額外檢測的優先順序或請具有低轉移風險之肺癌患者直接進行確認性支氣管鏡活體組織切片，緊接著進行支氣管鏡遞送療法。

【0006】 本文揭示一種判定一對象的轉移性癌症之一對象水準風險之方法，該方法包含：獲得從該對象擷取的一或多個影像，該一或多個影像包含該對象之複數個淋巴結；及藉由應用一風險模型至所獲得之該一或多個影像之所

提取特徵來預測轉移性癌症之該對象水準風險，該風險模型經訓練以預測在影像中淋巴結的轉移性癌症之淋巴結(LN)水準風險，其中針對一參考個體之一組淋巴結，至少使用導出自轉移性癌症之經推算LN水準風險的標籤來訓練該風險模型，該組淋巴結之該經推算LN水準風險係使用辨別訓練影像之癌性與非癌性淋巴結的至少一推算模型來判定。在各種實施例中，預測轉移性癌症之該對象水準風險進一步包含：基於由該風險模型所預測的轉移性癌症之該複數個淋巴結之LN水準風險來選擇該複數個淋巴結中之一或多者；及使用針對該一或多個淋巴結所預測的轉移性癌症之LN水準風險來判定轉移性癌症之該對象水準風險。

【0007】 在各種實施例中，選擇該複數個淋巴結中之一或多者包含識別具有最高機率之LN水準風險的該淋巴結，且其中判定轉移性癌症之該對象水準風險包含指派該最高機率之LN水準風險作為轉移性癌症之該對象水準風險。在各種實施例中，該風險模型預測轉移性癌症之LN水準風險的機率範圍大於由該推算模型所預測的該等經推算LN水準風險。在各種實施例中，一參考個體之該組淋巴結係藉由下列來選擇：判定一中位數風險值；及比較該參考個體之該等淋巴結的經推算LN水準風險與該中位數風險值。在各種實施例中，一參考個體之該組淋巴結進一步藉由以下來選擇：在該組淋巴結中包括具有大於該中位數風險值之LN水準風險的一或多個淋巴結。在各種實施例中，從該組淋巴結中排除在該組淋巴結中具有小於該中位數風險值之LN水準風險的一或多個淋巴結。

【0008】 在各種實施例中，該中位數風險值係該參考個體的一中位數對象內LN水準風險。在各種實施例中，該一或多個影像包含一電腦斷層攝影(CT)影像。在各種實施例中，該一或多個影像獲自一胸腔CT掃描。在各種實施例

中，本文所揭露之方法進一步包含根據轉移性癌症之該預測對象水準風險來判定該癌症之一期別。在各種實施例中，本文所揭示之方法進一步包含根據轉移性癌症之該預測對象水準風險來選擇該對象之一診斷或治療。在各種實施例中，選擇該對象之一診斷或治療包含：當轉移性癌症之該預測對象水準風險低於一臨限風險值時，選擇手術腫瘤切除或合併的支氣管鏡或內視鏡診斷及治療。在各種實施例中，選擇該對象之一診斷或治療包含：當轉移性癌症之該預測對象水準風險高於一臨限風險值時，選擇執行一PET或PET-CT掃描或執行淋巴結活體組織切片中之一者或兩者的額外診斷檢測。

【0009】 在各種實施例中，該推算模型係使用獲自胸腔CT掃描之訓練影像來訓練。在各種實施例中，獲自胸腔CT掃描之該等訓練影像包含一或多個淋巴結。在各種實施例中，該推算模型係使用國家肺部篩選試驗(National Lung Screening Trial, NLST)之訓練影像來訓練。在各種實施例中，該推算模型係使用一自訂資料集之訓練影像來訓練。

【0010】 在各種實施例中，該推算模型係藉由從該等訓練影像之至少一者產生二或更多個影像圖譜(radiomic)板來訓練。在各種實施例中，該二或更多個影像圖譜板包含訓練影像之一高斯拉普拉斯變換(Laplacian of Gaussians transformation)及訓練影像之一小波變換。在各種實施例中，該推算模型進一步藉由下列訓練：從該二或更多個影像圖譜板之各者提取特徵；及至少使用該等所提取特徵來訓練該推算模型。在各種實施例中，該推算模型進一步藉由下列訓練：在該等影像圖譜板中界定一關注區域(ROI)，所界定之該關注區域包含一淋巴結；從該等影像圖譜板中之該ROI提取特徵；及至少使用該等所提取特徵來訓練該推算模型。在各種實施例中，該等所提取特徵包括一階特徵、形狀特

徵、灰階共生矩陣1 (GLCM)特徵、灰階級長矩陣(GLRLM)特徵、灰階級帶矩陣(GLSZM)特徵、及鄰域灰階差分矩陣(NGTDM)特徵中之一或多個特徵類別。

【0011】 在各種實施例中，該推算模型係進一步使用指示訓練影像中轉移的存在或不存在的參考基準真相值來訓練。在各種實施例中，該等參考基準真相值係導出自指示該癌症之該期別的淋巴結轉移期數共變異數值，其中「0」或「1」之一癌症期別指示轉移的不存在，且其中「2」或「3」之一癌症期別指示轉移的存在。在各種實施例中，該等參考基準真相值指示個體淋巴結為癌性或非癌性的。在各種實施例中，指示個體淋巴結是否是癌與非癌的該等參考基準真相值係基於該腫瘤或該等淋巴結之特性而從該等淋巴結轉移期數共變異數值反算。

【0012】 在各種實施例中，該推算模型係一隨機森林分類器或一LASSO分類器。在各種實施例中，該風險模型係一隨機森林分類器或一LASSO分類器。在各種實施例中，該風險模型係使用監督學習技術來訓練。在各種實施例中，複數個淋巴結包含1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、或21個淋巴結。在各種實施例中，該複數個淋巴結包含8至20個之間的淋巴結、10至19個之間的淋巴結、12至18個之間的淋巴結、或14與16個之間的淋巴結。在各種實施例中，該轉移性癌症係轉移性肺癌。在各種實施例中，該風險模型展現至少AUC = 78%、靈敏度= 82%、或特異度= 60%之效能。

【0013】 本文中額外揭示一種訓練如請求項1至34中任一項之風險模型之方法，該方法包含：應用一推算模型以產生用於在訓練影像中之複數個淋巴結的經推算LN水準風險；選擇一組淋巴結，其中該組中之該等淋巴結具有大於一

中位數風險值的LN水準風險；使用在該組淋巴結中的該等淋巴結之該等LN水準風險作為用於訓練該風險模型的參考基準真相，該風險模型能夠預測LN水準風險的機率範圍大於由該推算模型所預測的該等經推算LN水準風險。在各種實施例中，該風險模型進一步藉由下列訓練：從一或多個影像圖譜板提取特徵；及至少使用該等所提取特徵來訓練該風險模型。在各種實施例中，從一或多個影像圖譜板提取特徵包含：在該一或多個影像圖譜板中界定一關注區域(ROI)，所界定之該關注區域包含一淋巴結；從該一或多個影像圖譜板中之該ROI提取特徵。在各種實施例中，該等所提取特徵包括一階特徵、形狀特徵、灰階共生矩陣1 (GLCM)特徵、灰階級長矩陣(GLRLM)特徵、灰階級帶矩陣(GLSZM)特徵、及鄰域灰階差分矩陣(NGTDM)特徵中之一或多個特徵類別。

【0014】 本文中額外揭示一種用於執行上述方法之系統。在各種實施例中，一種用於判定一對象的轉移性癌症之一對象水準風險之系統包含：一成像裝置，其經組態以擷取該對象的一或多個影像；及一運算裝置，其經組態以執行下列步驟：獲得從該對象擷取的一或多個影像，該一或多個影像包含該對象之複數個淋巴結；及藉由應用一風險模型至所獲得之該一或多個影像之所提取特徵來預測轉移性癌症之該對象水準風險，該風險模型經訓練以預測在影像中淋巴結的轉移性癌症之淋巴結(LN)水準風險，其中針對一參考個體之一組淋巴結，至少使用導出自轉移性癌症之經推算LN水準風險的標籤來訓練該風險模型，該組淋巴結之該經推算LN水準風險係使用辨別訓練影像之癌性與非癌性淋巴結的至少一推算模型來判定。

【0015】 本文中額外揭示一種用於執行上述方法之非暫態電腦可讀取媒體。在各種實施例中，非暫態電腦可讀取媒體包含當由一處理器執行時造成

該處理器進行下列之指令：獲得從該對象擷取的一或多個影像，該一或多個影像包含該對象之複數個淋巴結；及藉由應用一風險模型至所獲得之該一或多個影像之所提取特徵來預測轉移性癌症之該對象水準風險，該風險模型經訓練以預測在影像中淋巴結的轉移性癌症之淋巴結(LN)水準風險，其中針對一參考個體之一組淋巴結，至少使用導出自轉移性癌症之經推算LN水準風險的標籤來訓練該風險模型，該組淋巴結之該經推算LN水準風險係使用辨別訓練影像之癌性與非癌性淋巴結的至少一推算模型來判定。

【圖式簡單說明】

【0016】 本發明之這些及其他特徵、態樣、及優點將隨著以下說明與隨附圖式更佳地理解。

〔圖1〕繪示用於產生用於應用程式開發的經訓練機器學習模型之系統的簡化例示性圖。

〔圖1A〕描繪根據一實施例之藉由判定一對象的轉移性癌症之一對象水準風險進行癌症分期的系統環境概觀。

〔圖1B〕描繪根據一實施例之癌症分期系統的方塊圖。

〔圖2A〕描繪根據一實施例之識別淋巴結站的實例淋巴結圖。

〔圖2B〕係根據一實施例之一實例淋巴結標註。

〔圖3A〕描繪根據一實施例之推算模型及風險模型之一實例實施方案。

〔圖3B〕描繪根據一實施例之進一步細節的風險模型之實施方案。

〔圖4〕繪示一實例電腦，其用於實施圖1A、圖1B、圖2A、圖2B、圖3A、及圖3B所示之實體。

〔圖5A〕至〔圖5R〕描繪不同淋巴結影像圖譜特徵之實例相關性熱圖。

〔圖6〕描繪推算模型之淋巴結水準效能。

〔圖7A〕及〔圖7B〕描繪藉由LN水準分類器所預測之最高機率評分之對象水準效能。

〔圖8〕描繪在兩步驟式拔靴法(bootstrap approach)中的推算模型之淋巴結水準效能。

〔圖9A〕及〔圖9B〕描繪使用拔靴法之對象水準效能。

【0017】 應注意，在切實可行的範圍內，類似或相似的參考符號可用於圖式中，且可指示類似或相似的功能。例如，一參考符號之後的字母（諸如「個體315A」）指示該文字具體地指具有特定參考符號的元件。本文中的參考符號無接在後面的字母（諸如「個體315」）指在圖式中有參考符號的任何或所有元件（例如，本文中的「個體315」指圖式中的參考符號「個體315A」、「個體315B」、「個體315C」、及「個體315D」）。

【實施方式】

I. 定義

【0018】 請求項及說明書中使用的用語係定義如下所述，除非另有指明。

【0019】 用語「對象(subject)」涵蓋細胞、組織、或生物體、人類或非人類，無論是體內、離體、或體外、男性或女性。

【0020】 用語「哺乳動物(mammal)」涵蓋人類及非人類，並且包括但不限於人類、非人類靈長動物、犬、貓、鼠、牛、馬、及豬。

【0021】 用語「樣本(sample)」或「測試樣本(test sample)」可包括藉由包括靜脈穿刺、排泄、射精、按摩、活體組織切片、針頭抽吸物、灌洗樣本、刮取、手術切口、或介入之手段或所屬技術領域中已知之其他手段而自一對象取得之單一細胞或多個細胞或細胞之片段或一等分試樣之體液（諸如血液樣本）。一等分試樣的體液實例包括羊水、前房液、膽汁、淋巴液、乳汁、組織間隙液、血液、血液血漿、耳垢（耳蠟）、考伯液（預射精液）、乳糜、食糜、女性潮射、月經、黏液、唾液、尿液、嘔吐物、淚液、陰道潤滑液、汗液、血清、精液、皮脂、膿汁、胸膜液、腦脊髓液、滑液、細胞內液、及玻璃體液。

【0022】 用語「參考個體(reference individual)」係指具有已知癌症分期的個體。例如，參考個體可包括健康個體（例如，對照個體）。作為另一實例，參考個體可包括先前經診斷有癌症及經分期之個體。作為另一實例，參考個體包括先前經診斷有已轉移至淋巴結之癌症（如第IV期癌症）的個體。

【0023】 用語「獲得一或多個影像(obtaining one or more images)」涵蓋獲得從對象擷取的一或多個影像，或獲得從對象獲得之樣本擷取之一或多個影像。獲得一或多個影像可涵蓋執行自對象或自獲自對象之樣本擷取一或多個影像之步驟。該用語亦可涵蓋例如自己實施自對象或自獲自對象之樣本擷取一或多個影像之步驟之第三方接收一或多個影像。一或多個影像可藉由所屬技術領域中具有通常知識者經由多種已知方式獲得，包括儲存於儲存記憶體上。

【0024】 用語「淋巴結侵犯(lymph node involvement)」、「淋巴結轉移(lymph node metastasis)」、或「淋巴結癌(lymph node cancer)」係指一對象之一或多個淋巴結中的轉移癌症之存在或不存在。

【0025】 用語「淋巴結轉移風險(lymph node metastatic risk)」、「淋巴結水準風險(lymph node level risk)」、及「LN水準風險(LN-level risk)」可互換使用，且通常指在特定淋巴結中轉移性癌症的風險。

【0026】 用語「轉移性癌症之對象水準風險(subject level risk of metastatic cancer)」或「轉移性癌症之患者水準風險(patient level risk of metastatic cancer)」可互換使用，且係指一對象中的癌症已轉移的可能性。在各種實施例中，對象中之癌症係在對象之一或多個淋巴結中的已轉移的肺癌。

【0027】 用語「癌症期別(cancer stage或stage of cancer)」係指對象中之癌症期別。癌症期別可指包括第0至4期的TNM分期系統。此外，癌症期別可指淋巴結癌0至3期。淋巴結癌0期指示無淋巴結侵犯，淋巴結癌1期指示肺門淋巴結侵犯，而淋巴結2或3期指示縱隔腔侵犯。

【0028】 應注意的是，除非上下文清楚地指明，否則本說明書中所用的單數形式「一(a/an)」及「該(the)」包括複數指代物。

II. 系統環境概觀

【0029】 圖1A描繪根據一實施例之藉由判定一對象的轉移性癌症之一對象水準風險進行癌症分期的系統環境概觀100。系統環境100提供脈絡以介紹對象110、成像裝置120、及用於判定癌症預測140之癌症分期系統130。在各種實施例中，藉由癌症分期系統130判定的癌症預測140用作為對象110的癌症分

期。在各種實施例中，藉由癌症分期系統130判定的癌症預測140係對象110的預測風險水準，並用作為用於判定對象110之癌症分期的基礎。

【0030】 在某些實施例中，該對象係健康的。因此，本文中所述之用於癌症分期之方法對於在健康對象中早期偵測癌症可為有益。在各種實施例中，對象先前經診斷為癌症。在各種實施例中，該對象係健康的，但疑似有某種形式的癌症（例如，已展現出與癌症相關聯之症狀的對象）。在具體實施例中，對象的癌症類型係肺癌。因此，本文中所述之用於癌症分期之方法對於識別及區分處於包括淋巴結侵犯的高轉移性癌症風險之對象與處於具有淋巴結轉移之低風險的疑似患有肺癌的患者可係有益的。處於高轉移性癌症風險的對象可接受額外檢測，諸如PET掃描及淋巴結活體組織切片，而處於具有淋巴結轉移之低風險的對象可接受立即治療，諸如肺癌的手術切除或合併的支氣管鏡診斷及治療。

【0031】 在各種實施例中，成像裝置120擷取對象110之影像以供進一步分析。在各種實施例中，成像裝置120自獲自對象110之測試樣本擷取影像。該影像及/或該樣本可由個體或由第三方（例如，健康照護提供者或專業醫療人員）獲得。醫療專業人員之實例包括醫師、緊急醫療技術人員、護理師、第一線應變人員、心理師、抽血師、醫學物理人員、護理人員、外科醫師、牙醫師、及將為所屬技術領域中具有通常知識者已知之任何其他顯而易見的醫療專業人員。

【0032】 在一些實施例中，成像裝置120擷取對象110之解剖學位置的影像。對象之實例解剖學位置可包括肺、胸腔、腎、肝臟、胰臟、腦、胃、小腸、髖部、膝蓋、腿、手臂、及臉。在各種實施例中，成像裝置120擷取對象

110之胸腔的影像。在各種實施例中，成像裝置120擷取對象110之一或多個淋巴結的影像。在各種實施例中，成像裝置120擷取對象110之胸腔（包括一或多個淋巴結）的影像。

【0033】 在各種實施例中，成像裝置120係電腦斷層攝影(CT)掃描器、磁共振造影(MRI)掃描器、正子發射斷層攝影(PET)掃描器、X射線掃描器、或超音波成像裝置中之一者。在具體實施例中，成像裝置120係CT掃描器，其擷取對象110之一或多個影像。在具體實施例中，成像裝置120係CT掃描器，其擷取對象110之胸腔（包括一或多個淋巴結）的一或多個CT影像。

【0034】 一般而言，癌症分期系統130分析一或多個影像（例如，由成像裝置120擷取的影像），且產生癌症預測140。在各種實施例中，癌症預測140可係對象110中的預測癌症期別。在各種實施例中，癌症預測140可係對象110中之癌淋巴結侵犯的預測可能性。在各種實施例中，癌症預測140可包括建議的臨床方法（例如，額外檢測，諸如PET掃描及淋巴結活體組織切片）或療法（例如，手術切除及/或藥物療法）。癌症分期系統130可包括體現為電腦系統400（如下文關於圖4所討論）的一或多個電腦。因此，在各種實施例中，參照癌症分期系統130所描述的步驟係於電腦上(*in silico*)執行。

【0035】 在各種實施例中，成像裝置120及癌症分期系統130係由不同方所採用。例如，第一方操作成像裝置120以自對象110擷取一或多個影像，且接著將所擷取之該一或多個影像提供給第二方，該第二方實施癌症分期系統130以判定癌症預測140。

【0036】 現參照描繪方塊圖之圖1B，該方塊圖繪示根據一實施例之癌症分期系統130的電腦邏輯組件。此處，癌症分期系統130包括淋巴結特徵模組

145、推算模組150、風險訓練模組155、風險部署模組160、癌症預測模組165、及訓練資料儲存170。

【0037】 一般而言，淋巴結特徵模組145從導出自對象或訓練對象之影像提取特徵。推算模組150訓練及部署一推算模型。該推算模型分析所提取特徵，並判定影像中之各淋巴結的淋巴結水準風險。風險訓練模組155使用訓練資料來訓練風險模型，該訓練資料包括由該推算模型所判定的至少該淋巴結水準風險。風險部署模組160實施風險模型以分析從一對象（例如，圖1A中之對象110）獲得之影像所提取之特徵，以判定癌症預測，諸如預測對象110之對象水準轉移風險或癌症期別。換言之，癌症分期系統130訓練並實施至少兩個模型（例如，一個推算模型及一個風險模型）。因此，透過涉及該至少兩個模型的此多步驟實施方案，癌症分期系統130產生一對象的癌症預測，諸如對象水準轉移風險或癌症期別之預測。

【0038】 推算模型與風險模型兩者的實施方案導致改善的癌症預測。具體而言，推算模型從來自訓練資料之對象水準淋巴結轉移期數(subject-level N-stage)中學習，並被應用至淋巴結之經提取特徵，以界定哪些淋巴結有可能促成陽性淋巴結轉移期數。接著，風險模型從經推算資料學習更精進的淋巴結轉移風險機率。此處，風險模型預測轉移性癌症之LN水準風險的機率範圍大於由推算模型所預測的經推算LN水準風險。這實現了轉移性癌症之對象水準風險的更準確判定。

【0039】 癌症分期系統130的組件在以下係參照二期來描述：1)訓練期及2)部署期。更具體而言，訓練期係指基於訓練資料（諸如從參考個體（例如，具有已知癌症分期之個體）擷取的訓練影像）來對一或多個模型進行的建置及

訓練。因此，模型係使用訓練資料來訓練，使得在部署期期間，模型的實施實現了一對象（例如圖1A中之對象110）之癌症分期的預測。在各種實施例中，推算模組150亦可經實施以訓練及部署一模型，該模型在下文稱為一推算模型。因此，推算模組150在訓練期期間訓練該推算模型及在部署期期間部署該推算模型。

【0040】 在一些實施例中，在訓練期及部署期中之一者期間應用癌症分期系統130的組件。例如，在訓練期期間應用風險訓練模組155及訓練資料儲存170以訓練風險模型。此外，在部署期期間應用風險部署模組160。在各種實施例中，取決於在訓練期或部署期期間應用癌症分期系統130的組件，該等組件可由不同方來執行。在此類情景中，預測模型的訓練及部署係由不同方執行。例如，在訓練期期間所應用之風險訓練模組155及訓練資料儲存170可由第一方採用（例如，以訓練風險模型），且在部署期期間所應用之風險部署模組160可由第二方執行（例如，以部署風險模型）。同樣地，可由一第一方應用推算模組150以訓練推算模型，且可由一第二方應用推算模組150以部署推算模型。

III. 用於癌症分期之方法

【0041】 本文中所述之實施例包括用於藉由下列方式進行癌症分期之方法：分析從對象擷取的一或多個影像；及應用一或多個模型以判定涉及淋巴結之轉移性癌症風險的對象水準。此類方法可藉由在圖1B中描述的癌症分期系統130來執行。

【0042】 參照圖1B中的癌症分期系統130之模組，淋巴結特徵模組145分析一或多個影像並從該一或多個影像提取特徵。在各種實施例中，淋巴結特徵

模組145從獲自參考個體的訓練影像提取特徵。從訓練影像提取的這些特徵可用於訓練一或多個模型（諸如推算模型），如下文進一步詳細描述。在各種實施例中，淋巴結特徵模組145從獲自對象影像提取特徵（在圖1A中之對象110）。從獲自對象之影像提取的這些特徵可藉由部署模型（諸如風險模型）來分析，如下文進一步詳細描述。

【0043】 在各種實施例中，一或多個影像或訓練影像包括在影像中之一或多個淋巴結。例如，影像或訓練影像可係包括一或多個淋巴結之胸腔的影像。在各種實施例中，影像包括至少2個淋巴結。在各種實施例中，影像包括1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、或21個淋巴結。在各種實施例中，影像包括1至10個之間的淋巴結、1至15個之間的淋巴結、2至10個之間的淋巴結、2至15個之間的淋巴結、2至20個之間的淋巴結、4至10個之間的淋巴結、4至15個之間的淋巴結、4至20個之間的淋巴結、5至10個之間的淋巴結、5至15個之間的淋巴結、5至20個之間的淋巴結、8至15個之間的淋巴結、8至20個之間的淋巴結、10至15個之間的淋巴結、10至20個之間的淋巴結、12至15個之間的淋巴結、12至20個之間的淋巴結、或15至20個之間的淋巴結。在各種實施例中，從單一對象或參考個體擷取多個影像。多個影像可擷取對象或參考個體之不同淋巴結。在各種實施例中，多個影像包括至少2個淋巴結。在各種實施例中，多個影像包括2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、或21個淋巴結。在各種實施例中，多個影像包括8至20個之間的淋巴結。在各種實施例中，多個影像包括10至19個之間的淋巴結、12至18個之間的淋巴結、或14與16之間的淋巴結。

【0044】 圖2A描繪根據一實施例之識別淋巴結站的實例淋巴結圖。更具體而言，圖2A係修改自美國胸腔學會映射方案(mapping scheme)的局部淋巴結分類之描繪。此處，淋巴結站經標註以描繪各淋巴結之位置及變化。具體而言，該等標註係如下：(1)鎖骨上結、(2R/2L)上氣管旁結、(3A)血管前結、(3P)脊椎前方結、(4R/4L)下管旁結、(5)主動脈下結、(6)側主動脈結、(7)氣管隆凸下結、(8)食道旁結、(9)肺韌帶結、(10)肺門結、(11)葉間結、(12)肺葉結、(13)肺小節結、及(14)肺次節葉結。

【0045】 淋巴結特徵模組145標註影像中的一或多個淋巴結。在各種實施例中，淋巴結特徵模組145藉由界定圍繞淋巴結的區域來標註一或多個淋巴結。在各種實施例中，圍繞淋巴結的經界定區域係球形區域。在各種實施例中，該球形區域具有 X mm之半徑。在各種實施例中， X 係5 mm。在各種實施例中， X 係7.5 mm。在一些實施例中， X 係0.5 mm、1 mm、1.5 mm、2 mm、2.5 mm、3 mm、3.5 mm、4 mm、4.5 mm、5.5 mm、6 mm、6.5 mm、7 mm、8 mm、8.5 mm、9 mm、9.5 mm、10 mm、10.5 mm、11 mm、11.5 mm、12 mm、12.5 mm、13 mm、13.5 mm、14 mm、14.5 mm、或15 mm。在各種實施例中， X 係1與15 mm之間。在各種實施例中， X 係2與12 mm之間。在各種實施例中， X 係3與10 mm之間。在各種實施例中， X 係4與8 mm之間。在各種實施例中， X 係5與7.5 mm之間。在各種實施例中， X 係4與6 mm之間。在各種實施例中， X 係6與8 mm之間。在各種實施例中， X 係7與8 mm之間。在各種實施例中， X 係10與15 mm之間。

【0046】 在各種實施例中，淋巴結特徵模組145藉由界定圍繞淋巴結的二或更多個球形區域來標註一或多個淋巴結。例如，淋巴結特徵模組145藉由界

第 16 頁，共 52 頁(發明說明書)

定圍繞淋巴結的5 mm之一第一球形區域及圍繞淋巴結的7.5 mm之一第二球形區域來標註一或多個淋巴結。圖2B係根據一實施例之一實例淋巴結標註。具體而言，圖2B描繪圍繞兩個2R上氣管旁結的兩個經界定球形區域。

【0047】 在各種實施例中，淋巴結特徵模組145從圍繞淋巴結的各經界定區域提取特徵。在各種實施例中，淋巴結特徵模組145執行原始輸入的變換以產生一經變換影像，然後從該經變換影像中提取額外特徵。在一實施例中，淋巴結特徵模組145執行原始影像之小波變換。在一實施例中，淋巴結特徵模組145執行原始影像之高斯拉普拉斯(LoG)。因此，淋巴結特徵模組145可從經變換影像中圍繞淋巴結的經界定區域提取特徵。在各種實施例中，淋巴結特徵模組145從原始影像及經變換影像中圍繞淋巴結的經界定區域提取特徵。

【0048】 在具體實施例中，淋巴結特徵模組145執行原始影像的兩個變換（例如，原始影像之小波變換及原始影像之高斯拉普拉斯(LoG)兩者）。因此，淋巴結特徵模組145可從第一經變換影像及第二經變換兩者中圍繞淋巴結的經界定區域提取特徵。在各種實施例中，淋巴結特徵模組145從原始影像、第一經變換影像、及第二經變換中圍繞淋巴結的經界定區域提取特徵。

【0049】 來自原始影像、第一變換影像、及/或第二變換影像之實例特徵類別可包括一階特徵、形狀特徵、灰階共生矩陣1 (GLCM)特徵、灰階級長矩陣(GLRLM)特徵、灰階級帶矩陣(GLSZM)特徵、及鄰域灰階差分矩陣(NGTDM)特徵中之一或多個特徵類別。一階特徵係指強度值的一階統計量之特徵。GLCM特徵在影像上以給定偏移從共生灰階值導出直方圖。GLRLM特徵涉及灰階級長矩陣(run length matrix)的統計。GLSDRUE特徵涉及灰階級帶矩陣(size zone matrix)的統計。NGTDM特徵涉及鄰域灰階差分矩陣(gray tone difference

matrix)的統計。不同特徵類別中之特徵的實例包括：相關特徵、LMC2、MCC、差熵、聯合熵、和熵、群聚突出、平均差、聯合平均、平均加總、自動相關度、變異性差別、對比、群聚傾向(cluster tendency)、差值平方和、LDMN、LDN、群聚陰暗度(cluster shade)、LMC1、逆方差、LD、LDM、聯合能量、及最大機率。

【0050】 在各種實施例中，淋巴結特徵模組145從影像提取至少50個特徵。在各種實施例中，淋巴結特徵模組145從影像提取至少100個特徵、至少150個特徵、至少200個特徵、至少250個特徵、至少300個特徵、至少350個特徵、至少400個特徵、至少450個特徵、至少500個特徵、至少550個特徵、至少600個特徵、至少650個特徵、至少700個特徵、至少750個特徵、至少800個特徵、至少850個特徵、至少900個特徵、至少950個特徵、至少1000個特徵、至少1100個特徵、至少1200個特徵、至少1300個特徵、至少1400個特徵、至少1500個特徵、至少1600個特徵、至少1700個特徵、至少1800個特徵、至少1900個特徵、至少2000個特徵、至少3000個特徵、至少4000個特徵、至少5000個特徵、至少6000個特徵、至少7000個特徵、至少8000個特徵、至少9000個特徵、或至少10,000個特徵。在各種實施例中，淋巴結特徵模組145提取介於10個特徵與10000個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於100個特徵與1000個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於300個特徵與900個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於500個特徵與1000個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於500個特徵與900個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於700個特徵與1000個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於700個特

徵與900個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於800個特徵與1000個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於800個特徵與900個特徵之間。

【0051】 在各種實施例中，淋巴結特徵模組145提供經提取特徵至推算模組150，以用於訓練或部署推算模型。例如，若提取特徵係導出自從參考個體所擷取之訓練影像，則淋巴結特徵模組145將所提取特徵提供至推算模組150。在各種實施例中，淋巴結特徵模組145將所提取特徵提供至風險訓練模組155，以用於訓練風險模型。例如，若所提取特徵係導出自從參考個體所擷取之訓練影像，則淋巴結特徵模組145將所提取特徵提供至風險訓練模組155，以用於訓練風險模型。在各種實施例中，淋巴結特徵模組145將所提取特徵提供至風險部署模組160，以用於部署風險模型。。例如，若所提取特徵係導出自從一對象（諸如圖1A中之對象110）所擷取之影像，則淋巴結特徵模組145將所提取特徵提供至風險部署模組160。

III. A.訓練一推算模型

【0052】 推算模組150訓練及部署一推算模型。一般而言，該推算模型經結構化使得其：分析從包括一或多個淋巴結之影像所提取的特徵；及針對該一或多個淋巴結預測該等淋巴結為癌性的機率。因此，基於來自包括一或多個淋巴結之一影像的所提取特徵，該推算模型推算該等淋巴結之各者的一機率，所推算機率指示該淋巴結為癌性（例如，轉移性癌症）的可能性。換言之，推算模型藉由判定淋巴結水準之轉移性癌症的風險來辨別癌性與非癌性淋巴結。

【0053】 在各種實施例中，推算模型係下列任一者：迴歸模型（例如，線性迴歸、邏輯迴歸、或多項式迴歸）、決策樹、隨機森林、支持向量機、單純貝氏模型、k平均值群集、或神經網路（例如，前饋網路、卷積神經網路(CNN)、深度神經網路(DNN)、自動編碼器神經網路、生成對抗網路、或遞迴網路（例如，長短期記憶網路(LSTM)、雙向遞迴網路、深度雙向遞迴網路）、或其任何組合。在具體實施例中，該推算模型係一隨機森林分類器或一最小絕對緊縮與選擇運算子(LASSO)分類器。在各種實施例中，該推算模型係一機器訓練模型。在各種實施例中，該推算模型係一梯度提升機器學習模型（例如，xgboost）。

【0054】 該推算模型可使用機器學習實施方法來訓練，諸如下列任一者：線性迴歸演算法、邏輯迴歸演算法、決策樹演算法、支持向量機分類、單純貝氏分類、K最近鄰分類、隨機森林演算法、深度學習演算法、梯度提升演算法、及降維技術（諸如流形學習(manifold learning)、主組分分析、因子分析、自動編碼器正則化、及獨立組分分析）、或其組合。在各種實施例中，該推算模型係使用監督學習演算法、無監督學習演算法、半監督學習演算法（例如，部分監督）、弱監督(weak supervision)、遷移(transfer)、多工學習、或其任何組合訓練。

【0055】 在各種實施例中，該推算模型具有一或多個參數，諸如超參數或模型參數。超參數通常在訓練之前建立。超參數之實例包括學習速率、決策樹的深度或樹葉、深度神經網路中的隱層數、k平均值群集中的聚類數、迴歸模型中的懲罰項、及與代價函數相關聯的正則化參數。模型參數通常在訓練期間調整。模型參數之實例包括與神經網路層中之節點相關聯的權重、支持向量

機中的支持向量、及迴歸模型中的係數。使用訓練資料訓練（例如，調整）該推算模型的模型參數，以改善該推算模型的預測能力。

【0056】 推算模組150使用訓練資料來訓練該推算模型。可儲存訓練資料及/或自訓練資料儲存170取得訓練資料。在各種實施例中，訓練資料包括來自包括獲自參考個體之一或多個淋巴結的訓練影像的所提取特徵。在各種實施例中，參考個體包括健康個體（例如，對照個體）。在各種實施例中，參考個體包括先前經診斷有癌症之個體。在各種實施例中，參考個體包括先前經診斷有已經轉移至淋巴結之癌症的個體。因此，從此類個體所擷取的影像包括一或多個癌性淋巴結。

