

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 639 559**

51 Int. Cl.:

G01N 33/574 (2006.01)

G06F 19/00 (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.12.2013 PCT/EP2013/077295**

87 Fecha y número de publicación internacional: **03.07.2014 WO14102130**

96 Fecha de presentación y número de la solicitud europea: **19.12.2013 E 13815724 (3)**

97 Fecha y número de publicación de la concesión europea: **05.07.2017 EP 2939026**

54 Título: **Análisis de imágenes para el pronóstico de cáncer de mama**

30 Prioridad:

28.12.2012 US 201261747148 P

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

27.10.2017

73 Titular/es:

**VENTANA MEDICAL SYSTEMS, INC. (33.3%)
1910 E. Innovation Park Drive
Tucson, Arizona 85755, US;
CLEVELAND CLINIC FOUNDATION (33.3%) y
THE UNIVERSITY OF MELBOURNE (33.3%)**

72 Inventor/es:

**CHUKKA, SRINIVAS;
SARKAR, ANINDYA;
SCHEMP, CRYSTAL;
SERTEL, OLCAY;
SINGH, SHALINI;
TUBBS, RAYMOND;
WARING, PAUL MICHAEL y
WICK, NIKOLAUS**

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 639 559 T3

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

DESCRIPCIÓN

Análisis de imágenes para el pronóstico de cáncer de mama

5 **Campo**

Esta solicitud se refiere a un procedimiento para determinar el pronóstico de cáncer de mama, tal como cáncer de mama en estadio temprano.

10 **Partes del convenio de investigación conjunta**

Ventana Medical Systems, Inc., Cleveland Clinic, y la Universidad de Melbourne son partes de convenios de investigación conjunta que regulan las invenciones divulgadas en el presente documento.

15 **Antecedentes**

Los pacientes con cáncer de mama localizado (en estadio temprano, resecable) que se someten a cirugía curativa tienen un riesgo subyacente de recidiva de cáncer local o a distancia, y las personas con recidiva muestran una tasa de mortalidad incrementada. Dependiendo de la envergadura del riesgo, existen diferentes opciones de tratamiento. Por tanto, se necesita un ensayo que pueda identificar de manera fiable a los pacientes con un riesgo bajo o alto de recidiva de cáncer. En consecuencia, también se necesitan tecnologías que puedan discriminar de manera fiable entre pacientes con alto y bajo riesgo y proporcionar a los profesionales sanitarios información adicional que se ha de considerar a la hora de determinar las opciones de tratamiento del paciente.

25 El documento JASON CHRISTIANSEN *ET AL*: "Validation of IHC4 algorithms for prediction of risk of recurrence in early breast cancer using both conventional and quantitative IHC approaches", 2012 ASCO ANNUAL MEETING PROCEEDINGS, vol. 30, 30 de mayo de 2012 (30-05-2012), XP055081113, divulga el uso de la puntuación IHC4 combinada para predecir el riesgo de recidiva de cáncer de mama.

30 **Sumario**

La presente solicitud proporciona procedimientos de pronóstico o determinación del pronóstico de cáncer de mama en un sujeto. Los procedimientos de la invención comprenden seleccionar en una muestra de cáncer de mama obtenida del sujeto al menos dos campos de visión (FOV) diferentes para cada uno de receptor de estrógenos (ER), receptor del factor de crecimiento epidérmico humano 2 (HER2), Ki-67 y receptor de progesterona (PR), en los que la muestra se marca de manera detectable con anticuerpos para cada uno de ER, HER2, Ki-67 y PR; medir la expresión de las proteínas ER, HER2, Ki-67 y PR en cada uno de los FOV seleccionados; determinar una puntuación inmunohistoquímica (IHC) combinada; medir la heterogeneidad de proteínas para ER y PR en cada uno de los FOV seleccionados; determinar una puntuación de heterogeneidad de proteínas para cada uno de ER y PR; combinar la puntuación de heterogeneidad de proteínas y la puntuación IHC combinada, generando así una puntuación de pronóstico de salida; y determinar que es probable que el cáncer de mama en el sujeto sea agresivo si la puntuación de pronóstico de salida alcanza un valor umbral o determinar que es improbable que el cáncer de mama en el sujeto sea agresivo si la puntuación de pronóstico de salida no alcanza el valor umbral.

45 Los campos de visión digitales en imágenes de una muestra de cáncer de mama de un sujeto marcada de manera detectable con anticuerpos para un biomarcador se pueden recibir y procesar para medir la heterogeneidad de proteínas para el biomarcador.

