(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 550 192**

(51) Int. CI.:

C12Q 1/68 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **24.05.2012 E 12723210 (6)**

(97) Fecha y número de publicación de la concesión europea: **22.07.2015 EP 2714926**

(54) Título: **Biomarcadores para el cáncer de pulmón**

(30) Prioridad:

25.05.2011 US 201161490021 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
05.11.2015

(73) Titular/es:

NOVARTIS AG (100.0%)
Lichtstrasse 35
4056 Basel, CH

(72) Inventor/es:

MISSIAGLIA, EDOARDO;
WIRAPATI, PRATYAKSHA;
ROSSI, SIMONA y
KROLL, WERNER

(74) Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

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

Biomarcadores para el cáncer de pulmón

Campo de la invención

La presente invención se refiere a un método de prognosis y terapia personalizada.

Antecedentes de la invención

El cáncer de pulmón es el diagnóstico de cáncer más común en el mundo con 1,5 millones de nuevos casos en 2007 (Salomon et al., Crit Rev Oncol Hematol., 19:183-232, 1995, base de datos SEER 05.2010). La elevada incidencia y las tasas de mortalidad hacen de él la cauda principal de muerte relacionada con el cáncer con más de 975.000 muertes al año y una tasa de supervivencia a los 5 años de 15% (Salomon et al., *supra*). El cáncer de pulmón se puede clasificar como cáncer de pulmón de células pequeñas, o cáncer de pulmón de células no pequeñas (CPCNP). El CPCNP representa aproximadamente 80% de los casos. El CPCNP se puede subdividir adicionalmente en varios tipos histológicos, los más comunes son el adenocarcinoma (40%) y el carcinoma de células escamosas (25%).

El tratamiento actual del CPCNP se basa principalmente en la morfología del tumor y el sistema de estadificación basado en tumor-nódulo-metástasis (TNM) que clasifica los tumores en categorías graduadas (Estadío IA, IB, IIA; IIB, IIIA, IIIB y IV) correspondientes al grado de progreso del tumor. Existen muchos sistemas de estadificación (p. ej., estadificación clínica versus patológica, así como diversas ediciones de pautas de estadificación tales como la expedida por la International Association for the Study of Lung Cancer (IASLC) (Goldstraw et al, J. Thorac. Oncol. 2, 706-714, 2007). La frecuencia de los estadíos, por ejemplo de acuerdo con la estadificación clínica y la 6ª edición IASLC de las recomendaciones de estadificación TNM, son 23% de estadío I, 19% estadío II, 37% estadío III y 21% estadío IV (Goldstraw et al. *supra*). Las proporciones relativas pueden cambiar sustancialmente dependiendo de las pautas y de si se utiliza la estadificación clínica o patológica.

La cirugía es el tratamiento convencional para los CPCNP de estadío temprano (estadío I y II), seguido de terapia coadyuvante tal como terapia de radiación, quimioterapia (para el estadío II y posterior), y bevacizumab e inhibidores de la tirosina quinasa (TKI) del receptor del factor de crecimiento epidérmico (EGFR) para los estadíos avanzados del CPCNP (Kutikova et al., Lung Cancer; 50(2):143-154, 2005). Las pautas clínicas en los Estados Unidos y Europa para el tratamiento del CPCNP apoyan estas opciones de tratamiento (Mendelssohn et al., 2000; NCCN-Guidelines NSCLC V1, 2010).

Basándose en el sistema de estadificación basado en TNM actual para el estadío I del cáncer de pulmón, los pacientes con CPCNP tienen 35% de posibilidades de recaída en el plazo de 5 años después de la cirugía (SEER-Database, 2008). Las pautas de tratamiento actuales no recomiendan una quimioterapia coadyuvante para estos pacientes. Mientras 30% de los pacientes con un estadío II basado en TNM no experimentará una recaída sin ninguna quimioterapia coadyuvante (SEER-Database) lo que significa que estos pacientes sufren un tratamiento excesivo basado en las actuales pautas de tratamiento (esto es, ESMO (D'Addario et al., Annals of Oncology. 2009; 20 (supl 4):iv68-iv70, 2009), NCCN VI-2010). Esto coincide con los informes que establecen que 60% de los pacientes con CPCNP temprano no tendrán recaída después de la cirugía (Arriagada et al., NEJM, 350:351-360, 2004). Basándose en los datos clínicos actuales, el tratamiento con quimioterapia coadyuvante del CPCNP temprano proporciona evidencia de que el beneficio medio para la terapia coadyuvante es de 4% (NSCLC Meta-Analysis Collaborative Group), mejorando de 60% a 64 % a los 5 años.

Basándose en los defectos actuales, existe una justificación médica para la necesidad de una firma genómica de prognóstico y/o predicción para pacientes con CPCNP.

Compendio de la invención

La presente invención se refiere, en parte, a métodos para determinar un pronóstico de cáncer de pulmón de estadificación temprana en un individuo utilizando uno o más biomarcadores descritos en la presente memoria. Estos descubrimientos se pueden utilizar para ayudar a determinar los tratamientos apropiados para los pacientes con cáncer de pulmón en estadío temprano, por ejemplo la identificación de aquellos pacientes que se beneficiarían de recibir la terapia coadyuvante.

En un aspecto, la invención incluye un método para pronosticar o clasificar a un sujeto con cáncer de pulmón de células no pequeñas (CPCNP) que incluye la obtención de una muestra de ensayo de un sujeto que padece CPCNP después de la resección quirúrgica; la determinación del nivel de expresión de al menos uno o más biomarcadores identificados en la Tabla 1, la Tabla 2 y/o la Tabla 3, o cualquier combinación de biomarcadores identificados en la Tabla 1, la Tabla 2 y/o la Tabla 3 en la muestra de ensayo; y el análisis del nivel de expresión para generar una puntuación del riesgo, en donde la puntuación del riesgo se puede utilizar para proporcionar un pronóstico o clasificar al sujeto.

En otro aspecto, la invención incluye un método para pronosticar o clasificar a un sujeto con cáncer de pulmón de células no pequeñas (CPCNP) que comprende: obtener una muestra de ensayo de un sujeto que padece CPCNP después de la resección quirúrgica; determinar el nivel de expresión de al menos un biomarcador a partir de la Tabla 1, la Tabla 2 y la Tabla 3 en la muestra de ensayo; y analizar el nivel de expresión para generar una puntuación del riesgo, en donde la puntuación del riesgo se puede utilizar para proporcionar un pronóstico o clasificar al sujeto. En una realización, el al menos un biomarcador identificado en la Tabla 1, la Tabla 2 y la Tabla 3 incluye CBX7, STX1A, y TPX2. En otra realización, el al menos un biomarcador identificado en la Tabla 1, la Tabla 2 y la Tabla 3 incluye CBX7, TMPRSS2, STX1A, KLK6, TPX2 y UCK. En otra realización más, el al menos un biomarcador identificado en la Tabla 1, la Tabla 2 y la Tabla 3 incluye CBX7, TMPRSS2, GPR116, STX1A, KLK6, SLC16A3, TPX2, UCK2, PHKA1. En otra realización adicional, el al menos un biomarcador identificado en la Tabla 1, la Tabla 2 y la Tabla 3 comprende CBX7, TMPRSS2, GPR116, KCNJ15, STX1A, KLK6, SLC16A3, PYGL, TPX2, UCK2, PHKA1, o EIF4A3. En otra realización más, el al menos un biomarcador identificado en la Tabla 1, la Tabla 2 y la Tabla 3 incluye CBX7, TMPRSS2, GPR116, KCNJ15, PTPN13, STX1A, KLK6, SLC16A3, PYGL, LDHA, TPX2, UCK2, PHKA1, EIF4A3 o TK1. En otra realización más, el al menos un biomarcador identificado en la Tabla 1, la Tabla 2 y la Tabla 3 comprende CBX7, TMPRSS2, GPR116, KCNJ15, PTPN13, CTSH, STX1A, KLK6, SLC16A3, PYGL, LDHA, ITGA5, TPX2, UCK2, PHKA1, EIF4A3, TK1, o CCNA2.

En una realización, la puntuación de riesgo de la invención se puede utilizar para el pronóstico mediante el mapeo de sujetos para determinar la probabilidad de muerte específica del tiempo debida a cáncer de pulmón, metástasis distante o recaída local.

En otra realización, la puntuación de riesgo puede clasificar al sujeto en un grupo de alto riesgo que se beneficiaría de recibir quimioterapia coadyuvante o en un grupo de riesgo bajo que no se beneficiaría de recibir quimioterapia coadyuvante.

En otro aspecto, la invención incluye un método de pronóstico predictivo en un sujeto con cáncer de pulmón de células no pequeñas (CPCNP) después de la resección quirúrgica, que comprende determinar el perfil de expresión del ARNm de muestras tumorales, o bien de material recién congelado (FF) o incluido en parafina fijado con formalina (FFPE). El perfil comprende uno o más biomarcadores enumerados en la Tabla 1, la Tabla 2 y/o la Tabla 3, en donde se utiliza un incremento en la expresión de uno o más biomarcadores enumerados en la Tabla 2 y/o la Tabla 3 y una disminución en la expresión de uno o más biomarcadores enumerados en la Tabla 1 en comparación con un control para pronosticar si el sujeto está en un grupo de alto riesgo que tiene una escasa supervivencia o en un grupo de bajo riesgo que tiene una buena supervivencia. En el método de la invención, se selecciona un sujeto del grupo de alto riesgo para la quimioterapia coadyuvante y el sujeto del grupo de bajo riesgo no se selecciona para la quimioterapia coadyuvante y a continuación se trataron en consecuencia.

En otra realización más, la invención incluye un método de selección de una terapia para un sujeto con CPCNP, que incluye obtener una muestra de ensayo de un sujeto que padece CPCNP que ha sufrido una resección; determinar el nivel de expresión de al menos dos o más biomarcadores identificados en la Tabla 2 en la muestra de ensayo para generar un valor de expresión para cada gen; y analizar el valor de expresión para generar una puntuación de riesgo, en donde la puntuación de riesgo se puede utilizar para clasificar si el sujeto se selecciona para recibir un inhibidor de la angiogénesis tal como la avastina.

En los métodos de la invención, la invención incluye la determinación de la expresión de al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9, al menos10, al menos 11, al menos 12, al menos 13, al menos 14 o al menos 15 biomarcadores identificados en la Tabla 1, la Tabla 2 y/o la Tabla 3, o cualquier combinación de los mismos. Por ejemplo, se selecciona la expresión de al menos 5 biomarcadores cualesquiera de cada una de la Tabla 1, la Tabla 2 y la Tabla 3 (la firma en esta realización incluiría al menos 15 biomarcadores).

En una realización, el CPCNP es CPCNP en estadío I, CPCNP en estadío II, o una combinación de los mismos. El CPCNP puede ser identificado en el grupo que consiste en carcinoma de células escamosas y/o adenocarcinoma.

En una realización, el sujeto es un ser humano. En otra realización, la muestra de ensayo pueden ser células frescas, congeladas, FFPE. En otra realización, el nivel de expresión se determina utilizando una PCR cuantitativa o una matriz.

En otra realización más, el análisis de la expresión para generar una puntuación de riesgo se realiza utilizando el análisis estadístico tal como la regresión de Cox o predictores de supervivencia paramétricos.

En otro aspecto, la invención incluye un método de tratamiento selectivo de un sujeto que tiene cáncer CPCNP que incluye obtener una muestra de ensayo de un sujeto que padece CPCNP después de la resección quirúrgica, determinar el nivel de expresión de al menos uno o más biomarcadores identificados en la Tabla 1, la Tabla 2 y/o la Tabla 3, o cualquier combinación de biomarcadores identificados en la Tabla 1, la Tabla 2 y/o la Tabla 3 en la muestra de ensayo para generar una puntuación de riesgo; clasificar el sujeto basándose en la puntuación de riesgo en un grupo de alto riesgo o un grupo de bajo riesgo; y administrar terapia coadyuvante al sujeto clasificado

perteneciente al grupo de alto riesgo o no administrar terapia coadyuvante al sujeto clasificado perteneciente al grupo de bajo riesgo.

En otro aspecto más, la invención incluye un kit que incluye una pluralidad de agentes para medir la expresión de uno o más biomarcadores identificados en la Tabla 1, la Tabla 2 y/o la Tabla 3 e instrucciones para su uso. En otro aspecto más, la invención incluye un kit para pronosticar si un sujeto con cáncer de pulmón se beneficiaría de la terapia coadyuvante, el kit incluye una pluralidad de agentes para medir la expresión de uno o más biomarcadores identificados en la Tabla 1, la Tabla 2 y/o la Tabla 3; medios para analizar la expresión y generar una puntuación de riesgo para predecir si un paciente se beneficiaría de la terapia coadyuvante. Los agentes para medir la expresión pueden incluir una matriz de polinucleótidos complementarios a los ARNm de los biomarcadores identificados. Los agentes que miden la expresión pueden incluir una pluralidad de sondas de PCR y/o cebadores para qRT-PCR. El kit puede incluir agentes para la medición de al menos un biomarcador identificado en la Tabla 1, la Tabla 2 y la Tabla 3 tal como CBX7, STX1A, o TPX2.

En otro aspecto, la invención incluye una matriz que comprende una o más sondas polinucleotídicas complementarias e hibridables con un producto de expresión de al menos dos biomarcadores etc. mostrados en la Tabla 1, la Tabla 2 y/o la Tabla 3.

En otro aspecto más, la invención incluye una composición que comprende una pluralidad de secuencias de ácido nucleico aisladas, en donde cada secuencia de ácido nucleico aislada hibrida con un producto de ARN de un biomarcador mostrado en la Tabla 1, p. ej., los biomarcadores CBX7, STX1A y TPX2, en donde la composición se utiliza para medir el nivel de expresión de ARN de los tres genes.

En otro aspecto más, la invención incluye un producto informático para predecir un pronóstico, o clasificar a un sujeto con CPCNP incluyendo los medios para recibir los datos correspondientes al nivel de expresión de uno o más biomarcadores en una muestra de un sujeto que tiene CPCNP, en donde los uno o más biomarcadores se identifican en la Tabla 1, la Tabla 2 y/o la Tabla 3, medios para generar un valor de expresión para cada gen; y medios para generar una puntuación de riesgo basada en la introducción del valor de expresión en una base de datos que comprende un perfil de expresión de referencia asociado con un pronóstico, en donde la puntuación de riesgo predice un pronóstico de supervivencia o clasifica al sujeto en un grupo de alto riesgo o un grupo de bajo riesgo.

En otro aspecto más, la invención incluye un producto informático para su uso con el método de uno cualquiera de los métodos descritos anteriormente.

Un "biomarcador" es una molécula útil como indicador de un estado biológico en un sujeto. Con referencia a la presente materia sujeto, los biomarcadores descritos en la presente memoria pueden ser moléculas que muestran un cambio en la expresión y cuya presencia se puede utilizar para pronosticar o predecir si un sujeto se beneficiaría de recibir un tratamiento concreto. Los biomarcadores de interés se pueden determinar detectando un cambio en la expresión del biomarcador. Un cambio en la expresión describe la conversión de la secuencia del gen de ADN en un ARN transcrito (el transcrito de ARN no unido inicial o el ARN maduro) o el producto de proteína codificado. Los biomarcadores descritos en la presente memoria incluyen cualquiera, o cualquier combinación de los biomarcadores enumerados en las Tablas 1, 2 y 3 y pueden ser ARN transcrito o un producto de proteína codificado.

Breve descripción de los dibujos

La **Fig. 1** representa gráficos de dispersión por pares de las puntuaciones normalizadas de los módulos 1, 2 y 3, que demuestran que éstos expresan diferente información acerca del tumor.

Las **Figs. 2A-D** representan curvas de análisis de la supervivencia global de Kaplan-Meier que muestran el rendimiento pronóstico de las tres firmas (módulo 1, 2, y 3) cuando se aplican por separado, o combinadas, para pacientes hasta cinco años después de la cirugía.

Las **Figs. 3A-F** representan curvas de análisis de supervivencia global de Kaplan-Meier que muestran el rendimiento pronóstico del módulo 1 bajo estratificación por el estadío, la histología y la edad, para pacientes hasta cinco años después de la cirugía.

Las **Figs. 4A-F** representan curvas de análisis de supervivencia de Kaplan-Meier que muestran el rendimiento pronóstico del módulo 2 bajo estratificación por el estadío, la histología y la edad, para pacientes hasta cinco años después de la cirugía.

Las **Figs. 5A-F** representan curvas de análisis de supervivencia de Kaplan-Meier que muestran el rendimiento pronóstico del módulo 3 bajo estratificación por el estadío, la histología y la edad, para pacientes hasta cinco años después de la cirugía.

Las **Figs. 6A-F** representan curvas de análisis de supervivencia de Kaplan-Meier que muestran el rendimiento pronóstico de la puntuación combinada bajo estratificación por el estadío, la histología y la edad, para pacientes hasta cinco años después de la cirugía.

Las **Figs. 7A-F** representan curvas de análisis de supervivencia de Kaplan-Meier que muestran el rendimiento pronóstico de numerosas firmas ilustrativas construidas como subgrupos de la lista de 38 genes.

La **Fig. 8A** representa el porcentaje de mortalidad del cáncer de pulmón (equivalente a 100% menos supervivencia) como una función de la puntuación de riesgo de la firma reivindicada y muestra semejante pronóstico a un tiempo de seguimiento de 1, 2, 3, 4 y 5 años. La **Fig. 8B** representa el porcentaje de mortalidad del cáncer de pulmón (equivalente a 100% menos supervivencia) como una función de la puntuación de riesgo de la firma reivindicada y muestra la mortalidad en 5 años como una función de la puntuación de riesgo, estratificada por el estadío 1 y 2 del tumor.

La **Fig. 9** representa una curva de análisis de supervivencia de Kaplan-Meier que muestra el rendimiento pronóstico en conjuntos de datos disponibles públicamente de M. D. Anderson Cancer Center que se obtuvieron a partir de material FFPE.

La **Fig. 10A-B** representa gráficos de dispersión por pares que comparan FF/Affymetrix vs FFPE/qNPA, así como FF/Affymetrix vs FFPE/nanoString.

Descripción detallada de la invención

La presente invención se basa en parte, en métodos que se pueden utilizar para el pronóstico o la clasificación de individuos que tienen cáncer de pulmón en estadío temprano. La invención incluye adicionalmente la identificación de aquellos pacientes que tienen un riesgo elevado de recurrencia de la enfermedad y para los cuales se podría recomendar una terapia coadyuvante, así como pacientes con un bajo riesgo de recurrencia, que no se beneficiarían de la terapia coadyuvante. En un ejemplo, los métodos de pronóstico y predicción descritos en la presente memoria se basan en la expresión diferencial de una pluralidad de biomarcadores en una muestra de ensayo de cáncer de pulmón. Los biomarcadores de la invención pueden incluir 38 genes (CBX7, TMPRSS2, GPR116, KCNJ15, PTPN13, CTSH, PPFIBP2, CD302, SFTPB, HSD17B6, DLC1, ADRB2, PARM1, KLRB1, MS4A1, STX1A, KLK6, SLC16A3, PYGL, LDHA, ITGA5, VEGFC, EEF1A2, TPX2, UCK2, PHKA1, EIF4A3, TK1, CCNA2, GGH, CCNB1, MELK, HMMR, EIF2S1, TEAD4, HMGA1, RIMS2, H2AFZ), o una combinación de los mismos, que pueden ser divididos en tres módulos (Tabla 1, 2, y 3) basándose en criterios que incluyen la función biológica. La Tabla 1 (que es referida en la presente memoria también como Módulo 1) incluye genes implicados en la supresión tumoral, la Tabla 2 (que es referida en la presente memoria también como Módulo 2) incluye genes implicados en la angiogénesis, y la Tabla 3 (que es referida en la presente memoria también como Módulo 3) incluye genes implicados en la proliferación.

Se descubrió que algunos biomarcadores son expresados en exceso en el cáncer de pulmón en estadío temprano tales como aquellos marcadores implicados en la angiogénesis o la proliferación (Tabla 2 y Tabla 3, respectivamente), mientras otros biomarcadores implicados en la supresión tumoral son expresados a la baja (Tabla 1) en comparación con un control (p. ej., la expresión promedio de estos genes en pacientes con cáncer de pulmón en estadío temprano (estadío I y II)).

Biomarcador

Los biomarcadores de la invención incluyen uno o más biomarcadores enumerados en la Tabla 1, la Tabla 2, y/o la Tabla 3, o sus productos génicos. La presente invención se basa en el descubrimiento de que los biomarcadores enumerados en la Tabla 1, la Tabla 2, y/o la Tabla 3 se expresan diferencialmente. Analizando los niveles del perfil de expresión de uno o más biomarcadores identificados en la Tabla 1, la Tabla 2, y/o la Tabla 3 es posible determinar el pronóstico de un individuo con cáncer de pulmón en estadío temprano.

En un ejemplo, el método de la invención incluye medir uno o más biomarcadores de la Tabla 1. Por ejemplo, el método de la invención mide al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez, al menos once, al menos doce, al menos trece, al menos catorce o al menos quince, biomarcadores de la Tabla 1. En un ejemplo, se mide el nivel de expresión de un gen CBX7 de la Tabla 1. En otro ejemplo, se mide el nivel de expresión de dos biomarcadores CBX7 y TMPRSS2 de la Tabla 1. En otro ejemplo, se mide el nivel de expresión de tres biomarcadores CBX7, TMPRSS2 y GPR116 de la Tabla 1. En otro ejemplo más, se mide el nivel de expresión de cuatro biomarcadores CBX7, TMPRSS2, GPR116 y KCNJ15 de la Tabla 1. En otro ejemplo más, se mide el nivel de expresión de cinco biomarcadores CBX7, TMPRSS2, GPR116, KCNJ15 y PTPN13 de la Tabla 1.

Tabla 1.