【0057】 在各種實施例中，用於訓練推算模型之訓練資料包括參考基準真相值，其指示淋巴結是否為癌性的（以下亦稱為「陽性(positive)」或「+」）或淋巴結是否為非癌性的（以下亦稱為「陰性(negative)」或「-」）。在各種實施例中，訓練資料中的參考基準真相值係二進位值，諸如「1」或「0」。例如，癌性的陽性淋巴結可在訓練資料中以「1」之值識別，而非癌性的陰性淋巴結可在訓練資料中以「0」之值識別。總之，推算模組150使用訓練資料來訓練推算模型，以最小化損失函數，使得推算模型可基於輸入（例如，淋巴結之所提取特徵）來更佳地預測結果（例如，淋巴結的淋巴結轉移性癌症之機率）。在各種實施例中，損失函數係針對最小絕對緊縮與選擇運算子(LASSO)迴歸、脊迴歸、或彈性網路(ElasticNet)迴歸中之任一者來建構。

【0058】 在各種實施例中，訓練資料可獲自及/或導出自公開資料庫。例如，訓練資料可獲自及/或導出自國家肺部篩檢試驗(NLST)，其包括參考個體之淋巴結期別（淋巴結轉移期數(N stage)）癌症資訊。在一些實施例中，訓練

資料可不依賴公開資料庫來獲得及收集，例如藉由從複數個參考個體來擷取一或多個淋巴結之影像。例如，此訓練資料可為自訂資料集。自訂資料集的訓練資料可在患者照護期間來獲取，因此可具有對應之淋巴結的手術採樣資訊。訓練資料可包括用於參考個體的淋巴結轉移期數癌症資訊，例如，基於對參考個體的成像及/或活體組織切片結果所先前診斷的淋巴結轉移期數資訊。

【0059】 在各種實施例中，指示淋巴結是否為癌性的參考基準真相值係導出自臨床期別資料的淋巴結期別（淋巴結轉移期數）資訊。例如，公開的NLST資料集包括參考個體之臨床期別資訊（例如，1a期、1b期、2a期、2b期、3a期等），但不包括淋巴結特異度資訊。換言之，NLST資料集無法區分個體淋巴結是否為癌性或非癌性的。在此類實施例中，指示淋巴結是否為癌性的參考基準真相值可基於參考個體癌症之淋巴結轉移期數資訊及其他特性（例如，腫瘤大小、腫瘤位置、腫瘤形狀、及/或腫瘤密度）及/或淋巴結（例如，淋巴結大小、淋巴結密度、及/或淋巴結形狀）來反算。作為第一實例，具有0或1之淋巴結轉移期數的參考個體尚未遇到癌症轉移，且因此，此參考個體之淋巴結係被指派有「陰性」或非癌性的參考基準真相值。作為另一實例，具有2或3之淋巴結轉移期數的參考個體已發生癌症轉移。因此，基於腫瘤及/或淋巴結的特性，淋巴結被指派「陽性」或「陰性」的參考基準真相值。

III. B. 部署—推算模型

【0060】 推算模組150部署經訓練之推算模型以：分析從包括一或多個淋巴結之影像所提取的特徵；及針對該一或多個淋巴結預測該等淋巴結為癌性的機率。在各種實施例中，經訓練之該推算模型係針對一測試資料集來部署。在

各種實施例中，用於訓練推算模型的測試資料集及訓練資料集可係導出自一共同資料集。例如，共同資料集可被分割分成兩個子集，以用於訓練及測試該推算模型。在一些實施例中，共同資料集可經歷50:50的訓練資料集:測試資料集分割。在一些實施例中，共同資料集可經歷60:40的訓練資料集:測試資料集分割。在一些實施例中，共同資料集可經歷80:20的訓練資料集:測試資料集分割。

【0061】 經訓練之該推算模型之部署產出淋巴結的機率（例如，0與1之間的連續機率），各機率指示對應的淋巴結是否可能為癌性的。在各種實施例中，推算模組150提供用於訓練風險模型的由推算模型所預測的淋巴結之淋巴結轉移風險中之一或多者。在各種實施例中，推算模組150提供用於訓練風險模型的由推算模型所預測的淋巴結之所有淋巴結轉移風險。在各種實施例中，推算模組150執行一選擇程序，以識別由推算模型所預測的淋巴結的淋巴結轉移風險之一子集。因此，淋巴結轉移風險之子集隨後用於訓練風險模型。

【0062】 在一些實施例中，推算模組150藉由比較淋巴結轉移風險與臨限值來選擇淋巴結轉移風險之子集。若淋巴結之淋巴結轉移風險高於一臨限值，則選擇將該淋巴結之該淋巴結轉移風險納入該子集，以用於訓練風險模型。若淋巴結之淋巴結轉移風險低於一臨限值，則將將淋巴結之將淋巴結轉移風險從該子集排除，並因此不將其用於訓練風險模型。

【0063】 在各種實施例中，臨限值係一固定值。在各種實施例中，臨限值係50%轉移性癌症風險。在各種實施例中，臨限值係30%、35%、40%、45%、55%、60%、65%、70%、75%、80%、85%、90%、95%、或99%轉移性癌症風險。在一些實施例中，臨限值係介於30%與70%之間。在一些實施例

中，臨限值係介於40%與60%之間。在一些實施例中，臨限值係介於45%與55%之間。在一些實施例中，臨限值係介於48%與52%之間。在一些實施例中，臨限值係介於49%與51%之間。在各種實施例中，臨限值係可變的。在各種實施例中，臨限值係具有癌症的中位數對象內機率。

【0064】 推算模組150提供淋巴結轉移風險之子集至風險訓練模組155，以用於訓練風險模型。

III. C.訓練一風險模型

【0065】 風險訓練模組155使用訓練資料（包括由推算模型所預測的淋巴結之淋巴結轉移風險中之至少一或多者）來訓練風險模型。在此，一或多個淋巴結轉移風險可係由推算模組150所選擇的淋巴結轉移風險之子集，如上文所述。相較於用來訓練推算模型的訓練資料集，用於訓練風險模型的淋巴結轉移風險之子集表示一經純化訓練資料集。具體而言，經純化訓練資料集包括較可能為癌性的淋巴結之子集，而將較不可能為癌性（但除此之外可能表示偽陽性，或者對於訓練風險模型而言係混淆的）的淋巴結從經純化訓練資料集移除。因此，藉由以此純化訓練資料集來訓練風險模型，風險模型能夠針對患有轉移性疾病之對象中的所有淋巴結均預測出較大的機率範圍。更具體而言，風險模型較為可能將較高淋巴結轉移風險指派給很可能有癌症的淋巴結，並將較低淋巴結轉移風險指派給不可能有癌症的淋巴結。

【0066】 在各種實施例中，該風險模型分析從包括一或多個淋巴結之影像所提取的特徵，並針對該一或多個淋巴結之各者預測該淋巴結為癌性的連續機率。因此，基於來自包括一或多個淋巴結之一影像的所提取特徵，該風險模

型判定該等淋巴結之各者的一機率，所判定機率指示該淋巴結為癌性（例如，轉移性癌症）的可能性。在此，風險模型判定淋巴結水準的轉移性癌症之風險。在各種實施例中，風險模型類似於推算模型，因為這兩個模型皆分析從影像所提取的特徵並預測淋巴結水準的轉移性癌症之風險。然而，風險模型係使用推算模型的預測（例如，由推算模型所預測的一或多個淋巴結水準風險）來訓練。總之，此使風險模型能夠預測轉移性癌症之LN水準風險的機率範圍大於由該推算模型所預測的該等經推算LN水準風險。

【0067】 風險模型係使用包括由推算模型所預測的至少一或多個淋巴結轉移風險的訓練資料來訓練。在此，一或多個淋巴結轉移風險可用作為參考基準真相值，以用於訓練風險模型。在各種實施例中，訓練資料可額外包括從獲自參考個體之一或多個淋巴結的訓練影像的所提取特徵。在一些實施例中，用於訓練風險模型的訓練影像可係與先前用於訓練推算模型相同的訓練影像。在一些實施例中，用於訓練風險模型的訓練影像係與先前用於訓練推算模型不同的訓練影像。

【0068】 在各種實施例中，訓練資料包括個體訓練實例，使得風險模型係使用各訓練實例來訓練。例如，各訓練實例可包括一特定訓練影像之一或多個淋巴結的所提取特徵以及參考基準真相值（例如，特定訓練影像的相同一或多個淋巴結之淋巴結轉移風險）。因此，對於各訓練實例，風險模型經訓練以最小化損失函數（例如，針對最小絕對緊縮與選擇運算子(LASSO)迴歸、脊迴歸、或彈性網路迴歸中之任一者的損失函數）。

【0069】 在各種實施例中，風險模型係下列任一者：迴歸模型（例如，線性迴歸、邏輯迴歸、或多項式迴歸）、決策樹、隨機森林、支持向量機、單

純貝氏模型、k平均值群集、或神經網路（例如，前饋網路、卷積神經網路(CNN)、深度神經網路(DNN)、自動編碼器神經網路、生成對抗網路、或遞迴網路（例如，長短期記憶網路(LSTM)、雙向遞迴網路、深度雙向遞迴網路）、或其任何組合。在具體實施例中，風險模型係一隨機森林分類器或一最小絕對緊縮與選擇運算子(LASSO)分類器。在各種實施例中，風險模型係一機器訓練模型。

【0070】 風險模型可使用機器學習實施方法來訓練，諸如下列任一者：線性迴歸演算法、邏輯迴歸演算法、決策樹演算法、支持向量機分類、單純貝氏分類、K最近鄰分類、隨機森林演算法、深度學習演算法、梯度提升演算法、及降維技術（諸如流形學習、主組分分析、因子分析、自動編碼器正則化、及獨立組分分析）、或其組合。在各種實施例中，風險模型係使用監督學習演算法、無監督學習演算法、半監督學習演算法（例如，部分監督）、弱監督、遷移、多工學習、或其任何組合訓練。

【0071】 在各種實施例中，風險模型具有一或多個參數，諸如超參數或模型參數。超參數通常在訓練之前建立。超參數之實例包括學習速率、決策樹的深度或樹葉、深度神經網路中的隱層數、k平均值群集中的聚類數、迴歸模型中的懲罰項、及與代價函數相關聯的正則化參數。模型參數通常在訓練期間調整。模型參數之實例包括與神經網路層中之節點相關聯的權重、支持向量機中的支持向量、及迴歸模型中的係數。使用訓練資料訓練（例如，調整）風險模型的模型參數，以改善風險模型的預測能力。

III. C.部署一風險模型

【0072】 風險部署模組160部署風險模型，以判定一對象（例如，上文在圖1A所述之對象110）的轉移性癌症之一對象水準風險。此處，風險部署模組160將從對象擷取的包括一或多個淋巴結之影像的所提取特徵作為輸入提供至經訓練之該風險模型。經訓練之該風險模型分析所提取特徵，並針對淋巴結輸出指示個體淋巴結是否可能為癌性的機率。

【0073】 在各種實施例中，在部署風險模型之後，風險部署模組160從由風險模型所預測的轉移性癌症之個體淋巴結水準風險變換成對象的轉移性癌症之對象水準風險。在各種實施例中，風險部署模組160組合轉移性癌症之個體淋巴結水準風險，以判定轉移性癌症之對象水準風險。在各種實施例中，風險部署模組160跨轉移性癌症之淋巴結水準風險來執行統計運算。統計運算可識別跨轉移性癌症之個體淋巴結水準風險的最大值、最小值、平均值、中位數值、或眾數值。在具體實施例中，統計運算係跨轉移性癌症之個體淋巴結水準風險的最大值。因此，具有轉移性疾病之最高淋巴結水準風險的淋巴結被指派為對象的轉移性疾病之對象水準風險。

【0074】 癌症分期模組165基於對象的轉移性疾病之對象水準風險來判定對象的癌症期別。在各種實施例中，該癌症分期模組165藉由比較該對象的轉移性疾病之對象水準風險與一或多個臨限風險值來判定該對象的癌症期別。這些臨限風險值可作為用於判定該對象中之癌症是否可能在0期、1期、2期、或3期中之任一者的截止點。

【0075】 在一些實施例中，若對象水準風險大於30%，則對象被分類為具有至少3期淋巴結癌。在一些實施例中，若對象水準風險大於40%、大於

50%、大於60%、大於70%、大於80%、或大於90%風險，則對象被分類為具有至少3期癌症。

【0076】 在各種實施例中，臨限風險值界定對應於一或多個癌症期別的風險範圍。因此，臨限風險值設定風險範圍的上限及下限。在一些實施例中，若對象水準風險係介於30%與100%之間，則將對象分類為具有3期淋巴結癌。在一些實施例中，若對象水準風險係介於40%與100%之間，則將對象分類為具有3期淋巴結癌。在一些實施例中，若對象水準風險係介於50%與100%之間，則將對象分類為具有3期淋巴結癌。

【0077】 在一些實施例中，若對象水準風險係介於10%與40%之間，則將對象分類為具有2期淋巴結癌。在一些實施例中，若對象水準風險係介於15%與40%之間，則將對象分類為具有2期淋巴結癌。在一些實施例中，若對象水準風險係介於20%與40%之間，則將對象分類為具有2期淋巴結癌。在一些實施例中，若對象水準風險係介於30%與40%之間，則將對象分類為具有2期淋巴結癌。在一些實施例中，若對象水準風險係介於15%與30%之間，則將對象分類為具有2期淋巴結癌。在一些實施例中，若對象水準風險係介於20%與30%之間，則將對象分類為具有2期淋巴結癌。

【0078】 在一些實施例中，若對象水準風險係介於5%與20%之間，則將對象分類為具有0期淋巴結癌或1期淋巴結癌。在一些實施例中，若對象水準風險係介於10%與20%之間，則將對象分類為具有0期淋巴結癌或1期淋巴結癌。在一些實施例中，若對象水準風險係介於5%與15%之間，則將對象分類為具有0期淋巴結癌或1期淋巴結癌。在一些實施例中，若對象水準風險係介於8%與12%之間，則將對象分類為具有0期淋巴結癌或1期淋巴結癌。

IV. 用於癌症分期之實例方法

【0079】 圖3A描繪根據一實施例之推算模型及風險模型之一實例實施方案。此處，參考個體305A及305B可經識別為具有特定淋巴結轉移期數。例如，參考個體305A被識別為陰性的淋巴結轉移期數，指示參考個體305A未展現淋巴結癌侵犯。參考個體305B被識別為陽性的淋巴結轉移期數，指示參考個體305B展現淋巴結癌侵犯。然而，對象水準淋巴結轉移期數不區分特定淋巴結，且不區分個體淋巴結為癌性或非癌性。

【0080】 推算模型310推算各淋巴結的轉移風險（例如，淋巴結轉移風險），從而實現具有轉移性癌症之淋巴結與非癌性淋巴結的區分。在各種實施例中，推算模型310針對各淋巴結預測表示該淋巴結為癌性之可能性的機率。

【0081】 如圖3A所示，推算模型310之輸出可識別出個體305C（對應於參考個體305A）未展現淋巴結侵犯。換言之，個體305C之淋巴結皆可能不是癌性的。此確認了陰性淋巴結轉移期數的識別。此外，推算模型310之輸出可識別個體305D（對應於參考個體305B）展現一或多個淋巴結的淋巴結侵犯。例如，推算模型310可推算機率使得在個體305D中只有一個淋巴結被識別為癌性的，而個體305D中的其他淋巴結則係非癌性的。

【0082】 在各種實施例中，由推算模型310所預測的淋巴結之淋巴結轉移風險中之一或多者係用於訓練風險模型320。在各種實施例中，由推算模型310所預測的淋巴結之所有淋巴結轉移風險係用於訓練風險模型320。在各種實施例中，對應於淋巴結之子集的淋巴結轉移風險之子集係用於訓練風險模型

320，從而使風險模型能夠針對患有轉移性疾病之對象中的所有淋巴結均預測出較大的機率範圍。

【0083】 經訓練後，風險模型320經部署以預測一或多個對象315之轉移性癌症之對象水準風險。因此，風險模型320經部署以從轉移性疾病之淋巴結水準風險轉變至轉移性疾病之對象水準風險。如圖3A所示，部署風險模型320經部署以分析從擷取自個體315之影像的所提取特徵。風險模型320預測個體315之各者的轉移性癌症之對象水準風險。例如，風險模型320預測個體315A具有20%轉移性癌症風險，個體315B具有30%轉移性癌症風險，個體315C具有50%轉移性癌症風險，且個體315D具有60%轉移性癌症風險。因此，個體315A及315B可被分類為具有早期癌症（例如，0期或1期），而個體315C及315D可被分類為具有末期癌症（例如，2期，3期、或4期）。

【0084】 個體315可依據其轉移性疾病之對象水準風險接收不同治療。例如，處於具有淋巴結轉移之低風險的個體315A及315B可直接進行至立即治療，諸如手術切除或合併的支氣管鏡診斷及治療。作為另一實例，處於具有淋巴結轉移之較高風險的個體315C及315D可接收額外檢測，諸如PET掃描及淋巴結活體組織切片。

【0085】 圖3B描繪根據一實施例之進一步細節的風險模型320之實施方案。此處，風險模型320分析從對象350擷取之影像的所提取特徵，並輸出對象之一或多個淋巴結之各者的機率。如圖3B所示，風險模型320輸出三個淋巴結的機率。在其他實施例中，風險模型320可輸出對象350之額外或更少的淋巴結機率。

【0086】 如圖3B所示，風險模型320預測第一淋巴結具有20%的機率為癌性，而反之，80%的機率為非癌性。此外，風險模型320預測第二淋巴結具有30%的機率為癌性，而反之，70%的機率為非癌性。此外，風險模型320預測第三淋巴結具有60%的機率為癌性，而反之，40%的機率為非癌性。個體淋巴結機率經分析以識別轉移性癌症之對象水準風險。在此實例中，將轉移性癌症的最大淋巴結機率當作是轉移性癌症之對象水準機率。具有60%的機率為癌性的第三淋巴結係淋巴結之中的最高機率。因此，對象350的轉移性癌症之對象水準機率係60%。

V. 癌症

【0087】 本文中所述之方法涉及藉由判定對象的癌症之對象水準風險來進行癌症分期。在各種實施例中，對象中的癌症可包括以下一或多者：淋巴瘤、B細胞淋巴瘤、T細胞淋巴瘤、蕈狀肉芽腫、何杰金氏病(Hodgkin's Disease)、骨髓性白血病、膀胱癌、腦癌、神經系統癌症、頭頸癌、頭頸鱗狀細胞癌、腎癌、肺癌、神經胚細胞瘤/神經膠質母細胞瘤、卵巢癌、胰臟癌、前列腺癌、皮膚癌、肝癌、黑色素瘤、口、咽喉、喉、肺部的鱗狀細胞癌、結腸癌、子宮頸癌(cervical cancer)、子宮頸癌(cervical carcinoma)、乳癌、及上皮癌症、腎癌、泌尿生殖道癌症、肺癌(pulmonary cancer)、食道癌、胃癌、甲狀腺癌、頭頸癌、大腸癌、造血性癌症、睪丸癌、結腸及/或直腸癌、子宮癌、或前列腺癌。在各種實施例中，癌症可係下列任一者：支氣管肺泡癌(bronchioloalveolar carcinoma, BAC)、膀胱癌、雌性生殖道惡性腫瘤（例如，子宮漿液性腺癌、子宮內膜癌、外陰鱗狀細胞癌、及子宮肉瘤）、卵巢表面上皮

第 31 頁，共 52 頁(發明說明書)

癌（例如，卵巢透明細胞癌、上皮卵巢癌、輸卵管癌、及原發性腹膜癌）、乳房癌、非小細胞肺癌(non-small cell lung cancer, NSCLC)、雄性生殖道惡性腫瘤（例如，睪丸癌）、後腹腔或腹膜癌、胃食道腺癌、胃食道接合部癌、肝細胞癌、食道胃接合部癌、子宮頸癌、膽管癌、胰臟腺癌、肝外膽管癌、小腸惡性腫瘤、胃癌、原發部位不明轉移癌(cancer of unknown primary, CUP)、結直腸腺癌、食道癌、攝護腺癌、腎癌、頭頸鱗狀細胞癌、胸腺癌、非黑色素瘤皮膚癌、甲狀腺癌（例如，乳突癌）、頭頸癌、肛癌、非上皮卵巢癌(non-epithelial ovarian cancer, non-EOC)、葡萄膜黑色素瘤、惡性肋膜間皮瘤、小細胞肺癌(small cell lung cancer, SCLC)、中樞神經系統癌、神經內分泌瘤、及軟組織腫瘤。在某些實施例中，該癌症係乳癌、非小細胞肺癌、膀胱癌、腎臟癌、結腸癌、及黑色素瘤。在一些實施例中，對象的癌症可係轉移性癌症，包括膀胱癌、乳癌、結腸癌、腎癌、肺癌、黑色素瘤、卵巢癌、胰臟癌、前列腺癌、直腸癌、胃癌、甲狀腺癌、或子宮癌中任一者。在具體實施例中，癌症係肺癌。在具體實施例中，癌症係一種肺癌，包括小細胞肺癌、非小細胞肺癌、腺癌、鱗狀細胞癌、大細胞癌、小細胞癌、組合式小細胞癌、肺部肉瘤、肺部淋巴瘤、支氣管類癌中任一者。

VI. 導引介入

【0088】 本文中所述之實施例涉及藉由判定對象中之轉移性癌症之對象水準風險來判定癌症分期。在各種實施例中，基於針對一對象所預測的癌症分期及/或轉移性癌症之對象水準風險而提供介入給該對象。此類介入（以下稱為導引介入）可為下列任一者：診斷施用、治療劑之施用、對象的後續檢測、後

續的活體組織切片（例如，癌活體組織切片或淋巴結活體組織切片、淋巴結的支氣管鏡採樣）、後續影像掃描（例如，PET掃描、MRI掃描超音波成像、或X射線成像）、腫瘤切除、支氣管鏡診斷、選擇及/或投予（多個）治療劑、選擇/投予醫藥組成物、納入臨床試驗、不納入臨床試驗、或其任何組合。

【0089】 在各種實施例中，針對對象所預測的轉移性癌症之對象水準風險或所預測的癌症期別指示該對象處於高轉移性癌症風險。在各種實施例中，處於高轉移性癌症風險的對象具有高於一臨限評分的轉移性癌症之所預測對象水準風險。在各種實施例中，臨限評分高於20%轉移性癌症風險。在各種實施例中，臨限評分高於25%轉移性癌症風險。在各種實施例中，臨限評分高於30%、高於35%、高於40%、高於45%、高於50%、高於55%、高於60%、高於65%、高於70%、高於75%、高於80%、高於85%、高於9%、高於95%、或高於99%轉移性癌症風險。在各種實施例中，臨限評分係介於20%與80%之間。在一些實施例中，臨限值係介於30%與70%之間。在一些實施例中，臨限值係介於40%與60%之間。在一些實施例中，臨限值係介於45%與55%之間。在一些實施例中，臨限值係介於20%與60%之間。在一些實施例中，臨限值係介於30%與55%之間。在一些實施例中，臨限值係介於35%與50%之間。在一些實施例中，臨限值係介於40%與45%之間。

【0090】 在各種實施例中，若對象之所預測癌症期別係至少2期（例如，2期或3期），則該對象處於高轉移性癌症風險。在此類實施例中，導引介入涉及對象之後續檢測、對象之後續活體組織切片、或對象之後續影像掃描中之一者。此允許臨床提供者對最高風險之對象排定後續檢測、活體組織切片、或影像掃描的優先順序

【0091】 在各種實施例中，若對象經預測處於高轉移性癌症風險，則可為該對象排除對於後期轉移性癌症不適當的醫療程序。所排除的此類醫療程序之實例包括腫瘤切除、肺切除、放射（例如，外部光束放射）、化療（例如，局部或全身化療）、免疫治療、或甚至全身麻醉。

【0092】 在具體實施例中，若對象經預測為處於高轉移性癌症風險，則該對象接受後續影像掃描以確認預測。例如，若對象經預測為處於高轉移性癌症風險，則該對象進一步接受PET掃描。在此類實施例中，若PET掃描回報陰性結果（例如，無淋巴結侵犯），則該對象可接受進一步檢測。例如，該對象可進一步接受後續活體組織切片以明確地採樣淋巴結。在一些實施例中，後續活體組織切片係淋巴結的支氣管鏡採樣。

【0093】 在具體實施例中，若對象經預測為處於高轉移性癌症風險，則該對象接受後續活體組織切片以確認預測。例如，該對象不需要接受後續影像掃描以確認預測，而是替代地直接接受後續活體組織切片。例如，若對象經預測為處於高轉移性癌症風險，則該對象進一步接受淋巴結的支氣管鏡採樣。

【0094】 如本文所述，轉移性癌症之對象水準風險可從跨淋巴結的個體淋巴結水準風險來判定。因此，對於基於對象水準風險而經預測具有高轉移性癌症風險的對象，可識別導致高對象水準風險的（多個）特定淋巴結。因此，該等特定淋巴結可係後續檢測或治療的目標。在具體實施例中，該等特定淋巴結可接受後續影像掃描（例如，PET掃描）以確認預測。在具體實施例中，該等特定淋巴結可接受後續活體組織切片以明確地採樣特定淋巴結。因此，後續活體組織切片可確認在該等特定淋巴結中的轉移性癌症之存在或不存在。

【0095】 在各種實施例中，若對象經預測為處於高轉移性癌症風險，則該對象可納入臨床試驗。例如，此一臨床試驗可係檢測對於轉移性癌症之治療療效，且因此，其可在其合格標準中指定納入者將展現出轉移性癌症。在各種實施例中，若對象經預測為處於高轉移性癌症風險，則該對象被中止納入臨床試驗。例如，此一臨床試驗可檢測對於一種非轉移性癌症的治療療效。

【0096】 在各種實施例中，針對對象所預測的轉移性癌症之對象水準風險指示該對象處於低轉移性癌症風險。在各種實施例中，處於低轉移性癌症風險的對象具有低於一臨限評分之所預測對象水準風險。在各種實施例中，臨限評分低於20%轉移性癌症風險。在各種實施例中，臨限評分低於25%轉移性癌症風險。在各種實施例中，該臨限評分低於30%、低於35%、低於40%、低於45%、低於50%、低於55%、低於60%、低於65%、低於70%、低於75%、低於80%、低於85%、低於90%、低於95%、或低於99%轉移性癌症風險。在一些實施例中，臨限值係介於5%與40%之間。在一些實施例中，臨限值係介於10%與30%之間。在一些實施例中，臨限值係介於15%與25%之間。在一些實施例中，臨限值係介於18%與22%之間。在各種實施例中，若對象之所預測癌症期別低於2期（例如，0期或1期），則該對象處於低轉移性癌症風險。在此類實施例中，導引介入涉及以下中之一者：腫瘤切除、支氣管鏡診斷、選擇及/或投予治療劑、及/或選擇/投予醫藥組成物。此實現快速識別及治療尚未發生轉移的對象。

【0097】 在各種實施例中，可基於轉移性癌症之對象水準風險來選擇及/或投予治療劑至對象，該經選擇之治療劑可能展現出對該癌症的療效。例示性治療劑包括化療、放射、抗原特異性單株抗體、或免疫療法。在各種實施例

中，提供治療劑給處於高轉移性癌症風險的對象。在各種實施例中，處於高轉移性癌症風險的對象具有高於一臨限評分的轉移性癌症之所預測對象水準風險。在各種實施例中，臨限評分高於20%轉移性癌症風險。在各種實施例中，臨限評分高於25%轉移性癌症風險。在各種實施例中，臨限評分高於30%、高於35%、高於40%、高於45%、高於50%、高於55%、高於60%、高於65%、高於70%、高於75%、高於80%、高於85%、高於90%、高於95%、或高於99%轉移性癌症風險。在各種實施例中，若對象之所預測癌症期別係至少2期（例如，2期、3期、或4期），則該對象處於高轉移性癌症風險。

【0098】 在各種實施例中，提供治療劑給處於低轉移性癌症風險的對象。在各種實施例中，處於低轉移性癌症風險的對象具有低於一臨限評分之所預測對象水準風險。在各種實施例中，臨限評分低於20%轉移性癌症風險。在各種實施例中，臨限評分低於25%轉移性癌症風險。在各種實施例中，該臨限評分低於30%、低於35%、低於40%、低於45%、低於50%、低於55%、低於60%、低於65%、低於70%、低於75%、低於80%、低於85%、低於90%、低於95%、或低於99%轉移性癌症風險。在各種實施例中，若對象之所預測癌症期別低於2期（例如，0期或1期），則該對象處於低轉移性癌症風險。

【0099】 在各種實施例中，若對象經預測為處於低轉移性癌症風險，則該對象可被導引到早期癌症治療。早期癌症治療之實例包括腫瘤切除、肺切除、化療（例如，局部或全身化療）、局部遞送治療（例如，支氣管針滴注(transbronchial needle instillation)或心臟電燒術(catheter based ablation)）、或放射。

【0100】 在具體實施例中，若對象經預測為處於低轉移性癌症風險，則該對象接受後續影像掃描以確認預測。例如，若對象經預測為處於低轉移性癌症風險，則該對象進一步接受PET掃描。在此類實施例中，若PET掃描回報陰性結果（例如，無淋巴結侵犯），則PET掃描結果符合預測。然而，若PET掃描回報陽性結果（例如，淋巴結侵犯），則該對象可接受進一步檢測。例如，該對象可進一步接受後續活體組織切片以明確地採樣淋巴結。在一些實施例中，後續活體組織切片係淋巴結的支氣管鏡採樣。

【0101】 在各種實施例中，治療劑係生物製劑，例如細胞介素、抗體、可溶性細胞介素受體、反義寡核苷酸、siRNA等。此類生物劑涵蓋生物劑之突變蛋白及衍生物，該衍生物可包括例如所屬技術領域中已知之融合蛋白、聚乙二醇化衍生物、膽固醇共軛衍生物、及類似物。亦包括細胞介素及細胞介素受體之拮抗劑，例如誘捕劑及單株拮抗劑。亦包括本文所述之活性劑之生物類似或生物等效藥物。

【0102】 用於肺癌的治療劑可包括化學治療劑，諸如多西紫杉醇、順鉑、卡鉑、吉西他濱(gemcitabine)、白蛋白結合型太平洋紫杉醇(Nab-paclitaxel)、太平洋紫杉醇、培美曲塞(pemetrexed)、吉非替尼(gefitinib)、埃羅替尼(erlotinib)、布格替尼(brigatinib) (Alunbrig[®])、卡馬替尼(capmatinib) (Tabrecta[®])、賽爾帕替尼(selpercatinib) (Retevmo[®])、恩曲替尼(entrectinib) (Rozlytrek[®])、洛拉替尼(lorlatinib) (Lorbrena[®])、拉羅替尼(larotrectinib) (Vitrakvi[®])、達可替尼(dacomitinib) (Vizimpro[®])、及長春瑞濱(vinorelbine)。用於肺癌的治療劑可包括抗體療法，諸如德瓦魯單抗(durvalumab) (Imfinzi[®])、尼沃

魯單抗(nivolumab) (Opdivo[®])、派姆單抗(pembrolizumab) (Keytruda[®])、阿特珠單抗(atezolizumab) (Tecentriq[®])、及雷莫蘆單抗(ramucirumab)。

【0103】 在各種實施例中，可基於轉移性癌症之對象水準風險來選擇及/或投予醫藥組成物至對象，該經選擇之治療劑可能展現出對該癌症的療效。向個體投予之醫藥組成物包括活性劑，諸如上述治療劑。活性成分係以治療有效量存在，亦即當經投予時足以治療疾病或由其所媒介之醫療病況的量。組成物亦可包括各種其他劑以增強遞送及療效，例如增強活性成分的遞送及穩定性。因此，例如，組成物取決於所欲之配方，亦可包括醫藥上可接受、無毒性載劑或稀釋劑，其定義為常用於調配用於動物或人類投予之醫藥組成物的媒劑。稀釋劑係經選擇以不影響該組合之生物活性。此類稀釋劑之實例係蒸餾水、緩衝水、生理食鹽水、PBS、林格氏(Ringer' s)液、右旋糖溶液、及Hank氏溶液。此外，醫藥組成物或配方可包括其他載劑、佐劑、或毒性、非治療性、非免疫原性穩定劑、賦形劑、及類似物。組成物亦可包括額外物質以近似生理條件，諸如pH調整及緩衝劑、毒性調整劑、潤濕劑、及清潔劑。組成物亦可包括多種穩定劑之任一者，諸如抗氧化劑。

【0104】 在本文中描述之醫藥組成物可以各種不同方式投予。實例包括經由口服、鼻內、直腸、外用、腹膜內、靜脈、肌肉、皮下、表皮下、經皮、鞘內或顱內方法投予含有醫藥上可接受之載劑的組成物。在各種實施例中，醫藥組成物可經由基於導管之部署（例如，使用可撓性或剛性機器性支氣管鏡檢查）來投予。

【0105】 可投予此類醫藥組成物以用於治療（例如，診斷患者患有肺癌之後）之目的。如在本發明的上下文中針對疾病或病症所使用的預防、疾病預

防、或預防措施係指投予組成物以預防肺癌或肺癌之所有症狀的發生、發作、惡化、或復發，或減輕肺癌發作的可能性。根據本發明，肺癌的治療(treating/treatment)或療法(therapy)應意指藉由投予治療來減緩、停止、或反轉癌症的惡化。在較佳實施例中，治療肺癌係指反轉癌症的惡化，理想的程度為至消除癌症本身的程度。