50 Las mediciones de heterogeneidad se pueden combinar con una puntuación inmunohistoquímica combinada para generar una puntuación de pronóstico de recidiva de cáncer de mama.

Dicha puntuación puede proporcionar más información que la puntuación inmunohistoquímica combinada por sí sola.

55 Se puede soportar un flujo de trabajo digital del patólogo para facilitar la selección de campos de visión en las imágenes.

Los anteriores y otros objetos y características de la divulgación serán más evidentes a partir de la siguiente descripción detallada, que procede con referencia a las figuras adjuntas.

60 **Breve descripción de los dibujos**

La FIG. 1 es un diagrama de bloques de un sistema ejemplar para llevar a cabo las tecnologías descritas en el presente documento.

65 La FIG. 2 es un diagrama de bloques que muestra cómo se usan imágenes de preparaciones teñidas para diferentes

proteínas para generar una puntuación de pronóstico de recidiva de cáncer de mama.

La FIG. 3 es un esquema que muestra una vista general de un procedimiento de ejemplo.

5 Las FIG. 4 A-B son imágenes que muestran anotaciones ejemplares en campos de visión.

La FIG. 5 es un esquema que muestra una vista general de un procedimiento de ejemplo.

La FIG. 6 es un dibujo esquemático que muestra etapas ejemplares para obtener imágenes de inmunohistoquímica (IHC) de una muestra de tejido mamario.

10 La FIG. 7 es un dibujo esquemático que muestra etapas ejemplares para determinar o medir la expresión de proteínas en un campo de visión.

15 La FIG. 8 es un dibujo esquemático que muestra etapas ejemplares para determinar o medir la heterogeneidad de proteínas.

La FIG. 9 es un diagrama de bloques de un sistema ejemplar que implementa las tecnologías para el pronóstico de cáncer de mama descritas en el presente documento.

20 La FIG. 10 es un diagrama de flujo de un procedimiento implementado por ordenador ejemplar que implementa las tecnologías para el pronóstico de cáncer de mama descritas en el presente documento.

La FIG. 11 es un diagrama de flujo de un procedimiento implementado por ordenador ejemplar para identificar núcleos en una preparación.

25 La FIG. 12 es un diagrama de bloques de un sistema ejemplar para la puntuación del campo de visión.

La FIG. 13 es un diagrama de bloques de otro sistema ejemplar para la puntuación del campo de visión.

30 La FIG. 14 es un diagrama de bloques de un clasificador ejemplar para determinar núcleos teñidos verdaderos positivos.

La FIG. 15 es un diagrama de bloques de un clasificador multietapa ejemplar para determinar núcleos teñidos verdaderos positivos.

35 La FIG. 16 es un diagrama de flujo de un procedimiento ejemplar de determinar una puntuación para HER2 en intervalos de clase.

La FIG. 17 es un diagrama de bloques de un sistema ejemplar que incluye una herramienta de heterogeneidad que implementa las tecnologías para el pronóstico de cáncer de mama descritas en el presente documento.

40 La FIG. 18 es un diagrama de flujo de un procedimiento implementado por ordenador ejemplar que implementa las tecnologías para el pronóstico de cáncer de mama descritas en el presente documento por medio de la determinación de una puntuación de heterogeneidad para biomarcadores respectivos.

45 La FIG. 19 es un diagrama de bloques de un sistema ejemplar que incluye una máquina diferencial para calcular una puntuación de heterogeneidad para un biomarcador.

La FIG. 20 es un diagrama de flujo de un procedimiento implementado por ordenador ejemplar para calcular una puntuación de heterogeneidad para su uso en las tecnologías para el pronóstico de cáncer de mama descritas en el presente documento.

50 La FIG. 21 es una captura de pantalla de una interfaz de usuario ejemplar para indicar campos de visión digitales dentro de una imagen.

55 La FIG. 22 es un diagrama de flujo de un procedimiento del flujo de trabajo digital del patólogo ejemplar.

La FIG. 23 es un diagrama de bloques de un sistema computacional ejemplar en el que se pueden implementar modos de realización descritos.

60 La FIG. 24 es un gráfico que muestra la capacidad de los procedimientos divulgados para clasificar una muestra de cáncer de mama en cuanto a mayor (evolución) o menor (sin evolución) agresividad. Los casos mostrados en el gráfico son los mismos casos usados para crear el algoritmo.