Símbolo del gen	Entrada ID	Descripción	Acceso	Alias Símbolo
CBX7	23492	cromobox homólogo 7	NM_175709.3	
TMPRSS2	7113	proteasa transmembrana, serina 2	NM_005656.3; NM_001135099.1	FLJ41954; PP9284; PRSS10
GPR116	221395	receptor acoplado a proteína G 116	NM_015234.4; NM_001098518.1	DKFZp564O192 FLJ90640; KIAA0758; KPG 001
KCNJ15	3772	canal de potasio rectificado internamente, subfamilia J, miembro 15	NM_170736.1; NM_170737.1; NM_002243.3	IRKK; KIR1.3; KIR4.2; MGC13584
PTPN13	5783	proteína tirosina fosfatasa, tipo 13 no receptor (APO-1/CD95 fosfatasa asociada a (Fas))	NM_080684.2; NM_006264.2; NM_080685.2; NM_080683.2	DKFZp686J1497; FAP-1; PNP1; PTP-BAS; PTP-BL; PTP1E; PTPL1; PTPLE
CTSH	1512	catepsina H	NM_004390.3	ACC-4; ACC-5; CPSB; DKFZp686B24257; MGC1519; minicadena
PPFIBP2	8495	proteína de interacción con PTPRF, proteína de unión 2 (liprina beta 2)	NM_003621.2	Cclp1; DKFZp781K06126; MGC42541
CD302	9936	molécula CD302	NM_001198763.1; NM_001198764.1; NM_014880.4	BIMLEC; CLEC13A; DCL-1; DCL1; FLJ43091; KIAA0022; MGC22301
SFTPB	6439	proteína B tensioactiva	NM_198843.2; NM_000542.3	PSP-B; SFTB3; SFTP3; SMDP1; SP-B
HSD17B6	8630	homólogo de hidroxiesteroide (17-beta) deshidrogenasa 6 (ratón)	NM_003725.2	HSE; RODH; SDR9C6
DLC1	10395	suprimido en cáncer de hígado 1	NM_182643.2; NM_024767.3; NM_006094.4; NM_001164271.1	ARHGAP7; FLJ21120; HP; p122-RhoGAP; STARD12
ADRB2	154	receptor beta-2 adrenérgico, superficie	NM_000024.5	ADRB2R; ADRBR; B2AR; BAR;BETA2AR
PARM1	25849	proteína 1 de tipo mucina regulada por andrógeno de próstata	NM_015393.3	Cipar1; DKFZP564O0823; PARM-1; WSC4
KLRB1	3820	subfamilia de receptores de tipo lectina de células asesinas, miembro 1	NM_002258.2	CD161; CLEC5B; hNKR-P1A; MGC138614; NKR; NKR-P1; NKR-P1A; NKRP1A
MS4A1	931	4 dominios que abarcan la membrana, subfamilia A, miembro 1	NM_152866.2; NM_021950.3	B1; Bp35; CD20; CVID5; LEU-16; MGC3969; MS4A2; S7

En otro ejemplo, el método de la invención incluye medir uno más biomarcadores de la Tabla 2. Por ejemplo, el método de la invención mide la expresión de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete o al menos ocho biomarcadores de la Tabla 2. En un ejemplo, se mide el nivel de expresión de un gen STX1A de la Tabla 2. En un ejemplo, se mide el nivel de expresión de dos biomarcadores STX1A y KLK6 de la Tabla 2. En otro ejemplo, se mide el nivel de expresión de tres biomarcadores STX1A, KLK6 y SLC16A3 de la Tabla 2. En otro ejemplo, se mide el nivel de expresión de cuatro biomarcadores STX1A, KLK6, SLC16A3 y PYGL de la Tabla 2. En otro ejemplo, se mide el nivel de expresión de cinco biomarcadores STX1A, KLK6, SLC16A3, PYGL y LDHA de la Tabla 2.

Tabla 2

Símbolo del Gen	Entrada ID	Descripción	Acceso	Alias Símbolo
STX1A	6804	sintaxina 1A (cerebro)	NM_004603.3; NM_001165903.1	HPC-1; P35-1; STX1; SYN1A
KLK6	5653	peptidasa 6 relacionada con calicreína	NM_002774.3; NM_001012964.1; NM_001012965.1	Bssp; hK6; Klk7; MGC9355; PRSS18; PRSS9; SP59
SLC16A3	9123	familia 16 de portador de soluto, miembro 3 (transportador de ácido monocarboxílico 4)	NM_001042422.1; NM_001042423.1; NM_004207.2	MCT3; MCT4; MGC138472; MGC138474
PYGL	5836	fosforilasa, glicógeno, hígado	NM_002863.4; NM_001163940.1	GSD6
LDHA	3939	lactato deshidrogenasa A	NR_028500.1; NM_001165415.1; NM_005566.3; NM_001165416.1; NM_001135239.1; NM_001165414.1	GSD11; LDH1; LDHM; PIG19
ITGA5	3678	integrina, alfa 5 (receptor de fibronectina, polipéptido alfa)	NM_002205.2	CD49e; FNRA; VLA5A
VEGFC	7424	factor de crecimiento endotelial vascular C	NM_005429.2	Flt4-L; VRP
EEF1A2	1917	factor 1 de elongación de la traducción eucariótica alfa 2	NM_001958.2	EEF1AL; EF-1-alfa-2; EF1A; FLJ41696; HS1; STN; STNL

En otro ejemplo, el método de la invención incluye medir uno o más biomarcadores de la Tabla 3. Por ejemplo, el método de la invención mide al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez, al menos once, al menos doce, al menos trece, al menos catorce o al menos quince, biomarcadores de la Tabla 3. En un ejemplo, se mide el nivel de expresión de un gen TPX2 de la Tabla 3. En otro ejemplo, se mide el nivel de expresión de dos biomarcadores TPX2 y UCK2 de la Tabla 3. En otro ejemplo, se mide el nivel de expresión de tres biomarcadores TPX2, UCK2 y PHKA1 de la Tabla 3. En otro ejemplo, se mide el nivel de expresión de cuatro biomarcadores TPX2, UCK2, PHKA1 y EIF4A3 de la Tabla 3. En otro ejemplo más, se mide el nivel de expresión de cinco biomarcadores TPX2, UCK2, PHKA1 EIF4A3 y TK1 de la Tabla 3.

Tabla 3

Símbolo del Gen	Entrada ID	Descripción	Acceso	Alias Símbolo
TPX2	22974	TPX2, microtúbulo asociado, homólogo (Xenopus laevis)	NM_012112.4	C20orf1; C20orf2; DIL-2; DIL2; FLS353; GD:C20orf1; HCA519; HCTP4; p100; REPP86
UCK2	7371	uridina-citidina quinasa 2	NM_012474.4	DKFZp686M04245; TSA903; UK; UMPK
PHKA1	5255	fosforilasa quinasa, alfa 1 (músculo)	NM_001172436.1; NM_002637.3; NM_001122670.1	MGC132604; PHKA
EIF4A3	9775	factor de iniciación de la traducción eucariótica 4A3	NM_014740.3	DDX48; DKFZp686O16189; eIF4AIII; KIAA0111; MGC10862; NMP265; NUK34
TK1	7083	timidina quinasa 1, soluble	NM_003258.4	TK2
CCNA2	890	ciclina A2	NM_001237.3	CCN1; CCNA
GGH	8836	gamma-glutamil hidrolasa (conjugasa, folilpoligammaglutamil hidrolasa)	NM_003878.2	GH

Símbolo del Gen	Entrada ID	Descripción	Acceso	Alias Símbolo
CCNB1	891	ciclina B1	NM_031966.2	CCNB
MELK	9833	quinasa de la cremallera de leucina embrionaria materna	NM_014791.2	HPK38; KIAA0175
HMMR	3161	receptor de motilidad mediada por hialuronano (RHAMM)	NM_012485.2; NM_001142557.1; NM_001142556.1; NM_012484.2	CD168; IHABP; MGC119494; MGC119495; RHAMM
EIF2S1	1965	factor de inicio de la traducción eucariótica 2, subunidad 1 alfa, 35kDa	NM_004094.4	EIF-2; EIF-2A; EIF-2alfa; EIF2; EIF2A
TEAD4	7004	miembro 4 de la familia de dominio TEA	NM_003213.3; NM_201441.2; NM_201443.2	EFTR-2; hRTEF-1B; MGC9014; RTEF1; TCF13L1; TEF-3; TEF3; TEFR-1
HMGA1	3159	garfio 1 de AT del grupo de alta movilidad	NM_145903.2; NM_002131.3; NM_145899.2; NM_145902.2; NM_145901.2; NM_145905.2	HMG-R; HMGA1A; HMGIY; MGC12816; MGC4242; MGC4854
RIMS2	9699	exocitosis de membrana sináptica reguladora 2	NM_001100117.2; NM_014677.4	DKFZp781A0653; KIAA0751; OBOE; RAB3IP3; RIM2
H2AFZ	3015	familia de histona H2A, miembro Z	NM_002106.3	H2A.z; H2A/z; H2AZ; MGC117173

Los biomarcadores de la invención también pueden incluir cualquier combinación de biomarcadores identificados en la Tabla 1, la Tabla 2 y la Tabla 3 cuyo nivel de expresión o producto génico sirve como un marcador o biomarcador predictivo para el pronóstico de un individuo con cáncer de pulmón en estadío temprano. En un ejemplo, se mide el nivel de expresión de un gen seleccionado de cada Tabla, Tabla 1, Tabla 2 y Tabla 3, p. ej., CBX7, STX1A y TPX2. En otro ejemplo, se mide el nivel de expresión de dos biomarcadores seleccionados de cada una de las Tablas, Tabla 1, Tabla 2 y Tabla 3, p. ej., CBX7 y TMPRSS2 de la Tabla 1, STX1A y KLK6 de la Tabla 2 y TPX2 y UCK2 de la Tabla 3. Véase la siguiente Tabla 4 para ejemplos de las diversas combinaciones de biomarcadores de las Tablas 1, 2 y 3. No se pretende que las combinaciones mostradas en la Tabla 4 sean consideradas limitantes y se puede realizar cualquier combinación de los biomarcadores mostrados en las Tablas 1-3.

Tabla 4

Biomarcadores y Combinaciones de Biomarcadores	valor p
CBX7	4,90E-06
CBX7 TMPRSS2	1,10E-07
CBX7 TMPRSS2 GPR116	1,90E-07
CBX7 TMPRSS2 GPR116 KCNJ15	1,10E-08
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13	5,50E-09
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH	2,60E-09
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2	3,60E-10
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP22 CD302	2,10E-10
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2 CD302 SFTPB	1,10E-10
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2 CD302 SFTPB HSD17B6	1,90E-10
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2 CD302 SFTPB HSD17B6 DLC1	1,60E-10
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2 CD302 SFTPB HSD17B6 DLC1 ADRB2	8,50E-11
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2 CD302 SFTPB HSD17B6 DLC1 ADRB2 PARM1	8,40E-11

Biomarcadores y Combinaciones de Biomarcadores	valor p
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2 CD302 SFTPB HSD17B6 DLC1 ADRB2 PARM1 KLRB1	3,60E-11
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2 CD302 SFTPB HSD17B6 DLC1 ADRB2 PARM1 KLRB1 MS4A1	5,90E-12
STX1A	3,20E-04
STX1A KLK6	1,40E-05
STX1A KLK6 SLC16A3	1,50E-06
STX1A KLK6 SLC16A3 PYGL	6,90E-08
STX1A KLK6 SLC16A3 PYGL LDHA	2,10E-08
STX1A KLK6 SLC16A3 PYGL LDHA ITGA5	3,00E-08
STX1A KLK6 SLC16A3 PYGL LDHA ITGA5 VEGFC	1,40E-08
STX1A KLK6 SLC16A3 PYGL LDHA ITGA5 VEGFC EEF1A2	6,90E-08
TPX2	2,20E-05
TPX2 UCK2	4,30E-07
TPX2 UCK2 PHKA1	5,40E-08
TPX2 UCK2 PHKA1 EIF4A3	1,30E-08
TPX2 UCK2 PHKA1 EIF4A3 TK1	2,60E-08
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2	6,60E-08
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2 GGH	4,40E-08
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2 GGH CCNB1	1,50E-07
TPX2 UCK2 PHKA EIF4A3 TK1 CCNA2 GGH CCNB1 MELK	2,60E-07
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2 GGH CCNB1 MELK HMMR	2,80E-07
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2 GGH CCNB1 MELK HMMR EIF2S1	1,70E-07
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2 GGH CCNB1 MELK HMMR EIF2S1 TEAD4	1,20E-07
TPX2 UCK2 PHKA EIF4A3 TK1 CCNA2 GGH CCNB1 MELK HMMR EIF2S1 TEAD4 HMGA1	1,60E-07
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2 GGH CCNB1 MELK HMMR EIF2S1 TEAD4 HMGA1 RIMS2	7,60E-08
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2 GGH CCNB1 MELK HMMR EIF2S1 TEAD4 HMGA1 RIMS2 H2AFZ	7,00E-08
CBX7	3,10E-09
STX1A	
TPX2	
CBX7 TMPRSS2	1,60E-10
STX1A KLK6	
TPX2 UCK2	
CBX7 TMPRSS2 GPR116	6,10E-11
STX1A KLK6 SLC16A3	
TPX2 UCK2 PHKA1	
CBX7 TMPRSS2 GPR116 KCNJ15	1,90E-12
STX1A KLK6 SLC16A3 PYGL	
TPX2 UCK2 PHKA1 EIF4A3	
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13	1,80E-12

Biomarcadores y Combinaciones de Biomarcadores	valor p
STX1A KLK6 SLC16A3 PYGL LDHA	
TPX2 UCK2 PHKA1 EIF4A3 TK1	
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH	
STX1A KLK6 SLC16A3 PYGL LDHA ITGA5	3,20E-12
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2	
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2	
STX1A KLK6 SLC16A3 PYGL LDHA ITGA5 VEGFC	1,40E-12
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2 GGH	
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2 CD302	
STX1A KLK6 SLC16A3 PYGL LDHA ITGA5 VEGFC EEF1A2	3,10E-12
TPX2 UCK2 PHKA1 E1F4A3 TK1 CCNA2 GGH CCNB1	
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2 CD302 SFTPB	
STX1A KLK6 SLC16A3 PYGL LDHA ITGA5 VEGFC EEF1A2	2,80E-12
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2 GGH CCNB1 MELK	
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2 CD302 SFTPB HSD17B6	
STX1A KLK6 SLC16A3 PYGL LDHA ITGA5 VEGFC EEF1A2	4,00E-12
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2 GGH CCNB1 MELK HMMR	
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2 CD302 SFTPB HSD17B6 DLC1	
STX1A KLK6 SLC16A3 PYGL LDHA ITGA5 VEGFC EEF1A2	3,50E-12
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2 GGH CCNB1 MELK HMMR EIF2S1	
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2 CD302 SFTPB HSD17B6 DLC1 ADRB2	
STX1A KLK6 SLC16A3 PYGL LDHA ITGA5 VEGFC EEF1A2	2,60E-12
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2 GGH CCNB1 MELK HMMR EIF2S1 TEAD4	
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2 CD302 SFTPB	
HSD17B6 DLC1 ADRB2 PARM1	
STX1A KLK6 SLC16A3 PYGL LDHA ITGA5 VEGFC EEF1A2	3,50E-12
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2 GGH CCNB1 MELK HMMR EIF2S1	
TEAD4 HMGA1	
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2 CD302 SFTPB	
HSD17B6DLC1 ADRB2 PARM1 KLRB1	
STX1A KLK6 SLC16A3 PYGL LDHA ITGA5 VEGFC EEF1A2	1,90E-12
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2 GGH CCNB1 MELK HMMR EIF2S1	
TEAD4 HMGA1 RIMS2	
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 CTSH PPFIBP2 CD302 SFTPB	
HSD17B6 DLC1 ADRB2 PARM1 KLRB1 MS4A1	
STX1A KLK6 SLC16A3 PYGL LDHA ITGA5 VEGFC EEF1A2	9,60E-13
TPX2 UCK2 PHKA1 EIF4A3 TK1 CCNA2 GGH CCNB1 MELK HMMR EIF2S1	
TEAD4 HMGA1 RIMS2 H2AFZ	

En otro ejemplo, se seleccionan al menos 3, 4, 5, 6, 7, 8, 9, 10 genes cualesquiera de cada Tabla 1, Tabla 2, y Tabla 3. Por ejemplo, en una realización, se seleccionan al menos 15 biomarcadores donde se seleccionan 5

biomarcadores cualesquiera de la Tabla 1 (p. ej., CBX7, TMPRSS2, GPR116, KCNJ15 y PTPN13 o CBX7, TMPRSS2, CTSH, PPFIBP2 y CD302; o SFTPB; HSD17B6; DLC1; ADRB2 y PARM1), 5 biomarcadores cualesquiera se seleccionan de la Tabla 2 (p. ej., STX1A, KLK6, SLC16A3, PYGL y LDHA o SLC16A3, PYGL, ITGA5, VEGFC, y EEF1A2 o STX1A, KLK6, SLC16A3, PYGL y LDHA) y 5 biomarcadores cualesquiera se seleccionan de la Tabla 3 (TPX2, UCK2, PHKA1, EIF4A3, y TK1 o CCNA2, GGH, CCNB1, MELK y HMMR o EIF2S1, TEAD4, HMGA1, RIMS2, y H2AFZ).

En otra realización, los biomarcadores de la invención incluyen uno cualquiera, o cualquier combinación, de los siguientes genes: CBX7, TMPRSS2, GPR116, KCNJ15, PTPN13, CTSH, PPFIBP2, CD302, SFTPB, HSD17B6, DLC1, ADRB2, PARM1, KLRB1, MS4A1, STX1A, KLK6, SLC16A3, PYGL, LDHA, ITGA5, VEGFC, EEF1A2, TPX2, UCK2, PHKA1, EIF4A3, TK1, CCNA2, GGH, CCNB1, MELK, HMMR, EIF2S1, TEAD4, HMGA1, RIMS2, H2AFZ. En otra realización, los biomarcadores de la invención incluyen al menos 15, 20, 25, 30, 35, 36, 37 o 38 de los siguientes genes: CBX7, TMPRSS2, GPR116, KCNJ15, PTPN13, CTSH, PPFIBP2, CD302, SFTPB, HSD17B6, DLC1, ADRB22, PARM1, KLRB1, MS4A1, STX1A, KLK6, SLC16A3, PYGL, LDHA, ITGA5, VEGFC, EEF1A2, TPX2, UCK2, PHKA1, EIF4A3, TK1, CCNA2, GGH, CCNB1, MELK, HMMR, EIF2S1, TEAD4, HMGA1, RIMS2, H2AFZ. En una realización concreta, se seleccionan los siguientes 37 biomarcadores: CBX7, TMPRSS2, GPR116, KCNJ15, PTPN13, CTSH, PPFIBP2, CD302, SFTPB, HSD17B6, DLC1, ADRB2, PARM1, KLRB1, MS4A1, STX1A, KLK6, SLC16A3, PYGL, LDHA, ITGA5, VEGFC, TPX2, UCK2, PHKA1, EIF4A3, TK1, CCNA2, GGH, CCNB1, MELK, HMMR, EIF2S1, TEAD4, HMGA1, RIMS2, H2AFZ.

En un ejemplo, el perfil de expresión puede ser un conjunto de valores que representan los niveles de ARNm de uno o más biomarcadores enumerados en la Tabla 1, la Tabla 2, y/o la Tabla 3. En otro ejemplo, el perfil de expresión puede incluir un conjunto de valores que representan una o más proteínas o polipéptidos codificados por los biomarcadores enumerados en la Tabla 1, la Tabla 2, y/o la Tabla 3.

Preparación de Muestras

Se puede utilizar cualquier muestra de ensayo apropiada de células tomadas de un individuo que tiene cáncer de pulmón en estadío temprano que ha experimentado una resección quirúrgica para determinar la expresión de una pluralidad de biomarcadores de la invención. El tipo y la clasificación del cáncer de pulmón en estadío temprano pueden variar. El cáncer de pulmón puede estar en estadío I y/o estadío II. La muestra de ensayo puede ser de un cáncer de pulmón de células no pequeñas (CPCNP) que incluye carcinoma de células escamosas, adenocarcinoma, carcinoma de células grandes, así como histotipos independientes del subgrupo.

Generalmente, la muestra de ensayo de células o tejido se obtendrá del sujeto con cáncer por medio de biopsia o resección quirúrgica. La resección quirúrgica puede ser curativa o no curativa/RO. Se puede retirar una muestra de células, tejido o fluido por medio de biopsia de aspiración con jeringa. Para esto, se inserta una aguja fina unida a una jeringa a través de la piel y en el órgano o tejido de interés. La aguja se guía típicamente a la región de interés utilizando técnicas de diagnóstico médico por ultrasonidos o por tomografía computarizada (TC). Una vez que la aguja se inserta en el tejido, se crea un vacío con la jeringa de manera que las células o el fluido pueden ser succionados a través de la aguja y recogidos en la jeringa. Una muestra de células o tejido también se puede separar mediante incisión o biopsia con aguja gruesa. Para esto, se retira un cono, un cilindro, o un fragmento diminuto de tejido de la región de interés. Generalmente se utiliza la técnica de diagnóstico médico por TC, ultrasonidos, o una endoscopia para guiar este tipo de biopsia. Más concretamente, la lesión cancerosa completa puede ser retirada por medio de biopsia excisional o resección quirúrgica. En la presente invención, la muestra de ensayo es típicamente una muestra de células retirada como parte de una resección quirúrgica.

La muestra de ensayo, por ejemplo, de tejido, también se puede almacenar, p. ej., en RNAlater (Ambion; Austin Tex.) o congelar instantáneamente y almacenar a -80°C para su uso posterior. La muestra de tejido de la que se ha tomado la biopsia también se puede fijar con un fijador, tal como formaldehído, paraformaldehído, o ácido acético/etanol. La muestra de tejido fijada se puede incluir en cera (parafina) o una resina plástica. La muestra de tejido incluida (o la muestra de tejido congelada) se puede cortar en secciones finas. El ARN o la proteína también se pueden extraer de una muestra de tejido fijada o incluida en cera.

El sujeto con cáncer será generalmente un sujeto mamífero tal como un primate. En una realización ilustrativa, el sujeto es un ser humano.

Una vez que una muestra de células o una muestra de tejido son retiradas del sujeto con cáncer, pueden ser procesadas para el aislamiento del ARN o las proteínas utilizando mecanismos bien conocidos en la técnica como se describe a continuación.