VII. 電腦實施方案

【0106】 本發明之方法（包括用於藉由判定對象的轉移性癌症之對象水準風險來進行癌症分期之方法）在一些實施例中係在一或多個電腦上執行。

【0107】 例如，可在硬體或軟體或兩者的組合中實施推算模型及/或風險模型以及資料庫儲存的建置及部署。在本發明之一實施例中，提供機器可讀取儲存媒體，該媒體包含經編碼有機器可讀取資料的資料儲存材料，當使用經指令程式化以使用該資料之機器時，該機器可讀取資料能夠顯示資料集以及本文所述之模型的執行及結果中之任何者。本發明可在執行於可程式化電腦上之電腦程式中實施，該等可程式化電腦包含處理器、資料儲存系統（包括揮發性及非揮發性記憶體及/或儲存元件）、圖形配接器、指向裝置、網路配接器、至少一個輸入裝置、及至少一個輸出裝置。顯示器係耦接至圖形配接器。程式碼經應用於輸入資料，以執行上文描述之功能並產生輸出資訊。輸出資訊係以已知方式應用至一或多個輸出裝置。電腦可為例如個人電腦、微電腦、或習知設計的工作站。

【0108】 各程式可實施於高階程序或物件導向程式設計語言以與電腦系統溝通。然而，若需要，程式可以組合或機器語言實施。在任何情況下，語言

可為編譯或解譯語言。各此類電腦程式較佳係儲存於可由一般或特殊目的可程式電腦所讀取之儲存媒體或裝置（例如，**ROM**或磁碟），用於當儲存媒體或裝置由電腦讀取以執行本文所述之程序時組態及操作電腦。系統亦可視為實施為經電腦程式組態之電腦可讀取儲存媒體，其中儲存媒體經組態造成電腦以特定及預定方式操作以執行本文所述之功能。

【0109】 簽章模式及其資料庫可提供於多種媒體以利其等之使用。「媒體(media)」係指含有本發明之簽章模式資訊的製品。本發明之資料庫可記錄於電腦可讀取媒體上，例如可由電腦直接讀取及存取的任何媒體。此類媒體包括但不限於：磁性儲存媒體，諸如軟碟、硬碟儲存媒體、及磁帶；光學儲存媒體，諸如**CD-ROM**；電性儲存媒體，諸如**RAM**及**ROM**；及這些類別的混合體，諸如磁性/光學儲存媒體。所屬技術領域中具有通常知識者可輕易明白可如何使用任何目前已知的電腦可讀取媒體以產生包含本資料庫資訊之記錄的製品。「記錄(recorded)」係指使用所屬技術領域中已知的任何此類方法將資訊儲存在電腦可讀取媒體上之過程。可基於用以存取所儲存資訊之構件，選擇任何便利的資料儲存結構。多種資料處理器程式及格式可用於儲存（例如文字處理文字檔）、資料庫格式等。

【0110】 在一些實施例中，本發明之方法（包括用於藉由判定對象的轉移性癌症之對象水準風險來進行癌症分期之方法）係在分散式運算系統環境（例如，雲端運算環境）中之一或多個電腦上執行。在此描述中，「雲端運算(cloud computing)」係定義為用於實現對於一共享之可組態運算資源集之隨選網路存取的模型。可採用雲端運算來對該共享之可組態運算資源集提供隨選存取。該共享之可組態運算資源集可經由虛擬化來快速提供，且以低管理努力或

服務提供者互動來釋放，並接著相應地按比例調整。雲端運算模型可由各種特性構成，諸如例如隨選自我服務、廣泛網路存取、資源集區、快速彈性、測量服務等。雲端運算模型亦可暴露各種服務模型，諸如例如軟體及服務

(「SaaS」)、平台即服務(「PaaS」)、及基礎設施即服務(「IaaS」)。雲端運算模型亦可使用不同的部署模型(諸如，私有雲端、社群雲端、公用雲端、混合雲端等)來部署。在本說明書及申請專利範圍中，「雲端運算環境(cloud-computing environment)」係採用雲端運算的環境。

【0111】 圖4繪示一實例電腦，其用於實施圖1A、圖1B、圖2A、圖2B、圖3A、及圖3B所示之實體。電腦400包括至少一個處理器402，其耦接至晶片組404。晶片組404包括記憶體控制器中樞420及輸入/輸出(I/O)控制器中樞422。記憶體406及圖形配接器412係耦接至記憶體控制器中樞420，且顯示器418係耦接至圖形配接器412。儲存裝置408、輸入裝置414、及網路配接器416係耦接至I/O控制器中樞422。電腦400的其他實施例具有不同架構。

【0112】 儲存裝置408係非暫態電腦可讀取儲存媒體，諸如硬碟、唯讀光碟記憶體(CD-ROM)、DVD、或固態記憶體裝置。記憶體406保存由處理器402使用之指令及資料。輸入介面414係觸控螢幕介面、滑鼠、軌道球、或其他類型之指向裝置、鍵盤、或其一些組合，且係用於輸入資料至電腦400中。在一些實施例中，電腦400可經組態以接受來自輸入介面414之經由使用者姿勢之輸入(例如，命令)。圖形配接器412顯示影像及其他資訊在顯示器418上。網路配接器416將電腦400耦接至一或多個電腦網路。

【0113】 電腦400經調適以執行用於提供本文所述功能的電腦程式模組。如本文中所使用，用語「模組(module)」係指用於提供指定功能的電腦程式邏

輯。因此，可在硬體、韌體、及/或軟體中實施模組。在一實施例中，程式模組係儲存在儲存裝置408上、裝載至記憶體406中、並由處理器402執行。

【0114】 由圖1A或圖1B之實體所使用的電腦400之類型可取決於實施例及實體所需之處理能力而變化。例如，癌症分期系統130可在單一電腦400中或在透過網路（諸如在伺服器場中）彼此溝通之多個電腦400中運行。電腦400可缺乏上述之一些組分，諸如圖形配接器412、及顯示器418。

VIII. 系統

【0115】 本文進一步揭示用於藉由判定一對象的轉移性癌症之一對象水準風險來進行癌症分期的系統。在各種實施例中，此一系統可至少包括上文在圖1A所述之癌症分期系統130。在各種實施例中，癌症分期系統130體現為電腦系統，諸如具有圖4所述之實例電腦400的電腦系統。

【0116】 在各種實施例中，系統包括成像裝置，諸如上文在圖1A所述之成像裝置120。在各種實施例中，系統包括癌症分期系統130（例如，電腦系統）及成像裝置兩者。在此類實施例中，癌症分期系統130可與成像裝置120可溝通地耦接，以接受自一對象擷取之影像（例如，CT掃描）。電腦系統於電腦上(*in silico*)實施推算模型與風險模型兩者中之一或兩者，以分析影像及判定對象之淋巴結疾病之對象水準風險。

實例

【0117】 以下係用於實行本發明之具體實施例之實例。實例的提供僅為了說明目的而非意欲以任何方式限制本發明之範疇。已致力於確保關於所使用數字（例如，量、溫度等）之精確性，但應允許一些實驗誤差及偏差。

實例1：研究族群及特徵提取

【0118】 此研究係使用來自國家肺部篩檢試驗(NLST)之CT組的資料來執行。此NLST之次世代包括來自658位研究參與者的資料，301位具有良性結節，而357位患有肺癌。658名研究參與者的各CT掃描均經過檢查，且高達21個淋巴結站（各人16+/-6）經標註（9,969個淋巴結經標註；4,631個淋巴結來自對照組，及5,318個淋巴結來自肺癌患者）。從此初步樣本，患有肺癌的288人具有可用於分析的分期資料。為了建立平衡的初始世代，識別出287名對照對象。淋巴結站係描繪在圖2A中且詳述於表1中。

【0119】

〔表1〕：各站的淋巴結標註

淋巴站	淋巴結數量（每對象之平均）	
	對照組	癌症組
1L	34	22
1R	29	13
2L	191	186
2R	732	879
3A	203	263
3P	133	139
4L	430	459
4R	837	987
5	386	457
6	189	132

7	477	541
8L	212	230
8R	140	215
9L	29	98
9R	8	14
10L	181	188
10R	92	60
11L	74	106
11R	57	59
12L	5	4
12R	5	8

【0120】 針對各淋巴結站，界定具有5 mm及7.5 mm半徑的一球形區域。由標註標籤所界定之球面中心。圖2B展示兩個淋巴結標註之一實例。

【0121】 這些較小(5 mm)及較大(7.5 mm)的球面ROI（關注區域）係用於探索用於模型建置之影像範圍。從這些球面ROI，提取3個分開的影像圖譜板，各有6種特徵類別。各板所使用的度量數目不同，且此係基於CT的不同變換。板1使用原始影像作為特徵提取之輸入。板2及3分別採用輸入影像之小波變換及高斯拉普拉斯(LoG)變換。特徵類別包括一階特徵、形狀特徵、灰階共生矩陣1 (GLCM)特徵、灰階級長矩陣(GLRLM)特徵、灰階級帶矩陣(GLSZM)特徵、及鄰域灰階差分矩陣(NGTDM)特徵。表2展示特徵細項。

【0122】

〔表2〕：來自CT掃描影像的特徵

特徵數目						
特徵類別						
影像類別	一階	形狀*	GLCM	GLRLM	GLSZM	NGTDM
原始	22	14	24	16	16	5

小波	144	0	192	128	128	40
LoG	18	0	24	16	16	5

*由於使用固定的球形形狀區域分析各站，所以形狀特徵不包括在淋巴結分析中

【0123】 圖5A至圖5R描繪跨GLC特徵類別之不同板的不同淋巴結影像圖譜特徵的實例相關性熱圖。此外，圖5A至圖5R各展示特定特徵。圖5A至圖5R中展示之此類特徵的實例包括相關特徵、LMC2、MCC、差熵、聯合熵、和熵、群聚突出、平均差、聯合平均、平均加總、自動相關度、變異性差別、對比、群聚傾向、差值平方和、LDMN、LDN、群聚陰暗度、LMC1、逆方差、LD、LDM、聯合能量、及最大機率。圖5C至圖5R展示相對於不同濾波器之特徵。具體而言，圖5C至圖5R提供「XYZ」之命名，其中「X」、「Y」、及「Z」各係指應用至影像的X、Y、或Z方向的濾波器。「H」係指高通濾波器，而「L」係指低通濾波器。因此，「HHH」之命名係指從在X、Y、及Z方向之各者中經受高通濾波的3D影像所導出的特徵。「HHL」之命名係指從在X及Y方向上經受高通濾波，且在Z方向經受低通濾波的特3D影像所導出的特徵。

實例2：用於癌症分期的推算模型

【0124】 模型化始於預測各淋巴結之癌症侵犯。若要這樣做，對各標註之淋巴結台推算（NLST中報告的）癌症期別（NLST中未報告個體淋巴結資料，所以將淋巴結侵犯從癌症期別「排除」）。接著將訓練:測試資料進行50:50分割，其具有表2中所提供之資料分布。

【0125】

〔表2〕：實例的50:50資料分割

	訓練		測試	
淋巴結轉移期數	對照組	癌症組	對照組	癌症組
0	2337	2179	2263	1713
1	0	127	0	73
2	0	125	0	342
3	0	40	0	42

【0126】 該模型之結果係二進位（使用NLST資料集中的淋巴結轉移期數共變異數，淋巴轉移「是/否」。「否」係N=0或N=1，且「是」係N=2或N=3）。將度量正規化（z-評分），並訓練及測試隨機森林及LASSO分類器。針對各淋巴結，產生其為癌性的連續機率。例如，淋巴結可能有20%的機率為癌性（藉由減法，意指其被指派80%的機率為無癌）。整個研究參與者內的所有淋巴結皆視為獨立資料。圖6展示初始模型之效能。具體而言，該模型達成曲線下方面積(AUC) = 65.3%、準確度(ACC) = 96.6%、靈敏度= 61.2%、及特異度= 60.8%。

【0127】 對NLST資料的限制在於淋巴結特異度資訊無法取得，且平均16個淋巴結被標註，且每個人接受基於影像圖譜之評估。因此，此模型假設在患有轉移性癌症之患者內的所有淋巴結為「+」，而在臨床醫學上，一般只有1個淋巴結具有癌細胞（癌症不會同時擴散至所有淋巴結，但通常擴散至僅一個或兩個淋巴結）。因此，從對象水準癌症期別推算個體淋巴結侵犯引入了偽陽性，該等偽陽性混淆了淋巴結水準的模型效能（同樣因為在人的CT掃描內的所有淋巴結係假設為「+」）。

【0128】 然後，針對給定對象之所有淋巴結站，使用由淋巴結水準分類器所預測的最高機率評分來判定淋巴結轉移期數可能性。圖7A至圖7B繪示基於淋巴結轉移期數可能性的對象水準效能（AUC = 69.1%、靈敏度= 72.7%、及特異度= 61.6%）。亦將評分（例如，最大機率）的分布橫跨淋巴結轉移期數來繪製，如在圖7B中所示。在此，針對具有3期癌症（左圖）的情況，觀察到較高的最大機率。

實例3：遞增性淋巴結分類方法：對患者內淋巴結進行拔靴抽樣

【0129】 執行兩步驟式拔靴法以從轉移性疾病之淋巴結水準風險轉變至轉移性疾病之對象水準風險。具體而言，兩步驟式拔靴法包括實施推算模型及/或風險模型，如上文參考圖3A所述。在此，挑戰在於：對於所有淋巴結站推算對象的淋巴結轉移期數，以訓練分類器。鑑於大多數淋巴結係無癌的，此推算在淋巴結站引入錯誤。

【0130】 兩個步驟如下：

1. 推算：對於各對象，很可能被癌症侵犯的淋巴結別為具有已知轉移性疾病。此係針對各對象藉由選擇具有高於中位數對象內值（涉及圖 3A 中所示之推算模型的推算程序）的為癌性之機率（從最後模型輸出）的淋巴結來完成。低於具有癌症之中位數對象內機率的淋巴結被分類為陰性。
2. 分類：世代被分割為 50:50（訓練:測試）。接著，使用淋巴結結果之推算集區來訓練第二分類器（例如，圖 3A 之風險模型 320）以預測各對象之淋巴結疾病存在。

【0131】 此模型之輸出係各淋巴結為癌性之連續機率，其類似於上文實例2中所述之模型之輸出。然而，額外推算及分類步驟的優點在於：具有轉移性疾病的對象中所有淋巴結現在經指派有較大的機率範圍。更具體而言，可能具有癌症的淋巴結具有較高之為「+」的機率，而這些不太可能具有癌症的淋巴結具有較低之為「+」的機率。對於預測步驟，基於具有最高機率（圖3B）的淋巴結而產生淋巴結侵犯的對象水準評分。具體而言，此涉及：採用淋巴結特定機率，並提供具有其肺癌之淋巴結侵犯的對象特定機率。在圖3B所示之檢測對象中，底部淋巴結被指派有60%機會患有癌症（例如，機率(+) = 0.6）。此係三個淋巴結之中具有最高「+」機率之淋巴結。因此，對象水準機率被指派0.6的值。

【0132】 在使用隨機森林分類器訓練該遞增方法之後，拔靴抽樣模型之對象水準效能達成AUC = 77.7%、靈敏度= 81.8%、及特異度= 59.9%，如在圖9A中所繪示。因此，相較於實例2中所述之方法，兩步驟式拔靴抽樣模型的效能獲得改善。使用LASSO分類器達成類似的效能。最後，該模型可用於繪製橫跨淋巴結轉移期數（圖9B）之評分（或最大機率）的分布。有趣的是，注意到，對於具有較高淋巴結轉移期數的案例，模型產出較高的最大機率。此機率評分可用於不同的臨床背景情境中，以實現不同的檢傷(triage)決策工具。用於訓練及測試之額外資料可提供預測N作為分類而非二進位共變異數的模型。

【0133】 所有在此申請案中引用之所有公開案、專利、專利申請案及其他文件全文皆特此為所有目的以引用方式併入本文中，如同各個別公開案、專利、專利申請案或其他文件係個別地指示為所有目的以引用方式併入本文中。

【0134】 雖然已經說明及描述各種具體實施例，但上述說明書不具限制性。將理解的是，在不偏離（多個）本揭露之精神和範圍下，可做出各種變化。許多變異在所屬技術領域中具有通常知識者檢視本說明書後將變得顯而易見。

實例4：使用以CT影像為基礎之預測裝置的實例導引介入

【0135】 實施本文中描述之方法（例如，部署用於預測淋巴結(LN)之轉移性癌症的淋巴結水準風險的風險模型）的以CT影像為基礎之裝置將影響臨床照護，此係因為涉及PET掃描的習知方法之不良效能。最近檢查接受系統性淋巴結切除/採樣的219名NSCLC患者檢查術前PET分期的效能（例如，Li等人的「Implications of False Negative and False Positive Diagnosis in Lymph Node Staging of NSCLC by Means of 18F-FDG PET/CT. PLoS ONE 8(10):e78552」，其全部內容以引用方式併入本文中）。用於偵測肺門及縱隔腔淋巴結之PET的靈敏度、特異度、PPV、及NPV係74.2%、54.4%、86.8%、及73.5%，其中假陰性率係13.2%及偽陽性率係45.5%。因此，本文所述之方法表示用於偵測已導致淋巴結侵犯之轉移性疾病的非侵入性技術，其將改善決策及臨床照護。此裝置將影響決策及臨床照護的實例包括：

【0136】 促使對患有大小屬轉移低風險之腫瘤的對象進行額外檢測。由於對於患有小肺腫瘤者使用術前PET掃描或淋巴結活體組織切片沒有注意標準，所以基於CT裝置所擷取之CT影像預測淋巴結侵犯促使經由對象之淋巴結的PET掃描或支氣管鏡採樣進行後續檢測。相較之下，習知方法將導引對象進行手術，且其中更惡化的疾病期將僅被認為要進行外科手術。

【0137】 促使對獲得陰性PET結果者的淋巴結進行以組織為基礎的採樣。雖然普遍認識到PET具有偽陽性及假陰性失敗率，但是診斷檢測受到PET掃描結果強烈的影響。患有較小腫瘤大小及陰性PET的患者經常進行手術而不會進行額外的淋巴結採樣。此患者之子集會受益於在手術前進行額外的淋巴結採樣。

【0138】 對具有偽陽性PET掃描的患者進行分期，並示範其等對手術介入的合格性。目前，基於PET掃描而強烈疑似患有惡化期別疾病的患者會進行非治癒性全身性/非手術療法，其等不含額外的確認性淋巴結採樣。接收淋巴結侵犯之預測（例如，基於淋巴結的CT影像）的患者及陰性PET結果的患者將替代地在決定手術之前進行更明確的淋巴結採樣。因此，這些患者可為外科切除的候選者，並且會增加其治癒機會。

【0139】 告知哪個淋巴結最可能具有轉移性疾病。處於高轉移性疾病風險的患者進行淋巴結的支氣管鏡或手術採樣。一般而言，此策略會強度聚焦在靠近腫瘤的淋巴結，而使更遠端淋巴結未被採樣或欠採樣。在此，本文所述之方法實現淋巴結侵犯之預測，並且此外，識別被侵犯的淋巴結。此導引臨床團隊適當地聚焦其等之採樣策略以獲得組織、最大化產率、及最小化誤診的風險。

【0140】 精進臨床試驗以著重於納入無淋巴結轉移的患者。患有局部疾病的患者通常尋求臨床試驗對於局部遞送的臨床試驗（化療之支氣管針滴注或燒術）的臨床試驗。不幸地，經受淋巴結採樣時，此類患者在手術期間經常發現具有更惡化期癌症。此延遲了適當的照護。因此，實施本文中描述之方法（例如，部署用於預測淋巴結(LN)之轉移性癌症的淋巴結水準風險的風險模

型)實現識別展現出淋巴結侵犯的患者，且因此，這些患者不需要被納入著重於無淋巴結轉移之患者的臨床試驗。

【符號說明】

【0141】

100:系統環境概觀；系統環境

110:對象

120:成像裝置

130:癌症分期系統

140:癌症預測

145:淋巴結特徵模組

150:推算模組

155:風險訓練模組

160:風險部署模組

165:癌症預測模組；癌症分期模組

170:訓練資料儲存

305A:參考個體

305B:參考個體

305C:個體

305D:個體

310:推算模型

315:對象；個體

315A:個體

315B:個體

315C:個體

315D:個體

320:風險模型

350:對象

400:電腦系統；電腦

402:處理器

404:晶片組

406:記憶體

408:儲存裝置

412:圖形配接器

414:輸入裝置；輸入介面

416:網路配接器

418:顯示器

420:記憶體控制器中樞

422:輸入/輸出(I/O)控制器中樞

【生物材料寄存】

無

【發明申請專利範圍】

【請求項1】 一種判定一對象的轉移性癌症之一對象水準風險之方法，該方法包含：

獲得從該對象擷取的一或多個影像，該一或多個影像包含該對象之複數個淋巴結；及

藉由應用一風險模型至所獲得之該一或多個影像之所提取特徵來預測轉移性癌症之該對象水準風險，該風險模型經訓練以預測在影像中淋巴結的轉移性癌症之淋巴結(LN)水準風險；

其中針對一參考個體之一組淋巴結，至少使用導出自轉移性癌症之經推算LN水準風險的標籤來訓練該風險模型，該組淋巴結之該經推算LN水準風險係使用辨別訓練影像之癌性與非癌性淋巴結的至少一推算模型來判定。

【請求項2】 如請求項1所述之方法，其中預測轉移性癌症之該對象水準風險進一步包含：

基於由該風險模型所預測的轉移性癌症之該複數個淋巴結之LN水準風險來選擇該複數個淋巴結中之一或多者；及

使用針對該一或多個淋巴結所預測的轉移性癌症之該LN水準風險來判定轉移性癌症之該對象水準風險。

【請求項3】 如請求項2所述之方法，其中選擇該複數個淋巴結中之一或多者包含識別具有最高機率之LN水準風險的該淋巴結，且其中判定轉移性癌症之該對象水準風險包含指派該最高機率之LN水準風險作為轉移性癌症之該對象水準風險。

【請求項4】 如請求項1至3中任一項所述之方法，其中該風險模型預測轉移性癌症之LN水準風險的機率範圍大於由該推算模型所預測的該經推算LN水準風險。

【請求項5】 如請求項1至4中任一項所述之方法，其中一參考個體之該組淋巴結係藉由下列來選擇：

判定一中位數風險值；及

比較該參考個體之該等淋巴結的經推算LN水準風險與該中位數風險值。

【請求項6】 如請求項5所述之方法，其中一參考個體之該組淋巴結係進一步藉由下列來選擇：

在該組淋巴結中包括具有大於該中位數風險值之LN水準風險的一或多個淋巴結。

【請求項7】 如請求項5或6所述之方法，其中從該組淋巴結中排除在該組淋巴結中具有小於該中位數風險值之LN水準風險的一或多個淋巴結。

【請求項8】 如請求項5至7中任一項所述之方法，其中該中位數風險值係該參考個體的一中位數對象內LN水準風險。

【請求項9】 如請求項1至8中任一項所述之方法，其中該一或多個影像包含一電腦斷層攝影(CT)影像。

【請求項10】 如請求項1至9中任一項所述之方法，其中該一或多個影像獲自一胸腔CT掃描。

【請求項11】 如請求項1至10中任一項所述之方法，其進一步包含根據轉移性癌症之該預測對象水準風險判定該癌症之一期別。

【請求項12】 如請求項1至11中任一項所述之方法，其進一步包含根據轉移性癌症之該預測對象水準風險來選擇該對象之一診斷或治療。

【請求項13】 如請求項12之所述方法，其中選擇該對象之一診斷或治療包含：當轉移性癌症之該預測對象水準風險低於一臨限風險值時，選擇手術腫瘤切除或合併的支氣管鏡或內視鏡診斷及治療。

【請求項14】 如請求項12所述之方法，其中選擇該對象之一診斷或治療包含：當轉移性癌症之該預測對象水準風險高於一臨限風險值時，從執行一PET或PET-CT掃描或執行淋巴結活體組織切片中之一者或兩者選擇額外的診斷檢測。

【請求項15】 如請求項12所述之方法，其中選擇該對象之一診斷或治療包含：當轉移性癌症之該預測對象水準風險低於一臨限風險值時，選擇一PET或PET-CT掃描的額外診斷檢測。

【請求項16】 如請求項1至15中任一項所述之方法，其中該推算模型係使用獲自胸腔CT掃描之訓練影像來訓練。

【請求項17】 如請求項16所述之方法，其中獲自胸腔CT掃描之該等訓練影像包含一或多個淋巴結。

【請求項18】 如請求項1至17中任一項所述之方法，其中該推算模型係使用國家肺部篩選試驗(NLST)之訓練影像來訓練。

【請求項19】 如請求項1至17中任一項所述之方法，其中該推算模型係使用一自訂資料集之訓練影像來訓練。

【請求項20】 如請求項16至19中任一項所述之方法，其中該推算模型係藉由從該等訓練影像之至少一者產生二或更多個影像圖譜板來訓練。

【請求項21】 如請求項20所述之方法，其中該二或更多個影像圖譜板包含該訓練影像之一高斯拉普拉斯變換及該訓練影像之一小波變換。

【請求項22】 如請求項20或21所述之方法，其中該推算模型進一步藉由下列訓練：從該二或更多個影像圖譜板之各者提取特徵；及至少使用該等所提取特徵來訓練該推算模型。

【請求項23】 如請求項20或21所述之方法，其中該推算模型進一步藉由下列訓練：

在該等影像圖譜板中界定一關注區域(ROI)，所界定之該關注區域包含一淋巴結；

從該等影像圖譜板中之該ROI提取特徵；及

至少使用該等所提取特徵來訓練該推算模型。

【請求項24】 如請求項22或23所述之方法，其中該等所提取特徵包括一階特徵、形狀特徵、灰階共生矩陣1 (GLCM)特徵、灰階級長矩陣(GLRLM)特徵、灰階級帶矩陣(GLSZM)特徵、及鄰域灰階差分矩陣(NGTDM)特徵中之一或多個特徵類別。

【請求項25】 如請求項16至24之任一項所述之方法，其中該推算模型係進一步使用指示訓練影像中轉移的存在或不存在的參考基準真相值來訓練。

【請求項26】 如請求項25所述之方法，其中該等參考基準真相值係導出自指示該癌症之該期別的淋巴結轉移期數共變異數值，其中「0」或「1」之一癌症期別指示轉移的不存在，且其中「2」或「3」之一癌症期別指示轉移的存在。

【請求項27】 如請求項26所述之方法，其中該等參考基準真相值指示個體淋巴結為癌性或非癌性。

【請求項28】 如請求項27所述之方法，其中指示個體淋巴結為癌性或非癌性的該等參考基準真相值係基於該腫瘤或該等淋巴結之特性而從該等淋巴結轉移期數共變異數值來反算。

【請求項29】 如請求項1至28中任一項所述之方法，其中該推算模型係一隨機森林分類器或一LASSO分類器。

【請求項30】 如請求項1至29中任一項所述之方法，其中該風險模型係一隨機森林分類器或一LASSO分類器。

【請求項31】 如請求項30所述之方法，其中該風險模型係使用監督學習技術來訓練。

【請求項32】 如請求項1至31中任一項所述之方法，其中該複數個淋巴結包含2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、或21個淋巴結。

【請求項33】 如請求項1至31中任一項所述之方法，其中該複數個淋巴結包含8至20個之間的淋巴結、10至19個之間的淋巴結、12至18個之間的淋巴結、或14與16個之間的淋巴結。

【請求項34】 如請求項1至33中任一項所述之方法，其中該轉移性癌症係轉移性肺癌。

【請求項35】 如請求項1至34中任一項所述之方法，其中該風險模型展現至少AUC = 78%、靈敏度= 82%、或特異度= 60%之效能。

【請求項36】 一種訓練如請求項1至35中任一項所述之風險模型之方法，該方法包含：

應用一推算模型以產生用於在訓練影像中之複數個淋巴結的經推算LN水準風險；

選擇一組淋巴結，其中該組中之該等淋巴結具有大於一中位數風險值的LN水準風險；

使用在該組淋巴結中的該等淋巴結之該等LN水準風險作為用於訓練該風險模型的參考基準真相，該風險模型能夠預測LN水準風險的機率範圍大於由該推算模型所預測的該等經推算LN水準風險。

【請求項37】 如請求項36所述之方法，其中該風險模型進一步藉由下列訓練：從一或多個影像圖譜板提取特徵；及至少使用該等所提取特徵來訓練該風險模型。

【請求項38】 如請求項37所述之方法，其中從二或更多個影像圖譜板提取特徵包含：

在該一或多個影像圖譜板中界定一關注區域(ROI)，所界定之該關注區域包含一淋巴結；

從該一或多個影像圖譜板中之該ROI提取特徵。

【請求項39】 如請求項37或38所述之方法，其中該等所提取特徵包括一階特徵、形狀特徵、灰階共生矩陣1 (GLCM)特徵、灰階級長矩陣(GLRLM)特徵、灰階級帶矩陣(GLSZM)特徵、及鄰域灰階差分矩陣(NGTDM)特徵中之一或多個特徵類別。

【請求項40】 一種用於判定一對象的轉移性癌症之一對象水準風險之系統，該系統包含：

一成像裝置，其經組態以擷取該對象的一或多個影像；及

一運算裝置，其經組態以執行下列步驟：

獲得從該對象擷取的一或多個影像，該一或多個影像包含該對象之複數個淋巴結；及

藉由應用一風險模型至所獲得之該一或多個影像之所提取特徵來預測轉移性癌症之該對象水準風險，該風險模型經訓練以預測在影像中淋巴結的轉移性癌症之淋巴結(LN)水準風險，

其中針對一參考個體之一組淋巴結，至少使用導出自轉移性癌症之經推算LN水準風險的標籤來訓練該風險模型，該組淋巴結之該等經推算LN水準風險係使用辨別訓練影像之癌性與非癌性淋巴結的至少一推算模型來判定。

【請求項41】 如請求項40所述之系統，其中預測轉移性癌症之該對象水準風險進一步包含：

基於由該風險模型所預測的轉移性癌症之該複數個淋巴結之LN水準風險來選擇該複數個淋巴結中之一或多者；及

使用針對該一或多個淋巴結所預測的轉移性癌症之該LN水準風險來判定轉移性癌症之該對象水準風險。

【請求項42】 如請求項41所述之系統，其中選擇該複數個淋巴結中之一或多者包含識別具有最高機率之LN水準風險的該淋巴結，且其中判定轉移性癌症

之該對象水準風險包含指派該最高機率之LN水準風險作為轉移性癌症之該對象水準風險。

【請求項43】 如請求項40至42中任一項所述之系統，其中該風險模型預測轉移性癌症之LN水準風險的機率範圍大於由該推算模型所預測的該等經推算LN水準風險。

【請求項44】 如請求項40至43中任一項所述之系統，其中一參考個體之該組淋巴結係藉由下列來選擇：

判定一中位數風險值；及

比較該參考個體之該等淋巴結的經推算LN水準風險與該中位數風險值。

【請求項45】 如請求項44所述之系統，其中一參考個體之該組淋巴結係進一步藉由下列來選擇：

在該組淋巴結中包括具有大於該中位數風險值之LN水準風險的一或多個淋巴結。

【請求項46】 如請求項44或45所述之系統，其中從該組淋巴結中排除在該組淋巴結中具有小於該中位數風險值之LN水準風險的一或多個淋巴結。

【請求項47】 如請求項44至46中任一項所述之系統，其中該中位數風險值係該參考個體的一中位數對象內LN水準風險。

【請求項48】 如請求項40至47中任一項所述之系統，其中該成像裝置係電腦斷層攝影(CT)掃描器、磁共振造影(MRI)掃描器、正子發射斷層攝影(PET)掃描器、X射線掃描器、或超音波成像裝置中之一者。

【請求項49】 如請求項40至48中任一項所述之系統，其中該成像裝置係一CT掃描器，且其中該一或多個影像包含一電腦斷層攝影(CT)影像。

【請求項50】 如請求項40至49中任一項所述之系統，其中該一或多個影像包含一胸腔CT影像。

【請求項51】 如請求項40至50中任一項所述之系統，其中該運算裝置進一步經組態以執行根據轉移性癌症之該預測對象水準風險來判定該癌症之一期別的步驟。

【請求項52】 如請求項40至51中任一項所述之系統，其中該運算裝置進一步經組態以執行根據轉移性癌症之該預測對象水準風險來選擇該對象之一診斷或治療的步驟。

【請求項53】 如請求項52所述之系統，其中選擇該對象之一診斷或治療包含：當轉移性癌症之該預測對象水準風險低於一臨限風險值時，選擇手術腫瘤切除或合併的支氣管鏡或內視鏡診斷及治療。

【請求項54】 如請求項53所述之系統，其中選擇該對象之一診斷或治療包含：當轉移性癌症之該預測對象水準風險高於一臨限風險值時，選擇執行一PET或PET-CT掃描或執行淋巴結活體組織切片中之一者或兩者的額外診斷檢測。

【請求項55】 如請求項53所述之系統，其中選擇該對象之一診斷或治療包含：當轉移性癌症之該預測對象水準風險低於一臨限風險值時，選擇一PET或PET-CT掃描的額外診斷檢測。

【請求項56】 如請求項40至54中任一項所述之系統，其中該推算模型係使用獲自胸腔CT掃描之訓練影像來訓練。

【請求項57】 如請求項56所述之系統，其中獲自胸腔CT掃描之該等訓練影像包含一或多個淋巴結。

【請求項58】 如請求項40至57中任一項所述之系統，其中該推算模型係使用國家肺部篩選試驗(NLST)之訓練影像來訓練。

【請求項59】 如請求項40至57中任一項所述之系統，其中該推算模型係使用一自訂資料集之訓練影像來訓練。

【請求項60】 如請求項56至59中任一項所述之系統，其中該推算模型係藉由從該等訓練影像之至少一者產生二或更多個影像圖譜板來訓練。

【請求項61】 如請求項60所述之系統，其中該二或更多個影像圖譜板包含該訓練影像之一高斯拉普拉斯變換及該訓練影像之一小波變換。

【請求項62】 如請求項60或61所述之系統，其中該推算模型進一步藉由下列訓練：從該二或更多個影像圖譜板之各者提取特徵；及至少使用該等所提取特徵來訓練該推算模型。