65 **Descripción detallada**

Abreviaturas y términos

Las siguientes explicaciones de términos y procedimientos se proporcionan para describir mejor la presente divulgación y para guiar a los expertos en la técnica en la práctica de la presente divulgación. Los datos de los números de acceso a GenBank a los que se hace referencia en el presente documento son las secuencias disponibles al menos tan pronto como desde el 28 de diciembre de 2012.

Anticuerpo: Moléculas de inmunoglobulina y porciones inmunológicamente activas de moléculas de inmunoglobulina, es decir, moléculas que contienen un sitio de unión a antígeno que se une específicamente (inmunorreacciona con) un antígeno (tal como ER, PR, Ki-67 o HER2). Los anticuerpos ejemplares incluyen anticuerpos monoclonales, policlonales y humanizados.

Un anticuerpo natural (tal como IgG, IgM, IgD) incluye cuatro cadenas polipeptídicas, dos cadenas pesadas (H) y dos cadenas ligeras (L), interconectadas mediante enlaces disulfuro. Tal como se usa en el presente documento, el término anticuerpo también incluye anticuerpos recombinantes producidos mediante la expresión de un ácido nucleico que codifica una o más cadenas de anticuerpo en una célula (por ejemplo, véase la patente de EE. UU. n.º 4.745.055, la patente de EE. UU. n.º 4.444.487, el documento WO 88/03565; el documento EP 256.654; el documento EP 120.694; el documento EP 125.023; Faoukner *et al.*, *Nature* 298:286, 1982; Morrison, *J. Immunol.* 123:793, 1979; Morrison *et al.*, *Ann Rev. Immunol* 2:239, 1984).

El término anticuerpo también incluye un fragmento de unión a antígeno de un anticuerpo natural o recombinante. Los ejemplos no limitantes específicos de fragmentos de unión abarcados por el término anticuerpo incluyen Fab, (Fab')₂, Fv y Fv monocatenario (scFv). Fab es el fragmento que contiene un fragmento monovalente de unión a antígeno de una molécula de anticuerpo producida mediante digestión de anticuerpo completo con la enzima papaína para producir una cadena ligera intacta y una porción de una cadena pesada o de manera equivalente mediante ingeniería genética. Fab' es el fragmento de una molécula de anticuerpo obtenida tratando el anticuerpo completo con pepsina, seguido de reducción, para producir una cadena ligera intacta y una porción de la cadena pesada; se obtienen dos fragmentos Fab' por cada molécula de anticuerpo. (Fab')₂ es el fragmento del anticuerpo obtenido al tratar el anticuerpo completo con la enzima pepsina sin reducción posterior o de manera equivalente mediante ingeniería genética. F(Ab')₂ es un dímero de dos fragmentos FAb' unidos entre sí mediante enlaces disulfuro. Fv es un fragmento genomanipulado que contiene la región variable de la cadena ligera y la región variable de la cadena pesada expresadas como dos cadenas. El anticuerpo monocatenario ("SCA") es una molécula genomanipulada que contiene la región variable de la cadena ligera y la región variable de la cadena pesada unidas mediante un enlazador polipeptídico adecuado como una molécula monocatenaria fusionada genéticamente. Los procedimientos para preparar estos fragmentos son rutinarios en la técnica.

Afinidad de unión: Afinidad de un anticuerpo por un antígeno, tal como la afinidad de un anticuerpo por un péptido ER, PR, Ki-67 o HER2. Los procedimientos de determinación de la afinidad de anticuerpos se conocen en la técnica e incluyen el cálculo mediante una modificación del procedimiento de Scatchard descrito por Frankel *et al.*, *Mol. Immunol.* 16:101-106, 1979, medición mediante una tasa de disociación antígeno/anticuerpo, o mediante un radioinmunoensayo de competición. Una afinidad de unión alta puede ser de al menos aproximadamente 1×10^{-8} M, al menos aproximadamente $1,5 \times 10^{-8}$, al menos aproximadamente $2,0 \times 10^{-8}$, al menos aproximadamente $2,5 \times 10^{-8}$, al menos aproximadamente $3,0 \times 10^{-8}$, al menos aproximadamente $3,5 \times 10^{-8}$, al menos aproximadamente $4,0 \times 10^{-8}$, al menos aproximadamente $4,5 \times 10^{-8}$ o al menos aproximadamente $5,0 \times 10^{-8}$ M.