En un ejemplo, el ARN puede ser extraído de las muestras de tejido o células mediante una variedad de métodos, por ejemplo, lisis con tiocianato de guanidio seguida de centrifugación en CsCl (Chirgwin, et al., Biochemistry 18:5294-5299, 1979). Se puede obtener ARN de células individuales como se describe en los métodos para preparar bibliotecas de ADNc a partir de células individuales (véanse, p. ej., Dulac, Curr. Top. Dev. Biol. 36:245, 1998; Jena, et al., J. Immunol. Methods 190:199, 1996). La muestra de ARN puede ser adicionalmente enriquecida para una especie concreta. En una realización, por ejemplo, se puede aislar ARN poli (A)+ a partir de una muestra

de ARN. En particular, se pueden inmovilizar oligonucleótidos poli-T sobre un soporte sólido para que sirvan como ligandos de afinidad para ARNm. Los kits para este propósito se encuentran disponibles en el mercado, por ejemplo, el MessageMaker kit (Life Technologies, Grand Island, N. Y.). En una realización, la población de ARN puede ser enriquecida para secuencias de interés, como se detalla en las Tablas 1-3. El enriquecimiento se puede lograr, por ejemplo, mediante síntesis de ADNc específica del cebador, o múltiples rondas de amplificación lineal basándose en síntesis de ADNc y transcripción in vitro dirigida por molde (véase, p. ej., Wang, et al., Proc. Natl. Acad. Sci. USA 86:9717, 1989).

Detección de la expresión del biomarcador

En un ejemplo, el método incluye determinar la expresión de uno o más biomarcadores enumerados en la Tabla 1, la Tabla 2, y/o la Tabla 3, o sus productos génicos a partir de una muestra tumoral de un paciente de ensayo con cáncer. Las secuencias génicas de cada uno de los biomarcadores enumerados en la Tabla 1, la Tabla 2, y/o la Tabla 3 se pueden detectar utilizando métodos conocidos en la técnica, p. ej., agentes que se pueden utilizar para detectar específicamente el gen o los productos génicos del mismo.

Los agentes de detección ilustrativos son sondas de ácido nucleico, que hibridan con los ácidos nucleicos correspondientes a los genes descritos en la presente memoria, y anticuerpos que se unen a los productos codificados por estos genes. Se pretende que los biomarcadores enumerados en la Tabla 1, la Tabla 2, y/o la Tabla 3 también incluyan secuencias de origen natural incluyendo variantes alélicas y otros miembros de la familia. Los biomarcadores de la invención también incluyen secuencias que son complementarias a las secuencias enumeradas resultantes de la degeneración del código y también las secuencias que son secuencias suficientemente homólogas que hibridan en condiciones restrictivas con los biomarcadores enumerados en la Tabla 1, la Tabla 2, y/o la Tabla 3.

En una realización, el método incluye: proporcionar una sonda de ácido nucleico que comprende una secuencia de nucleótidos, por ejemplo, al menos 10, 15, 25 o 40 nucleótidos, y hasta la totalidad o casi la totalidad de la secuencia codificante que es complementaria a una porción de la secuencia codificante de una secuencia de ácido nucleico enumerada en la Tabla 1, la Tabla 2, y/o la Tabla 3; obtener una muestra de tejido de un mamífero que tiene una célula cancerosa; poner en contacto la sonda de ácido nucleico en condiciones restrictivas con ARN obtenido de una biopsia tomada de un paciente con CPCNP (p. ej., en una transferencia Northern o un análisis de hibridación in situ); y determinar la cantidad de hibridación de la sonda con el ARN.

Las condiciones para la hibridación son conocidas por los expertos en la técnica y se pueden encontrar en Current Protocols in Molecular Biology, John Wiley and Sons, N.Y. (1989), 6.3.1-6.3.6. Un ejemplo preferido, no limitante de condiciones de hibridación altamente restrictivas es hibridación en 6 X cloruro de sodio/citrato de sodio (SSC) a aproximadamente 45 grados centígrados seguido de uno o más lavados en 0,2 X SSC, SDS al 0,1 por ciento a 50-65 grados centígrados. Por "suficientemente homóloga" se quiere significar una secuencia de aminoácidos o nucleótidos de un biomarcador que contiene un número suficiente o mínimo de residuos de aminoácido o nucleótidos idénticos o equivalentes (p. ej., un residuo de aminoácido que tiene una cadena lateral similar) a una segunda secuencia de aminoácidos o nucleótidos de manera que la primera y segunda secuencias de aminoácidos o nucleótidos comparten dominios estructurales o motivos comunes y/o una actividad funcional común. Por ejemplo, las secuencias de aminoácidos o nucleótidos que comparten dominios estructurales que tienen una homología de al menos aproximadamente 50 por ciento, una homología de al menos aproximadamente 60 por ciento, al menos aproximadamente 70 por ciento, al menos aproximadamente 80 por ciento, y una homología de al menos aproximadamente 90-95 por ciento a través de las secuencias de aminoácidos de los dominios se definen en la presente memoria como suficientemente homólogas. Además, las secuencias de aminoácidos o nucleótidos con una homología de al menos aproximadamente 50 por ciento, una homología de al menos aproximadamente 60-70 por ciento, al menos aproximadamente 70-80 por ciento, al menos aproximadamente 80-90 por ciento, y al menos aproximadamente 90-95 por ciento que comparten una actividad funcional común se definen en la presente memoria como suficientemente homólogas.

La comparación de secuencias y la determinación del porcentaje de homología entre dos secuencias se pueden completar utilizando un algoritmo matemático. Un ejemplo preferido, no limitante de un algoritmo matemático utilizado para la comparación de secuencias es el algoritmo de Karin y Altschul (1990) Proc. Natl. Acad. Sci. USA 87:2264-68, modificado como en Karlin y Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873-77. Semejante algoritmo se incorpora a los programas NBLAST y XBLAST (versión 2.0) de Altschul, et al. (1990) J. Mol. Biol. 215:403-10. Las búsquedas de nucleótidos BLAST se pueden llevar a cabo con el programa NBLAST, puntuación=100, longitud de palabra=12 para obtener secuencias de nucleótidos homólogas a las moléculas de ácido nucleico TRL de la invención. Las búsquedas de proteínas BLAST se pueden llevar a cabo con el programa XBLAST, puntuación=50, longitud de palabra=3 para obtener secuencias de aminoácidos homólogas a las secuencias de proteínas codificadas por los biomarcadores enumerados en la Tabla 1, la Tabla 2, y/o la Tabla 3. Para obtener alineamientos con espacios con fines de comparación, se puede utilizar Gapped BLAST como describen Altschul et al., (1997) Nucleic Acids Research 25(17):3389-3402. Cuando se utilizan los programas BLAST y Gapped BLAST, pueden usarse los parámetros por defecto de los programas respectivos (por ejemplo, XBLAST y NBLAST). Véase http://www.ncbi.nlm.nih.gov. Otro ejemplo no limitante, preferido de un algoritmo matemático utilizado para la comparación de secuencias es el algoritmo ALIGN de Myers y Miller, CABIOS (1989). Cuando se

utiliza el programa ALIGN para comparar secuencias de aminoácidos, se pueden utilizar una tabla de residuos de peso PAM120, una penalización de longitud del espacio de 12, y una penalización del espacio de 4.

Los ácidos nucleicos pueden ser marcados durante o después del enriquecimiento y/o la amplificación de los ARN. Por ejemplo, se puede llevar a cabo la transcripción inversa en presencia de un dNTP conjugado con una marca detectable, por ejemplo, un dNTP marcado fluorescentemente. En otra realización, la sonda de ADNc o ARN se puede sintetizar en ausencia de una marca detectable y se puede marcar con posterioridad, por ejemplo, incorporando dNTP o rNTP biotinilados, o algún medio similar (p. ej., fotoentrecruzamiento de un derivado psoraleno de biotina con ARN), seguido de la adición de estreptavidina marcada (p. ej., estreptavidina conjugada con ficoeritrina) o un equivalente.

Los radicales fluorescentes o marcas de interés incluyen cumarina y sus derivados (p. ej., 7-amino-4-metilcumarina, aminocumarina); colorantes Bodipy tales como azul cascada Bodipy FLy; fluoresceína y sus derivados (p. ej., isotiocianato de fluoresceína, verde Oregon); colorantes de rodamina (p. ej., rojo Texas, tetrametilrrodamina); eosinas y eritrosinas; derivados de cianina (p. ej., Cy2, Cy3, Cy3.5, Cy5, Cy5.5, Cy7); FluorX, quelatos macrocíclicos de iones lantánidos (p. ej., colorante Quantum.TM.); colorantes de transferencia de energía fluorescente tales como heterodímero de naranja de tiazol-etidio, TOTAB, dansilo, etc. También se pueden utilizar compuestos fluorescentes individuales que tienen funcionalidades para la unión a un elemento detectado deseablemente en un aparato o análisis de la invención, o que pueden ser modificados para incorporar tales funcionalidades (véase, p. ej., Kricka, 1992, Nonisotopic DNA Probe Techniques, Academic Press San Diego; Calif.). Las marcas quimioluminiscentes incluyen luciferina y 2,3-dihidroftalazinodionas, por ejemplo, luminol.

Detección de la expresión del producto del gen biomarcador

La detección de la presencia de un producto de proteína codificado por uno o más de los biomarcadores enumerados en la Tabla 1, la Tabla 2 y/o la Tabla 3 se puede llevar a cabo utilizando cualquier método apropiado conocido en la técnica. Por ejemplo, se puede utilizar un agente de interés para detectar una proteína concreta de interés, por ejemplo utilizando un anticuerpo. El método para producir anticuerpos policlonales y/o monoclonales que se unen específicamente a polipéptidos útiles en la presente invención es conocido por los expertos en la técnica y se puede encontrar, por ejemplo, en Dymecki, et al., (J. Biol. Chem. 267:4815, 1992); Boersma y Van Leeuwen, (J. Neurosci. Methods 51:317, 1994); Green, et al., (Cell 28:477, 1982); y Arnheiter, et al., (Nature 294:278, 1981). En una realización, se puede utilizar un inmunoanálisis para cuantificar los niveles de proteínas en las muestras celulares. La invención no está limitada a un procedimiento de análisis concreto, y por lo tanto, se pretende que incluya procedimientos tanto homogéneos como heterogéneos. Los inmunoanálisis ilustrativos que se pueden llevar a cabo de acuerdo con la invención incluyen inmunoanálisis de polarización de fluorescencia (FPIA), inmunoanálisis de fluorescencia (FIA), inmunoanálisis enzimático (EIA), inmunoanálisis de inhibición nefelométrica (NIA), análisis de inmunoabsorción con enzima ligada (ELISA), y radioinmunoanálisis (RIA). Se puede anclar un radical indicador, o un grupo marcador, a los anticuerpos sujeto y seleccionarlo con el fin de satisfacer las necesidades de diversos usos del método que a menudo están dictadas por la disponibilidad de un equipo de análisis y de procedimientos de inmunoanálisis compatibles. Los mecanismos generales que se van a utilizar en la realización de los diversos inmunoanálisis indicados anteriormente son conocidos por los expertos normales en la técnica. Alternativamente se pueden utilizar otros métodos tales como análisis de transferencia Western que incluye separar electroforéticamente proteínas sobre un gel de poliacrilamida, y después de teñir las proteínas separadas, se puede medir la cantidad relativa de cada proteína evaluando su densidad óptica. Alternativamente, se pueden utilizar otros métodos tales como análisis de transferencia puntual, FACS o inmunohistoquímica.

Las muestras de tejido se fijan mediante tratamiento con un reactivo tal como formalina, glutaraldehído, metanol, o similar. Las muestras se incuban a continuación con un anticuerpo (p. ej., un anticuerpo monoclonal) con una especificidad de unión para los polipéptidos marcadores. Este anticuerpo se puede conjugar con una marca para la posterior detección de la unión. Las muestras se incuban durante un tiempo suficiente para la formación de los inmunocomplejos. La unión del anticuerpo se detecta a continuación en virtud de la marca conjugada con este anticuerpo. Cuando el anticuerpo no está marcado, se puede emplear un segundo anticuerpo marcado, por ejemplo, que es específico para el isotipo del anticuerpo anti-polipéptido marcador. Los ejemplos de las marcas que se pueden emplear incluyen radionúclidos, compuestos fluorescentes, compuestos quimioluminiscentes, enzimas, y similares.

Cuando se emplean enzimas, el sustrato para la enzima se puede añadir a las muestras para proporcionar un producto coloreado o fluorescente. Los ejemplos de las enzimas adecuadas para su uso en productos conjugados incluyen peroxidasa de rábano picante, fosfatasa alcalina, malato deshidrogenasa, y similares. Cuando no están disponibles comercialmente, dichos conjugados anticuerpo-enzima se producen fácilmente mediante mecanismos conocidos por los expertos en la técnica.

En otra realización más, la invención contempla la utilización de un panel de anticuerpos que se generan contra los polipéptidos marcadores de esta invención.

Detección de ARNm

Un aspecto importante de la presente invención es la medición del nivel de expresión de uno o más biomarcadores identificados en la Tabla 1, la Tabla 2 y/o la Tabla 3 en un biopsia tumoral de cáncer de pulmón tomada de un sujeto que padece cáncer de pulmón en estadío temprano después de la resección quirúrgica. Los niveles de expresión se pueden analizar y utilizar para generar una puntuación de riesgo.

En un ejemplo, se puede utilizar la PCR con Transcriptasa inversa (RT-PCR) para el perfilado de la expresión génica para comparar los niveles de ARNm en diferentes poblaciones de muestras. El método incluye aislar el ARNm utilizando cualquier mecanismo conocido en la técnica, p. ej., utilizando un kit de purificación, un grupo de tampones y proteasa de fabricantes comerciales, tales como Qiagen, de acuerdo con las instrucciones del fabricante o un kit de Purificación de ADN y ARN completo (EPICENTRE (R), Madison, WT). La etapa de transcripción inversa se ceba típicamente utilizando cebadores específicos, hexámeros al azar, o cebadores oligo-dT, dependiendo de las circunstancias y el objetivo del perfilado de la expresión y a continuación se puede utilizar el ADNc derivado como molde en la siguiente reacción de PCR. Después se puede llevar a cabo una RT-PCR TaqMan(R) utilizando, p. ej., el equipo disponible en el mercado.

Una variación más reciente de la técnica de RT-PCR es la PCR cuantitativa en tiempo real, que mide la acumulación del producto de la PCR por medio de una sonda fluorogénica doblemente marcada (esto es, una sonda TaqMan(R)). La PCR en tiempo real es compatible tanto con la PCR competitiva cuantitativa, donde se utiliza un competidor interno para cada secuencia diana para la normalización, como con la PCR comparativa cuantitativa utilizando un gen de normalización contenido dentro de la muestra, o un gen constitutivo para la RT-PCR. Para más detalles véase, p. ej., Held et al, Genome Research 6:986-994 (1996).

En otro ejemplo, se utilizan micromatrices que incluyen una o más sondas correspondientes a uno o más biomarcadores identificados en la Tabla 1, la Tabla 2 y/o la Tabla 3. El método descrito anteriormente da como resultado la producción de patrones de hibridación de ácidos nucleicos diana marcados sobre la superficie de la matriz. Los patrones de hibridación resultantes de los ácidos nucleicos marcados se pueden visualizar o detectar en una variedad de formas, basándose la forma concreta de detección seleccionada en la marca concreta del ácido nucleico diana. Los métodos de detección representativos incluyen recuento de centelleo, autorradiografía, medición de la fluorescencia, medición calorimétrica, medición de la emisión de luz, dispersión de luz, y similares.

En un ejemplo, el método de detección utiliza un escáner para la lectura de matrices que se encuentra disponible en el mercado (Affymetrix, Santa Clara, Calif.), por ejemplo, el 417Arrayer Scanner, el 418Array Scanner, o el Agilent GeneArray.Scanner. Este escáner está controlado desde un sistema informático con una interfaz y herramientas de soporte lógico fáciles de utilizar. El resultado puede ser importado directamente o leído directamente por una variedad de aplicaciones de soporte lógico.

Según se utiliza en la presente memoria, el control para la comparación puede ser determinado por un experto en la técnica. En un aspecto, el control se determina seleccionando un valor que sirve como valor de corte. Por ejemplo, el valor puede ser un valor que diferencia entre p. ej., aquellas muestras de ensayo que tienen una buena supervivencia y aquellas que tienen una mala supervivencia; o entre aquellas muestras de ensayo en las que el individuo se beneficiaría de la terapia coadyuvante y aquellas en las que no; o entre aquellas muestras en las que el individuo se beneficiaría de la administración de un fármaco concreto tal como un inhibidor de la angiogénesis o un inhibidor de la proliferación. Un paciente que se podría beneficiar de la terapia coadyuvante representa una mejora en cualquiera de las mediciones del estado del paciente incluyendo aquellas mediciones utilizadas normalmente en la técnica tales como la supervivencia global, la supervivencia a largo plazo, la supervivencia libre de recurrencia, y la supervivencia libre de recurrencia distante.

Otros métodos para determinar los niveles de expresión génica incluyen el método de perfilado de la expresión génica basado en MassARRAY, desarrollado por Sequenom, Inc. (San Diego, CA) y el análisis seriado de la expresión génica (SAGE) (Velculescu et al, Science 270:484-487 (1995); y Velculescu et al., Cell 88:243-51 (1997).

Otros métodos más para determinar los niveles de expresión génica en materiales FFPE son la teconología qNPA™ (HTG Molecular Diagnostics, Inc., Arizona) y nanoString™ Technologies (Seattle), en la que no se requiere ni extracción de ARN ni amplificación. Utilizando la tecnología qNPA la muestra de FFPE se expone primero a tampón de lisis HTG y las sondas de protección de nucleasa complementarias al ARNm de los biomarcadores descritos en la presente memoria se añaden a continuación a la solución. Las sondas hibridan con todos los biomarcadores de ARN de interés, solubles y entrecruzados. Después de la hibridación, se añade la nucleasa S1 que destruye todos los ácidos nucleicos de hebra sencilla, no específicos, produciendo una cantidad estequiométrica de dúplex de biomarcador-sonda de ARNm. La hidrólisis con álcali libera a continuación la sonda de los dúplex. Las sondas se pueden transferir a continuación a un ArrayPlate programado, se añade conector de detección, y tanto las sondas como los conectores de detección son capturados sobre la matriz. El ArrayPlate se lava a continuación y una sonda de detección marcada con HRP añadida, se incuba. La placa con la matriz se lava a continuación y se añade un sustrato quimioluminiscente. Finalmente, se toman imágenes del ArrayPlate y se mide la expresión de cada uno de los biomarcadores de todos los pocillos. Utilizando las tecnologías Nanostring se emplean dos sondas de ~50 bases por ARNm biomarcador que hibridan con el ARNm en solución. La sonda informadora porta la señal, mientras la sonda de captura permite que el complejo sea inmovilizado para la recogida de datos. Después de la hibridación, las sondas en exceso se retiran y los complejos de sonda/diana se alinean y se inmovilizan en un Counter Cartridge.

En el método de la invención se mide y se analiza el nivel de expresión de uno o más biomarcadores como se describe anteriormente y se utiliza para generar una puntuación de riesgo como se describe más abajo. El umbral de expresión se puede utilizar como pronóstico, p. ej., para seleccionar aquellos individuos que tienen una buena supervivencia y aquellos que tienen una mala supervivencia.

Es necesario corregir (normalizar) tanto las diferencias en la cantidad de ARN analizado como la variabilidad en la calidad del ARN utilizado. Por lo tanto, el análisis mide e incorpora típicamente la expresión de ciertos genes normalizadores, incluyendo genes constitutivos bien conocidos, tales como GAPDH y Cypl. Alternativamente, la normalización se puede basar en la señal media o mediana (Ct) de todos los biomarcadores analizados o un gran subgrupo de los mismos (enfoque de normalización global). En una base de gen a gen, se comparan la cantidad normalizada medida de un ARNm de un tumor de un paciente con la cantidad encontrada en un grupo de referencia de tejidos de cáncer de pulmón. El número (N) de tejidos de cáncer de pulmón en esta referencia debe ser suficientemente elevado como para garantizar que los diferentes grupos de referencia (en su conjunto) se comportan esencialmente de la misma manera. Si se cumple esta condición, la identidad del tejido de cáncer de pulmón individual presente en un grupo concreto no tendrá un impacto significativo sobre las cantidades relativas de los biomarcadores analizados.

En los métodos de la invención, se mide la expresión de cada biomarcador y se convierte típicamente en un valor de expresión. Estos valores de expresión se utilizarán para generar una puntuación de riesgo calculando el promedio ponderado. La puntuación de riesgo se asocia con el riesgo de muerte, metástasis o recaída por medio de una base de datos de calibración, a través de una fórmula paramétrica o de modelos basados en datos, no paramétricos. Esta base de datos se construye a partir de un grupo de muestras de referencia con valores de expresión conocidos, puntuaciones de riesgo y seguimiento clínico. La calibración de la puntuación de riesgo puede estar disponible por separado para cada módulo (tabla 1, 2 ó 3) o ser específica para un subtipo de enfermedad concreta como define la histología, la estadificación del tumor u otras características tales como la edad del paciente. Para la predicción de la respuesta al tratamiento, se construyen fórmulas de calibración separadas para los pacientes tratados por medio de terapias específicas (o sin tratamiento). También se puede utilizar una puntuación de riesgo de un compuesto (combinando los módulos de la Tabla 1, 2, o 3 con cierta ponderación), con su propia fórmula de calibración o base de datos. El protocolo de elaboración de una decisión clínica se puede efectuar de acuerdo con la puntuación de riesgo calibrada o la supervivencia pronosticada o el tiempo para los eventos (recaída o metástasis) como se ha descrito anteriormente.

Una vez calculada la puntuación de riesgo, se puede utilizar para decidir sobre el curso apropiado del tratamiento para el sujeto. Un sujeto que tiene una puntuación de riesgo elevada (esto es, un tiempo de supervivencia corto o un mal pronóstico) se puede beneficiar de la recepción de una terapia coadyuvante. La terapia coadyuvante puede incluir agentes quimioterapéuticos apropiados, p. ej., Paraplatino (carboplatino), Platinol (cisplatino), Taxótero (docetaxel), Adriamicina (doxorrubicina), VePesid (etopósido), Gemzar (gemcitabina), Ifex (ifosfamida), Camptosar (irinotecán), Taxol (paclitaxel), Alimta (pemetrexed) e Hycamtin (topotecán) y/o terapia de radiación. Un sujeto que tiene una puntuación de riesgo negativa (esto es, un tiempo de supervivencia largo o un buen pronóstico) puede no beneficiarse de un tratamiento adicional.

