

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 995 959**

51 Int. Cl.:

G01N 33/48 (2006.01)

G16B 5/20 (2009.01)

G01N 33/68 (2006.01)

G01N 33/574 (2006.01)

G16B 40/10 (2009.01)

G16B 20/00 (2009.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **09.02.2018** **PCT/US2018/017704**

87 Fecha y número de publicación internacional: **16.08.2018** **WO18148600**

96 Fecha de presentación y número de la solicitud europea: **09.02.2018** **E 18752036 (6)**

97 Fecha y número de publicación de la concesión europea: **02.10.2024** **EP 3580560**

54 Título: **Métodos para la detección y tratamiento del cáncer de pulmón**

30 Prioridad:

09.02.2017 US 201762456731 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
11.02.2025

73 Titular/es:

**BOARD OF REGENTS, THE UNIVERSITY OF
TEXAS SYSTEM (100.00%)
210 West 7th Street
Austin, TX 78701, US**

72 Inventor/es:

**HANASH, SAMIR;
TAGUCHI, AYUMU y
FENG, ZIDING**

74 Agente/Representante:

ELZABURU, S.L.P

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 995 959 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Métodos para la detección y tratamiento del cáncer de pulmón

Antecedentes

El cáncer de pulmón es el cáncer más frecuente en los Estados Unidos, con una tasa de supervivencia a los cinco años de menos del 15 %. La terapia para el cáncer de pulmón está cambiando del uso de una selección limitada de terapias que consisten en radiación, metabolismo de folato, fármacos basados en platino y/o fármacos basados en taxol a tratamientos más dirigidos que requieren la caracterización histológica del tumor y/o la presencia o ausencia de biomarcadores clave o proteínas diana terapéuticas.

El cáncer de pulmón es el cáncer principal causante de muerte en todo el mundo y representa una de cuatro de todas las muertes por cáncer en los Estados Unidos. Los datos del Ensayo Nacional de Cribado de Pulmón (NLST) sugieren que el cribado anual con LDCT torácica para fumadores actuales de alto riesgo y exfumadores reduce la mortalidad por cáncer de pulmón en un 20 % y la mortalidad total en un 7 %. Como resultado, el Grupo de Trabajo de Servicios Preventivos de los Estados Unidos (USPSTF) ha recomendado el cribado por LDCT para cáncer de pulmón en fumadores ocasionales de 55-80 años de edad que han fumado 30 paquetes al año con no más de 15 años desde haberlo dejado. Sin embargo, el estudio de NLST resaltó varios aspectos negativos importantes asociados con el cribado por CT en términos de morbilidad asociada con un sobrediagnóstico, tratamiento de nódulos benignos y costes financieros. Es importante destacar que, basándose en los criterios de cribado actuales, se estima que el número de vidas que se salvarían mediante cribado como porcentaje del número total de muertes debidas a cáncer de pulmón es relativamente modesto. Además, los datos de más de un millón de sujetos reclutados de una cohorte de población general (Biobank del Reino Unido y el estudio Europeo de Investigación Prospectiva en Cáncer y Nutrición (EPIC)) indican que menos del 50 % de los casos de cáncer de pulmón incidentes habrían sido elegibles para el cribado basándose en los criterios de USPSTF.

Existe una abundancia de literatura sobre la predicción del riesgo de cáncer de pulmón que proporciona argumentos convincentes sobre el beneficio de reemplazar los criterios de cribado dicotómicos de USPSTF con un modelo basado en el riesgo cuando se identifican sujetos para cribado por CT. Por ejemplo, recientemente se estimó que se podían evitar un 20 % muertes adicionales por cáncer de pulmón usando un criterio de cribado basado en la evaluación de riesgos individuales. Considerando que la información requerida para utilizar tal herramienta de predicción de riesgo puede ser fácilmente determinada por un profesional general - o incluso autoevaluarse usando un calculador de riesgo en línea - parecería probable que los programas de cribado de cáncer de pulmón futuros implementaran tales herramientas cuando se evalúa la elegibilidad de cribado.

Una alternativa a los criterios de cribado de USPSTF sería un criterio de cribado basado en riesgo a nivel individual que estime con precisión el riesgo de cáncer de pulmón en el futuro cercano (por ejemplo, 1-3 años) para cada individuo. Se han publicado varios modelos de predicción de riesgo que se basan en datos demográficos (edad y sexo) y datos de factores de riesgo de cuestionarios (principalmente tabaquismo). Sin embargo, se obtendría una mejora sustancial incorporando biomarcadores que incorporen información que no puede capturarse con cuestionarios.

Actualmente, se están llevando a cabo esfuerzos para desarrollar biomarcadores basados en sangre para la detección temprana de cáncer de pulmón usando una variedad de plataformas analíticas que se dirigen a ácidos nucleicos, proteínas y metabolitos, así como a células y micropartículas tumorales circulantes. Por ejemplo, los estudios proteómicos han llevado a la identificación de varios marcadores circulantes candidatos, incluyendo proteína tensioactiva B (SFTPb) y su forma precursora (Pro-SFTPb) que se ha demostrado que mejora sustancialmente la predicción del riesgo de cáncer de pulmón en combinación con la información de tabaquismo cuando se analiza usando muestras de sangre recogidas en el momento de la LDCT (mejora del AUC: 0,67 a 0,74), así como en muestras de sangre antes del diagnóstico del Estudio de Salud de los Médicos. Se ha encontrado que otras proteínas circulantes exhiben niveles aumentados en el cáncer de pulmón, pero con sensibilidad y especificidad limitadas para tener utilidad como marcadores únicos para el cribado del cáncer de pulmón.

Dado el rendimiento de SFTPb y su forma precursora Pro-SFTPb en muestras recogidas antes del diagnóstico del cáncer de pulmón, y en un entorno de cribado, se investigó el potencial de mejorar los criterios de USPSTF usando un panel de proteínas seleccionadas relacionadas con tumores. El estudio se basó en una cohorte prospectiva de cáncer de pulmón de alto riesgo, el estudio del Ensayo de Eficacia de Betacaroteno y Retinol (CARET) para construir el panel, y casos y controles seleccionados de dos cohortes grandes de población general, el estudio Europeo de Investigación Prospectiva en Cáncer y Nutrición (EPIC) y el estudio de Salud y Enfermedad del Norte de Suecia (NSHDS) que implicaba 367.000 participantes para la validación.

El documento US 2013/196868 enseña el uso de un panel que usa tres marcadores de proteínas tumorales y tres marcadores de autoanticuerpos, pero no dice nada sobre la necesidad de medir los niveles de Pro-SFTPb para evaluar el riesgo de cáncer de pulmón. El documento WO 2011/140234 enseña el uso de tres marcadores para predecir la sensibilidad de un cáncer a ABT-869, pero tampoco dice nada sobre la necesidad de medir los niveles de Pro-SFTPb para evaluar el riesgo de cáncer de pulmón. Los documentos, tales como, WILLIAM R. WIKOFF ET AL: 20 de noviembre de 2015 (20-11-2015), DON D. SIN ET AL: 20 de diciembre de 2013 (20-12-2013), el documento WO

2007/076439 y US 2016/060329 divulgan el biomarcador clave Pro-SFTPb del cáncer de pulmón. El documento US 2015/072890 enseña el uso de tres marcadores de proteína tumoral con marcadores de miARN y opcionalmente tres marcadores de autoanticuerpos, pero no dice nada sobre la necesidad de medir los niveles de Pro-SFTPb para evaluar el riesgo de cáncer de pulmón. Cedrés Susana ET AL: 1 de mayo de 2011 (01-05-2011) solo se ocupa del "pronóstico empeorado" del paciente, sin embargo, no dice nada sobre la necesidad de medir los niveles de Pro-SFTPb para evaluar el riesgo de cáncer de pulmón. El documento CN 102 321 760 se refiere a la validación del uso de STAT3 como biomarcador. MOLINA R ET AL.: 1 de agosto de 2003 (01-08-2003) enseña que "usar más de dos o tres marcadores no aumenta significativamente la sensibilidad", otros documentos contradicen esto.

Resumen

El alcance de protección se define por las reivindicaciones adjuntas.

La presente divulgación proporciona métodos y kits para la detección temprana del cáncer de pulmón. Los métodos y kits usan múltiples ensayos de biomarcadores contenidos dentro de una muestra biológica obtenida de un sujeto. El análisis combinado de al menos cuatro biomarcadores: CEA, CA125, CYFRA21-1 y Pro-SFTPb, proporciona un diagnóstico de alta precisión del cáncer de pulmón cuando se criba contra cohortes con estado de cáncer de pulmón conocido, es decir, positivo para cáncer de pulmón o negativo para cáncer de pulmón. Opcionalmente, también se puede analizar el nivel de diacetilpermina (DAS).

Se identificó un modelo de regresión que puede predecir el estado de cáncer de pulmón para un sujeto basándose en los niveles de CEA, CA125, CYFRA21-1 y Pro-SFTPb que se encuentran en una muestra biológica del sujeto.

Se desarrolló una puntuación de biomarcador basada en cuatro biomarcadores separados en una cohorte de Estados Unidos, y la validación externa en dos cohortes europeas confirmó que incorporar la puntuación de biomarcador con información de tabaquismo en un modelo integrado de predicción de riesgo proporcionaba una mejora notable en la discriminación general entre casos y controles en comparación con un modelo basado solo en información de tabaquismo. El modelo integrado de predicción de riesgo identificó el 76 % de casos futuros de cáncer de pulmón en la muestra de validación, en comparación con el 42 % de los criterios de cribado de USPSTF, sin aumentar el número de falsos positivos.

Por consiguiente, en la presente memoria se proporcionan:

- métodos para la detección de cáncer de pulmón y riesgo de cáncer de pulmón en un sujeto;
- métodos para determinar y/o cuantificar el riesgo aumentado de presencia de cáncer de pulmón en un sujeto;
- métodos para determinar el riesgo de un sujeto de albergar cáncer de pulmón;

comprendiendo cada uno medir el nivel de CEA, CA125, CYFRA21-1 y Pro-SFTPb, y opcionalmente, diacetilpermina (DAS), en una muestra del sujeto.

También se proporcionan métodos de tratamiento o prevención de la progresión del cáncer de pulmón en un sujeto en donde los niveles de CEA, CA125, CYFRA21-1 y pro-SFTPb y, opcionalmente, diacetilpermina (DAS), clasifica al sujeto como que tiene o está en riesgo de albergar cáncer de pulmón.

También se proporcionan kits correspondientes para determinar la presencia de indicadores de cáncer de pulmón en una muestra del sujeto, para la detección de cáncer de pulmón y riesgo de cáncer de pulmón en un sujeto, y para determinar y/o cuantificar el aumento del riesgo de la presencia de cáncer de pulmón en un sujeto, que comprende materiales para medir CEA, CA125, CYFRA21-1 y pro-SFTPb, y opcionalmente, diacetilpermina (DAS) en la muestra.

En algunas realizaciones, los biomarcadores se miden en muestras de sangre extraídas de sujetos. En algunas realizaciones, se puede determinar la presencia o ausencia de biomarcadores en una muestra biológica. En algunas realizaciones, se puede cuantificar el nivel de biomarcadores en una muestra biológica.

Se proporciona una superficie para analizar una muestra biológica. En algunas realizaciones, los biomarcadores de interés se adsorben de forma no específica en esta superficie. En algunas realizaciones, se incorporan receptores específicos para biomarcadores de interés sobre esta superficie. En algunas realizaciones, la superficie está asociada con una partícula, por ejemplo, una perla.

El biomarcador se une a una molécula receptora particular, y se puede determinar la presencia o ausencia del complejo biomarcador-receptor. En algunas realizaciones, se puede cuantificar la cantidad de complejo biomarcador-receptor. La molécula receptora está unida a una enzima para facilitar la detección y cuantificación.

El biomarcador se une a una molécula transmisora particular, y el complejo biomarcador-molécula transmisora se une a su vez a una molécula receptora. En algunas realizaciones, se puede determinar la presencia o ausencia del complejo biomarcador-retransmisor-receptor. La cantidad de complejo biomarcador-retransmisor-receptor puede cuantificarse. En algunas realizaciones, la molécula receptora está unida a una enzima para facilitar la detección y cuantificación.

Se analiza una muestra biológica secuencialmente para determinar biomarcadores individuales. Una muestra biológica se divide en porciones separadas para permitir el análisis simultáneo de múltiples biomarcadores. En algunas realizaciones, se analiza una muestra biológica en un único proceso para múltiples biomarcadores.

5 La ausencia o presencia de biomarcador puede determinarse mediante inspección visual. En algunas realizaciones, la cantidad de biomarcador puede determinarse mediante el uso de una técnica espectroscópica. La técnica espectroscópica es espectrometría de masas. La técnica espectroscópica es la espectrometría UV/Vis. En algunas realizaciones, la técnica espectroscópica es una técnica de excitación/emisión tal como espectrometría de fluorescencia.

10 El análisis de biomarcadores CEA, CA125, CYFRA21-1 y Pro-SFTPb se puede combinar con el análisis de biomarcadores adicionales. En algunas realizaciones, los biomarcadores adicionales pueden ser biomarcadores proteicos. Los biomarcadores adicionales pueden ser biomarcadores no proteicos. Los biomarcadores no proteicos pueden ser ADN tumoral circulante (ADNct). Dicho método puede comprender además medir el nivel de un metabolito, por ejemplo, diacetilpermina (DAS), en la muestra biológica; en donde la cantidad de diacetilpermina (DAS) clasifica al sujeto como que está en riesgo de albergar cáncer de pulmón o no en riesgo de albergar cáncer de pulmón. Pueden incorporarse metabolitos adicionales según sea necesario.

15 En otro aspecto, la divulgación proporciona un método para determinar el riesgo de un sujeto de albergar cáncer de pulmón, que comprende obtener una muestra biológica del sujeto; medir el nivel de diacetilpermina (DAS) en la muestra biológica; en donde la cantidad de diacetilpermina (DAS) clasifica al sujeto como que está en riesgo de albergar cáncer de pulmón o no está en riesgo de albergar cáncer de pulmón.

20 En otro aspecto, la divulgación proporciona un método para determinar el riesgo de un sujeto de albergar cáncer de pulmón, que comprende un panel de biomarcadores derivados de plasma y un panel de marcadores proteicos en donde el panel de biomarcadores derivados de plasma comprende diacetilpermina (DAS); en donde el panel de biomarcadores proteicos comprende CEA, CA125, CYFRA21-1 y pro-SFTPb; en donde el método comprende obtener una muestra biológica del sujeto; medir los niveles de los biomarcadores derivados de plasma y los biomarcadores proteicos en la muestra biológica; en donde la cantidad de los biomarcadores derivados de plasma y los biomarcadores proteicos clasifica al sujeto como que está en riesgo de albergar cáncer de pulmón o no en riesgo de albergar cáncer de pulmón.

25 En otro aspecto, la divulgación proporciona un método para determinar el riesgo de un sujeto de albergar cáncer de pulmón, que comprende determinar los niveles de uno o más biomarcadores proteicos y uno o más marcadores de metabolitos, comprendiendo dicho método obtener una muestra biológica del sujeto; poner en contacto la muestra con una primera molécula informadora que se une al antígeno CEA; poner en contacto la muestra con una segunda molécula informadora que se une al antígeno CA125; poner en contacto la muestra con una tercera molécula informadora que se une al antígeno CYFRA21-1; y poner en contacto la muestra con una cuarta molécula informadora que se une al antígeno pro-SFTPb; y determinar los niveles del uno o más biomarcadores, en donde el uno o más biomarcadores se seleccionan del grupo que consiste en diacetilpermina (DAS); en donde la cantidad de la primera molécula informadora, la segunda molécula informadora, la tercera molécula informadora, la cuarta molécula informadora y el uno o más biomarcadores clasifica al sujeto como que está en riesgo de albergar cáncer de pulmón o no en riesgo de albergar cáncer de pulmón.

