

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2011-516046

(P2011-516046A)

(43) 公表日 平成23年5月26日(2011.5.26)

(51) Int.Cl.	F I	テーマコード (参考)
C 1 2 Q 1/68 (2006.01)	C 1 2 Q 1/68 A	4 B 0 2 4
G O 1 N 33/574 (2006.01)	G O 1 N 33/574 A	4 B 0 6 3
C 1 2 N 15/113 (2010.01)	C 1 2 N 15/00 G	

審査請求 未請求 予備審査請求 未請求 (全 19 頁)

(21) 出願番号	特願2011-502136 (P2011-502136)	(71) 出願人	301069856 トラスティーズ オブ ボストン ユニバーシティ アメリカ合衆国 マサチューセッツ 02215, ボストン, ワン シルバー ウェイ
(86) (22) 出願日	平成21年3月30日 (2009. 3. 30)	(71) 出願人	510257776 アレグロ ダイアグノスティックス, インコーポレイテッド アメリカ合衆国 マサチューセッツ 02218, ボストン, アルバニー ストリート 600
(85) 翻訳文提出日	平成22年11月24日 (2010. 11. 24)	(74) 代理人	100078282 弁理士 山本 秀策
(86) 国際出願番号	PCT/US2009/038836		
(87) 国際公開番号	W02009/121070		
(87) 国際公開日	平成21年10月1日 (2009. 10. 1)		
(31) 優先権主張番号	61/040, 434		
(32) 優先日	平成20年3月28日 (2008. 3. 28)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 肺障害の検出のための多元的方法

(57) 【要約】

本明細書に記載するのは、肺障害または疾患、例えば肺癌を検出し、診断し、またはその診断を援助するための多元的方法である。本発明の方法は、複数の診断パラダイムを利用して、例えばそれぞれのパラダイム単独より診断感度、特異度、陰性的中率および/または陽性的中率を向上させる。例えば、臨床的因子と遺伝子発現、特に感度および特異的遺伝子発現バイオマーカーを組み合わせた、肺癌診断の臨床ゲノムモデルを開示する。好ましい実施形態では、2つ以上の肺癌関連の診断パラダイムは、2つ以上のパラダイムの少なくとも1つのみを超える（例えば、2つ以上のパラダイムのいずれかのみを超える）特異度、陽性的中率、陰性的中率および/または感度をもたらす。

【特許請求の範囲】**【請求項 1】**

肺疾患を有する疑いがある患者における肺疾患の診断を援助する方法であって、
評価される患者における 2 つ以上の独立した肺癌関連の診断パラダイムを分析する工程と、該患者の複合的分類を肺疾患を有するかまたは肺疾患を有しないと判定する工程を含む方法。

【請求項 2】

前記肺疾患が肺癌である、請求項 1 に記載の方法。

【請求項 3】

前記患者が喫煙者または元喫煙者である、請求項 1 に記載の方法。

10

【請求項 4】

前記患者が、異常な X 線写真所見、または診断不能な気管支鏡検査を有していた、請求項 1 に記載の方法。

【請求項 5】

前記 2 つ以上の肺癌関連の診断パラダイムが、前記患者における 1 つまたは複数の肺癌関連遺伝子の発現の分析、該患者の 1 つまたは複数の肺癌関連の臨床的因子または変数の分析、該患者の血液中の 1 つまたは複数の肺癌関連抗体の有無に関する試験、該患者の血液中の 1 つまたは複数の肺癌関連タンパク質の有無に関する試験、および 1 つまたは複数の肺癌関連マイクロ RNA の発現の分析からなる群から選択される、請求項 1 に記載の方法。

20

【請求項 6】

前記 2 つ以上の肺癌関連の診断パラダイムが、前記患者における 1 つまたは複数の肺癌関連遺伝子の発現の分析を含む、請求項 1 に記載の方法。

【請求項 7】

前記 2 つ以上の肺癌関連の診断パラダイムが、前記患者における 1 つまたは複数の肺癌関連遺伝子の発現の分析、ならびに該患者の 1 つまたは複数の肺癌関連の臨床的因子または変数の分析、該患者の血液中の 1 つまたは複数の肺癌関連抗体の有無に関する試験、該患者の血液中の 1 つまたは複数の肺癌関連タンパク質の有無に関する試験、および 1 つまたは複数の肺癌関連マイクロ RNA の発現の分析からなる群から選択される 1 つまたは複数の肺癌関連の診断パラダイムを含む、請求項 1 に記載の方法。

30

【請求項 8】

前記 1 つまたは複数の肺癌関連遺伝子が、その発現データが遺伝子発現オムニバスのアクセッション番号 G S E 4 1 1 5 に含有される遺伝子の全部またはサブセットである、請求項 6 に記載の方法。

【請求項 9】

前記 1 つまたは複数の肺癌関連遺伝子が、その発現データが遺伝子発現オムニバスのアクセッション番号 G S E 4 1 1 5 に含有される遺伝子の全部またはサブセットである、請求項 7 に記載の方法。

【請求項 10】

前記 2 つ以上の肺癌関連の診断パラダイムが、前記患者における 1 つまたは複数の肺癌関連遺伝子の発現の分析、および該患者の 1 つまたは複数の肺癌関連の臨床的因子または変数の分析を含む、請求項 1 に記載の方法。

40

【請求項 11】

前記 1 つまたは複数の肺癌関連遺伝子が、その発現データが遺伝子発現オムニバスのアクセッション番号 G S E 4 1 1 5 に含有される遺伝子の全部またはサブセットである、請求項 10 に記載の方法。

【請求項 12】

肺癌を有する疑いがある患者の追跡治療レジメンを決定する方法であって、
評価される患者における 2 つ以上の独立した肺癌関連の診断パラダイムを分析する工程と、

50

該分析に基づいて癌を有するかまたは癌を有しないと、該患者を分類する工程とを含み、癌を有すると分類した患者を、侵襲性試験および/または治療レジメンの開始について選択し、癌を有しないと分類した患者を、侵襲性試験または治療レジメンの開始を行わずにモニターする方法。

【請求項 13】

前記患者が喫煙者または元喫煙者である、請求項 12 に記載の方法。

【請求項 14】

前記患者が、異常な X 線写真所見、または診断不能な気管支鏡検査を有していた、請求項 12 に記載の方法。

【請求項 15】

前記 2 つ以上の肺癌関連の診断パラダイムが、前記患者における 1 つまたは複数の肺癌関連遺伝子の発現の分析、該患者の 1 つまたは複数の肺癌関連の臨床的因子または変数の分析、該患者の血液中の 1 つまたは複数の肺癌関連抗体の有無に関する試験、該患者の血液中の 1 つまたは複数の肺癌関連タンパク質の有無に関する試験、および 1 つまたは複数の肺癌関連マイクロ RNA の発現の分析からなる群から選択される、請求項 12 に記載の方法。

10

【請求項 16】

前記 2 つ以上の肺癌関連の診断パラダイムが、前記患者における 1 つまたは複数の肺癌関連遺伝子の発現の分析を含む、請求項 12 に記載の方法。

【請求項 17】

前記 2 つ以上の肺癌関連の診断パラダイムが、前記患者における 1 つまたは複数の肺癌関連遺伝子の発現の分析、ならびに該患者の 1 つまたは複数の肺癌関連の臨床的因子または変数の分析、該患者の血液中の 1 つまたは複数の肺癌関連抗体の有無に関する試験、該患者の血液中の 1 つまたは複数の肺癌関連タンパク質の有無に関する試験、および 1 つまたは複数の肺癌関連マイクロ RNA の発現の分析からなる群から選択される 1 つまたは複数の肺癌関連の診断パラダイムを含む、請求項 12 に記載の方法。

20

【請求項 18】

前記 1 つまたは複数の肺癌関連遺伝子が、その発現データが遺伝子発現オムニバスのアクセッション番号 G S E 4 1 1 5 に含有される遺伝子の全部またはサブセットである、請求項 16 に記載の方法。

30

【請求項 19】

前記 2 つ以上の肺癌関連の診断パラダイムが、前記患者における 1 つまたは複数の肺癌関連遺伝子の発現の分析、および該患者の 1 つまたは複数の肺癌関連の臨床的因子または変数の分析を含む、請求項 12 に記載の方法。

【請求項 20】

肺癌を有する疑いがある患者における肺癌の診断を援助する方法であって、該患者から生物サンプルを得、該サンプル中の 1 つまたは複数の肺癌関連遺伝子の発現を分析する工程であって、該 1 つまたは複数の肺癌関連遺伝子が、その発現データが遺伝子発現オムニバスのアクセッション番号 G S E 4 1 1 5 に含有される遺伝子の全部またはサブセットである、工程と、