En otro ejemplo, la puntuación de riesgo generada utilizando un grupo de genes de una Tabla concreta se puede utilizar para determinar el transcurso de una terapia específica. Por ejemplo, si un individuo tiene una puntuación de riesgo elevada basada en el análisis de los biomarcadores de la Tabla 2, ese individuo se puede beneficiar de recibir terapia coadyuvante que incluye un inhibidor de la angiogénesis tal como avastina, srafinib, sunitinib o pazopanib. Alternativamente, si un individuo tiene una puntuación de riesgo elevada basada en el análisis de biomarcadores de la Tabla 3, ese individuo se puede beneficiar de recibir una terapia coadyuvante que incluye un agente anti-proliferativo tal como un inhibidor de la topoisomerasa (I y II), Taxano, antraciclina, antitubulina, antimetabolito o agentes alqulilantes.

Análisis de los Datos

Para facilitar la operación de análisis de la muestra, se pueden analizar los datos obtenidos por el lector a partir del dispositivo utilizando un ordenador digital. Típicamente, el ordenador se programará apropiadamente para recibir y almacenar los datos desde el dispositivo, así como para el análisis y la presentación de los datos reunidos, por ejemplo, la sustracción del fondo, la verificación de que los controles han funcionado apropiadamente, la normalización de las señales, la interpretación de los datos de fluorescencia para determinar la cantidad de diana hibridada, la normalización del fondo, y similares.

Kits

La invención proporciona adicionalmente kits para determinar el nivel de expresión de los biomarcadores descritos en la presente memoria. Los kits pueden ser útiles para determinar el pronóstico de sujetos con cáncer de pulmón. Un kit puede comprender una micromatriz que comprende sondas de cualquiera, de cualquier combinación de biomarcadores, identificados en las Tablas 1-3 y/o cualquier otro soporte sólido al cual se pueden anclar las sondas y el soporte sólido se puede utilizar para medir la expresión génica de una muestra de ensayo. En una realización, el kit comprende un medio legible por ordenador que incluye un soporte lógico para el análisis del perfil de expresión

susceptible de ser cargado en la memoria del sistema informático y que puede convertir los valores de expresión medidos en una puntuación de riesgo. Un kit puede comprender adicionalmente controles de ácido nucleico, tampones, e instrucciones para su uso.

Un experto en la técnica reconocerá muchos métodos y materiales similares o equivalentes a los descritos en la presente memoria, que se podrían utilizar en la práctica de la presente invención. En efecto, la presente invención no está en modo alguno limitada a los métodos y materiales descritos. Para los propósitos de la presente invención, los términos siguientes se definen más adelante.

Ejemplos

Ejemplo 1:

Se estudió una combinación de un grupo de datos de Novartis y varios grupos de datos públicos de cohortes de pacientes de pulmón. La descripción de los criterios de selección de los pacientes y las características clínicas se puede encontrar en los respectivos artículos originales para los grupos de datos públicos (véase más abajo). Para el grupo de datos de Novartis, se recogieron 412 muestras de pacientes a partir de pacientes con CPCNP que habían experimentado resección quirúrgica. Se realizaron procedimientos de estadificación convencionales incluyendo TC-Scans, FDG-PET de ganglios linfáticos sospechosos (> 1 cm en TC) y MRI. Se realizaron histologías de CPCNP para determinar si el CPCNP era Escamoso, Adeno-Carcinoma u otro tal como BAC. Asimismo se realizó la estadificación basada en TNM para definir si el CPCNP se encontraba en estadío I o estadío II. El tejido recién congelado se acumuló para el análisis genómico. El criterio de valoración primario fue la supervivencia global. La supervivencia global se refiere al tiempo (en años) desde la primera cirugía y se puede definir por un período tal como al menos 3 años, por ejemplo un período de 5 años, que está libre de recaída o recurrencia.

Grupos de datos públicos utilizados en este ejemplo:

Tabla 5

Nombre del grupo de Datos	Institución	Referencia	Fuente	Palataforma de Micromatriz de Expresión
DFCI	Dana-Farber Cancer Institute	Shedden et al. 2008	caArray	Affymetrix U133A
			https://array.nci.nih.gov/caarray/project/jacob-00182	
HLM	Moffitt Cancer Center	Shedden et al. 2008	caArray	Affymetrix U133A
			https://array.nci.nih.gov/caarray/project/jacob-00182	
JBR	Ontario Cancer Institute, Prince Margaret Hospital	Zhu et al. 2010	GEO GSE14814	Affymetrix U133A
			http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE14814	
MI	University of Michigan Cancer Center	Shedden et al. 2008	caArray	Affymetrix U133A
			https://array.nci.nih.gov/caarray/project/jacob-00182	
MIT	Massachusetts Institute of Technology	Bhattacharjee et al. 2001	http://broad.institute.org/mpr/lung	Affymetrix U95A
MSKCC	Memorial Sloan-Kettering Cancer Center	Shedden et al. 2008	caArray	Affymetrix U133A
			https://array.nci.nih.gov/caarray/project/jacob-00182	
NCCH	University of Caroline at Chapel Hill	Wilkerson et al. 2010	GEO GSE17710	Agilent UNC matriz adaptada 44k
			http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE17710	
NU1	Nagoya University	Takeuchi et al. 2006	GEO GSE11969	Agilent matriz adaptada 21.6k
			http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE11969	

Nombre del grupo de Datos	Institución	Referencia	Fuente	Palataforma de Micromatriz de Expresión
VRX	Veridex, LLC	Raponi et al. 2006	GEO GSE4573	Affymetrix U133A
			http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE4573	
WU	Washington University	Lu et al. 2006	GEO GSE6253	Affymetrix U133A, U133B, U95Av2
			http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE6253	

Las referencias de grupos de datos incluyen:

Bhattacharjee A et al. Proc Natl Acad Sci USA **98**:13790-5, 2001; Takeuchi et al..(2006) J Clin Oncol **24**:1679-88; Raponi et al. (2006) Cancer Res**66**:7466-72; Lu et al. (2006) PLoS Med **3**:e467; Shedden et al. (2008) Nat Med **14**:822-7; Hou et al. (2010) PLoS One **5**:e10312; Wilkerson et al., (2010) Cancer Res. **16**: 4864-75. Zhu et al..(2010) J Clin Oncol **28**:4417-24;

Procesamiento de la micromatriz de expresión génica de los grupos de datos públicos

El procedimiento para preparar los grupos de datos para el análisis fue similar al descrito por Wirapati et. al. 2008 en Breast Cancer Research **10**:R65, y brevemente se esboza como sigue:

1. Para los datos de expresión génica, los valores de expresión normalizados o pre-procesados que fueron proporcionados por los autores originales se utilizaron sin modificación

2. Las sondas o grupos de sondas de la plataforma de la micromatriz se volvieron a mapear frente a la misma base de datos de secuencias de referencia (RefSeq versión 39)

3. Solamente se utilizaron en este ejemplo pacientes con un estadío temprano (estadío I y II), sumando un total de n=834 de todos los grupos de datos (con n=571 de estadío I, y n=263 de estadío II). Se incluyen todos los tipos histológicos, aunque la mayor parte son adenocarcinoma (n=585) o carcinoma de células escamosas (n=226), siendo solo n=23 de otros tipos histológicos.

Los genes de la firma se identificaron por medio de un análisis integrado a gran escala de una base de datos de expresión génica y clínica completa que consiste en grupos de datos de cáncer de pulmón recién generados por Novartis (dos cohortes que suman 412 pacientes, no publicadas) y grupos de datos de expresión génica disponibles públicamente. Los genes de la firma se seleccionaron y se agruparon en los tres módulos (Tablas 1, 2, y 3) basándose en los criterios que incluyen la similitud de patrones de expresión con los de otros tipos de cáncer y función biológica. Los grupos de datos disponibles públicamente se seleccionaron de manera que el rendimiento pronóstico pudiera ser verificado independientemente utilizando los métodos esbozados más abajo.

Cuando se aplica una firma de biomarcador (un grupo de biomarcadores como se especifica en las Tablas 1, 2, y/o 3), los genes que se habían perdido se ignoraban. Se asignó una puntuación bruta a cada firma promediando los valores de expresión para los genes que estaban presentes en un plataforma concreta. La puntuación estandarizada se produjo sustrayendo la media de la puntuación bruta, y dividiendo por la desviación típica. La media y la desviación típica se determinaron por separado para cada grupo de datos. Se produjeron tres puntuaciones diferentes para cada paciente. Los autores de la presente invención se referirán a ellas como módulos 1, 2, o 3, correspondientes a los grupos de genes de las Tabla 1, 2, o 3, respectivamente.

Para demostrar que las puntuaciones de los tres módulos no estaban proporcionando una información similar, los autores de la presente invención realizaron gráficos de dispersión de distribuciones por pares de las puntuaciones en la Fig. 1. El módulo 1 está actuando más o menos en dirección contraria al módulo 3, de acuerdo con su implicación biológica en la supresión del tumor (módulo 1) y en la proliferación (módulo 3). Aunque estos dos módulos pueden proporcionar una información de pronóstico muy similar, la reciprocidad de la expresión fue importante técnicamente para una detección exacta. De este modo un valor bajo de un módulo solamente se interpretó como una expresión promedio baja (en lugar de como un fallo técnico en la detección) si estaba acompañado por un valor elevado del otro módulo. El módulo 2 mostró una ligera correlación con el módulo 1 y 3, pero con patrones asimétricos interesantes: aunque era posible tener ambos valores de los módulos 1 y 2 bajos, era poco probable tener valores elevados en ambos. Por otra parte, los valores elevados del módulo 2 parecen requerir un valor elevado del módulo 3, aunque lo contrario no es cierto (esto es, algunos tumores con un valor elevado del módulo 3 pueden tener un valor bajo del módulo 2).

Evaluación del rendimiento pronóstico de las firmas

Para demostrar la utilidad clínica de cada una de las puntuaciones de los módulos, los autores de la presente invención realizaron análisis de supervivencia utilizando curvas de Kaplan-Meier (Kaplan y Meier J. Am. Statist. Assoc. 53:457-481, 1953), utilizando las puntuaciones cuantitativas para dividir los pacientes en cuartiles (grupos que contienen 25% de los pacientes).

Las Figs. 2A-C representan el rendimiento pronóstico de las tres firmas (módulo 1, 2, y 3) cuando se aplican por separado, mostradas por análisis de supervivencia de Kaplan-Meier de la supervivencia global de los pacientes hasta cinco años después de la cirugía. Los pacientes se clasificaron en cuatro grupos Q1, Q2, Q3, y Q4, de acuerdo con los cuartiles de las puntuaciones, correspondientes al 1er, 2º, 3er, y 4º cuartiles, respectivamente. Esta clasificación permite el examen del cambio de riesgo como una función de la puntuación. Se utiliza el análisis de regresión de Cox para comparar pares de curvas seleccionados, mostrando los resultados la razón de riesgo (HR) y su intervalo de confianza del 95%, y el valor de p del ensayo de hipótesis frente a la ausencia de efecto (HR = 1). Los números de la derecha de cada curva son los cinco años de supervivencia en tanto por ciento. Los tres paneles demuestran que cada firma tiene un valor pronóstico por sí mismo. La Fig. 2D ("Combinación") muestra los resultados de una puntuación de riesgo obtenida mediante una simple combinación de los tres módulos (-módulo 1 + módulo 2 + módulo 3)

Un ejemplo de un sistema de pronóstico que combina las tres puntuaciones también se muestra en la Fig. 2D. Las puntuaciones simplemente se resumieron, con una multiplicación por -1 para el módulo 1 para invertir la dirección del efecto. Aunque no se observó una mejora drástica sobre algunos módulos individuales, la gradación del riesgo como una función de la puntuación fue más uniforme.

En resumen, cada firma individual (módulo 1, 2 o 3), así como su combinación (módulo 1, 2 y 3), mostraron la capacidad de distinguir entre la supervivencia (o equivalentemente, la mortalidad relacionada con la enfermedad) de subgrupos de pacientes. En todos los casos se observan diferencias sustanciales y estadísticamente significativas al menos entre los cuartiles extremos.

Rendimiento de las firmas en el contexto de los factores clínico-patológicos

Para demostrar que las firmas propuestas añaden nueva información de pronóstico a factores bien establecidos tales como histología, estadificación del tumor y edad en el diagnóstico, los autores de la presente invención realizaron análisis similares a los anteriores, pero estratificando los datos en grupos. En este ejemplo ilustrativo, los autores de la presente invención mostraron la estratificación por cada factor por separado, dividiéndolos en dos grupos principales con el fin de que tuvieran un tamaño de muestra suficiente en cada grupo. Los factores considerados son:

- Estadificación del tumor: estadío I versus estadío II
- Histología: adenocarcinoma versus carcinoma de células escamosas
- Edad en el diagnóstico: menos de o igual a 65 años versus más de 65 años (esta es la mediana de la edad de los pacientes en los datos)

Cada una de las firmas (módulo 1, 2, y 3) y la combinación se muestran por separado en las Fig. 3A-F, Fig. 4A-F, Fig. 5A-Fy Fig. 6A-F, respectivamente. Específicamente, la Fig. 3A-F muestra el rendimiento pronóstico del módulo 1 en la estratificación por estadío, histología y edad en el diagnóstico. Los cuartiles son los mismos que en la Figura 2 (esto es, el agrupamiento se realiza primero, antes de la estratificación). Aquí, como se muestra, el poder pronóstico permanece dentro de cada uno de los estratos. La Fig. 4A-F muestra el rendimiento pronóstico del módulo 2 en la estratificación por estadío, histología y edad en el diagnóstico. La Fig. 5A-F muestra el rendimiento pronóstico del módulo 3 en la estratificación por estadío, histología y edad en el diagnóstico. La Fig. 6A-F muestra el rendimiento pronóstico de la puntuación combinada del módulo 1, 2 y 3 en la estratificación por estadío, histología y edad en el diagnóstico.

En mayor parte de los casos, todavía se observaba el poder pronóstico de las firmas (individualmente y combinadas). Esto indica que las firmas propuestas proporcionaban un pronóstico adicional más allá de los factores tradicionales. Esto es, las firmas no son simplemente marcadores sustitutos que están muy correlacionados con factores existentes. En particular, los pacientes con el mismo estadío de tumor se pueden distinguir adicionalmente en un intervalo de riesgos. Esto fue no solamente un perfeccionamiento del sistema de estadificación, ya que la clasificación de riesgo puede ser invertida. Por ejemplo, el grupo de pacientes con el riesgo más alto en el estadío I está teniendo realmente una peor supervivencia que la supervivencia media de los pacientes en estadío II.

Los análisis en varios contextos de los factores clínicos-patológicos existentes también destacan que estos factores también pueden ser incorporados en la aplicación de las firmas. Por ejemplo, para el carcinoma de células escamosas, las firmas individuales no mostraron una discriminación del riesgo sustancial, pero la puntuación combinada muestra el cuartil superior que tiene sustancialmente un resultado peor que el resto.

La Figura 7A-F muestra el rendimiento pronóstico de diversas variaciones de los ejemplos de la firma reivindicada obtenido seleccionando subgrupos de genes de la Tabla 1, 2, y/o 3. La capacidad para separar a los pacientes de acuerdo con la supervivencia no está comprendida al no tener el grupo completo de genes. Una firma mínima formada por tres genes (uno de cada Tabla 1, 2 y 3) ya funciona bastante bien (panel A). Las separaciones entre grupos mejoran a medida que el número de genes aumenta (panel B-D). Quince genes seleccionados arbitrariamente funcionan bien (panel E y F).

Capacidad de decisión de terapia individualizada

El sistema de predicción del riesgo utilizará la base de datos de los datos clínicos y de expresión génica, similar a la utilizada en este ejemplo para permitir la proyección del riesgo en tratamientos alternativos (incluyendo la ausencia de tratamiento). Este sistema es similar a AdjuvantOnline (Ravdin et al 2001, J. Clin. Oncol. 19:980), excepto que incluye también las puntuaciones de la invención reivindicada.

Ejemplo 2:

Aplicación típica de los biomarcadores descritos en las Tablas 1, 2 y/o 3 para un paciente con cáncer de pulmón.

1. Un paciente diagnosticado de cáncer de pulmón con tumores primarios pequeños y operables experimenta cirugía para retirar la lesión. Una parte del tejido tumoral se examina mediante un procedimiento patológico convencional tal como la determinación del tamaño del tumor, los tipos histológicos del tumor (tales como adenocarcinoma, carcinoma de células escamosas, u otros tipos). La estadificación del tumor se determina utilizando pautas convencionales, basadas en el tamaño del tumor, la presencia de metástasis en los ganglios linfáticos u otros sitios distantes de metástasis. La información obtenida a partir de las mediciones clínico-patológicas convencionales puede modificar y aumentar el pronóstico y la aplicación predictiva de la invención, pero no es un requerimiento ni una parte integrante.

2. Una parte del tejido tumoral se utiliza como material de origen para la invención reivindicada, ya sea tejido congelado, incluido en parafina o fresco. La extracción de todo el ARN del transcriptoma se realiza sobre el tejido.

3. La medición de la cantidad relativa del ARN para el grupo específico de genes (reivindicada en la invención) se realiza por medio de cualquiera de estos procedimientos:

- Tecnología qNPA o tecnología nanoString
- RT-PCR cuantitativa de genes específicos
- Hibridación de micromatrices que contienen sondas seleccionadas o transcriptoma completo
- Secuenciación de alto rendimiento del ARN (RNA-sec), seguida de selección informática y cuantificación de los genes relevantes

4. La abundancia relativa del ARN para cada una se calibra por medio de un procedimiento de normalización computacional frente a un grupo de genes de control que se espera que tengan una expresión constante a través de todos los pacientes (Popovici et al. (2009) Selecting control genes for RT-QPCR using public microarray data. BMC Bioinformatics. 2009 Feb 2; 10:42), o por medio de normalización frente a una base de datos de mediciones similares a partir de tumores de pulmón existentes (McCall et al. (2010) Biostatistics. 2010 Abr; 11 (2):242, -53. Epub 2010 Ener 22.)

5. La abundancia relativa es un log transformado, y se computan los promedios ponderados de múltiples genes en un grupo. Los pesos pueden ser tan simples como +1 o -1 (signo positivo o negativo, dependiendo del efecto del cambio de la expresión génica sobre el resultado) (Sotiriou et al. (2006) J Natl Cancer Inst. 2006 Feb 15; 98(4):262-72.) o pueden ser números calibrados para las técnicas de medición específicas. Se considera que estos promedios ponderados son las puntuaciones de riesgo brutas que se van a utilizar como entrada en la siguiente calculadora de riesgo. Cada módulo de firma tiene su propia puntuación de riesgo.

6. La calculadora de perfiles de riesgo proporciona una proyección/predicción de la probabilidad de un estado libre de enfermedad, un estado libre de metástasis o supervivencia de un paciente individual, para cualquier punto temporal en el futuro (medido en años después de la cirugía). La fiabilidad de la proyección se caracteriza por un intervalo de confianza específico del tiempo que depende de la cantidad y la calidad de la base de datos de referencia. La proyección a corto plazo está apoyada típicamente por más datos y por lo tanto es más fiable (como se indica por un intervalo de confianza más estrecho) en comparación con la predicción a largo plazo. Un ejemplo de un diagrama calculador del perfil de riesgo se muestra en la Fig. 8A, donde una puntuación del riesgo medido para un paciente específico se traduce a una mortalidad por cáncer de pulmón (basándose en una base de datos de referencia, utilizando estimaciones de Kaplan-Meier del vecino más próximo).

7. La proyección del riesgo se puede modular por la información distinta de las puntuaciones de riesgo proporcionadas por la invención, tales como la edad y el género de los pacientes, la puntuación funcional de

Karnofsky, el tamaño del tumor y el estado de los ganglios linfáticos. La integración de todos estos factores en un perfil de riesgo específico para un paciente dado se puede realizar utilizando un método de regresión de supervivencia bien establecido (Therneau TM, Grambsch PM (2000) Modelling survival data: extending the Cox model. Springer, Nueva York), tal como

- modelos de regresión paramétrica (p. ej. utilizando modelos exponenciales o de Weibull)
- método semi-paramétrico (regresión de Cox)
- estimadores de curva de supervivencia empíricos tales como métodos de Kaplan-Meier o actuariales para cada posible combinación de factores

Un ejemplo de cómo la predicción del riesgo es modulada por el estadío del tumor se muestra en la Fig. 8B. Los pacientes en estadío 2 muestran una mortalidad en 5 años mayor que los pacientes en estadío 1, como cabía esperar. Sin embargo, un subgrupo de pacientes en estadío 1 que tiene una puntuación de riesgo elevada (mayor de cero, en unidad de puntuación de riesgo normalizada), muestra una tasa de mortalidad muy similar a la de pacientes en estadío 2 promedio (curva discontinua). Esto indica que la quimioterapia coadyuvante por estadificación solamente podría ser inadecuada porque algunos pacientes en estadío 1 (no tratados con las pautas actuales) tienen de hecho un riesgo similar al de muchos pacientes en estadío 2 (tratados con las pautas actuales).

La proyección de riesgo se puede calibrar frente a una base de datos de observaciones del pasado de pacientes con registros del resultado de la enfermedad, variables clínico-patológicas y mediciones de la invención reivindicada. La base de datos se puede actualizar periódicamente con información de nuevos pacientes. Este sistema es similar a AdjuvantOnline (Ravdin et al. (2001) J Clin Oncol. 2001 Feb 15; 19(4):980-91), una herramienta ampliamente utilizada para proyectar la supervivencia de pacientes con cáncer como una función de diversas variables clinico-patológicas.

Los autores de la presente invención extienden el sistema para incluir puntuaciones multi-modales derivadas de la tecnología genómica y transcriptómica.

8. Los factores moduladores colocados en la calculadora de riesgo pueden incluir la terapia coadyuvante utilizada comúnmente (tal como la quimioterapia basada en platino) o fármacos anti-angiogénesis. En este escenario de aplicaciones de respuesta-predicción, se presentarán dos o más perfiles de riesgo, correspondientes a la probabilidad proyectada de resultados bajo tratamientos alternativos, o sin tratamiento.

Ejemplo 3

Aplicación de la firma a materiales tumorales incluidos en parafina fijados con formalina (FFPE).

