

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 694 409**

51 Int. Cl.:

C12Q 1/6886 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.11.2013 PCT/EP2013/074270**

87 Fecha y número de publicación internacional: **30.05.2014 WO14079874**

96 Fecha de presentación y número de la solicitud europea: **20.11.2013 E 13799501 (5)**

97 Fecha y número de publicación de la concesión europea: **08.08.2018 EP 2922970**

54 Título: **Método de pronóstico para individuos con cáncer de próstata**

30 Prioridad:

20.11.2012 SE 1251310
16.05.2013 SE 1350602

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
20.12.2018

73 Titular/es:

PHADIA AB (100.0%)
Box 6460
751 37 Uppsala, SE

72 Inventor/es:

GRÖNBERG, HENRIK y
EKLUND, MARTIN

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 694 409 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

Método de pronóstico para individuos con cáncer de próstata

Campo de la invención

5 La presente invención se refiere, en general, a la detección y la identificación de diversas formas de marcadores genéticos, y diversas formas de proteínas que tienen utilidad potencial como marcadores de pronóstico. En particular, la presente invención se refiere al uso simultáneo de múltiples marcadores de pronóstico para valorar mejor si un individuo que padece cáncer de próstata necesitará tratamiento para el cáncer de próstata en el futuro.

Antecedentes de la invención

10 La medición del antígeno específico de próstata ("prostate specific antigen", PSA) en suero se emplea con profusión para la selección y la detección temprana del cáncer de próstata (PCa). Tal como se analiza en el informe público "Polygenic Risk Score Improves Prostate Cancer Risk Prediction: Results from the Stockholm-1 Cohort Study", de Markus Aly y coautores publicado en EUROPEAN UROLOGY, 60 (2011), 21-28, el PSA en suero que puede medirse por medio de los inmunoensayos clínicos actuales está presente principalmente como la forma "no complejada" libre (PSA libre), o como un complejo con la a-lantiquimotripsina (ACT). Se ha demostrado que la
15 proporción de PSA libre a total en el suero mejora significativamente la detección del PCa. Otros factores, tales como la edad y una historia familiar documentada, también pueden mejorar aún más la detección del PCa. La medición de marcadores genéticos relacionados con el PCa, en particular de polimorfismos de un único nucleótido ("single nucleotide polymorphisms", SNP), es una modalidad emergente para la selección y la detección temprana del cáncer de próstata. El análisis de múltiples SNP relacionados con el PCa, en combinación con biomarcadores
20 tales como el PSA y con la información general sobre el paciente, puede mejorar la evaluación del riesgo a través de una combinación de varios SNP en una puntuación genética.

En el pasado se han descrito intentos de combinar la información procedente de múltiples fuentes en un modelo algorítmico para la predicción de un criterio de valoración diferente, el riesgo de PCa, comparado con la presente invención. En el informe público "Blood Biomarker Levels to Aid Discovery of Cancer-Related Single-Nucleotide Polymorphisms: Kallikreins and Prostate Cancer", de Robert Klein y coautores, publicado en Cancer Prey Res.
25 (2010), 3(5):611-619, los autores analizan el modo en que biomarcadores de la sangre pueden ayudar al descubrimiento de nuevos SNP, pero también sugieren la posible utilidad de incorporar los niveles de genotipo y de marcador en modelos predictivos para la valoración del riesgo de PCa. Además, este informe proporciona pruebas de que la combinación no aditiva de marcadores genéticos y biomarcadores de forma concertada puede tener un valor predictivo en la valoración del riesgo de PCa. Posteriormente, Xu y coinventores han descrito un método para evaluar el riesgo de un sujeto que padece PCa en la solicitud de patente WO2012/031207A2. Esta descripción describe un método para predecir si un individuo está en riesgo de padecer cáncer de próstata mediante el uso de la información genética en 33 SNP definidos, que puede usarse de modo implícito para predecir si el individuo
30 ensayado es adecuado para someterse a una terapia quimiopreventiva. La terapia quimiopreventiva es una medicación proactiva suministrada antes del diagnóstico del cáncer con el fin de reducir la probabilidad de aparición de un cáncer.

Aunque el PSA se emplea predominantemente para el diagnóstico del PCa, también se ha descrito como marcador de pronóstico para individuos diagnosticados con PCa. Un posible método para valorar la prognosis del PCa en un individuo es seguir el avance del valor de PSA, según describen Collette y coautores en el informe público "Prostate specific antigen: a prognostic marker of survival in good prognosis metastatic prostate cancer?", publicado en Eur Urol., agosto de 2003, 44(2):182-9; análisis 189.
40

Existen otros marcadores adecuados para valorar la prognosis de un diagnóstico de PCa, según describen E.P. Gelmann y S.M. Henshall en el informe público "Clinically Relevant Prognostic Markers for Prostate Cancer: The Search Goes On", publicado en Ann Intern Med., 5 de mayo de 2009, 150(9):647-649. En este informe, se analiza el grado histológico (puntuación de Gleason), la expresión de P53, la expresión de BCL2 y la densidad de microvasos como marcadores de pronóstico potenciales, aunque todos presentan importantes inconvenientes para dicho objetivo.
45

La práctica clínica actual (en Suecia) consiste en emplear la puntuación de Gleason como elemento de juicio principal para decidir si se aplica un tratamiento activo (cirugía o terapia de radiación) para un cáncer de próstata que está confinado a la glándula prostática. Otros factores, tales como la edad, enfermedades no relacionadas, extensión calculada del tumor y la opinión del paciente también son importantes para tomar esta decisión. Desde un punto de vista práctico, la inmensa mayoría de los pacientes con tumores de Gleason 8+ son tratados de una manera activa. Para pacientes con tumores de Gleason 6, una fracción más pequeña se trata de manera activa, pero a la mayoría se les somete a un control activo. Se sabe que, puesto que el paciente influye en este proceso de decisión, esta tiene una naturaleza subjetiva. Un método de pronóstico en el campo de la decisión de tratar o no a un paciente de una manera activa sería muy beneficioso, si se proporciona apoyo para la decisión en casos que están en el límite, es decir, para pacientes con tumores de Gleason 6-7.
50
55

Por tanto, la valoración de la prognosis es una tarea difícil para la cual una mejora en la técnica actual conduciría a un gran beneficio para la sociedad. De particular importancia resulta valorar si un individuo diagnosticado con PCa requerirá una terapia avanzada (cirugía o radiación) o si la enfermedad se controlará de modo activo. La terapia avanzada tiene una serie de efectos secundarios graves, que incluyen impotencia (predominantemente en la cirugía), incontinencia y problemas gastrointestinales (estos dos últimos predominantemente en la terapia de radiación). Sin embargo, esta invención proporciona modelos predictivos para la prognosis del PCa a través del análisis de biomarcadores y del perfil genético del individuo diagnosticado con PCa.

Sumario de la invención

La presente invención se basa en el descubrimiento de que la combinación de marcadores de pronóstico de origen diferente puede mejorar la capacidad para determinar si un individuo diagnosticado con PCa requerirá de terapia activa o avanzada. Esto puede generar un gran beneficio para la sociedad, porque los cánceres agresivos que se identifican en un estadio temprano pueden tratarse con más facilidad.

En la solicitud de patente citada anteriormente WO2012/031207A2, no se describe el modo en que el método se aplica en casos en que están disponibles datos de SNP solo para un subconjunto de los 33 SNP, en particular cuando faltan los datos de más del 5% o más del 10% o más del 20% de los SNP. Esto significa que el método en la solicitud de patente WO2012/031207A2 puede requerir que el individuo ensayado suministre una segunda muestra para volver a ensayar la información genética que no pudo obtenerse en el primer ensayo, si el primer ensayo condujo a un resultado parcial. Tampoco se indica el modo en que puede utilizarse el método para predecir la elección de una terapia después del diagnóstico.

Un aspecto de la presente invención proporciona un método basado en una combinación de datos diseñada de modo redundante para valorar si un individuo diagnosticado con cáncer de próstata requerirá una terapia activa, que comprende las etapas de:

1. En al menos una muestra biológica procedente de dicho individuo, analizar una categoría de SNP relacionados con el PCa (SNPpc) midiendo la presencia o la ausencia de uno o dos alelos de riesgo en cada uno de una pluralidad de SNPpc;

2. Combinar los datos relativos a dicha categoría de SNPpc para formar un valor compuesto de SNPpc, en el que el método permite desestimar los datos de un subconjunto de al menos 10% de los SNPpc de la categoría de SNPpc cuando se forma el valor compuesto de SNPpc, en el que el valor compuesto de SNPpc se basa en una oportunidad relativa ("odds ratio") predeterminada para cada SNP individual incluido en la categoría de SNPpc, y en el que el valor compuesto de SNPpc se forma a partir de los datos de al menos 90 SNPpc;

3. Correlacionar dicho valor compuesto de SNPpc con la probabilidad de que un individuo requiera de terapia activa comparando el valor compuesto de SNPpc con un valor de corte predeterminado establecido con muestras control, en el que se sabe si los individuos de los cuales se obtienen las muestras control requirieron una terapia activa o no requirieron una terapia activa.

Según un aspecto de la invención, se proporcionan una o más etapas del método, generalmente las etapas 2 y 3, por medio de un producto de programa informático cuando se ejecuta en un ordenador que comprende un procesador y una memoria.

En una realización, la etapa 2 del método descrito anteriormente se realiza con un ordenador programado para formar o calcular un valor compuesto de SNPpc a partir de los datos de la etapa 1 y/o la etapa 3 para correlacionar el valor compuesto de SNPpc con la probabilidad de que un individuo requiera de terapia activa comparando el valor compuesto de SNPpc con un valor de corte predeterminado establecido con muestras control, en el que se sabe si los individuos de los cuales se obtienen las muestras control requirieron una terapia activa o no requirieron una terapia activa. Además, la presente invención se refiere a un medio de almacenamiento legible por ordenador tangible no transitorio que contiene instrucciones ejecutables para realizar dichos cálculos o formar dichos valores compuestos y/o para realizar la etapa de correlación según se describió anteriormente.

La elección de un valor de corte (o nivel de corte) depende de muchos factores que incluyen, pero no se limitan al riesgo de la enfermedad como tal y al riesgo asociado a diagnosticar erróneamente como positivo a un individuo que no padece la enfermedad (falso positivo). La elección del valor de corte se describe con más detalle a continuación.

En una realización de la presente invención, los SNP relacionados con el PCa (SNPpc) incluyen al menos dos de rs11672691, rs11704416, rs3863641, rs12130132, rs4245739, rs3771570, rs7611694, rs1894292, rs6869841, rs2018334, rs16896742, rs2273669, rs1933488, rs11135910, rs3850699, rs11568818, rs1270884, rs8008270, rs4643253, rs684232, rs11650494, rs7241993, rs6062509, rs1041449, rs2405942, rs12621278, rs9364554, rs10486567, rs6465657, rs2928679, rs6983561, rs16901979, rs16902094, rs12418451, rs4430796, rs11649743, rs2735839, rs9623117, y rs138213197.

En una realización, el método comprende además analizar, en dicha muestra biológica, una categoría de biomarcadores del PCa midiendo la presencia o la concentración de cada uno de una pluralidad de biomarcadores

5 del PCa de dicha categoría de biomarcadores del PCa; combinar los datos relativos a dicha categoría de biomarcadores del PCa para formar un valor compuesto de biomarcador; combinar el valor compuesto de biomarcador y el valor compuesto de SNPpc para formar un valor compuesto global; y correlacionar dicho valor compuesto global con la probabilidad de que dicho individuo requiera de terapia activa comparando el valor compuesto global con un valor predeterminado establecido con muestras control, en el que se sabe si los individuos de los cuales se obtienen las muestras control requirieron una terapia activa o no requirieron una terapia activa.

Preferiblemente, el método comprende medir la presencia o la concentración de biomarcadores del PCa al menos parcialmente redundantes, y en el que al menos uno, tal como dos, de los biomarcadores del PCa se selecciona del grupo que consiste en (i) PSA, (ii) PSA total (PSAt), (iii) PSA intacto (PSAi), (iv) PSA libre (PSAI), y (v) hK2.

10 Más en concreto, el método permite desestimar los datos de un subconjunto de al menos uno de dichos biomarcadores del PCa (i)-(v) de la categoría de biomarcadores del PCa cuando se forma dicho valor compuesto de biomarcador, tal como un subconjunto de uno, dos, tres o cuatro de dichos biomarcadores del PCa (i)-(v).

Además, en una realización, el método permite desestimar al menos 10%, tal como 15%, tal como 20%, tal como 30% de los SNPpc de la categoría de SNPpc cuando se forma el valor compuesto de SNPpc.

15 Preferiblemente, los datos relativos a la categoría de biomarcadores del PCa se combinan según una ecuación predeterminada para formar dicho valor compuesto de biomarcador y/o los datos relativos a la categoría de SNPpc se combinan según una ecuación predeterminada para formar dicho valor compuesto de SNPpc. Además, dicho valor compuesto de biomarcador y dicho valor compuesto de SNPpc se combinan preferiblemente según una ecuación predeterminada para formar dicho valor compuesto global.

20 En una realización, el anterior método comprende además una etapa de recomendar al individuo para recibir una terapia activa si el valor compuesto global es mayor que el valor de corte.