【請求項63】 如請求項60或61所述之系統，其中該推算模型進一步藉由下列訓練：

在該等影像圖譜板中界定一關注區域(ROI)，所界定之該關注區域包含一淋巴結；

從該等影像圖譜板中之該ROI提取特徵；及

至少使用該等所提取特徵來訓練該推算模型。

【請求項64】 如請求項62或63所述之系統，其中該等所提取特徵包括一階特徵、形狀特徵、灰階共生矩陣1 (GLCM)特徵、灰階級長矩陣(GLRLM)特徵、灰階級帶矩陣(GLSZM)特徵、及鄰域灰階差分矩陣(NGTDM)特徵中之一或多個特徵類別。

【請求項65】 如請求項56至64中任一項所述之系統，其中該推算模型係進一步使用指示訓練影像中轉移的存在或不存在的參考基準真相值來訓練。

【請求項66】 如請求項65所述之系統，其中該等參考基準真相值係導出自指示該癌症之該期別的淋巴結轉移期數共變異數值，其中「0」或「1」之一癌症期別指示轉移的不存在，且其中「2」或「3」之一癌症期別指示轉移的存在。

【請求項67】 如請求項66所述之系統，其中該等參考基準真相值指示個體淋巴結為癌性或非癌性。

【請求項68】 如請求項67所述之系統，其中指示個體淋巴結為癌性或非癌性的該等參考基準真相值係基於該腫瘤或該等淋巴結之特性而從該等淋巴結轉移期數共變異數值來反算。

【請求項69】 如請求項40至68中任一項所述之系統，其中該推算模型係一隨機森林分類器或一LASSO分類器。

【請求項70】 如請求項40至69中任一項所述之系統，其中該風險模型係一隨機森林分類器或一LASSO分類器。

【請求項71】 如請求項70所述之系統，其中該風險模型係使用監督學習技術來訓練。

【請求項72】 如請求項40至71中任一項所述之系統，其中該複數個淋巴結包含2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、或21個淋巴結。

【請求項73】 如請求項40至72中任一項所述之系統，其中該複數個淋巴結包含8至20個之間的淋巴結、10至19個之間的淋巴結、12至18個之間的淋巴結、或14與16個之間的淋巴結。

【請求項74】 如請求項40至73中任一項所述之系統，其中該轉移性癌症係轉移性肺癌。

【請求項75】 如請求項40至74中任一項所述之系統，其中該風險模型展現至少AUC = 78%、靈敏度= 82%、或特異度= 60%之效能。

【請求項76】 如請求項40至75中任一項所述之系統，其中該風險模型藉由下列訓練：

應用一推算模型以產生用於在訓練影像中之複數個淋巴結的經推算LN水準風險；

選擇一組淋巴結，其中該組中之該等淋巴結具有大於一中位數風險值的LN水準風險；

使用在該組淋巴結中的該等淋巴結之該等LN水準風險作為用於訓練該風險模型的參考基準真相值，該風險模型能夠預測LN水準風險的機率範圍大於由該推算模型所預測的該等經推算LN水準風險。

【請求項77】 如請求項76所述之系統，其中該風險模型進一步藉由下列訓練：從一或多個影像圖譜板提取特徵；及至少使用該等所提取特徵來訓練該風險模型。

【請求項78】 如請求項77所述之系統，其中從二或更多個影像圖譜板提取特徵包含：

在該一或多個影像圖譜板中界定一關注區域(ROI)，所界定之該關注區域包含一淋巴結；

從該一或多個影像圖譜板中之該ROI提取特徵。

【請求項79】 如請求項77或78所述之系統，其中該等所提取特徵包括一階特徵、形狀特徵、灰階共生矩陣1 (GLCM)特徵、灰階級長矩陣(GLRLM)特徵、灰階級帶矩陣(GLSZM)特徵、及鄰域灰階差分矩陣(NGTDM)特徵中之一或多個特徵類別。

【請求項80】 一種非暫態電腦可讀取媒體，其包含當由一處理器執行時造成該處理器進行下列之指令：

獲得從該對象擷取的一或多個影像，該一或多個影像包含該對象之複數個淋巴結；及

藉由應用一風險模型至所獲得之該一或多個影像之所提取特徵來預測轉移性癌症之該對象水準風險，該風險模型經訓練以預測在影像中淋巴結的轉移性癌症之淋巴結(LN)水準風險，

其中針對一參考個體之一組淋巴結，至少使用導出自轉移性癌症之經推算LN水準風險的標籤來訓練該風險模型，該組淋巴結之該等經推算LN水準風險係使用辨別訓練影像之癌性與非癌性淋巴結的至少一推算模型來判定。

【請求項81】 如請求項80所述之非暫態電腦可讀取媒體，其中造成該處理器預測轉移性癌症之該對象水準風險的該等指令進一步包含當由該處理器執行時造成該處理器進行下列之指令：

基於由該風險模型所預測的轉移性癌症之該複數個淋巴結之LN水準風險來選擇該複數個淋巴結中之一或多者；及

第 13 頁，共 19 頁(發明申請專利範圍)

使用針對該一或多個淋巴結所預測的轉移性癌症之該LN水準風險來判定轉移性癌症之該對象水準風險。

【請求項82】 如請求項81所述之非暫態電腦可讀取媒體，其中造成該處理器選擇該複數個淋巴結中之一或多者的該等指令進一步包含當由該處理器執行時造成該處理器識別具有最高機率之LN水準風險的該淋巴結的指令，且其中造成該處理器判定轉移性癌症之該對象水準風險之該等指令包含當由該處理器執行時造成該處理器指派該最高機率之LN水準風險作為轉移性癌症之該對象水準風險的指令。

【請求項83】 如請求項80至82中任一項所述之非暫態電腦可讀取媒體，其中該風險模型預測轉移性癌症之LN水準風險的機率範圍大於由該推算模型所預測的該等經推算LN水準風險。

【請求項84】 如請求項80至83中任一項所述之非暫態電腦可讀取媒體，其中一參考個體之該組淋巴結係藉由下列來選擇：

判定一中位數風險值；及

比較該參考個體之該等淋巴結的經推算LN水準風險與該中位數風險值。

【請求項85】 如請求項84所述之非暫態電腦可讀取媒體，其中一參考個體之該組淋巴結進一步藉由以下來選擇：在該組淋巴結中包括具有大於該中位數風險值之LN水準風險的一或多個淋巴結。

【請求項86】 如請求項84或85所述之非暫態電腦可讀取媒體，其中從該組淋巴結中排除在該組淋巴結中具有小於該中位數風險值之LN水準風險的一或多個淋巴結。

【請求項87】 如請求項84至86中任一項所述之非暫態電腦可讀取媒體，其中該中位數風險值係該參考個體的一中位數對象內LN水準風險。

【請求項88】 如請求項80至87中任一項所述之非暫態電腦可讀取媒體，其中該一或多個影像包含一電腦斷層攝影(CT)影像。

【請求項89】 如請求項80至88中任一項所述之非暫態電腦可讀取媒體，其中該一或多個影像獲自一胸腔CT掃描。

【請求項90】 如請求項80至89中任一項所述之非暫態電腦可讀取媒體，其進一步包含當由該處理器執行時造成該處理器根據轉移性癌症之該預測對象水準風險來判定該癌症之一期別的指令。

【請求項91】 如請求項80至90中任一項所述之非暫態電腦可讀取媒體，其進一步包含當由該處理器執行時造成該處理器根據轉移性癌症之該預測對象水準風險來選擇該對象之一診斷或治療的指令。

【請求項92】 如請求項91所述之非暫態電腦可讀取媒體，其中造成該處理器選擇該對象之一診斷或治療的該等指令進一步包含當由該處理器執行時造成該處理器進行下列之指令：當轉移性癌症之該預測對象水準風險低於一臨限風險值時，選擇手術腫瘤切除或合併的支氣管鏡或內視鏡診斷及治療。

【請求項93】 如請求項91所述之非暫態電腦可讀取媒體，其中造成該處理器選擇該對象之一診斷或治療的該等指令進一步包含當由該處理器執行時造成該處理器進行下列之指令：當轉移性癌症之該預測對象水準風險高於一臨限風險值時，從執行一PET或PET-CT掃描或執行淋巴結活體組織切片中之一者或兩者選擇額外的診斷檢測。

【請求項94】 如請求項91所述之非暫態電腦可讀取媒體，其中造成該處理器選擇該對象之一診斷或治療的該等指令進一步包含當由該處理器執行時造成該處理器進行下列之指令：當轉移性癌症的該預測對象水準風險低於一臨限風險值時，選擇一PET或PET-CT掃描的額外診斷檢測。

【請求項95】 如請求項80至94中任一項所述之非暫態電腦可讀取媒體，其中該推算模型係使用獲自胸腔CT掃描之訓練影像來訓練。

【請求項96】 如請求項95所述之非暫態電腦可讀取媒體，其中獲自胸腔CT掃描之該等訓練影像包含一或多個淋巴結。

【請求項97】 如請求項80至96中任一項所述之非暫態電腦可讀取媒體，其中該推算模型係使用國家肺部篩檢試驗(NLST)之訓練影像來訓練。

【請求項98】 如請求項80至96中任一項所述之非暫態電腦可讀取媒體，其中該推算模型係使用一自訂資料集之訓練影像來訓練。

【請求項99】 如請求項95至98中任一項所述之非暫態電腦可讀取媒體，其中該推算模型係藉由從該等訓練影像之至少一者產生二或更多個影像圖譜板來訓練。

【請求項100】 如請求項99所述之非暫態電腦可讀取媒體，其中該二或更多個影像圖譜板包含該訓練影像之一高斯拉普拉斯變換及該訓練影像之一小波變換。

【請求項101】 如請求項99或100所述之非暫態電腦可讀取媒體，其中該推算模型進一步藉由下列訓練：從該二或更多個影像圖譜板之各者提取特徵；及至少使用該等所提取特徵來訓練該推算模型。

【請求項102】 如請求項99或100所述之非暫態電腦可讀取媒體，其中該推算模型進一步藉由下列訓練：

在該等影像圖譜板中界定一關注區域(ROI)，所界定之該關注區域包含一淋巴結；

從該等影像圖譜板中之該ROI提取特徵；及

至少使用該等所提取特徵來訓練該推算模型。

【請求項103】 如請求項101或102所述之非暫態電腦可讀取媒體，其中該等所提取特徵包括一階特徵、形狀特徵、灰階共生矩陣1 (GLCM)特徵、灰階級長矩陣(GLRLM)特徵、灰階級帶矩陣(GLSZM)特徵、及鄰域灰階差分矩陣(NGTDM)特徵中之一或多個特徵類別。

【請求項104】 如請求項95至103所述之非暫態電腦可讀取媒體，其中該推算模型係進一步使用指示訓練影像中轉移的存在或不存在的參考基準真相值來訓練。

【請求項105】 如請求項104所述之非暫態電腦可讀取媒體，其中該等參考基準真相值係導出自指示該癌症之該期別的淋巴結轉移期數共變異數值，其中「0」或「1」之一癌症期別指示轉移的不存在，且其中「2」或「3」之一癌症期別指示轉移的存在。

【請求項106】 如請求項105所述之非暫態電腦可讀取媒體，其中該等參考基準真相值指示個體淋巴結為癌性或非癌性。

【請求項107】 如請求項106所述之非暫態電腦可讀取媒體，其中指示個體淋巴結為癌性或非癌性的該等參考基準真相值係基於該腫瘤或該等淋巴結之特性而從該等淋巴結轉移期數共變異數值來反算。

【請求項108】 如請求項80至107中任一項所述之非暫態電腦可讀取媒體，其中該推算模型係一隨機森林分類器或一LASSO分類器。

【請求項109】 如請求項80至108中任一項所述之非暫態電腦可讀取媒體，其中該風險模型係一隨機森林分類器或一LASSO分類器。

【請求項110】 如請求項109所述之非暫態電腦可讀取媒體，其中該風險模型係使用監督學習技術來訓練。

【請求項111】 如請求項80至110中任一項所述之非暫態電腦可讀取媒體，其中該複數個淋巴結包含2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、或21個淋巴結。

【請求項112】 如請求項80至111中任一項所述之非暫態電腦可讀取媒體，其中該複數個淋巴結包含8至20個之間的淋巴結、10至19個之間的淋巴結、12至18個之間的淋巴結、或14與16個之間的淋巴結。

【請求項113】 如請求項80至112中任一項所述之非暫態電腦可讀取媒體，其中該轉移性癌症係轉移性肺癌。

【請求項114】 如請求項80至113中任一項所述之非暫態電腦可讀取媒體，其中該風險模型展現至少AUC = 78%、靈敏度= 82%、或特異度= 60%之效能。

【請求項115】 如請求項80至114中任一項所述之非暫態電腦可讀取媒體，其中該風險模型藉由下列訓練：

應用一推算模型以產生用於在訓練影像中之複數個淋巴結的經推算LN水準風險；

選擇一組淋巴結，其中該組中之該等淋巴結具有大於一中位數風險值的LN水準風險；

使用在該組淋巴結中的該等淋巴結之該等LN水準風險作為用於訓練該風險模型的參考基準真相，該風險模型能夠預測LN水準風險的機率範圍大於由該推算模型所預測的該等經推算LN水準風險。

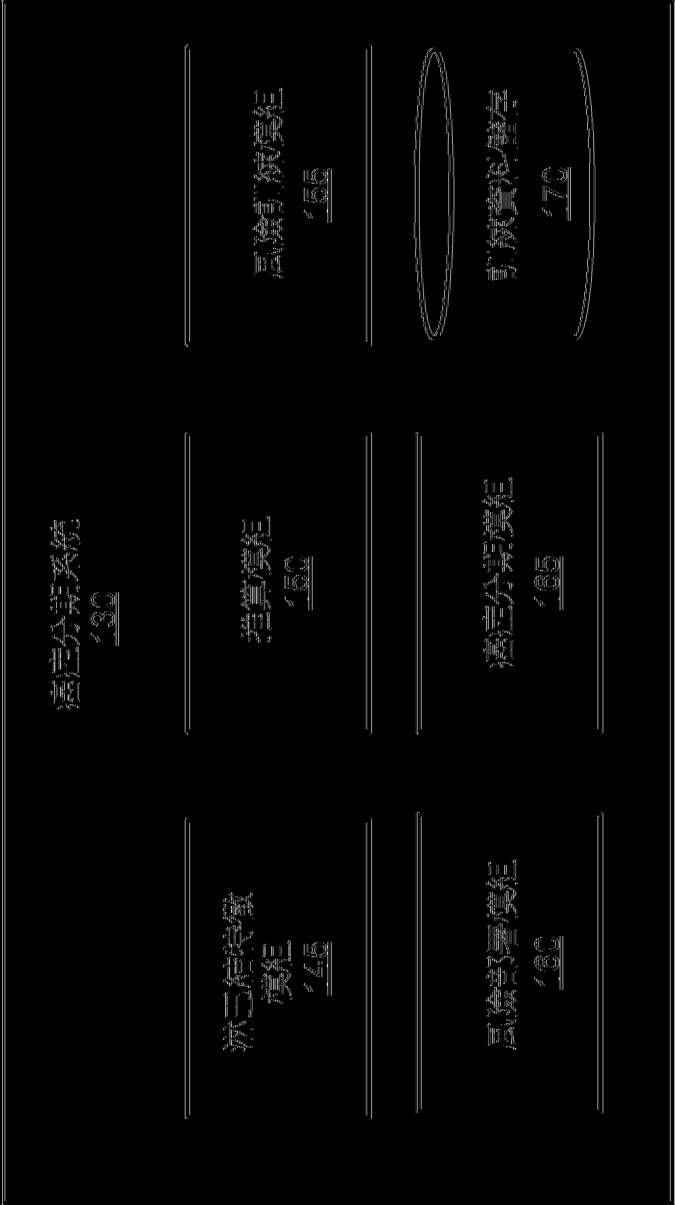
【請求項116】 如請求項115所述之非暫態電腦可讀取媒體，其中該風險模型進一步藉由下列訓練：從一或多個影像圖譜板提取特徵；及至少使用該等所提取特徵來訓練該風險模型。

【請求項117】 如請求項116所述之非暫態電腦可讀取媒體，其中從二或更多個影像圖譜板提取特徵包含：

在該一或多個影像圖譜板中界定一關注區域(ROI)，所界定之該關注區域包含一淋巴結；

從該一或多個影像圖譜板中之該ROI提取特徵。

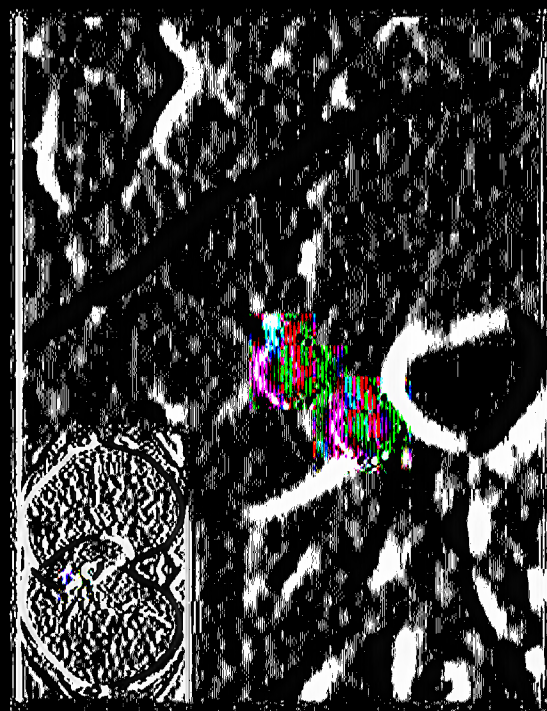
【請求項118】 如請求項115或116所述之非暫態電腦可讀取媒體，其中該等所提取特徵包括一階特徵、形狀特徵、灰階共生矩陣1 (GLCM)特徵、灰階級長矩陣(GLRLM)特徵、灰階級帶矩陣(GLSZM)特徵、及鄰域灰階差分矩陣(NGTDM)特徵中之一或多個特徵類別。



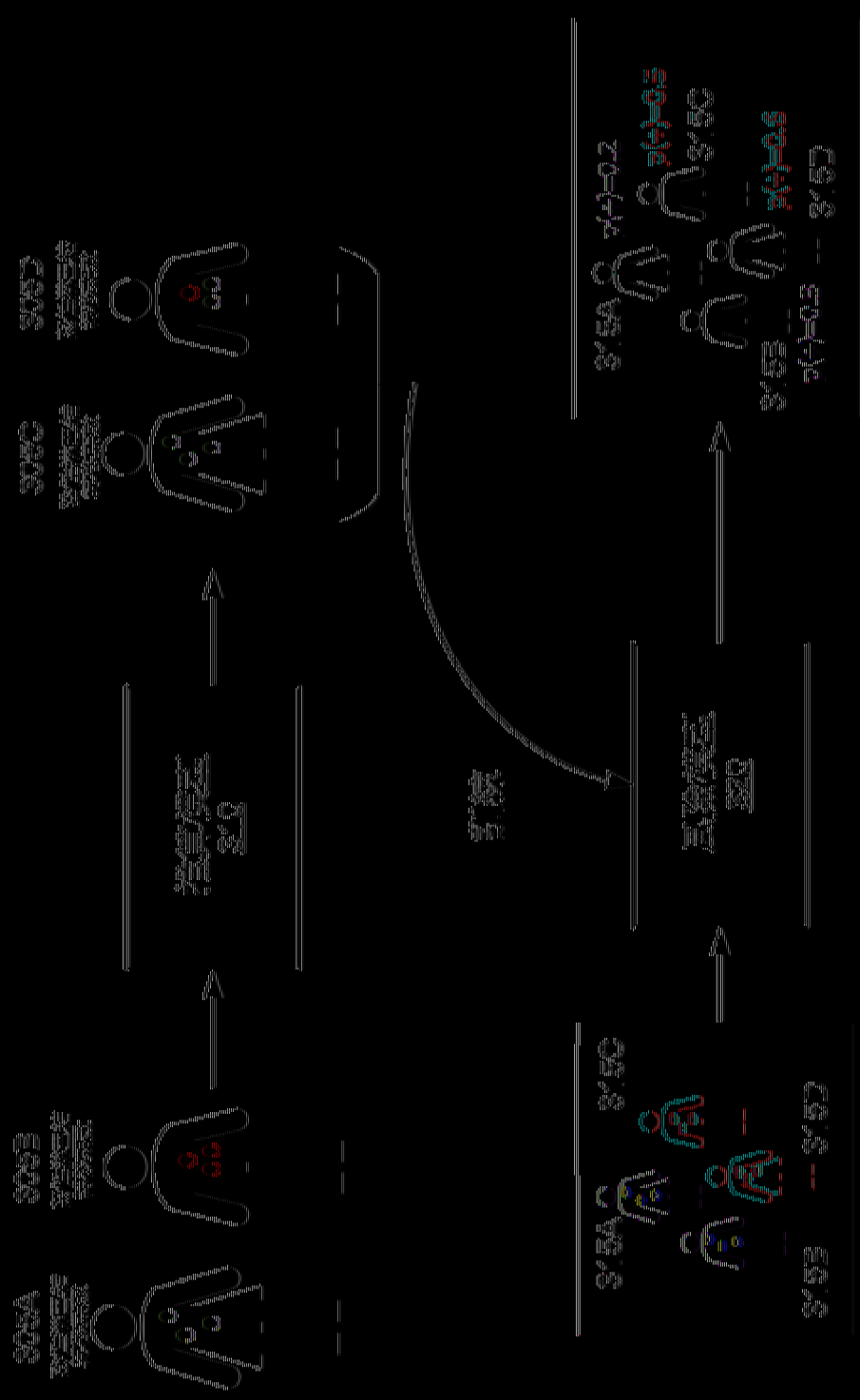
[(Fig)113]



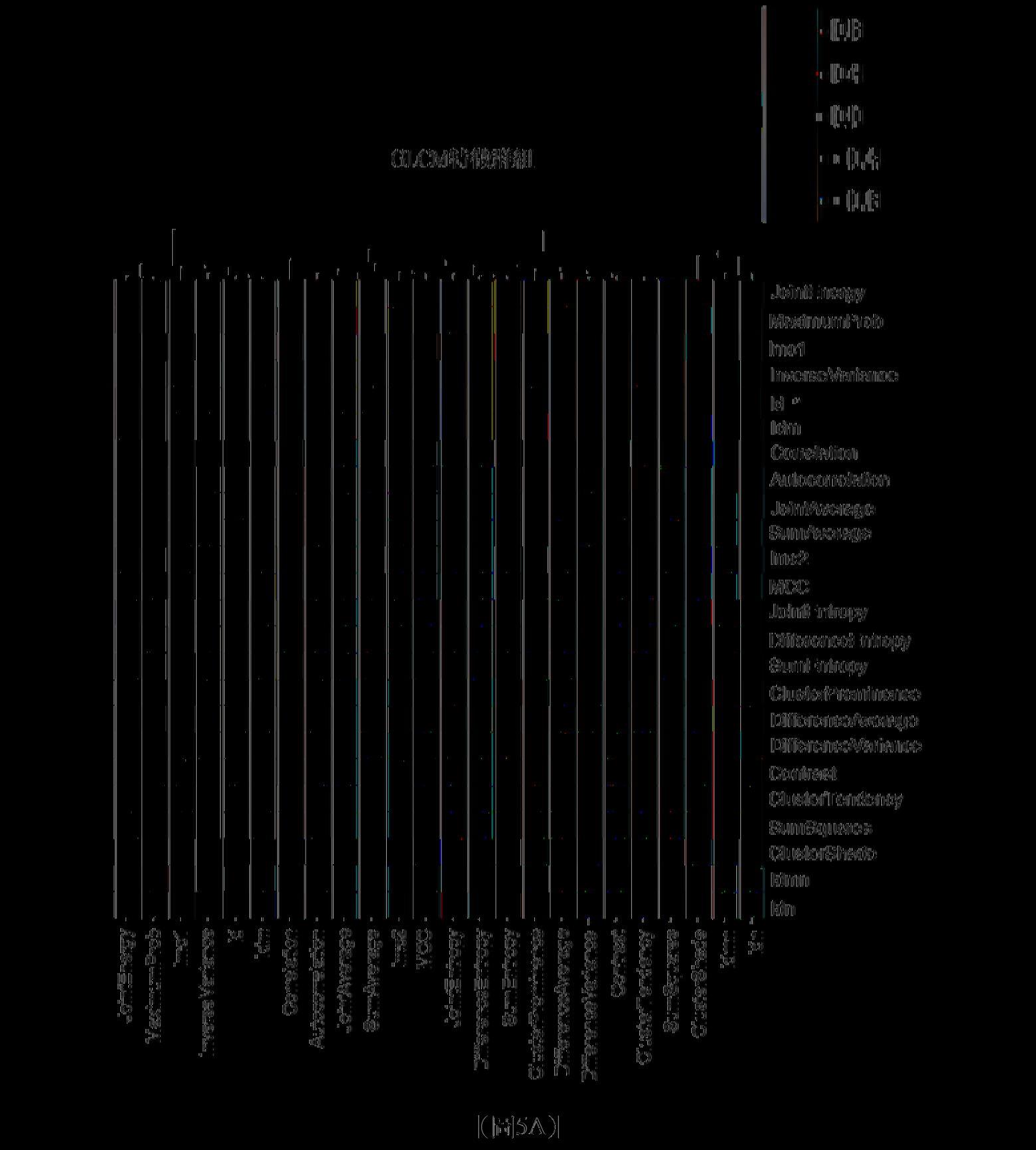
[(圖)2A]

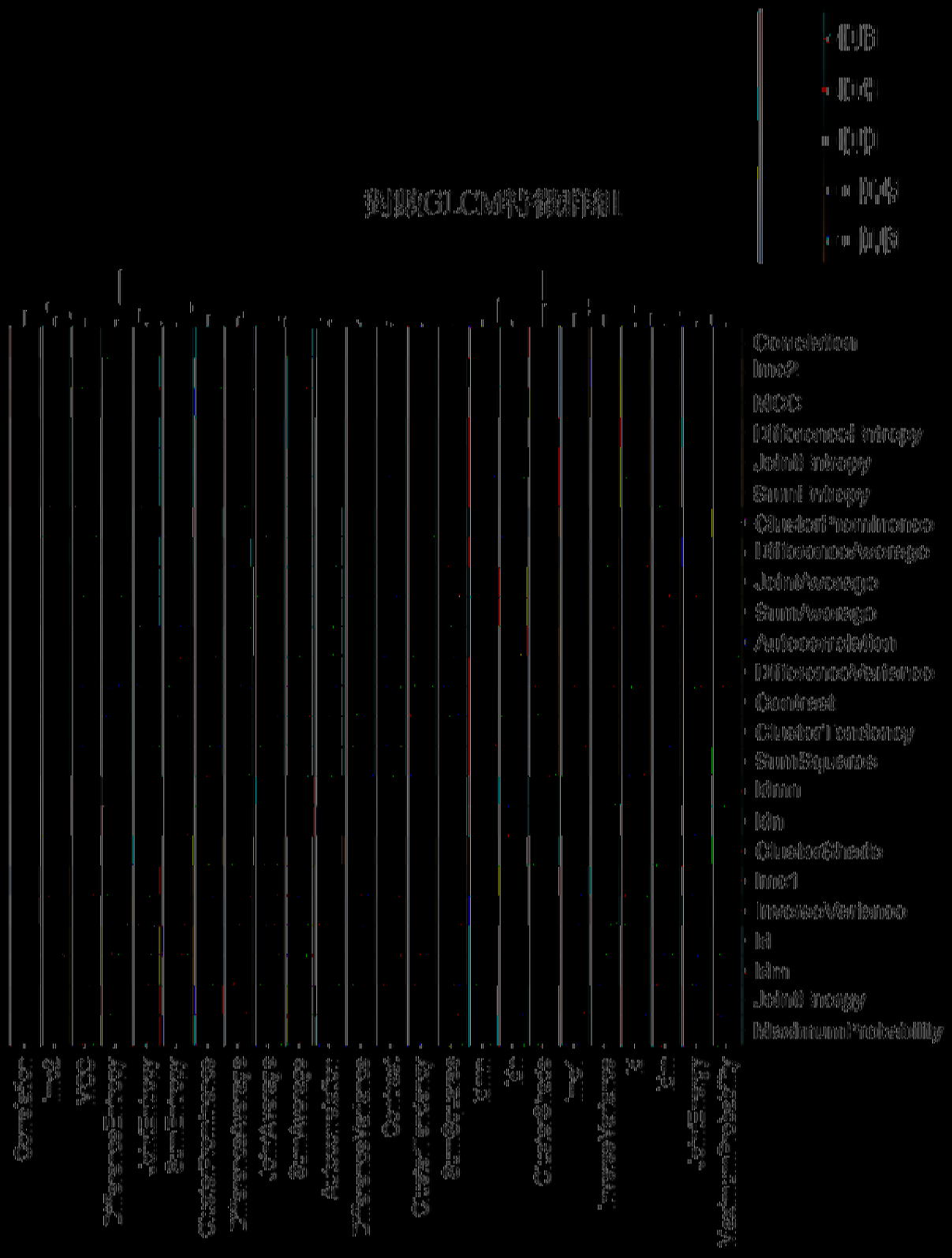


[(圖)2B]



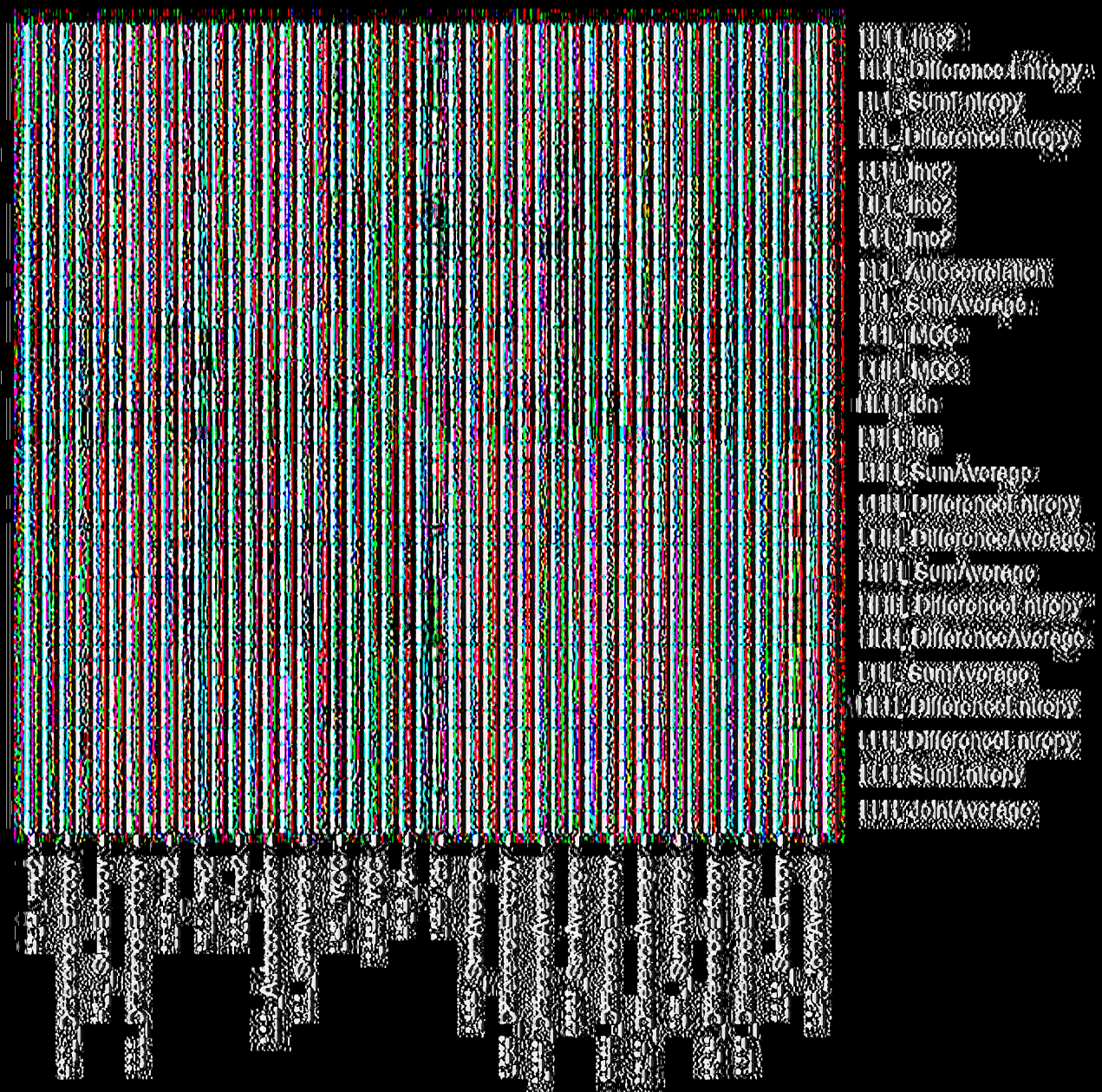
[圖3A]





(表53)