40

該患者の 1 つまたは複数の肺癌関連の臨床的因子または変数を分析する工程と、

該患者の複合的分類を、癌を有するかまたは癌を有しないと判定する工程と

を含む方法。

【発明の詳細な説明】

【技術分野】

【0001】

(関連出願)

本願は、2008年5月28日に出願された、米国仮特許出願第61/040,434号の利益を主張し、この出願の全体の教示は、本明細書中に参考として援用される。

【0002】

50

(政府支援)

本発明は、国立衛生研究所(国立癌研究所)からの助成金 R 2 1 C A 1 0 6 5 0 6、および R 0 1 C A 1 2 4 6 4 0 により、その全てが、または一部が成された。合衆国政府は、本発明に一定の権利を有する。

【背景技術】

【0003】

肺癌は、部分的には初期診断ツールの欠如による、癌死の主な原因である。喫煙者は、異常な X 線写真所見および/または肺癌に特異的でない症状に基づいて、肺癌を有することがしばしば疑われる。ファイバー気管支鏡検査は、疾患の疑いがある喫煙者において比較的侵襲的な初期診断試験となり、被験部の気管支内ブラッシング、気管支肺胞洗浄、ならびに気管支内および経気管支生検によって得た材料の細胞学的検査を可能にする。残念ながらこの方法は、感度が比較的低い。コストが高く、危険があり、肺癌が疑わしい患者の診断評価を延長する追加のより侵襲的な診断試験が通常必要とされる。気管支鏡検査では癌陰性である肺癌が疑わしい患者の誰が、これらの追加の診断試験を受けるべきかの決定は、現在は臨床判断の問題である。したがって、肺癌などの肺疾患の診断に関する向上した感度ならびに陽性および陰性的中率をもたらす、追加の診断ツールを有することは有益となる。

10

【発明の概要】

【課題を解決するための手段】

【0004】

本明細書に記載する本発明は、肺障害または疾患、例えば肺癌を検出し、診断し、またはその診断を援助するための多元的方法に関する。本発明の方法は、複数(すなわち、2つ以上の)診断パラダイムを利用して、例えばそれぞれのパラダイム単独より診断感度、特異度、陰性的中率および/または陽性的中率を向上させる。好ましい実施形態では、診断パラダイムは互いに独立している。

20

【0005】

例えば、一実施形態では本発明は、臨床的因子と遺伝子発現、特に感度および特異的遺伝子発現バイオマーカーを組み合わせた、肺癌診断の臨床ゲノムモデルに関する。本明細書に記載する研究では、遺伝子発現バイオマーカー、臨床的因子、およびこれらのデータの組合せを使用して、肺癌の疑いで気管支鏡検査を受けた一連の喫煙者において癌の可能性を分析した(臨床ゲノムモデル)。臨床ゲノムモデルの成績の有意な差は、臨床的因子単独と比較して確認した。実際、臨床ゲノムモデルは感度および陰性的中率を100%まで増大させ、他のモデルと比較してより高い特異度および陽性的中率をもたらす。したがって、臨床ゲノムモデルの使用により、肺癌を有する個体にはより侵襲的な試験および確定的な治療法を促進することができ、肺癌を有さない個体には侵襲的な診断手順を減らすことができる。

30

【0006】

一実施形態では、本発明は、肺疾患を有する疑いがある患者における肺疾患の診断を援助する方法であって、評価する患者における2つ以上の独立した肺癌関連の診断パラダイムを分析する工程と、肺疾患を有するかまたは肺疾患を有しないと患者の複合的分類を判定する工程とを含む方法に関する。一実施形態では、肺疾患は肺癌である。一態様では、患者は喫煙者または元喫煙者である。別の態様では、患者は、異常な X 線写真所見または診断不能な気管支鏡検査を有していた。

40

【0007】

好ましい実施形態では、本発明は、2つ以上の肺癌関連の診断パラダイムが、患者における1つまたは複数の肺癌関連遺伝子の発現の分析、患者の1つまたは複数の肺癌関連の臨床的因子または変数の分析、患者の血液中の1つまたは複数の肺癌関連抗体の有無に関する試験、患者の血液中の1つまたは複数の肺癌関連タンパク質の有無に関する試験、および1つまたは複数の肺癌関連マイクロRNAの発現の分析からなる群から選択される方法に関する。別の態様では、2つ以上の肺癌関連の診断パラダイムは、患者における1つ

50

または複数の肺癌関連遺伝子の発現の分析を含む。一実施形態では、2つ以上の肺癌関連の診断パラダイムは、患者における1つまたは複数の肺癌関連遺伝子の発現の分析、ならびに患者の1つまたは複数の肺癌関連の臨床的因子または変数の分析、患者の血液中の1つまたは複数の肺癌関連抗体の有無に関する試験、患者の血液中の1つまたは複数の肺癌関連タンパク質の有無に関する試験、および1つまたは複数の肺癌関連マイクロRNAの発現の分析からなる群から選択される1つまたは複数の肺癌関連の診断パラダイムを含む。特定の実施形態では、1つまたは複数の肺癌関連遺伝子は、その発現データが遺伝子発現オムニバスのアクセッション番号GSE4115に含有される遺伝子の全部またはサブセットである。特定の実施形態では、2つ以上の肺癌関連の診断パラダイムは、患者にお

10

【0008】

本発明は、肺癌を有する疑いがある患者の追跡治療レジメンを決定する方法であって、評価する患者における2つ以上の独立した肺癌関連の診断パラダイムを分析する工程と、分析に基づいて癌を有するかまたは癌を有しないと患者を分類する工程とを含み、癌を有すると分類した患者を、侵襲性試験および/または治療レジメンの開始について選択し、癌を有しないと分類した患者を、侵襲性試験または治療レジメンの開始を行わずにモニターする方法にも関する。

【0009】

一実施形態では、患者は喫煙者または元喫煙者である。別の態様では、患者は、異常なX線写真所見または診断不能な気管支鏡検査を有していた。特定の実施形態では、2つ以上の肺癌関連の診断パラダイムは、患者における1つまたは複数の肺癌関連遺伝子の発現の分析、患者の1つまたは複数の肺癌関連の臨床的因子または変数の分析、患者の血液中の1つまたは複数の肺癌関連抗体の有無に関する試験、患者の血液中の1つまたは複数の肺癌関連タンパク質の有無に関する試験、および1つまたは複数の肺癌関連マイクロRNAの発現の分析からなる群から選択される。別の実施形態では、2つ以上の肺癌関連の診断パラダイムは、患者における1つまたは複数の肺癌関連遺伝子の発現の分析を含む。本発明の一態様では、2つ以上の肺癌関連の診断パラダイムは、患者における1つまたは複数の肺癌関連遺伝子の発現の分析、ならびに患者の1つまたは複数の肺癌関連の臨床的因子または変数の分析、患者の血液中の1つまたは複数の肺癌関連抗体の有無に関する試験、患者の血液中の1つまたは複数の肺癌関連タンパク質の有無に関する試験、および1つまたは複数の肺癌関連マイクロRNAの発現の分析からなる群から選択される1つまたは複数の肺癌関連の診断パラダイムを含む。本発明の特定の態様では、1つまたは複数の肺癌関連遺伝子が、その発現データが遺伝子発現オムニバスのアクセッション番号GSE4115に含有される遺伝子の全部またはサブセットである。本発明の別の実施形態では、2つ以上の肺癌関連の診断パラダイムは、患者における1つまたは複数の肺癌関連遺伝子の発現の分析、および患者の1つまたは複数の肺癌関連の臨床的因子または変数の分析を含む。

20

30

【0010】

本発明は、肺癌を有する疑いがある患者における肺癌の診断を援助する方法であって、患者から生物サンプルを得、サンプル中の1つまたは複数の肺癌関連遺伝子（これは、発現データが遺伝子発現オムニバスのアクセッション番号GSE4115に含有される遺伝子の全部またはサブセットである）の発現を分析する工程、と、患者の1つまたは複数の肺癌関連の臨床的因子または変数を分析する工程と、癌を有するかまたは癌を有しないと患者の複合的分類を決定する工程とを含む方法にも関する。

40

【0011】

好ましい実施形態では、2つ以上の肺癌関連の診断パラダイムは、2つ以上のパラダイムの少なくとも1つのみを超える（例えば、2つ以上のパラダイムのいずれかのみを超える）特異度、陽性的中率、陰性的中率および/または感度をもたらす。