25 En una realización, el método comprende además analizar, en dicha muestra biológica, una categoría de SNP relacionados con una concentración de biomarcador de PCa (SNPbm) midiendo la presencia o la ausencia de al menos un SNPbm; combinar los datos relativos a dicho SNPbm para formar un valor compuesto de SNPbm; e incluir el valor compuesto de SNPbm en el valor compuesto global.

En una realización, el SNPbm incluye al menos uno de rs3213764, rs1354774, rs1227732, rs2736098, rs401681, rs10788160, rs11067228, rs1363120, rs888663, y rs1054564.

30 En una realización de la invención, el método comprende además analizar, en dicha muestra biológica, una categoría de SNP relacionados con el índice de masa corporal de dicho individuo (SNPbmi) midiendo la presencia o la ausencia de al menos un SNPbmi; combinar los datos relativos a dicha categoría de SNPbmi para formar un valor compuesto de SNPbmi; e incluir dicho valor compuesto de SNPbmi en el valor compuesto global.

En una realización, el SNPbmi incluye al menos uno de rs3817334, rs10767664, rs2241423, rs7359397, rs7190603, rs571312, rs29941, rs2287019, rs2815752, rs713586, rs2867125, rs9816226, rs10938397, y rs1558902.

35 En otra realización de la invención, el método comprende además recopilar la historia familiar con respecto al PCa, la historia de tratamiento y los datos físicos de dicho individuo; y en el que dicha historia familiar, historia de tratamiento y/o datos físicos se incluyen en el valor compuesto global.

40 En otra realización, el método comprende además analizar una categoría adicional de biomarcadores del PCa midiendo la presencia o la concentración de uno o cada uno de una pluralidad de biomarcadores del PCa de dicha categoría adicional de biomarcadores; combinar los datos relativos a dicha categoría adicional de biomarcadores del PCa para formar un valor compuesto de biomarcador adicional para dicha categoría adicional de biomarcadores del PCa; e incluir dicho valor compuesto de biomarcador adicional en un valor compuesto global; en el que la combinación de datos para formar el valor compuesto de biomarcador adicional se diseña de modo redundante cuando la categoría adicional de biomarcadores del PCa comprende más de un biomarcador del PCa.

45 En una realización preferida, la categoría adicional de biomarcadores del PCa comprende el biomarcador MIC-1 y opcionalmente otros biomarcadores relacionados con MIC-1, o el biomarcador MSMB y opcionalmente otros biomarcadores relacionados con MSMB.

50 En otra realización, el método comprende analizar cada una de pluralidad de categorías adicionales de biomarcadores del PCa y formar un valor compuesto de biomarcador adicional para cada una de las categorías de biomarcadores del PCa, según el procedimiento descrito anteriormente. Preferiblemente, se analizan al menos dos categorías adicionales de biomarcadores del PCa, en las que una categoría adicional de biomarcadores del PCa comprende el biomarcador MIC-1 y opcionalmente otros biomarcadores relacionados con MIC-1, y otra categoría adicional comprende el biomarcador MSMB y opcionalmente otros biomarcadores relacionados con MSMB.

En una realización, la muestra biológica es una muestra de sangre.

En una realización de la invención, el valor compuesto global se calcula empleando un método en el que se utiliza el efecto no aditivo de un SNPbm y la correspondiente concentración de biomarcador del PCa.

En una realización del método, la medición de la presencia o la ausencia de los SNP se realiza utilizando la espectrometría de masas MALDI.

- 5 En una realización del método, la medición de la presencia o la concentración de los biomarcadores del PCa se realiza utilizando la tecnología de micromatrices.

La medición de la presencia o la ausencia de un SNP (que pertenece a cualquier categoría de SNP) comprende medir el número de alelos de dicho SNP. En una realización, uno o dos alelos se corresponde con la presencia de dicho SNP, y cero alelos se corresponde con la ausencia de dicho SNP en dicho individuo; en el que cero alelos se
10 corresponde con homocigótico negativo para dicho SNP, un alelo se corresponde con heterocigótico positivo, y dos alelos se corresponde con homocigótico positivo.

- 15 En una realización, el método descrito anteriormente comprende la utilización de un dispositivo de ensayo de ELISA, un dispositivo de ensayo de micromatrices, un dispositivo de ensayo de inmunoprecipitación, un dispositivo de ensayo de inmunofluorescencia, un dispositivo de radioinmunoensayo, o un dispositivo de espectrometría de masas que emplea la ionización/desorción láser asistida por matriz ("matrix-assisted laser desorption/ionization", MALDI), para la medición de la presencia o la concentración de un biomarcador del PCa.

En una realización, que puede combinarse con la realización mencionada anteriormente, el método descrito anteriormente puede comprender la utilización de un dispositivo de espectrometría de masas que emplea la ionización/desorción láser asistida por matriz (MALDI) para la medición de la presencia o la ausencia de un SNP.

- 20 En la presente se describe además un dispositivo de ensayo para realizar la etapa 1 (es decir, la medición de la presencia o la ausencia de cada uno de una pluralidad de SNPpc) del método descrito anteriormente para valorar si un individuo diagnosticado con cáncer de próstata requerirá una terapia activa, que comprende una fase sólida sobre la que se encuentra inmovilizada una categoría de ligandos que se unen específicamente a un SNPpc, y que incluye una pluralidad de ligandos diferentes que se unen específicamente a cada uno de una pluralidad de SNPpc,
25 tal como al menos uno de rs11672691, rs11704416, rs3863641, rs12130132, rs4245739, rs3771570, rs7611694, rs1894292, rs6869841, rs2018334, rs16896742, rs2273669, rs1933488, rs11135910, rs3850699, rs11568818, rs1270884, rs8008270, rs4643253, rs684232, rs11650494, rs7241993, rs6062509, rs1041449, o rs2405942, rs12621278, rs9364554, rs10486567, rs6465657, rs2928679, rs6983561, rs16901979, rs16902094, rs12418451, rs4430796, rs11649743, rs2735839, rs9623117 y rs138213197.

- 30 En la presente se indica además que el dispositivo de ensayo puede adaptarse además para medir la presencia o la concentración de al menos un biomarcador del PCa, en el que la fase sólida contiene además una segunda categoría de ligandos inmovilizados que se unen específicamente a un biomarcador del PCa, y que incluye una pluralidad de ligandos diferentes que se unen específicamente a cada uno de una pluralidad de biomarcadores del PCa, tal como al menos uno de PSA, PSAi, PSA_t, PSA_i, y hK2, y opcionalmente MSMB y/o MIC-1.

- 35 En la presente se indica además que el dispositivo de ensayo puede adaptarse además para medir la presencia o la ausencia de un SNPbm, en cuyo caso la fase sólida contiene además una tercera categoría de ligandos inmovilizados que se unen específicamente a un SNPbm, tal como al menos uno de rs1227732, rs3213764, rs1354774, rs2736098, rs401681, rs10788160, rs11067228, rs1363120, rs888663, y rs1054564.

- 40 En la presente se indica además que el dispositivo de ensayo puede adaptarse además para medir la presencia o la ausencia de un SNPbmi, en cuyo caso la fase sólida contiene además una cuarta categoría de ligandos inmovilizados que se unen específicamente a un SNPbmi, tal como al menos uno de rs3817334, rs10767664, rs2241423, rs7359397, rs7190603, rs571312, rs29941, rs2287019, rs2815752, rs713586, rs2867125, rs9816226, rs10938397, y rs1558902.

- 45 El dispositivo de ensayo descrito anteriormente puede comprender un dispositivo de ensayo de ELISA, un dispositivo de ensayo de micromatrices, un dispositivo de ensayo de inmunoprecipitación, un dispositivo de ensayo de inmunofluorescencia, un dispositivo de radioinmunoensayo, o un dispositivo de espectrometría de masas que emplea la ionización/desorción láser asistida por matriz (MALDI), para la medición de la presencia o la concentración de un biomarcador del PCa.

- 50 Como alternativa o además, el dispositivo de ensayo descrito anteriormente puede comprender un dispositivo de espectrometría de masas que emplea la ionización/desorción láser asistida por matriz (MALDI) para la medición de la presencia o la ausencia de un SNP.

- 55 En la presente se describe además un kit de ensayo para realizar la etapa 1 (es decir, la medición de la presencia de cada uno de una pluralidad de SNPpc) del método descrito anteriormente para valorar si un individuo diagnosticado con cáncer de próstata requerirá una terapia activa, que comprende un correspondiente dispositivo de ensayo según se describió anteriormente, y una categoría de moléculas de detección, que son capaces de detectar un SNPpc, tal como al menos uno de rs11672691, rs11704416, rs3863641, rs12130132, rs4245739, rs3771570, rs7611694,

rs1894292, rs6869841, rs2018334, rs16896742, rs2273669, rs1933488, rs11135910, rs3850699, rs11568818, rs1270884, rs8008270, rs4643253, rs684232, rs11650494, rs7241993, rs6062509, rs1041449, o rs2405942, rs12621278, rs9364554, rs10486567, rs6465657, rs2928679, rs6983561, rs16901979, rs16902094, rs12418451, rs4430796, rs11649743, rs2735839, rs9623117 y rs138213197.

5 En la presente se indica además que el kit de ensayo puede comprender un dispositivo de ensayo que está adaptado también para medir la presencia o la concentración de al menos un biomarcador del PCa, y una segunda categoría de moléculas de detección, que son capaces de detectar un biomarcador del PCa, tal como al menos uno de PSA, PSAi, PSA_t, PSAI, y hK2, y opcionalmente MSMB y/o MIC-1.

10 En la presente se indica además que el kit de ensayo puede comprender un dispositivo de ensayo que está adaptado también para medir la presencia o la ausencia de al menos un SNP_{bmi}, y una tercera categoría de moléculas de detección, que son capaces de detectar un SNP_{bmi}, tal como al menos uno de rs1227732, rs3213764, rs1354774, rs2736098, rs401681, rs10788160, rs11067228, rs1363120, rs888663, y rs1054564.

15 En la presente se indica además que el kit de ensayo puede comprender un dispositivo de ensayo que está adaptado también para medir la presencia o la ausencia de al menos un SNP_{bmi}, y una cuarta categoría de moléculas de detección, que son capaces de detectar un SNP_{bmi}, tal como al menos uno de rs3817334, rs10767664, rs2241423, rs7359397, rs7190603, rs571312, rs29941, rs2287019, rs2815752, rs713586, rs2867125, rs9816226, rs10938397, y rs1558902.

20 En la presente se describe además un dispositivo de ensayo que comprende una fase sólida sobre la que se encuentra inmovilizada una categoría de ligandos que se unen específicamente a un SNP_{pc}, y que incluye una pluralidad de ligandos diferentes que se unen específicamente a cada uno de una pluralidad de SNP_{pc} diferentes, seleccionados de al menos uno de rs11672691, rs11704416, rs3863641, rs12130132, rs4245739, rs3771570, rs7611694, rs1894292, rs6869841, rs2018334, rs16896742, rs2273669, rs1933488, rs11135910, rs3850699, rs11568818, rs1270884, rs8008270, rs4643253, rs684232, rs11650494, rs7241993, rs6062509, rs1041449, o rs2405942, rs12621278, rs9364554, rs10486567, rs6465657, rs2928679, rs6983561, rs16901979, rs16902094, rs12418451, rs4430796, rs11649743, rs2735839, rs9623117 y rs138213197.

25 En la presente se indica además que, con respecto al dispositivo de ensayo, la fase sólida puede contener además una segunda categoría de ligandos inmovilizados que se unen específicamente a un biomarcador del PCa, y que incluye una pluralidad de ligandos diferentes que se unen específicamente a cada uno de una pluralidad de biomarcadores del PCa diferentes seleccionados de al menos uno de PSA, PSAi, PSA_t, PSAI, y hK2, y opcionalmente MSMB y/o MIC-1.

30 En la presente se indica además que, con respecto al dispositivo de ensayo, la fase sólida puede contener además una tercera categoría de ligandos inmovilizados que se unen específicamente a un SNP_{bmi}, y que incluye uno o una pluralidad de ligandos diferentes que se unen específicamente a uno o a cada uno de una pluralidad de SNP_{bmi} seleccionados de al menos uno de rs1227732, rs3213764, rs1354774, rs2736098, rs401681, rs10788160, rs11067228, rs1363120, rs888663, y rs1054564.

35 En la presente se indica además que, con respecto al dispositivo de ensayo, la fase sólida puede contener además una cuarta categoría de ligandos inmovilizados que se unen específicamente a un SNP_{bmi}, y que incluye uno o una pluralidad de ligandos diferentes que se unen específicamente a uno o a cada uno de una pluralidad de SNP_{bmi} diferentes seleccionados de al menos uno de rs3817334, rs10767664, rs2241423, rs7359397, rs7190603, rs571312, rs29941, rs2287019, rs2815752, rs713586, rs2867125, rs9816226, rs10938397, y rs1558902.