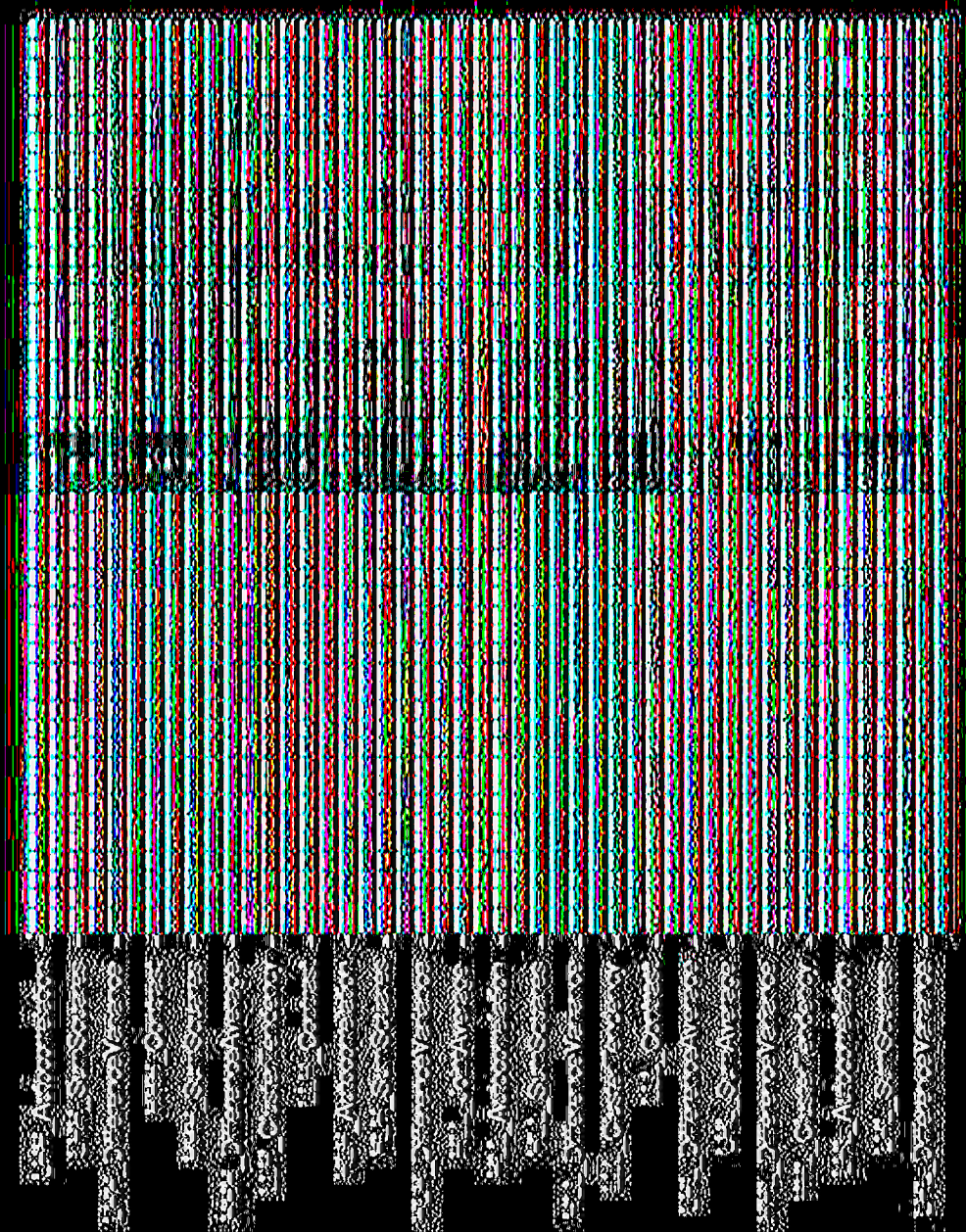
小波GLCM特徵群組1



[(圖)5C]

小波包分解係數

0.8
0.4
0.0
-0.4
-0.8

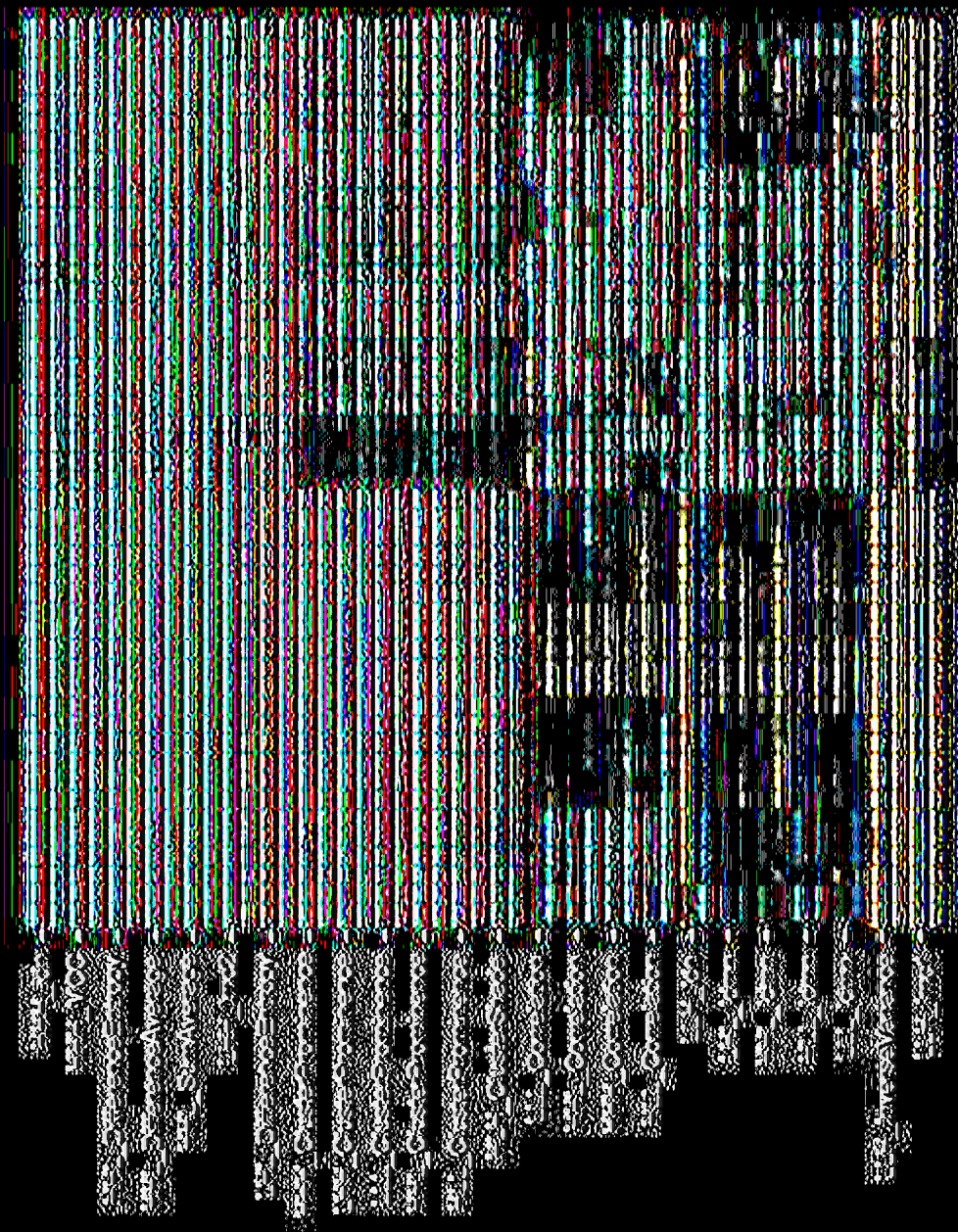


1111 Ime2
1111 differenceEntropy
1111 SumEntropy
1111 differenceEntropy
1111 Ime2
1111 Ime2
1111 Ime2
1111 Autocorrelation
1111 SumAverage
1111 MCC
1111 MCC
1111 Jm
1111 Jm
1111 SumAverage
1111 differenceEntropy
1111 differenceAverage
1111 SumAverage
1111 differenceEntropy
1111 differenceAverage
1111 SumAverage
1111 differenceEntropy
1111 differenceEntropy
1111 SumEntropy
1111 JmAverage

(1111SD)

小波GLOM特徵子組

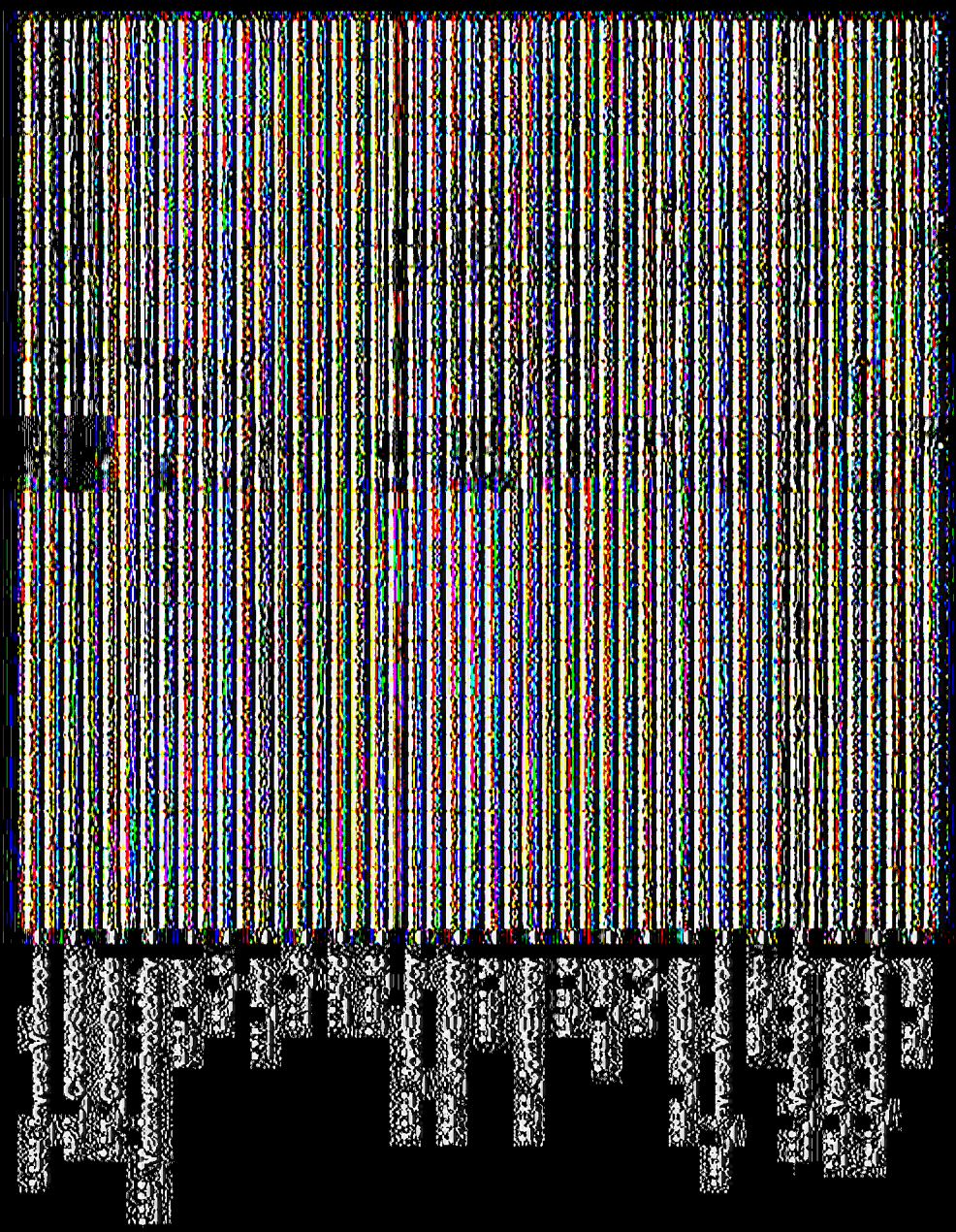
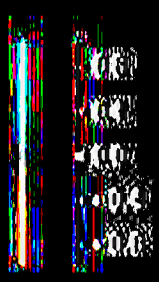
- C6
- C6
- C6
- C4
- C6



- 1111 Imc2
- 1111 DifferenceEntropy
- 1111 SumEntropy
- 1111 DifferenceEntropy
- 1111 Imc2
- 1111 Imc2
- 1111 Imc2
- 1111 AllCorrelation
- 1111 SumAverage
- 1111 MCO
- 1111 MCO
- 1111 Len
- 1111 Len
- 1111 SumAverage
- 1111 DifferenceEntropy
- 1111 DifferenceAverage
- 1111 SumAverage
- 1111 DifferenceEntropy
- 1111 DifferenceAverage
- 1111 SumAverage
- 1111 DifferenceEntropy
- 1111 DifferenceEntropy
- 1111 SumEntropy
- 1111 SumAverage

[(H)S(E)]

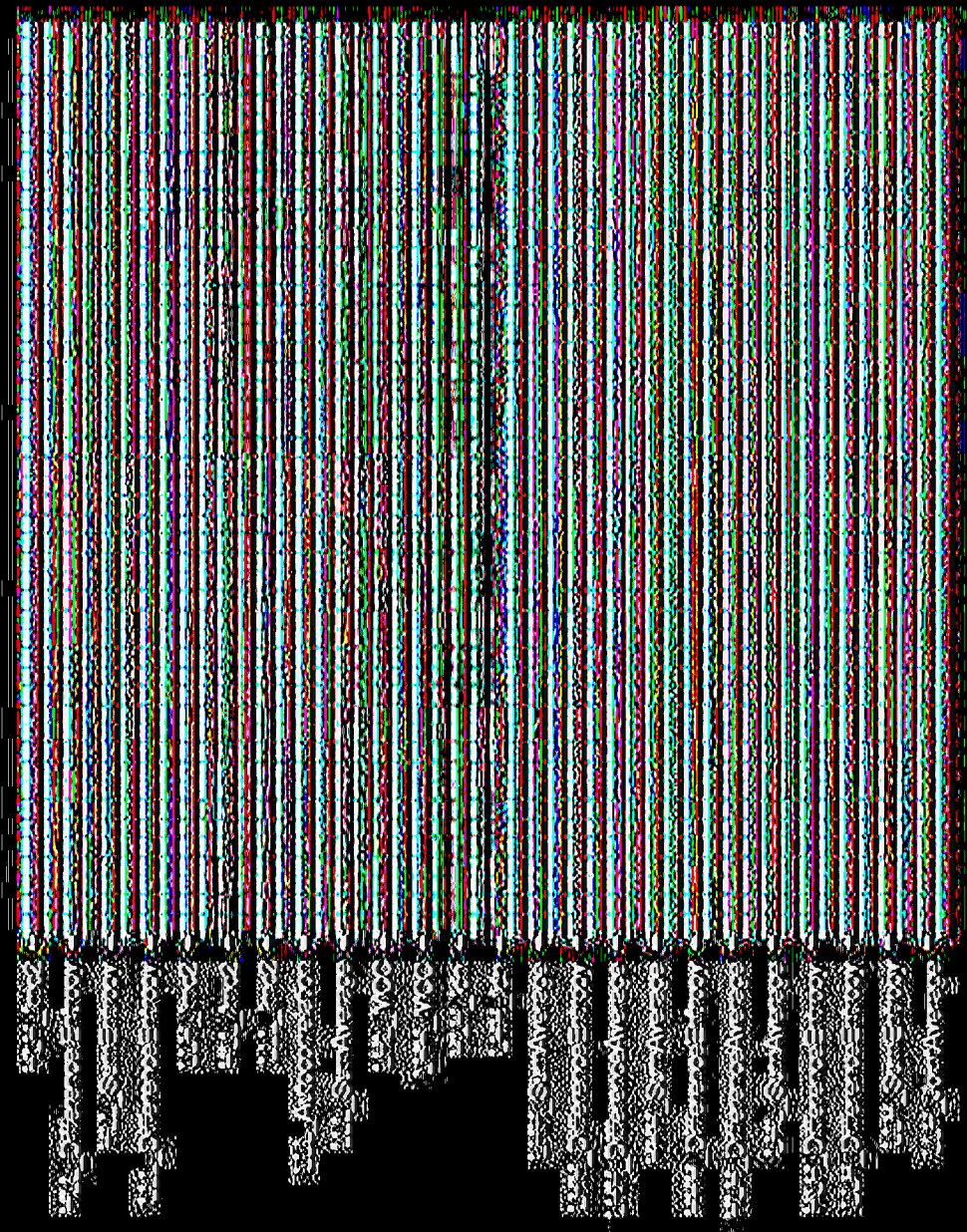
小波CIOM4特徵群組



- 1111:Info2
- 1111:DifferenceEntropy
- 1111:SumEntropy
- 1111:DifferenceEntropy
- 1111:Info2
- 1111:Info2
- 1111:Info2
- 1111:Autocorrelation
- 1111:SumAverage
- 1111:NCC
- 1111:NCC
- 1111:Info
- 1111:Info
- 1111:SumAverage
- 1111:DifferenceEntropy
- 1111:DifferenceAverage
- 1111:SumAverage
- 1111:DifferenceEntropy
- 1111:DifferenceAverage
- 1111:SumAverage
- 1111:DifferenceEntropy
- 1111:DifferenceEntropy
- 1111:SumEntropy
- 1111:JoinAverage

$$|H[S]|$$

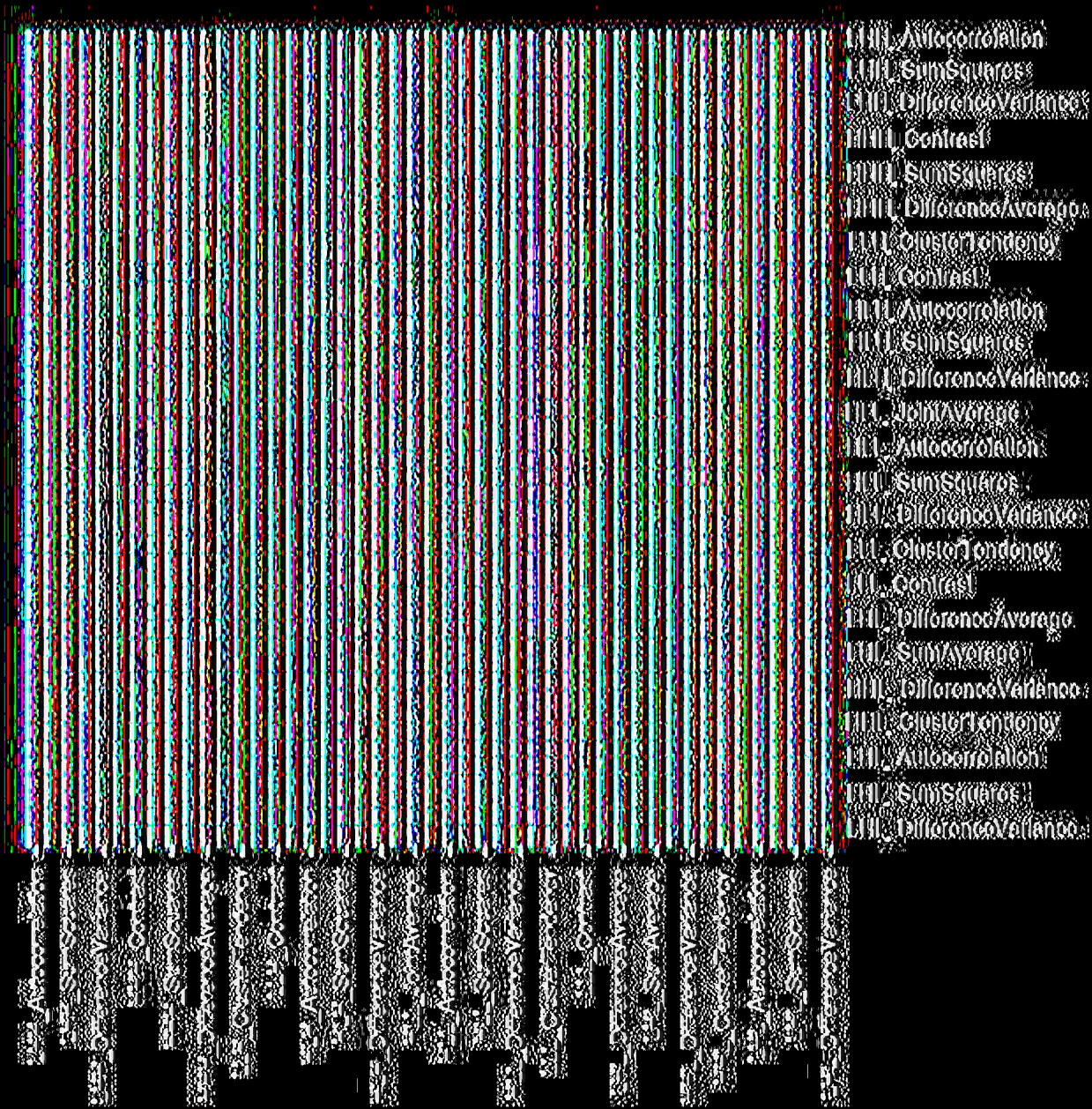
小波GLOM4特徵係數



- 1111 Autocorrelation
- 1111 SumSquares
- 1111 DifferenceVariance
- 1111 Contrast
- 1111 SumSquares
- 1111 DifferenceVariance
- 1111 ClustorTendency
- 1111 Contrast
- 1111 Autocorrelation
- 1111 SumSquares
- 1111 DifferenceVariance
- 1111 Sum/Average
- 1111 Autocorrelation
- 1111 SumSquares
- 1111 DifferenceVariance
- 1111 ClustorTendency
- 1111 Contrast
- 1111 DifferenceVariance
- 1111 Sum/Average
- 1111 DifferenceVariance
- 1111 ClustorTendency
- 1111 Autocorrelation
- 1111 SumSquares
- 1111 DifferenceVariance

(1111SG)

小波GLCM特徵非矩組



[FIG. 500]

小波GLCM特徵群組

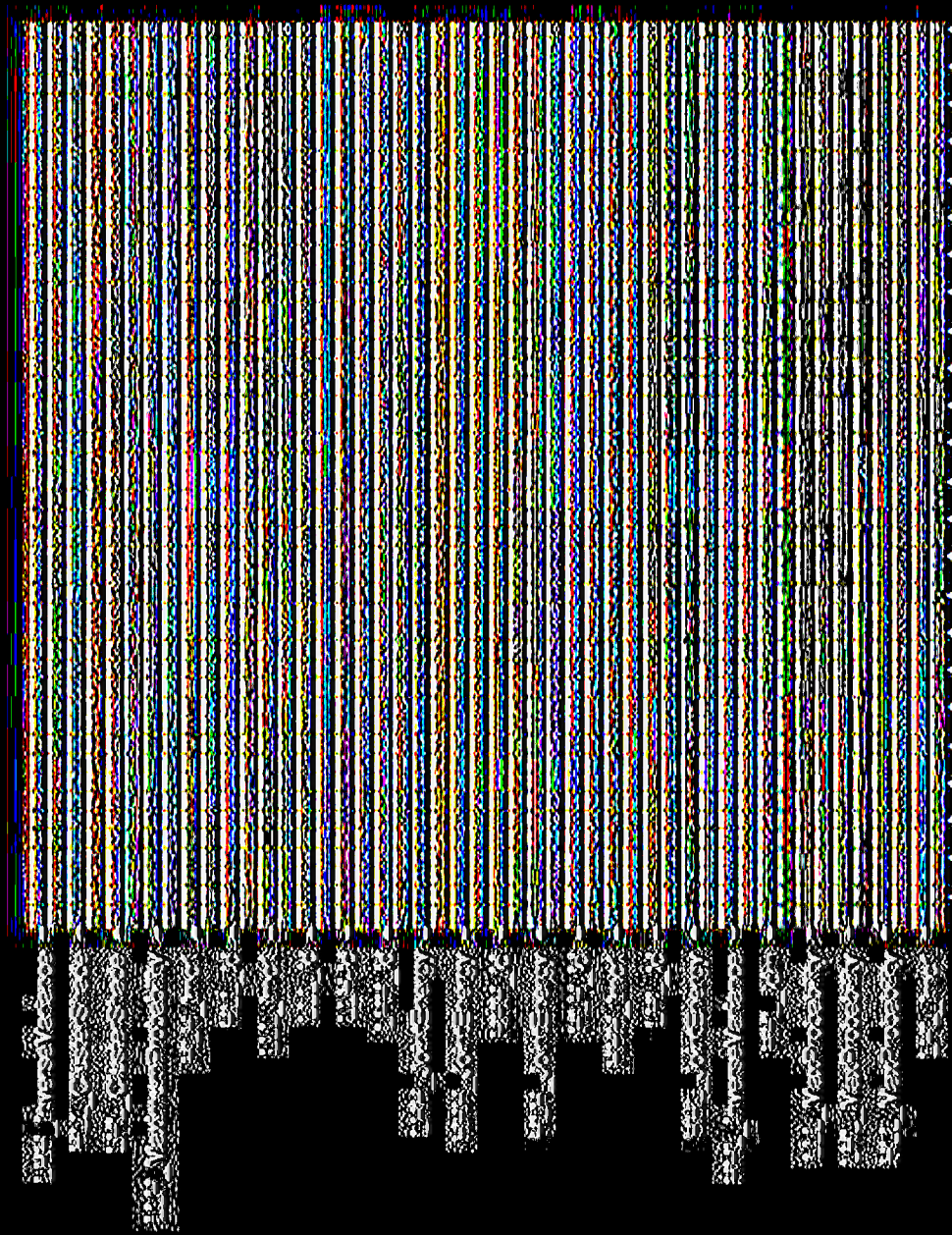
00
04
06
04
06



1111 AutoCorrelation
1111 SumSquares
1111 DifferenceVariance
1111 Contrast
1111 SumSquares
1111 DifferenceAverage
1111 ClusterTendency
1111 Contrast
1111 AutoCorrelation
1111 SumSquares
1111 DifferenceVariance
1111 SumAverage
1111 AutoCorrelation
1111 SumSquares
1111 DifferenceVariance
1111 ClusterTendency
1111 Contrast
1111 DifferenceAverage
1111 SumAverage
1111 DifferenceVariance
1111 ClusterTendency
1111 AutoCorrelation
1111 SumSquares
1111 DifferenceVariance

|(H)S|

小波GLCM特徵群組



- 1111 Autocorrelation
- 1111 SumSquares
- 1111 DifferenceVariance
- 1111 Contrast
- 1111 SumSquares
- 1111 DifferenceAverage
- 1111 ClusterTendency
- 1111 Contrast
- 1111 Autocorrelation
- 1111 SumSquares
- 1111 DifferenceVariance
- 1111 SumAverage
- 1111 Autocorrelation
- 1111 SumSquares
- 1111 DifferenceVariance
- 1111 ClusterTendency
- 1111 Contrast
- 1111 DifferenceAverage
- 1111 SumAverage
- 1111 DifferenceVariance
- 1111 ClusterTendency
- 1111 Autocorrelation
- 1111 SumSquares
- 1111 DifferenceVariance

$|([i][j])|$

- 1000. idn
- 1001. MCO
- 1002. DifferenceofEntropy
- 1003. DifferenceOfAverage
- 1004. SumAverage
- 1005. Info?
- 1006. DifferenceofEntropy
- 1007. ClusterProminence
- 1008. ClusterProminence
- 1009. ClusterProminence
- 1010. ClusterProminence
- 1011. ClusterProminence
- 1012. ClusterProminence
- 1013. ClusterProminence
- 1014. ClusterProminence
- 1015. ClusterProminence
- 1016. Correlation
- 1017. Correlation
- 1018. Correlation
- 1019. Correlation
- 1020. idn
- 1021. idn
- 1022. idn
- 1023. idn
- 1024. idn
- 1025. InverseVariance
- 1026. idn

[illegible]

小波GLCM特徵提取

- 06
- 04
- 00
- +04
- +08



- 1111: Ent
- 1111: MCG
- 1111: DifferentialEntropy
- 1111: DifferentialAverage
- 1111: SumAverage
- 1111: Ino2
- 1111: DifferentialEntropy
- 1111: ClusterProminence
- 1111: ClusterProminence
- 1111: ClusterProminence
- 1111: ClusterProminence
- 1111: ClusterProminence
- 1111: ClusterShade
- 1111: Correlation
- 1111: Correlation
- 1111: Correlation
- 1111: Correlation
- 1111: Ent
- 1111: Ent
- 1111: Ent
- 1111: Ent
- 1111: Ent
- 1111: InverseVariance
- 1111: Ino1

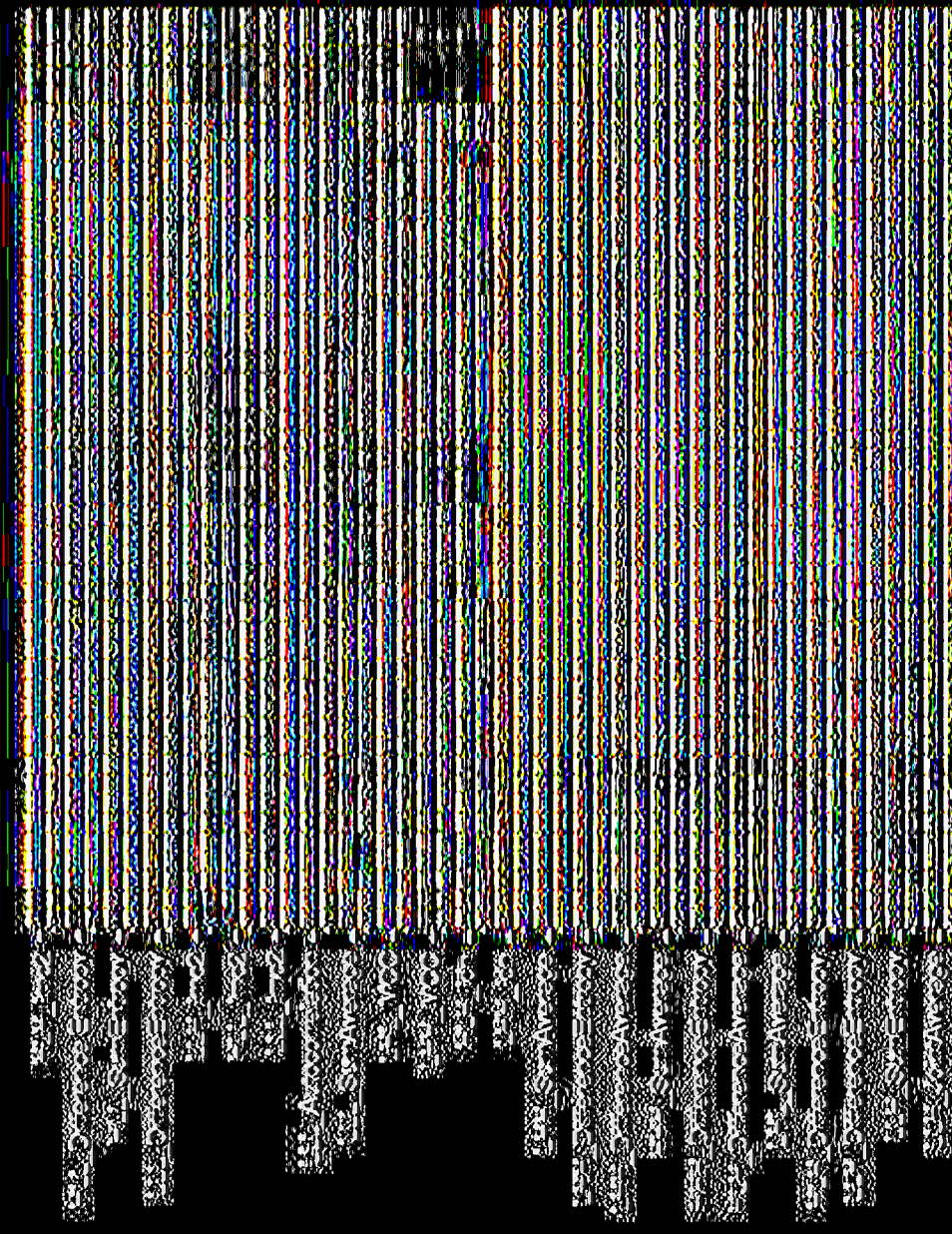
$|H(SM)|$

The image is a high-contrast, black-and-white scan of a textured surface, likely a book cover or endpaper. It features a dense, irregular pattern of black, white, and gray pixels, creating a noisy, almost abstract appearance. A prominent vertical crease or fold line runs down the center of the image, suggesting a hinge or a point of damage. The overall texture is highly irregular and noisy, with many small, dark, irregular shapes scattered throughout, possibly representing dust, wear, or the underlying material's grain. The lighting is uneven, with some areas appearing brighter than others, contributing to the high-contrast effect.