【図面の簡単な説明】

50

【 0 0 1 2 】

【図 1】図 1 A および 1 B は、実施例中で使用した訓練および試験サンプルセットを示す図である。訓練および試験サンプルは、临床上の肺癌の疑いで気管支鏡検査を受けた現喫煙者および元喫煙者からの気道上皮遺伝子の発現をアッセイした、以前に公開された試験から抽出した。図 1 A は、77 名の患者の訓練セットを使用して肺癌の存在を予測する、以前に構築された遺伝子発現バイオマーカーを示す図である。実施例中に記載した試験用では、これらのサンプルの 1 つは不完全な喫煙歴のために除去し、76 名の患者からのデータで訓練したロジスティック回帰モデルをもたらした。これらのモデルは、細胞病理学的検査で肺癌と診断不能であった訓練サンプルのサブセットに対して後で試験した ($n = 56$)。図 1 B 中に示すように、バイオマーカーも、以前の試験中で使用された組合せ試験およびプロスペクティブパリエーションのサンプルセット ($n = 87$) からの、細胞病理検査で診断不能であった独立サンプルのサブセット ($n = 62$) に対して試験した。

【図 2】図 2 A ~ 2 C は、異なるサンプルセット全体の臨床モデルおよび臨床ゲノムモデルに関する ROC 曲線を示す図である。臨床モデル (赤線) は以下の変数: 年齢、塊サイズ、およびリンパ節症を含み、臨床およびバイオマーカーモデル (黒線) は前述の変数およびバイオマーカースコアを含む。訓練セットサンプル ($n = 76$) を使用して両方のモデルを誘導した。図 2 A は、診断が不十分であった訓練セットサンプル ($n = 56$) の ROC 分析を示す図である。臨床および臨床ゲノムモデルに関する濃度曲線下面積は、それぞれ 0.84 および 0.90 である。図 2 B は、試験サンプル ($n = 62$) の ROC 分析を示す図である。臨床および臨床ゲノムモデルに関する濃度曲線下面積は、それぞれ 0.94 および 0.97 である。図 2 C は、組み合わせた訓練および試験セット ($n = 118$) の ROC 分析を示す図である。臨床および臨床ゲノムモデルに関する濃度曲線下面積はそれぞれ 0.89 および 0.94 であり、これは 2 曲線間の有意な差を表す ($P < 0.05$)。

【図 3】図 3 A ~ 3 C は、試験セットサンプル全体の 3 つのロジスティック回帰モデルの成績を示す図である。モデルから導いた、肺癌を有する確率が 0.5 以上のサンプルは癌性と分類し、確率が 0.5 未満のサンプルは非癌性と分類した。オレンジ; 癌性という最終診断を受けたサンプル。青色; 非癌性という最終診断を受けたサンプル。色の彩度は、それぞれのモデルにより癌を有するまたは癌を有しないと分類した、それぞれの最終診断群の比率を表す。それぞれのモデルに関して、感度 (S e n s)、特異度 (S p e c)、陽性的中率 (P P V)、および陰性的中率 (N P V) を示す。図 3 A は臨床モデルを示し、図 3 B はバイオマーカーモデルを示し、図 3 C は臨床ゲノムモデルを示す。臨床モデルとバイオマーカーモデルは、それぞれ 84% と 87% の精度で、それぞれ同様に機能する。臨床ゲノムモデルは、他の 2 つのモデルのいずれかより高い精度 (94%)、特異度、および陽性的中率を有する。

【図 4】図 4 は、臨床モデルによって予測した肺癌を有する確率と、試験セットサンプル全体 ($n = 62$) の医師の主観的評価の間の関連性を示す図である。モデルから導いた確率は y 軸上に示し、主観的臨床評価は x 軸上に示す。赤色の円、3 人の臨床医間の完全な一致。黒色の円、2 人の臨床医間の一致。緑色の円、一致せず。低位群対中位群、中位群対高位群、および低位群対高位群の確率の間には有意な差が存在する ($P < 0.01$ 、ウィルコクソンの検定)。主観的リスク評価により層別化したそれぞれの被験体の癌の状態は図 5 中に示す。

【図 5】図 5 は、試験セットサンプル全体 ($n = 62$) の癌の状態および医師の主観的評価により層別化した、臨床ゲノムモデルから導いた肺癌の予測を示す図である。ダークグレー、癌性という最終診断。ライトグレー、非癌性という最終診断。四角形、正確な臨床ゲノムモデルの予測。円、不正確なモデルの予測。医師により中程度の肺癌のリスクを有すると分類したサンプルのそれぞれは、臨床ゲノムモデルによって正確に予測された。

【図 6】図 6 は、訓練または試験セット中の癌の状態および帰属関係により層別化した、人工統計および臨床的特徴、ならびにバイオマーカースコアに関する平均および S D を示す図である (表 1)。

10

20

30

40

50

【図7】図7は、癌患者における腫瘍の細胞型、段階、および位置、ならびにそれぞれの亜群に関する気管支鏡検査で診断された分画に関する情報を示す図である（表2）。

【図8】図8は、それぞれ3つのロジスティック回帰モデル中の変数に関する効果推定値および誘導オッズ比を示す図である（表3）。

【図9】図9は、臨床ゲノムモデルが、試験セット中の3cm未満の塊サイズを有する病変ならびに明確に定義されていないX線写真の浸潤物も、正確に予測したことを示す図である（表4）。

【発明を実施するための形態】

【0013】

本明細書に記載する本発明は、肺障害または疾患、例えば肺癌を検出し、診断し、またはその診断を援助するための多元的方法に関する。本発明の方法は、複数（すなわち、2つ以上の）診断パラダイムを利用して、例えばそれぞれのパラダイム単独より診断感度、特異度、陰性的中率および/または陽性的中率を向上させる。これは、多元的方法中で使用するそれぞれのパラダイム下でなされた予測が互いに独立している場合、特に強力な手法である。本発明の方法は、肺障害（例えば、肺癌）を有する疑いがあるが気管支鏡検査で癌陰性である被験体（患者）を評価する際に特に有用であるが、この方法は、肺癌または他の肺障害を有する疑いがある任意の患者を診断する際に有利に利用することができる。

10

【0014】

本発明中で有用なパラダイムには、1つまたは複数の癌関連遺伝子の発現、1つまたは複数の癌関連の臨床的因子または変数の有無または重度、被験体の血液中の1つまたは複数の癌関連抗体の有無、被験体の血液中の1つまたは複数の癌関連タンパク質の有無、および1つまたは複数の癌関連マイクロRNAの発現があるが、これらだけには限られない。遺伝子発現（例えば、プローブおよびプライマー、マイクロアレイなどを使用するアッセイ）、タンパク質の有無および抗体の有無を測定する多くの具体的な方法は、当技術分野で周知である。さらに、肺癌関連の臨床的変数を測定する方法も当技術分野で公知である。

20

【0015】

本明細書で使用する「癌関連の」は、「癌の有無と関連があること」を意味するものとする。例えば、癌関連遺伝子は、癌を有さない個体と比較して癌を有する個体中で、（例えば、タイミング、レベルまたは位置（例えば、組織または細胞型）において）差次的に発現される遺伝子である。特定の実施形態では、癌関連の実体は肺癌関連の実体である。本明細書に記載するように、多元的方法は、2つ以上のパラダイム、3つ以上のパラダイム、4つ以上のパラダイムなどを、非限定的に利用することができる。

30

【0016】

本発明の1つの例示的な実施形態を以下に記載する。この実施形態は、遺伝子発現データ（すなわち、肺癌関連遺伝子の特定のセットからの遺伝子発現プロファイル；遺伝子発現バイオマーカー）の特定のセット、および肺癌関連の臨床的因子の特定のセットを利用する。しかしながら、本発明が、これらの具体的な臨床的因子または遺伝子発現データが抽出される遺伝子の特定のセットのいずれかに限られないことは明らかであるはずである。さらに本発明は、これら2つの特定のパラダイム（遺伝子発現プロファイルおよび臨床的因子）の使用に限られない。

40

【0017】

例えば、他の類似したデータまたは因子と組み合わせて使用するサブセットを含めた、いずれかのパラメーターのサブセットは、本発明により包含されるものとする。一実施形態では、遺伝子発現オムニバスのアクセッション番号GSE4115に含有される発現データの全部またはサブセットを、例示的な実施形態中に開示する臨床的因子の全部またはサブセットと組み合わせて使用することができる。さらに、例示的な実施形態中で使用する遺伝子発現データおよび/または臨床的因子の全部またはサブセットは、当技術分野で公知である追加の遺伝子発現データおよび/または臨床的因子と組み合わせて使用するこ