30 La divulgación proporciona un método para determinar el riesgo de un sujeto de albergar cáncer de pulmón, que comprende obtener una muestra biológica del sujeto; medir los niveles de los antígenos CEA, CA125, CYFRA21-1 y pro-SFTPb en la muestra biológica; y medir los niveles de uno o más marcadores de metabolitos seleccionados del grupo que consiste en diacetilpermina (DAS) en la muestra biológica; asignar la afección del sujeto como en riesgo de albergar cáncer de pulmón o no en riesgo de albergar cáncer de pulmón, tal como se determina mediante análisis estadístico de los niveles de antígeno CEA, antígeno CA125, antígeno CYFRA21-1, antígeno pro-SFTPb y diacetilpermina (DAS) en la muestra biológica.

35 La detección de la cantidad de CEA, CA125, CYFRA21-1, pro-SFTPb, y diacetilpermina (DAS) comprende el uso de una partícula sólida. En otra realización, la partícula sólida es una perla. En otra realización, al menos una de las moléculas informadoras está unida a una enzima. En otra realización, al menos uno de los marcadores de proteína o metabolito genera una señal detectable. En otra realización, la señal detectable es detectable mediante un método espectrométrico. En otra realización, el método espectrométrico es espectrometría de masas. El método puede comprender la inclusión de la información del historial del paciente en la asignación de tener cáncer de pulmón o no tener cáncer de pulmón. El método puede comprender administrar al menos un ensayo de diagnóstico alternativo para un paciente asignado como que tiene cáncer de pulmón. El al menos un ensayo diagnóstico alternativo comprende un ensayo o secuenciación de al menos un ADNct.

Breve descripción de los dibujos

55 La **FIG. 1** representa curvas ROC para los 5 biomarcadores en CARET junto con el panel de 4 marcadores.

La **FIG. 2** representa curvas ROC para los 5 biomarcadores en CARET junto con el panel de 4 marcadores para 0-6 meses desde el diagnóstico.

La **FIG. 3** representa curvas ROC para los 5 biomarcadores en CARET junto con el panel de 4 marcadores para ≥ 6 -12 meses desde el diagnóstico.

La **FIG. 4** representa curvas ROC para el panel de 4 marcadores (línea gris oscuro), el panel de 4 marcadores con la adición de diacetilpermina (DAS) (línea discontinua) y DAS solo (línea gris claro) para todas las muestras analizadas (panel superior izquierdo), muestras de adenocarcinoma (panel superior derecho), cáncer de pulmón de células escamosas (panel inferior izquierdo) y muestras distintas de cáncer de pulmón de células escamosas y adenocarcinoma (panel inferior derecho).

Descripción detallada

En un aspecto, la invención proporciona un método para determinar el riesgo de un sujeto de albergar cáncer de pulmón, que comprende obtener una muestra biológica del sujeto; medir el nivel de CEA en la muestra biológica; medir el nivel de CA125 en la muestra biológica; medir el nivel de CYFRA21-1 en la muestra biológica; y medir el nivel de Pro-SFTPb en la muestra biológica; en donde la cantidad de CEA, CA125, CYFRA21-1 y Pro-SFTPb clasifica al sujeto como que está en riesgo de albergar cáncer de pulmón o no está en riesgo de albergar cáncer de pulmón.

En una realización, las cantidades de CEA, CA125, CYFRA21-1 y pro-SFTPb o las moléculas informadoras unidas a los mismos se elevan en el sujeto en relación con un sujeto sano.

En otra realización, dichos métodos comprenden, además: medir el nivel de diacetilpermina (DAS) en la muestra biológica; en donde la cantidad de diacetilpermina (DAS) clasifica al paciente como que está en riesgo de albergar cáncer de pulmón o no en riesgo de albergar cáncer de pulmón. En otra realización, la muestra comprende una muestra biológica seleccionada de sangre, plasma y suero. En otra realización, la muestra biológica es suero. En otra realización, se cuantifica la cantidad de CEA, CA125, CYFRA21-1 y pro-SFTPb. En otra realización, se miden las concentraciones de CEA, CA125, CYFRA21-1, Pro-SFTPb y diacetilpermina (DAS). En otra realización, se determina que el sujeto tiene cáncer de pulmón basándose en las concentraciones medidas de CEA, CA125, CYFRA21-1, y pro-SFTPb, y opcionalmente DAS. En otra realización, las concentraciones medidas se usan para calcular una puntuación de biomarcador basándose en los valores de sensibilidad y especificidad en un punto de corte expuesto en la Tabla 10. En otra realización, dichos métodos comprenden además las etapas de: comparar las concentraciones medidas de cada biomarcador en la muestra biológica con la predicción de un modelo estadístico. En otra realización, el panel se selecciona del grupo que consiste en: a. el panel que consiste en CEA, CA125, CYFRA21-1 y Pro-SFTPb; o b. el panel que consiste en CEA, CA125, CYFRA21-1, Pro-SFTPb y diacetilpermina (DAS).

En otra realización, la determinación de los niveles de CEA, CA125, CYFRA21-1 y pro-SFTPb se realiza sustancialmente al mismo tiempo. En otra realización, la determinación de los niveles de CEA, CA125, CYFRA21-1 y pro-SFTPb se realiza de manera escalonada.

En otra realización, dichos métodos comprenden además comparar la cantidad de CEA, CA125, CYFRA21-1 y pro-SFTPb con un valor de punto de corte como se ejemplifica en la Tabla 10.

En otra realización, al menos una de las superficies comprende además ensayar al menos una molécula receptora que se une selectivamente a un biomarcador o antígeno seleccionado de CEA, CA125, CYFRA21-1, pro-SFTPb y diacetilpermina (DAS); o en donde la cantidad de CEA, CA125, CYFRA21-1, pro-SFTPb y diacetilpermina (DAS) se detecta con el uso de una partícula sólida; o en donde al menos una de las moléculas informadoras está unida a una enzima; o en donde al menos una de las moléculas informadoras proporciona una señal detectable; o en donde el método comprende la inclusión de la información del historial del sujeto en la asignación de tener cáncer de pulmón o no tener cáncer de pulmón.

En otra realización, la señal detectable es detectable mediante un método seleccionado de espectroscopia UV-visible, espectrometría de masas, espectroscopia de resonancia magnética nuclear (RMN), espectroscopia de RMN de protones, espectrometría de resonancia magnética nuclear (RMN), cromatografía de gases-espectrometría de masas (CG-MS), cromatografía líquida-espectrometría de masas (LC-MS), espectroscopia de correlación (COSY), espectroscopia de efecto nuclear de Overhauser (NOESY), espectroscopia de efecto nuclear de Overhauser de marco rotatorio (ROESY), LC-TOF-MS, LC-MS/MS y electroforesis capilar-espectrometría de masas.

Los métodos presentados en la presente memoria permiten el cribado de sujetos de alto riesgo, tales como los que tienen antecedentes familiares de cáncer de pulmón, o sujetos con otros factores de riesgo tales como obesidad, tabaquismo intenso y posiblemente diabetes. El modelo de regresión logística descrito en la presente memoria puede incorporar estos factores en su método de clasificación.

Como se usa en la presente memoria, "estado de cáncer de pulmón" se refiere a la clasificación de un individuo, sujeto o paciente como que tiene cáncer de pulmón o que no tiene cáncer de pulmón. En algunas realizaciones, un individuo que tiene cáncer de pulmón puede denominarse "positivo para cáncer de pulmón". En otras realizaciones, un individuo que no tiene cáncer de pulmón puede denominarse "negativo para cáncer de pulmón". Para sujetos que se clasifican como positivos para cáncer de pulmón, se pueden proporcionar métodos adicionales para aclarar el estado del cáncer de pulmón. La clasificación como positivo para cáncer de pulmón puede ir seguida de métodos que incluyen, pero no se limitan a, tomografía computarizada (CT).

La detección de CEA puede realizarse por contacto con una biomolécula con la secuencia expuesta en la SEQ ID NO.: 1.

La detección de CA125 puede realizarse por contacto con una biomolécula con la secuencia expuesta en la SEQ ID NO.: 2.

- 5 La detección de CYFRA21-1 puede realizarse por contacto con una biomolécula con la secuencia expuesta en la SEQ ID NO.: 3.

La detección de Pro-SFTPb puede realizarse por contacto con una biomolécula con la secuencia expuesta en la SEQ ID NO.: 4.

- 10 Una combinación de al menos los cuatro biomarcadores CEA, CA125, CYFRA21-1 y Pro-SFTPb puede proporcionar un poder predictivo de cáncer de pulmón altamente fiable, previamente no visible. El modelo integrado de predicción de riesgo dio como resultado un AUC de 0,83 (IC del 95 %: 0,77-0,89), en comparación con un AUC de 0,72 (IC del 95 %: 0,65-0,79) de un modelo que solo incluía información de tabaquismo (valor p para la diferencia en AUC: 0,001). A la especificidad de USPSTF de 0,78 (IC del 95 %: 0,64-0,87), la sensibilidad para el modelo integrado de predicción de riesgo fue de 0,76 (IC del 95 %: 0,57-0,86), en comparación con 0,41 (IC del 95 %: 0,28-0,66) para el modelo de tabaquismo. Por el contrario, a la sensibilidad de USPSTF de 0,42 (IC del 95 %: 0,26-0,54), la especificidad del modelo integrado de predicción de riesgo fue de 0,94 (IC del 95 %: 0,88-0,98), en comparación con 0,78 (IC del 95 %: 0,70-0,90) del modelo de tabaquismo. Estas mejoras en las estimaciones de AUC, sensibilidad y especificidad se observaron de manera consistente a través de estratos relevantes definidos por el sexo y el estado de tabaquismo.

- 20 La divulgación no se limita a las biomoléculas específicas que se informan en la presente memoria para la detección de los biomarcadores. Se pueden elegir otras moléculas para su uso en otras realizaciones, incluyendo, pero sin limitación, biomoléculas basadas en proteínas, anticuerpos, ácidos nucleicos, aptámeros y compuestos orgánicos sintéticos. Otras moléculas pueden demostrar ventajas en términos de sensibilidad, eficiencia, velocidad de ensayo, coste, seguridad o facilidad de fabricación o almacenamiento.

- 25 En algunas realizaciones, se miden los niveles de CEA, CA125, CYFRA21-1 y Pro-SFTPb en una muestra biológica. En algunas realizaciones, CEA, CA125, CYFRA21-1 y Pro-SFTPb se ponen en contacto con moléculas informadoras, y se miden los niveles de las moléculas informadoras respectivas. En algunas realizaciones, se proporcionan cuatro moléculas informadoras que se unen específicamente a CEA, CA125, CYFRA21-1 y Pro-SFTPb, respectivamente. El uso de moléculas informadoras puede proporcionar ganancias en conveniencia y sensibilidad para el ensayo.

- 30 En algunas realizaciones, CEA, CA125, CYFRA21-1 y Pro-SFTPb se adsorben sobre una superficie que se proporciona en un kit. En algunas realizaciones, las moléculas informadoras se unen a CEA, CA125, CYFRA21-1 y Pro-SFTPb adsorbidos en la superficie. La adsorción de biomarcadores puede ser no selectiva o selectiva. En algunas realizaciones, la superficie comprende una funcionalidad de receptor para aumentar la selectividad hacia la adsorción de uno o más biomarcadores.

- 35 En algunas realizaciones, CEA, CA125, CYFRA21-1 y Pro-SFTPb se adsorben sobre cuatro superficies que son selectivas para uno o más de los biomarcadores. Una molécula informadora o múltiples moléculas informadoras pueden unirse entonces a biomarcadores adsorbidos en la superficie, y el nivel de molécula o moléculas informadoras asociadas con una superficie particular puede permitir una cuantificación fácil del biomarcador particular que está presente en esa superficie.

- 40 En algunas realizaciones, CEA, CA125, CYFRA21-1 y Pro-SFTPb se adsorben sobre una superficie que se proporciona en un kit; las moléculas transmisoras que son específicas para uno o más de estos biomarcadores se unen a los biomarcadores adsorbidos en la superficie; y las moléculas receptoras que son específicas para una o más moléculas transmisoras se unen a las moléculas transmisoras. Las moléculas transmisoras pueden proporcionar especificidad para ciertos biomarcadores, y las moléculas receptoras pueden permitir la detección.

- 45 En algunas realizaciones, se proporcionan cuatro moléculas transmisoras que se unen específicamente a CEA, CA125, CYFRA21-1 y Pro-SFTPb, respectivamente. Las moléculas transmisoras pueden diseñarse intencionadamente para que tengan especificidad hacia un biomarcador, o pueden seleccionarse de un conjunto de candidatos debido a sus propiedades de unión.

- 50 En algunas realizaciones, CEA, CA125, CYFRA21-1 y Pro-SFTPb se adsorben en cuatro superficies discretas que se proporcionan en un kit; las moléculas transmisoras que son específicas para uno o más de estos biomarcadores se unen a los biomarcadores adsorbidos en superficie; y las moléculas receptoras se unen a las moléculas transmisoras. El análisis de las superficies puede realizarse de una manera escalonada o concurrente.

En algunas realizaciones, la molécula informadora está unida a una enzima, facilitando la cuantificación de la molécula informadora. En algunas realizaciones, la cuantificación puede lograrse mediante la producción catalítica de una sustancia con propiedades espectroscópicas deseables.

- 55 En algunas realizaciones, la cantidad de biomarcador se determina con espectroscopía. En algunas realizaciones, la

espectroscopia que se utiliza es espectroscopia UV-visible. En algunas realizaciones, la espectroscopia que se utiliza es espectrometría de masas. En otras realizaciones, la espectroscopia que se utiliza es espectroscopia de resonancia magnética nuclear (RMN), tal como, incluyendo, pero sin limitarse a, espectroscopia de RMN de protones, espectrometría de resonancia magnética nuclear (RMN), cromatografía de gases-espectrometría de masas (CG-MS), cromatografía líquida-espectrometría de masas (LC-MS), espectroscopia de correlación (COSY), espectroscopia de efecto nuclear de Overhauser (NOESY), espectroscopia de efecto nuclear de Overhauser de marco rotatorio (ROESY), LC-TOF-MS, LC-MS/MS y electroforesis capilar-espectrometría de masas.

La cantidad de biomarcador o biomarcadores que se encuentra en un ensayo particular puede informarse directamente a un operador, o alternativamente puede almacenarse digitalmente y hacerse fácilmente disponible para el procesamiento matemático. Se puede proporcionar un sistema para realizar el análisis matemático, y puede informar además de la clasificación como positivo para cáncer de pulmón o negativo para cáncer de pulmón a un operador.

El cáncer de pulmón es cáncer de pulmón de células no pequeñas (NSCLC).

Los diversos componentes del sistema analizados en la presente memoria pueden incluir uno o más de los siguientes: un ordenador que comprende uno o más procesadores para procesar datos digitales; memoria digital a corto o largo plazo; un convertidor analógico-digital de entrada para proporcionar datos digitalizados; un programa de aplicación puesto a disposición del procesador para dirigir el procesamiento de datos digitales por el procesador; un dispositivo de entrada para recoger información del sujeto u operador, y un dispositivo de salida para mostrar la información al sujeto u operador.