第 20 頁，共 30 頁(發明圖式)

小波GLCM特徵提取

0.8
0.4
0.0
-0.4
-0.8

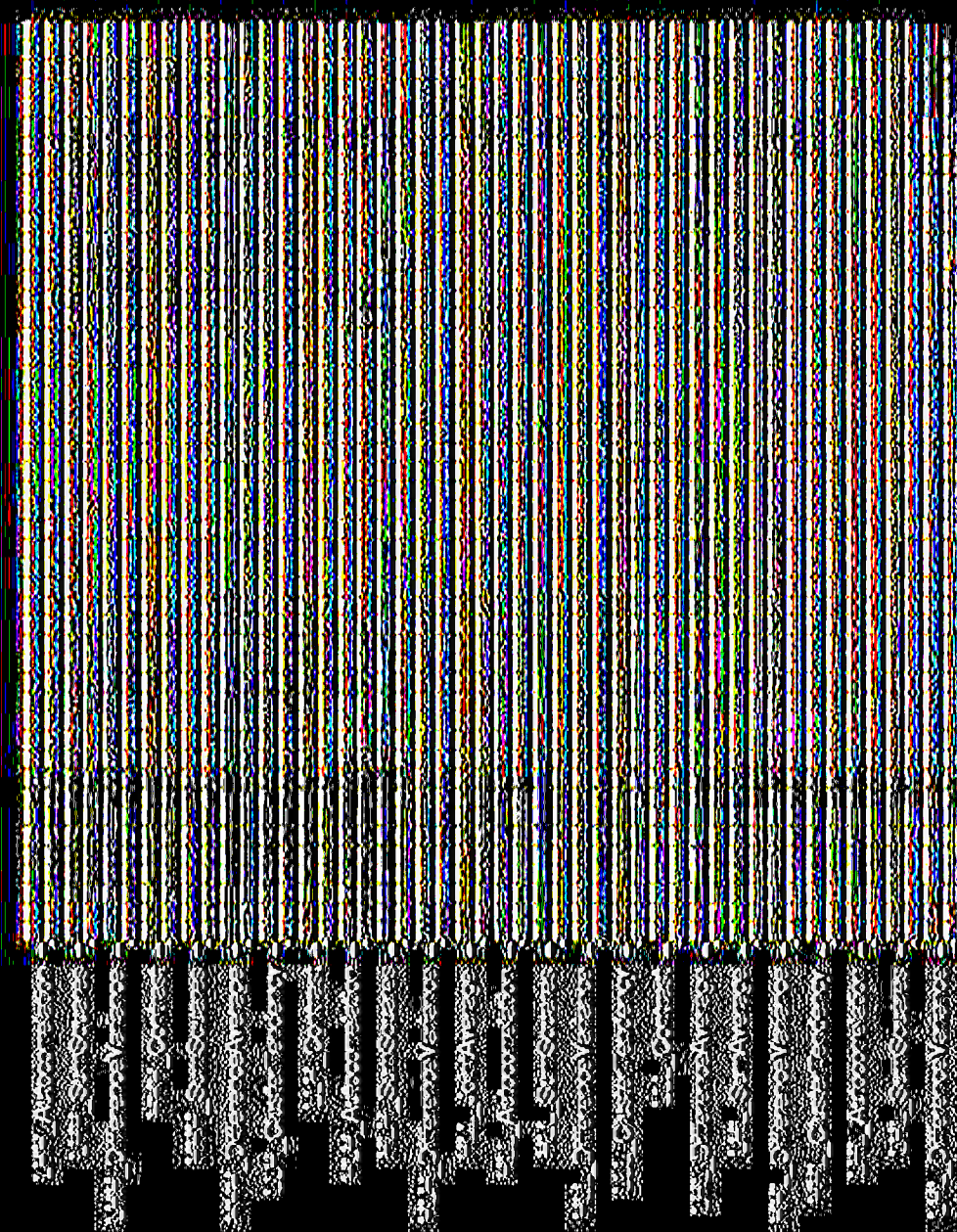


1111 InverseVariance
1111 ClusterShade
1111 ClusterShade
1111 MaximumReliability
1111 Incof
1111 Id
1111 Incof
1111 Id
1111 Id
1111 JointEnergy
1111 JointEnergy
1111 Id
1111 JointEnergy
1111 Id
1111 Incof
1111 Id
1111 JointEnergy
1111 InverseVariance
1111 Id
1111 MaxReliability
1111 MaxReliability
1111 MaxReliability
1111 Incof

(150)

小波GLCM特徵係數

06
04
10
004
008



- 1101 InverseVariance
- 1102 ClusterShade
- 1103 ClusterShade
- 1104 MaximumProbability
- 1105 Inm
- 1106 In
- 1107 Inm
- 1108 In
- 1109 In
- 1110 JointEnergy
- 1111 JointEnergy
- 1112 In
- 1113 JointEnergy
- 1114 In
- 1115 Inm
- 1116 In
- 1117 JointEnergy
- 1118 InverseVariance
- 1119 Inm
- 1120 MaxProbability
- 1121 MaxProbability
- 1122 MaxProbability
- 1123 Inm

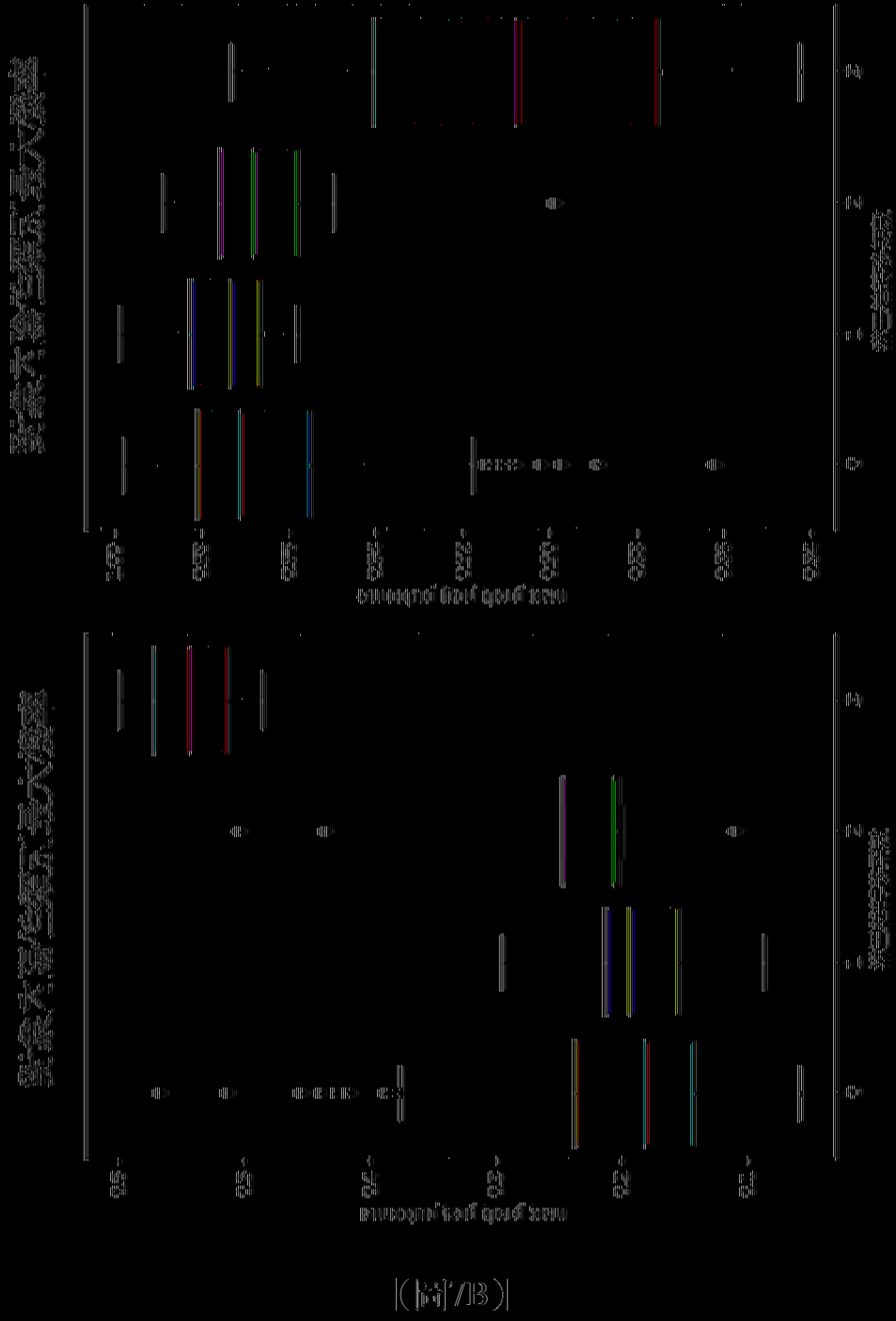
[[H]SP]

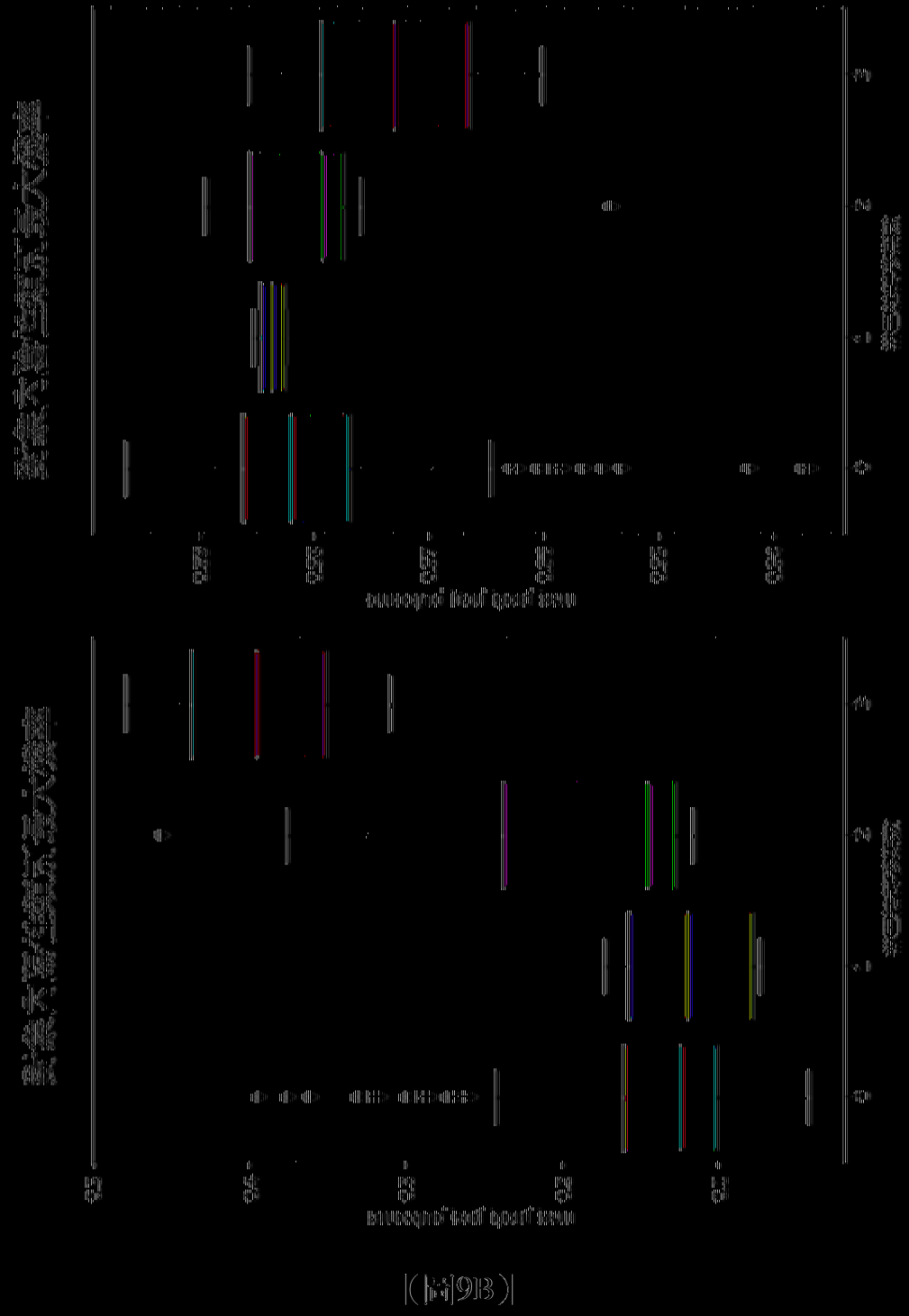
第 23 頁，共 30 頁(發明圖式)

10

1111	Inverse Variance
1111	ClusterShade
1111	ClusterShade
1111	MaximumProbability
1111	Inter
1111	Id
1111	Inter
1111	Id
1111	Id
1111	Id
1111	JointEnergy
1111	JointEnergy
1111	Id
1111	JointEnergy
1111	Id
1111	Inter
1111	Id
1111	JointEnergy
1111	InverseVariance
1111	Idm
1111	MaxProbability
1111	MaxProbability
1111	MaxProbability
1111	Inter

$$|(\mathbb{H}^{\mathbb{R}}]5\mathbb{R})|$$





【發明說明書】

【中文發明名稱】 機器學習預測侵犯淋巴結的癌症

【英文發明名稱】 MACHINE LEARNING TO PREDICT CANCER

INVOLVEMENT OF LYMPH NODES

【技術領域】

相關申請案之交互參照

【0001】 本申請案主張2020年12月1日申請之美國臨時專利申請案第63/120,102號之優先權及權益，該案之整體揭露出於所有目的以全文以引用方式併入本文中。

【先前技術】

【0002】 胸部的電腦斷層攝影(computed tomographic, CT)成像已成為用於評估處於肺癌風險者的臨床標準。這些努力著重於偵測肺結節，該等肺結節可表示在其最早及最容易治癒期的疾病。此一肺結節的偵測觸發更廣泛的臨床評估，以判定其是否為癌性的，並且若是，則判定其是否已擴散至胸部或甚至身體的遠端區域中的其他結構。此係用以判定肺癌之期別，其繼而決定治療介入之計劃。

【0003】 肺癌分期的一個態樣係判定其在胸腔淋巴結中存在或不存在（淋巴結狀態）。目前標準照護涉及分期方案，其中首先將結節判定為良性或惡性，接著完成額外檢測，諸如針對淋巴結的正子發射斷層攝影(PET)掃描及手術、內視鏡、或基於支氣管鏡之採樣。組織採樣程序增加住院診斷工作(work up)的風險、費用、及時間。需要新的非侵入性方法以減輕這些風險並縮短介於

診斷與治療之間的時間。如本文進一步詳細說明者，應用於用於肺癌篩檢之初始胸腔CT掃描的機器學習及基於深度學習之技術可解決這些需求。

【0004】 肺癌偵測、分期及治療規劃係可費時數週或數月的多步驟程序。除了此一臨床住院診斷工作的風險及費用外，此程序所需的額外時間亦會導致顯著患者焦慮及提高間隔轉移(interval metastasis)。事實上，較長的診斷至治療時間與更糟的整體存活期相關聯。減輕這些疑慮的例示性方法是允許幾乎同時的癌症診斷、分期、及（在支氣管鏡療法的情況下）治療。

【發明內容】

【0005】 本文中所揭露之本發明的實施例係基於用於篩檢高風險族群中肺癌的胸腔CT掃描之分析。在一些態樣中，識別處於癌症已轉移至胸部中之至少一個淋巴結(LN)之高風險且因此需要額外檢測（諸如PET掃描及淋巴結活體組織切片）的患者。此允許臨床提供者為最需要檢測者排定優先順序。相反地，處於具有淋巴結轉移之低風險的疑似患有肺癌的患者可直接進行至立即治療，諸如手術切除或合併的支氣管鏡診斷及治療。此將消除疾病診斷及療法之間的時間。具體而言，本文中所揭露之本發明實施例實現識別出具有淋巴結疾病之高或低風險的患者，從而允許臨床醫師對高風險個體排定額外檢測的優先順序或請具有低轉移風險之肺癌患者直接進行確認性支氣管鏡活體組織切片，緊接著進行支氣管鏡遞送療法。

【0006】 本文揭示一種判定一對象的轉移性癌症之一對象水準風險之方法，該方法包含：獲得從該對象擷取的一或多個影像，該一或多個影像包含該對象之複數個淋巴結；及藉由應用一風險模型至所獲得之該一或多個影像之所

提取特徵來預測轉移性癌症之該對象水準風險，該風險模型經訓練以預測在影像中淋巴結的轉移性癌症之淋巴結(LN)水準風險，其中針對一參考個體之一組淋巴結，至少使用導出自轉移性癌症之經推算LN水準風險的標籤來訓練該風險模型，該組淋巴結之該經推算LN水準風險係使用辨別訓練影像之癌性與非癌性淋巴結的至少一推算模型來判定。在各種實施例中，預測轉移性癌症之該對象水準風險進一步包含：基於由該風險模型所預測的轉移性癌症之該複數個淋巴結之LN水準風險來選擇該複數個淋巴結中之一或多者；及使用針對該一或多個淋巴結所預測的轉移性癌症之LN水準風險來判定轉移性癌症之該對象水準風險。

【0007】 在各種實施例中，選擇該複數個淋巴結中之一或多者包含識別具有最高機率之LN水準風險的該淋巴結，且其中判定轉移性癌症之該對象水準風險包含指派該最高機率之LN水準風險作為轉移性癌症之該對象水準風險。在各種實施例中，該風險模型預測轉移性癌症之LN水準風險的機率範圍大於由該推算模型所預測的該等經推算LN水準風險。在各種實施例中，一參考個體之該組淋巴結係藉由下列來選擇：判定一中位數風險值；及比較該參考個體之該等淋巴結的經推算LN水準風險與該中位數風險值。在各種實施例中，一參考個體之該組淋巴結進一步藉由以下來選擇：在該組淋巴結中包括具有大於該中位數風險值之LN水準風險的一或多個淋巴結。在各種實施例中，從該組淋巴結中排除在該組淋巴結中具有小於該中位數風險值之LN水準風險的一或多個淋巴結。

【0008】 在各種實施例中，該中位數風險值係該參考個體的一中位數對象內LN水準風險。在各種實施例中，該一或多個影像包含一電腦斷層攝影(CT)影像。在各種實施例中，該一或多個影像獲自一胸腔CT掃描。在各種實施例

中，本文所揭露之方法進一步包含根據轉移性癌症之該預測對象水準風險來判定該癌症之一期別。在各種實施例中，本文所揭示之方法進一步包含根據轉移性癌症之該預測對象水準風險來選擇該對象之一診斷或治療。在各種實施例中，選擇該對象之一診斷或治療包含：當轉移性癌症之該預測對象水準風險低於一臨限風險值時，選擇手術腫瘤切除或合併的支氣管鏡或內視鏡診斷及治療。在各種實施例中，選擇該對象之一診斷或治療包含：當轉移性癌症之該預測對象水準風險高於一臨限風險值時，選擇執行一PET或PET-CT掃描或執行淋巴結活體組織切片中之一者或兩者的額外診斷檢測。

【0009】 在各種實施例中，該推算模型係使用獲自胸腔CT掃描之訓練影像來訓練。在各種實施例中，獲自胸腔CT掃描之該等訓練影像包含一或多個淋巴結。在各種實施例中，該推算模型係使用國家肺部篩選試驗(National Lung Screening Trial, NLST)之訓練影像來訓練。在各種實施例中，該推算模型係使用一自訂資料集之訓練影像來訓練。

【0010】 在各種實施例中，該推算模型係藉由從該等訓練影像之至少一者產生二或更多個影像圖譜(radiomic)板來訓練。在各種實施例中，該二或更多個影像圖譜板包含訓練影像之一高斯拉普拉斯變換(Laplacian of Gaussians transformation)及訓練影像之一小波變換。在各種實施例中，該推算模型進一步藉由下列訓練：從該二或更多個影像圖譜板之各者提取特徵；及至少使用該等所提取特徵來訓練該推算模型。在各種實施例中，該推算模型進一步藉由下列訓練：在該等影像圖譜板中界定一關注區域(ROI)，所界定之該關注區域包含一淋巴結；從該等影像圖譜板中之該ROI提取特徵；及至少使用該等所提取特徵來訓練該推算模型。在各種實施例中，該等所提取特徵包括一階特徵、形狀特

徵、灰階共生矩陣1 (GLCM)特徵、灰階級長矩陣(GLRLM)特徵、灰階級帶矩陣(GLSZM)特徵、及鄰域灰階差分矩陣(NGTDM)特徵中之一或多個特徵類別。

【0011】 在各種實施例中，該推算模型係進一步使用指示訓練影像中轉移的存在或不存在的參考基準真相值來訓練。在各種實施例中，該等參考基準真相值係導出自指示該癌症之該期別的淋巴結轉移期數共變異數值，其中「0」或「1」之一癌症期別指示轉移的不存在，且其中「2」或「3」之一癌症期別指示轉移的存在。在各種實施例中，該等參考基準真相值指示個體淋巴結為癌性或非癌性的。在各種實施例中，指示個體淋巴結是否是癌與非癌的該等參考基準真相值係基於該腫瘤或該等淋巴結之特性而從該等淋巴結轉移期數共變異數值反算。

【0012】 在各種實施例中，該推算模型係一隨機森林分類器或一LASSO分類器。在各種實施例中，該風險模型係一隨機森林分類器或一LASSO分類器。在各種實施例中，該風險模型係使用監督學習技術來訓練。在各種實施例中，複數個淋巴結包含1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、或21個淋巴結。在各種實施例中，該複數個淋巴結包含8至20個之間的淋巴結、10至19個之間的淋巴結、12至18個之間的淋巴結、或14與16個之間的淋巴結。在各種實施例中，該轉移性癌症係轉移性肺癌。在各種實施例中，該風險模型展現至少AUC = 78%、靈敏度= 82%、或特異度= 60%之效能。

【0013】 本文中額外揭示一種訓練如請求項1至34中任一項之風險模型之方法，該方法包含：應用一推算模型以產生用於在訓練影像中之複數個淋巴結的經推算LN水準風險；選擇一組淋巴結，其中該組中之該等淋巴結具有大於一

中位數風險值的LN水準風險；使用在該組淋巴結中的該等淋巴結之該等LN水準風險作為用於訓練該風險模型的參考基準真相，該風險模型能夠預測LN水準風險的機率範圍大於由該推算模型所預測的該等經推算LN水準風險。在各種實施例中，該風險模型進一步藉由下列訓練：從一或多個影像圖譜板提取特徵；及至少使用該等所提取特徵來訓練該風險模型。在各種實施例中，從一或多個影像圖譜板提取特徵包含：在該一或多個影像圖譜板中界定一關注區域(ROI)，所界定之該關注區域包含一淋巴結；從該一或多個影像圖譜板中之該ROI提取特徵。在各種實施例中，該等所提取特徵包括一階特徵、形狀特徵、灰階共生矩陣1 (GLCM)特徵、灰階級長矩陣(GLRLM)特徵、灰階級帶矩陣(GLSZM)特徵、及鄰域灰階差分矩陣(NGTDM)特徵中之一或多個特徵類別。

【0014】 本文中額外揭示一種用於執行上述方法之系統。在各種實施例中，一種用於判定一對象的轉移性癌症之一對象水準風險之系統包含：一成像裝置，其經組態以擷取該對象的一或多個影像；及一運算裝置，其經組態以執行下列步驟：獲得從該對象擷取的一或多個影像，該一或多個影像包含該對象之複數個淋巴結；及藉由應用一風險模型至所獲得之該一或多個影像之所提取特徵來預測轉移性癌症之該對象水準風險，該風險模型經訓練以預測在影像中淋巴結的轉移性癌症之淋巴結(LN)水準風險，其中針對一參考個體之一組淋巴結，至少使用導出自轉移性癌症之經推算LN水準風險的標籤來訓練該風險模型，該組淋巴結之該經推算LN水準風險係使用辨別訓練影像之癌性與非癌性淋巴結的至少一推算模型來判定。

【0015】 本文中額外揭示一種用於執行上述方法之非暫態電腦可讀取媒體。在各種實施例中，非暫態電腦可讀取媒體包含當由一處理器執行時造成

〔圖5A〕至〔圖5R〕描繪不同淋巴結影像圖譜特徵之實例相關性熱圖。

〔圖6〕描繪推算模型之淋巴結水準效能。

〔圖7A〕及〔圖7B〕描繪藉由LN水準分類器所預測之最高機率評分之對象水準效能。

〔圖8〕描繪在兩步驟式拔靴法(bootstrap approach)中的推算模型之淋巴結水準效能。

〔圖9A〕及〔圖9B〕描繪使用拔靴法之對象水準效能。

【0017】 應注意，在切實可行的範圍內，類似或相似的參考符號可用於圖式中，且可指示類似或相似的功能。例如，一參考符號之後的字母（諸如「個體315A」）指示該文字具體地指具有特定參考符號的元件。本文中的參考符號無接在後面的字母（諸如「個體315」）指在圖式中有參考符號的任何或所有元件（例如，本文中的「個體315」指圖式中的參考符號「個體315A」、「個體315B」、「個體315C」、及「個體315D」）。

【實施方式】

I. 定義

【0018】 請求項及說明書中使用的用語係定義如下所述，除非另有指明。

【0019】 用語「對象(subject)」涵蓋細胞、組織、或生物體、人類或非人類，無論是體內、離體、或體外、男性或女性。

【0020】 用語「哺乳動物(mammal)」涵蓋人類及非人類，並且包括但不限於人類、非人類靈長動物、犬、貓、鼠、牛、馬、及豬。

【0021】 用語「樣本(sample)」或「測試樣本(test sample)」可包括藉由包括靜脈穿刺、排泄、射精、按摩、活體組織切片、針頭抽吸物、灌洗樣本、刮取、手術切口、或介入之手段或所屬技術領域中已知之其他手段而自一對象取得之單一細胞或多個細胞或細胞之片段或一等分試樣之體液（諸如血液樣本）。一等分試樣的體液實例包括羊水、前房液、膽汁、淋巴液、乳汁、組織間隙液、血液、血液血漿、耳垢（耳蠟）、考伯液（預射精液）、乳糜、食糜、女性潮射、月經、黏液、唾液、尿液、嘔吐物、淚液、陰道潤滑液、汗液、血清、精液、皮脂、膿汁、胸膜液、腦脊髓液、滑液、細胞內液、及玻璃體液。

【0022】 用語「參考個體(reference individual)」係指具有已知癌症分期的個體。例如，參考個體可包括健康個體（例如，對照個體）。作為另一實例，參考個體可包括先前經診斷有癌症及經分期之個體。作為另一實例，參考個體包括先前經診斷有已轉移至淋巴結之癌症（如第IV期癌症）的個體。

【0023】 用語「獲得一或多個影像(obtaining one or more images)」涵蓋獲得從對象擷取的一或多個影像，或獲得從對象獲得之樣本擷取之一或多個影像。獲得一或多個影像可涵蓋執行自對象或自獲自對象之樣本擷取一或多個影像之步驟。該用語亦可涵蓋例如自己實施自對象或自獲自對象之樣本擷取一或多個影像之步驟之第三方接收一或多個影像。一或多個影像可藉由所屬技術領域中具有通常知識者經由多種已知方式獲得，包括儲存於儲存記憶體上。

【0024】 用語「淋巴結侵犯(lymph node involvement)」、「淋巴結轉移(lymph node metastasis)」、或「淋巴結癌(lymph node cancer)」係指一對象之一或多個淋巴結中的轉移癌症之存在或不存在。

【0025】 用語「淋巴結轉移風險(lymph node metastatic risk)」、「淋巴結水準風險(lymph node level risk)」、及「LN水準風險(LN-level risk)」可互換使用，且通常指在特定淋巴結中轉移性癌症的風險。

【0026】 用語「轉移性癌症之對象水準風險(subject level risk of metastatic cancer)」或「轉移性癌症之患者水準風險(patient level risk of metastatic cancer)」可互換使用，且係指一對象中的癌症已轉移的可能性。在各種實施例中，對象中之癌症係在對象之一或多個淋巴結中的已轉移的肺癌。

【0027】 用語「癌症期別(cancer stage或stage of cancer)」係指對象中之癌症期別。癌症期別可指包括第0至4期的TNM分期系統。此外，癌症期別可指淋巴結癌0至3期。淋巴結癌0期指示無淋巴結侵犯，淋巴結癌1期指示肺門淋巴結侵犯，而淋巴結癌症期別2或3期指示縱隔腔侵犯。

【0028】 應注意的是，除非上下文清楚地指明，否則本說明書中所用的單數形式「一(a/an)」及「該(the)」包括複數指代物。

II. 系統環境概觀

【0029】 圖1A描繪根據一實施例之藉由判定一對象的轉移性癌症之一對象水準風險進行癌症分期的系統環境概觀100。系統環境100提供脈絡以介紹對象110、成像裝置120、及用於判定癌症預測140之癌症分期系統130。在各種實施例中，藉由癌症分期系統130判定的癌症預測140用作為對象110的癌症分

期。在各種實施例中，藉由癌症分期系統130判定的癌症預測140係對象110的預測風險水準，並用作為用於判定對象110之癌症分期的基礎。

【0030】 在某些實施例中，該對象係健康的。因此，本文中所述之用於癌症分期之方法對於在健康對象中早期偵測癌症可為有益。在各種實施例中，對象先前經診斷為癌症。在各種實施例中，該對象係健康的，但疑似有某種形式的癌症（例如，已展現出與癌症相關聯之症狀的對象）。在具體實施例中，對象的癌症類型係肺癌。因此，本文中所述之用於癌症分期之方法對於識別及區分處於包括淋巴結侵犯的高轉移性癌症風險之對象與處於具有淋巴結轉移之低風險的疑似患有肺癌的患者可係有益的。處於高轉移性癌症風險的對象可接受額外檢測，諸如PET掃描及淋巴結活體組織切片，而處於具有淋巴結轉移之低風險的對象可接受立即治療，諸如肺癌的手術切除或合併的支氣管鏡診斷及治療。

【0031】 在各種實施例中，成像裝置120擷取對象110之影像以供進一步分析。在各種實施例中，成像裝置120自獲自對象110之測試樣本擷取影像。該影像及/或該樣本可由個體或由第三方（例如，健康照護提供者或專業醫療人員）獲得。醫療專業人員之實例包括醫師、緊急醫療技術人員、護理師、第一線應變人員、心理師、抽血師、醫學物理人員、護理人員、外科醫師、牙醫師、及將為所屬技術領域中具有通常知識者已知之任何其他顯而易見的醫療專業人員。

【0032】 在一些實施例中，成像裝置120擷取對象110之解剖學位置的影像。對象之實例解剖學位置可包括肺、胸腔、腎、肝臟、胰臟、腦、胃、小腸、髖部、膝蓋、腿、手臂、及臉。在各種實施例中，成像裝置120擷取對象

110之胸腔的影像。在各種實施例中，成像裝置120擷取對象110之一或多個淋巴結的影像。在各種實施例中，成像裝置120擷取對象110之胸腔（包括一或多個淋巴結）的影像。

【0033】 在各種實施例中，成像裝置120係電腦斷層攝影(CT)掃描器、磁共振造影(MRI)掃描器、正子發射斷層攝影(PET)掃描器、X射線掃描器、或超音波成像裝置中之一者。在具體實施例中，成像裝置120係CT掃描器，其擷取對象110之一或多個影像。在具體實施例中，成像裝置120係CT掃描器，其擷取對象110之胸腔（包括一或多個淋巴結）的一或多個CT影像。

【0034】 一般而言，癌症分期系統130分析一或多個影像（例如，由成像裝置120擷取的影像），且產生癌症預測140。在各種實施例中，癌症預測140可係對象110中的預測癌症期別。在各種實施例中，癌症預測140可係對象110中之癌淋巴結侵犯的預測可能性。在各種實施例中，癌症預測140可包括建議的臨床方法（例如，額外檢測，諸如PET掃描及淋巴結活體組織切片）或療法（例如，手術切除及/或藥物療法）。癌症分期系統130可包括體現為電腦系統400（如下文關於圖4所討論）的一或多個電腦。因此，在各種實施例中，參照癌症分期系統130所描述的步驟係於電腦上(*in silico*)執行。

【0035】 在各種實施例中，成像裝置120及癌症分期系統130係由不同方所採用。例如，第一方操作成像裝置120以自對象110擷取一或多個影像，且接著將所擷取之該一或多個影像提供給第二方，該第二方實施癌症分期系統130以判定癌症預測140。

【0036】 現參照描繪方塊圖之圖1B，該方塊圖繪示根據一實施例之癌症分期系統130的電腦邏輯組件。此處，癌症分期系統130包括淋巴結特徵模組

145、推算模組150、風險訓練模組155、風險部署模組160、癌症預測模組165、及訓練資料儲存170。

【0037】 一般而言，淋巴結特徵模組145從導出自對象或訓練對象之影像提取特徵。推算模組150訓練及部署一推算模型。該推算模型分析所提取特徵，並判定影像中之各淋巴結的淋巴結水準風險。風險訓練模組155使用訓練資料來訓練風險模型，該訓練資料包括由該推算模型所判定的至少該淋巴結水準風險。風險部署模組160實施風險模型以分析從一對象（例如，圖1A中之對象110）獲得之影像所提取之特徵，以判定癌症預測，諸如預測對象110之對象水準轉移風險或癌症期別。換言之，癌症分期系統130訓練並實施至少兩個模型（例如，一個推算模型及一個風險模型）。因此，透過涉及該至少兩個模型的此多步驟實施方案，癌症分期系統130產生一對象的癌症預測，諸如對象水準轉移風險或癌症期別之預測。

【0038】 推算模型與風險模型兩者的實施方案導致改善的癌症預測。具體而言，推算模型從來自訓練資料之對象水準淋巴結轉移期數(subject-level N-stage)中學習，並被應用至淋巴結之經提取特徵，以界定哪些淋巴結有可能促成陽性淋巴結轉移期數。接著，風險模型從經推算資料學習更精進的淋巴結轉移風險機率。此處，風險模型預測轉移性癌症之LN水準風險的機率範圍大於由推算模型所預測的經推算LN水準風險。這實現了轉移性癌症之對象水準風險的更準確判定。

【0039】 癌症分期系統130的組件在以下係參照二期來描述：1)訓練期及2)部署期。更具體而言，訓練期係指基於訓練資料（諸如從參考個體（例如，具有已知癌症分期之個體）擷取的訓練影像）來對一或多個模型進行的建置及

訓練。因此，模型係使用訓練資料來訓練，使得在部署期期間，模型的實施實現了一對象（例如圖1A中之對象110）之癌症分期的預測。在各種實施例中，推算模組150亦可經實施以訓練及部署一模型，該模型在下文稱為一推算模型。因此，推算模組150在訓練期期間訓練該推算模型及在部署期期間部署該推算模型。