50

とができる。

【0018】

いくつかの実施形態では、肺障害の検出と関連があることが当技術分野で公知である、異なる遺伝子発現プロファイル（すなわち、遺伝子発現オムニバスのアクセッション番号 G S E 4 1 1 5 に含有される発現データではない）を、例示的な実施形態中に開示する肺疾患を検出するための臨床的因子の全部またはサブセットと組み合わせることができる。他の実施形態では、肺障害の検出と関連があることが当技術分野で公知である、異なる臨床的因子（すなわち、例示的な実施形態中に開示する臨床的因子のセットではない）を、例示的な実施形態中に開示する遺伝子発現プロファイルの全部またはサブセットと組み合わせることができる。

10

【0019】

さらなる実施形態では、肺障害の検出と関連があることが当技術分野で公知である異なる遺伝子発現プロファイル（すなわち、遺伝子発現オムニバスのアクセッション番号 G S E 4 1 1 5 に含有される発現データではない異なる遺伝子から決定した）を、肺障害の検出と関連があることが当技術分野で公知である異なる臨床的因子（すなわち、例示的な実施形態中に開示する臨床的因子のセットではない）と組み合わせることができる。例示的な実施形態中に記載する方法およびアルゴリズムを、任意のパラダイム、例えば任意の肺癌関連バイオマーカーまたは任意の臨床的因子から得たデータと共に使用して、肺の障害を予測または検出することができる。これらの方法およびアルゴリズムを最適化して、より高い的中率を有する（1つまたは複数の）パラダイムにより高い重みを、より低い的中率を有する（1つまたは複数の）パラダイムにより低い重みを与えることができる。

20

【0020】

例えば、本発明中で使用するための代替の遺伝子発現バイオマーカーは、米国特許公開 2 0 0 7 - 0 1 4 8 6 5 0、米国特許公開 2 0 0 6 - 0 1 5 4 2 7 8、米国特許出願第 1 1 / 9 1 8 , 5 8 8 号（2007年10月15日に出願）、米国仮出願第 6 0 / 9 9 4 , 6 3 7 号（2007年9月19日に出願）、米国仮出願第 6 0 / 9 9 4 , 6 4 3 号（2007年9月19日に出願）、および P C T 公開 W O 0 7 / 1 0 3 5 4 1 中で見ることができる。これらの特許出願の全ての教示は、その全体が参照により本明細書に組み込まれる。肺障害の診断と関連がある代替の遺伝子発現バイオマーカーも、当技術分野で公知である。

30

【0021】

本発明中で使用する臨床的因子には、記載するとおりに実施した臨床ゲノム診断試験中で使用しようとしなくとも、例示的な実施形態中に記載する全ての臨床的因子が挙げられるが、これらだけには限られない。本発明中で使用するための特定の臨床的因子には、年齢、喫煙歴（喫煙のパックイヤー数、開始年齢、強度および禁煙してからの年数を含む）、アスベスト曝露歴、喀血および体重減少を含めた臨床症状、節または塊のサイズ、および胸部イメージングにおけるX線像、リンパ節症の存在、転移性疾患の臨床またはX線上の証拠、肺活量測定における気道閉塞の証拠、ポジトロン放射断層撮影スキャンにおけるフルオロデオキシグルコースの撮取、任意の公知または疑いのある発癌物質への曝露、被験体によって使用されたたばこ製品の型、被験体における胸部の疼痛の有無、被験体における息切れの有無、被験体における一過性の息切れの有無、被験体の痰中の血液の有無、被験体における咳の有無、被験体における一過性の咳の有無、被験体の血液中の肺癌と関連がある1つまたは複数の抗体の存在、不在または量（例えば、Z h o n g ら、A m . J . R e s p r C r i t C a r e M e d 1 7 2 巻：1 3 0 8 ~ 1 3 1 4 頁（2005年）；Z h o n g ら、J T h o r a c O n c o l 1 巻：5 1 3 ~ 5 1 9 頁（2006年））、およびそれらの組合せがあるが、これらだけには限られない。臨床的因子は有無に基づいてスコア付けすることができ、または重度もしくは頻度に基づいて調整することができることが想定される。

40

【0022】

本発明の別の実施形態では、上記多元的診断法は、被験体の血液中の肺癌と関連がある

50

1つまたは複数の抗体の存在、不在または量（例えば、Zhongら、Am. J. Respr Crit Care Med 172巻：1308～1314頁（2005年）；Zhongら、J Thorac Oncol 1巻：513～519頁（2006年））を、遺伝子発現データ（免疫ゲノム学的診断を与えるため）と共に、または臨床的変数（免疫臨床的診断を与えるため）と共に、または遺伝子発現データと臨床的変数の両方（免疫学的臨床ゲノム診断を与えるため）と組み合わせて利用する。本発明のいくつかの実施形態では、本発明の方法は、被験体の血液中の肺癌と関連がある1つまたは複数の抗体の存在、不在または量を、1つまたは複数の追加の診断パラダイムと共に利用する。

【0023】

本発明の他の実施形態では、上記多元的診断法は、被験体の血液中の1つまたは複数の癌関連タンパク質の存在、不在または量を利用する。癌関連タンパク質には、ヒトアスパルチル - ヒドロキシラーゼ（HAAH）、癌胎児性抗原（CEA）、レチノール結合タンパク質（RBP）、 α -1-アンチトリプシン（AAT）、扁平上皮細胞癌抗原（SCCA）、血清アミロイドA、および腫瘍関連NADHオキシダーゼ（tNOX）があるが、これらだけには限られない。本発明のいくつかの実施形態では、この方法は、被験体の血液中の1つまたは複数の癌関連タンパク質の存在、不在または量を、1つまたは複数の追加の診断パラダイムと共に利用する。

【0024】

本発明の他の実施形態では、上記多元的診断法は、1つまたは複数の癌関連マイクロRNAの発現を、1つまたは複数の追加の診断パラダイムと共に利用する。例えば、喫煙者および非喫煙者中で差次的に発現されるマイクロRNA（miRNA）が記載されている（Schembriら、Proc Natl Acad Sci USA 106巻：2319～2324頁（2009年））。一実施形態では、1つまたは複数の肺癌関連miRNAは、miR-337、miR-18a、miR-189、miR-365、miR-181d、miR-10b、miR-150、miR-218、miR-338、miR-362、miR-17-3p、miR-15a、miR-652、miR-106b、miR-19b、miR-106a、miR-128a、miR-30a-3p、miR-128b、miR-130a、miR-500、miR-363、miR-199b、miR-223、miR-625、miR-99a、miR-125b、およびmiR-146aからなる群から選択される。特定の実施形態では、miRNAは、miR-218、miR-128b、miR-500およびmiR-181dの1つまたは複数である。

【0025】

独立した肺癌関連の診断パラダイムは、本明細書に記載するとおりに組合せて使用して、パラダイムの感度、特異度、陽性的中率および/または陰性的中率を個別に向上させることが好ましい。パラダイムの適切な組合せは、感度、特異度、陽性的中率および/または陰性的中率の全部またはサブセットを向上させることができる。特定のパラダイムは独立していることが当技術分野で公知である可能性があり、あるいはパラダイムのセットを以下に記載するように評価して、それらの互いの独立性を決定することができる。

【0026】

本発明の文脈では、診断判定（例えば、癌性/非癌性）を、それらが当技術分野で行われるように、それぞれのパラダイムにおいて行う。例えば、1つまたは複数の癌関連遺伝子の遺伝子発現プロファイルを評価する患者の生物学的サンプルから入手し、発現プロファイルを対照または標準と比較して、その遺伝子発現プロファイルに基づいて患者が癌を有するかまたは有さないかを判定する。利用したパラダイムのそれぞれからの診断判定を組み合わせ、全体のスコア付けまたは分類をもたらす。これらのステップのそれぞれに関する統計法は本明細書に記載し、他の方法は当技術分野で公知である。

【0027】

以前の試験から、肺癌を有する喫煙者と肺癌を有さない喫煙者由来の細胞学的に正常な

10

20

30

40

50

大気道の上皮細胞を区別することができる、遺伝子発現バイオマーカーが同定された (Spiral, Nat Med 13巻: 361~366頁 (2007年))。これらの細胞は、肺癌の疑いで気管支鏡検査を受けた患者の気管支気道のブラッシングから、比較的侵襲的な形式で回収することができる。気管支鏡検査中に得た細胞の細胞病理学的検査は肺癌に100%特異的であるが、癌の段階および位置に応じて30%と80%の間の限られた感度を有し、初期段階の疾患および末梢癌は最も低い感度を有する (SchreiberおよびMcCrorry, Chest 123巻: 115~285頁 (2003年))。