Con respecto a la detección de los biomarcadores detallados en la presente memoria, la divulgación no se limita a las biomoléculas específicas notificadas en la presente memoria. En algunas realizaciones, se pueden elegir otras biomoléculas para la detección y análisis de los biomarcadores divulgados que incluyen, pero no se limitan a, biomoléculas basadas en proteínas, anticuerpos, ácidos nucleicos, aptámeros y compuestos orgánicos sintéticos. Otras moléculas pueden demostrar ventajas en términos de sensibilidad, eficiencia, velocidad de ensayo, coste, seguridad o facilidad de fabricación o almacenamiento. A este respecto, los expertos en la técnica apreciarán que el poder de predicción y diagnóstico de los biomarcadores divulgados en la presente memoria puede extenderse al análisis no solo de la forma proteica de estos biomarcadores, sino también de otras representaciones de los biomarcadores (por ejemplo, ácido nucleico). Además, los expertos en la técnica apreciarán que el poder de predicción y diagnóstico de los biomarcadores divulgados en la presente memoria también puede usarse en combinación con un análisis de otros biomarcadores asociados con el cáncer de pulmón. En algunas realizaciones, otros biomarcadores asociados con el cáncer de pulmón pueden ser biomarcadores basados en proteínas. En algunas realizaciones, otros biomarcadores asociados con el cáncer de pulmón pueden ser biomarcadores no basados en proteínas, tales como, por ejemplo, ADNct.

Lo anterior ha esbozado de manera bastante amplia las características y beneficios técnicos de la divulgación con el fin de que la descripción detallada pueda entenderse mejor. Los expertos en la técnica deben apreciar que las realizaciones específicas divulgadas pueden utilizarse fácilmente como base para modificar o diseñar otras estructuras o procesos para llevar a cabo los mismos propósitos de la divulgación.

Definiciones

Como se usa en la presente memoria, el término "cáncer de pulmón" se refiere a un neoplasma maligno del pulmón caracterizado por la proliferación anormal de células, excediendo el crecimiento de estas células, y de manera descoordinada con, el de los tejidos normales que las rodean.

Como se usa en la presente memoria, el término "positivo para cáncer de pulmón" se refiere a la clasificación de un sujeto como que tiene cáncer de pulmón.

Como se usa en la presente memoria, el término "negativo para cáncer de pulmón" se refiere a la clasificación de un sujeto como que no tiene cáncer de pulmón.

Como se usa en la presente memoria, los términos "sujeto" o "paciente" se refieren a un mamífero, preferiblemente un ser humano, para el que se desea una clasificación como positivo para cáncer de pulmón o negativo para cáncer de pulmón, y para el que puede proporcionarse tratamiento adicional.

Como se usa en la presente memoria, un "paciente de referencia", "sujeto de referencia" o "grupo de referencia" se refiere a un grupo de pacientes o sujetos con los que se puede comparar una muestra de ensayo de un paciente o sujeto sospechoso de tener o estar en riesgo de albergar cáncer de pulmón. En algunas realizaciones, dicha comparación puede usarse para determinar si el sujeto de ensayo tiene cáncer de pulmón. Un paciente o grupo de referencia puede servir como control para fines de ensayo o diagnóstico. Como se describe en la presente memoria, un paciente o grupo de referencia puede ser una muestra obtenida de un único paciente, o puede representar un grupo de muestras, tal como un grupo combinado de muestras.

Como se usa en la presente memoria, "sano" se refiere a un individuo en el que no se encuentra evidencia de cáncer de pulmón, es decir, el individuo no tiene cáncer de pulmón. Dicho individuo puede clasificarse como "negativo para

cáncer de pulmón" o como que tiene pulmones sanos, o función pulmonar normal, no comprometida. Un paciente o sujeto sano no tiene síntomas de cáncer de pulmón u otra enfermedad pulmonar. En algunas realizaciones, un paciente o sujeto sano puede usarse como paciente de referencia para la comparación con muestras enfermas o sospechosas de ser enfermas para la determinación de cáncer de pulmón en un paciente o un grupo de pacientes.

- 5 Como se usa en la presente memoria, los términos "tratamiento" o "tratar" se refieren a la administración de medicina o la realización de procedimientos médicos con respecto a un sujeto, para profilaxis (prevención) o para curar o reducir el grado o probabilidad de aparición o recurrencia de la insuficiencia o enfermedad o afección o evento en el caso en donde el sujeto o paciente lo padezca. En relación con la presente divulgación, el término también puede significar la administración de sustancias o formulaciones farmacológicas, o la realización de métodos no farmacológicos que incluyen, pero no se limitan a, radioterapia y cirugía. Las sustancias farmacológicas como se usan en la presente memoria pueden incluir, pero no se limitan a, agentes quimioterapéuticos que están establecidos en la técnica, tales como Erlotinib (TARCEVA y otros), Afatinib (GILOTRIF), Gefitinib (IRESSA), Bevacizumab (AVASTIN), Crizotinib (XALKORI), Ceritinib (ZYKADIA), cisplatino (PLATINOL), carboplatino (PARAPLATIN), docetaxel (TAXOTERE), gemcitabina (GEMZAR), paclitaxel (TAXOL y otros), vinorelbina (NAVELBINE y otros), o pemetrexed (ALIMTA). Las sustancias farmacológicas pueden incluir sustancias usadas en inmunoterapia, tales como inhibidores de puntos de control. El tratamiento puede incluir una multiplicidad de sustancias farmacológicas, o una multiplicidad de métodos de tratamiento, incluyendo, pero sin limitarse a, cirugía y quimioterapia.

Como se usa en la presente memoria, el término "CARET" se refiere al estudio del Ensayo de Eficacia de Betacaroteno y Retinol.

- 20 Como se usa en la presente memoria, el término "NLST" se refiere al Ensayo Nacional de Cribado Pulmonar.

Como se utiliza en la presente memoria, el término "USPSTF" se refiere al Grupo de Trabajo de Servicios Preventivos de los Estados Unidos.

Como se usa en la presente memoria, el término "EPIC" se refiere a la Investigación Prospectiva Europea sobre el Cáncer y la Nutrición.

- 25 Como se usa en la presente memoria, el término "NSHDS" se refiere al Estudio de Salud y Enfermedades del Norte de Suecia.

Como se usa en la presente memoria, el término "ICD-O-2" se refiere a la Clasificación Internacional de Enfermedades para Oncología, Segunda Edición.

- 30 Como se usa en la presente memoria, el término "ELISA" se refiere a un ensayo inmunoabsorbente ligado a enzimas. Este ensayo implica generalmente poner en contacto una muestra de proteínas marcada fluorescentemente con anticuerpos que tienen afinidad específica por esas proteínas. La detección de estas proteínas puede realizarse con una variedad de medios, incluyendo, pero sin limitación, fluorimetría láser.

- 35 Como se usa en la presente memoria, el término "regresión" se refiere a un método estadístico que puede asignar un valor predictivo para una característica subyacente de una muestra basándose en un rasgo observable (o conjunto de rasgos observables) de dicha muestra. En algunas realizaciones, la característica no es observable directamente. Por ejemplo, los métodos de regresión usados en la presente memoria pueden vincular un resultado cualitativo o cuantitativo de un ensayo de biomarcador particular, o conjunto de ensayos de biomarcadores, en un determinado sujeto, a una probabilidad de que dicho sujeto sea positivo para cáncer de pulmón.

- 40 Como se usa en la presente memoria, el término "regresión logística" se refiere a un método de regresión en donde la asignación de una predicción a partir del modelo puede tener uno de varios valores discretos permitidos. Por ejemplo, los modelos de regresión logística usados en la presente memoria pueden asignar una predicción, para un determinado sujeto, de positivo para cáncer de pulmón o negativo para cáncer de pulmón.

- 45 Como se usa en la presente memoria, el término "puntuación de biomarcador" se refiere a una puntuación numérica para un sujeto particular que se calcula introduciendo los niveles de biomarcador particular para dicho sujeto en un método estadístico.

- 50 Como se usa en la presente memoria, el término "puntuación compuesta" se refiere a una suma de los valores normalizados para los marcadores predeterminados medidos en la muestra del sujeto. En una realización, los valores normalizados se informan como una puntuación de biomarcador y esos valores de puntuación de biomarcador se suman después para proporcionar una puntuación compuesta para cada sujeto sometido a ensayo. Cuando se usa en el contexto de la tabla de categorización de riesgos y se correlaciona con una agrupación estratificada basada en un rango de puntuaciones compuestas en la Tabla de Categorización de Riesgos, la "puntuación compuesta" se usa para determinar la "puntuación de riesgo" para cada sujeto sometido a ensayo, en donde el multiplicador que indica una probabilidad aumentada de tener el cáncer para la agrupación estratificada se convierte en la "puntuación de riesgo".

- 55 Como se usa en la presente memoria, el término "puntuación de riesgo" se refiere a un único valor numérico que indica un aumento del riesgo de un sujeto humano asintomático de albergar un cáncer en comparación con la prevalencia

- conocida de cáncer en la cohorte de enfermedad. En ciertas realizaciones, la puntuación compuesta se calcula para un sujeto humano y se correlaciona con un multiplicador que indica un mayor riesgo de albergar el cáncer, en donde la puntuación compuesta se correlaciona basándose en el rango de puntuaciones compuestas para cada agrupación estratificada en la tabla de categorización de riesgos. De esta manera, la puntuación compuesta se convierte en una puntuación de riesgo basándose en el multiplicador que indica una mayor probabilidad de tener el cáncer para la agrupación que tiene la mejor coincidencia para la puntuación compuesta.
- 5 Como se usa en la presente memoria, el término "corte" o "punto de corte" se refiere a un valor matemático asociado con un método estadístico específico que puede usarse para asignar una clasificación de positivo para cáncer de pulmón o negativo para cáncer de pulmón a un sujeto, basándose en la puntuación del biomarcador de dicho sujeto.
- 10 Como se usa en la presente memoria, cuando un valor numérico por encima o por debajo de un valor de corte "es característico del cáncer de pulmón", lo que significa es que el sujeto, cuyo análisis de la muestra produjo el valor, tiene cáncer de pulmón o está en riesgo de albergar cáncer de pulmón.
- Como se usa en la presente memoria, un sujeto que está "en riesgo de albergar cáncer de pulmón" es uno que todavía puede no evidenciar síntomas manifiestos de cáncer de pulmón, pero que está produciendo niveles de biomarcadores que indican que el sujeto tiene cáncer de pulmón, o puede desarrollarlo a corto plazo. Un sujeto que tiene cáncer de pulmón o se sospecha que alberga cáncer de pulmón puede tratarse para el cáncer o cáncer sospechoso.
- 15 Como se usa en la presente memoria, el término "clasificación" se refiere a la asignación de un sujeto como positivo para cáncer de pulmón o negativo para cáncer de pulmón, basándose en el resultado de la puntuación del biomarcador que se obtiene para dicho sujeto.
- 20 Como se usa en la presente memoria, el término "positivo para cáncer de pulmón" se refiere a una indicación de que se predice que un sujeto está en riesgo de albergar cáncer de pulmón, basándose en los resultados del resultado de los métodos de la divulgación.
- Como se usa en la presente memoria, el término "negativo para cáncer de pulmón" se refiere a una indicación de que se predice que un sujeto no está en riesgo de albergar cáncer de pulmón, basándose en los resultados del resultado de los métodos de la divulgación.
- 25 Como se usa en la presente memoria, el término "prueba de suma de rangos de Wilcoxon", también conocido como la prueba de la U de Mann-Whitney, la prueba de Mann-Whitney-Wilcoxon o la prueba de Wilcoxon-Mann-Whitney, se refiere a un método estadístico específico usado para la comparación de dos poblaciones. Por ejemplo, la prueba puede usarse en la presente memoria para vincular un rasgo observable, en particular un nivel de biomarcador, con la ausencia o presencia de cáncer de pulmón en sujetos de una cierta población.
- 30 Como se usa en la presente memoria, el término "tasa positiva verdadera" se refiere a la probabilidad de que un sujeto dado clasificado como positivo por un cierto método sea verdaderamente positivo.
- Como se usa en la presente memoria, el término "tasa de falsos positivos" se refiere a la probabilidad de que un sujeto dado clasificado como positivo por un cierto método sea verdaderamente negativo.
- 35 Como se usa en la presente memoria, el término "sensibilidad" se refiere, en el contexto de diversos ensayos bioquímicos, a la capacidad de un ensayo para identificar correctamente aquellos con una enfermedad (es decir, la tasa positiva verdadera). En comparación, como se usa en la presente memoria, el término "especificidad" se refiere, en el contexto de diversos ensayos bioquímicos, a la capacidad de un ensayo para identificar correctamente aquellos sin la enfermedad (es decir, la tasa negativa verdadera). La sensibilidad y la especificidad son medidas estadísticas del rendimiento de una prueba de clasificación binaria (es decir, función de clasificación). La sensibilidad cuantifica la evitación de falsos negativos, y la especificidad hace lo mismo para falsos positivos.
- 40 Como se usa en la presente memoria, una "muestra" se refiere a una sustancia de ensayo que se va a ensayar para determinar la presencia de, y niveles o concentraciones de la misma, de un biomarcador como se describe en la presente memoria. Una muestra puede ser cualquier sustancia apropiada según la presente divulgación, incluyendo, pero sin limitarse a, sangre, suero sanguíneo, plasma sanguíneo o cualquier parte de los mismos.
- 45 Como se usa en la presente memoria, un "antígeno" se refiere a una proteína, metabolito u otra molécula a la que un anticuerpo o reactivo o fragmento de unión a antígeno puede unirse para la detección de un biomarcador como se describe en la presente memoria. En algunas realizaciones, un biomarcador puede servir como antígeno. En otras realizaciones, una porción de un biomarcador puede servir como antígeno. En algunas realizaciones, se puede usar un anticuerpo para la detección de un antígeno como se describe en la presente memoria. En otras realizaciones, puede usarse un ácido nucleico, tal como ADN, ARN, híbridos de ADN/ARN, anticuerpos, fragmentos de anticuerpo o cualquier otro compuesto o molécula capaz de unirse a un antígeno, para detectar un antígeno, tal como un biomarcador como se describe en la presente memoria. Un antígeno como se describe en la presente memoria puede servir como base para la detección de los niveles, concentraciones o cantidades de un marcador de proteína o metabolito para su uso con los métodos como se describe en la presente memoria.
- 50
- 55

Como se usa en la presente memoria, el término "CEA" se refiere al antígeno carcinoembrionario.

Como se usa en la presente memoria, el término "CA125" se refiere al antígeno 125 de cáncer.

Como se usa en la presente memoria, el término "CYFRA21-1", también conocido como Cyfra 21-1, se refiere al fragmento de citoqueratina 19, también conocido como fragmento de citoqueratina-19.

5 Como se usa en la presente memoria, el término "SFTPB" se refiere a la proteína tensioactiva B.

Como se usa en la presente memoria, el término "Pro-SFTPB", se refiere a la Pro-proteína tensioactiva B, que es una forma precursora de SFTPB.

Como se usa en la presente memoria, el término "HE4", también conocido como WFDC2, se refiere a la proteína 4 de epididimitis humana.