【0040】 在一些實施例中，在訓練期及部署期中之一者期間應用癌症分期系統130的組件。例如，在訓練期期間應用風險訓練模組155及訓練資料儲存170以訓練風險模型。此外，在部署期期間應用風險部署模組160。在各種實施例中，取決於在訓練期或部署期期間應用癌症分期系統130的組件，該等組件可由不同方來執行。在此類情景中，預測模型的訓練及部署係由不同方執行。例如，在訓練期期間所應用之風險訓練模組155及訓練資料儲存170可由第一方採用（例如，以訓練風險模型），且在部署期期間所應用之風險部署模組160可由第二方執行（例如，以部署風險模型）。同樣地，可由一第一方應用推算模組150以訓練推算模型，且可由一第二方應用推算模組150以部署推算模型。

III. 用於癌症分期之方法

【0041】 本文中所述之實施例包括用於藉由下列方式進行癌症分期之方法：分析從對象擷取的一或多個影像；及應用一或多個模型以判定涉及淋巴結之轉移性癌症風險的對象水準。此類方法可藉由在圖1B中描述的癌症分期系統130來執行。

【0042】 參照圖1B中的癌症分期系統130之模組，淋巴結特徵模組145分析一或多個影像並從該一或多個影像提取特徵。在各種實施例中，淋巴結特徵

模組145從獲自參考個體的訓練影像提取特徵。從訓練影像提取的這些特徵可用於訓練一或多個模型（諸如推算模型），如下文進一步詳細描述。在各種實施例中，淋巴結特徵模組145從由對象獲得的影像中提取特徵（在圖1A中之對象110）。從獲自對象之影像提取的這些特徵可藉由部署模型（諸如風險模型）來分析，如下文進一步詳細描述。

【0043】 在各種實施例中，一或多個影像或訓練影像包括在影像中之一或多個淋巴結。例如，影像或訓練影像可係包括一或多個淋巴結之胸腔的影像。在各種實施例中，影像包括至少2個淋巴結。在各種實施例中，影像包括1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、或21個淋巴結。在各種實施例中，影像包括1至10個之間的淋巴結、1至15個之間的淋巴結、2至10個之間的淋巴結、2至15個之間的淋巴結、2至20個之間的淋巴結、4至10個之間的淋巴結、4至15個之間的淋巴結、4至20個之間的淋巴結、5至10個之間的淋巴結、5至15個之間的淋巴結、5至20個之間的淋巴結、8至15個之間的淋巴結、8至20個之間的淋巴結、10至15個之間的淋巴結、10至20個之間的淋巴結、12至15個之間的淋巴結、12至20個之間的淋巴結、或15至20個之間的淋巴結。在各種實施例中，從單一對象或參考個體擷取多個影像。多個影像可擷取對象或參考個體之不同淋巴結。在各種實施例中，多個影像包括至少2個淋巴結。在各種實施例中，多個影像包括2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、或21個淋巴結。在各種實施例中，多個影像包括8至20個之間的淋巴結。在各種實施例中，多個影像包括10至19個之間的淋巴結、12至18個之間的淋巴結、或14與16之間的淋巴結。

【0044】 圖2A描繪根據一實施例之識別淋巴結站的實例淋巴結圖。更具體而言，圖2A係修改自美國胸腔學會映射方案(mapping scheme)的局部淋巴結分類之描繪。此處，淋巴結站經標註以描繪各淋巴結之位置及變化。具體而言，該等標註係如下：(1)鎖骨上結、(2R/2L)上氣管旁結、(3A)血管前結、(3P)脊椎前方結、(4R/4L)下管旁結、(5)主動脈下結、(6)側主動脈結、(7)氣管隆凸下結、(8)食道旁結、(9)肺韌帶結、(10)肺門結、(11)葉間結、(12)肺葉結、(13)肺小節結、及(14)肺次節葉結。

【0045】 淋巴結特徵模組145標註影像中的一或多個淋巴結。在各種實施例中，淋巴結特徵模組145藉由界定圍繞淋巴結的區域來標註一或多個淋巴結。在各種實施例中，圍繞淋巴結的經界定區域係球形區域。在各種實施例中，該球形區域具有 X mm之半徑。在各種實施例中， X 係5 mm。在各種實施例中， X 係7.5 mm。在一些實施例中， X 係0.5 mm、1 mm、1.5 mm、2 mm、2.5 mm、3 mm、3.5 mm、4 mm、4.5 mm、5.5 mm、6 mm、6.5 mm、7 mm、8 mm、8.5 mm、9 mm、9.5 mm、10 mm、10.5 mm、11 mm、11.5 mm、12 mm、12.5 mm、13 mm、13.5 mm、14 mm、14.5 mm、或15 mm。在各種實施例中， X 係1與15 mm之間。在各種實施例中， X 係2與12 mm之間。在各種實施例中， X 係3與10 mm之間。在各種實施例中， X 係4與8 mm之間。在各種實施例中， X 係5與7.5 mm之間。在各種實施例中， X 係4與6 mm之間。在各種實施例中， X 係6與8 mm之間。在各種實施例中， X 係7與8 mm之間。在各種實施例中， X 係10與15 mm之間。

【0046】 在各種實施例中，淋巴結特徵模組145藉由界定圍繞淋巴結的二或更多個球形區域來標註一或多個淋巴結。例如，淋巴結特徵模組145藉由界

定圍繞淋巴結的5 mm之一第一球形區域及圍繞淋巴結的7.5 mm之一第二球形區域來標註一或多個淋巴結。圖2B係根據一實施例之一實例淋巴結標註。具體而言，圖2B描繪圍繞兩個2R上氣管旁結的兩個經界定球形區域。

【0047】 在各種實施例中，淋巴結特徵模組145從圍繞淋巴結的各經界定區域提取特徵。在各種實施例中，淋巴結特徵模組145執行原始輸入的變換以產生一經變換影像，然後從該經變換影像中提取額外特徵。在一實施例中，淋巴結特徵模組145執行原始影像之小波變換。在一實施例中，淋巴結特徵模組145執行原始影像之高斯拉普拉斯(LoG)。因此，淋巴結特徵模組145可從經變換影像中圍繞淋巴結的經界定區域提取特徵。在各種實施例中，淋巴結特徵模組145從原始影像及經變換影像中圍繞淋巴結的經界定區域提取特徵。

【0048】 在具體實施例中，淋巴結特徵模組145執行原始影像的兩個變換（例如，原始影像之小波變換及原始影像之高斯拉普拉斯(LoG)兩者）。因此，淋巴結特徵模組145可從第一經變換影像及第二經變換兩者中圍繞淋巴結的經界定區域提取特徵。在各種實施例中，淋巴結特徵模組145從原始影像、第一經變換影像、及第二經變換中圍繞淋巴結的經界定區域提取特徵。

【0049】 來自原始影像、第一變換影像、及/或第二變換影像之實例特徵類別可包括一階特徵、形狀特徵、灰階共生矩陣1 (GLCM)特徵、灰階級長矩陣(GLRLM)特徵、灰階級帶矩陣(GLSZM)特徵、及鄰域灰階差分矩陣(NGTDM)特徵中之一或多個特徵類別。一階特徵係指強度值的一階統計量之特徵。GLCM特徵在影像上以給定偏移從共生灰階值導出直方圖。GLRLM特徵涉及灰階級長矩陣(run length matrix)的統計。GLSDRUE特徵涉及灰階級帶矩陣(size zone matrix)的統計。NGTDM特徵涉及鄰域灰階差分矩陣(gray tone difference

matrix)的統計。不同特徵類別中之特徵的實例包括：相關特徵、LMC2、MCC、差熵、聯合熵、和熵、群聚突出、平均差、聯合平均、平均加總、自動相關度、變異性差別、對比、群聚傾向(cluster tendency)、差值平方和、LDMN、LDN、群聚陰暗度(cluster shade)、LMC1、逆方差、LD、LDM、聯合能量、及最大機率。

【0050】 在各種實施例中，淋巴結特徵模組145從影像提取至少50個特徵。在各種實施例中，淋巴結特徵模組145從影像提取至少100個特徵、至少150個特徵、至少200個特徵、至少250個特徵、至少300個特徵、至少350個特徵、至少400個特徵、至少450個特徵、至少500個特徵、至少550個特徵、至少600個特徵、至少650個特徵、至少700個特徵、至少750個特徵、至少800個特徵、至少850個特徵、至少900個特徵、至少950個特徵、至少1000個特徵、至少1100個特徵、至少1200個特徵、至少1300個特徵、至少1400個特徵、至少1500個特徵、至少1600個特徵、至少1700個特徵、至少1800個特徵、至少1900個特徵、至少2000個特徵、至少3000個特徵、至少4000個特徵、至少5000個特徵、至少6000個特徵、至少7000個特徵、至少8000個特徵、至少9000個特徵、或至少10,000個特徵。在各種實施例中，淋巴結特徵模組145提取介於10個特徵與10000個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於100個特徵與1000個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於300個特徵與900個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於500個特徵與1000個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於500個特徵與900個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於700個特徵與1000個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於700個特

徵與900個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於800個特徵與1000個特徵之間。在各種實施例中，淋巴結特徵模組145提取介於800個特徵與900個特徵之間。

【0051】 在各種實施例中，淋巴結特徵模組145提供經提取特徵至推算模組150，以用於訓練或部署推算模型。例如，若提取特徵係導出自從參考個體所擷取之訓練影像，則淋巴結特徵模組145將所提取特徵提供至推算模組150。在各種實施例中，淋巴結特徵模組145將所提取特徵提供至風險訓練模組155，以用於訓練風險模型。例如，若所提取特徵係導出自從參考個體所擷取之訓練影像，則淋巴結特徵模組145將所提取特徵提供至風險訓練模組155，以用於訓練風險模型。在各種實施例中，淋巴結特徵模組145將所提取特徵提供至風險部署模組160，以用於部署風險模型。。例如，若所提取特徵係導出自從一對象（諸如圖1A中之對象110）所擷取之影像，則淋巴結特徵模組145將所提取特徵提供至風險部署模組160。

III. A.訓練一推算模型

【0052】 推算模組150訓練及部署一推算模型。一般而言，該推算模型經結構化使得其：分析從包括一或多個淋巴結之影像所提取的特徵；及針對該一或多個淋巴結預測該等淋巴結為癌性的機率。因此，基於來自包括一或多個淋巴結之一影像的所提取特徵，該推算模型推算該等淋巴結之各者的一機率，所推算機率指示該淋巴結為癌性（例如，轉移性癌症）的可能性。換言之，推算模型藉由判定淋巴結水準之轉移性癌症的風險來辨別癌性與非癌性淋巴結。

【0053】 在各種實施例中，推算模型係下列任一者：迴歸模型（例如，線性迴歸、邏輯迴歸、或多項式迴歸）、決策樹、隨機森林、支持向量機、單純貝氏模型、k平均值群集、或神經網路（例如，前饋網路、卷積神經網路(CNN)、深度神經網路(DNN)、自動編碼器神經網路、生成對抗網路、或遞迴網路（例如，長短期記憶網路(LSTM)、雙向遞迴網路、深度雙向遞迴網路）、或其任何組合。在具體實施例中，該推算模型係一隨機森林分類器或一最小絕對緊縮與選擇運算子(LASSO)分類器。在各種實施例中，該推算模型係一機器訓練模型。在各種實施例中，該推算模型係一梯度提升機器學習模型（例如，xgboost）。

【0054】 該推算模型可使用機器學習實施方法來訓練，諸如下列任一者：線性迴歸演算法、邏輯迴歸演算法、決策樹演算法、支持向量機分類、單純貝氏分類、K最近鄰分類、隨機森林演算法、深度學習演算法、梯度提升演算法、及降維技術（諸如流形學習(manifold learning)、主組分分析、因子分析、自動編碼器正則化、及獨立組分分析）、或其組合。在各種實施例中，該推算模型係使用監督學習演算法、無監督學習演算法、半監督學習演算法（例如，部分監督）、弱監督(weak supervision)、遷移(transfer)、多工學習、或其任何組合訓練。

【0055】 在各種實施例中，該推算模型具有一或多個參數，諸如超參數或模型參數。超參數通常在訓練之前建立。超參數之實例包括學習速率、決策樹的深度或樹葉、深度神經網路中的隱層數、k平均值群集中的聚類數、迴歸模型中的懲罰項、及與代價函數相關聯的正則化參數。模型參數通常在訓練期間調整。模型參數之實例包括與神經網路層中之節點相關聯的權重、支持向量

機中的支持向量、及迴歸模型中的係數。使用訓練資料訓練（例如，調整）該推算模型的模型參數，以改善該推算模型的預測能力。

【0056】 推算模組150使用訓練資料來訓練該推算模型。可儲存訓練資料及/或自訓練資料儲存170取得訓練資料。在各種實施例中，訓練資料包括來自包括獲自參考個體之一或多個淋巴結的訓練影像的所提取特徵。在各種實施例中，參考個體包括健康個體（例如，對照個體）。在各種實施例中，參考個體包括先前經診斷有癌症之個體。在各種實施例中，參考個體包括先前經診斷有已經轉移至淋巴結之癌症的個體。因此，從此類個體所擷取的影像包括一或多個癌性淋巴結。

【0057】 在各種實施例中，用於訓練推算模型之訓練資料包括參考基準真相值，其指示淋巴結是否為癌性的（以下亦稱為「陽性(positive)」或「+」）或淋巴結是否為非癌性的（以下亦稱為「陰性(negative)」或「-」）。在各種實施例中，訓練資料中的參考基準真相值係二進位值，諸如「1」或「0」。例如，癌性的陽性淋巴結可在訓練資料中以「1」之值識別，而非癌性的陰性淋巴結可在訓練資料中以「0」之值識別。總之，推算模組150使用訓練資料來訓練推算模型，以最小化損失函數，使得推算模型可基於輸入（例如，淋巴結之所提取特徵）來更佳地預測結果（例如，淋巴結的淋巴結轉移性癌症之機率）。在各種實施例中，損失函數係針對最小絕對緊縮與選擇運算子(LASSO)迴歸、脊迴歸、或彈性網路(ElasticNet)迴歸中之任一者來建構。

【0058】 在各種實施例中，訓練資料可獲自及/或導出自公開資料庫。例如，訓練資料可獲自及/或導出自國家肺部篩檢試驗(NLST)，其包括參考個體之淋巴結期別（淋巴結轉移期數(N stage)）癌症資訊。在一些實施例中，訓練

資料可不依賴公開資料庫來獲得及收集，例如藉由從複數個參考個體來擷取一或多個淋巴結之影像。例如，此訓練資料可為自訂資料集。自訂資料集的訓練資料可在患者照護期間來獲取，因此可具有對應之淋巴結的手術採樣資訊。訓練資料可包括用於參考個體的淋巴結轉移期數癌症資訊，例如，基於對參考個體的成像及/或活體組織切片結果所先前診斷的淋巴結轉移期數資訊。

【0059】 在各種實施例中，指示淋巴結是否為癌性的參考基準真相值係導出自臨床期別資料的淋巴結期別（淋巴結轉移期數）資訊。例如，公開的 NLST 資料集包括參考個體之臨床期別資訊（例如，1a 期、1b 期、2a 期、2b 期、3a 期等），但不包括淋巴結特異度資訊。換言之，NLST 資料集無法區分個體淋巴結是否為癌性或非癌性的。在此類實施例中，指示淋巴結是否為癌性的參考基準真相值可基於參考個體癌症之淋巴結轉移期數資訊及其他特性（例如，腫瘤大小、腫瘤位置、腫瘤形狀、及/或腫瘤密度）及/或淋巴結（例如，淋巴結大小、淋巴結密度、及/或淋巴結形狀）來反算。作為第一實例，具有 0 或 1 之淋巴結轉移期數的參考個體尚未遇到癌症轉移，且因此，此參考個體之淋巴結係被指派有「陰性」或非癌性的參考基準真相值。作為另一實例，具有 2 或 3 之淋巴結轉移期數的參考個體已發生癌症轉移。因此，基於腫瘤及/或淋巴結的特性，淋巴結被指派「陽性」或「陰性」的參考基準真相值。

III. B. 部署—推算模型

【0060】 推算模組 150 部署經訓練之推算模型以：分析從包括一或多個淋巴結之影像所提取的特徵；及針對該一或多個淋巴結預測該等淋巴結為癌性的機率。在各種實施例中，經訓練之該推算模型係針對一測試資料集來部署。在

各種實施例中，用於訓練推算模型的測試資料集及訓練資料集可係導出自一共同資料集。例如，共同資料集可被分割分成兩個子集，以用於訓練及測試該推算模型。在一些實施例中，共同資料集可經歷50:50的訓練資料集:測試資料集分割。在一些實施例中，共同資料集可經歷60:40的訓練資料集:測試資料集分割。在一些實施例中，共同資料集可經歷80:20的訓練資料集:測試資料集分割。

【0061】 經訓練之該推算模型之部署產出淋巴結的機率（例如，0與1之間的連續機率），各機率指示對應的淋巴結是否可能為癌性的。在各種實施例中，推算模組150提供用於訓練風險模型的由推算模型所預測的淋巴結之淋巴結轉移風險中之一或多者。在各種實施例中，推算模組150提供用於訓練風險模型的由推算模型所預測的淋巴結之所有淋巴結轉移風險。在各種實施例中，推算模組150執行一選擇程序，以識別由推算模型所預測的淋巴結的淋巴結轉移風險之一子集。因此，淋巴結轉移風險之子集隨後用於訓練風險模型。

【0062】 在一些實施例中，推算模組150藉由比較淋巴結轉移風險與臨限值來選擇淋巴結轉移風險之子集。若淋巴結之淋巴結轉移風險高於一臨限值，則選擇將該淋巴結之該淋巴結轉移風險納入該子集，以用於訓練風險模型。若淋巴結之淋巴結轉移風險低於一臨限值，則將將淋巴結之將淋巴結轉移風險從該子集排除，並因此不將其用於訓練風險模型。

【0063】 在各種實施例中，臨限值係一固定值。在各種實施例中，臨限值係50%轉移性癌症風險。在各種實施例中，臨限值係30%、35%、40%、45%、55%、60%、65%、70%、75%、80%、85%、90%、95%、或99%轉移性癌症風險。在一些實施例中，臨限值係介於30%與70%之間。在一些實施例

中，臨限值係介於40%與60%之間。在一些實施例中，臨限值係介於45%與55%之間。在一些實施例中，臨限值係介於48%與52%之間。在一些實施例中，臨限值係介於49%與51%之間。在各種實施例中，臨限值係可變的。在各種實施例中，臨限值係具有癌症的中位數對象內機率。

【0064】 推算模組150提供淋巴結轉移風險之子集至風險訓練模組155，以用於訓練風險模型。

III. C.訓練一風險模型

【0065】 風險訓練模組155使用訓練資料（包括由推算模型所預測的淋巴結之淋巴結轉移風險中之至少一或多者）來訓練風險模型。在此，一或多個淋巴結轉移風險可係由推算模組150所選擇的淋巴結轉移風險之子集，如上文所述。相較於用來訓練推算模型的訓練資料集，用於訓練風險模型的淋巴結轉移風險之子集表示一經純化訓練資料集。具體而言，經純化訓練資料集包括較可能為癌性的淋巴結之子集，而將較不可能為癌性（但除此之外可能表示偽陽性，或者對於訓練風險模型而言係混淆的）的淋巴結從經純化訓練資料集移除。因此，藉由以此純化訓練資料集來訓練風險模型，風險模型能夠針對患有轉移性疾病之對象中的所有淋巴結均預測出較大的機率範圍。更具體而言，風險模型較為可能將較高淋巴結轉移風險指派給很可能有癌症的淋巴結，並將較低淋巴結轉移風險指派給不可能有癌症的淋巴結。

【0066】 在各種實施例中，該風險模型分析從包括一或多個淋巴結之影像所提取的特徵，並針對該一或多個淋巴結之各者預測該淋巴結為癌性的連續機率。因此，基於來自包括一或多個淋巴結之一影像的所提取特徵，該風險模

型判定該等淋巴結之各者的一機率，所判定機率指示該淋巴結為癌性（例如，轉移性癌症）的可能性。在此，風險模型判定淋巴結水準的轉移性癌症之風險。在各種實施例中，風險模型類似於推算模型，因為這兩個模型皆分析從影像所提取的特徵並預測淋巴結水準的轉移性癌症之風險。然而，風險模型係使用推算模型的預測（例如，由推算模型所預測的一或多個淋巴結水準風險）來訓練。總之，此使風險模型能夠預測轉移性癌症之LN水準風險的機率範圍大於由該推算模型所預測的該等經推算LN水準風險。

【0067】 風險模型係使用包括由推算模型所預測的至少一或多個淋巴結轉移風險的訓練資料來訓練。在此，一或多個淋巴結轉移風險可用作為參考基準真相值，以用於訓練風險模型。在各種實施例中，訓練資料可額外包括從獲自參考個體之一或多個淋巴結的訓練影像的所提取特徵。在一些實施例中，用於訓練風險模型的訓練影像可係與先前用於訓練推算模型相同的訓練影像。在一些實施例中，用於訓練風險模型的訓練影像係與先前用於訓練推算模型不同的訓練影像。

【0068】 在各種實施例中，訓練資料包括個體訓練實例，使得風險模型係使用各訓練實例來訓練。例如，各訓練實例可包括一特定訓練影像之一或多個淋巴結的所提取特徵以及參考基準真相值（例如，特定訓練影像的相同一或多個淋巴結之淋巴結轉移風險）。因此，對於各訓練實例，風險模型經訓練以最小化損失函數（例如，針對最小絕對緊縮與選擇運算子(LASSO)迴歸、脊迴歸、或彈性網路迴歸中之任一者的損失函數）。

【0069】 在各種實施例中，風險模型係下列任一者：迴歸模型（例如，線性迴歸、邏輯迴歸、或多項式迴歸）、決策樹、隨機森林、支持向量機、單

純貝氏模型、 k 平均值群集、或神經網路（例如，前饋網路、卷積神經網路(CNN)、深度神經網路(DNN)、自動編碼器神經網路、生成對抗網路、或遞迴網路（例如，長短期記憶網路(LSTM)、雙向遞迴網路、深度雙向遞迴網路）、或其任何組合。在具體實施例中，風險模型係一隨機森林分類器或一最小絕對緊縮與選擇運算子(LASSO)分類器。在各種實施例中，風險模型係一機器訓練模型。

【0070】 風險模型可使用機器學習實施方法來訓練，諸如下列任一者：線性迴歸演算法、邏輯迴歸演算法、決策樹演算法、支持向量機分類、單純貝氏分類、 K 最近鄰分類、隨機森林演算法、深度學習演算法、梯度提升演算法、及降維技術（諸如流形學習、主組分分析、因子分析、自動編碼器正則化、及獨立組分分析）、或其組合。在各種實施例中，風險模型係使用監督學習演算法、無監督學習演算法、半監督學習演算法（例如，部分監督）、弱監督、遷移、多工學習、或其任何組合訓練。

【0071】 在各種實施例中，風險模型具有一或多個參數，諸如超參數或模型參數。超參數通常在訓練之前建立。超參數之實例包括學習速率、決策樹的深度或樹葉、深度神經網路中的隱層數、 k 平均值群集中的聚類數、迴歸模型中的懲罰項、及與代價函數相關聯的正則化參數。模型參數通常在訓練期間調整。模型參數之實例包括與神經網路層中之節點相關聯的權重、支持向量機中的支持向量、及迴歸模型中的係數。使用訓練資料訓練（例如，調整）風險模型的模型參數，以改善風險模型的預測能力。

III. C.部署一風險模型

【0072】 風險部署模組160部署風險模型，以判定一對象（例如，上文在圖1A所述之對象110）的轉移性癌症之一對象水準風險。此處，風險部署模組160將從對象擷取的包括一或多個淋巴結之影像的所提取特徵作為輸入提供至經訓練之該風險模型。經訓練之該風險模型分析所提取特徵，並針對淋巴結輸出指示個體淋巴結是否可能為癌性的機率。

【0073】 在各種實施例中，在部署風險模型之後，風險部署模組160從由風險模型所預測的轉移性癌症之個體淋巴結水準風險變換成對象的轉移性癌症之對象水準風險。在各種實施例中，風險部署模組160組合轉移性癌症之個體淋巴結水準風險，以判定轉移性癌症之對象水準風險。在各種實施例中，風險部署模組160跨轉移性癌症之淋巴結水準風險來執行統計運算。統計運算可識別跨轉移性癌症之個體淋巴結水準風險的最大值、最小值、平均值、中位數值、或眾數值。在具體實施例中，統計運算係跨轉移性癌症之個體淋巴結水準風險的最大值。因此，具有轉移性疾病之最高淋巴結水準風險的淋巴結被指派為對象的轉移性疾病之對象水準風險。

【0074】 癌症分期模組165基於對象的轉移性疾病之對象水準風險來判定對象的癌症期別。在各種實施例中，該癌症分期模組165藉由比較該對象的轉移性疾病之對象水準風險與一或多個臨限風險值來判定該對象的癌症期別。這些臨限風險值可作為用於判定該對象中之癌症是否可能在0期、1期、2期、或3期中之任一者的截止點。

【0075】 在一些實施例中，若對象水準風險大於30%，則對象被分類為具有至少3期淋巴結癌。在一些實施例中，若對象水準風險大於40%、大於

50%、大於60%、大於70%、大於80%、或大於90%風險，則對象被分類為具有至少3期癌症。

【0076】 在各種實施例中，臨限風險值界定對應於一或多個癌症期別的風險範圍。因此，臨限風險值設定風險範圍的上限及下限。在一些實施例中，若對象水準風險係介於30%與100%之間，則將對象分類為具有3期淋巴結癌。在一些實施例中，若對象水準風險係介於40%與100%之間，則將對象分類為具有3期淋巴結癌。在一些實施例中，若對象水準風險係介於50%與100%之間，則將對象分類為具有3期淋巴結癌。

【0077】 在一些實施例中，若對象水準風險係介於10%與40%之間，則將對象分類為具有2期淋巴結癌。在一些實施例中，若對象水準風險係介於15%與40%之間，則將對象分類為具有2期淋巴結癌。在一些實施例中，若對象水準風險係介於20%與40%之間，則將對象分類為具有2期淋巴結癌。在一些實施例中，若對象水準風險係介於30%與40%之間，則將對象分類為具有2期淋巴結癌。在一些實施例中，若對象水準風險係介於15%與30%之間，則將對象分類為具有2期淋巴結癌。在一些實施例中，若對象水準風險係介於20%與30%之間，則將對象分類為具有2期淋巴結癌。

【0078】 在一些實施例中，若對象水準風險係介於5%與20%之間，則將對象分類為具有0期淋巴結癌或1期淋巴結癌。在一些實施例中，若對象水準風險係介於10%與20%之間，則將對象分類為具有0期淋巴結癌或1期淋巴結癌。在一些實施例中，若對象水準風險係介於5%與15%之間，則將對象分類為具有0期淋巴結癌或1期淋巴結癌。在一些實施例中，若對象水準風險係介於8%與12%之間，則將對象分類為具有0期淋巴結癌或1期淋巴結癌。

IV. 用於癌症分期之實例方法

【0079】 圖3A描繪根據一實施例之推算模型及風險模型之一實例實施方案。此處，參考個體305A及305B可經識別為具有特定淋巴結轉移期數。例如，參考個體305A被識別為陰性的淋巴結轉移期數，指示參考個體305A未展現淋巴結癌侵犯。參考個體305B被識別為陽性的淋巴結轉移期數，指示參考個體305B展現淋巴結癌侵犯。然而，對象水準淋巴結轉移期數不區分特定淋巴結，且不區分個體淋巴結為癌性或非癌性。

【0080】 推算模型310推算各淋巴結的轉移風險（例如，淋巴結轉移風險），從而實現具有轉移性癌症之淋巴結與非癌性淋巴結的區分。在各種實施例中，推算模型310針對各淋巴結預測表示該淋巴結為癌性之可能性的機率。

【0081】 如圖3A所示，推算模型310之輸出可識別出個體305C（對應於參考個體305A）未展現淋巴結侵犯。換言之，個體305C之淋巴結皆可能不是癌性的。此確認了陰性淋巴結轉移期數的識別。此外，推算模型310之輸出可識別個體305D（對應於參考個體305B）展現一或多個淋巴結的淋巴結侵犯。例如，推算模型310可推算機率使得在個體305D中只有一個淋巴結被識別為癌性的，而個體305D中的其他淋巴結則係非癌性的。

【0082】 在各種實施例中，由推算模型310所預測的淋巴結之淋巴結轉移風險中之一或多者係用於訓練風險模型320。在各種實施例中，由推算模型310所預測的淋巴結之所有淋巴結轉移風險係用於訓練風險模型320。在各種實施例中，對應於淋巴結之子集的淋巴結轉移風險之子集係用於訓練風險模型

320，從而使風險模型能夠針對患有轉移性疾病之對象中的所有淋巴結均預測出較大的機率範圍。

【0083】 經訓練後，風險模型320經部署以預測一或多個對象315之轉移性癌症之對象水準風險。因此，風險模型320經部署以從轉移性疾病之淋巴結水準風險轉變至轉移性疾病之對象水準風險。如圖3A所示，部署風險模型320經部署以分析從擷取自個體315之影像的所提取特徵。風險模型320預測個體315之各者的轉移性癌症之對象水準風險。例如，風險模型320預測個體315A具有20%轉移性癌症風險，個體315B具有30%轉移性癌症風險，個體315C具有50%轉移性癌症風險，且個體315D具有60%轉移性癌症風險。因此，個體315A及315B可被分類為具有早期癌症（例如，0期或1期），而個體315C及315D可被分類為具有末期癌症（例如，2期，3期、或4期）。

【0084】 個體315可依據其轉移性疾病之對象水準風險接收不同治療。例如，處於具有淋巴結轉移之低風險的個體315A及315B可直接進行至立即治療，諸如手術切除或合併的支氣管鏡診斷及治療。作為另一實例，處於具有淋巴結轉移之較高風險的個體315C及315D可接收額外檢測，諸如PET掃描及淋巴結活體組織切片。

【0085】 圖3B描繪根據一實施例之進一步細節的風險模型320之實施方案。此處，風險模型320分析從對象350擷取之影像的所提取特徵，並輸出對象之一或多個淋巴結之各者的機率。如圖3B所示，風險模型320輸出三個淋巴結的機率。在其他實施例中，風險模型320可輸出對象350之額外或更少的淋巴結機率。

【0086】 如圖3B所示，風險模型320預測第一淋巴結具有20%的機率為癌性，而反之，80%的機率為非癌性。此外，風險模型320預測第二淋巴結具有30%的機率為癌性，而反之，70%的機率為非癌性。此外，風險模型320預測第三淋巴結具有60%的機率為癌性，而反之，40%的機率為非癌性。個體淋巴結機率經分析以識別轉移性癌症之對象水準風險。在此實例中，將轉移性癌症的最大淋巴結機率當作是轉移性癌症之對象水準機率。具有60%的機率為癌性的第三淋巴結係淋巴結之中的最高機率。因此，對象350的轉移性癌症之對象水準機率係60%。

V. 癌症

【0087】 本文中所述之方法涉及藉由判定對象的癌症之對象水準風險來進行癌症分期。在各種實施例中，對象中的癌症可包括以下一或多者：淋巴瘤、B細胞淋巴瘤、T細胞淋巴瘤、蕈狀肉芽腫、何杰金氏病(Hodgkin's Disease)、骨髓性白血病、膀胱癌、腦癌、神經系統癌症、頭頸癌、頭頸鱗狀細胞癌、腎癌、肺癌、神經胚細胞瘤/神經膠質母細胞瘤、卵巢癌、胰臟癌、前列腺癌、皮膚癌、肝癌、黑色素瘤、口、咽喉、喉、肺部的鱗狀細胞癌、結腸癌、子宮頸癌(cervical cancer)、子宮頸癌(cervical carcinoma)、乳癌、及上皮癌症、腎癌、泌尿生殖道癌症、肺癌(pulmonary cancer)、食道癌、胃癌、甲狀腺癌、頭頸癌、大腸癌、造血性癌症、睪丸癌、結腸及/或直腸癌、子宮癌、或前列腺癌。在各種實施例中，癌症可係下列任一者：支氣管肺泡癌(bronchioloalveolar carcinoma, BAC)、膀胱癌、雌性生殖道惡性腫瘤（例如，子宮漿液性腺癌、子宮內膜癌、外陰鱗狀細胞癌、及子宮肉瘤）、卵巢表面上皮

第 31 頁，共 52 頁(發明說明書)

癌（例如，卵巢透明細胞癌、上皮卵巢癌、輸卵管癌、及原發性腹膜癌）、乳房癌、非小細胞肺癌(non-small cell lung cancer, NSCLC)、雄性生殖道惡性腫瘤（例如，睪丸癌）、後腹腔或腹膜癌、胃食道腺癌、胃食道接合部癌、肝細胞癌、食道胃接合部癌、子宮頸癌、膽管癌、胰臟腺癌、肝外膽管癌、小腸惡性腫瘤、胃癌、原發部位不明轉移癌(cancer of unknown primary, CUP)、結直腸腺癌、食道癌、攝護腺癌、腎癌、頭頸鱗狀細胞癌、胸腺癌、非黑色素瘤皮膚癌、甲狀腺癌（例如，乳突癌）、頭頸癌、肛癌、非上皮卵巢癌(non-epithelial ovarian cancer, non-EOC)、葡萄膜黑色素瘤、惡性肋膜間皮瘤、小細胞肺癌(small cell lung cancer, SCLC)、中樞神經系統癌、神經內分泌瘤、及軟組織腫瘤。在某些實施例中，該癌症係乳癌、非小細胞肺癌、膀胱癌、腎臟癌、結腸癌、及黑色素瘤。在一些實施例中，對象的癌症可係轉移性癌症，包括膀胱癌、乳癌、結腸癌、腎癌、肺癌、黑色素瘤、卵巢癌、胰臟癌、前列腺癌、直腸癌、胃癌、甲狀腺癌、或子宮癌中任一者。在具體實施例中，癌症係肺癌。在具體實施例中，癌症係一種肺癌，包括小細胞肺癌、非小細胞肺癌、腺癌、鱗狀細胞癌、大細胞癌、小細胞癌、組合式小細胞癌、肺部肉瘤、肺部淋巴瘤、支氣管類癌中任一者。