40 Otro aspecto de la invención proporciona un producto de programa informático que puede cargarse directamente en la memoria interna de un ordenador digital, que se caracteriza porque dicho producto comprende un medio de código de software para realizar al menos la etapa 2 (es decir, combinar los datos relativos a dicha categoría de SNP_{pc} para formar un valor compuesto de SNP_{pc}) y la etapa 3 (es decir, correlacionar dicho valor compuesto de SNP_{pc} con la probabilidad de que un individuo requiera de terapia activa comparando el valor compuesto de SNP_{pc} con un valor de corte predeterminado establecido con muestras control, en el que se sabe si los individuos de los cuales se obtienen las muestras control requirieron una terapia activa o no requirieron una terapia activa) del método descrito anteriormente para valorar si un individuo diagnosticado con cáncer de próstata requerirá una terapia activa, tal como la etapa 1 (es decir, en al menos una muestra biológica procedente de dicho individuo, analizar una categoría de SNP_{pc} midiendo la presencia o la ausencia de uno o dos alelos de riesgo en cada uno de una pluralidad de SNP_{pc}), la etapa 2 y la etapa 3 de dicho método.

45 En una realización, el producto de programa informático comprende además un medio de código de software para determinar la presencia o la concentración de al menos un biomarcador del PCa.

50 En una realización, el producto de programa informático comprende además un medio de código de software para analizar una categoría de SNP_{bmi} midiendo la presencia o la ausencia de al menos un SNP_{bmi}.

55 En una realización, el producto de programa informático comprende además un medio de código de software para analizar una categoría de SNP_{bmi} midiendo la presencia o la ausencia de al menos un SNP_{bmi}.

En la presente también se describe un aparato que comprende un dispositivo de ensayo tal como se describió anteriormente y un correspondiente producto de programa informático tal como se describió anteriormente.

Breve descripción de los dibujos

5 La figura 1 muestra las curvas de ROC para el modelo lineal del ejemplo 1 que ilustran las diferencias en la actuación entre PSA (101) y un modelo de puntuación genética (102) en la predicción de la necesidad de un tratamiento activo.

La figura 2 muestra las curvas de ROC para el modelo lineal del ejemplo 1 que ilustran las diferencias en la actuación entre PSA (201) y un modelo multiparamétrico (202) en la predicción de la necesidad de un tratamiento activo.

10 **Descripción detallada de la invención**

Para los objetivos de esta solicitud y para que sea más clara, se ofrecen las siguientes definiciones.

15 El término "PSA" se refiere al antígeno específico de próstata en suero en general. El PSA existe en diferentes formas. La expresión "PSA libre" se refiere a un PSA que no está unido a otra molécula, la expresión "PSA unido" se refiere a un PSA que está unido o complejoado con otra molécula y, por último, la expresión "PSA total" se refiere a la suma del PSA libre y PSA unido (complejado). La expresión "PSA L/T" es la proporción de PSA no unido al PSA total. También existen derivados moleculares del PSA, en los que el término "proPSA" se refiere a una forma inactiva precursora del PSA y la expresión "PSA intacto" se refiere a otra forma de proPSA que está intacta e inactiva.

20 La expresión "ensayo de diagnóstico" se refiere a la detección de la presencia o la naturaleza de un trastorno patológico. Puede emplearse de modo intercambiable con "método de diagnóstico". Los ensayos de diagnóstico se diferencian según su sensibilidad y especificidad.

25 La expresión "ensayo de pronóstico" se refiere a la predicción del desarrollo de un trastorno patológico existente. Puede emplearse de modo intercambiable con "método de pronóstico". Los ensayos de pronóstico, cuando proporcionan una prognosis sobre si un acontecimiento concreto se va a producir, son similares a los ensayos de diagnóstico y, en estos casos, se diferencian según su sensibilidad y especificidad. Un ejemplo es el ensayo de pronóstico que predice si será necesaria una terapia activa.

La expresión "tratamiento activo" indica el tratamiento de un paciente con PCa mediante cirugía, mediante radiación externa, mediante radioterapia dirigida, mediante quimioterapia, terapia hipotérmica, terapia hipertérmica o mediante cualquier otro procedimiento médico realizado con el fin de tratar el PCa.

30 Una medición de la utilidad de una herramienta de diagnóstico es el "área bajo la curva característica de receptor-operador", que se conoce habitualmente como estadística de ROC-AUC. Esta medición ampliamente aceptada toma en cuenta la sensibilidad y la especificidad de la herramienta. La medición de ROC-AUC generalmente varía de 0,5 a 1,0, en la que un valor de 0,5 indica que la herramienta no tiene un valor de diagnóstico, y un valor de 1,0 indica que la herramienta tiene 100% de sensibilidad y 100% de especificidad.

35 El término "sensibilidad" se refiere a la proporción de todos los sujetos que requieren de tratamiento activo que han sido correctamente identificados como tales (que es igual al número de positivos verdaderos dividido entre la suma del número de positivos verdaderos y falsos negativos).

El término "especificidad" se refiere a la proporción de todos los sujetos que no requieren tratamiento activo (es decir, adecuados para una terapia expectante) que han sido correctamente identificados como tales (que es igual al número de negativos verdaderos dividido entre la suma del número de negativos verdaderos y falsos positivos).

40 El término "biomarcador" se refiere a una proteína, una parte de una proteína, un péptido o un polipéptido que puede utilizarse como marcador biológico, por ejemplo, para objetivos de diagnóstico.

45 La expresión "biomarcador similar a la calicreína" se refiere a biomarcadores de proteínas que pertenecen o que están relacionados con la familia de proteínas de la calicreína, que incluyen, pero no se limitan al antígeno específico de próstata (PSA) en forma libre o en forma complejada, proPSA (una colección de isoformas de PSA) y en particular la forma truncada (-2) proPSA, PSA intacto, fosfatasa ácida prostática humana (PAP), y calicreína 2 humana (abreviada como hK2 o HK2 o hk2 en la presente solicitud).

50 La expresión "polimorfismo de un único nucleótido" (SNP) se refiere a las propiedades genéticas de un locus definido en el código genético de un individuo. Un SNP puede estar relacionado con un aumento del riesgo de PCa y, por tanto, puede utilizarse para evaluaciones de diagnóstico o pronóstico de un individuo. La base de datos de polimorfismos de un único nucleótido ("Single Nucleotide Polymorphism Database", dbSNP) es un archivo para la variación genética dentro y a través de diferentes especies desarrollada y alojada en the National Center for Biotechnology Information (NCBI) en colaboración con the National Human Genome Research Institute (NHGRI), ambas localizadas en EE. UU. Aunque el nombre de la base de datos insinúa que es una colección de una sola

clase de polimorfismos (es decir, polimorfismos de un único nucleótido (SNP)), de hecho contiene una gama de variación molecular. Cada registro de SNP exclusivo enviado recibe un número de referencia de ID de SNP ("rs#"; "agrupación de refSNP"). En esta solicitud, los SNP se identifican principalmente empleando los números de rs#. Por consiguiente, dentro de la presente solicitud, SNP se emplea para indicar la gama de variación molecular según se incluye en la dbSNP, en lugar de solo los polimorfismos de un único nucleótido. Para los objetivos de la presente solicitud, el término "SNP" puede usarse para describir el singular y/o el plural de "polimorfismo de un único nucleótido".

La expresión "índice de masa corporal" ("body-mass index", BMI) se refiere a una representación heurística de la grasa corporal humana basada en el peso y la altura de un individuo, según la fórmula $BMI = \text{peso}/(\text{altura}^2)$, en la que peso es el peso de un individuo expresado en kilogramos, y altura es la altura de un individuo expresada en metros. Se considera generalmente que un valor de BMI sano normal está en el intervalo de 18,5 y 25, y los individuos que tienen un $BMI > 30$ se consideran generalmente obesos.

La expresión "historia médica" se refiere a la información relativa a los exámenes históricos, los diagnósticos y/o las terapias para cualquier enfermedad cancerosa. Un ejemplo no limitante de una historia médica sería si un sujeto ha sido examinado previamente para determinar la presencia del PCa mediante una biopsia de la próstata.

La expresión "categoría de parámetros" se refiere a un grupo o a una familia de parámetros relacionados, tales como biomarcadores relacionados o SNP relacionados, que son parcial o completamente redundantes en términos de la actuación predictiva. Un ejemplo de una categoría de parámetros son los "biomarcadores similares a calicreína", una categoría que incluye, por ejemplo, PSA, PSA total (PSAt), PSA intacto (PSAi), PSA libre (PSAI), y hk2. Otro ejemplo de una categoría de parámetros son los "SNP relacionados con BMI", una categoría que incluye los SNP que están relacionados con el BMI de un individuo. En los modelos de predicción de la presente invención puede resultar suficiente poseer los resultados de mediciones (datos) para un subconjunto de los miembros de cada categoría para que cada categoría quede representada en el modelo de predicción, aunque se emplee solo un subconjunto de los miembros de las respectivas categorías. La expresión "categoría de parámetros" a veces se denomina solo "categoría" en la presente solicitud.

La expresión "valor compuesto" se refiere a la combinación de datos relacionados con una categoría de parámetros para obtener un valor representativo de dicha categoría de parámetros. La combinación de datos puede realizarse generalmente según una o más ecuaciones predeterminadas. Un valor compuesto es el resultado de la combinación de datos según una o más ecuaciones predeterminadas. Las diferentes ecuaciones son aplicables a diferentes resultados de mediciones (es decir, datos), dependiendo de los subconjuntos de los miembros de la categoría de parámetros para los cuales existen datos disponibles. Un ejemplo no limitante de un método para formar un valor compuesto para una categoría de parámetros concreta es emplear el promedio de los resultados disponibles para los miembros de dicha categoría. La expresión "valor compuesto" a veces se denomina "puntuación" en la presente solicitud. Un ejemplo no limitante de un valor compuesto es el "valor compuesto de biomarcador". Otro ejemplo no limitante de un valor compuesto es el "valor compuesto genético" (o "puntuación genética") y, de modo más específico, el "valor compuesto de SNP".

La expresión "combinación de datos diseñada de modo redundante" se refiere a una combinación de datos obtenidos por medio de un pluralidad de mediciones para formar un valor compuesto para una o más categorías de parámetros, o sus subconjuntos, en la que la combinación de datos se realiza de modo que puede producirse un valor compuesto que representa a una categoría de parámetros basándose en un subconjunto de datos, por ejemplo, en los que algunos datos faltan o son erróneos, o en el conjunto total de datos.

La expresión "una pluralidad", tal como se emplea en la presente solicitud, significa "dos o más".

La presente invención proporciona métodos de pronóstico para ayudar a indicar, valorar, detectar y/o determinar si se recomendará a un individuo para que reciba una terapia activa para el PCa. La presente invención puede ajustarse, si se desea, a subpoblaciones definidas para aumentar la actuación y la utilidad de la invención dentro de dicha subpoblación. Aunque la presente invención puede aplicarse a la población general de hombres, es posible construir métodos para indicar, valorar, detectar o determinar si se recomendará a un individuo para que reciba una terapia activa para el PCa con una actuación potenciada para subpoblaciones definidas, que incluyen, pero no se limitan a individuos que presentan un valor de PSA menor que aproximadamente 7 ng/ml (es decir, menor que un valor predeterminado de entre 1 ng/ml y 30 ng/ml) o una concentración de PSA libre menor que aproximadamente 0,91 ng/ml (es decir, menor que un valor predeterminado de entre 0,1 ng/ml y 3 ng/ml).

El principio básico de la invención es el uso de combinaciones de biomarcadores y de información genética de tal modo que el uso combinatorio de la información evaluada acerca del individuo mejora la calidad de la prognosis.

- Se recopila la historia familiar con respecto al PCa de dicho paciente (categoría HIST).

- Se recopilan los datos físicos del paciente, tales como altura, BMI, edad y similares (categoría PPD)

- Se obtienen de una serie de muestras biológicas de dicho paciente.

- En dichas muestras biológicas, se mide o se cuantifica la presencia o la concentración de una pluralidad de biomarcadores definidos (categoría biomarcador), seguido de la combinación de los datos con respecto a dichos biomarcadores para formar un valor compuesto de biomarcador.

5 - En dichas muestras biológicas, se mide o se cuantifica el estado genético de dichos pacientes con respecto a una pluralidad de SNP definidos relacionados con el PCa (categoría SNPpc), midiendo o cuantificando la presencia o la ausencia de una pluralidad de SNP definidos relacionados con el PCa (SNPpc), y seguido de la combinación de los datos con respecto a los SNP relacionados con el PCa para formar un valor compuesto de SNPpc.

10 - En dichas muestras biológicas, se mide o se cuantifica el estado genético de dichos pacientes con respecto a una pluralidad de SNP definidos relacionados con el nivel de expresión de un biomarcador o la concentración de un biomarcador (categoría SNPbm), midiendo o cuantificando la presencia o la ausencia de una pluralidad de SNP definidos relacionados con el nivel de expresión de un biomarcador o la concentración de un biomarcador (SNPbm) para formar un valor compuesto de SNPbm.

15 - Se emplea el valor compuesto de al menos la categoría de SNPpc, según se definió anteriormente, para evaluar la prognosis del cáncer de próstata. Habitualmente, los valores compuestos de al menos dos de las categorías definidas anteriormente se combinan para formar un valor compuesto global para su uso en la evaluación de la prognosis del cáncer de próstata.

- Se determina, empleando dicho valor compuesto de categoría o valor compuesto global, por sí solos o en combinación con otros datos, si es probable que el paciente requiera un tratamiento avanzado para el PCa.

20 Con más detalle, la etapa que comprende la recopilación de la historia familiar incluye, pero no se limita a la identificación de cualquier miembro masculino de la familia más cercana (tal como el padre, el hermano o el hijo del paciente) que padezca o haya padecido PCa.