【0028】

結果として、医師は、気管支鏡検査が異常な細胞病理学的所見を有するいかなる細胞にも行き着かない場合、潜在的に初期段階の治療可能な疾患を有する患者のケア管理の仕方 10
に関して、困難な決定に直面する。しばしば、より高感度およびしばしばより侵襲的な診断手順で進行するかどうかに関する決定、または初期の疑わしいX線写真所見が後の反復イメージング研究において解明されるかを判定するかどうかの決定は、肺癌に関する患者の臨床的危険因子およびX線写真の危険因子の主観的評価に基づく。大気道の遺伝子発現バイオマーカーは、気管支鏡検査時に容易に回収することができる材料を使用するので (わずかにさらに2~3分手順を延長する)、他では分からない肺癌リスクに関する情報をバイオマーカーが得る場合、この試験は意思決定過程の有用な要素であり得る。

【0029】

本明細書で記載する結果は、大気道の上皮細胞中での遺伝子発現のパターンは、他の臨床危険因子と独立した肺癌の存在に関する情報を反映することを示唆する。この解釈は、 20
臨床的変数またはバイオマーカーのいずれかを含有するモデルと、組合せ臨床ゲノムモデルとの比較に起因する。この比較は、バイオマーカーはバイオマーカーモデルと臨床ゲノムモデルの両方において肺癌を有する確率と有意に関係があること、および組合せ臨床ゲノムモデル中のそれぞれの変数の重要性は、最初の非組合せモデル中のそれらの重要性と同様であることを示す。

【0030】

臨床ゲノムモデルは、独立した試験セット中のいずれかの最初のモデルより、肺癌の優れた予測因子である。ROC曲線の分析は、臨床ゲノムモデルは臨床モデルより有意に優れて機能することを示す。さらに、臨床ゲノムモデルは、臨床モデルの感度、特異度、陽性的中率および陰性的中率を増大させ、その精度が病変のサイズまたは位置によって影響 30
を受けることはないようである。

【0031】

小さなサンプルサイズの制約および限られた臨床パラメーターにもかかわらず、患者の完全な医療記録に基づく主観的臨床評価は臨床モデルの確率と関係があることは有望である。これは、これらの試験はコホート中の少数の被験体のみを実施したので、ポジトロン放射断層撮影スキャンの所見などの特定の変数が臨床モデル中に含まれなかったことを考慮すると、特に重要である。しかしながら、ポジトロン放射断層撮影スキャンの所見などの全ての利用可能なデータは、肺疾患の医師によって、彼らの肺癌の可能性の主観的評価の一部と考えられた。さらに、臨床ゲノムモデルは、主観的臨床評価によって中程度のリスク亜群に割り当てられた患者を正確に分類するようである。患者のこの亜群は、これら 40
の患者のほぼ3分の1が肺癌という最終診断を有していたので、おそらく臨床的に管理するのが特に困難である亜群である。

【0032】

本明細書で開示するデータは、遺伝子発現と肺癌に関する臨床危険因子を組み合わせた臨床ゲノムモデルが、さらなる侵襲性試験 (例えば肺生検) から恩恵を被る可能性がある患者を同定して、推定肺癌診断を確認し、それによってその根底にある悪性腫瘍に関する診断および治療を促進するために働くことができることを示唆する。さらに、臨床ゲノム診断の使用は、追加のより侵襲的な手順に曝される肺癌を有さない個体数の減少をもたらして、診断不能な気管支鏡検査後の肺癌診断を排除することができる。臨床医は、より確信的に低侵襲性および低コストの手法 (例えば、3~6カ月中の反復CTスキャン) を使 50

用して、低い臨床ゲノムの肺癌リスク値を有する患者を追跡することができる。

【0033】

肺癌のバイオマーカーとして働く細胞学的に正常な気道の上皮内の遺伝子発現プロファイルの能力は、これらの細胞中で観察される癌特異的な分子の変化の根底にある生物学に関する問題を提議する。小さな末梢肺病変の環境におけるバイオマーカーの高い診断精度は、肺癌を有する喫煙者と肺癌を有さない喫煙者の間の気道の遺伝子発現の変化が、腫瘍の直接の影響であるとは考えられないことを示唆する。遺伝子発現バイオマーカー中の酸化剤および炎症関連遺伝子の存在は、バイオマーカーはたばこの煙への曝露に応答して気道全体の癌特異的な差を検出する可能性を提議する。したがって、遺伝子発現の変化は肺癌の発生より先に生じ、その感度と比較して幾分低いバイオマーカーの特異度を説明することができる。これが本当である場合、バイオマーカーはおそらく、化学的予防戦略から恩恵を被る可能性があり疾患のリスクが最も高い喫煙者を同定するための有用なツールであり得る。

10

【0034】

本発明を、以下の非限定的な実施形態によってさらに記載する。全ての引用した参考文献の教示は、その全体が参照により本明細書に組み込まれる。

【実施例】

【0035】

材料および方法

患者集団

20

本試験のコホートは、大気道の遺伝子発現バイオマーカーを開発するための以前の試験に参加した患者からなる (Spiral, Nat Med 13 巻: 361~366 頁 (2007年))。その試験では、臨床上の肺癌の疑いで軟性気管支鏡検査を受けた現喫煙者および元喫煙者を、以前に記載されたのと同様に (Spiral, Nat Med 13 巻: 361~366 頁 (2007年)) 2003年1月と2005年4月の間に4カ所の三次医療センターで動員した。全ての被験体は21才の年齢を超えており、軟性気管支鏡検査法に対する禁忌はなかった。喫煙歴のない者、および葉巻のみを吸った者は、この試験から除外した。肺癌という最終診断または代替の診断を行うまで、気管支鏡検査後全ての被験体を追跡した (平均追跡時間、52日)。2005年5月に最終診断を達成し高品質のマイクロアレイデータを得た129人の被験体 (肺癌を有する60人の喫煙者および肺癌を有さない69人の喫煙者) を、第一のサンプルセット中に含めた。77人のこれらのサンプルを訓練セットにランダムに割り当てた。この試験の訓練セット (n=76) は、不完全な喫煙歴のため、これらの訓練セットサンプルの1つは除外した (図1A~1B)。一次試験の終了後、第二のサンプルセット (n=35) を、2005年6月と2006年1月の間に5カ所の医療センターで、臨床上の肺癌の疑いで軟性気管支鏡検査を受けた喫煙者から予め回収した。組み入れ基準および除外基準は、第一のサンプルセットと同じであった。この試験中の試験セットサンプル (n=87) は、第一のサンプルセットからの残りのサンプル (n=52) とこのプロスペクティブ試験セット (n=35) の両方と組み合わせた。図1A~1B中に示し以下でより詳細に記載するように、試験セットは気管支鏡検査後に確定診断がなかった患者のサブセットに限られた (n=62)。全ての被験体に関する人工統計情報は表1中に詳述し (図6)、研究のコホートにおける肺腫瘍 (n=78) の細胞型、段階、および位置に関する情報は表2 (図7) 中に示す。この試験は患者が動員された5カ所の医療センターの Institutional Review Boards (Boston University Medical Center, Boston, MA; Boston Veterans Administration, West Roxbury, MA; Lahey Clinic, Burlington, MA; St. James's Hospital, Dublin, Ireland; および St. Elizabeth's Medical Center, Boston, MA) によって承認され、全ての参加者は書面によるインフォームドコンセントを提出した。

30

40

50

【0036】

肺癌についての大気道の遺伝子発現バイオマーカー

Affymetrix HG-U133Aマイクロアレイを使用して、肺癌についての遺伝子発現バイオマーカーを、肺癌の疑いで気管支鏡検査を受けた喫煙者の右主気管支のブラッシングから回収した細胞学的に正常な大気道の上皮細胞における遺伝子発現プロファイルを使用して事前に開発した（遺伝子発現オムニバスのアクセッション番号GSE4115；Spiralら、Nat Med 13巻：361～366頁（2007年））。バイオマーカーは、完全な喫煙歴を有していなかった1つの追加のサンプルを加え、この試験の訓練セット（ $n = 76$ ）を使用して開発した（図1A～1B）。バイオマーカーは、バイオマーカースコアにこれらの発現レベルを組み合わせた加重投票アルゴリズム（Golubら、Science 286巻：531～537頁（1999年））を使用して、80個のプロブセット（72個の特有遺伝子、7個の注釈付けされていない転写産物、および1個の冗長プロブセット）の発現レベルから構築した。正の値は癌性を示し、負の値は非癌性を示す。