10 Como se usa en la presente memoria, el término "ADNct" se refiere a ADN tumoral libre de células o circulante. El ADNct es ADN tumoral que se encuentra circulando libremente en la sangre de un paciente con cáncer. Sin estar limitado por la teoría, se cree que el ADNct se origina a partir de células tumorales moribundas y puede estar presente en una amplia gama de cánceres, pero a niveles variables y fracciones alélicas mutantes. Generalmente, el ADNct porta mutaciones somáticas únicas formadas en la célula tumoral de origen y que no se encuentran en las células sanas del huésped.

15 Como tales, las mutaciones somáticas del ADNct pueden actuar como biomarcadores específicos de cáncer.

Como se usa en la presente memoria, un "metabolito" se refiere a moléculas pequeñas que son intermedios y/o productos del metabolismo celular. Los metabolitos pueden realizar una variedad de funciones en una célula, por ejemplo, efectos estructurales, de señalización, estimuladores y/o inhibidores sobre enzimas. En algunas realizaciones, un metabolito puede ser un marcador de metabolitos no proteicos derivados del plasma, tal como, incluyendo, pero sin limitarse a, acetilsermidina, diacetilsermina, lisofosfatidilcolina (18:0), lisofosfatidilcolina (20:3) y un derivado de indol.

20 Como se usa en la presente memoria, el término "ROC" se refiere a la característica operativa del receptor, que es una representación gráfica usada en la presente memoria para calibrar el rendimiento de un determinado método de diagnóstico en diversos puntos de corte. Se puede construir un gráfico ROC a partir de la fracción de positivos verdaderos y falsos positivos en diversos puntos de corte.

25 Como se usa en la presente memoria, el término "AUC" se refiere al área bajo la curva del gráfico ROC. El AUC puede usarse para estimar la potencia predictiva de un determinado ensayo diagnóstico. Generalmente, un AUC mayor corresponde a un aumento de la potencia predictiva, con una disminución de la frecuencia de errores de predicción. Los valores posibles de AUC varían de 0,5 a 1,0, siendo este último valor característico de un método de predicción sin errores.

30 Como se usa en la presente memoria, el término "valor p" o "p" se refiere a la probabilidad de que las distribuciones de las puntuaciones de biomarcadores para sujetos positivos para cáncer de pulmón y negativos para cáncer de pulmón sean idénticas en el contexto de una prueba de suma de rangos de Wilcoxon. Generalmente, un valor p cercano a cero indica que un método estadístico particular tendrá un alto poder predictivo en la clasificación de un sujeto.

35 Como se usa en la presente memoria, el término "CI" se refiere a un intervalo de confianza, es decir, un intervalo en donde se puede predecir que un cierto valor se encuentra con un cierto nivel de confianza. Como se usa en la presente memoria, el término "CI del 95 %" se refiere a un intervalo en donde se puede predecir que un cierto valor se encuentra con un nivel de confianza del 95 %.

40 El término "AIC" se refiere al Criterio de Información de Akaike, un método basado en la teoría de la información que se puede usar para evaluar los méritos relativos de varios modelos de regresión para un conjunto de datos.

Ejemplos

Los siguientes ejemplos se incluyen para demostrar realizaciones de la divulgación. Los siguientes ejemplos se presentan solo a modo de ilustración y para ayudar a un experto en la técnica a usar la divulgación. Los ejemplos no pretenden, de ninguna manera, limitar de otra forma el alcance de la divulgación.

45 Ejemplo 1: Conjuntos de muestras de sangre.

Las muestras de plasma de todos los participantes del estudio se enviaron en hielo seco al laboratorio en MD Anderson, TX, EE. UU., donde se mantuvieron por debajo de -80 °C hasta el análisis.

Cohorte de entrenamiento: Estudio del Ensayo de Eficacia de Betacaroteno y Retinol (CARET).

50 Se utilizaron muestras de suero de antes del diagnóstico de participantes en CARET (fumadores empedernidos actuales o anteriores) para entrenar la puntuación de biomarcadores de posibles marcadores de predicción de riesgo basados en proteínas. El CARET fue un ensayo aleatorizado, doble ciego, controlado con placebo que evaluó la

eficacia de prevención del cáncer y la seguridad de la suplementación diaria con 30 mg de betacaroteno y 25.000 UI de palmitato de retinol en 18.314 personas con alto riesgo de cáncer de pulmón. Los participantes elegibles incluyeron dos poblaciones de alto riesgo: 14.254 hombres y mujeres de 50 a 69 años de edad que eran fumadores actuales o anteriores (que lo dejaron dentro de los 6 años previos) que estaban expuestos a al menos 20 paquetes por año de fumar cigarrillos, y 4.060 hombres de 45 a 69 años de edad que eran fumadores actuales o anteriores (que lo dejaron no más de 15 años antes del inicio del estudio) y tenían un historial sustancial de exposición al asbesto ocupacional. Los participantes se incluyeron en 6 centros de EE. UU. de 1985 a 1994 y se siguieron para determinar los resultados de cáncer y mortalidad hasta 2005. En total, se ensayaron muestras de 108 sujetos que posteriormente desarrollaron cáncer de pulmón de células no pequeñas (NSCLC) en el plazo de 12 meses después de proporcionar una muestra de sangre, y 216 controles coincidentes con cada caso basándose en la edad en la línea base (grupos de 5 años), el sexo, el estado de tabaquismo en la línea base (actual frente a anterior) y el periodo de registro del estudio.

Se utilizaron muestras de suero de antes del diagnóstico de participantes en CARET (fumadores empedernidos actuales o anteriores), que consistían en 108 sujetos que posteriormente desarrollaron cáncer de pulmón de células no pequeñas (NSCLC) y 216 controles coincidentes para ensayar un conjunto de marcadores de predicción de riesgo potenciales basados en proteínas que incluyen Pro-SFTPB. Ninguna de estas muestras fue parte de ningún ensayo previo de los marcadores candidatos a ensayar y, por tanto, representaban un conjunto independiente para ensayos de marcadores que consistían en 108 NSCLC y 216 controles coincidentes. Se recogieron muestras de casos hasta un año antes de un diagnóstico de NSCLC.

Cohorte de validación: Estudios EPIC y NSHDS.

La cohorte de validación usada en este estudio se obtuvo de dos estudios prospectivos europeos: el estudio de Investigación Prospectiva Europea de Cáncer y Nutrición (EPIC) y el Estudio de Salud y Enfermedad del Norte de Suecia (NSHDS).

Investigación Prospectiva Europea de Cáncer y Nutrición

El estudio EPIC es una cohorte prospectiva multicentro en curso que reclutaba 521.330 participantes entre 1992 y 1998 de 23 centros en diez países en Europa. El estudio actual de biomarcadores implicaba a participantes de EPIC de 7 países (Grecia, Países Bajos, Reino Unido, Francia, Alemania, España e Italia) que donaron una muestra de sangre en el reclutamiento del estudio, que era parte de un repositorio biológico, mantenida en nitrógeno líquido en la Agencia Internacional para Investigación sobre el Cáncer (IARC) en Lyon.

Estudio de Salud y Enfermedad del Norte de Suecia

El NSHDS es una cohorte prospectiva en curso de la población general del Condado de Västerbotten en el norte de Suecia. Después del inicio del estudio en 1985, todos los residentes del condado han sido invitados desde entonces a participar atendiendo a un control de salud a los 40, 50 y 60 años de edad. A partir de 2014, la cohorte había reclutado 99.404 participantes del estudio que donaron una muestra de sangre para investigación futura. Los detalles sobre los procedimientos de reclutamiento de EPIC y NSHDS, la recogida de cuestionarios y datos antropométricos, y la recogida y almacenamiento de muestras de sangre se han descrito en detalle en otra parte. Las características para las cohortes de validación de EPIC y NSHDS se presentan a continuación en la Tabla 1.

Las características en la línea base de los casos y controles en el estudio de entrenamiento CARET y los estudios de validación EPIC y NSHDS se presentan en la Tabla 1. Las características de la muestra de estudio de EPIC y NSHDS usada para entrenar el modelo de tabaquismo se presentan en la Tabla 2.

Tabla 1. Características clínicas de pacientes y controles en cohortes de entrenamiento y validación.

		Estudio de entrenamiento (CARET)		Estudio de validación (EPIC y NSHDS)	
		Casos	Controles	Casos	Controles
Globalmente	N (%)	108	216	83	158
Sexo	Hombre	75 (69,4)	150 (69,4)	50 (60,2)	93 (58,9)
	Mujer	33 (30,6)	66 (30,6)	33 (39,8)	65 (41,1)
Raza	Blanco	99	200		
	Negro	6	8		
	Otros	3	8		
Edad, años	≤40			3 (3,6)	6 (3,8)
	40-50	2 (1,9)	4 (1,9)	7 (8,4)	13 (8,2)

		Estudio de entrenamiento (CARET)		Estudio de validación (EPIC y NSHDS)	
	N (%)	Casos	Controles	Casos	Controles
	50-60	35 (32,4)	72 (33,3)	40 (48,2)	78 (49,4)
	60-70	69 (63,9)	136 (63,0)	29 (34,9)	54 (34,2)
	>70	2 (1,9)	4 (1,9)	4 (4,8)	7 (4,4)
Paquetes-año	Media	54	49		
	SD	23	20		
Edad en el diagnóstico, años	Media	65,1			
	SD	6,3			
	≤40	-	-		
	40-50	2	4		
	50-60	35	72		
	60-70	69	136		
	>70	2	4		
Años desde la recogida de sangre hasta el diagnóstico	0-0.5	40 (37,0)	-	31 (37,3)	
	0.5-1	68 (63,0)	-	33 (39,8)	
	1-2	-	-	19 (22,9)	
	2-5	-	-	-	
	5-10	-	-	-	
Estado de tabaquismo	Nunca	-	-	9 (10,8)	50 (31,7)
	Anterior	36 (33,3)	72 (33,3)	26 (31,3)	53 (33,5)
	Actual	72 (66,7)	144 (66,7)	48 (57,8)	55 (34,8)
Subtipo histológico	ADC	40 (37,0)	-	32 (38,6)	
	SCC	38 (35,2)	-	21 (25,3)	
	Otros	30 (27,8)	-	30 (36,1)	
Estadio	I y II	26			
	III y IV	64			
	Desconocido	18			
Elegible para el cribado del cáncer de pulmón (USPSTF)	No elegible	29 (26,9)	57 (26,4)	51 (63,0)	132 (85,2)
	Elegible	79 (73,1)	159 (73,6)	30 (37,0)	23 (14,8)
	N/A	-	-	2	3

Abreviaturas: NSCLC, cáncer de pulmón de células no pequeñas; SCC, carcinoma de células escamosas; ADC, adenocarcinoma; SD, desviación estándar

Tabla 2. Características en la línea base del estudio EPIC y NSHDS con las que se usaron para entrenar el modelo de tabaquismo.

	N (%)	Muestra de entrenamiento para el modelo de tabaquismo (EPIC+NSHDS)	
		Casos	Controles
Globalmente		1.008	1.873
Sexo	Hombre	605 (60)	1.088 (58,1)
	Mujer	403 (40)	785 (41,9)
Edad, años	≤40	20 (2,0)	45 (2,4)
	40-50	167 (16,6)	315 (16,8)
	50-60	430 (42,7)	787 (42,0)
	60-70	315 (31,2)	599 (32,0)
	>70	76 (7,5)	127 (6,8)
Años desde la recogida de sangre hasta el diagnóstico	0-0.5		
	0.5-1		
	1-2		
	2-5	351 (34,8)	
	5-10	657 (65,2)	
Estado de tabaquismo	Nunca	122 (12,1)	524 (28,0)
	Anterior	296 (29,4)	606 (32,3)
	Actual	590 (58,5)	743 (39,7)
Subtipo histológico	Adenocarcinoma	366 (36,3)	
	Carc. de células escamosas.	200 (19,8)	
	Otros	442 (43,6)	
Elegible para el cribado del cáncer de pulmón (USPSTF)	No elegible	688 (68,7)	1.599 (86,6)
	Elegible	313 (31,3)	248 (13,4)
	N/A	7	26

Seguimiento y selección de casos y controles

- 5 El seguimiento de casos de cáncer incidente se realizó usando una combinación de métodos, incluyendo el vínculo de registros con registros regionales o nacionales de cáncer, así como registros de seguros de salud, registros de cáncer y patología, y seguimiento activo a través de sujetos de estudio y sus parientes más próximos. Los casos de cáncer de pulmón se definieron basándose en la Clasificación Internacional de Enfermedades para Oncología, Segunda Edición (ICD-O-2), e incluyeron todos los cánceres invasivos que se codificaron como C34. Después de
- 10 excluir los casos que tenían un historial de otro cáncer (excepto cáncer de piel no melanoma) y los casos con falta de muestras de sangre o información de tabaquismo, 48 casos incidentes diagnosticados dentro de 1 año de extracción de sangre en EPIC y 35 casos incidentes diagnosticados dentro de 2 años de extracción de sangre en NSHDS estaban disponibles para el estudio actual. En Italia, Países Bajos, España, Suecia, y el Reino Unido, se identificaron casos de cáncer incidente a través de un vínculo de registro con registros de cáncer regionales o nacionales. En Francia,
- 15 Alemania y Grecia, el seguimiento se basó en una combinación de métodos, incluyendo registros de seguros de salud, registros de cáncer y patología, y seguimiento activo a través de sujetos de estudio y sus parientes más próximos.

Para cada caso índice, se eligieron dos controles al azar de conjuntos de riesgo que consistían en todos los miembros de la cohorte vivos y libres de cáncer (excepto cáncer de piel no melanoma) en el momento del diagnóstico del caso índice. Los criterios coincidentes fueron centro de estudio, sexo, fecha de extracción de sangre (± 1 mes, relajado a ± 12 meses para conjuntos sin controles disponibles), tiempo en la extracción de sangre (± 1 hora, relajado a ± 12 horas)

20

y edad en la extracción de sangre (± 3 meses, relajado a ± 5 años). Con el fin de mejorar la potencia estadística en los análisis estratificados con tabaquismo, cada uno de los controles se coincidió adicionalmente basándose en el estado de tabaquismo del caso índice de 5 categorías; nunca fumadores, que lo habían dejado a corto y largo plazo entre fumadores anteriores (<10 años, ≥ 10 años desde que lo dejaron), y fumadores ligeros y empedernidos entre fumadores actuales (<15 años, ≥ 15 cigarrillos por día).

El estudio de validación combinada final de EPIC (casos diagnosticados dentro de 1 año de la extracción de sangre) y NSHDS (casos diagnosticados dentro de 2 años de la extracción de sangre) incluyó 83 casos de cáncer de pulmón incidente y 158 controles coincidentes. Todos los participantes del estudio dieron su consentimiento informado por escrito para participar en el estudio y la investigación fue aprobada por los comités de ética locales en los países participantes, así como los Comités de Revisión Ética de IARC y MD Anderson.

Para cada caso de índice de EPIC y NSHDS, se eligieron dos controles al azar de conjuntos de riesgo que consistían en todos los miembros de la cohorte vivos y libres de cáncer (excepto cáncer de piel no melanoma) en el momento del diagnóstico del caso de índice. Para cada caso, se eligió aleatoriamente un control de conjuntos de riesgo que consistían en todos los miembros de la cohorte vivos y libres de cáncer (excepto cáncer de piel no melanoma) en el momento del diagnóstico del caso índice. Los criterios coincidentes fueron centro de estudio, sexo, fecha de recogida de sangre (± 1 mes, relajado a ± 3 meses para conjuntos sin controles disponibles) y fecha de nacimiento (± 1 año, relajado a ± 3 años). Con el fin de mejorar la potencia estadística en los análisis estratificados con tabaquismo, uno de los controles coincidía adicionalmente con el estado de tabaquismo del caso índice de 5 categorías; nunca fumadores, que lo habían dejado a corto y largo plazo entre fumadores anteriores (<10 años, ≥ 10 años desde que lo dejaron), y fumadores ligeros y empedernidos entre fumadores actuales (<15 años, ≥ 15 cigarrillos por día).