Las firmas reivindicadas (el grupo de genes y la fórmula para derivar las puntuaciones de riesgo) se pueden aplicar directamente a los datos de expresión de plataformas de tecnología tales como Affymetrix, qNPA o nanoString, después de la preparación del tejido y el procedimiento de pre-procesamiento de los datos brutos adecuados para cada plataforma respectiva.

La Fig. 9 muestra el resultado de supervivencia en grupos de datos disponibles públicamente de M. D. Anderson Cancer Center (Xie et al. 2011 *Clin Cancer Res* **17**:5705-14; Gene Expression Omnibus acceso GSE29013) que se obtuvieron de material FFPE. Aquí, todavía se observa un poder pronóstico significativo, a pesar de la calidad generalmente baja (nivel de ruido mayor, mayor número de asignaciones ausentes de valor de expresión) en datos Affymetrix a partir de FFPE.

Para evaluar si las puntuaciones de riesgo de la firma pueden proporcionar potencialmente el mismo valor pronóstico en datos qNPA y nanoString con FFPE como en Affymetrix con tejido congelado fresco, se realizaron comparaciones sobre materiales de los mismos pacientes con cáncer de pulmón (datos no publicados). La Fig. 10A-B muestra gráficos de dispersión por pares que comparan FF/Affymetrix vs FFPE/qNPA, así como FF/Affymetrix vs FFPE/nanoString. Se observaron unas correlaciones elevadas (0,86 y 0,87 respectivamente), que demostraban que la firma reivindicada no es dependiente de la elección de las plataformas de tecnología, ni de los métodos de conservación del tejido.

REIVINDICACIONES

1. Un método para pronosticar en un sujeto cáncer de pulmón de células no pequeñas (CPCNP) que comprende:

utilizar una muestra de ensayo obtenida de un sujeto que padece CPCNP después de la resección quirúrgica;

determinar el nivel de expresión de una combinación de biomarcadores, en donde la combinación de biomarcadores comprende CBX7, TMPRSS2, GPR116, STX1A, KLK6, SLC16A3, TPX2, UCK2, y PHKA1; y

analizar el nivel de expresión para generar una puntuación del riesgo, en donde la puntuación del riesgo se puede utilizar para proporcionar un pronóstico del sujeto.

2. Un método para pronosticar en un sujeto cáncer de pulmón de células no pequeñas (CPCNP) que comprende:

utilizar una muestra de ensayo obtenida de un sujeto que padece CPCNP después de la resección quirúrgica;

determinar el nivel de expresión de una combinación de biomarcadores en donde la combinación de biomarcadores comprende CBX7, STX1A, y TPX2; y

analizar el nivel de expresión para generar una puntuación del riesgo, en donde la puntuación del riesgo se puede utilizar para proporcionar un pronóstico del sujeto.

3. El método de la reivindicación 1 o la reivindicación 2, en donde la puntuación de riesgo clasifica al sujeto en un grupo de alto riesgo y se beneficiaría de recibir quimioterapia coadyuvante o en un grupo de riesgo bajo y no se beneficiaría de recibir quimioterapia coadyuvante.

4. El método de la reivindicación 1 o la reivindicación 2, en donde el sujeto tiene un pronóstico de escasa supervivencia o un pronóstico de buena supervivencia.

5. El método de la reivindicación 2, en donde la combinación de biomarcadores comprende adicionalmente TMPRSS2, KLK6, y UCK2.

6. El método de la reivindicación 2, en donde la combinación de biomarcadores comprende adicionalmente TMPRSS2, GPR116, KLK6, SLC16A3, UCK2, y PHKA1.

7. El método de la reivindicación 1, en donde la combinación de biomarcadores comprende adicionalmente KCNJ15, PYGL, y EIF43.

8. El método de la reivindicación 1, en donde la combinación de biomarcadores comprende adicionalmente KCNJ15, PTPN13, PYGL, LDHA, EIF4A3 y TK1.

9. El método de la reivindicación 1, en donde la combinación de biomarcadores comprende adicionalmente KCNJ15, PTPN13, CTSH, PYGL, LDHA, ITGA5, EIF4A3, TK1 y CCNA2.

10. El método de una cualquiera de las reivindicaciones anteriores, en donde la muestra de ensayo son células frescas, congeladas, incluidas fijadas en parafina.

11. El método de una cualquiera de las reivindicaciones anteriores, en donde el nivel de expresión se determina utilizando qNPA, nanoString, PCR cuantitativa o una matriz.

12. El método de la reivindicación 1 o la reivindicación 2, en donde el análisis de la expresión para generar una puntuación de riesgo se realiza utilizando el análisis estadístico.

13. El método de la reivindicación 12, en donde el análisis estadístico comprende el análisis de regresión de Cox o predictores de supervivencia paramétricos.

14. El método de las reivindicaciones anteriores, en donde el sujeto es un ser humano.

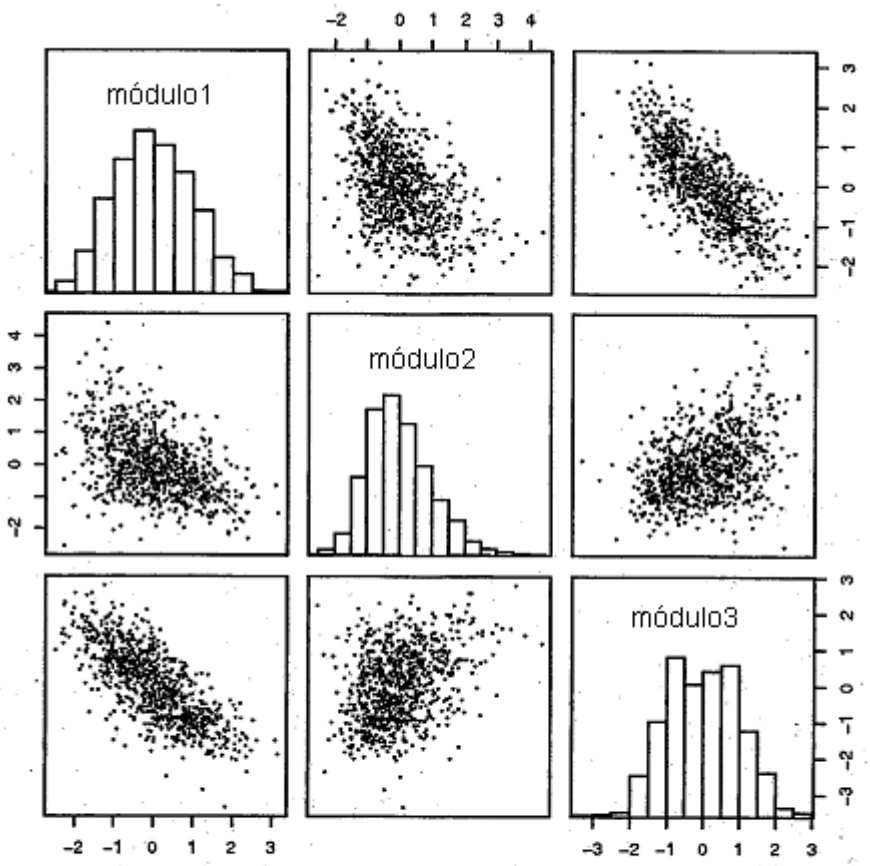


Fig. 1

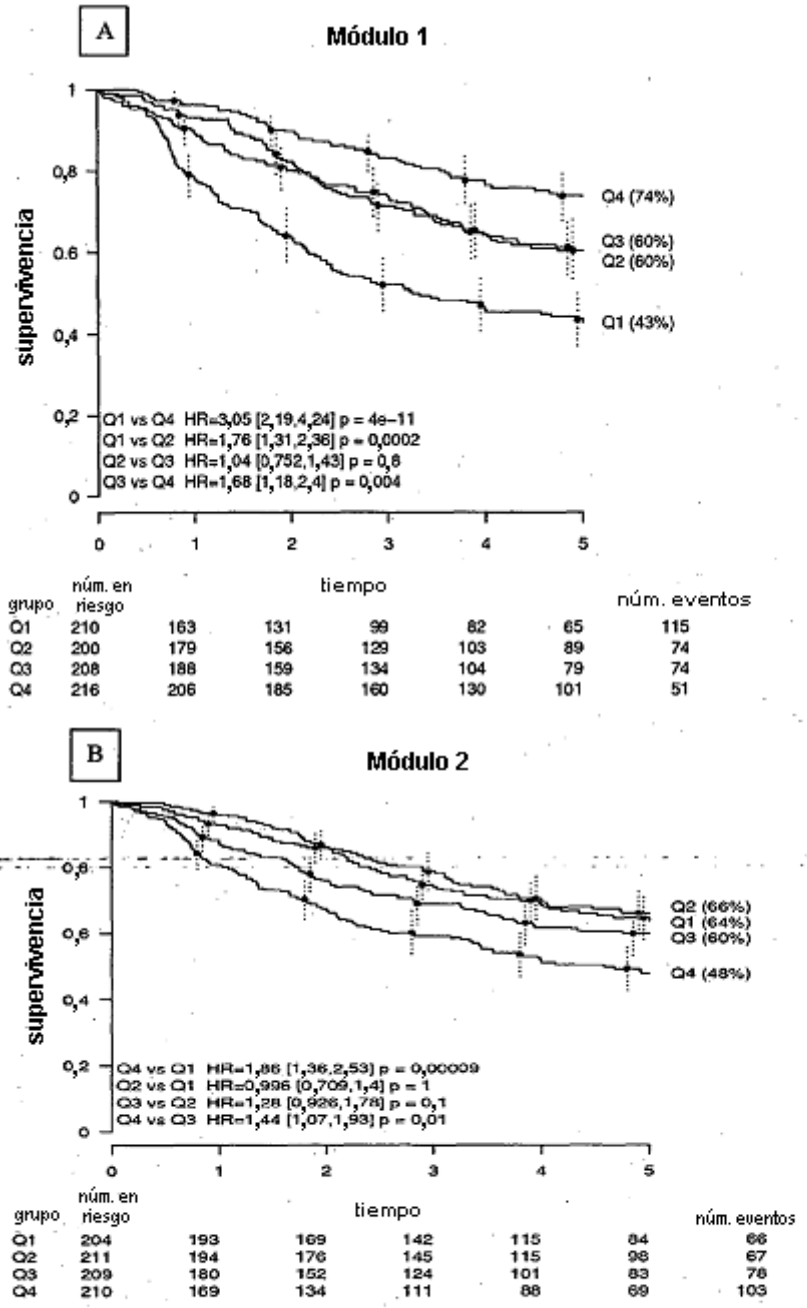

A
Módulo 1
supervivencia
1
0,8
0,6
0,4
0,2
0
Q4 (74%)
Q3 (60%)
Q2 (60%)
Q1 (43%)
Q1 vs Q4 HR=3,05 [2,19,4,24] p = 4e−11
Q1 vs Q2 HR=1,76 [1,31,2,36] p = 0,0002
Q2 vs Q3 HR=1,04 [0,752,1,43] p = 0,8
Q3 vs Q4 HR=1,68 [1,18,2,4] p = 0,004
0 1 2 3 4 5
tiempo
grupo núm. en riesgo
Q1 210 163 131 99 82 65 115
Q2 200 179 156 129 103 89 74
Q3 208 188 159 134 104 79 74
Q4 216 206 185 160 130 101 51
núm. eventos
B
Módulo 2
supervivencia
1
0,8
0,6
0,4
0,2
0
Q2 (66%)
Q1 (64%)
Q3 (60%)
Q4 (48%)
Q4 vs Q1 HR=1,86 [1,36,2,53] p = 0,00009
Q2 vs Q1 HR=0,996 [0,709,1,4] p = 1
Q3 vs Q2 HR=1,28 [0,926,1,78] p = 0,1
Q4 vs Q3 HR=1,44 [1,07,1,93] p = 0,01
0 1 2 3 4 5
tiempo
grupo núm. en riesgo
Q1 204 193 169 142 115 84 66
Q2 211 194 176 145 115 98 67
Q3 209 180 152 124 101 83 76
Q4 210 169 134 111 88 69 103
núm. eventos

Fig. 2

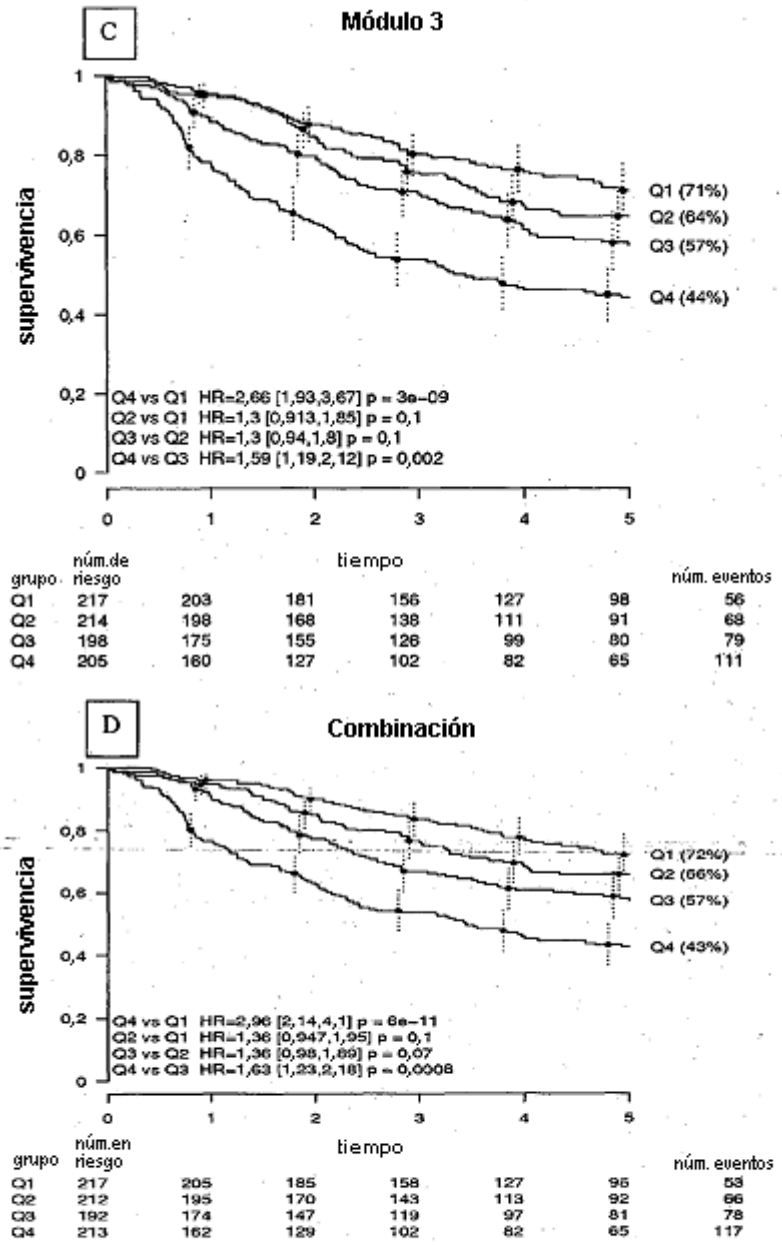

C
Módulo 3
supervivencia
1
0,8
0,6
0,4
0,2
0
Q1 (71%)
Q2 (64%)
Q3 (57%)
Q4 (44%)
Q4 vs Q1 HR=2,66 [1,93,3,67] p = 3e-09
Q2 vs Q1 HR=1,3 [0,913,1,85] p = 0,1
Q3 vs Q2 HR=1,3 [0,94,1,8] p = 0,1
Q4 vs Q3 HR=1,59 [1,19,2,12] p = 0,002
0 1 2 3 4 5
tiempo
grupo núm.de riesgo
Q1 217 203 181 156 127 98 56
Q2 214 198 168 138 111 91 68
Q3 198 175 155 126 99 80 79
Q4 205 160 127 102 82 65 111
núm. eventos
D
Combinación
supervivencia
1
0,8
0,6
0,4
0,2
0
Q1 (72%)
Q2 (66%)
Q3 (57%)
Q4 (43%)
Q4 vs Q1 HR=2,96 [2,14,4,1] p = 6e-11
Q2 vs Q1 HR=1,36 [0,947,1,95] p = 0,1
Q3 vs Q2 HR=1,36 [0,98,1,89] p = 0,07
Q4 vs Q3 HR=1,63 [1,23,2,16] p = 0,0008
0 1 2 3 4 5
tiempo
grupo núm.en riesgo
Q1 217 205 185 158 127 96 53
Q2 212 195 170 143 113 92 66
Q3 192 174 147 119 97 81 78
Q4 213 162 129 102 82 65 117
núm. eventos

Fig 2

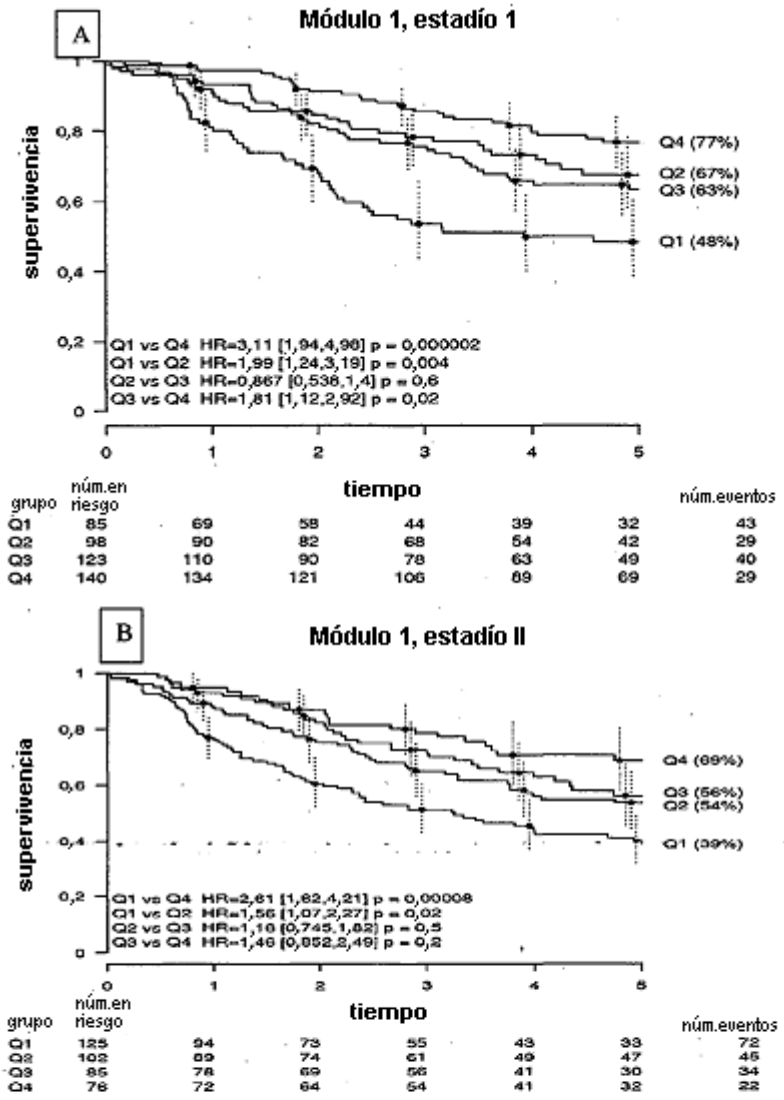

A
Módulo 1, estadío I
supervivencia
Q4 (77%)
Q2 (67%)
Q3 (63%)
Q1 (48%)
Q1 vs Q4 HR=3,11 [1,94,4,96] p = 0,000002
Q1 vs Q2 HR=1,99 [1,24,3,19] p = 0,004
Q2 vs Q3 HR=0,867 [0,536,1,4] p = 0,6
Q3 vs Q4 HR=1,81 [1,12,2,92] p = 0,02
tiempo
grupo núm.en riesgo núm.eventos
Q1 85 69 58 44 39 32 43
Q2 98 90 82 68 54 42 29
Q3 123 110 90 78 63 49 40
Q4 140 134 121 106 89 69 29
B
Módulo 1, estadío II
supervivencia
Q4 (69%)
Q3 (56%)
Q2 (54%)
Q1 (39%)
Q1 vs Q4 HR=2,61 [1,62,4,21] p = 0,00008
Q1 vs Q2 HR=1,56 [1,07,2,27] p = 0,02
Q2 vs Q3 HR=1,16 [0,745,1,82] p = 0,5
Q3 vs Q4 HR=1,46 [0,952,2,49] p = 0,2
tiempo
grupo núm.en riesgo núm.eventos
Q1 125 94 73 55 43 33 72
Q2 102 89 74 61 49 47 45
Q3 85 78 69 56 41 30 34
Q4 78 72 64 54 41 32 22