La información física con respecto al paciente generalmente se obtiene mediante un examen físico frecuente mediante el cual se recopilan los datos de la edad, peso, altura, BMI y datos físicos similares.

25 La recogida de muestras biológicas de un paciente incluye, pero no se limita a muestras de plasma, suero, ADN de leucocitos periféricos y orina.

30 La cuantificación de la presencia o la concentración de biomarcadores en una muestra biológica puede realizarse de muchas formas diferentes. Un método habitual es el uso de los ensayos inmunoabsorbentes ligados a enzimas (ELISA), que emplean anticuerpos y una curva de calibración para evaluar la presencia y la concentración (cuando sea posible) de un biomarcador seleccionado. Los ensayos ELISA son habituales y conocidos en la técnica, tal como evidencia la publicación "Association between saliva PSA and serum PSA in conditions with prostate adenocarcinoma", de Shiiki N. y coautores, publicada en Biomarkers, septiembre de 2011, 16(6):498-503. Otro método habitual es el uso de un ensayo de micromatrices para la cuantificación de la presencia o la concentración de biomarcadores en una muestra biológica. Un ensayo de micromatrices típico comprende un portaobjetos de vidrio plano sobre el cual se unen una pluralidad de diferentes reactivos de captura (generalmente un anticuerpo) seleccionados cada uno para capturar específicamente un tipo de biomarcador, en áreas no solapantes sobre un lado del portaobjetos. Se deja que la muestra biológica se ponga en contacto, durante un periodo de tiempo definido, con el área en que están localizados los reactivos de captura, seguido del lavado del área de los reactivos de captura. En este punto, en el caso de que el biomarcador buscado estuviera presente en la muestra biológica, el correspondiente reactivo de captura habrá capturado una fracción del biomarcador buscado y lo mantendrá unido al portaobjetos de vidrio después del lavado. Después se añade un conjunto de reactivos de detección al área de los reactivos de captura (que ahora mantiene unidos potencialmente a los biomarcadores), y dichos reactivos de detección son capaces de (i) unirse al biomarcador presentado sobre el portaobjetos de vidrio, y (ii) producir una señal detectable (normalmente a través de la conjugación con un tinte fluorescente). Generalmente, es necesario añadir al portaobjetos de vidrio un reactivo de detección por biomarcador. Existen muchos otros métodos capaces de cuantificar la presencia o la concentración de un biomarcador que incluyen, pero no se limitan a ensayos de inmunoprecipitación, ensayos de inmunofluorescencia, radioinmunoensayos, y la espectrometría de masas que emplea la ionización/desorción láser asistida por matriz (MALDI), por mencionar unos pocos ejemplos.

45 La cuantificación de la presencia de SNP mediante el análisis de una muestra biológica generalmente implica un análisis de espectrometría de masas MALDI basado en extensiones de cebadores específicos de alelos, aunque también pueden aplicarse otros métodos. Esto se aplica a cualquier tipo de SNP, es decir, SNP relacionados con el PCa (SNPpc), SNP relacionados con el BMI (SNPbmi), y SNP relacionados con la expresión/concentración de biomarcadores (SNPbm).

55 La combinación de los datos puede ser cualquier tipo de combinación algorítmica de resultados, tales como una combinación lineal de datos, en la que la combinación lineal mejora la actuación de diagnóstico (por ejemplo, según se mide empleando ROC-AUC). Otros métodos posibles para ser combinados en un modelo capaz de producir una valoración diagnóstica incluyen, pero no se limitan a polinomios no lineales, máquinas de vectores de apoyo, clasificadores de redes neurales, análisis discriminante, bosques aleatorios, potenciación de gradientes, mínimos

cuadrados parciales, regresión contraída, lasso, redes elásticas, vecinos más cercanos a k. Además, el libro "The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition" de T. Hastie, R. Tibshirani y J. Friedman publicado por Springer Series en Statistics, ISBN 978-0387848570, describe muchos métodos adecuados para combinar datos para predecir o clasificar un resultado concreto.

- 5 Los biomarcadores adecuados para realizar pronósticos del PCa incluyen, pero no se limitan al antígeno específico de próstata (PSA) en forma libre o en forma complejada, proPSA (una colección de isoformas de PSA) y en particular la forma truncada (-2) proPSA, PSA intacto, fosfatasa ácida prostática humana (PAP), calicreína 2 humana (hK2), antígeno del cáncer de próstata temprano (EPCA), proteína secretora de la próstata (PSP94; también conocida como beta-microseminoproteína y MSMB), glutatión S-transferasa π (GSTP1), y α -metilcil coenzima A racemasa (AMACR). Los biomarcadores relacionados que pueden ser útiles para mejorar la precisión de diagnóstico del método incluyen la citoquina 1 inhibidora de macrófagos (MIC-1; también conocida como GDF-15).

- 15 Los SNP adecuados relacionados con el PCa incluyen, pero no se limitan a rs12621278 (cromosoma 2, locus 2q31.1), rs9364554 (cromosoma 6, locus 6q25.3), rs10486567 (cromosoma 7, locus 7p15.2), rs6465657 (cromosoma 7, locus 7q21.3), rs2928679 (cromosoma 8, locus 8p21), rs6983561 (cromosoma 8, locus 8q24.21), rs16901979 (cromosoma 8, locus 8q24.21), rs16902094 (cromosoma 8, locus 8q24.21), rs12418451 (cromosoma 11, locus 11q13.2), rs4430796 (cromosoma 17, locus 17q12), rs11649743 (cromosoma 17, locus 17q12), rs2735839 (cromosoma 19, locus 19q13.33), rs9623117 (cromosoma 22, locus 22q13.1), y rs138213197 (cromosoma 17, locus 17q21).

- 20 Los SNP adecuados relacionados con el PCa incluyen también, pero no se limitan a rs11672691, rs11704416, rs3863641, rs12130132, rs4245739, rs3771570, rs7611694, rs1894292, rs6869841, rs2018334, rs16896742, rs2273669, rs1933488, rs11135910, rs3850699, rs11568818, rs1270884, rs8008270, rs4643253, rs684232, rs11650494, rs7241993, rs6062509, rs1041449, y rs2405942.

- 25 Los SNP adecuados relacionados con el PCa incluyen también, pero no se limitan a rs138213197, tal como se describe en el artículo "Germline mutations in HOXB13 and prostate-cancer risk", de Ewing C.M. y coautores, publicado en N. Engl. J. Med., enero de 2012, 12, 366(2):141-9, 1100delC (22q12.1) y I157T (22q12.1), tal como se describe en el artículo "A novel founder CHEK2 mutation is associated with increased prostate cancer risk", de Cybulski C. y coautores, publicado en Cancer Res., abril de 2004, 15, 64(8):2677-9, y 657de15 (8q21), tal como se describe en el artículo "NBS1 is a prostate cancer susceptibility gene", de Cybulski C. y coautores, publicado en Cancer Res., febrero de 2004, 15, 64(4):1215-9.

- 30 Es posible definir una categoría de parámetros como "SNP relacionados con el PCa", que incluye los SNP relacionados con el PCa. Los miembros adecuados incluyen, pero no se limitan a los SNP listados anteriormente. Un subconjunto de los miembros de esta categoría sería suficiente para representar la categoría en un modelo predictivo.

- 35 Los SNP adecuados relacionados con otros procesos distintos del PCa incluyen también, pero no se limitan a rs3213764, rs1354774, rs2736098, rs401681, rs10788160, rs11067228, estando todos estos relacionados con el nivel de expresión del PSA. Es posible definir una categoría de parámetros como "SNP relacionados con la concentración de PSA" o "SNP relacionados con el nivel de expresión de PSA", que incluye los SNP relacionados con la concentración o el nivel de expresión del PSA. Un subconjunto de los miembros de esta categoría sería suficiente para representar la categoría en un modelo predictivo. Los SNP rs3213764 y rs1354774 se refieren en particular al nivel de expresión del PSA libre.

- 40 Los SNP adecuados relacionados con otros procesos distintos del PCa incluyen, pero no se limitan a rs1363120, rs888663, rs1227732, rs1054564, estando todos estos relacionados con el nivel de expresión del biomarcador de la citoquina de inflamación MIC1. Es posible definir una categoría de parámetros como "SNP relacionados con la concentración de MIC1" o "SNP relacionados con el nivel de expresión de MIC1", que incluye los SNP relacionados con la concentración o el nivel de expresión de MIC1. Un subconjunto de los miembros de esta categoría sería suficiente para representar la categoría en un modelo predictivo.

- 45 Es posible definir una categoría de parámetros como "SNP relacionados con la concentración de un biomarcador del PCa" o "SNP relacionados con el nivel de expresión de un biomarcador del PCa", que incluye los SNP relacionados con la concentración o el nivel de expresión de biomarcadores pertinentes, tales como el antígeno específico de próstata (PSA) en forma libre o en forma complejada, proPSA (una colección de isoformas de PSA) y en particular la forma truncada (-2) proPSA, PSA intacto, fosfatasa ácida prostática humana (PAP), calicreína 2 humana (hK2), antígeno del cáncer de próstata temprano (EPCA), proteína secretora de la próstata (PSP94; también conocida como beta-microseminoproteína, glutatión S-transferasa π (GSTP1), α -metilcil coenzima A racemasa (AMACR), y citoquina 1 inhibidora de macrófagos (MIC-1; también conocida como GDF-15). Un subconjunto de los miembros de esta categoría sería suficiente para representar la categoría en un modelo predictivo.

- 50 Los SNP adecuados relacionados con otros procesos distintos del PCa incluyen también, pero no se limitan a rs3817334, rs10767664, rs2241423, rs7359397, rs7190603, rs571312, rs29941, rs2287019, rs2815752, rs713586, rs2867125, rs9816226, rs10938397, y rs1558902, estando todos estos relacionados con el BMI de un individuo.

Otros SNP adecuados relacionados con BMI se describen en el artículo "Contribution of 32 GWAS-identified common variants to severe obesity in European adults referred for bariatric surgery", de Magi y coautores, publicado en PLoS One, agosto de 2013, 7, 8(8):e70735. Es posible definir una categoría de parámetros como "SNP relacionados con el nivel de expresión de BMI", que incluye los SNP relacionados con el BMI del individuo. Un subconjunto de los miembros de esta categoría sería suficiente para representar la categoría en un modelo predictivo.

Una colección preferida de SNP para ser empleados en la evaluación de la presencia o no presencia de un cáncer de próstata agresivo en un sujeto es rs582598, rs439378, rs2207790, rs1046011, rs10458360, rs7525167, rs10489871, rs7529518, rs4245739, rs4512641, rs10178804, rs11900952, rs1873555, rs10191478, rs6755901, rs6545962, rs721048, rs2710647, rs12612891, rs2028900, rs1009, rs12233245, rs6760417, rs10496470, rs10199796, rs12475433, rs16860513, rs12151618, rs3765065, rs13017302, rs12988652, rs871688, rs749264, rs3771570, rs4346531, rs6770955, rs12637074, rs2660753, rs13319878, rs6437715, rs2162185, rs1515542, rs2270785, rs9830294, rs1439024, rs6762443, rs888507, rs6794467, rs12490248, rs1477886, rs4833103, rs3796547, rs17779822, rs2366711, rs16849146, rs1894292, rs12640320, rs3805284, rs12500426, rs4699312, rs17021918, rs7679673, rs2047408, rs2647262, rs12506850, rs7658048, rs2078277, rs12505546, rs13113975, rs4246742, rs2736098, rs401681, rs11134144, rs10060513, rs40485, rs2087724, rs1482679, rs16901841, rs1295683, rs2070874, rs7752029, rs2018334, rs9358913, rs1140809, rs409558, rs3096702, rs9267911, rs2025645, rs9359428, rs6569371, rs2813532, rs1933488, rs712242, rs6934898, rs9456490, rs651164, rs3120137, rs9364554, rs9457937, rs10486562, rs10807843, rs7801918, rs6962297, rs2465796, rs6957416, rs7777631, rs2272316, rs6961773, rs2132276, rs13265330, rs16887736, rs2911756, rs2272668, rs2339654, rs1380862, rs9297746, rs12543663, rs10086908, rs16901922, rs1016343, rs17832285, rs16901979, rs4871779, rs10107982, rs16902094, rs620861, rs17467139, rs6983267, rs9297756, rs10094059, rs7818556, rs1992833, rs986472, rs12552397, rs4273907, rs4237185, rs753032, rs11253002, rs2386841, rs10795841, rs10508422, rs7075945, rs10508678, rs539357, rs10826398, rs3818714, rs7090755, rs10993994, rs4382847, rs1891158, rs10887926, rs10788160, rs6579002, rs10832514, rs7358335, rs1944047, rs3019779, rs10896437, rs12793759, rs7106762, rs7102758, rs2449600, rs585197, rs2509867, rs11568818, rs7125415, rs11601037, rs11222496, rs4570588, rs6489721, rs3213764, rs17395631, rs4423250, rs11168936, rs10875943, rs3759129, rs902774, rs1827611, rs4760442, rs11610799, rs6539333, rs11067228, rs7485441, rs6489794, rs4119478, rs17070292, rs2293710, rs17256058, rs1950198, rs2331780, rs7141529, rs12880777, rs17123359, rs785437, rs524908, rs12903579, rs7178085, rs7164364, rs896615, rs11634741, rs9972541, rs12594014, rs11631109, rs1558902, rs8044335, rs2738571, rs885479, rs385894, rs684232, rs4925094, rs17138478, rs11649743, rs2107131, rs7213769, rs12946864, rs306801, rs138213197, rs1863610, rs17224342, rs9911515, rs12947919, rs966304, rs17744022, rs7234917, rs1943821, rs2227270, rs1363120, rs888663, rs1227732, rs1054564, rs4806120, rs11672691, rs758643, rs3745233, rs6509345, rs2659051, rs2735839, rs1354774, rs2691274, rs6090461, rs2297434, rs6062509, rs2315654, rs2823118, rs2838053, rs398146, rs16988279, rs2269640, rs4822763, rs132774, rs747745, rs5978944, rs6530238, rs5934705, rs5935063, rs4830488, rs17318620, rs5945619, rs5945637, rs11091768, rs2473057, rs5918762, rs4844228, rs6625760 y rs17324573. Aunque el uso de la lista completa es preferible, cualquier subconjunto de esta lista es adecuado para su uso en la evaluación de la presencia o no presencia de un cáncer de próstata agresivo en un sujeto. Los SNP en esta lista (todos o un subconjunto que comprende aproximadamente 95%, o 90%, o 85%, o 80%, o 75%, o 70% de los SNP en esta lista) pueden colocarse sobre el mismo soporte sólido, por ejemplo, el mismo portaobjetos de vidrio, para la detección simultánea en un instrumento analítico adecuado.