La cohorte de validación combinada total de EPIC y NSHDS incluyó 74 casos de cáncer de pulmón incidente y 109 controles coincidentes después de excluir los casos y controles de nunca fumadores. Todos los participantes dieron su consentimiento informado por escrito para participar en el estudio y la investigación fue aprobada por los comités de ética locales en los países participantes y el Comité de Revisión Institucional de la IARC.

Ejemplo 2: Selección de candidatos de panel de biomarcadores.

Se realizó el cribado inicial de biomarcadores en un conjunto de cinco candidatos: CEA, CA125, CYFRA 21-1, Pro-SFTPB y HE4. CEA y CA125 son biomarcadores tumorales bien conocidos y se usan rutinariamente como marcadores de pronóstico y diagnóstico de cáncer colorrectal y de ovario, respectivamente. Desde los 90, CYFRA 21-1 se ha implicado como marcador de cáncer de pulmón de células no pequeñas. SFTPB y HE4 están significativamente asociados con el cáncer de pulmón tanto en modelos de cáncer de pulmón en ratón como en muestras de sangre recogidas hasta un año antes del diagnóstico clínico, en comparación con controles coincidentes (Bach et al., J Natl Cancer Inst 95:470-8, 2003). Una forma precursora de SFTPB (Pro-SFTPB) era predictiva del riesgo de cáncer de pulmón en estudios de cohorte prospectivos que englobaban el Estudio Canadiense de la Detección Temprana General de Cáncer de Pulmón, el Estudio de Salud de los Médicos y el estudio CARET (Cassidy et al., Br J Cancer 98:270-6, 2008; Hoggart et al., Cancer Prev Res (Phila) 5:834-46, 2012; Spitz et al., Cancer Prev Res (Phila) 1:250-4, 2008). Originalmente, se encontró que pro-SFTPB estaba elevado en plasma de ratones con cáncer de pulmón.

De este conjunto, se seleccionaron cuatro biomarcadores para un panel basándose en su rendimiento en la cohorte de entrenamiento: CA-125, CYFRA 21-1, CEA y Pro-SFTPB.

Los niveles observados de biomarcadores para la población de control se exponen en la Tabla 3. También se proporcionan rangos típicos para estos valores en sujetos sanos.

Tabla 3. Niveles de biomarcadores observados en la población de control.

Biomarcador	Mín*	Mediana	Máx	Rango típico
CEA	0,47	1,96	10	0-5
CA125	4,33	6,7	12,13	0-35
CYFRA	2,51	4,24	15,89	0,16-2,95
Pro-SFTPB	0	5,2	93,23	N/A

*Unidades en ng/ml para CEA, CYFRA y Pro-SFTPB; Unidades en U/ml para CA125.

Los niveles observados de biomarcadores para la población de cáncer se exponen en la Tabla 4. También se proporcionan rangos típicos para estos valores en sujetos afectados por cáncer.

Tabla 4. Niveles de biomarcadores observados en la población de cáncer.

Biomarcador	Mín*	Mediana	Máx	Rango típico
CEA	0,62	9,26	120,6	0-20
CA125	4,24	7,29	20,51	>35
CYFRA	3,3	5,66	20,93	0,93-221,6
Pro-SFTPb	1,08	22,66	233,11	N/A

*Unidades en ng/ml para CEA, CYFRA y Pro-SFTPb; Unidades en U/ml para CA125.

Ejemplo 3: Análisis de muestras.

Los cinco marcadores candidatos se analizaron por inmunoensayo usando muestras de suero de la cohorte de entrenamiento. Las concentraciones para CA125, CEA, Pro-SFTPb, CYFRA 21-1 y HE4 se determinaron usando inmunoensayos basados en perlas en el instrumento MAGPIX(R) (Luminex Corporation, Austin TX). Las muestras se analizaron en lotes de 36 muestras por duplicado, y los procedimientos de control de calidad incluyeron 7 muestras de calibración, 2 muestras de control de calidad, y 1 muestra de blanco en cada lote. Los coeficientes de variación (CV) dentro y entre lotes fueron, respectivamente, 6,86 % y 15,54 % para CA125, 1,45 % y 9,32 % para CEA, 6,55 % y 17,26 % para Pro-SFTPb, y 5,56 % y 28,71 % para CYFRA 21-1. Todos los casos de cáncer de pulmón y sus controles individualmente coincidentes se analizaron juntos dentro de los mismos lotes en orden aleatorio. El personal del laboratorio estaba ciego para el estado de control de casos de las muestras de sangre.

Ejemplo 4: Modelo de regresión.

Cada biomarcador evaluado se transformó logarítmicamente inicialmente y se estandarizó dentro de cada muestra de estudio por separado. Los biomarcadores CA125, CEA, Pro-SFTPb y CYFRA se combinaron en una puntuación de riesgo basada en biomarcadores (puntuación de biomarcadores) ajustando modelos de regresión logística. CYFRA en CARET2 carecía de datos para el 47,53 % de las muestras. Con el fin de usar tantas observaciones como fuera posible para la construcción del modelo, se empleó un método de dos etapas: en la primera etapa, se seleccionó un panel de biomarcadores óptimo basándose en el Criterio de Información de Akaike usando todos los biomarcadores aparte de CYFRA usando regresión logística; en la segunda etapa, la puntuación combinada de la primera etapa se fijó y se combinó con CYFRA para obtener el modelo final. Tal método de dos etapas permite que las observaciones también contribuyan al entrenamiento del modelo. Toda la inferencia se basa en el análisis de remuestreo que emplea 1.000 muestras con reemplazo por separado para la sana y la enferma.

Con el fin de evaluar el grado en que la puntuación del biomarcador podría mejorar en un modelo de predicción de riesgo basado en el historial de exposición al tabaco, se ajustó un modelo de tabaquismo usando datos de EPIC y NSHDS que no se usaron en el estudio de validación, definido por casos diagnosticados de 2 a 10 años después del reclutamiento del estudio con controles que se aparearon individualmente con los mismos criterios de coincidencia que en el estudio de validación (1.008 casos y 1.873 controles. Con el uso de regresión logística condicional y el acondicionamiento en los conjuntos de casos individuales, el modelo de tabaquismo incluyó el estado de tabaquismo (anterior frente a nunca, actual frente a nunca), el número de cigarrillos por día para los fumadores actuales (continuo [no disponible en los fumadores anteriores]), la duración del tabaquismo (continuo en los fumadores anteriores y actuales), y el tiempo desde el abandono en los fumadores anteriores [continuo].

Se evaluó la puntuación de biomarcador para su potencial para mejorar un modelo de predicción de riesgo basado en variables de tabaquismo solas usando la muestra combinada de EPIC y NSHDS, primero ajustando un modelo (modelo de tabaquismo) que incluye variables de tabaquismo usando regresión logística incondicional. Este modelo incluía edad (continua), sexo (dicotómico), cohorte (dicotómico), estado de tabaquismo (anterior/actual), cigarrillos al día en fumadores actuales (continuo [no disponible en fumadores anteriores]), duración de tabaquismo (continuo) y tiempo desde el abandono en fumadores anteriores [continuo]. Las estimaciones de parámetros para el modelo de tabaquismo se presentan en la Tabla 5.

El grado en que la puntuación del biomarcador y el modelo de tabaquismo podrían discriminar entre casos de cáncer de pulmón incidente y controles se evaluó posteriormente de manera externa y no paramétrica asignando los modelos respectivos a cada participante en el estudio de validación (casos diagnosticados de 0 a 2 años después de extracción de sangre). Además, con el fin de evaluar el potencial de combinar la puntuación del biomarcador con el modelo de tabaquismo, se desarrolló un modelo de predicción de riesgo integrado ajustando un modelo de regresión logística condicional usando la puntuación de riesgo basada en el modelo de tabaquismo y la puntuación del biomarcador como dos covariables separadas en el estudio de validación.

Tabla 5. Especificaciones del modelo de tabaquismo desarrollado en EPIC y NSHDS basadas en casos diagnosticados entre 2 y 10 años desde la extracción de sangre.

Variables incluidas en la puntuación de tabaquismo	Estimaciones beta para la puntuación de tabaquismo	OR	IC 95 %
Anterior frente a nunca	1,417089	4,13	[2,73-6,22]
Actual frente a nunca	2,084509	8,04	[5,45-11,9]
Duración del tabaquismo (años) entre fumadores ocasionales	0,038906	1,04	[1,02-1,06]
Tiempo desde el cese de fumar (años) para fumadores anteriores	- 0,027166	0,97	[0,95-0,99]
Número de cigarrillos fumados por día por los fumadores actuales	0,066884	1,07	[1,05-1,09]

Los coeficientes del logaritmo de las concentraciones de biomarcadores fueron los siguientes: CA-125, 0,4730; CYFRA 21-1, 0,2612; CEA, 0,6531; y Pro-SFTPB, 0,9238; constante, -8,4927.

$$\logit(p) = -8,4927 + 0,4730 \times \log CA125 + 0,2612 \times \log CYFRA211 + 0,6531 \times \log CEA + 0,9238 \times \log ProSFTPB$$

Ejemplo 5: Rendimiento del modelo de regresión.

Cada biomarcador candidato discriminaba entre casos y controles en el estudio de entrenamiento CARET (valor de $p < 0,05$), con estimaciones de AUC que variaban de 0,60 (IC del 95 %: 0,53-0,67, CA125) a 0,70 (IC del 95 %: 0,64-0,76, Pro-SFTPB) (Tabla 6). Basándose en AIC, se excluyó HE4 del modelo, y la puntuación final del biomarcador se definió por CA125, CEA, Pro-SFTPB y CYFRA 21-1 y produjo un AUC global de 0,80 (IC del 95 % 0,72-0,87) en la muestra de entrenamiento (FIG. 1).

Tabla 6. Rendimiento de biomarcadores individuales.

	Rendimiento discriminativo		Sensibilidad	Especificidad	Especificación del modelo	
	AUC	IC de 95 %	al 95 % de especificidad	al 95 % de sensibilidad	Estimación beta	IC del 95 %
CA125	0,597	[0,53-0,67]	0,1250	0,0979	0,4730	[0,0886-0,08583]
CEA	0,688	[0,62-0,75]	0,2708	0,1563	0,6531	[0,1364-1,1698]
Pro-SFTPB	0,704	[0,64-0,76]	0,1667	0,2272	0,9238	[0,3627-1,4849]
CYFRA21-1	0,658	[0,56-0,75]	0,1042	0,1672	0,2612	[-0,1601-0,6825]
HE4	0,645	[0,58-0,71]	0,1262	0,1053	N/A	

La puntuación de biomarcadores discriminaba de manera similar para los dos tipos histológicos más prevalentes de cáncer de pulmón en el estudio de entrenamiento, con un AUC de 0,79 para adenocarcinoma de pulmón (IC del 95 %: 0,67-0,92) y 0,79 en carcinoma escamoso de pulmón (IC del 95 %: 0,62-0,96) (Tabla 7). El AUC fue mayor para casos de cáncer de pulmón diagnosticados en los 6 meses desde la extracción de sangre (AUC: 0,86, IC del 95 %: 0,76-0,96) en comparación con casos diagnosticados entre 6 meses y 12 meses después de la extracción de sangre (AUC: 0,77, IC del 95 %: 0,66-0,88) (Tabla 7).

Tabla 7. Rendimiento de la puntuación de biomarcadores en el estudio de entrenamiento CARET

	AUC	IC del 95 %		valor <i>p</i>	Sensibilidad al 95 % de especificidad	Especificidad al 95 % de sensibilidad
TODOS los individuos	0,798	0,72	0,87	<0,0001	0,2000	0,1875
Hombres	0,783	0,68	0,88	<0,0001	0,2581	0,1486
Mujeres	0,829	0,70	0,96	<0,0001	0,1429	0,2105
Estadio I-II	0,677	0,45	0,90	0,0599	0,2000	0,1538
Estadio III-IV	0,831	0,75	0,91	<0,0001	0,1724	0,5857
Tiempo entre extracción de sangre y diagnóstico						
0-6 meses	0,858	0,76	0,96	<0,0001	0,3810	0,6957
6-12 Meses	0,770	0,66	0,88	<0,0001	0,2083	0,1667
Subtipo histológico:						
ADC*	0,794	0,67	0,92	<0,0001	0,2143	0,4444
Otros NSCLC	0,804	0,68	0,93	<0,0001	0,3333	0,1316
SCC*	0,790	0,62	0,96	0,0004	0,1538	0,1579

ADC: Adenocarcinoma; SCC: Carcinoma de células escamosas; NSCLC, cáncer de pulmón de células no pequeñas

El rendimiento discriminativo de la puntuación de biomarcadores en el estudio de validación produjo un AUC de 0,89 (IC del 95 %: 0,84-0,93). el modelo de tabaquismo a un AUC de 0,79 (IC del 95 %: 0,73-0,85, con un valor *P* para la diferencia entre las dos estimaciones 0,002, FIG. 1, Tabla 6. El AUC para la puntuación de biomarcadores fue consistentemente mayor que para el modelo de tabaquismo a través de estratos relevantes, variando entre 0,83 para fumadores anteriores y 0,93 para fumadores ligeros (sujetos con menos de 30 paquetes al año de historial de tabaquismo, Tabla 8). La combinación de la puntuación de biomarcadores con el modelo de tabaquismo en el modelo de predicción de riesgo integrado no mejoró notablemente el AUC en comparación con la puntuación de biomarcadores solo (AUC para el modelo de predicción de riesgo integrado: 0,90, IC del 95 %: 0,86-0,94, FIG. 1).

Los rendimientos discriminativos del modelo de tabaquismo, la puntuación de biomarcadores y el modelo de predicción de riesgo integrado se evaluaron usando análisis ROC en el estudio de validación. Con el fin de estimar la fracción de casos futuros de cáncer de pulmón que se habrían identificado usando los diferentes modelos, estimamos la sensibilidad de cada modelo a un nivel de especificidad correspondiente al proporcionado aplicando los criterios de elegibilidad de cribado de USPSTF a cada sujeto en el estudio de validación. A la inversa, también evaluamos la especificidad del modelo al nivel de sensibilidad obtenido usando los criterios USPSTF con el fin de estimar el grado en que se podría evitar el cribado para sujetos que no están destinados a desarrollar cáncer de pulmón.

La aplicación de los criterios de cribado dicotómicos de USPSTF al estudio de validación dio como resultado una sensibilidad de 0,37 (IC del 95 %: 0,23-0,48) y una especificidad de 0,85 (IC del 95 %: 0,72-0,91) (Tabla 8). La puntuación de biomarcadores a la especificidad de los criterios de cribado de USPSTF de 0,85, produjo una sensibilidad de 0,78 (IC del 95 %: 0,65-0,87), mientras que el modelo de tabaquismo produjo una sensibilidad de 0,58 (IC del 95 %: 0,31-0,71). La sensibilidad de la puntuación de biomarcadores varió entre 0,65 y 0,85 a través de los estratos (Tabla 8). En comparación, la sensibilidad del modelo de tabaquismo a la misma especificidad varió entre 0,12 y 0,73 a través de los estratos. También se evaluó el potencial para mejorar la especificidad en comparación con los criterios de USPSTF, con el fin de evitar el cribado de sujetos que es poco probable que desarrollen cáncer de pulmón. A la sensibilidad global de 0,37 basándose en los criterios de USPSTF, la especificidad correspondiente en el estudio de validación usando la puntuación de biomarcadores fue de 0,98 (IC del 95 %: 0,96-1,00), en comparación con 0,90 (IC del 95 %: 0,84-0,94) para el modelo de tabaquismo. La especificidad de la puntuación de biomarcadores varió entre 0,96 y 1,00 a través de los estratos y fue consistente y notablemente mayor que la del modelo de tabaquismo para fumadores actuales (Tablas 3 y 9).