10

【0037】

この試験では、バイオマーカースコアは、以下の統計分析：（a）臨床的危険因子のみ、バイオマーカーのみを使用した肺癌の可能性、または組み合わせた臨床的危険因子とバイオマーカーを使用した癌の可能性を決定するための3つのロジスティック回帰モデルの構築、（b）初期モデル構築相に使用しなかった試験セットの患者における的中率の比較、および（c）熟練した臨床医によりなされた評価を用いる臨床モデルの比較の出発点として使用した。

20

【0038】

ロジスティック回帰モデルの構築

肺癌を有する患者の可能性を定量化するためのロジスティック回帰モデルを、訓練セットサンプル（ $n = 76$ ）を使用して作製した。この訓練セットは、肺癌または代替の非癌性病状のいずれかとの診断を確定した細胞病理学的所見を有した患者を含んでいた。気管支鏡検査で診断された患者は、サンプルの数を最大にするため、また診断不能な気管支鏡検査で患者の肺癌状態を正確に予測することができるモデルを開発するのにこれらのサンプルの除外は不必要であったので、訓練セット中に含めた。

【0039】

臨床および臨床ゲノムモデルに関して、利用可能な臨床的変数（表1；図6）は、年齢、喫煙のパックイヤー、および以下の二値変数、性別（男性、1；女性、0）、人種（1、白人；0、それ以外）、喀血（1、存在；0、それ以外）、リンパ節症（1、縦隔または肺門リンパ節。胸部CTスキャンで1cm；0、それ以外）、および塊サイズ（1、3cmを超える塊サイズを有する；0、それ以外）を含んでいた。ポジトロン放射断層撮影スキャンの情報は15人の患者にのみ利用可能であり、このモデル中には含まれなかった。Akaikeの情報基準（Akaike、IEEE Trans Automatic Control 19巻：716～723頁（1974年））を使用するバックワードステップワイズモデル選択を使用して、肺癌を有する患者の確率の最適臨床モデルを選択した。

30

【0040】

臨床的変数の影響に合わせて調節した後に、統合型臨床ゲノムモデルを作製し、遺伝子発現バイオマーカーの独立性および貢献度を判定するために、このバイオマーカーを最初に最適臨床モデルに加えた。バイオマーカースコアおよび利用可能な臨床的変数の全てを、次いでAkaikeの情報基準によるバックワードステップワイズモデル選択を使用して最適モデルを選択した。両方の手法が同じ組合せモデルをもたらした。以前の試験（Spiralら、Nat Med 13巻：361～366頁（2007年））において使用された加重投票予測アルゴリズム中と同様にロジスティック回帰中で、バイオマーカースコアが同様に機能することを確認するために、精度、感度、特異度、陽性的中率、および陰性的中率を、加重投票予測、および独立した試験サンプル全体のバイオマーカースコアの

40

50

みを含んだロジスティック回帰モデルによりなされた予測で比較した。

【0041】

独立した患者におけるモデルの成績の比較

ロジスティック回帰モデル（臨床、バイオマーカー、および臨床ゲノム）の成績を、気管支鏡検査法で得た材料の細胞病理検査で診断不能であった、訓練セット（ $n = 76$ ）中の患者のサブセット（ $n = 56$ ；図1A～1B）に対して最初に評価した。本発明者らは、診断不能な気管支鏡検査に焦点を置くことを選択して、肺癌に関するさらなる診断評価を必要とする患者の状況で、遺伝子発現バイオマーカーおよび臨床パラメーターの有用性を具体的に評価した。さらに重要なことに、本発明者らは、診断不能な気管支鏡検査の試験セットにおけるモデルも試験した（ $n = 62$ ；図1A～1B）。それぞれのモデルに関して、0.5以上の肺癌の確率を有していた患者は肺癌を有すると分類し、0.5未満の確率を有していた患者は肺癌を有しないと分類した。受信者動作特性（ROC）曲線も使用して、気管支鏡検査で診断不能である訓練セットの患者、独立した試験セット、および気管支鏡検査で診断不能である全患者の組合せセット（ $n = 118$ ）において、臨床モデルと臨床ゲノムモデルを比較した。同じセットのサンプルに基づく2つのROC曲線が有意に異なっただろうかを評価するために、同じ症例から誘導したROC曲線を比較するために開発された方法を使用した（HanleyおよびMcNeil、Radiology 143巻：29～36頁（1982年）；HanleyおよびMcNeil、Radiology 148巻：839～843頁（1983年））。異なるサンプルセットに基づくROC曲線を比較するために、2標本のz検定を使用した。2つの異なる変数のオッズ比は比較することができないので、ROC曲線はモデルに加えた変数の追加の利点を評価するための共通の尺度として働く（Sullivanら、J Natl Cancer Inst 93巻：1054～1061頁（2001年））。精度、感度、特異度、陽性的中率、および陰性的中率は、臨床モデル、バイオマーカーモデル、および臨床ゲノムモデルの独立した試験セット全体でも計算した。

10

20

【0042】

主観的臨床評価

三次医療センターで診療している3人の独立した肺疾患の臨床医は最終診断を見ずに、気管支鏡検査時にそれぞれの患者の臨床歴を評価した。臨床歴は、年齢、喫煙状態、たばこへの累積曝露時間、併存疾患、症状/兆候、X線写真所見、および入手可能な場合はポジトロン放射断層撮影スキャンの結果を含んでいたが、これらだけに限られなかった。この情報に基づいて、臨床医は3つのリスク群：低位群（肺癌の確率が10%未満と評価された）、中位群（肺癌の確率が10～50%と評価された）、および高位群（肺癌の確率が50%を超えると評価された）のうちの1つにそれぞれの患者を分類した。それぞれの被験体に関する最終的な主観的割り当ては中間意見を選択することによって決定した。患者の診断不能な気管支鏡検査の臨床的分類に関する評定者間信頼性は有意であり、臨床医間の一致レベルは、K統計値により測定して偶然に予想され得るレベルより高かったことが示された（ $K = 0.57$ ； $P < 0.001$ ；ref. 28）。

30

【0043】

主観的臨床評価と臨床ゲノムモデルの比較

肺癌を有するリスクを予測するための包括的臨床モデルの構築用のサンプルサイズは、臨床および臨床ゲノムモデルに含めるために利用可能であった変数の範囲と同様に限られていた。したがって本発明者らは、臨床モデルが肺疾患の専門家によりなされた主観的臨床評価と同様に機能するかどうか決定しようとした。なぜならば、この評価は、（a）それぞれの臨床医のキャリアにわたって経験した多数の患者に「訓練され」、（b）患者の医療記録内に含まれた全ての情報を考慮しているからである。ウィルコクソンの検定を使用して、臨床モデルから導いた、肺癌を有する確率が、臨床医によって低い、中程度、または高い癌のリスクと分類したサンプル間で異なっていたかどうか評価した。

40

【0044】

統計分析

50

全ての統計分析は、R統計ソフトウェアバージョン2.2.1を使用して実施した。

【0045】

結果

肺癌の独立した予測因子としての遺伝子発現バイオマーカーの評価

訓練または試験セット中の癌の状態および帰属関係により層別化した、人工統計および臨床的特徴、ならびにバイオマーカースコアに関する平均およびSDを、表1(図6)中に示す。年齢、人種、喫煙のバックイヤー数、リンパ節症、塊サイズ、およびバイオマーカースコアは、肺癌を有する患者と肺癌を有さない患者の間で有意に異なった($P < 0.001$)。しかしながら、試験および訓練セットは、分析中で使用した変数に関して十分バランスがとれていた(ただし、3cmを超える塊サイズを有する発生率は、訓練セットと比較して試験セット中で幾分低かった; $P = 0.047$)。癌患者における腫瘍の細胞型、段階、および位置、ならびにそれぞれの亜群に関する気管支鏡検査で診断された画分に関する情報を、表2(図7)中に示す。それぞれ3つのロジスティック回帰モデル中の変数に関する効果推定値および誘導オッズ比を、表3(図8)中に示す。本発明者らは、このコホートの最適臨床モデルが、バックイヤー数を含んでいなかったことを発見した。これはおそらく、年齢とバックイヤー数の間の強い相関関係によるものである。禁煙してからの時間で二分したかどうかと無関係に、最適臨床モデルは喫煙状態(元喫煙者対現喫煙者)を含んでいなかった。さらに、(3cmの代わりに)2cmの閾値を使用した塊サイズの二分によって、類似した全体精度を有する臨床および臨床ゲノムモデルがもたらされた。