Una consecuencia inevitable de las dificultades para obtener una valoración precisa y comparable de la actuación predictiva de cualquier modelo concreto de diagnóstico o pronóstico en la selección del PCa es que cuando se calcula la mejora relativa de un método nuevo, en comparación con la utilización del PSA por sí solo, la mejora relativa calculada variará dependiendo de muchos factores. Un factor importante que influye en la mejora relativa calculada es la forma en que se obtiene el grupo control (es decir, los negativos conocidos). Por ejemplo, puesto que no resulta ético realizar biopsias en sujetos que no presentan indicaciones de PCa, el grupo control a menudo será seleccionado con un sesgo. Así, la mejora relativa de un nuevo método depende en la forma en que se selecciona el grupo control. Por tanto, cualquier mejora calculada indicada debe considerarse a la luz de dicha varianza. Según la experiencia de los inventores, para los ensayos de diagnóstico se calcula que si se indica que la mejora relativa de un nuevo método es 15%, comparado con el valor del PSA por sí solo empleando un método imparcial para seleccionar el grupo control, dicho método será al menos 10% mejor que el valor del PSA por sí solo empleando cualquier otro método imparcial para seleccionar el grupo control. Para los ensayos de pronóstico, la comparación con un criterio de referencia es igualmente difícil, y cualquier declaración sobre la actuación del ensayo de pronóstico (en este documento y en otros) debe considerarse a la luz de la varianza inducida por la elección del grupo control.

Un posible método para obtener un método de selección para el PCa que cumpla con los requisitos de un uso generalizado es combinar la información procedente de múltiples fuentes. A un nivel global, esto comprende combinar los valores obtenidos del análisis de los biomarcadores (por ejemplo, los valores del PSA), los perfiles genéticos (por ejemplo, el perfil de SNP), la historia familiar y otras fuentes. Esta combinación tiene la posibilidad de producir una declaración de diagnóstico mejor que cualquiera de los factores incluidos por sí solos. Se han descrito intentos de combinar valores en un modelo multiparamétrico para producir mejores declaraciones de diagnóstico, en

otro punto de la presente solicitud. Puede aplicarse la misma estrategia para los métodos de pronóstico

El algoritmo que transforma los datos procedentes de las diferentes categorías en un único valor que es indicativo de la probabilidad del paciente de padecer PCa es preferiblemente una función no lineal, en la que se emplea la dependencia de las diferentes categorías para aumentar aún más la actuación de diagnóstico del método. Por ejemplo, una dependencia importante es el nivel medido de un biomarcador seleccionado combinado con cualquier marcador genético asociado relacionado con el nivel de expresión esperado de dicho biomarcador. En los casos en que se encuentra una elevada concentración del biomarcador en una muestra de un paciente y, al mismo tiempo, dicho paciente está genéticamente predispuesto a presentar niveles menores de dichos biomarcadores, la importancia del nivel elevado del biomarcador aumenta. De forma similar, si el nivel de un biomarcador es claramente menor que el normal en un paciente que está genéticamente predispuesto a presentar niveles mayores de dichos biomarcadores, este descubrimiento contradictorio aumenta la importancia de la interpretación del nivel del biomarcador.

El algoritmo utilizado para predecir una terapia preferida puede aprovechar el uso de variables transformadas, por ejemplo, usando el valor de $\log_{10}(\text{PSA})$. La transformación es particularmente beneficiosa para las variables con una distribución que se desvía claramente de la distribución normal. Las posibles transformaciones de variables incluyen, pero no se limitan al logaritmo, la inversa, el cuadrado y la raíz cuadrada. También es habitual centrar cada variable al promedio nulo y varianza unitaria.

Aunque la combinación de los datos puede realizarse de diferentes formas, puede ilustrarse un procedimiento típico según la presente invención de la siguiente manera no limitante.

En un caso típico, los datos relativos a los biomarcadores que pertenecen a una categoría de parámetros se combinarán según una ecuación predeterminada para formar un valor compuesto que está relacionado con el riesgo relacionado con la categoría de parámetros. Un ejemplo no limitante consiste en calcular el valor promedio de todos los valores de mediciones disponibles (datos) procedentes de los miembros de una categoría de biomarcadores y emplear dicho valor promedio como el valor compuesto que representa a dicha categoría de biomarcadores. Este procedimiento claramente puede aplicarse independientemente del número de miembros de biomarcadores que pertenecen a la categoría. Si solo están disponibles los datos de uno de los biomarcadores incluidos en la categoría, este puede emplearse en sí mismo para representar a la categoría de biomarcadores. Para los biomarcadores, el valor medido que se emplea habitualmente en la etapa de combinación de los datos es la concentración de dicho biomarcador en la muestra biológica. Por ejemplo, para los biomarcadores PSA y HK2, esta es, de modo más habitual, la concentración del biomarcador en una muestra de sangre expresada en unidades de ng/ml.

El cálculo de la puntuación genética (es decir, el valor compuesto genético o, más específicamente, el valor compuesto de SNP) se basa generalmente en una oportunidad relativa predeterminada para cada SNP individual incluido en una categoría de parámetros. Para cada SNP, la oportunidad relativa, es decir, la probabilidad de que un individuo que porta un SNP (es decir, presenta el riesgo de alelo definido por el SNP) padezca la enfermedad o el trastorno en estudio, se determina con antelación. La determinación de la oportunidad relativa para un SNP normalmente se realiza en grandes estudios prospectivos que implican a miles de sujetos con un trastorno o enfermedad conocido.

La puntuación genética para cada individuo puede calcularse, en un ejemplo no limitante, según el siguiente algoritmo: Para el individuo ensayado, cada SNP se procesa de la siguiente manera. Para cada SNP, el individuo puede portar dos alelos de riesgo de SNP (homocigótico positivo para dicho SNP), o un alelo de riesgo (heterocigótico positivo para dicho SNP) o cero alelos de riesgo (homocigótico negativo para dicho SNP). El número de alelos para un SNP se multiplica por el logaritmo natural de la oportunidad relativa para dicho SNP para formar un valor de evaluación de riesgo para ese SNP concreto. Esto significa que en un individuo que es negativo para un SNP concreto (es decir, tiene cero alelos de riesgo de SNP), dicho SNP concreto no contribuirá al riesgo. Este procedimiento se repite para todos los SNP cuyos datos de mediciones estén disponibles. Cuando se han calculado todos los valores de evaluación del riesgo, se calcula el promedio de la contribución al riesgo para los SNP cuyos datos de mediciones estén disponibles y se emplea como la puntuación genética para dicho individuo, es decir, el valor compuesto genético con respecto a cierta categoría de SNP. Este procedimiento claramente puede aplicarse independientemente del número de miembros de SNP que pertenecen a la categoría de SNP.

Para ilustrar más a fondo un procedimiento típico según la presente invención cuando se aplica a un individuo, se realizan las siguientes suposiciones. Se definen dos categorías de parámetros, en primer lugar una categoría de biomarcadores de proteínas (o categoría de biomarcadores) que contiene los miembros Prot1 y Prot2, y en segundo lugar una categoría genética (o, más específicamente, una categoría de SNP) que contiene los miembros Snp1, Snp2, y Snp3. En un experimento que implica a 100 individuos con la condición conocida C y 100 individuos que no tienen la condición C, se establece la relación de Prot1, Prot2, Snp1, Snp2, y Snp3 con la condición C y se formula como un valor compuesto de biomarcador de proteína para Prot1 y Prot2, y un valor compuesto genético para Snp1, Snp2, y Snp3, y también un valor compuesto global que, a su vez, está relacionado con el riesgo de presentar la condición C. El valor compuesto para la categoría de biomarcadores de proteínas se calcula empleando las siguientes ecuaciones predeterminadas:

ES 2 694 409 T3

$P = (\text{Prot1} + 2 \cdot \text{Prot2})/3$ [si están disponibles los datos con respecto a ambos Prot1 y Prot2 (es decir, el valor de Prot1 y el valor de Prot2)]

$P' = \text{Prot1}$ [en el caso de que solo estén disponibles los datos con respecto a Prot1 (es decir, el valor de Prot1)]

$P'' = \text{Prot2}$ [en el caso de que solo estén disponibles los datos con respecto a Prot2 (es decir, el valor de Prot2)]

- 5 Por tanto, en este caso hipotético, en el experimento se descubrió que (a) Prot1 y Prot2 tienen la misma escala, y (b) el valor de Prot2 es dos veces más importante para evaluar si el individuo presenta la condición C que Prot1. Si solo están disponibles los datos de uno de los biomarcadores de proteínas, este puede emplearse en sí mismo para representar a la categoría de biomarcadores de proteínas.

- 10 La oportunidad relativa para los miembros de la categoría genética se determinó con antelación y es la siguiente: Snp1 = 1,1; Snp2 = 1,2; y Snp3 = 1,3. El valor compuesto para la categoría genética se calcula como la puntuación genética descrita anteriormente.

El valor compuesto de biomarcador de proteína y la puntuación genética (que en este caso es equivalente al valor compuesto de la categoría genética, o el valor compuesto de SNP) entonces se combinan en un valor compuesto global según la siguiente ecuación predeterminada:

- 15 $Y = P + 10 \cdot \text{puntuación}$

- en la que Y se relaciona con el riesgo de presentar la condición C, P es el valor compuesto de biomarcador de proteína (y P puede sustituirse por P' o P'' según se definieron anteriormente), y puntuación es la puntuación genética. Todas las ecuaciones deben desarrollarse basándose en un grupo grande de individuos, en este caso hipotético 100 + 100 individuos, de los cuales se deriva la relación entre Y y la enfermedad o trastorno en estudio.
20 En este caso hipotético se supone que si $Y > 5$, el riesgo de que un individuo presente la condición C es elevada, y si $Y > 10$, el riesgo es muy alto.

Ahora supongamos que un primer individuo A está siendo ensayado para Prot1, Prot2, Snp1, Snp2, y Snp3. En este caso concreto, todas las mediciones tuvieron éxito y produjeron los siguientes resultados:

Prot1 = 3 ng/ml

- 25 Prot2 = 6 ng/ml

Snp1 = homocigótico negativo, es decir, sin alelos de riesgo = 0

Snp2 = heterocigótico positivo, es decir, un alelo de riesgo = 1

Snp3 = homocigótico positivo, es decir, dos alelos de riesgo = 2

- 30 El valor compuesto para la categoría de biomarcadores de proteínas será, en este caso, $P = (3 + 2 \cdot 6)/3 = 5$. El valor compuesto para la categoría genética, también conocido como la puntuación genética, se convierte en la puntuación $= (0 \cdot \log(1,1) + 1 \cdot \log(1,2) + 2 \cdot \log(1,3))/3 = 0,2357$. El valor compuesto global se convierte en $Y = 5 + 10 \cdot 0,2357 = 7,357$. Por tanto, se calcula que el riesgo de presentar la condición C para el individuo A es elevado, pero no muy alto.

- 35 Ahora supongamos que un segundo individuo B está siendo ensayado para Prot1, Prot2, Snp1, Snp2, y Snp3. En este caso concreto, tres mediciones tuvieron éxito y produjeron los siguientes resultados:

Prot1 = 2 ng/ml

Prot2 = FALTAN DATOS

Snp1 = homocigótico positivo, es decir, dos alelos de riesgo = 2

Snp2 = FALTAN DATOS

- 40 Snp3 = heterocigótico positivo, es decir, un alelo de riesgo = 1

- El valor compuesto para la categoría de biomarcadores de proteínas será, en este caso, $P' = 2$, porque solo están disponibles los resultados de Prot1. El valor compuesto para la categoría genética, también conocido como la puntuación genética, se convierte en la puntuación $= (2 \cdot \log(1,1) + 1 \cdot \log(1,3))/2 = 0,2264$. El valor compuesto global se convierte en $Y = 2 + 10 \cdot 0,2264 = 4,264$. Por tanto, se calcula que el riesgo de presentar la condición C para el individuo B es bajo.