Tabla 8. Rendimiento del panel de 4 marcadores en la cohorte CARET.

	Casos	Controles	Especificidad de los criterios de USPSTF	IC del 95 %	Sensibilidad de los criterios USPSTF	IC del 95 %	Modelo de riesgo	AUC	IC del 95 %	Sensibilidad a la especificidad de USPSTF	IC del 95 %	Especificidad a la sensibilidad de USPSTF	IC del 95 %
Todos	83	158	0.85	[0.72-0.91]	0.37	[0.23-0.48]	Tabaquismo	0.79	[0.73-0.85]	0.58	[0.31-0.71]	0.90	[0.84-0.94]
							Biomarcadores	0.89	[0.84-0.93]	0.78	[0.65-0.87]	0.98	[0.96-1.00]
							Tabaquismo + biomarcadores	0.90	[0.86-0.94]	0.80	[0.70-0.89]	0.97	[0.95-1.00]
Hombres	50	93	0.80	[0.61-0.88]	0.53	[0.33-0.67]	Tabaquismo	0.84	[0.77-0.90]	0.72	[0.56-0.84]	0.94	[0.84-0.98]
							Biomarcadores	0.87	[0.81-0.94]	0.76	[0.64-0.90]	0.98	[0.90-1.00]
							Tabaquismo + biomarcadores	0.89	[0.84-0.95]	0.82	[0.70-0.94]	0.96	[0.90-1.00]
Mujeres	33	65	0.92	[0.8-0.98]	0.13	[0.03-0.26]	Tabaquismo	0.71	[0.61-0.82]	0.12	[0.00-0.30]	0.89	[0.77-0.98]
							Biomarcadores	0.91	[0.84-0.97]	0.70	[0.52-0.88]	1.00	[0.97-1.00]
							Tabaquismo + biomarcadores	0.91	[0.85-0.97]	0.67	[0.48-0.88]	0.98	[0.95-1.00]
Nunca fumadores	9	50	N/A		N/A		Tabaquismo	N/A		N/A		N/A	
							Biomarcadores	0.89	[0.71-1.00]	N/A		N/A	
							Tabaquismo + biomarcadores	N/A		N/A		N/A	
Fumadores anteriores	26	53	0.90	[0.71-0.97]	0.29	[0.1-0.48]	Tabaquismo	0.81	[0.70-0.91]	0.58	[0.23-0.77]	0.96	[0.89-1.00]
							Biomarcadores	0.83	[0.73-0.93]	0.65	[0.31-0.85]	0.96	[0.89-1.00]
							Tabaquismo + biomarcadores	0.87	[0.79-0.99]	0.69	[0.35-0.91]	0.96	[0.91-1.00]

	Casos	Controles	Especificidad de los criterios de USPSTF	IC del 95 %	Sensibilidad de los criterios USPSTF	IC del 95 %	Modelo de riesgo	AUC	IC del 95 %	Sensibilidad a la especificidad de USPSTF	IC del 95 %	Especificidad a la sensibilidad de USPSTF	IC del 95 %
Fumadores actuales							biomarcadores		0,95]		0,88]		1,00]
	48	55	0,67	[0,49-0,81]	0,48	[0,3-0,63]	Tabaquismo	0,69	[0,59-0,79]	0,73	[0,31-0,85]	0,75	[0,60-0,85]
							Biomarcadores	0,88	[0,82-0,95]	0,85	[0,75-0,98]	0,98	[0,93-1,00]
							Tabaquismo + biomarcadores	0,88	[0,81-0,95]	0,88	[0,77-0,98]	0,96	[0,89-1,00]
Casos de NSCLC	70	158	0,85	[0,72-0,91]	0,35	[0,21-0,47]	Tabaquismo	0,78	[0,72-0,84]	0,54	[0,29-0,69]	0,90	[0,84-0,94]
							Biomarcadores	0,88	[0,83-0,93]	0,77	[0,64-0,86]	0,98	[0,96-1,00]
Fumadores empedernidos							Tabaquismo + biomarcadores	0,90	[0,85-0,94]	0,79	[0,69-0,89]	0,98	[0,96-1,00]
	25	19	N/A		N/A		Tabaquismo	0,65	[0,49-0,82]	N/A			
							Biomarcadores	0,84	[0,72-0,96]	N/A			
							Tabaquismo + biomarcadores	0,83	[0,70-0,95]	N/A			
Fumadores ligeros	23	36	N/A		N/A		Tabaquismo	0,69	[0,55-0,83]				
							Biomarcadores	0,93	[0,87-0,99]				
							Tabaquismo + biomarcadores	0,92	[0,85-0,99]				

Tabla 9. Rendimiento discriminativo de tres modelos de predicción de riesgo en el estudio de validación entre fumadores ocasionales

	Casos	Controles	Especificidad de los criterios de USPSTF	IC del 95 %	Sensibilidad de los criterios USPSTF	IC del 95 %	Modelo de riesgo	AUC	IC del 95 %	Sensibilidad a la especificidad de USPSTF	IC del 95 %	Especificidad a la sensibilidad de USPSTF	IC del 95 %
Todos	71	103	0,77	[0,62-0,86]	0,41	[0,26-0,53]	Tabaquismo	0,77	[0,70-0,84]	0,58	[0,42-0,73]	0,88	[0,79-0,96]
							Biomarcadores	0,86	[0,81-0,92]	0,75	[0,63-0,89]	0,97	[0,92-1,00]
							Tabaquismo + biomarcadores	0,88	[0,83-0,93]	0,85	[0,69-0,94]	0,95	[0,90-0,99]
Hombres	44	69	0,73	[0,51-0,84]	0,58	[0,37-0,72]	Tabaquismo	0,83	[0,75-0,90]	0,77	[0,59-0,91]	0,87	[0,74-0,94]
							Biomarcadores	0,87	[0,80-0,93]	0,75	[0,61-0,93]	0,96	[0,81-1,00]
							Tabaquismo + biomarcadores	0,89	[0,83-0,95]	0,89	[0,73-0,98]	0,96	[0,88-1,00]
Mujeres	27	34	0,85	[0,7-0,97]	0,12	[0,02-0,28]	Tabaquismo	0,63	[0,49-0,77]	0,26	[0,07-0,48]	0,94	[0,76-1,00]
							Biomarcadores	0,86	[0,77-0,96]	0,70	[0,48-0,93]	0,97	[0,91-1,00]
							Tabaquismo + biomarcadores	0,87	[0,78-0,96]	0,81	[0,33-0,96]	0,94	[0,85-1,00]
Fumadores anteriores	25	50	0,89	[0,7-0,97]	0,30	[0,11-0,49]	Tabaquismo	0,83	[0,72-0,93]	0,56	[0,32-0,80]	0,96	[0,90-1,00]
							Biomarcadores	0,84	[0,74-0,93]	0,64	[0,40-0,80]	0,98	[0,92-1,00]
							Tabaquismo + biomarcadores	0,89	[0,82-0,97]	0,68	[0,40-0,92]	0,98	[0,90-1,00]
Fumadores actuales	46	53	0,66	[0,49-0,8]	0,46	[0,28-0,62]	Tabaquismo	0,70	[0,59-0,80]	0,61	[0,41-0,80]	0,79	[0,62-0,91]
							Biomarcadores	0,88	[0,81-0,95]	0,89	[0,74-1,00]	0,94	[0,89-1,00]
							Tabaquismo + biomarcadores	0,87	[0,80-0,94]	0,91	[0,76-0,98]	0,94	[0,87-1,00]
Casos de NSCLC	58	103	0,77	[0,61-0,86]	0,39	[0,23-0,53]	Tabaquismo	0,77	[0,69-0,84]	0,59	[0,41-0,74]	0,90	[0,80-0,97]

	Casos	Controles	Especificidad de los criterios de USPSTF	IC del 95 %	Sensibilidad de los criterios USPSTF	IC del 95 %	Modelo de riesgo	AUC	IC del 95 % I	Sensibilidad a la especificidad de USPSTF	IC del 95 %	Especificidad a la sensibilidad de USPSTF	IC del 95 %
							Biomarcadores	0,86	[0,80- 0,92]	0,76	[0,64- 0,88]	0,98	[0,94- 1,00]
							Tabaquismo + biomarcadores	0,88	[0,82- 0,93]	0,81	[0,67- 0,91]	0,95	[0,91- 0,99]
Fumadores empedernidos	23	18	N/A		N/A		Tabaquismo	0,65	[0,48- 0,83]				
							Biomarcadores	0,86	[0,74- 0,97]				
							Tabaquismo + biomarcadores	0,86	[0,75- 0,97]				
Fumadores ligeros	20	35	N/A		N/A		Tabaquismo	0,73	[0,60- 0,86]				
							Biomarcadores	0,91	[0,83- 0,98]				
							Tabaquismo + biomarcadores	0,89	[0,80- 0,97]				

La sensibilidad del panel al 95 % de especificidad fue del 38 % para sujetos a los que se les diagnosticó cáncer de pulmón en los 6 meses desde la extracción de sangre (FIG. 2) y del 21 % para sujetos que desarrollaron cáncer de pulmón más de 6 meses y hasta 12 meses después de la extracción de sangre (FIG. 3) (AUC = 0,858, IC del 95 %: 0,76-0,96; AUC = 0,770, IC del 95 %: 0,66-0,88, respectivamente). El panel ofreció una clasificación comparable para los dos tipos histológicos más frecuentes de cáncer de pulmón: un AUC de 0,794 para adenocarcinoma de pulmón (IC del 95 %: 0,67-0,92; SN=21 %, SP=95 %), un AUC de 0,790 en carcinoma escamoso de pulmón (IC del 95 %: 0,62-0,96; SN=15 %, SP=95 %). El resto de NSCLC se clasificó con un AUC de 0,804 (IC del 95 %: 0,68-0,93) con una sensibilidad del 33 % (SP = 95 %). El rendimiento del panel para las mujeres fue (AUC=0.829; IC del 95 %: 0,70-0,96) y hombres (AUC=0.783; IC del 95 %: 0,69-0,88).

El grado en que la puntuación de biomarcadores podría discriminar entre casos de cáncer de pulmón incidente y controles se evaluó externamente y no paramétricamente combinando muestras de EPIC y NSHDS y asignando a cada participante la puntuación del biomarcador, y evaluando el rendimiento usando análisis de curvas de características operativas del receptor (ROC).

La puntuación de biomarcadores se evaluó para determinar su potencial para mejorar un modelo de predicción de riesgo basado en variables de tabaquismo solas usando la muestra combinada de EPIC y NSHDS, primero ajustando un modelo (modelo de tabaquismo) que incluye variables de tabaquismo usando regresión logística incondicional. Este modelo incluía edad (continua), sexo (dicotómico), cohorte (dicotómico), estado de tabaquismo (anterior/actual), cigarrillos al día en fumadores actuales (continuo [no disponible en fumadores anteriores]), duración de tabaquismo (continuo) y tiempo desde el abandono en fumadores anteriores (continuo). Posteriormente, se determinó un modelo de predicción de riesgo integrado ajustando un nuevo modelo que también incluye la puntuación de biomarcadores con el modelo de tabaquismo. El rendimiento discriminativo del modelo de predicción de riesgo integrado en comparación con el modelo de tabaquismo se evaluó principalmente analizando las muestras de EPIC y NSHDS combinadas.

Con el fin de estimar la fracción de casos futuros de cáncer de pulmón que se habrían identificado usando los diferentes modelos, se estableció la sensibilidad de cada modelo a una especificidad correspondiente a la proporcionada aplicando los criterios de elegibilidad de cribado de USPSTF. A la inversa, la especificidad del modelo también se examinó a la sensibilidad de USPSTF con el fin de estimar el grado en que se podría evitar el cribado en sujetos que no desarrollarán cáncer de pulmón.

En el análisis ROC, el rendimiento global del modelo de tabaquismo (sólo variables de tabaquismo basadas en demografía y en cuestionarios) en casos discriminativos y controles en la muestra de validación de EPIC y NSHDS combinada se estimó a un AUC de 0,72 (IC del 95 %: 0,65-0,79, FIG. 1). El rendimiento discriminativo correspondiente de la puntuación de biomarcadores solo se estimó a un AUC de 0,80 (IC del 95 %: 0,74-0,87, valor p que compara el AUC con el modelo de tabaquismo: 0,07). La combinación de la puntuación de biomarcadores con el modelo de tabaquismo en el modelo de predicción de riesgo integrado dio como resultado un AUC de 0,83 (IC del 95 %: 0,77-0,89, valor de p comparando el AUC con el modelo de tabaquismo: 0,001).

La estimación de AUC proporcionada por el modelo de predicción de riesgo integrado fue similar cuando se estratifica por sexo y estado de tabaquismo (rango: 0,79 a 0,84), pero superior en participantes del estudio que fumaban menos de 30 paquetes-año de cigarrillos (AUC: 0,90, IC del 95 %: 0,83-0,96) en comparación con participantes del estudio que fumaban más de 30 paquetes-año (AUC: 0,67, IC del 95 %: 0,53-0,81, valor p: 0,0053).

La aplicación de los criterios de cribado dicotómicos de USPSTF en la muestra de validación dio como resultado una sensibilidad global de 0,42 (IC del 95 %: 0,26-0,54) y una especificidad de 0,78 (IC del 95 %: 0,64-0,87). Con el fin de evaluar el potencial de mejorar la sensibilidad cuando se hace referencia a casos futuros de cáncer de pulmón para el cribado -en lugar de aplicar los criterios de USPSTF usados actualmente- se evaluó la sensibilidad del modelo de predicción de riesgo integrado a la especificidad de los criterios de USPSTF. Por consiguiente, a una especificidad de 0,78, la sensibilidad para el modelo de predicción de riesgo integrado fue de 0,76 (IC del 95 %: 0,57-0,86) y varió entre 0,61 y 0,83 a través de los estratos, pero fue consistentemente mayor que la del criterio de USPSTF (Tabla 9). En comparación, la sensibilidad global del modelo de tabaquismo a la misma especificidad fue de 0,41 (IC del 95 %: 0,28-0,66) y varió entre 0,21 y 0,57 a través de los estratos. A la inversa, se emprendieron estudios para mejorar la especificidad de los criterios de USPSTF, con el fin de evitar sujetos referentes que no desarrollarán cáncer de pulmón para el cribado. Aquí, a la sensibilidad global de 0,42 - la de los criterios de USPSTF - la especificidad correspondiente usando el modelo de predicción de riesgo integrado fue de 0,94 (IC del 95 %: 0,88-0,98), en contraposición a 0,78 (IC del 95 %: 0,70-0,90) del modelo de tabaquismo. De nuevo, la especificidad del modelo de predicción de riesgo integrado fue consistentemente y notablemente mayor que la del modelo de tabaquismo y varió entre 0,88 y 1,00 a través de los estratos.