10

20

【0046】

バイオマーカースコアから導いた、肺癌を有する可能性を記載するロジスティック回帰モデルは、以前(Postmus, Chest 128巻:16~18頁(2005年))の加重投票アルゴリズムによる肺癌状態の予測と同等の結果を生み出し、8対7の不正確な分類をもたらし、バイオマーカースコアは臨床ゲノムモデルにおいて原型バイオマーカー予測アルゴリズムをモデル化するための正確な方法であることを示した。バイオマーカースコアは、バイオマーカー単独モデル($P < 0.001$)および臨床ゲノムモデル($P < 0.005$)との両方において、肺癌の可能性の有意な予測因子である。臨床ゲノムモデルにおいて、臨床的変数の係数は、大半は臨床モデルと変わらず、バイオマーカーの係数は大半はバイオマーカー単独モデルと変わらない。これらのデータは、遺伝子発現バイオマーカーおよび臨床的変数が肺癌リスクの独立した予測因子であることを示唆する。

30

【0047】

臨床ゲノムモデルの成績の評価

3つのモデルを使用して、気管支鏡検査で診断不能である訓練サンプル($n = 56$)、独立した試験サンプル($n = 62$)、および組み合わせたこの2つのセット($n = 118$)のサブセットの癌の状態を予測した。ROC曲線を使用して、臨床モデルの成績と臨床ゲノムモデルの成績を比較した(図2A~2C)。臨床ゲノムモデルは、3つ全てのサンプルセット中で臨床モデルより良い成績を有していた。この成績の差は試験セット中で統計的に有意に達しないが、訓練および試験セットを組み合わせると、臨床ゲノムモデルと臨床モデルの間の濃度曲線下面積の有意な差が存在した($P < 0.05$)。ROCの濃度曲線下面積の差に関しては、訓練セットサンプル中のモデルの成績は試験セットサンプル中より決して優れているようではなく($P = 0.25$)、その濃度曲線下面積の差は0.065である。95%信頼区間、 $-0.046 \sim 0.174$)。これは、これらのモデルが訓練セットに過剰適合しないこと、したがって訓練および試験セットを組み合わせ、臨床モデルと臨床ゲノムモデルの成績の差の有意性を評価するのが妥当であることを示唆する。

40

【0048】

3つのモデルのそれぞれに関する感度、特異度、陽性的中率、および陰性的中率を、試験セット全体で評価した(図3A~3C)。組合せ臨床ゲノムモデルは感度および陰性的中率を100%まで増大させ、他のモデルと比較して高い特異度および陽性的中率をもた

50

らす。末梢性病変を有する癌被験体は試験セット中で十分示され(70.6%)、臨床ゲノムモデルは末梢または中枢肺腫瘍間で同様に正確であった。臨床ゲノムモデルは、試験セット中の3cm未満の塊サイズを有する病変ならびに明確に定義されていないX線写真の浸潤物も正確に予測した(表4;図9)。さらに、臨床および臨床ゲノムモデルの成績は、気管支鏡検査で診断不能であるサンプルに特異的であるようではない。これらのモデルは、診断が十分な気管支鏡検査法所見を有する独立サンプル(n=25)において90%および95%の感度を有していたからである。最後に、気管支鏡検査で診断不能である訓練サンプル単独全体(n=56)の臨床および臨床ゲノムモデルの訓練は、試験セット中で類似した精度(それぞれ82%と91%)およびモデル間のROCの濃度曲線下面積の有意な差をもたらした(P<0.05)。

10

【0049】

臨床ゲノムモデルと主観的臨床評価の比較

それが含有する比較的少数の変数を考慮して、臨床モデルが包括的であるかどうか評価するために、本発明者らは、それが3人の肺疾患の医師の平均主観的評価と関連があるかどうかを評価した。試験セットサンプル全体で臨床モデルの予測と主観的臨床評価の間には関連があった(図4)。臨床モデルの確率は、3人の医師が評価したリスク群間で有意に異なった(P<0.01)。

【0050】

臨床モデルと主観的臨床評価の間の関連を考慮して、本発明者らは、試験セットサンプル中で癌の状態および主観的臨床評価したカテゴリーにより層別化した臨床ゲノムモデルによってなされた予測を調べた(図5)。医師の意見は、中程度のリスクカテゴリー中の11サンプルに関する全臨床データに基づくと最も不確かである。臨床モデルは11サンプルのうち7サンプルを正確に分類することができるが、臨床ゲノムモデルは11サンプル全てを正確に分類する。