En general, en los modelos que predicen el riesgo de desarrollar PCa, a menudo se definen uno o más valores de corte. La elección del valor de corte (o nivel de corte) depende de muchos factores que incluyen, pero no se limitan al riesgo de la enfermedad como tal y al riesgo asociado a diagnosticar erróneamente como positivo a un individuo

que no padece la enfermedad (falso positivo). En el caso general, un modelo predictivo es habitualmente una función monótonica $Y = f(x_1, x_2, \dots, x_N)$, en la que el riesgo calculado de padecer la enfermedad se correlaciona con el valor creciente de Y . Esto significa que si el valor de corte se ajusta a un nivel bajo, el ensayo producirá un número grande de resultados falsos positivos, pero, por otra parte, detectará a la mayoría de los individuos que realmente padecen la enfermedad. Si el valor de corte se ajusta a un nivel alto ocurre lo contrario, y los individuos que presentan un valor de Y por encima del nivel de corte tendrán muy probablemente la enfermedad, pero un gran número de individuos con la enfermedad recibirán unos resultados del ensayo negativos (es decir, un gran número de resultados falsos negativos). La elección del nivel de corte depende de muchos factores, que incluyen las consecuencias socioeconómicas de sopesar (a) no detectar los individuos con la enfermedad, y (b) tratar individuos sin la enfermedad.

Cuando se aplica en la práctica, a veces puede suceder que una o unas pocas mediciones fallan debido, por ejemplo, a problemas técnicos imprevistos, al error humano o a cualquier otra razón inesperada y poco frecuente. En estos casos, el conjunto de datos obtenidos para un individuo será incompleto. Generalmente, dicho conjunto de datos incompletos será muy difícil o incluso imposible de evaluar. Sin embargo, la presente invención se basa en las mediciones de un gran número de características, de las cuales muchas son parcialmente redundantes. Esto significa que, también en el caso de individuos para los cuales el conjunto de datos sea incompleto, en muchas ocasiones será posible producir una evaluación de alta calidad según la invención. Esto resulta particularmente verdadero dentro de las categorías, en las que, por ejemplo, los biomarcadores similares a calicreína se correlacionan y son parcialmente redundantes. Por tanto, desde el punto de vista técnico, es posible aplicar una estrategia algorítmica de dos etapas, en la que la contribución del biomarcador de calicreína se resume en una puntuación de calicreína (o valor de calicreína). Esta puntuación de calicreína después, en una segunda etapa, se combina con otros datos (tales como la puntuación genética, la edad y la historia familiar, por mencionar algunos ejemplos no limitantes) para producir una declaración de diagnóstico o pronóstico sobre el PCa. Pueden aplicarse otros procedimientos en dos etapas similares para otras clases de marcadores, tales como marcadores genéticos relacionados con BMI o biomarcadores de proteínas relacionados con la superfamilia del factor de crecimiento transformante beta (una gran familia de proteínas reguladoras celulares relacionadas desde el punto de vista estructural, que incluye MIC-1), por mencionar dos ejemplos no limitantes.

El aspecto de la redundancia puede expresarse de muchas maneras diferentes. Un modo posible de aplicar el aspecto de la redundancia es definir un conjunto de biomarcadores que representan biomarcadores relacionados con una familia o un campo común. Un ejemplo no limitante de dicha familia o campo son los biomarcadores similares a calicreína. Puede determinarse más de un conjunto definido (o categoría) de biomarcadores y, además, otros biomarcadores pueden aplicarse fuera de dicho conjunto. Generalmente, las categorías no se solapan, es decir, cualquier biomarcador definido solo es miembro de una categoría definida o se emplea de manera solitaria. Después, se intenta determinar la presencia o la concentración de todos los biomarcadores. En la mayoría de los casos, la determinación de todos los biomarcadores se realizará con éxito, pero a veces faltarán uno o unos pocos valores. Para inducir la robustez del modelo frente a los valores que faltan, es posible definir un valor compuesto de categoría de biomarcadores que puede determinarse empleando todos los miembros o un subconjunto de los miembros de la categoría definida. En la práctica, esto requiere que los miembros de la categoría definida de biomarcadores sean al menos parcialmente redundantes. En la siguiente etapa, el valor compuesto de categoría de biomarcadores se combina con otros valores de biomarcadores, otros valores compuestos de categorías de biomarcadores (si se definen dos o más categorías de biomarcadores), la puntuación genética relacionada con el riesgo de PCa, la puntuación genética relacionada con otras características (tales como BMI o concentración del biomarcador, por mencionar dos ejemplos no limitantes), la historia familiar, la edad y otros portadores de información relacionados con la valoración de una terapia preferida para conformar un valor compuesto global. Por último, el valor compuesto global se emplea para la valoración de la terapia preferida.

Por tanto, el objetivo del valor compuesto de categoría de biomarcadores es actuar como un valor intermedio que puede calcularse empleando datos incompletos. Se supone que una categoría definida de biomarcadores comprende N biomarcadores diferentes denominados $B_1, B_2, B_3, \dots, B_N$, todos relacionados con la familia de biomarcadores B . En este caso, puede haber N modelos diferentes disponibles para calcular el valor compuesto de la familia B de biomarcadores C :

$$C = f_1(B_1, B_2, B_3, \dots, B_N)$$

$$C = f_2(B_2, B_3, \dots, B_N)$$

$$C = f_3(B_1, B_3, \dots, B_N)$$

...

$$C = f_N(B_1, B_2, B_3, \dots, B_{N-1})$$

en los que $f_1()$, $f_2()$... $f_N()$ son funciones matemáticas que emplean los valores para los biomarcadores B_1, \dots, B_N como entradas y, de alguna manera, producen un único resultado C que representa al valor compuesto de la familia B de biomarcadores. Un ejemplo no limitante de las funciones $f_1()$, ... $f_N()$ incluye combinaciones lineales de los

presentes argumentos. Con dicho conjunto de múltiples funciones capaz de calcular C para todos los casos de la ausencia del valor de un único biomarcador, el cálculo del valor compuesto global es menos sensible a los datos que faltan. Se entiende que el cálculo de C puede ser de menor calidad cuando no están presentes todos los datos, pero sigue siendo lo bastante bueno para emplearlo en la valoración de la terapia preferida. Así, empleando dicha estrategia, solo deben tener éxito N-1 determinaciones de biomarcadores para producir un cálculo de C. También es posible desarrollar cálculos para cualquier número de datos perdidos, es decir, si han tenido éxito N-2 determinaciones de biomarcadores, puede desarrollarse otro conjunto de funciones f() y aplicarse al cálculo de C.

Así, con respecto a los biomarcadores del PCa, la presente invención se refiere a un método basado en una combinación de datos diseñada de modo redundante, según se ha definido en otro punto de la presente solicitud. De modo más específico, el método comprende medir la presencia o la concentración de biomarcadores del PCa al menos parcialmente redundantes, y en el que al menos uno, tal como dos, de los biomarcadores del PCa se selecciona del grupo que consiste en (i) PSA, (ii) PSA total (PSAt), (iii) PSA intacto (PSAi), (iv) PSA libre (PSAI), y (v) hK2. El método permite desestimar un subconjunto de al menos uno de dichos biomarcadores del PCa (i)-(v) cuando se forma dicho valor compuesto de biomarcador. En otras palabras, el método permite formar el valor compuesto de biomarcador a partir de datos relativos a menos de todos los biomarcadores del PCa de la categoría de biomarcadores, de modo más específico a partir de datos relativos a un subconjunto de un máximo de cuatro de dichos biomarcadores del PCa. Tal como apreciarán los expertos en la técnica, esto es equivalente a un método en el que se necesitan los datos relativos a un subconjunto de un máximo de cuatro de dichos biomarcadores del PCa para formar dicho valor compuesto de biomarcador. Una ventaja del método según la presente invención es que es aceptable la omisión, falta o pérdida de datos relativos a un subconjunto de dichos biomarcadores del PCa cuando se forma el valor compuesto de biomarcador.

Tal como apreciarán los expertos en la técnica, la presente invención incluye que el método comprenda formar el valor compuesto de biomarcador a partir de datos relativos a todos los biomarcadores de la categoría de biomarcadores, con la condición de que estén disponibles datos relativos a todos los biomarcadores.

En una realización, el método permite desestimar un subconjunto de uno, dos, tres o cuatro de los biomarcadores del PCa (i) PSA, (ii) PSA total (PSAt), (iii) PSA intacto (PSAi), (iv) PSA libre (PSAI), y (v) hK2. En otras palabras, el método permite formar dicho valor compuesto de biomarcador a partir de datos relativos a un subconjunto de cuatro, tres, dos o uno de los biomarcadores del PCa (i)-(v), respectivamente.

Tal como se mencionó anteriormente en la presente solicitud, el método puede comprender además analizar una o cada una de pluralidad de categorías adicionales de biomarcadores del PCa, en el que la combinación de datos para formar cada valor compuesto de biomarcador adicional se diseña de modo redundante cuando la categoría adicional de biomarcadores del PCa comprende más de un biomarcador del PCa. El método permite desestimar un subconjunto de los biomarcadores del PCa cuando se forma el valor compuesto de biomarcador. En otras palabras, el método permite formar el valor compuesto de biomarcador a partir de datos relativos a menos de todos los biomarcadores del PCa de la categoría adicional de biomarcadores, tales como los datos relativos a un subconjunto de 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, o 90% de los biomarcadores del PCa del biomarcador del PCa adicional. Tal como apreciarán los expertos en la técnica, la presente invención incluye que el método comprenda formar cada valor compuesto de biomarcador adicional a partir de datos relativos a todos los biomarcadores del PCa de la categoría de biomarcadores del PCa, con la condición de que estén disponibles datos relativos a todos los biomarcadores del PCa.

Las puntuaciones de riesgo genéticas (es decir, las puntuaciones genéticas, o valores compuestos genéticos, más en concreto los valores compuestos de SNP) también son insensibles a pequeñas pérdidas de datos debidas, por ejemplo, a problemas técnicos imprevistos, al error humano o a cualquier otra razón inesperada y poco frecuente. La contribución de un SNP a la puntuación de riesgo generalmente no se correlaciona con ningún otro SNP. En el caso de los SNP, el cambio en el riesgo debido a cada SNP es pequeño, pero mediante el uso de múltiples SNP relacionados con un trastorno de forma concertada, el cambio en el riesgo para dicho trastorno aumenta lo suficiente como para tener un impacto sobre la actuación del modelo. El número preferido de SNP para formar una puntuación genética es de al menos 3 SNP, más preferiblemente 10 SNP, más preferiblemente 25 SNP, aún más preferiblemente 50 SNP, más preferiblemente 60 SNP, aún más preferiblemente 70 SNP, aún más preferiblemente 80 SNP, más preferiblemente 90 SNP, aún más preferiblemente 100 SNP, aún más preferiblemente 150 SNP, aún más preferiblemente 200 SNP, aún más preferiblemente 250, y aún más preferiblemente 300 SNP. Esto significa que el impacto de cualquier SNP individual en el resultado total es generalmente pequeño, y que la omisión de unos pocos SNP generalmente no alterará la evaluación de riesgo de la puntuación genética global en gran medida, es decir, generalmente no alterará el valor compuesto de SNP en un grado significativo. En el estado actual de la técnica, la pérdida de datos típica en las mediciones genéticas a gran escala es del orden del 1-2%, lo cual significa que si una puntuación genética está compuesta por 100 SNP diferentes, la caracterización genética típica de un individuo proporcionará información acerca de aproximadamente 98-99 de estos SNP. El presente modelo, tal como se ha descubierto en los trabajos de la presente invención, puede no obstante soportar una pérdida mayor o una falta de datos, tal como una pérdida del 5-7% de la información, o 7-15%, o incluso 15-30%. En este sentido, la combinación de datos con respecto a SNPpc es al menos parcialmente redundante.

Por consiguiente, también con respecto a los marcadores (SNP), la presente invención se refiere a un método que

se basa en una combinación de datos diseñada de modo redundante, según se ha definido en otro punto de la presente solicitud. El método permite desestimar al menos 5% de los SNPpc cuando se forma el valor compuesto de SNP. En otras palabras, el método permite formar dicho valor compuesto de SNPpc a partir de datos relativos a menos de todos los SNPpc de la categoría de SNPpc, de modo más específico a partir de datos relativos a un subconjunto de un máximo de 95% de dichos SNPpc. Tal como apreciarán los expertos en la técnica, esto es equivalente a un método en el que se necesitan los datos relativos a un subconjunto de un máximo de 95% de dichos SNPpc para formar dicho valor compuesto de SNPpc. Una ventaja del método según la presente invención es que es aceptable la omisión, falta o pérdida de datos relativos a un subconjunto de dichos SNPpc cuando se forma el valor compuesto de SNPpc.

10 Tal como apreciarán los expertos en la técnica, la presente invención incluye que el método comprenda formar el valor compuesto de SNPpc a partir de datos relativos a todos los SNPpc de la categoría de SNPpc, si están disponibles datos relativos a todos los SNPpc. De modo similar, la presente invención incluye que el método comprenda formar el valor compuesto de SNPpc a partir de datos relativos a un subconjunto de 99%, 98%, 97%, o 96% de dichos SNPpc.