Ejemplo 6: Expansión del panel con biomarcadores adicionales.

En un análisis exploratorio, también se analizaron 16 biomarcadores prometedores adicionales contra la cohorte de validación para evaluar su potencial en la predicción del riesgo. Basándose en los datos de estos 16 biomarcadores, se continuó con una mejora adicional del rendimiento discriminativo de la puntuación de biomarcadores desarrollada en CARET. Basándose en la regresión logística penalizada (lasso), se seleccionaron 2 biomarcadores adicionales de los 16 biomarcadores analizados además de la puntuación de biomarcadores, dando como resultado una AUC

nominalmente mayor de 0,83 (IC del 95 %: 0,77-0,90) en general, en comparación con 0,80 (IC del 95 %: 0,74-0,87) de la puntuación de biomarcadores sola (valor p: 0,11). La estimación de AUC corregida por la optimización del remuestreo fue de 0,80, lo que indica que el potencial de los 16 biomarcadores adicionales para mejorar aún más la predicción del riesgo de cáncer de pulmón es limitado.

5 **Ejemplo 7: Especificidad y sensibilidad en el rango de puntuaciones diagnósticas del modelo de regresión.**

Los expertos en la técnica apreciarán que diferentes métodos o ensayos de detección, cuantificación y análisis de biomarcadores, que pueden incluir usar diferentes reactivos, producirán diferentes resultados que pueden requerir la modificación del modelo de regresión. En particular, diferentes ensayos pueden producir resultados expresados, por ejemplo, en diferentes unidades. Además, las reacciones duplicadas en ensayos duplicados de las mismas muestras también pueden producir resultados brutos diferentes. Sin embargo, es la detección, cuantificación y análisis combinados de al menos los cuatro biomarcadores CEA, CA125, CYFRA21-1 y Pro-SFTPb lo que, cuando se incorporan en un modelo de regresión como se divulga en la presente memoria, producen un diagnóstico definitivo del cáncer de pulmón.

Un rango en los resultados informados para cada ensayo particular usado para detectar, cuantificar y analizar los cuatro biomarcadores tendrá un rango en la puntuación predictiva de cáncer de pulmón resultante que depende, en parte, del grado de sensibilidad o especificidad (Tabla 10; donde el corte preferido basado en el índice de Youden es -1,073 con una especificidad de 0,714 y sensibilidad de 0,822). El modelo de regresión usado para generar la puntuación predictiva de cáncer de pulmón puede depender de los ensayos específicos usados para ensayar los marcadores. Como entienden los expertos en la técnica, diferentes ensayos pueden dirigirse a diferentes epítomos de los cuatro biomarcadores o tener diferentes afinidades y sensibilidades. Como tal, el algoritmo del modelo de regresión utilizado para generar la puntuación predictiva de cáncer de pulmón puede modificarse para tener en cuenta estas variaciones de los ensayos.

Ejemplo 8: Ensayo de muestras y diagnóstico de pacientes con cáncer de pulmón.

En un ejemplo, a un paciente que se está cribando para cáncer de pulmón - basado en el panel de cuatro biomarcadores divulgado en la presente memoria - se le ha extraído una muestra de sangre (u otro fluido o biopsia de tejido) y se ha ensayado por ELISA (u otro ensayo) para cuantificar los niveles de CEA, CA125, CYFRA21-1 y Pro-SFTPb en el paciente. Los valores normalizados para al menos estos biomarcadores que tienen en cuenta el ensayo específico usado podrían ser, por ejemplo, CEA = 3,2 ng/ml; CA125 = 3,5 U/ml; CYFRA21-1 = 2,1 ng/ml; y Pro-SFTPb = 6 ng/ml. Los datos brutos del ensayo se transforman luego en log2, calculando la media y la desviación estándar para las muestras sanas en cada cohorte. Los datos se estandarizan entonces de modo que las muestras sanas tengan una media de 0 y una desviación estándar de 1: donde (Lectura_j - mediana_{sano}) / (estdsano), donde j es la j-ésima muestra.

Cuando se analiza usando el siguiente modelo de regresión:

$$\logit(p) = -8,4927 + 0,4730 \times \log CA125 + 0,2612 \times \log CYFRA211 + 0,6531 \times \log CEA$$

el paciente anterior tendría una puntuación combinada de 1,344. En vista del punto de corte preferido para la consideración tanto de la especificidad como de la sensibilidad (Tabla 10), un paciente con dicha puntuación combinada tendría cáncer de pulmón casi con certeza y, en consecuencia, será dirigido a ensayos de seguimiento y tratamiento para el cáncer de pulmón usando otras modalidades analizadas en la presente memoria y conocidas por los expertos en la técnica. Usando el modelo de regresión descrito en la presente memoria, cuanto más positiva sea la puntuación predictiva combinada de cáncer de pulmón, más certeza habrá de que el paciente tiene cáncer de pulmón. Por el contrario, cuanto más negativa sea la puntuación predictiva combinada de cáncer de pulmón, más certeza habrá de que el paciente no tiene cáncer de pulmón.

Por el contrario, en otro ejemplo, los valores normalizados para los biomarcadores CEA, CA125, CYFRA21-1 y Pro-SFTPb que tienen en cuenta el ensayo específico usado podrían ser, por ejemplo, CEA = 0,1 ng/ml; CA125 = 2,0 U/ml; CYFRA21-1 = -1,2 ng/ml; y Pro-SFTPb = 4,0 ng/ml. Cuando se analiza usando el mismo modelo de regresión como anteriormente, dicho paciente tendría una puntuación combinada de -4,3347. En vista del punto de corte preferido para la consideración tanto de la especificidad como de la sensibilidad (Tabla 10), un paciente con dicha puntuación combinada no tendría, casi con certeza, cáncer de pulmón y, por lo tanto, necesitaría o no seguirse para ensayos adicionales basados en la fuerza de cualquier otra afección clínica.

Tabla 10. Sensibilidad y especificidad en diferentes puntos de corte de las puntuaciones basadas en panel de biomarcadores (CEA, CA125, CYFRA21-1 y Pro-SFTPB) en el conjunto de validación de combinación.

Corte	Especificidad	Sensibilidad	Youden	0,083	0,911	0,356	0,266
2,252	1,000	0,000	0,000	0,053	0,911	0,378	0,288
2,252	1,000	0,022	0,022	0,037	0,902	0,378	0,280
1,978	1,000	0,044	0,044	-0,019	0,893	0,378	0,271
1,942	1,000	0,067	0,067	-0,027	0,884	0,378	0,262
1,925	1,000	0,089	0,089	-0,030	0,875	0,378	0,253
1,693	1,000	0,111	0,111	-0,043	0,866	0,378	0,244
1,560	1,000	0,133	0,133	-0,052	0,866	0,400	0,266
1,548	1,000	0,156	0,156	-0,055	0,866	0,422	0,288
1,120	1,000	0,178	0,178	-0,098	0,866	0,444	0,311
0,786	0,991	0,178	0,169	-0,127	0,866	0,467	0,333
0,786	0,991	0,200	0,191	-0,169	0,857	0,467	0,324
0,779	0,982	0,200	0,182	-0,174	0,857	0,489	0,346
0,579	0,973	0,200	0,173	-0,197	0,857	0,511	0,368
0,546	0,964	0,200	0,164	-0,231	0,848	0,511	0,359
0,526	0,955	0,200	0,155	-0,233	0,839	0,511	0,350
0,507	0,946	0,200	0,146	-0,257	0,839	0,533	0,373
0,464	0,946	0,222	0,169	-0,302	0,839	0,556	0,395
0,426	0,946	0,244	0,191	-0,348	0,830	0,556	0,386
0,407	0,938	0,244	0,182	-0,404	0,830	0,578	0,408
0,400	0,938	0,267	0,204	-0,466	0,830	0,600	0,430
0,357	0,938	0,289	0,226	-0,491	0,821	0,600	0,421
0,249	0,938	0,311	0,249	-0,515	0,813	0,600	0,413
0,230	0,929	0,311	0,240	-0,521	0,804	0,600	0,404
0,191	0,920	0,311	0,231	-0,552	0,795	0,600	0,395
0,138	0,920	0,333	0,253	-0,586	0,795	0,622	0,417
0,120	0,920	0,356	0,275	-0,629	0,786	0,622	0,408
-0,692	0,777	0,622	0,399	-1,339	0,607	0,844	0,452
-0,700	0,777	0,644	0,421	-1,378	0,598	0,844	0,443
-0,707	0,777	0,667	0,443	-1,384	0,598	0,867	0,465
-0,708	0,777	0,689	0,466	-1,390	0,589	0,867	0,456
-0,735	0,777	0,711	0,488	-1,399	0,580	0,867	0,447
-0,773	0,768	0,711	0,479	-1,401	0,571	0,867	0,438
-0,794	0,768	0,733	0,501	-1,407	0,563	0,867	0,429
-0,796	0,768	0,756	0,523	-1,438	0,563	0,889	0,451
-0,822	0,759	0,756	0,514	-1,447	0,554	0,889	0,442
-0,930	0,750	0,756	0,506	-1,449	0,545	0,889	0,434
-0,949	0,750	0,778	0,528	-1,459	0,536	0,889	0,425
-0,994	0,741	0,778	0,519	-1,464	0,527	0,889	0,416
-1,033	0,732	0,778	0,510	-1,470	0,518	0,889	0,407
-1,055	0,723	0,778	0,501	-1,551	0,518	0,911	0,429
-1,065	0,714	0,778	0,492	-1,556	0,509	0,911	0,420
-1,066	0,714	0,800	0,514	-1,612	0,500	0,911	0,411

ES 2 995 959 T3

-1,073	0,714	0,822	0,537	-1,614	0,500	0,933	0,433
-1,084	0,705	0,822	0,528	-1,626	0,491	0,933	0,424
-1,093	0,696	0,822	0,519	-1,627	0,482	0,933	0,415
-1,101	0,688	0,822	0,510	-1,631	0,473	0,933	0,407
-1,121	0,679	0,822	0,501	-1,672	0,464	0,933	0,398
-1,145	0,679	0,844	0,523	-1,694	0,455	0,933	0,389
-1,157	0,670	0,844	0,514	-1,769	0,446	0,933	0,380
-1,168	0,661	0,844	0,505	-1,772	0,438	0,933	0,371
-1,168	0,652	0,844	0,496	-1,816	0,429	0,933	0,362
-1,186	0,643	0,844	0,487	-1,826	0,420	0,933	0,353
-1,186	0,634	0,844	0,478	-1,846	0,411	0,933	0,344
-1,194	0,625	0,844	0,469	-1,864	0,402	0,933	0,335
-1,306	0,616	0,844	0,461	-1,872	0,393	0,933	0,326
-1,882	0,384	0,933	0,317	-2,966	0,143	0,978	0,121
-1,895	0,375	0,933	0,308	-2,976	0,134	0,978	0,112
-1,921	0,366	0,933	0,299	-3,031	0,134	1,000	0,134
-1,960	0,357	0,933	0,290	-3,039	0,125	1,000	0,125
-2,021	0,348	0,933	0,282	-3,051	0,116	1,000	0,116
-2,143	0,339	0,933	0,273	-3,083	0,107	1,000	0,107
-2,180	0,330	0,933	0,264	-3,231	0,098	1,000	0,098
-2,189	0,321	0,933	0,255	-3,262	0,089	1,000	0,089
-2,201	0,313	0,933	0,246	-3,414	0,080	1,000	0,080
-2,211	0,304	0,933	0,237	-3,521	0,071	1,000	0,071
-2,212	0,295	0,933	0,228	-3,645	0,063	1,000	0,063
-2,259	0,286	0,933	0,219	-3,705	0,054	1,000	0,054
-2,306	0,277	0,933	0,210	-3,774	0,045	1,000	0,045
-2,389	0,268	0,933	0,201	-3,882	0,036	1,000	0,036
-2,454	0,259	0,933	0,192	-3,936	0,027	1,000	0,027
-2,479	0,250	0,933	0,183	-4,225	0,018	1,000	0,018
-2,485	0,241	0,933	0,174	-4,356	0,009	1,000	0,009
-2,502	0,232	0,933	0,165	-4,595	0,000	1,000	0,000
-2,511	0,223	0,933	0,157				
-2,556	0,214	0,933	0,148				
-2,569	0,205	0,933	0,139				
-2,588	0,196	0,933	0,130				
-2,589	0,188	0,933	0,121				
-2,619	0,188	0,956	0,143				
-2,755	0,179	0,956	0,134				
-2,806	0,170	0,956	0,125				
-2,845	0,161	0,956	0,116				
-2,870	0,152	0,956	0,107				
-2,873	0,152	0,978	0,130				

Ejemplo 9: Efecto aditivo de marcadores de metabolitos y proteínas.

Se realizó una investigación para evaluar la contribución de añadir un metabolito tal como diacetilespermina (DAS), que se identificó previamente como un marcador para cáncer de pulmón de células no pequeñas, al panel de cáncer de pulmón de cuatro marcadores. Los datos indicaron un rendimiento aditivo, particularmente en casos que no son adenocarcinomas (FIG. 4).

La combinación de marcadores de metabolitos y proteínas en muestras de sangre antes del diagnóstico de un grupo de 100 casos de cáncer de pulmón y el doble de controles mostró un rendimiento mejorado en comparación con el panel de cuatro marcadores solo, con un rendimiento particularmente mejorado para sujetos que posteriormente se diagnosticaron con cáncer de pulmón escamoso o con otro diferente de escamoso o adenocarcinoma. Por lo tanto, la adición de un marcador de metabolito tal como DAS a un panel de marcadores de proteínas permite una detección temprana mejorada para el cáncer de pulmón en comparación con los marcadores de proteínas solos.

Otras realizaciones

La descripción detallada expuesta anteriormente se proporciona para ayudar a los expertos en la técnica a poner en práctica la presente divulgación. Sin embargo, la divulgación descrita y reivindicada en la presente memoria no está limitada en su alcance por las realizaciones específicas divulgadas en la presente memoria porque estas realizaciones están destinadas a ilustrar varios aspectos de la divulgación. Se pretende que cualquier realización equivalente esté dentro del alcance de esta divulgación. De hecho, diversas modificaciones de la divulgación, además de las mostradas y descritas en la presente memoria, resultarán evidentes para los expertos en la técnica a partir de la descripción anterior, que no se apartan del espíritu o alcance del presente descubrimiento inventivo. También se pretende que tales modificaciones se encuentren dentro del alcance de las reivindicaciones adjuntas.