20

【図1】

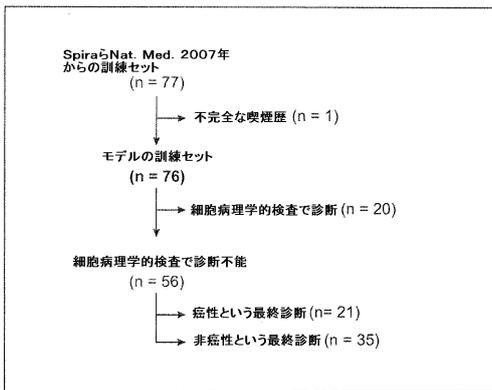


FIG. 1A

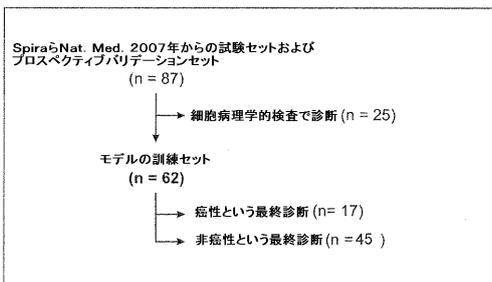


FIG. 1B

【図2】

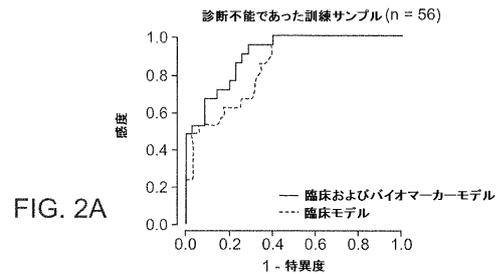


FIG. 2A

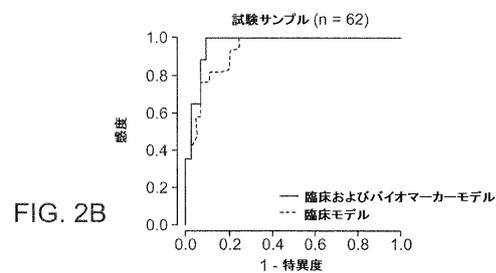


FIG. 2B

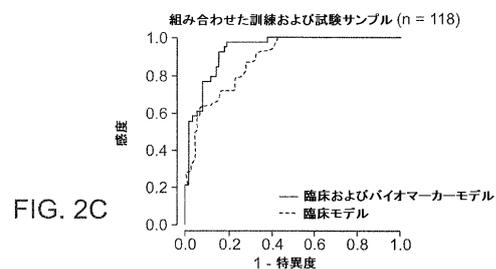


FIG. 2C

【 図 3 】

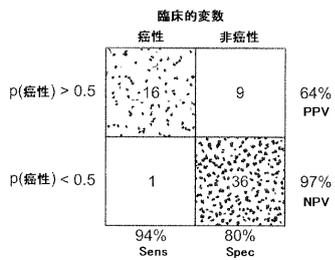


FIG. 3A

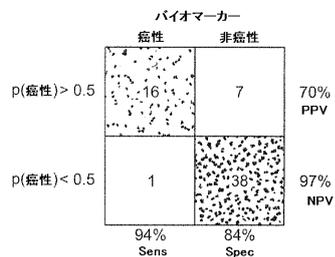


FIG. 3B

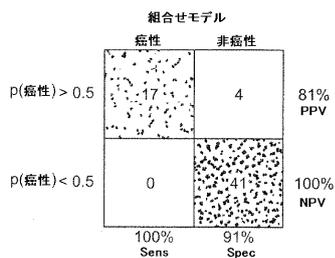


FIG. 3C

【 図 5 】

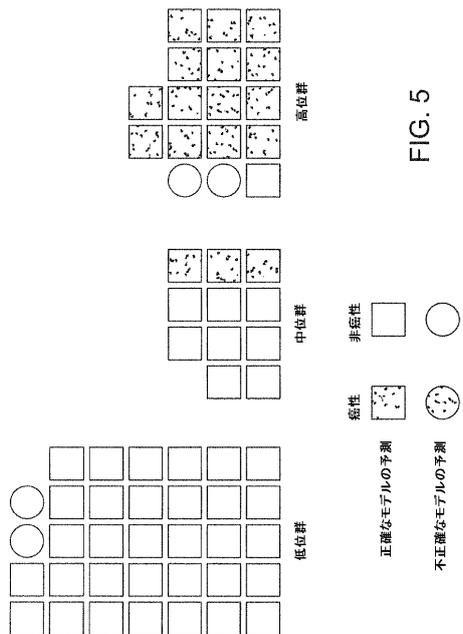


FIG. 5

【 図 4 】

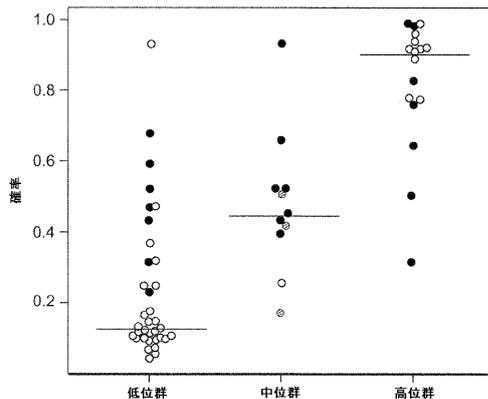


FIG. 4

【 図 6 】

表1. 訓練および試験セット中の癌の状態または癌属性により層別化した、人工脳波の特徴、臨床特徴およびバイオマーカーの特徴

因子	全体 (n=163)	癌性 (n=78)	非癌性 (n=85)	P*	訓練 (n=76)	試験 (n=62)	P*
年齢	58.1 ± 14.3	64.5 ± 9.6	52.3 ± 15.4	<0.001	57.3 ± 14.0	57.6 ± 15.3	0.91
男性	122/163 (74.8)	60/78 (76.9)	62/85 (72.9)	0.59	59/76 (77.6)	42/62 (67.7)	0.25
白人	110/163 (67.5)	67/78 (85.9)	43/85 (50.6)	<0.001	52/76 (68.4)	36/62 (58.1)	0.22
10年以内の喫煙	130/163 (79.8)	60/78 (76.9)	70/85 (82.4)	0.44	62/76 (81.6)	47/62 (75.8)	0.53
バックブレイク	44.9 ± 32.0	54.9 ± 26.8	35.7 ± 33.7	<0.001	45.8 ± 30.2	39.9 ± 35.4	0.3
気管支鏡検査法による診断	45/163 (27.6)	40/78 (51.3)	5/85 (5.9)	<0.001	20/76 (26.3)	0/62 (0)	<0.001
癌	78/163 (47.9)	78/78 (100.0)	0/85 (0.0)	<0.001	40/76 (52.6)	17/62 (27.4)	0.003
リンパ節症	43/163 (26.4)	38/78 (48.7)	5/85 (5.9)	<0.001	17/76 (22.4)	10/62 (16.1)	0.4
咳血	15/163 (9.2)	6/78 (7.7)	9/85 (10.6)	0.6	10/76 (13.2)	2/62 (3.2)	0.07
3mmを超える塊サイズ	48/163 (29.4)	43/78 (55.1)	5/85 (5.9)	<0.001	24/76 (31.6)	10/62 (16.1)	0.047
バイオマーカー	-0.35 ± 8.93	4.65 ± 7.04	-4.94 ± 7.98	<0.001	0.34 ± 8.97	-2.72 ± 9.12	0.05

注: P-値は、連続変数および二値変数の比較に関する平均土SDによる。P-値は、全変数変数および二値変数の比較に関する平均土SDによる。試験セットの比較に関するものである。不均衡なデータは、連続変数に等化して使用し、フィッシャーの正確確率検定は二値変数に対して使用した。

FIG. 6

【 図 7 】

表2. 肺癌サンプル(n=78)に関する細胞型、段階、および位置の情報

細胞型	n	気管支鏡検査で診断されたサンプルの割合(%)
SCLC	14	64.3
NSCLC (非公知の亜型)	15	60.0
扁平上皮	27	55.6
腺癌	18	33.3
大細胞癌	4	25.0
段階		
非公知	1	0.0
I	14	35.7
II	2	50.0
III	25	52.0
IV	22	54.5
位置		
中枢	28	71.4
末梢	49	40.8
その他*	1	0.0

注: 気管支鏡検査法から、細胞病理学的検査で肺癌の診断が得られた各群中のサンプルの割合を報告する。

略語: SCLC、小細胞肺癌、NSCLC、非公知細胞肺癌。

* 中枢 対 末梢として特徴付けすることができない症例。

FIG. 7

【 図 8 】

表3. 訓練セットサンプルに適合したロジスティック回帰モデル

モデル	範囲	係数	OR (95% CI)	P
バイオマーカ単独				
切片	NA	0.07	NA	0.776
バイオマーカ	-18.88 ~ 16.91	0.13	1.14 (1.06-1.21)	0.00017
臨床的変数単独				
切片	NA	-5.01	NA	0.003
年齢	23-79	0.07	1.07 (1.02-1.13)	0.008
塊サイズ	0-1	2.19	8.91 (2.08-38.25)	0.003
リンパ節症	0-1	2.09	8.12 (1.45-45.63)	0.017
バイオマーカ+臨床的変数				
切片	NA	-4.9	NA	0.014
バイオマーカ	-18.88 ~ 16.91	0.13	1.13 (1.04-1.24)	0.005
年齢	23-79	0.07	1.07 (1.00-1.14)	0.036
塊サイズ	0-1	1.85	6.38 (1.39-29.34)	0.017
リンパ節症	0-1	2.75	15.69 (2.23-110.28)	0.006

注: 訓練セットサンプル全体(n=76)の範囲、回帰係数、オッズ比(OR)、オッズ比の95%信頼区間(95%CI)、および変数のP値を報告する。

FIG. 8

【 図 9 】

表4. 訓練セット中で癌の求能および塊サイズまたは腫瘍の位置により層別化した臨床ゲノムモデルの精度(n=62)

塊サイズ (cm)	癌性		非癌性	
	n	精度 (%)	n	精度 (%)
>3	9	100.0	1	0.0
<3	5	100.0	37	91.9
明確に定義されていない浸潤物	3	100.0	7	100.0
位置				
中枢	5	100.0	3	100.0
末梢	12	100.0	17	76.5
その他*	0	NA	25	100.0

* 中枢 対 末梢として特徴付けすることができない症例。

FIG. 9

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US 09/38836
A. CLASSIFICATION OF SUBJECT MATTER IPC(8) - C12Q 1/68 (2009.01) USPC - 435/6 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) USPC - 435/6 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched USPC - 435/330,464; 536/23.1; 702/20 Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) PubWEST(DB=PGPB,USPT,USOC,EPAB,JPAB), Google Scholar(lung cancer clinico-genomic, lung cancer clinical-genetic, positron emission tomography gene marker, positron emission tomography markers, lung cancer clinico-genomic diagnosis, lung cancer clinical genetic diagnosis, lung cancer clinicogenomic diagnosis)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SPIRA et al, Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. NATURE MEDICINE, 4 March 2007 (04.03.2007), vol 13, no 3, pp 361-366; Abstract; pg 362, right col, para 4; pg 363, left col para 1; pg 364, right col, para 2-3; pg 365, left col para 1, Accession codes.	1-20
A	POTTI et al, A Genomic Strategy to Refine Prognosis in Early-Stage Non Small-Cell Lung Cancer. The New England Journal of Medicine, 10 August 2006 (10.08.2006), vol 335, no 6, pp 570-580	1-20
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/>		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 1 June 2009 (01.06.2009)		Date of mailing of the international search report 10 JUN 2009
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201		Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-1774

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(74)代理人 100062409

弁理士 安村 高明

(74)代理人 100113413

弁理士 森下 夏樹

(72)発明者 スピラ, アブラム

アメリカ合衆国 マサチューセッツ 02465, ニュートン, ランドレット パーク 25

(72)発明者 レンバーグ, マーク イー.

アメリカ合衆国 カリフォルニア 94705, バークレー, プリンズ ストリート 2136

(72)発明者 ビーン - エーベル, ジェニファー イー.

アメリカ合衆国 ニューメキシコ 87144, リオ ランチョ, コントレラス ロード エヌイー 2017

(72)発明者 リッピー, ダニエル

アメリカ合衆国 マサチューセッツ 01776, サドバリー, カドワース レーン 61

Fターム(参考) 4B024 AA12 CA09 CA11 HA14

4B063 QA01 QA19 QQ52 QR55 QS34