VI. 導引介入

【0088】 本文中所述之實施例涉及藉由判定對象中之轉移性癌症之對象水準風險來判定癌症分期。在各種實施例中，基於針對一對象所預測的癌症分期及/或轉移性癌症之對象水準風險而提供介入給該對象。此類介入（以下稱為導引介入）可為下列任一者：診斷施用、治療劑之施用、對象的後續檢測、後

續的活體組織切片（例如，癌活體組織切片或淋巴結活體組織切片、淋巴結的支氣管鏡採樣）、後續影像掃描（例如，PET掃描、MRI掃描超音波成像、或X射線成像）、腫瘤切除、支氣管鏡診斷、選擇及/或投予（多個）治療劑、選擇/投予醫藥組成物、納入臨床試驗、不納入臨床試驗、或其任何組合。

【0089】 在各種實施例中，針對對象所預測的轉移性癌症之對象水準風險或所預測的癌症期別指示該對象處於高轉移性癌症風險。在各種實施例中，處於高轉移性癌症風險的對象具有高於一臨限評分的轉移性癌症之所預測對象水準風險。在各種實施例中，臨限評分高於20%轉移性癌症風險。在各種實施例中，臨限評分高於25%轉移性癌症風險。在各種實施例中，臨限評分高於30%、高於35%、高於40%、高於45%、高於50%、高於55%、高於60%、高於65%、高於70%、高於75%、高於80%、高於85%、高於90%、高於95%、或高於99%轉移性癌症風險。在各種實施例中，臨限評分係介於20%與80%之間。在一些實施例中，臨限值係介於30%與70%之間。在一些實施例中，臨限值係介於40%與60%之間。在一些實施例中，臨限值係介於45%與55%之間。在一些實施例中，臨限值係介於20%與60%之間。在一些實施例中，臨限值係介於30%與55%之間。在一些實施例中，臨限值係介於35%與50%之間。在一些實施例中，臨限值係介於40%與45%之間。

【0090】 在各種實施例中，若對象之所預測癌症期別係至少2期（例如，2期或3期），則該對象處於高轉移性癌症風險。在此類實施例中，導引介入涉及對象之後續檢測、對象之後續活體組織切片、或對象之後續影像掃描中之一者。此允許臨床提供者對最高風險之對象排定後續檢測、活體組織切片、或影像掃描的優先順序

【0091】 在各種實施例中，若對象經預測處於高轉移性癌症風險，則可為該對象排除對於後期轉移性癌症不適當的醫療程序。所排除的此類醫療程序之實例包括腫瘤切除、肺切除、放射（例如，外部光束放射）、化療（例如，局部或全身化療）、免疫治療、或甚至全身麻醉。

【0092】 在具體實施例中，若對象經預測為處於高轉移性癌症風險，則該對象接受後續影像掃描以確認預測。例如，若對象經預測為處於高轉移性癌症風險，則該對象進一步接受PET掃描。在此類實施例中，若PET掃描回報陰性結果（例如，無淋巴結侵犯），則該對象可接受進一步檢測。例如，該對象可進一步接受後續活體組織切片以明確地採樣淋巴結。在一些實施例中，後續活體組織切片係淋巴結的支氣管鏡採樣。

【0093】 在具體實施例中，若對象經預測為處於高轉移性癌症風險，則該對象接受後續活體組織切片以確認預測。例如，該對象不需要接受後續影像掃描以確認預測，而是替代地直接接受後續活體組織切片。例如，若對象經預測為處於高轉移性癌症風險，則該對象進一步接受淋巴結的支氣管鏡採樣。

【0094】 如本文所述，轉移性癌症之對象水準風險可從跨淋巴結的個體淋巴結水準風險來判定。因此，對於基於對象水準風險而經預測具有高轉移性癌症風險的對象，可識別導致高對象水準風險的（多個）特定淋巴結。因此，該等特定淋巴結可係後續檢測或治療的目標。在具體實施例中，該等特定淋巴結可接受後續影像掃描（例如，PET掃描）以確認預測。在具體實施例中，該等特定淋巴結可接受後續活體組織切片以明確地採樣特定淋巴結。因此，後續活體組織切片可確認在該等特定淋巴結中的轉移性癌症之存在或不存在。

【0095】 在各種實施例中，若對象經預測為處於高轉移性癌症風險，則該對象可納入臨床試驗。例如，此一臨床試驗可係檢測對於轉移性癌症之治療療效，且因此，其可在其合格標準中指定納入者將展現出轉移性癌症。在各種實施例中，若對象經預測為處於高轉移性癌症風險，則該對象被中止納入臨床試驗。例如，此一臨床試驗可檢測對於一種非轉移性癌症的治療療效。

【0096】 在各種實施例中，針對對象所預測的轉移性癌症之對象水準風險指示該對象處於低轉移性癌症風險。在各種實施例中，處於低轉移性癌症風險的對象具有低於一臨限評分之所預測對象水準風險。在各種實施例中，臨限評分低於20%轉移性癌症風險。在各種實施例中，臨限評分低於25%轉移性癌症風險。在各種實施例中，該臨限評分低於30%、低於35%、低於40%、低於45%、低於50%、低於55%、低於60%、低於65%、低於70%、低於75%、低於80%、低於85%、低於90%、低於95%、或低於99%轉移性癌症風險。在一些實施例中，臨限值係介於5%與40%之間。在一些實施例中，臨限值係介於10%與30%之間。在一些實施例中，臨限值係介於15%與25%之間。在一些實施例中，臨限值係介於18%與22%之間。在各種實施例中，若對象之所預測癌症期別低於2期（例如，0期或1期），則該對象處於低轉移性癌症風險。在此類實施例中，導引介入涉及以下中之一者：腫瘤切除、支氣管鏡診斷、選擇及/或投予治療劑、及/或選擇/投予醫藥組成物。此實現快速識別及治療尚未發生轉移的對象。

【0097】 在各種實施例中，可基於轉移性癌症之對象水準風險來選擇及/或投予治療劑至對象，該經選擇之治療劑可能展現出對該癌症的療效。例示性治療劑包括化療、放射、抗原特異性單株抗體、或免疫療法。在各種實施例

中，提供治療劑給處於高轉移性癌症風險的對象。在各種實施例中，處於高轉移性癌症風險的對象具有高於一臨限評分的轉移性癌症之所預測對象水準風險。在各種實施例中，臨限評分高於20%轉移性癌症風險。在各種實施例中，臨限評分高於25%轉移性癌症風險。在各種實施例中，臨限評分高於30%、高於35%、高於40%、高於45%、高於50%、高於55%、高於60%、高於65%、高於70%、高於75%、高於80%、高於85%、高於90%、高於95%、或高於99%轉移性癌症風險。在各種實施例中，若對象之所預測癌症期別係至少2期（例如，2期、3期、或4期），則該對象處於高轉移性癌症風險。

【0098】 在各種實施例中，提供治療劑給處於低轉移性癌症風險的對象。在各種實施例中，處於低轉移性癌症風險的對象具有低於一臨限評分之所預測對象水準風險。在各種實施例中，臨限評分低於20%轉移性癌症風險。在各種實施例中，臨限評分低於25%轉移性癌症風險。在各種實施例中，該臨限評分低於30%、低於35%、低於40%、低於45%、低於50%、低於55%、低於60%、低於65%、低於70%、低於75%、低於80%、低於85%、低於90%、低於95%、或低於99%轉移性癌症風險。在各種實施例中，若對象之所預測癌症期別低於2期（例如，0期或1期），則該對象處於低轉移性癌症風險。

【0099】 在各種實施例中，若對象經預測為處於低轉移性癌症風險，則該對象可被導引到早期癌症治療。早期癌症治療之實例包括腫瘤切除、肺切除、化療（例如，局部或全身化療）、局部遞送治療（例如，支氣管針滴注(transbronchial needle instillation)或心臟電燒術(catheter based ablation)）、或放射。

【0100】 在具體實施例中，若對象經預測為處於低轉移性癌症風險，則該對象接受後續影像掃描以確認預測。例如，若對象經預測為處於低轉移性癌症風險，則該對象進一步接受PET掃描。在此類實施例中，若PET掃描回報陰性結果（例如，無淋巴結侵犯），則PET掃描結果符合預測。然而，若PET掃描回報陽性結果（例如，淋巴結侵犯），則該對象可接受進一步檢測。例如，該對象可進一步接受後續活體組織切片以明確地採樣淋巴結。在一些實施例中，後續活體組織切片係淋巴結的支氣管鏡採樣。

【0101】 在各種實施例中，治療劑係生物製劑，例如細胞介素、抗體、可溶性細胞介素受體、反義寡核苷酸、siRNA等。此類生物劑涵蓋生物劑之突變蛋白及衍生物，該衍生物可包括例如所屬技術領域中已知之融合蛋白、聚乙二醇化衍生物、膽固醇共軛衍生物、及類似物。亦包括細胞介素及細胞介素受體之拮抗劑，例如誘捕劑及單株拮抗劑。亦包括本文所述之活性劑之生物類似或生物等效藥物。

【0102】 用於肺癌的治療劑可包括化學治療劑，諸如多西紫杉醇、順鉑、卡鉑、吉西他濱(gemcitabine)、白蛋白結合型太平洋紫杉醇(Nab-paclitaxel)、太平洋紫杉醇、培美曲塞(pemetrexed)、吉非替尼(gefitinib)、埃羅替尼(erlotinib)、布格替尼(brigatinib) (Alunbrig[®])、卡馬替尼(capmatinib) (Tabrecta[®])、賽爾帕替尼(selpercatinib) (Retevmo[®])、恩曲替尼(entrectinib) (Rozlytrek[®])、洛拉替尼(lorlatinib) (Lorbrena[®])、拉羅替尼(larotrectinib) (Vitrakvi[®])、達可替尼(dacomitinib) (Vizimpro[®])、及長春瑞濱(vinorelbine)。用於肺癌的治療劑可包括抗體療法，諸如德瓦魯單抗(durvalumab) (Imfinzi[®])、尼沃

魯單抗(nivolumab) (Opdivo[®])、派姆單抗(pembrolizumab) (Keytruda[®])、阿特珠單抗(atezolizumab) (Tecentriq[®])、及雷莫蘆單抗(ramucirumab)。

【0103】 在各種實施例中，可基於轉移性癌症之對象水準風險來選擇及/或投予醫藥組成物至對象，該經選擇之治療劑可能展現出對該癌症的療效。向個體投予之醫藥組成物包括活性劑，諸如上述治療劑。活性成分係以治療有效量存在，亦即當經投予時足以治療疾病或由其所媒介之醫療病況的量。組成物亦可包括各種其他劑以增強遞送及療效，例如增強活性成分的遞送及穩定性。因此，例如，組成物取決於所欲之配方，亦可包括醫藥上可接受、無毒性載劑或稀釋劑，其定義為常用於調配用於動物或人類投予之醫藥組成物的媒劑。稀釋劑係經選擇以不影響該組合之生物活性。此類稀釋劑之實例係蒸餾水、緩衝水、生理食鹽水、PBS、林格氏(Ringer' s)液、右旋糖溶液、及Hank氏溶液。此外，醫藥組成物或配方可包括其他載劑、佐劑、或毒性、非治療性、非免疫原性穩定劑、賦形劑、及類似物。組成物亦可包括額外物質以近似生理條件，諸如pH調整及緩衝劑、毒性調整劑、潤濕劑、及清潔劑。組成物亦可包括多種穩定劑之任一者，諸如抗氧化劑。

【0104】 在本文中描述之醫藥組成物可以各種不同方式投予。實例包括經由口服、鼻內、直腸、外用、腹膜內、靜脈、肌肉、皮下、表皮下、經皮、鞘內或顱內方法投予含有醫藥上可接受之載劑的組成物。在各種實施例中，醫藥組成物可經由基於導管之部署（例如，使用可撓性或剛性機器性支氣管鏡檢查）來投予。

【0105】 可投予此類醫藥組成物以用於治療（例如，診斷患者患有肺癌之後）之目的。如在本發明的上下文中針對疾病或病症所使用的預防、疾病預

防、或預防措施係指投予組成物以預防肺癌或肺癌之所有症狀的發生、發作、惡化、或復發，或減輕肺癌發作的可能性。根據本發明，肺癌的治療(treating/treatment)或療法(therapy)應意指藉由投予治療來減緩、停止、或反轉癌症的惡化。在較佳實施例中，治療肺癌係指反轉癌症的惡化，理想的程度為至消除癌症本身的程度。

VII. 電腦實施方案

【0106】 本發明之方法（包括用於藉由判定對象的轉移性癌症之對象水準風險來進行癌症分期之方法）在一些實施例中係在一或多個電腦上執行。

【0107】 例如，可在硬體或軟體或兩者的組合中實施推算模型及/或風險模型以及資料庫儲存的建置及部署。在本發明之一實施例中，提供機器可讀取儲存媒體，該媒體包含經編碼有機器可讀取資料的資料儲存材料，當使用經指令程式化以使用該資料之機器時，該機器可讀取資料能夠顯示資料集以及本文所述之模型的執行及結果中之任何者。本發明可在執行於可程式化電腦上之電腦程式中實施，該等可程式化電腦包含處理器、資料儲存系統（包括揮發性及非揮發性記憶體及/或儲存元件）、圖形配接器、指向裝置、網路配接器、至少一個輸入裝置、及至少一個輸出裝置。顯示器係耦接至圖形配接器。程式碼經應用於輸入資料，以執行上文描述之功能並產生輸出資訊。輸出資訊係以已知方式應用至一或多個輸出裝置。電腦可為例如個人電腦、微電腦、或習知設計的工作站。

【0108】 各程式可實施於高階程序或物件導向程式設計語言以與電腦系統溝通。然而，若需要，程式可以組合或機器語言實施。在任何情況下，語言

可為編譯或解譯語言。各此類電腦程式較佳係儲存於可由一般或特殊目的可程式電腦所讀取之儲存媒體或裝置（例如，**ROM**或磁碟），用於當儲存媒體或裝置由電腦讀取以執行本文所述之程序時組態及操作電腦。系統亦可視為實施為經電腦程式組態之電腦可讀取儲存媒體，其中儲存媒體經組態造成電腦以特定及預定方式操作以執行本文所述之功能。

【0109】 簽章模式及其資料庫可提供於多種媒體以利其等之使用。「媒體(media)」係指含有本發明之簽章模式資訊的製品。本發明之資料庫可記錄於電腦可讀取媒體上，例如可由電腦直接讀取及存取的任何媒體。此類媒體包括但不限於：磁性儲存媒體，諸如軟碟、硬碟儲存媒體、及磁帶；光學儲存媒體，諸如**CD-ROM**；電性儲存媒體，諸如**RAM**及**ROM**；及這些類別的混合體，諸如磁性/光學儲存媒體。所屬技術領域中具有通常知識者可輕易明白可如何使用任何目前已知的電腦可讀取媒體以產生包含本資料庫資訊之記錄的製品。「記錄(recorded)」係指使用所屬技術領域中已知的任何此類方法將資訊儲存在電腦可讀取媒體上之過程。可基於用以存取所儲存資訊之構件，選擇任何便利的資料儲存結構。多種資料處理器程式及格式可用於儲存（例如文字處理文字檔）、資料庫格式等。

【0110】 在一些實施例中，本發明之方法（包括用於藉由判定對象的轉移性癌症之對象水準風險來進行癌症分期之方法）係在分散式運算系統環境（例如，雲端運算環境）中之一或多個電腦上執行。在此描述中，「雲端運算(cloud computing)」係定義為用於實現對於一共享之可組態運算資源集之隨選網路存取的模型。可採用雲端運算來對該共享之可組態運算資源集提供隨選存取。該共享之可組態運算資源集可經由虛擬化來快速提供，且以低管理努力或

服務提供者互動來釋放，並接著相應地按比例調整。雲端運算模型可由各種特性構成，諸如例如隨選自我服務、廣泛網路存取、資源集區、快速彈性、測量服務等。雲端運算模型亦可暴露各種服務模型，諸如例如軟體及服務（「SaaS」）、平台即服務（「PaaS」）、及基礎設施即服務（「IaaS」）。雲端運算模型亦可使用不同的部署模型（諸如，私有雲端、社群雲端、公用雲端、混合雲端等）來部署。在本說明書及申請專利範圍中，「雲端運算環境 (cloud-computing environment)」係採用雲端運算的環境。

【0111】 圖4繪示一實例電腦，其用於實施圖1A、圖1B、圖2A、圖2B、圖3A、及圖3B所示之實體。電腦400包括至少一個處理器402，其耦接至晶片組404。晶片組404包括記憶體控制器中樞420及輸入/輸出(I/O)控制器中樞422。記憶體406及圖形配接器412係耦接至記憶體控制器中樞420，且顯示器418係耦接至圖形配接器412。儲存裝置408、輸入裝置414、及網路配接器416係耦接至I/O控制器中樞422。電腦400的其他實施例具有不同架構。

【0112】 儲存裝置408係非暫態電腦可讀取儲存媒體，諸如硬碟、唯讀光碟記憶體(CD-ROM)、DVD、或固態記憶體裝置。記憶體406保存由處理器402使用之指令及資料。輸入介面414係觸控螢幕介面、滑鼠、軌道球、或其他類型之指向裝置、鍵盤、或其一些組合，且係用於輸入資料至電腦400中。在一些實施例中，電腦400可經組態以接受來自輸入介面414之經由使用者姿勢之輸入（例如，命令）。圖形配接器412顯示影像及其他資訊在顯示器418上。網路配接器416將電腦400耦接至一或多個電腦網路。

【0113】 電腦400經調適以執行用於提供本文所述功能的電腦程式模組。如本文中所使用，用語「模組(module)」係指用於提供指定功能的電腦程式邏

輯。因此，可在硬體、韌體、及/或軟體中實施模組。在一實施例中，程式模組係儲存在儲存裝置408上、裝載至記憶體406中、並由處理器402執行。

【0114】 由圖1A或圖1B之實體所使用的電腦400之類型可取決於實施例及實體所需之處理能力而變化。例如，癌症分期系統130可在單一電腦400中或在透過網路（諸如在伺服器場中）彼此溝通之多個電腦400中運行。電腦400可缺乏上述之一些組分，諸如圖形配接器412、及顯示器418。

VIII. 系統

【0115】 本文進一步揭示用於藉由判定一對象的轉移性癌症之一對象水準風險來進行癌症分期的系統。在各種實施例中，此一系統可至少包括上文在圖1A所述之癌症分期系統130。在各種實施例中，癌症分期系統130體現為電腦系統，諸如具有圖4所述之實例電腦400的電腦系統。

【0116】 在各種實施例中，系統包括成像裝置，諸如上文在圖1A所述之成像裝置120。在各種實施例中，系統包括癌症分期系統130（例如，電腦系統）及成像裝置兩者。在此類實施例中，癌症分期系統130可與成像裝置120可溝通地耦接，以接受自一對象擷取之影像（例如，CT掃描）。電腦系統於電腦上(*in silico*)實施推算模型與風險模型兩者中之一或兩者，以分析影像及判定對象之淋巴結疾病之對象水準風險。

實例

【0117】 以下係用於實行本發明之具體實施例之實例。實例的提供僅為了說明目的而非意欲以任何方式限制本發明之範疇。已致力於確保關於所使用數字（例如，量、溫度等）之精確性，但應允許一些實驗誤差及偏差。

實例1：研究族群及特徵提取

【0118】 此研究係使用來自國家肺部篩檢試驗(NLST)之CT組的資料來執行。此NLST之次世代包括來自658位研究參與者的資料，301位具有良性結節，而357位患有肺癌。658名研究參與者的各CT掃描均經過檢查，且高達21個淋巴結站（各人16+/-6）經標註（9,969個淋巴結經標註；4,631個淋巴結來自對照組，及5,318個淋巴結來自肺癌患者）。從此初步樣本，患有肺癌的288人具有可用於分析的分期資料。為了建立平衡的初始世代，識別出287名對照對象。淋巴結站係描繪在圖2A中且詳述於表1中。

【0119】
〔表1〕：各站的淋巴結標註

淋巴站	淋巴結數量（每對象之平均）	
	對照組	癌症組
1L	34	22
1R	29	13
2L	191	186
2R	732	879
3A	203	263
3P	133	139
4L	430	459
4R	837	987
5	386	457
6	189	132

7	477	541
8L	212	230
8R	140	215
9L	29	98
9R	8	14
10L	181	188
10R	92	60
11L	74	106
11R	57	59
12L	5	4
12R	5	8

【0120】 針對各淋巴結站，界定具有5 mm及7.5 mm半徑的一球形區域。由標註標籤所界定之球面中心。圖2B展示兩個淋巴結標註之一實例。

【0121】 這些較小(5 mm)及較大(7.5 mm)的球面ROI（關注區域）係用於探索用於模型建置之影像範圍。從這些球面ROI，提取3個分開的影像圖譜板，各有6種特徵類別。各板所使用的度量數目不同，且此係基於CT的不同變換。板1使用原始影像作為特徵提取之輸入。板2及3分別採用輸入影像之小波變換及高斯拉普拉斯(LoG)變換。特徵類別包括一階特徵、形狀特徵、灰階共生矩陣1 (GLCM)特徵、灰階級長矩陣(GLRLM)特徵、灰階級帶矩陣(GLSZM)特徵、及鄰域灰階差分矩陣(NGTDM)特徵。表2展示特徵細項。

【0122】

〔表2〕：來自CT掃描影像的特徵

特徵數目						
特徵類別						
影像類別	一階	形狀*	GLCM	GLRLM	GLSZM	NGTDM
原始	22	14	24	16	16	5

小波	144	0	192	128	128	40
LoG	18	0	24	16	16	5

*由於使用固定的球形形狀區域分析各站，所以形狀特徵不包括在淋巴結分析中

【0123】 圖5A至圖5R描繪跨GLC特徵類別之不同板的不同淋巴結影像圖譜特徵的實例相關性熱圖。此外，圖5A至圖5R各展示特定特徵。圖5A至圖5R中展示之此類特徵的實例包括相關特徵、LMC2、MCC、差熵、聯合熵、和熵、群聚突出、平均差、聯合平均、平均加總、自動相關度、變異性差別、對比、群聚傾向、差值平方和、LDMN、LDN、群聚陰暗度、LMC1、逆方差、LD、LDM、聯合能量、及最大機率。圖5C至圖5R展示相對於不同濾波器之特徵。具體而言，圖5C至圖5R提供「XYZ」之命名，其中「X」、「Y」、及「Z」各係指應用至影像的X、Y、或Z方向的濾波器。「H」係指高通濾波器，而「L」係指低通濾波器。因此，「HHH」之命名係指從在X、Y、及Z方向之各者中經受高通濾波的3D影像所導出的特徵。「HHL」之命名係指從在X及Y方向上經受高通濾波，且在Z方向經受低通濾波的特3D影像所導出的特徵。

實例2：用於癌症分期的推算模型

【0124】 模型化始於預測各淋巴結之癌症侵犯。若要這樣做，對各標註之淋巴結台推算（NLST中報告的）癌症期別（NLST中未報告個體淋巴結資料，所以將淋巴結侵犯從癌症期別「排除」）。接著將訓練:測試資料進行50:50分割，其具有表2中所提供之資料分布。

【0125】

〔表2〕：實例的50:50資料分割

淋巴結轉移期數	訓練		測試	
	對照組	癌症組	對照組	癌症組
0	2337	2179	2263	1713
1	0	127	0	73
2	0	125	0	342
3	0	40	0	42

【0126】 該模型之結果係二進位（使用NLST資料集中的淋巴結轉移期數共變異數，淋巴轉移「是/否」。「否」係N=0或N=1，且「是」係N=2或N=3）。將度量正規化（z-評分），並訓練及測試隨機森林及LASSO分類器。針對各淋巴結，產生其為癌性的連續機率。例如，淋巴結可能有20%的機率為癌性（藉由減法，意指其被指派80%的機率為無癌）。整個研究參與者內的所有淋巴結皆視為獨立資料。圖6展示初始模型之效能。具體而言，該模型達成曲線下方面積(AUC) = 65.3%、準確度(ACC) = 96.6%、靈敏度= 61.2%、及特異度= 60.8%。

【0127】 對NLST資料的限制在於淋巴結特異度資訊無法取得，且平均16個淋巴結被標註，且每個人接受基於影像圖譜之評估。因此，此模型假設在患有轉移性癌症之患者內的所有淋巴結為「+」，而在臨床醫學上，一般只有1個淋巴結具有癌細胞（癌症不會同時擴散至所有淋巴結，但通常擴散至僅一個或兩個淋巴結）。因此，從對象水準癌症期別推算個體淋巴結侵犯引入了偽陽性，該等偽陽性混淆了淋巴結水準的模型效能（同樣因為在人的CT掃描內的所有淋巴結係假設為「+」）。

【0128】 然後，針對給定對象之所有淋巴結站，使用由淋巴結水準分類器所預測的最高機率評分來判定淋巴結轉移期數可能性。圖7A至圖7B繪示基於淋巴結轉移期數可能性的對象水準效能（AUC = 69.1%、靈敏度= 72.7%、及特異度= 61.6%）。亦將評分（例如，最大機率）的分布橫跨淋巴結轉移期數來繪製，如在圖7B中所示。在此，針對具有3期癌症（左圖）的情況，觀察到較高的最大機率。

實例3：遞增性淋巴結分類方法：對患者內淋巴結進行拔靴抽樣

【0129】 執行兩步驟式拔靴法以從轉移性疾病之淋巴結水準風險轉變至轉移性疾病之對象水準風險。具體而言，兩步驟式拔靴法包括實施推算模型及/或風險模型，如上文參考圖3A所述。在此，挑戰在於：對於所有淋巴結站推算對象的淋巴結轉移期數，以訓練分類器。鑑於大多數淋巴結係無癌的，此推算在淋巴結站引入錯誤。

【0130】 兩個步驟如下：

1. 推算：對於各對象，很可能被癌症侵犯的淋巴結別為具有已知轉移性疾病。此係針對各對象藉由選擇具有高於中位數對象內值（涉及圖 3A 中所示之推算模型的推算程序）的為癌性之機率（從最後模型輸出）的淋巴結來完成。低於具有癌症之中位數對象內機率的淋巴結被分類為陰性。
2. 分類：世代被分割為 50:50（訓練:測試）。接著，使用淋巴結結果之推算集區來訓練第二分類器（例如，圖 3A 之風險模型 320）以預測各對象之淋巴結疾病的存在。

【0131】 此模型之輸出係各淋巴結為癌性之連續機率，其類似於上文實例2中所述之模型之輸出。然而，額外推算及分類步驟的優點在於：具有轉移性疾病的對象中所有淋巴結現在經指派有較大的機率範圍。更具體而言，可能具有癌症的淋巴結具有較高之為「+」的機率，而這些不太可能具有癌症的淋巴結具有較低之為「+」的機率。對於預測步驟，基於具有最高機率（圖3B）的淋巴結而產生淋巴結侵犯的對象水準評分。具體而言，此涉及：採用淋巴結特定機率，並提供具有其肺癌之淋巴結侵犯的對象特定機率。在圖3B所示之檢測對象中，底部淋巴結被指派有60%機會患有癌症（例如，機率(+) = 0.6）。此係三個淋巴結之中具有最高「+」機率之淋巴結。因此，對象水準機率被指派0.6的值。

【0132】 在使用隨機森林分類器訓練該遞增方法之後，拔靴抽樣模型之對象水準效能達成AUC = 77.7%、靈敏度= 81.8%、及特異度= 59.9%，如在圖9A中所繪示。因此，相較於實例2中所述之方法，兩步驟式拔靴抽樣模型的效能獲得改善。使用LASSO分類器達成類似的效能。最後，該模型可用於繪製橫跨淋巴結轉移期數（圖9B）之評分（或最大機率）的分布。有趣的是，注意到，對於具有較高淋巴結轉移期數的案例，模型產出較高的最大機率。此機率評分可用於不同的臨床背景情境中，以實現不同的檢傷(triage)決策工具。用於訓練及測試之額外資料可提供預測N作為分類而非二進位共變異數的模型。

【0133】 所有在此申請案中引用之所有公開案、專利、專利申請案及其他文件全文皆特此為所有目的以引用方式併入本文中，如同各個別公開案、專利、專利申請案或其他文件係個別地指示為所有目的以引用方式併入本文中。

【0134】 雖然已經說明及描述各種具體實施例，但上述說明書不具限制性。將理解的是，在不偏離（多個）本揭露之精神和範圍下，可做出各種變化。許多變異在所屬技術領域中具有通常知識者檢視本說明書後將變得顯而易見。

實例4：使用以CT影像為基礎之預測裝置的實例導引介入

【0135】 實施本文中描述之方法（例如，部署用於預測淋巴結(LN)之轉移性癌症的淋巴結水準風險的風險模型）的以CT影像為基礎之裝置將影響臨床照護，此係因為涉及PET掃描的習知方法之不良效能。最近檢查接受系統性淋巴結切除/採樣的219名NSCLC患者檢查術前PET分期的效能（例如，Li等人的「Implications of False Negative and False Positive Diagnosis in Lymph Node Staging of NSCLC by Means of 18F-FDG PET/CT. PLoS ONE 8(10):e78552」，其全部內容以引用方式併入本文中）。用於偵測肺門及縱隔腔淋巴結之PET的靈敏度、特異度、PPV、及NPV係74.2%、54.4%、86.8%、及73.5%，其中假陰性率係13.2%及偽陽性率係45.5%。因此，本文所述之方法表示用於偵測已導致淋巴結侵犯之轉移性疾病的非侵入性技術，其將改善決策及臨床照護。此裝置將影響決策及臨床照護的實例包括：

【0136】 促使對患有大小屬轉移低風險之腫瘤的對象進行額外檢測。由於對於患有小肺腫瘤者使用術前PET掃描或淋巴結活體組織切片沒有注意標準，所以基於CT裝置所擷取之CT影像預測淋巴結侵犯促使經由對象之淋巴結的PET掃描或支氣管鏡採樣進行後續檢測。相較之下，習知方法將導引對象進行手術，且其中更惡化的疾病期將僅被認為要進行外科手術。

【0137】 促使對獲得陰性PET結果者的淋巴結進行以組織為基礎的採樣。雖然普遍認識到PET具有偽陽性及假陰性失敗率，但是診斷檢測受到PET掃描結果強烈的影響。患有較小腫瘤大小及陰性PET的患者經常進行手術而不會進行額外的淋巴結採樣。此患者之子集會受益於在手術前進行額外的淋巴結採樣。

【0138】 對具有偽陽性PET掃描的患者進行分期，並示範其等對手術介入的合格性。目前，基於PET掃描而強烈疑似患有惡化期別疾病的患者會進行非治癒性全身性/非手術療法，其等不含額外的確認性淋巴結採樣。接收淋巴結侵犯之預測（例如，基於淋巴結的CT影像）的患者及陰性PET結果的患者將替代地在決定手術之前進行更明確的淋巴結採樣。因此，這些患者可為外科切除的候選者，並且會增加其治癒機會。

【0139】 告知哪個淋巴結最可能具有轉移性疾病。處於高轉移性疾病風險的患者進行淋巴結的支氣管鏡或手術採樣。一般而言，此策略會強度聚焦在靠近腫瘤的淋巴結，而使更遠端淋巴結未被採樣或欠採樣。在此，本文所述之方法實現淋巴結侵犯之預測，並且此外，識別被侵犯的淋巴結。此導引臨床團隊適當地聚焦其等之採樣策略以獲得組織、最大化產率、及最小化誤診的風險。

【0140】 精進臨床試驗以著重於納入無淋巴結轉移的患者。患有局部疾病的患者通常尋求臨床試驗對於局部遞送的臨床試驗（化療之支氣管針滴注或燒術）的臨床試驗。不幸地，經受淋巴結採樣時，此類患者在手術期間經常發現具有更惡化期癌症。此延遲了適當的照護。因此，實施本文中描述之方法（例如，部署用於預測淋巴結(LN)之轉移性癌症的淋巴結水準風險的風險模

型)實現識別展現出淋巴結侵犯的患者，且因此，這些患者不需要被納入著重於無淋巴結轉移之患者的臨床試驗。

【符號說明】

【0141】

100:系統環境概觀；系統環境

110:對象

120:成像裝置

130:癌症分期系統

140:癌症預測

145:淋巴結特徵模組

150:推算模組

155:風險訓練模組

160:風險部署模組

165:癌症預測模組；癌症分期模組

170:訓練資料儲存

305A:參考個體

305B:參考個體

305C:個體

305D:個體

310:推算模型

315:對象；個體

315A:個體

315B:個體

315C:個體

315D:個體

320:風險模型

350:對象

400:電腦系統；電腦

402:處理器

404:晶片組

406:記憶體

408:儲存裝置

412:圖形配接器

414:輸入裝置；輸入介面

416:網路配接器

418:顯示器

420:記憶體控制器中樞

422:輸入/輸出(I/O)控制器中樞

【生物材料寄存】

無