Fig. 3

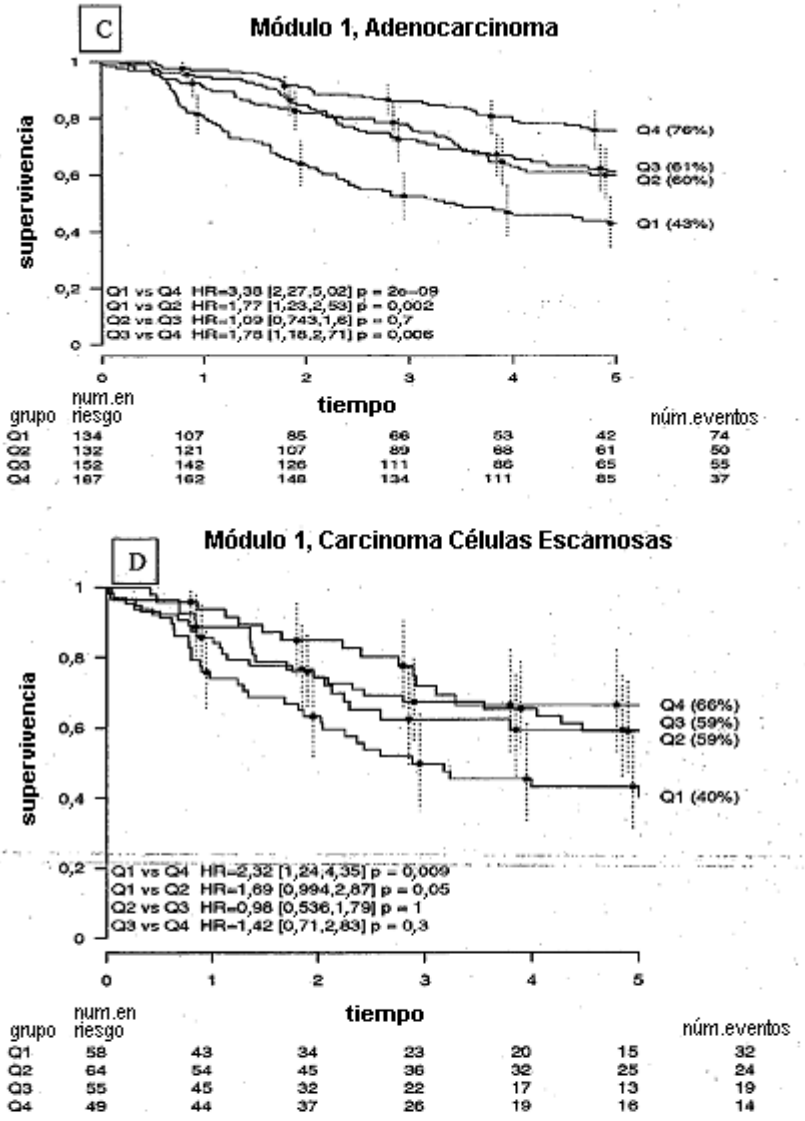

C
Módulo 1, Adenocarcinoma
supervivencia
Q4 (76%)
Q3 (61%)
Q2 (60%)
Q1 (43%)
Q1 vs Q4 HR=3,38 [2,27,5,02] p = 2e−09
Q1 vs Q2 HR=1,77 [1,23,2,53] p = 0,002
Q2 vs Q3 HR=1,09 [0,743,1,6] p = 0,7
Q3 vs Q4 HR=1,75 [1,18,2,71] p = 0,006
tiempo
grupo num.en riesgo núm.eventos
Q1 134 107 85 66 53 42 74
Q2 132 121 107 89 68 61 50
Q3 152 142 126 111 86 65 55
Q4 167 162 148 134 111 85 37
D
Módulo 1, Carcinoma Células Escamosas
supervivencia
Q4 (66%)
Q3 (59%)
Q2 (59%)
Q1 (40%)
Q1 vs Q4 HR=2,32 [1,24,4,35] p = 0,009
Q1 vs Q2 HR=1,69 [0,994,2,87] p = 0,05
Q2 vs Q3 HR=0,98 [0,536,1,79] p = 1
Q3 vs Q4 HR=1,42 [0,71,2,83] p = 0,3
tiempo
grupo num.en riesgo núm.eventos
Q1 58 43 34 23 20 15 32
Q2 64 54 45 36 32 25 24
Q3 55 45 32 22 17 13 19
Q4 49 44 37 26 19 16 14

Fig. 3

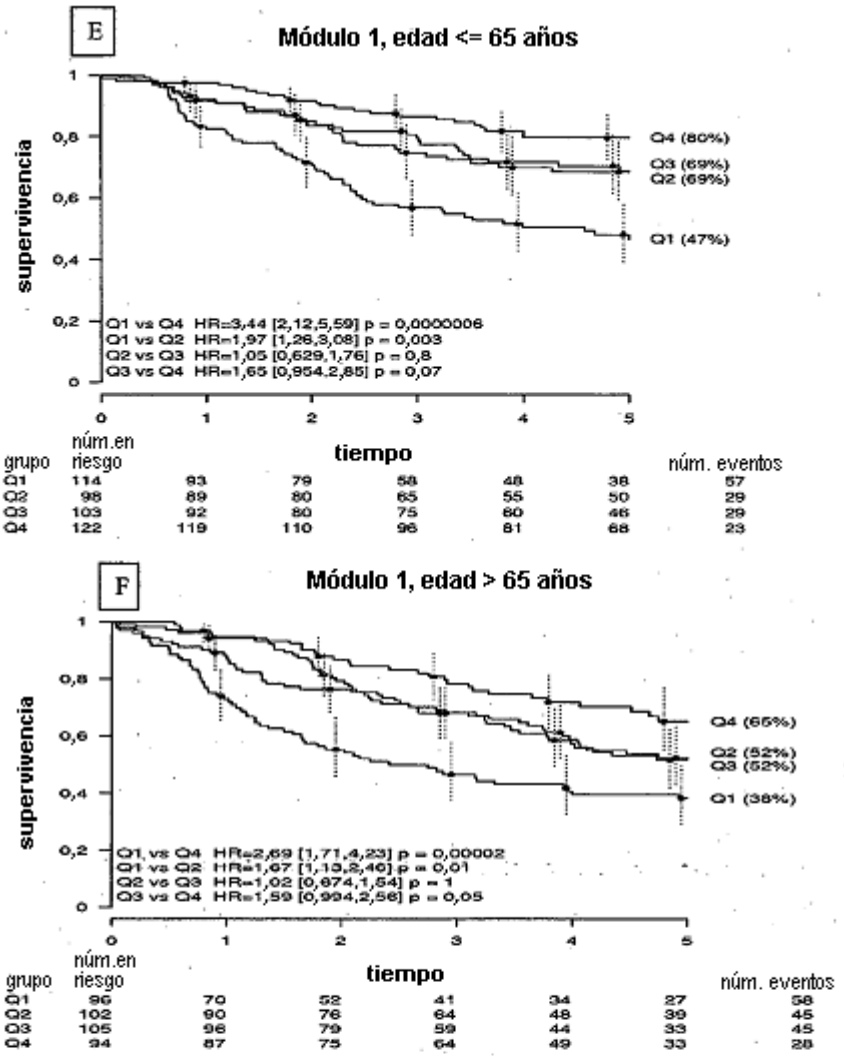

E
Módulo 1, edad <= 65 años
supervivencia
Q4 (80%)
Q3 (69%)
Q2 (69%)
Q1 (47%)
Q1 vs Q4 HR=3,44 [2,12,5,59] p = 0,0000006
Q1 vs Q2 HR=1,97 [1,26,3,08] p = 0,003
Q2 vs Q3 HR=1,05 [0,629,1,76] p = 0,8
Q3 vs Q4 HR=1,65 [0,954,2,85] p = 0,07
tiempo
grupo núm.en riesgo núm. eventos
Q1 114 93 79 58 48 38 57
Q2 98 89 80 65 55 50 29
Q3 103 92 80 75 60 46 29
Q4 122 119 110 98 81 68 23
F
Módulo 1, edad > 65 años
supervivencia
Q4 (65%)
Q2 (52%)
Q3 (52%)
Q1 (38%)
Q1 vs Q4 HR=2,69 [1,71,4,23] p = 0,00002
Q1 vs Q2 HR=1,67 [1,13,2,46] p = 0,01
Q2 vs Q3 HR=1,02 [0,674,1,54] p = 1
Q3 vs Q4 HR=1,59 [0,994,2,56] p = 0,05
tiempo
grupo núm.en riesgo núm. eventos
Q1 96 70 52 41 34 27 58
Q2 102 90 76 64 48 39 45
Q3 105 96 79 59 44 33 45
Q4 94 87 75 64 49 33 28

Fig. 3

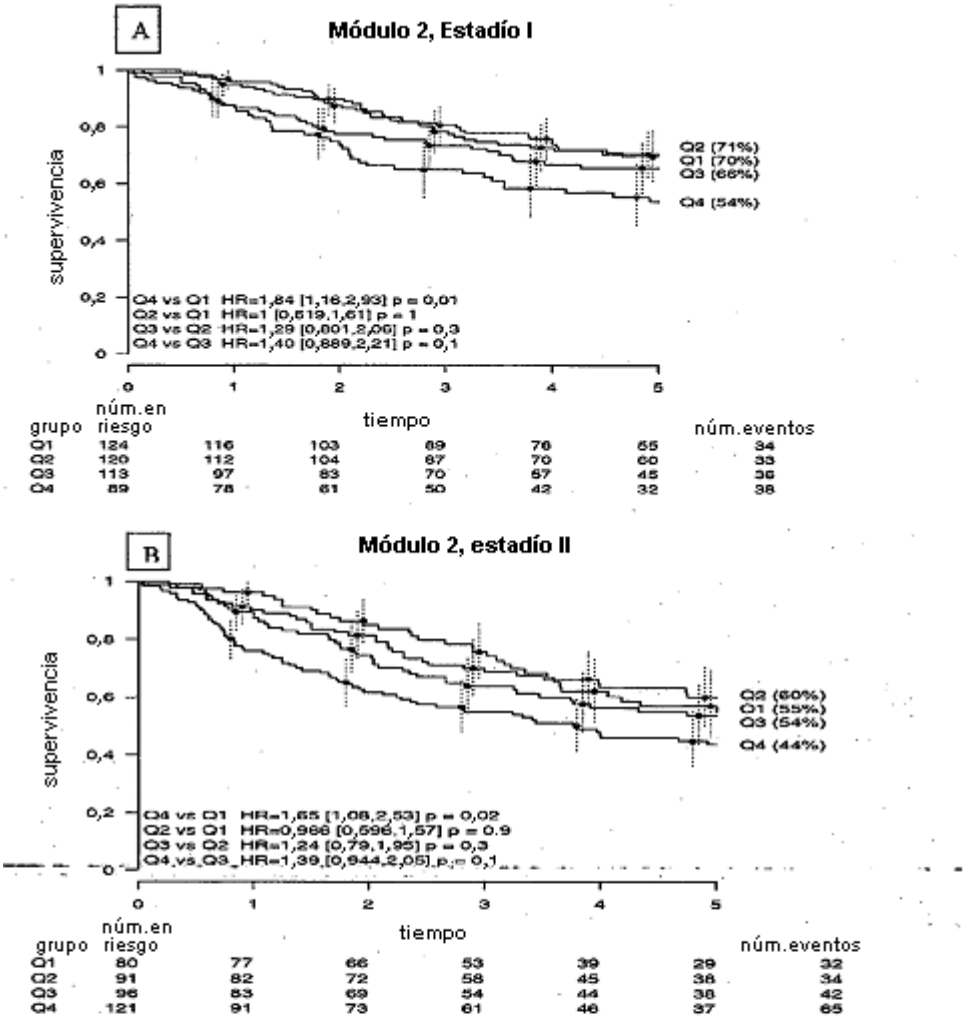

A
Módulo 2, Estadío I
supervivencia
Q2 (71%)
Q1 (70%)
Q3 (66%)
Q4 (54%)
Q4 vs Q1 HR=1,84 [1,16,2,93] p = 0,01
Q2 vs Q1 HR=1 [0,619,1,61] p = 1
Q3 vs Q2 HR=1,29 [0,801,2,06] p = 0,3
Q4 vs Q3 HR=1,40 [0,889,2,21] p = 0,1
tiempo
grupo núm.en riesgo núm.eventos
Q1 124 116 103 89 76 65 34
Q2 120 112 104 87 70 60 33
Q3 113 97 83 70 57 45 36
Q4 89 78 61 50 42 32 38
B
Módulo 2, estadío II
supervivencia
Q2 (60%)
Q1 (55%)
Q3 (54%)
Q4 (44%)
Q4 vs Q1 HR=1,65 [1,08,2,53] p = 0,02
Q2 vs Q1 HR=0,986 [0,596,1,57] p = 0,9
Q3 vs Q2 HR=1,24 [0,79,1,95] p = 0,3
Q4 vs Q3 HR=1,39 [0,944,2,05] p = 0,1
tiempo
grupo núm.en riesgo núm.eventos
Q1 80 77 66 53 39 29 32
Q2 91 82 72 58 45 38 34
Q3 96 83 69 54 44 38 42
Q4 121 91 73 61 46 37 65

Fig. 4

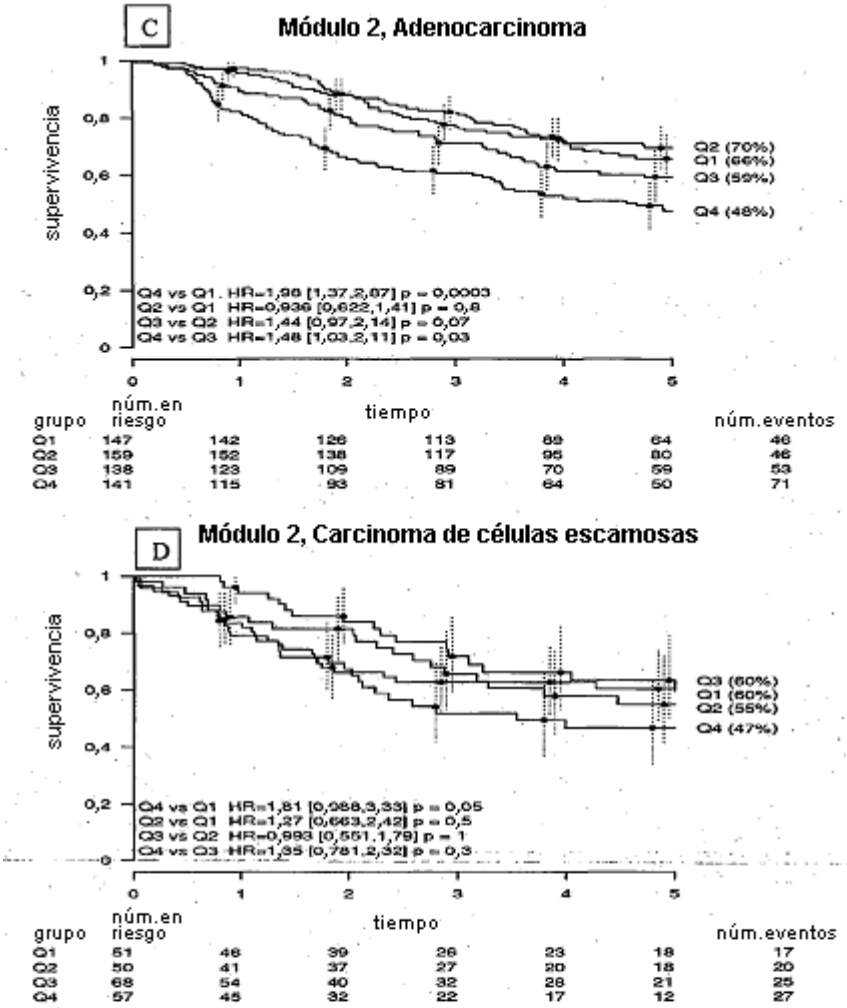
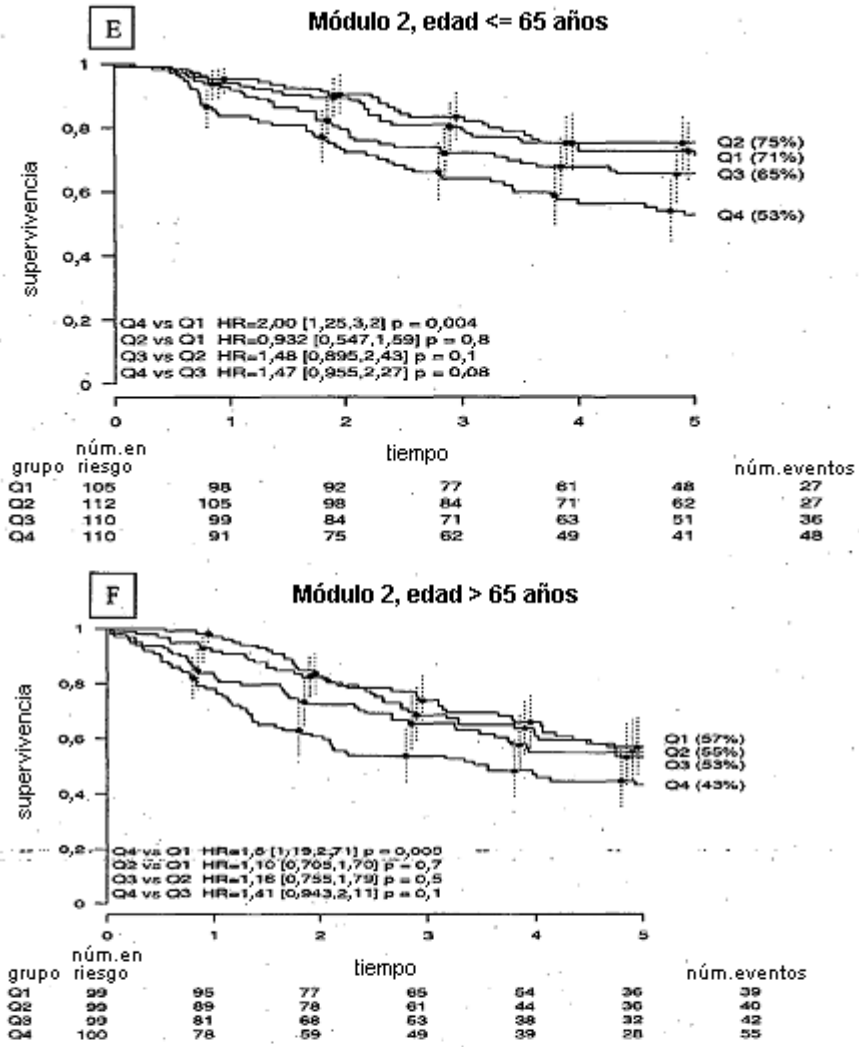

C
Módulo 2, Adenocarcinoma
supervivencia
Q2 (70%)
Q1 (66%)
Q3 (59%)
Q4 (48%)
Q4 vs Q1. HR=1,98 [1,37,2,87] p = 0,0003
Q2 vs Q1 HR=0,936 [0,622,1,41] p = 0,8
Q3 vs Q2 HR=1,44 [0,97,2,14] p = 0,07
Q4 vs Q3 HR=1,48 [1,03,2,11] p = 0,03
tiempo
grupo
núm.en riesgo
núm.eventos
Q1 147 142 126 113 89 64 46
Q2 159 152 138 117 95 80 46
Q3 138 123 109 89 70 59 53
Q4 141 115 93 81 64 50 71
D
Módulo 2, Carcinoma de células escamosas
supervivencia
Q3 (60%)
Q1 (60%)
Q2 (55%)
Q4 (47%)
Q4 vs Q1 HR=1,81 [0,988,3,33] p = 0,05
Q2 vs Q1 HR=1,27 [0,663,2,42] p = 0,5
Q3 vs Q2 HR=0,993 [0,551,1,79] p = 1
Q4 vs Q3 HR=1,35 [0,781,2,32] p = 0,3
tiempo
grupo
núm.en riesgo
núm.eventos
Q1 51 46 39 26 23 18 17
Q2 50 41 37 27 20 18 20
Q3 68 54 40 32 28 21 25
Q4 57 45 32 22 17 12 27

Fig. 4


E
Módulo 2, edad <= 65 años
supervivencia
1
0,8
0,6
0,4
0,2
0
Q2 (75%)
Q1 (71%)
Q3 (65%)
Q4 (53%)
Q4 vs Q1 HR=2,00 [1,25,3,2] p = 0,004
Q2 vs Q1 HR=0,932 [0,547,1,59] p = 0,8
Q3 vs Q2 HR=1,48 [0,895,2,43] p = 0,1
Q4 vs Q3 HR=1,47 [0,955,2,27] p = 0,08
0 1 2 3 4 5
tiempo
grupo núm.en riesgo
Q1 105 98 92 77 61 48
Q2 112 105 98 84 71 62
Q3 110 99 84 71 63 51
Q4 110 91 75 62 49 41
núm.eventos
27
27
36
48

F
Módulo 2, edad > 65 años
supervivencia
1
0,8
0,6
0,4
0,2
0
Q1 (57%)
Q2 (55%)
Q3 (53%)
Q4 (43%)
Q4 vs Q1 HR=1,6 [1,19,2,71] p = 0,005
Q2 vs Q1 HR=1,10 [0,705,1,70] p = 0,7
Q3 vs Q2 HR=1,16 [0,755,1,79] p = 0,5
Q4 vs Q3 HR=1,41 [0,943,2,11] p = 0,1
0 1 2 3 4 5
tiempo
grupo núm.en riesgo
Q1 99 95 77 65 54 36
Q2 99 89 78 61 44 30
Q3 99 81 68 53 38 32
Q4 100 78 59 49 39 28
núm.eventos
39
40
42
55