15 En una realización, el método permite desestimar 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, o 30% de los SNPpc cuando se forma el valor compuesto de SNPpc. En otras palabras, el método permite formar dicho valor compuesto de SNPpc partir de datos relativos a un subconjunto de 94%, 93%, 92%, 91%, 90%, 85%, 80%, 75%, o 70% de los SNPpc, respectivamente.

20 Un ejemplo no limitante de dicha combinación de datos diseñada de modo redundante es el cálculo del promedio del riesgo relacionado con cada SNP para el cual existen datos de medición. Otro ejemplo no limitante de dicha combinación de datos diseñada de modo redundante es la proporción de múltiples ecuaciones independientes para calcular el valor compuesto, una ecuación para cada subconjunto de datos que puede utilizarse para producir dicho valor compuesto.

25 Un método adecuado para asociar un SNP con un trastorno (por ejemplo, PCa, o una concentración elevada del biomarcador hk2 en sangre) se ha descrito en el informe público "Blood Biomarker Levels to Aid Discovery of Cancer-Related Single-Nucleotide Polymorphisms: Kallikreins and Prostate Cancer", de Robert Klein y coautores, publicado en *Cancer Prey Res.*, 2010, 3:611-619. En este informe, los autores describen el modo en que asociaron el SNP rs2735839 con un valor elevado de (PSA libre)/(PSA total). Además, pudieron asociar el SNP rs10993994 con un riesgo elevado de PCa, un valor de PSA total elevado, un valor de PSA libre elevado y un valor de hk2 elevado y, por último, el SNP rs198977 fue asociado con un riesgo elevado de PCa, un valor elevado de (PSA libre)/(PSA total) y un valor elevado de hk2.

35 En la práctica, un método habitual para asociar un SNP con un trastorno se basa en el acceso a un ensayo clínico de casos y controles que compara dos grandes grupos de individuos, un grupo control sano y un grupo de casos que padecen el trastorno que se está estudiando. Todos los individuos en cada grupo se genotipificaron para la mayoría de los SNP conocidos habituales. Cuando todos los datos de la genotipificación estuvieron disponibles, se investigó si la frecuencia alélica se encontraba significativamente alterada entre el grupo de casos y el grupo control. En estos sistemas, la unidad típica para indicar el tamaño del efecto es la oportunidad relativa. La oportunidad relativa indica la relación entre dos proporciones: la proporción de individuos en el grupo de casos que poseen un alelo específico y la proporción de individuos en el grupo control que poseen el mismo alelo. Si la frecuencia alélica en el grupo de casos es significativamente mayor que la frecuencia alélica en el grupo control, la oportunidad relativa será mayor que 1. Si la frecuencia alélica en el grupo de casos es significativamente menor que la frecuencia alélica en el grupo control, la oportunidad relativa será menor que 1.

45 Un método preferido para combinar la información procedente de múltiples fuentes ha sido descrito en el informe público "Polygenic Risk Score Improves Prostate Cancer Risk Prediction: Results from the Stockholm-1 Cohort Study", de Markus Aly y coautores, publicado en *EUROPEAN UROLOGY*, 60 (2011), 21-28. Se evaluaron las asociaciones entre cada SNP y el PCa en una biopsia empleando un ensayo de tendencias de Cochran-Armitage. Se calculó la oportunidad relativa (OR) alélica con un intervalo de confianza del 95% empleando modelos de regresión logística. Se creó para cada paciente una puntuación de riesgo genética sumando el número de alelos de riesgo (0, 1 o 2) en cada uno de los SNP multiplicado por el logaritmo de la OR de ese SNP. Se exploraron las asociaciones entre el diagnóstico de PCa y los factores de riesgo evaluados en un análisis de regresión logística. La porción del modelo relacionada con la información no genética incluía el PSA total transformado logarítmicamente, la proporción de PSA libre al total transformada logarítmicamente, la edad en el momento de la biopsia y la historia familiar de PCa (sí o no). Se empleó una validación cruzada en 10 veces repetida para calcular las probabilidades predichas de PCa en la biopsia. Se construyeron intervalos de confianza del 95% para los valores de ROC-AUC empleando una aproximación normal. Todos los valores de p indicados se basan en hipótesis bilaterales. Aunque el método de Aly y coautores se describe en el contexto de la selección de PCa, puede aplicarse la misma estrategia para los métodos de pronóstico.

60 En la mayoría de los casos, el cáncer de próstata es una enfermedad de avance lento. Así, la capacidad para recomendar a un individuo la terapia activa, ya antes de realizar una biopsia, permite, por ejemplo, motivar al individuo para que cambie su estilo de vida en preparación para la terapia activa, si esta fuera necesaria. Dejar de

fumar, alcanzar un valor de BMI menor que 30 y realizar ejercicio de modo habitual (aproximadamente 30 minutos 3-6 días a la semana) son todos factores que, en general, estimulan la supervivencia en condiciones de enfermedad grave, que incluyen el cáncer de próstata. Por tanto, si se descubre que un individuo es adecuado para recibir una terapia activa, resulta razonable sugerir a dicho individuo que deje de fumar, que intente alcanzar un BMI < 30 y que comience a realizar ejercicio para soportar mejor los efectos secundarios de la terapia y para aumentar las probabilidades de recuperarse de la enfermedad. Otro aspecto importante son las cuestiones dietéticas. A través de un cambio en la dieta puede reducirse o retrasarse el desarrollo del PCa. Existen pruebas que sugieren que un menor consumo de alimentos lácteos puede reducir el riesgo de aparición del PCa, tal como indican Song y coautores en la publicación "Whole milk intake is associated with prostate cancer-specific mortality among U.S. male physicians", publicado en J. Nutr., febrero de 2013, 143(2):189-96. Existen pruebas similares de los efectos positivos del consumo de té verde y el consumo de productos de la soja. Por tanto, si se descubre que un individuo es adecuado para recibir una terapia activa, resulta razonable sugerir a dicho individuo que disminuya el consumo de productos lácteos y/o que aumente el consumo de té verde y productos basados en la soja para retrasar o incluso prevenir someterse a una terapia activa.

15 Ejemplo 1

Para ilustrar la presente invención, se extrajo un conjunto de datos que comprende 172 casos (sujetos que se sabe que padecen PCa y han recibido un tratamiento activo) y 79 controles (sujetos que se sabe que padecen PCa para los cuales se ha decidido una terapia expectante) a partir del conjunto de datos de STHLM2. El conjunto de datos de STHLM2 ha sido analizado en el dominio público como evidente en la página web <http://sthlm2.se/>. En resumen, durante 2010-2012, aproximadamente 26000 hombres que se sometieron a un ensayo de PSA en el área de Estocolmo se incluyeron en el estudio de STHLM2. Los 172 + 79 = 251 sujetos se caracterizaron con respecto a los siguientes biomarcadores y SNP:

Biomarcadores:

Antígeno específico de próstata total (PSAt) [ng/ml]

25 Antígeno específico de próstata intacto (PSAi) [ng/ml]

Antígeno específico de próstata libre (PSAl) [ng/ml]

Calicreína humana 2 (hK2) [ng/ml]

Citoquina inhibidora de macrófagos 1 (MIC-1) [ng/ml]

Beta-microseminoproteína (MSMB) [ng/ml]

30 SNP:

657de15, rs10086908, rs1016343, rs10187424, rs1041449, rs10486567, rs1054564, rs10875943, rs10896449, rs10934853, rs10993994, rs11067228, rs11135910, rs11228565, rs11568818, rs11649743, rs11650494, rs11672691, rs11704416, rs12130132, rs12409639, rs12418451, rs12500426, rs12543663, rs12621278, rs12653946, rs1270884, rs130067, rs13252298, rs13385191, rs1354774, rs1363120, rs137853007, rs138213197, rs1447295, rs1465618, rs1512268, rs1571801, rs16901979, rs16902094, rs17021918, rs17632542, rs17879961, rs1859962, rs1894292, rs1933488, rs1983891, rs2018334, rs2121875, rs2242652, rs2273669, rs2292884, rs2405942, rs2660753, rs2735839, rs2736098, rs2928679, rs3213764, rs339331, rs3771570, rs3850699, rs3863641, rs401681, rs4245739, rs4430796, rs445114, rs4643253, rs4857841, rs4962416, rs5759167, rs5919432, rs5945619, rs6062509, rs620861, rs6465657, rs6763931, rs684232, rs6869841, rs6983267, rs6983561, rs7127900, rs7210100, rs721048, rs7241993, rs7611694, rs7679673, rs7931342, rs8008270, rs8102476, rs888663, rs902774, rs9364554, rs9600079, rs9623117.

Se recopiló la información básica para cada sujeto, que incluye la edad y la historia familiar (sí o no). La edad se expresa en unidades de años.

45 Para predecir los sujetos a los que se les recomendaría una terapia activa, es posible basarse solo en la puntuación genética, tal como se ilustra en la siguiente ecuación predeterminada:

$$y = 0,63 + 0,039 * \text{puntuación}$$

En esta ecuación, 'puntuación' es la variable de puntuación genética calculada según se describe en el informe público "Polygenic Risk Score Improves Prostate Cancer Risk Prediction: Results from the Stockholm-1 Cohort Study", de Markus Aly y coautores, publicado en EUROPEAN UROLOGY, 60 (2011), 21-28, que contiene los SNP de susceptibilidad al cáncer de próstata validados (dichos SNP están relacionados con la susceptibilidad al cáncer de próstata o están relacionados con los niveles plasmáticos de los biomarcadores PSA, PSA libre, MSMB y MIC-1) listados en el presente ejemplo.

El valor 'y' resultante se correlaciona en gran medida con la necesidad de una terapia activa, tal como se ilustra en la

figura 1. Las curvas de ROC en la figura 1 representan la predicción de la necesidad de una terapia activa empleando solo el valor de PSA (101; línea continua) y el modelo descrito en este ejemplo (102; línea discontinua). Si y es mayor que un valor de corte, se recomienda al hombre que consulte a un urólogo para el examen de la próstata con biopsias.

- 5 El valor de corte depende del compromiso entre la sensibilidad y la especificidad del ensayo. Si, por ejemplo, se emplea un valor de corte de 1,04, este ensayo concreto producirá una sensibilidad del ensayo de 0,5 y una especificidad de 0,67. Esto puede compararse empleando solo el valor de PSA como ensayo de selección, que produce una sensibilidad de 0,5 y una especificidad de 0,46. Es importante advertir que, para este método de pronóstico concreto, la actuación más valiosa es apoyar la decisión de los individuos en la zona gris, es decir, a un nivel de sensibilidad de 0,5-0,6.

Ejemplo 2

- 15 Para ilustrar aún más la presente invención, el conjunto de datos descritos en el ejemplo 1 se sometió a otros análisis. Se combinó la información procedente de varias categorías, que incluyen la puntuación genética y las concentraciones de biomarcadores, para formar un modelo multiparamétrico que incluye los anteriores biomarcadores y la puntuación genética según se describe en el ejemplo 1. El modelo se derivó empleando una regresión lineal.

- 20 El valor 'y' resultante de este modelo se correlaciona en gran medida con la necesidad de una terapia activa, tal como se ilustra en la figura 2. Las curvas de ROC en la figura 2 representan la predicción de la necesidad de una terapia activa empleando solo el valor de PSA (201) y el modelo descrito en este ejemplo (202). Si y es mayor que un valor de corte, se recomienda al hombre que consulte a un urólogo para el examen de la próstata con biopsias.

- 25 El valor de corte depende del compromiso entre la sensibilidad y la especificidad del ensayo. Si, por ejemplo, se emplea un valor de corte de 0,72, este ensayo concreto producirá una sensibilidad del ensayo de 0,5 y una especificidad de 0,78. Esto puede compararse empleando solo el valor de PSA como ensayo de selección, que produce una sensibilidad de 0,5 y una especificidad de 0,46. Es importante advertir que, para este método de pronóstico concreto, la actuación más valiosa es apoyar la decisión de los individuos en la zona gris, es decir, a un nivel de sensibilidad de 0,5-0,6.

Ejemplo 3

- 30 Para ilustrar aún más la presente invención, se extrajo un subconjunto del conjunto de datos del ejemplo 1 omitiendo todos los individuos con PSA > 7 ng/ml o, en el caso de que falte un valor de PS, se omitió a los individuos con PSA libre > 0,91 ng/ml.

El subconjunto contenía 144 individuos. Se derivaron cuatro modelos diferentes para intentar predecir si el individuo recibió un tratamiento activo.

$$Y1 = 1,4076640 + 0,0352188 * PSA - 0,0159339 * edad - 0,0005042 * PSA * edad$$

- 35 $Y2 = 6,420955 - 0,897462 * puntuación - 0,445241 * PSA - 0,067727 * edad + 0,081441 * puntuación * PSA + 0,009316 * puntuación * edad + 0,000848 * PSA * edad$

$$Y3 = 3,901443 - 0,408785 * puntuación90 - 0,262197 * PSA - 0,040232 * edad + 0,045336 * puntuación90 * PSA + 0,003840 * puntuación90 * edad + 0,001104 * PSA * edad$$

- 40 $Y4 = 3,503319 - 0,337091 * puntuación - 2,747459 * PSAI - 0,042007 * edad + 1,243013 * intacto + 0,242149 * puntuación * PSAI - 0,070408 * puntuación * intacto + 0,019290 * PSAI * edad + 0,205893 * PSAI * intacto - 0,014322 * edad * intacto$

en los que PSA es la concentración de PSA, PSAI es la concentración de PSA libre, intacto es la concentración de PSA intacto, edad es la edad del individuo, puntuación es la puntuación genética, y puntuación90 es la puntuación genética calculada empleando solo 90% (seleccionados aleatoriamente) de los SNP. El modelo Y1 tiene un ROC-AUC = 0,60, el modelo Y2 tiene un ROC-AUC = 0,69, Y3 tiene un ROC-AUC = 0,65, e Y4 tiene un ROC-AUC = 0,68.