REIVINDICACIONES

1. Un método para determinar el riesgo de un sujeto de albergar cáncer de pulmón, que comprende:
 - obtener una muestra biológica del sujeto;
 - medir el nivel de CEA en la muestra biológica;
 - 5 medir el nivel de CA125 en la muestra biológica;
 - medir el nivel de CYFRA21-1 en la muestra biológica; y
 - medir el nivel de Pro-SFTPb en la muestra biológica;
 - en donde la cantidad de CEA, CA125, CYFRA21-1 y Pro-SFTPb clasifica al sujeto como que está en riesgo de albergar cáncer de pulmón o no en riesgo de albergar cáncer de pulmón.
- 10 2. El método de la reivindicación 1, en donde las cantidades de CEA, CA125, CYFRA21-1 y pro-SFTPb, o moléculas informadoras unidas a los mismos, están elevadas en el sujeto con respecto a un sujeto sano.
3. El método de la reivindicación 1, que comprende, además:
 - medir el nivel de diacetilpermina (DAS) en la muestra biológica;
 - en donde la cantidad de diacetilpermina (DAS) clasifica al paciente como que está en riesgo de albergar cáncer de pulmón o no en riesgo de albergar cáncer de pulmón.
- 15 4. El método de cualquiera de las reivindicaciones 1 y 3, en donde la muestra comprende una muestra biológica seleccionada de sangre, plasma y suero.
5. El método de la reivindicación 4, en donde se determina que el sujeto tiene cáncer de pulmón en base a las concentraciones medidas de CEA, CA125, CYFRA21-1, y pro-SFTPb, y opcionalmente DAS.
- 20 6. El método de la reivindicación 5, en donde las concentraciones medidas se usan para calcular una puntuación de biomarcador basándose en los valores de sensibilidad y especificidad en un punto de corte expuesto en la Tabla 10.
7. El método de la reivindicación 6, que comprende además la etapa de: comparar las concentraciones medidas de cada biomarcador en la muestra biológica con la predicción de un modelo estadístico.
- 25 8. El método de la reivindicación 7, en donde las etapas enumeradas comprenden ensayar un panel seleccionado del grupo que consiste en:
 - a. el panel que consiste en CEA, CA125, CYFRA21-1 y Pro-SFTPb; o
 - b. el panel que consiste en CEA, CA125, CYFRA21-1, Pro-SFTPb y diacetilpermina (DAS).
9. El método de la reivindicación 1, en donde la determinación de los niveles de CEA, CA125, CYFRA21-1 y pro-SFTPb se realiza sustancialmente al mismo tiempo, o de una manera escalonada.
- 30 10. El método de la reivindicación 1, en donde la clasificación del sujeto como que tiene cáncer de pulmón tiene una sensibilidad de 0,76 y 0,42 a una especificidad del 78 % y 94 %, respectivamente.
11. El método de la reivindicación 1, en donde los niveles de antígeno CEA, antígeno CA125, antígeno CYFRA21-1, y antígeno pro-SFTPb son elevados en comparación con los niveles de antígeno CEA, antígeno CA125, antígeno CYFRA21-1, y antígeno pro-SFTPb en un sujeto o grupo de referencia que tiene adenocarcinoma o cáncer de células escamosas.
- 35 12. El método de la reivindicación 1, que comprende además comparar la cantidad de CEA, CA125, CYFRA21-1 y pro-SFTPb con un valor de corte como se ejemplifica en la Tabla 10.
13. El método de la reivindicación 3, en donde el valor de corte comprende un AUC (IC del 95 %) de al menos 0,83 o al menos 0,80.
- 40 14. El método de la reivindicación 1, en donde el cáncer de pulmón se diagnostica en o antes del estadio resecable límite o en el estadio resecable.
15. El método de la reivindicación 3, que comprende ensayar al menos una molécula receptora que se une selectivamente a un biomarcador o antígeno seleccionado del grupo que consiste en CEA, CA125, CYFRA21-1, pro-SFTPb y diacetilpermina (DAS); o
- 45 en donde la cantidad de CEA, CA125, CYFRA21-1, pro-SFTPb y diacetilpermina (DAS) se detecta con el uso

de una partícula sólida; o

en donde al menos una de las moléculas informadoras está unida a una enzima; o

en donde al menos una de las moléculas informadoras proporciona una señal detectable; o

5 en donde el método comprende la inclusión de información del historial del sujeto en la asignación de tener cáncer de pulmón o no tener cáncer de pulmón.

10 16. El método de la reivindicación 15, en donde la señal detectable es detectable por un método seleccionado de espectroscopia UV-visible, espectrometría de masas, espectroscopia de resonancia magnética nuclear (RMN), espectroscopia de RMN de protones, espectrometría de resonancia magnética nuclear (RMN), cromatografía de gases-espectrometría de masas (GC-MS), cromatografía líquida-espectrometría de masas (LC-MS), espectroscopia de correlación (COSY), espectroscopia de efecto nuclear de Overhauser (NOESY), espectroscopia de efecto nuclear de Overhauser de marco rotatorio (ROESY), LC-TOF-MS, LC-MS/MS y electroforesis capilar-espectrometría de masas.

FIG. 1

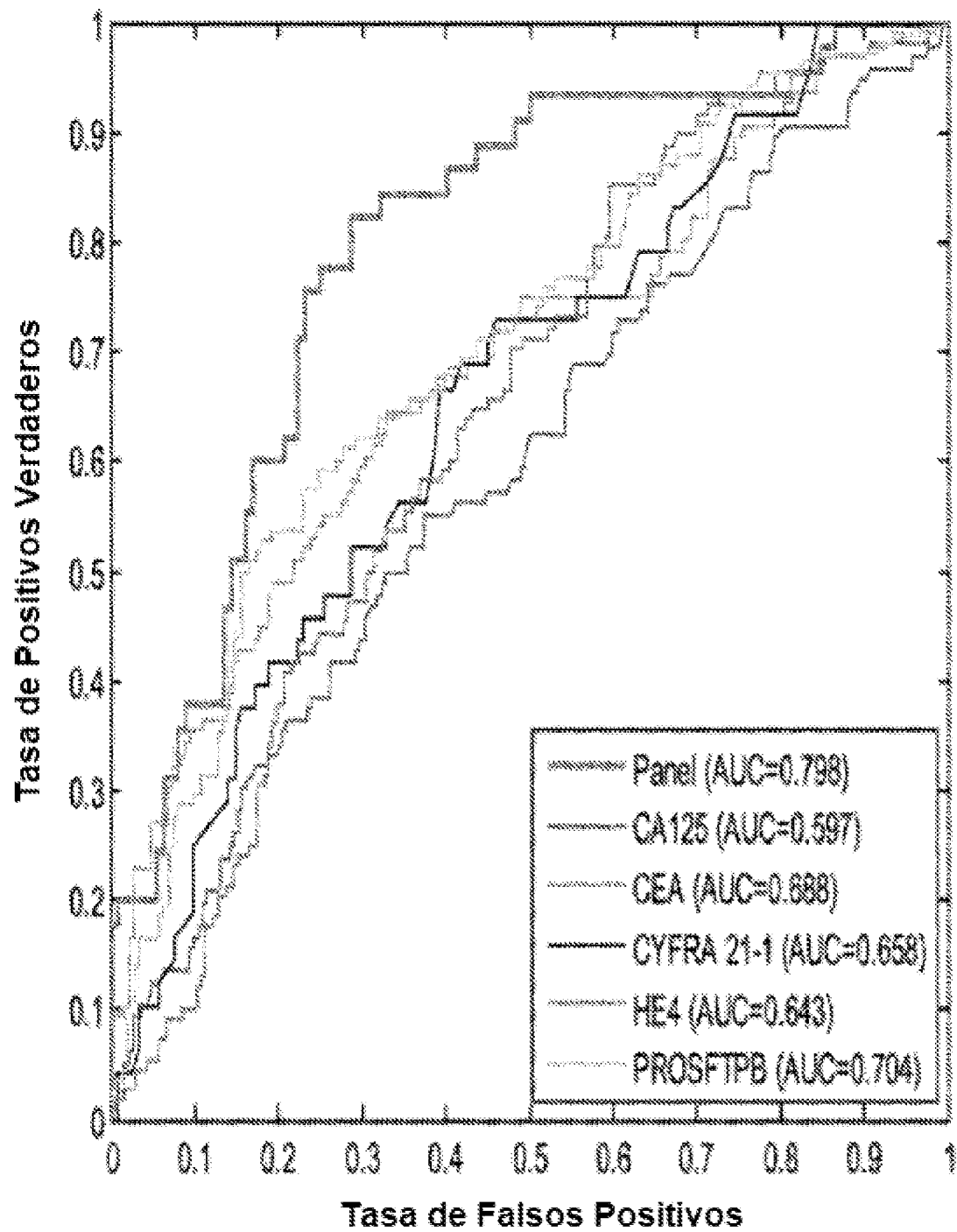


FIG. 2

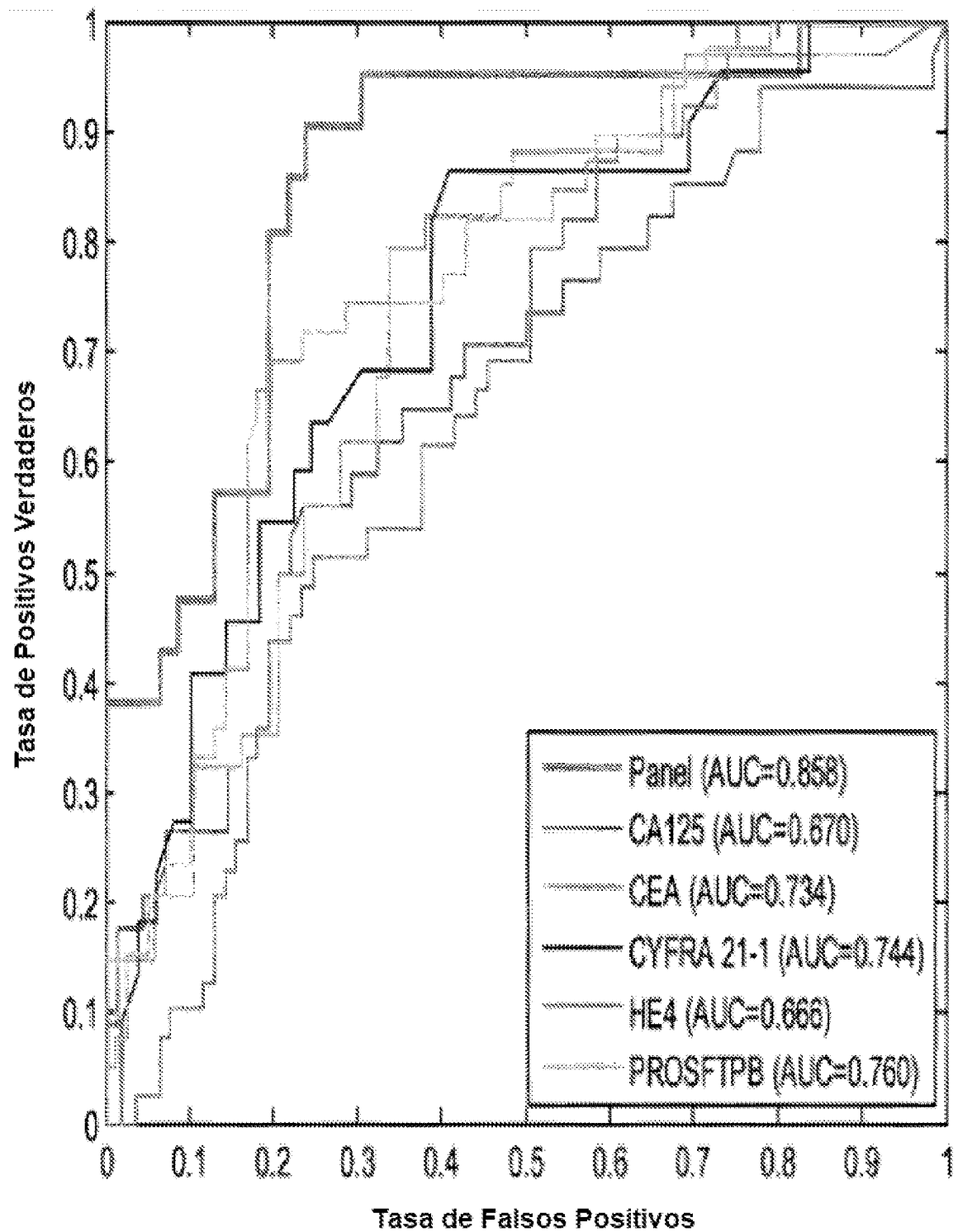


FIG. 3

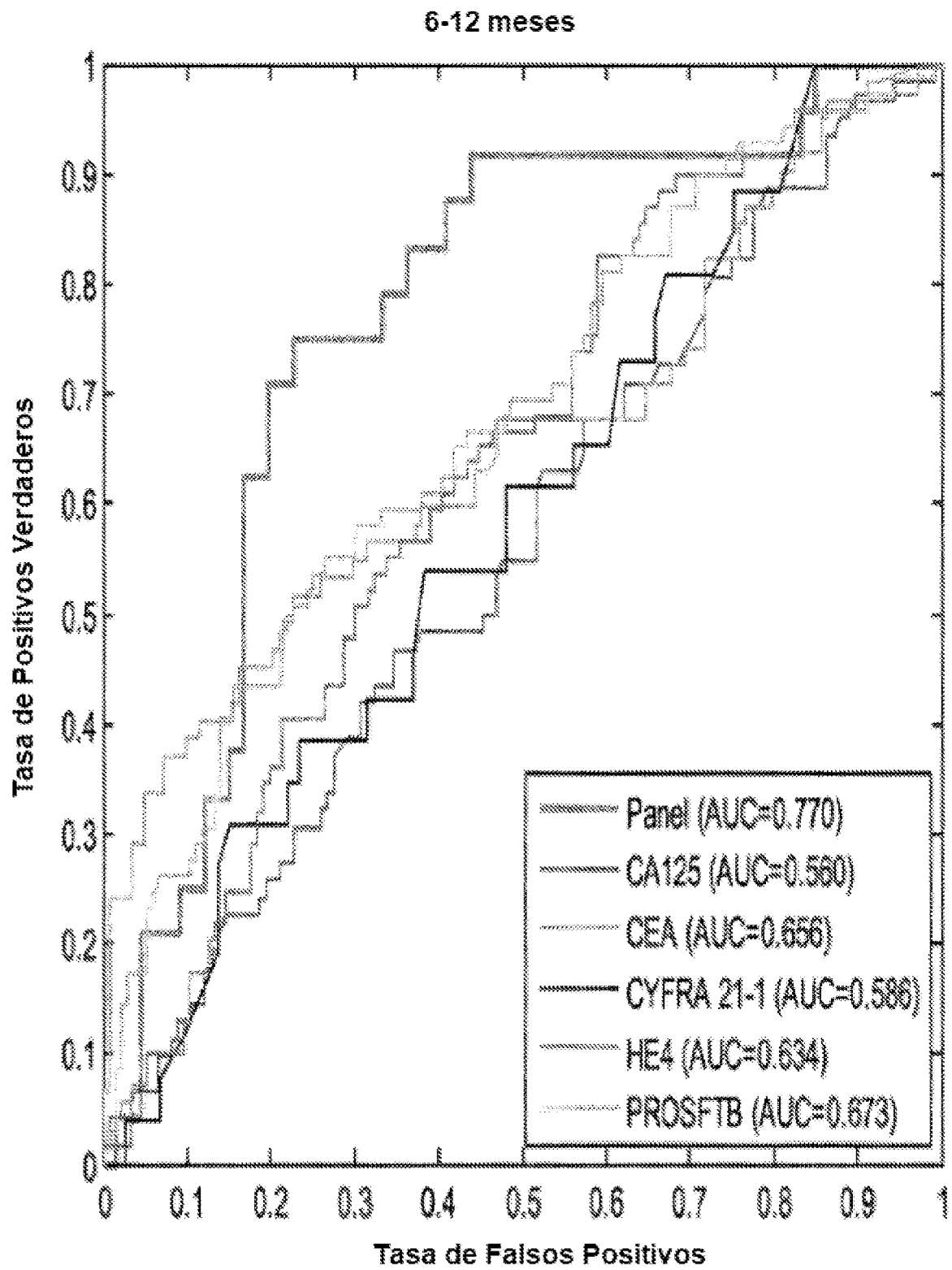


FIG. 4