Fig. 4

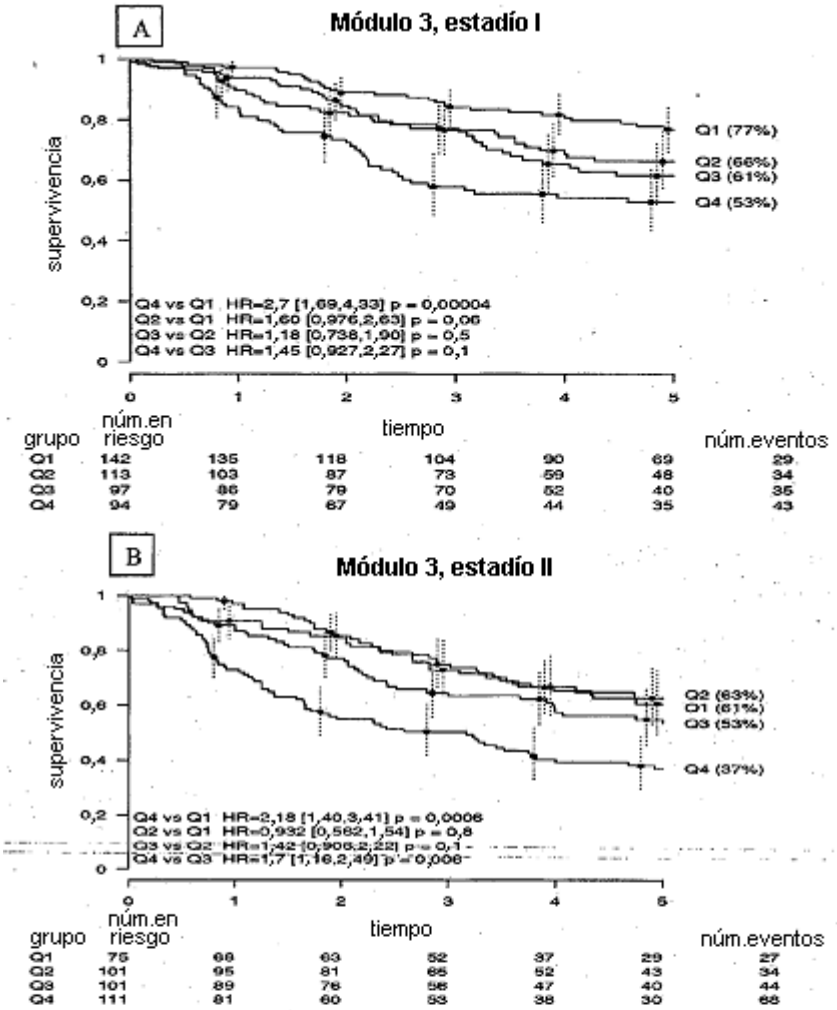
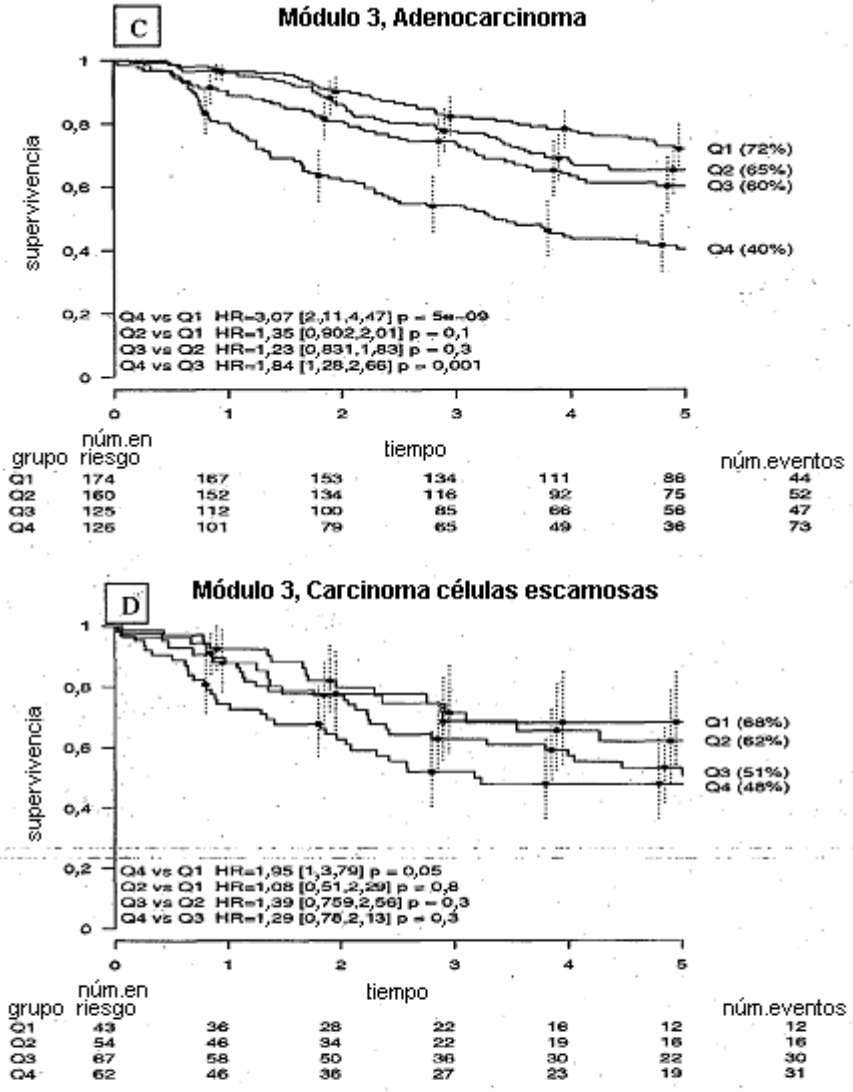

A
Módulo 3, estadío I
supervivencia
Q1 (77%)
Q2 (66%)
Q3 (61%)
Q4 (53%)
Q4 vs Q1 HR=2,7 [1,69,4,33] p = 0,00004
Q2 vs Q1 HR=1,60 [0,976,2,63] p = 0,06
Q3 vs Q2 HR=1,18 [0,738,1,90] p = 0,5
Q4 vs Q3 HR=1,45 [0,927,2,27] p = 0,1
tiempo
grupo núm.en riesgo núm.eventos
Q1 142 135 118 104 90 69 29
Q2 113 103 87 78 69 48 34
Q3 97 86 79 70 52 40 35
Q4 94 79 67 49 44 35 43
B
Módulo 3, estadío II
supervivencia
Q2 (63%)
Q1 (61%)
Q3 (53%)
Q4 (37%)
Q4 vs Q1 HR=2,18 [1,40,3,41] p = 0,0006
Q2 vs Q1 HR=0,932 [0,582,1,54] p = 0,8
Q3 vs Q2 HR=1,42 [0,906,2,22] p = 0,1
Q4 vs Q3 HR=1,7 [1,16,2,49] p = 0,006
tiempo
grupo núm.en riesgo núm.eventos
Q1 75 66 63 52 37 29 27
Q2 101 95 81 65 52 43 34
Q3 101 89 76 58 47 40 44
Q4 111 91 60 53 38 30 68

Fig. 5


C
Módulo 3, Adenocarcinoma
supervivencia
1
0,8
0,6
0,4
0,2
0
Q1 (72%)
Q2 (65%)
Q3 (60%)
Q4 (40%)
Q4 vs Q1 HR=3,07 [2,11,4,47] p = 5e-09
Q2 vs Q1 HR=1,35 [0,902,2,01] p = 0,1
Q3 vs Q2 HR=1,23 [0,831,1,83] p = 0,3
Q4 vs Q3 HR=1,84 [1,28,2,66] p = 0,001
0 1 2 3 4 5
tiempo
grupo núm.en riesgo
Q1 174 167 153 134 111 88
Q2 160 152 134 116 92 75
Q3 125 112 100 85 66 58
Q4 126 101 79 65 49 36
núm.eventos
44
52
47
73

D
Módulo 3, Carcinoma células escamosas
supervivencia
1
0,8
0,6
0,4
0,2
0
Q1 (68%)
Q2 (62%)
Q3 (51%)
Q4 (48%)
Q4 vs Q1 HR=1,95 [1,3,79] p = 0,05
Q2 vs Q1 HR=1,08 [0,51,2,29] p = 0,8
Q3 vs Q2 HR=1,39 [0,759,2,56] p = 0,3
Q4 vs Q3 HR=1,29 [0,76,2,13] p = 0,3
0 1 2 3 4 5
tiempo
grupo núm.en riesgo
Q1 43 36 28 22 16 12
Q2 54 46 34 22 19 16
Q3 67 58 50 36 30 22
Q4 62 46 36 27 23 19
núm.eventos
12
16
30
31

Fig. 5

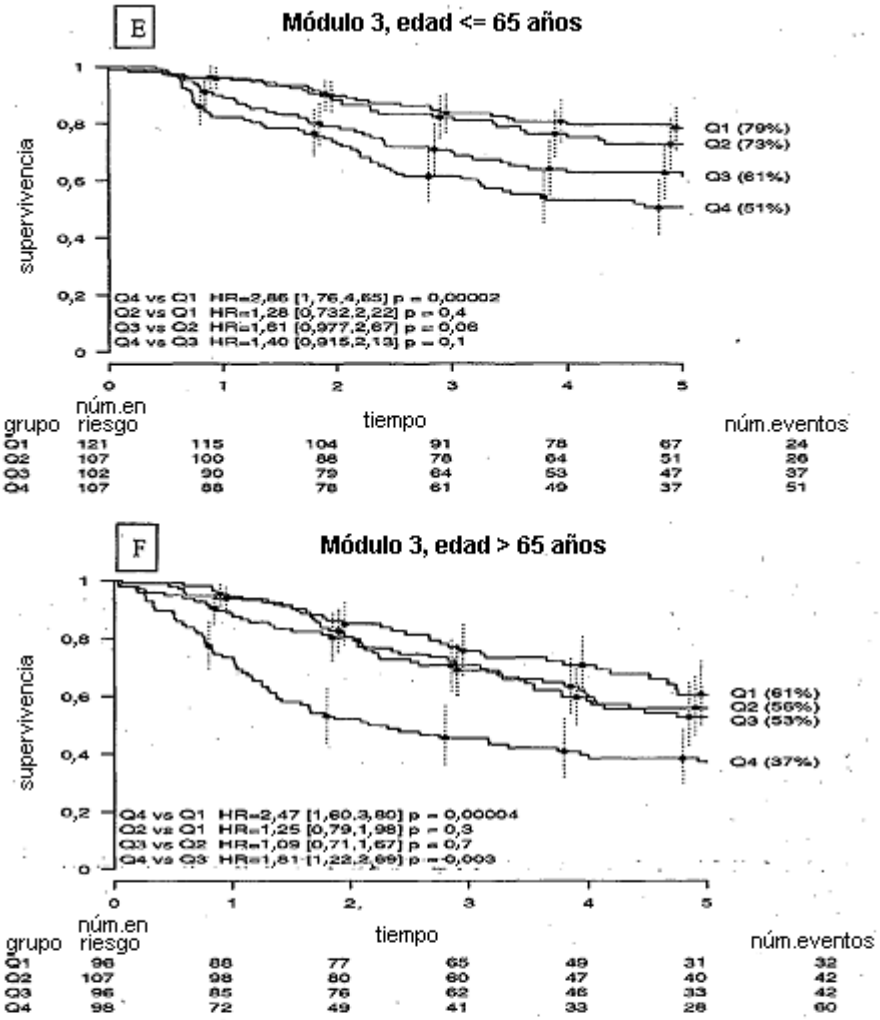

E
Módulo 3, edad <= 65 años
supervivencia
Q1 (79%)
Q2 (73%)
Q3 (61%)
Q4 (51%)
Q4 vs Q1 HR=2,86 [1,76,4,65] p = 0,00002
Q2 vs Q1 HR=1,28 [0,732,2,22] p = 0,4
Q3 vs Q2 HR=1,61 [0,977,2,67] p = 0,06
Q4 vs Q3 HR=1,40 [0,915,2,13] p = 0,1
tiempo
grupo
núm.en riesgo
núm.eventos
Q1 121 115 104 91 78 67 24
Q2 107 100 88 76 64 51 26
Q3 102 90 79 64 53 47 37
Q4 107 88 78 61 49 37 51
F
Módulo 3, edad > 65 años
supervivencia
Q1 (61%)
Q2 (56%)
Q3 (53%)
Q4 (37%)
Q4 vs Q1 HR=2,47 [1,60,3,80] p = 0,00004
Q2 vs Q1 HR=1,25 [0,79,1,98] p = 0,3
Q3 vs Q2 HR=1,09 [0,71,1,67] p = 0,7
Q4 vs Q3 HR=1,81 [1,22,2,69] p = 0,003
tiempo
grupo
núm.en riesgo
núm.eventos
Q1 96 88 77 65 49 31 32
Q2 107 98 80 60 47 40 42
Q3 96 85 76 62 46 33 42
Q4 98 72 49 41 33 28 60

Fig. 5

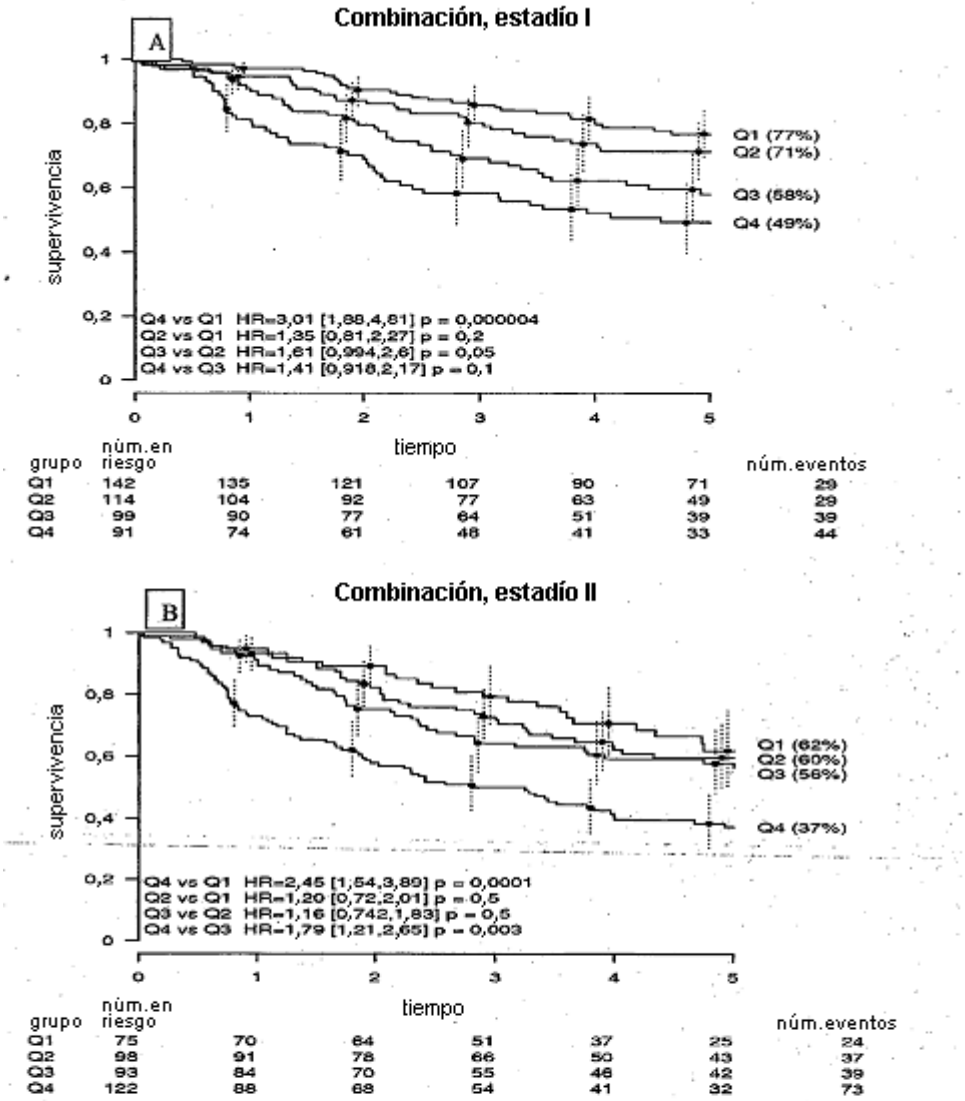
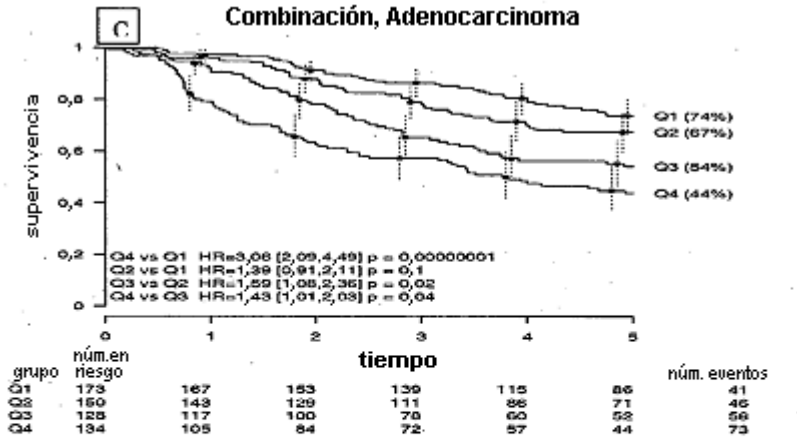
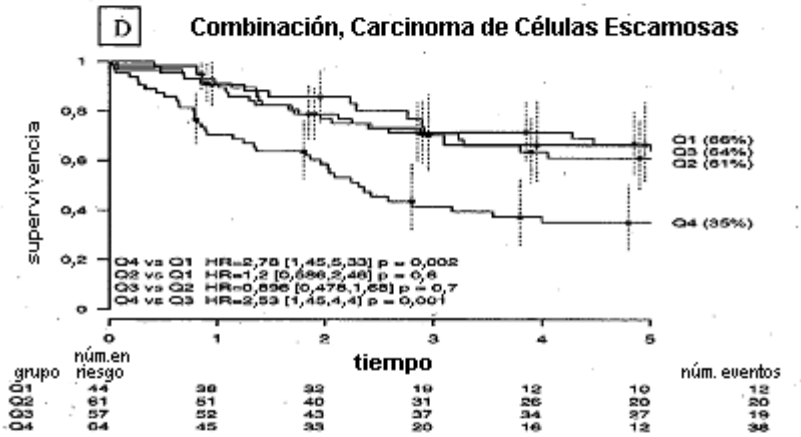

A
Combinación, estadío I
supervivencia
1
0,8
0,6
0,4
0,2
0
Q1 (77%)
Q2 (71%)
Q3 (58%)
Q4 (49%)
Q4 vs Q1 HR=3,01 [1,88,4,81] p = 0,000004
Q2 vs Q1 HR=1,35 [0,81,2,27] p = 0,2
Q3 vs Q2 HR=1,61 [0,994,2,6] p = 0,05
Q4 vs Q3 HR=1,41 [0,916,2,17] p = 0,1
0 1 2 3 4 5
tiempo
grupo núm.en riesgo núm.eventos
Q1 142 135 121 107 90 71 29
Q2 114 104 92 77 63 49 29
Q3 99 90 77 64 51 39 39
Q4 91 74 61 48 41 33 44
B
Combinación, estadío II
supervivencia
1
0,8
0,6
0,4
0,2
0
Q1 (62%)
Q2 (60%)
Q3 (56%)
Q4 (37%)
Q4 vs Q1 HR=2,45 [1,54,3,89] p = 0,0001
Q2 vs Q1 HR=1,20 [0,72,2,01] p = 0,5
Q3 vs Q2 HR=1,16 [0,742,1,83] p = 0,5
Q4 vs Q3 HR=1,79 [1,21,2,65] p = 0,003
0 1 2 3 4 5
tiempo
grupo núm.en riesgo núm.eventos
Q1 75 70 64 51 37 25 24
Q2 98 91 78 66 50 43 37
Q3 93 84 70 55 46 42 39
Q4 122 88 68 54 41 32 73

Fig. 6


C
Combinación, Adenocarcinoma
supervivencia
Q1 (74%)
Q2 (67%)
Q3 (54%)
Q4 (44%)
Q4 vs Q1 HR=3,06 [2,09,4,49] p = 0,00000001
Q2 vs Q1 HR=1,39 [0,91,2,11] p = 0,1
Q3 vs Q2 HR=1,59 [1,08,2,36] p = 0,02
Q4 vs Q3 HR=1,43 [1,01,2,03] p = 0,04
tiempo
grupo núm.en riesgo núm. eventos
Q1 173 167 153 139 115 86 41
Q2 150 143 129 111 86 71 46
Q3 128 117 100 78 60 52 58
Q4 134 105 84 72 57 44 73


D
Combinación, Carcinoma de Células Escamosas
supervivencia
Q1 (66%)
Q3 (64%)
Q2 (61%)
Q4 (35%)
Q4 vs Q1 HR=2,78 [1,45,5,33] p = 0,002
Q2 vs Q1 HR=1,2 [0,686,2,46] p = 0,6
Q3 vs Q2 HR=0,896 [0,476,1,65] p = 0,7
Q4 vs Q3 HR=2,53 [1,45,4,4] p = 0,001
tiempo
grupo núm.en riesgo núm. eventos
Q1 44 38 32 19 12 10 12
Q2 61 51 40 31 26 20 20
Q3 57 52 43 37 34 27 19
Q4 64 45 33 20 16 12 38

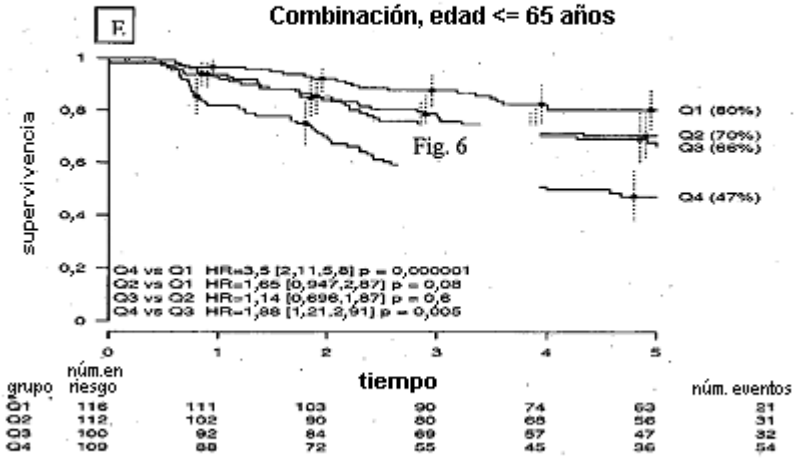
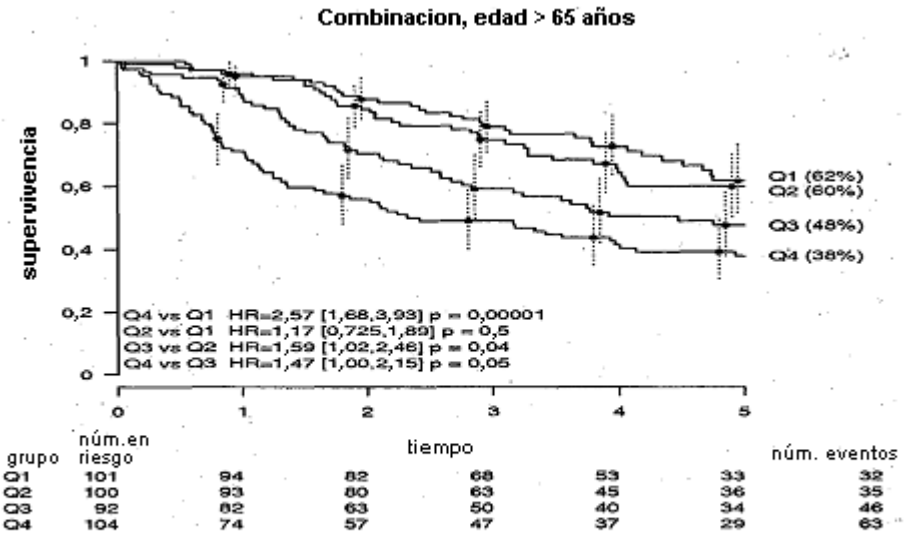

E
Combinación, edad <= 65 años
supervivencia
Q1 (80%)
Q2 (70%)
Q3 (66%)
Q4 (47%)
Q4 vs Q1 HR=3,5 [2,11,5,8] p = 0,000001
Q2 vs Q1 HR=1,65 [0,947,2,87] p = 0,08
Q3 vs Q2 HR=1,14 [0,696,1,87] p = 0,6
Q4 vs Q3 HR=1,88 [1,21,2,91] p = 0,005
tiempo
grupo núm.en riesgo núm. eventos
Q1 116 111 103 90 74 63 21
Q2 112 102 90 80 65 56 31
Q3 100 92 84 69 57 47 32
Q4 109 88 72 55 45 36 54