- 45 Este ejemplo ilustra cuatro aspectos diferentes de la invención:

Un primer aspecto es la capacidad para predecir si un paciente en efecto recibió un beneficio del tratamiento activo a partir de la inclusión de la puntuación genética (que se observa en la comparación de los modelos Y1 e Y2).

- 50 Un segundo aspecto es que el modelo Y2 tiene una redundancia inherente en el parámetro de puntuación y realizará una buena actuación en el modelo Y3 en comparación con el modelo Y1 (que no emplea información genética) aunque falta 10% de la información de los SNP en Y3. Los datos que faltan pueden ser debidos a una multitud de razones, tales como problemas técnicos, la falta de material de muestra o el error humano, por mencionar algunos ejemplos no limitantes.

Un tercer aspecto es que el biomarcador de proteína PSA (que pertenece a la familia de biomarcadores similares a la calicreína) puede sustituirse por otros miembros de la familia de biomarcadores de calicreína (ilustrado en el modelo Y4) conservando una buena actuación, comparado con el modelo Y1, más simple. Nótese que el criterio de inclusión de los individuos que carecen de información acerca del valor de PSA es PSA libre < 0,91.

- 5 Un cuarto aspecto es que este ejemplo ilustra que es posible proporcionar información a la categoría de individuos en los que la decisión de realizar el tratamiento y cuándo realizarlo resulta difícil. Mientras que los individuos con un elevado valor de PSA (mayor que 7 ng/ml, o incluso mayor que 30 ng/ml) habitualmente son candidatos para una terapia activa, los individuos con un valor más bajo de PSA están en una zona gris en la que la proporción de riesgo a beneficio no siempre es fácil de determinar.
- 10 Aunque la invención se ha descrito con respecto a su realización preferida, que constituye el mejor modo conocido en la actualidad por el inventor, debe entenderse que pueden realizarse diversos cambios y modificaciones, tal como será obvio para los expertos en la técnica, sin apartarse del alcance de la invención según se define en las reivindicaciones adjuntas.

REIVINDICACIONES

1.- Un método basado en una combinación de datos diseñada de modo redundante para valorar si un individuo diagnosticado con cáncer de próstata requerirá una terapia activa, que comprende las etapas de:

5 1. En al menos una muestra biológica procedente de dicho individuo, analizar una categoría de SNP relacionados con el PCa (SNPpc) midiendo la presencia o la ausencia de uno o dos alelos de riesgo en cada uno de una pluralidad de SNPpc;

10 2. Combinar los datos relativos a dicha categoría de SNPpc para formar un valor compuesto de SNPpc, en el que el método permite desestimar los datos de un subconjunto de al menos 10% de los SNPpc de la categoría de SNPpc cuando se forma el valor compuesto de SNPpc, en el que el valor compuesto de SNPpc se basa en una oportunidad relativa ("odds ratio") predeterminada para cada SNP individual incluido en la categoría de SNPpc, y en el que el valor compuesto de SNPpc se forma a partir de los datos de al menos 90 SNPpc;

15 3. Correlacionar dicho valor compuesto de SNPpc con la probabilidad de que un individuo requiera de terapia activa comparando el valor compuesto de SNPpc con un valor de corte predeterminado establecido con muestras control, en el que se sabe si los individuos de los cuales se obtienen las muestras control requirieron una terapia activa o no requirieron una terapia activa.

20 2.- El método de la reivindicación 1, en el que el SNPpc en la etapa 1 consiste en 657de15, rs10086908, rs1016343, rs10187424, rs1041449, rs10486567, rs1054564, rs10875943, rs10896449, rs10934853, rs10993994, rs11067228, rs11135910, rs11228565, rs11568818, rs11649743, rs11650494, rs11672691, rs11704416, rs12130132, rs12409639, rs12418451, rs12500426, rs12543663, rs12621278, rs12653946, rs1270884, rs130067, rs13252298, rs13385191, rs1354774, rs1363120, rs137853007, rs138213197, rs1447295, rs1465618, rs1512268, rs1571801, rs16901979, rs16902094, rs17021918, rs17632542, rs17879961, rs1859962, rs1894292, rs1933488, rs1983891, rs2018334, rs2121875, rs2242652, rs2273669, rs2292884, rs2405942, rs2660753, rs2735839, rs2736098, rs2928679, rs3213764, rs339331, rs3771570, rs3850699, rs3863641, rs401681, rs4245739, rs4430796, rs445114, rs4643253, rs4857841, rs4962416, rs5759167, rs5919432, rs5945619, rs6062509, rs620861, rs6465657, rs6763931, rs684232, rs6869841, rs6983267, rs6983561, rs7127900, rs7210100, rs721048, rs7241993, rs7611694, rs7679673, rs7931342, rs8008270, rs8102476, rs888663, rs902774, rs9364554, rs9600079, y rs9623117.

30 3.- El método de la reivindicación 1 o 2, que comprende además analizar, en dicha muestra biológica, una categoría de biomarcadores del PCa midiendo la presencia o la concentración de cada uno de una pluralidad de biomarcadores del PCa de dicha categoría de biomarcadores del PCa; combinar los datos relativos a dicha categoría de biomarcadores del PCa para formar un valor compuesto de biomarcador; combinar el valor compuesto de biomarcador y el valor compuesto de SNPpc para formar un valor compuesto global; y correlacionar dicho valor compuesto global con la probabilidad de que dicho individuo requiera de terapia activa comparando el valor compuesto global con un valor predeterminado establecido con muestras control, en el que se sabe si los individuos de los cuales se obtienen las muestras control requirieron una terapia activa o no requirieron una terapia activa.

35 4.- El método de la reivindicación 3, que comprende medir la presencia o la concentración de biomarcadores del PCa al menos parcialmente redundantes, y en el que al menos uno, tal como dos, de los biomarcadores del PCa se selecciona del grupo que consiste en (i) PSA, (ii) PSA total (PSAt), (iii) PSA intacto (PSAi), (iv) PSA libre (PSAI), y (v) hK2.

40 5.- El método de la reivindicación 4, en el que el método permite desestimar los datos de un subconjunto de al menos uno de dichos biomarcadores del PCa (i)-(v) de la categoría de biomarcadores del PCa cuando se forma dicho valor compuesto de biomarcador, tal como un subconjunto de uno, dos, tres o cuatro de dichos biomarcadores del PCa (i)-(v).

45 6.- El método de una cualquiera de las reivindicaciones anteriores, en el que dichos datos relativos a dicha categoría de SNPpc se combinan según una ecuación predeterminada para formar dicho valor compuesto de SNPpc.

7.- El método de una cualquiera de las reivindicaciones 3-6, en el que dichos datos relativos a dicha categoría de biomarcadores del PCa se combinan según una ecuación predeterminada para formar dicho valor compuesto de biomarcador.

50 8.- El método de una cualquiera de las reivindicaciones 3-7, en el que dicho valor compuesto de biomarcador y dicho valor compuesto de SNPpc se combinan según una ecuación predeterminada para formar dicho valor compuesto global.

9.- El método de una cualquiera de las reivindicaciones anteriores, que comprende además recomendar al individuo para que reciba una terapia activa si el valor compuesto de SNPpc o el valor compuesto global es mayor que el valor de corte.

55 10.- El método de una cualquiera de las reivindicaciones anteriores, que comprende además recopilar la historia

familiar con respecto al PCa, la historia de tratamiento y los datos físicos de dicho individuo; y en el que dicha historia familiar, historia de tratamiento y/o datos físicos se incluyen en el valor compuesto global.

5 11.- El método de una cualquiera de las reivindicaciones 3-10, que comprende además analizar una categoría adicional de biomarcadores del PCa midiendo la presencia o la concentración de uno o cada uno de una pluralidad de biomarcadores del PCa de dicha categoría adicional de biomarcadores; combinar los datos relativos a dicha categoría adicional de biomarcadores del PCa para formar un valor compuesto de biomarcador adicional para dicha categoría adicional de biomarcadores del PCa; e incluir dicho valor compuesto de biomarcador adicional en un valor compuesto global; en el que la combinación de datos para formar el valor compuesto de biomarcador adicional se diseña de modo redundante cuando la categoría adicional de biomarcadores del PCa comprende más de un biomarcador del PCa, y en el que la categoría adicional de biomarcadores del PCa comprende el biomarcador MIC-1 y opcionalmente otros biomarcadores relacionados con MIC-1, o el biomarcador MSMB y opcionalmente otros biomarcadores relacionados con MSMB.

10 12.- El método de una cualquiera de las reivindicaciones anteriores, en el que la medición de la presencia o la concentración de dichos biomarcadores del PCa se realiza utilizando la tecnología de micromatrices.

15 13.- Un producto de programa informático que puede cargarse directamente en la memoria interna de un ordenador digital, que se caracteriza porque dicho producto comprende un medio de código de software para realizar al menos las etapas 2 y 3 de la reivindicación 1, tal como las etapas 1-3 de la reivindicación 1.

20 14.- El producto de programa informático de la reivindicación 14, que comprende además un medio de código de software para realizar el método de la reivindicación 3.

FIGURA 1

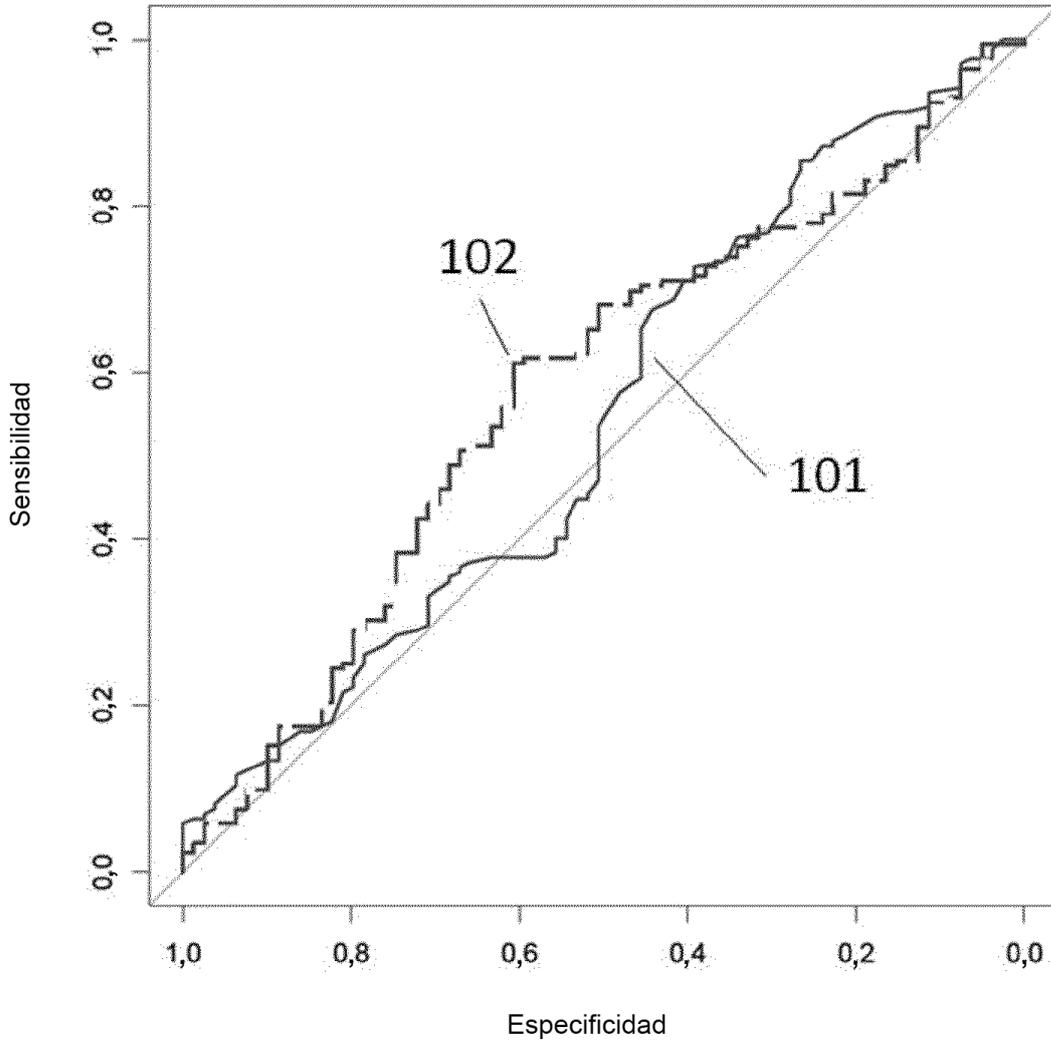


FIGURA 2