Fig. 6

F



grupo	núm.en riesgo						núm. eventos
Q1	101	94	82	68	53	33	32
Q2	100	93	80	63	45	36	35
Q3	92	82	63	50	40	34	46
Q4	104	74	57	47	37	29	63

Fig. 6

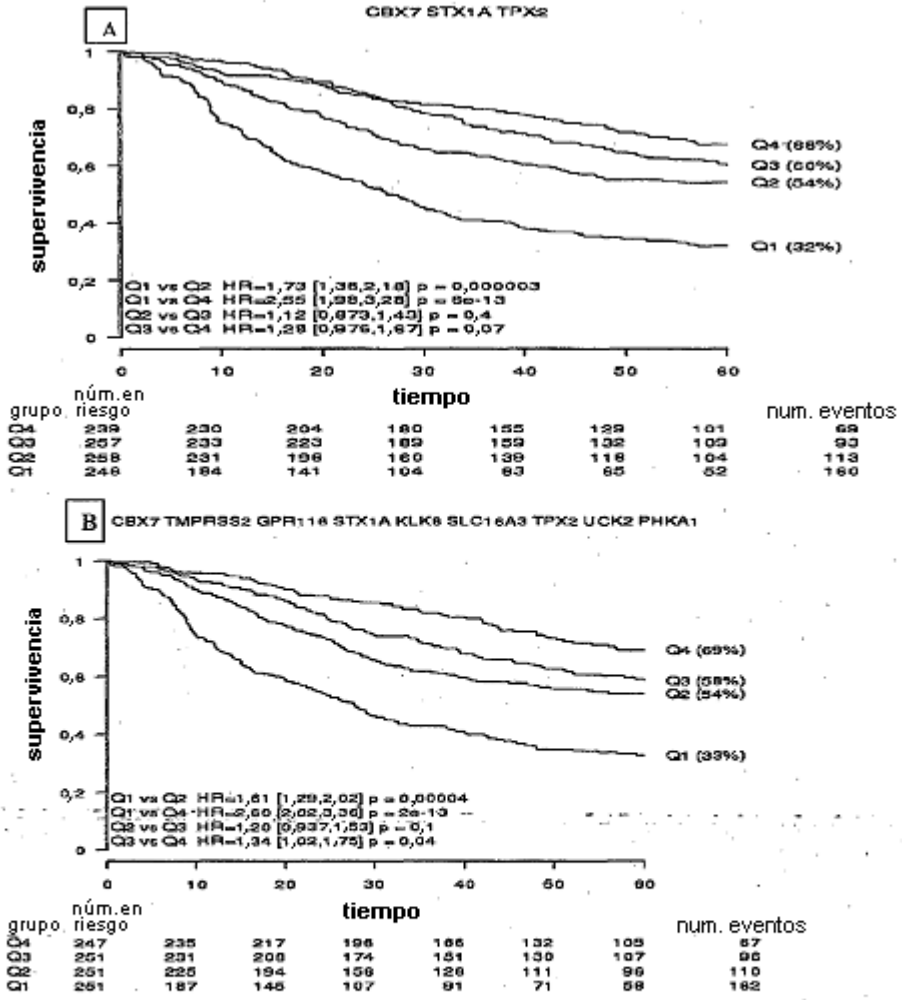

A
CBX7 STX1A TPX2
supervivencia
Q4 (68%)
Q3 (60%)
Q2 (54%)
Q1 (32%)
Q1 vs Q2 HR=1,73 [1,35,2,18] p = 0,000003
Q1 vs Q4 HR=2,55 [1,98,3,28] p = 6e-13
Q2 vs Q3 HR=1,12 [0,873,1,43] p = 0,4
Q3 vs Q4 HR=1,28 [0,976,1,67] p = 0,07
tiempo
grupo núm.en riesgo num. eventos
Q4 239 230 204 180 155 129 101 69
Q3 257 233 223 182 159 132 109 93
Q2 258 231 196 160 139 118 104 113
Q1 248 184 141 104 83 65 52 160
B
CBX7 TMPRSS2 GPR116 STX1A KLK6 SLC16A3 TPX2 UCK2 PHKA1
supervivencia
Q4 (69%)
Q3 (58%)
Q2 (54%)
Q1 (33%)
Q1 vs Q2 HR=1,61 [1,29,2,02] p = 0,00004
Q1 vs Q4 HR=2,60 [2,02,3,36] p = 2e-13
Q2 vs Q3 HR=1,20 [0,937,1,53] p = 0,1
Q3 vs Q4 HR=1,34 [1,02,1,75] p = 0,04
tiempo
grupo núm.en riesgo num. eventos
Q4 247 235 217 196 166 132 105 67
Q3 251 231 208 174 151 130 107 96
Q2 251 225 194 159 129 111 90 110
Q1 251 187 146 107 91 71 58 182

Fig. 7

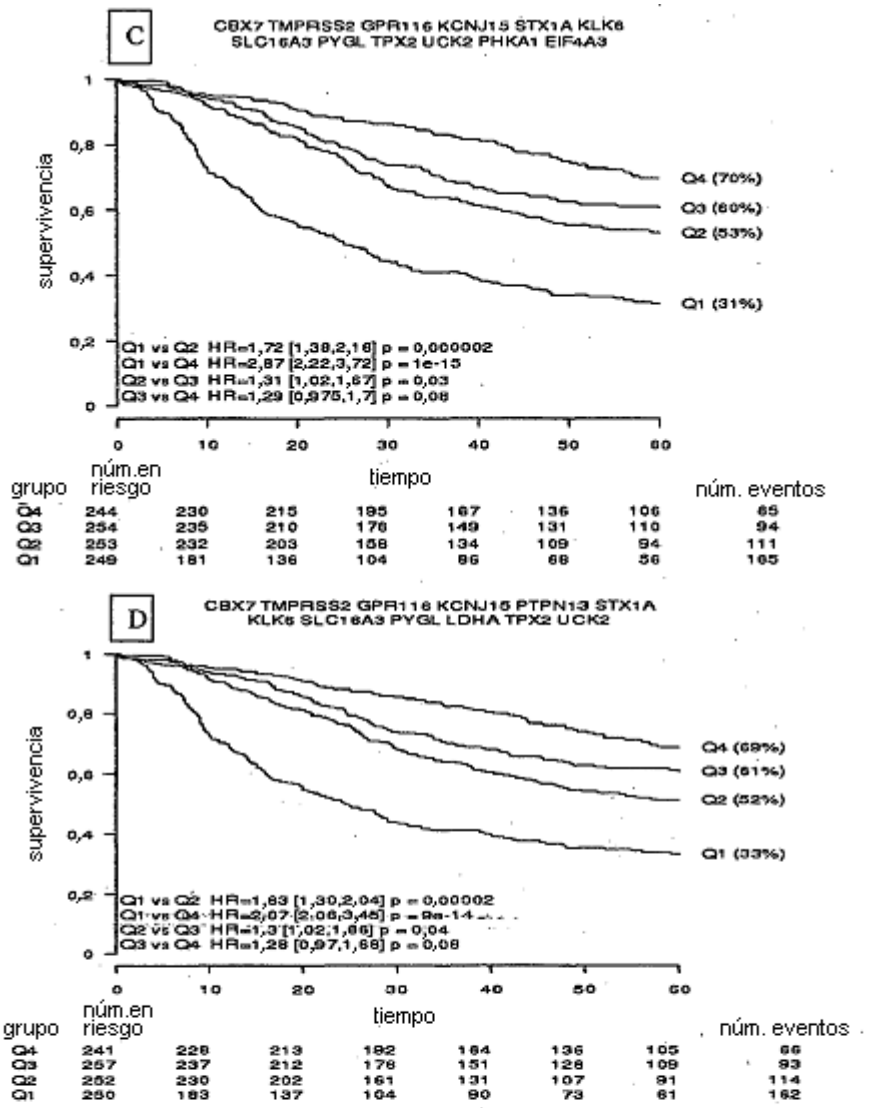

C
CBX7 TMPRSS2 GPR116 KCNJ15 STX1A KLK6
SLC16A3 PYGL TPX2 UCK2 PHKA1 EIF4A3
supervivencia
Q4 (70%)
Q3 (60%)
Q2 (53%)
Q1 (31%)
Q1 vs Q2 HR=1,72 [1,38,2,16] p = 0,000002
Q1 vs Q4 HR=2,87 [2,22,3,72] p = 1e-15
Q2 vs Q3 HR=1,31 [1,02,1,67] p = 0,03
Q3 vs Q4 HR=1,29 [0,975,1,7] p = 0,08
tiempo
grupo núm.en riesgo núm. eventos
Q4 244 230 215 195 167 136 106 65
Q3 254 235 210 176 149 131 110 94
Q2 253 232 203 158 134 109 94 111
Q1 249 181 136 104 86 68 56 165
D
CBX7 TMPRSS2 GPR116 KCNJ15 PTPN13 STX1A
KLK6 SLC16A3 PYGL LDHA TPX2 UCK2
supervivencia
Q4 (69%)
Q3 (61%)
Q2 (52%)
Q1 (33%)
Q1 vs Q2 HR=1,63 [1,30,2,04] p = 0,00002
Q1 vs Q4 HR=2,07 [2,06,3,45] p = 9e-14
Q2 vs Q3 HR=1,3 [1,02,1,66] p = 0,04
Q3 vs Q4 HR=1,28 [0,97,1,68] p = 0,08
tiempo
grupo núm.en riesgo núm. eventos
Q4 241 228 213 182 164 136 105 66
Q3 257 237 212 176 151 128 109 93
Q2 252 230 202 161 131 107 91 114
Q1 250 183 137 104 90 73 61 162

Fig. 7

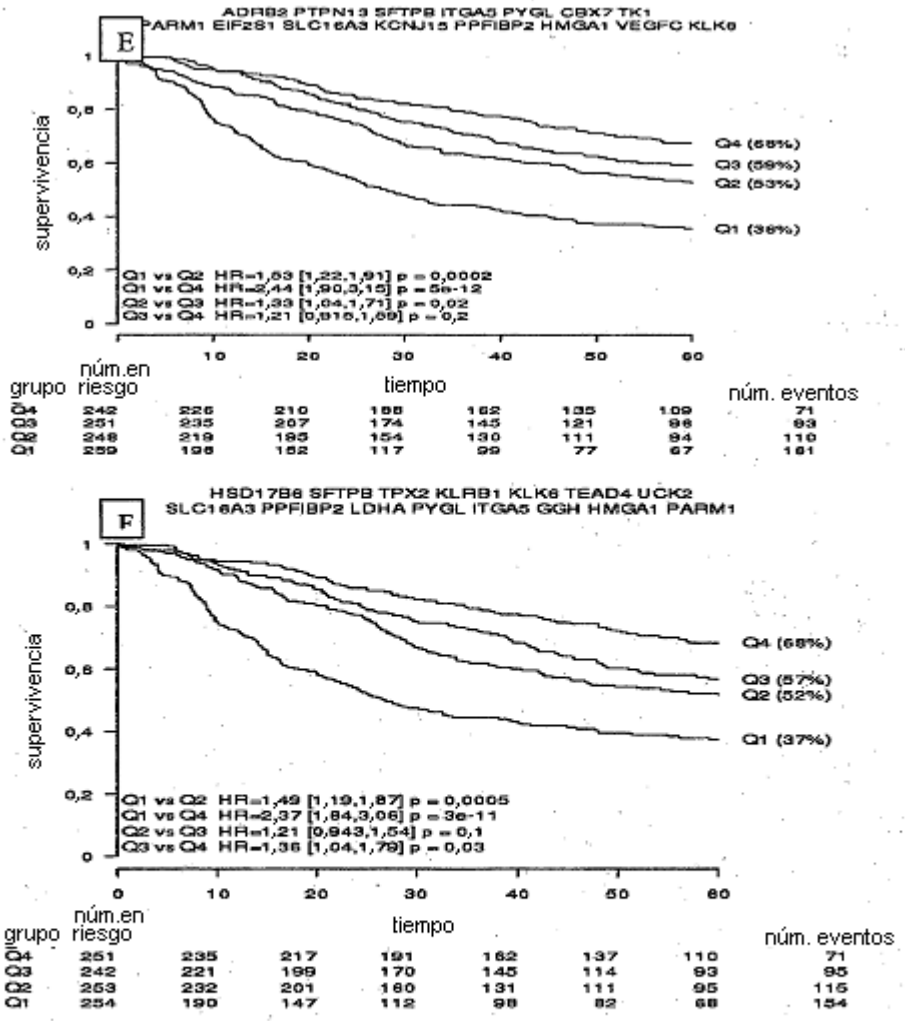

ADRB2 PTPN13 SFTPB ITGA5 PYGL CBX7 TK1
PARM1 EIF2S1 SLC16A3 KCNJ15 PPFIBP2 HMGA1 VEGFC KLK6
E
supervivencia
Q4 (68%)
Q3 (59%)
Q2 (53%)
Q1 (36%)
Q1 vs Q2 HR=1,63 [1,22,1,91] p = 0,0002
Q1 vs Q4 HR=2,44 [1,90,3,15] p = 5e-12
Q2 vs Q3 HR=1,33 [1,04,1,71] p = 0,02
Q3 vs Q4 HR=1,21 [0,910,1,69] p = 0,2
tiempo
grupo
núm.en riesgo
núm. eventos
Q4 242 226 210 188 162 135 109 71
Q3 251 235 207 174 145 121 96 83
Q2 248 219 185 154 130 111 94 110
Q1 259 196 152 117 99 77 67 161
HSD17B6 SFTPB TPX2 KLRB1 KLK6 TEAD4 UCK2
SLC16A3 PPFIBP2 LDHA PYGL ITGA5 GGH HMGA1 PARM1
F
supervivencia
Q4 (68%)
Q3 (57%)
Q2 (52%)
Q1 (37%)
Q1 vs Q2 HR=1,49 [1,19,1,87] p = 0,0005
Q1 vs Q4 HR=2,37 [1,84,3,06] p = 3e-11
Q2 vs Q3 HR=1,21 [0,943,1,54] p = 0,1
Q3 vs Q4 HR=1,36 [1,04,1,79] p = 0,03
tiempo
grupo
núm.en riesgo
núm. eventos
Q4 251 235 217 191 162 137 110 71
Q3 242 221 199 170 145 114 93 95
Q2 253 232 201 160 131 111 95 115
Q1 254 190 147 112 98 82 68 154

Fig 7

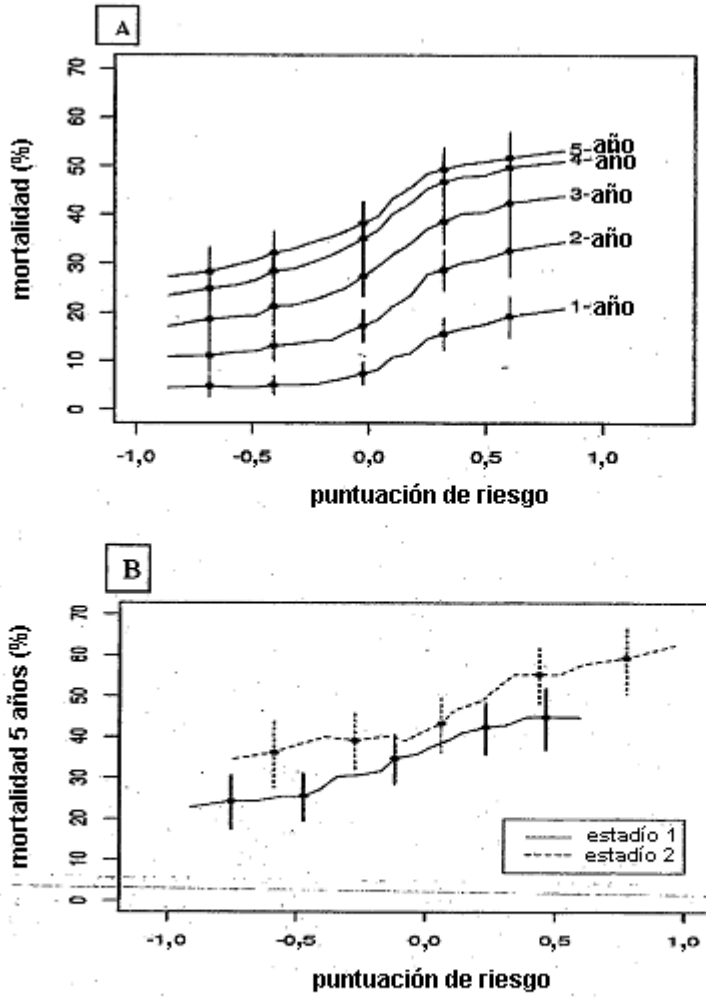

A
mortalidad (%)
0
10
20
30
40
50
60
70
-1,0
-0,5
0,0
0,5
1,0
puntuación de riesgo
5-año
4-año
3-año
2-año
1-año
B
mortalidad 5 años (%)
estadío 1
estadío 2

Fig. 8

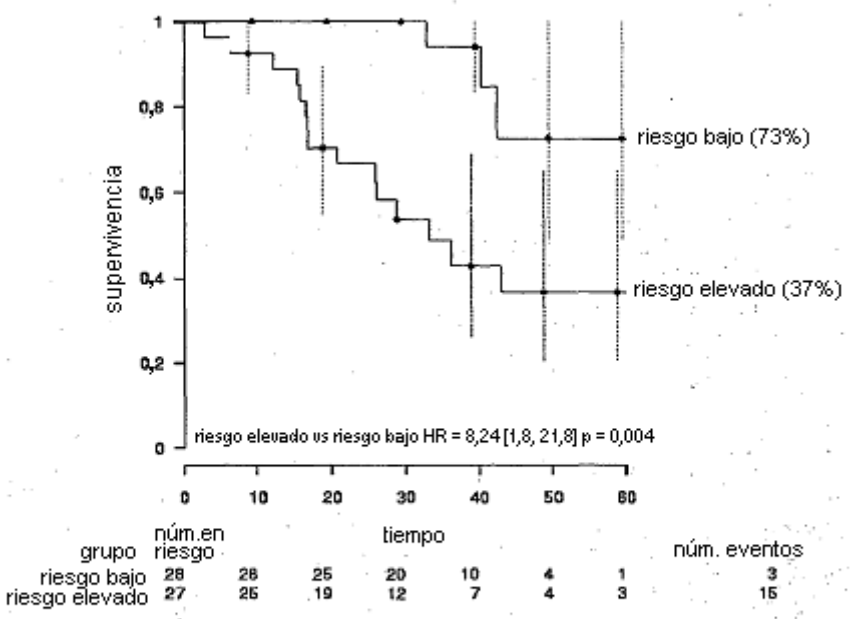

supervivencia
1
0,8
0,6
0,4
0,2
0
riesgo bajo (73%)
riesgo elevado (37%)
riesgo elevado vs riesgo bajo HR = 8,24 [1,8, 21,8] p = 0,004
0
10
20
30
40
50
60
tiempo
grupo
núm.en riesgo
núm. eventos
riesgo bajo 28 28 25 20 10 4 1 3
riesgo elevado 27 25 19 12 7 4 3 15

Fig. 9

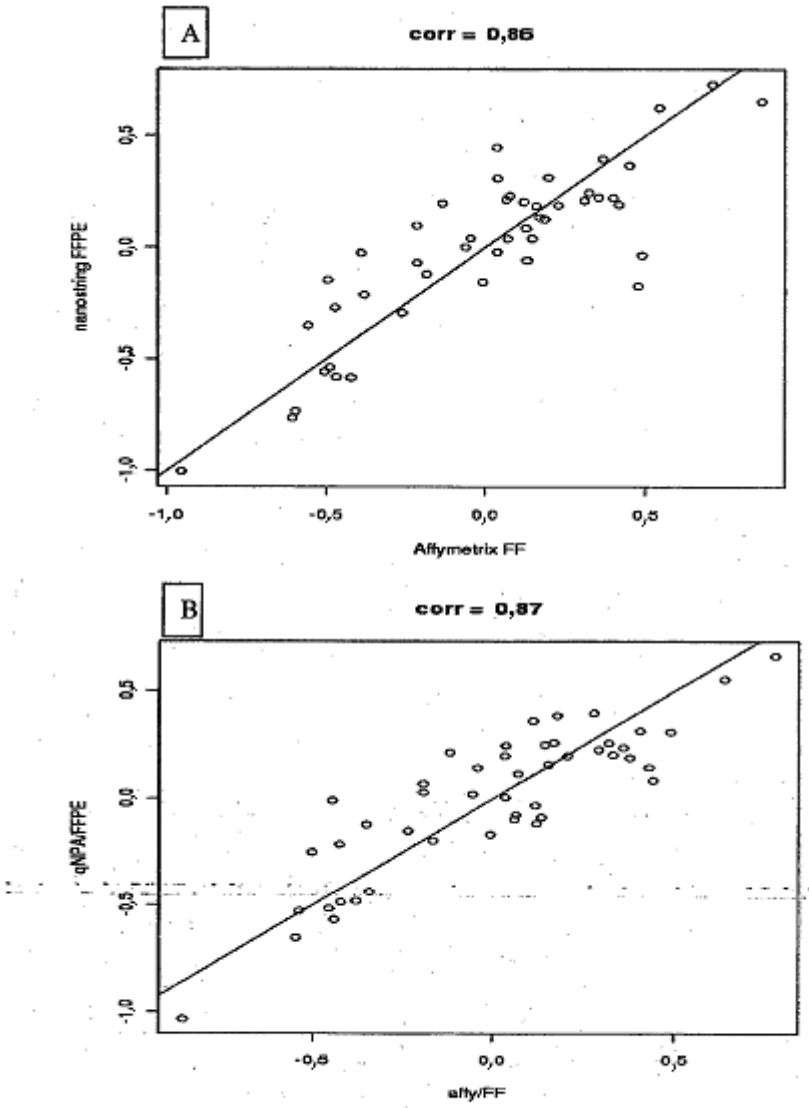

A
corr = 0,86
nanostring FFPE
Affymetrix FF
B
corr = 0,87
qNPA/FFPE
affy/FF

Fig. 10