Cáncer de mama: Incluye cualquier tumor de la mama, tal como tumores en, cerca de o incluyendo tejido mamario epitelial (carcinoma) o estromal (sarcoma). El carcinoma ductal *in situ* (DCIS) es una afección neoplásica no invasiva de los conductos mamarios. El carcinoma lobulillar no es una enfermedad invasiva, pero es un indicador de que se puede desarrollar un carcinoma. El carcinoma infiltrante (maligno) de la mama se puede dividir en estadios (I, IIA, IIB, IIIA, IIIB y IV). Véase, por ejemplo, Bonadonna *et al.*, (eds). *Textbook of Breast Cancer: A Clinical Guide to the Therapy*, 3.ª; London, Taylor & Francis, 2006. El DCIS se denomina en ocasiones cáncer de mama en estadio 0 porque no es invasivo. Los carcinomas de mama invasivos ejemplares incluyen carcinoma NOS (no especificado de otro modo), carcinoma lobulillar, carcinoma tubular/cribriforme, carcinoma mucinoso (coloide), carcinoma medular, carcinoma papilar y carcinoma metaplásico. Un sarcoma de mama ejemplar es un tumor filodes.

Un **cáncer de mama en estadio temprano** es uno que está en estadio I o II. Una muestra de tejido mamario, por ejemplo, que sea positiva para ER, con ganglios linfáticos negativos y, en algunos ejemplos, negativa para HER2 también se puede caracterizar como cáncer de mama en estadio temprano.

Los tratamientos ejemplares para el cáncer de mama incluyen cirugía (por ejemplo, extracción de parte o todo el tumor), tratamiento de bloqueo hormonal (por ejemplo, tamoxifeno), radiación, ciclofosfamida más doxorubicina (adriamicina), taxano (por ejemplo, docetaxel) y anticuerpos monoclonales tales como trastuzumab (Herceptin) o pertuzumab, o combinaciones de los mismos.

Control: Una muestra o patrón usado para su comparación con una muestra experimental o de prueba (tal como una muestra mamaria). En algunos modos de realización, el control es una muestra normal obtenida de un paciente

sano (o pluralidad de pacientes), tal como una muestra mamaria normal o pluralidad de muestras. En algunos ejemplos, el control es una muestra de tejido no tumoral obtenida de un paciente al que se ha diagnosticado cáncer de mama, tal como tejido mamario normal. En algunos modos de realización, el control es una muestra (o pluralidad de muestras) de cáncer de mama en estadio temprano conocida, tal como una muestra conocida por ser ER+, PR+, Ki-67+ y HER2-.

En algunos modos de realización, el control es un control histórico o un valor de referencia estándar o un intervalo de valores (tal como una(s) muestra(s) de control previamente sometida(s) a prueba, tal como un cáncer de mama conocido, muestra mamaria normal, muestra mamaria benigna, epitelio o estroma). En algunos modos de realización, el control es un valor estándar que representa el valor promedio (o intervalo promedio de valores) obtenido a partir de una pluralidad de muestras de pacientes, tales como muestras mamarias normales conocidas o muestras de cáncer de mama temprano conocidas.

Se pueden usar muestras de control para el control con tinción. Dicho enfoque puede ser pertinente para identificar la proporción señal/ruido de la muestra.

Poner en contacto: Poner un agente en estrecha proximidad con otro agente, permitiendo así que los agentes interactúen. Por ejemplo, se puede aplicar un anticuerpo a un portaobjetos de microscopio u otra superficie que contenga una muestra biológica, permitiendo así la detección de proteínas en la muestra que son reconocidas específicamente por el anticuerpo.

Detectar: Determinar si un agente está presente o ausente, y puede incluir determinar un patrón. En algunos ejemplos, esto puede incluir adicionalmente la cuantificación. Por ejemplo, el uso de un anticuerpo específico para una proteína particular (por ejemplo, Ki-67, ER, PR o HER2) permite la detección de la proteína en una muestra, tal como una muestra que contiene tejido de cáncer de mama. En ejemplos particulares, se detecta una señal de emisión de un marcador detectable (tal como un incremento en la señal si está presente la diana).

La detección puede ser en masa, de modo que se pueda observar simultáneamente un número macroscópico de moléculas. La detección también puede incluir la identificación de señales de moléculas individuales usando microscopía y técnicas tales como reflexión interna total para reducir el ruido de fondo.

Receptor de estrógenos (ER): Un miembro de la familia de hormonas nucleares de receptores intracelulares que se activa por 17 β -estradiol. Los receptores de estrógenos se sobreexpresan en alrededor de un 70 % de los casos de cáncer de mama, denominados "positivos para ER" (ER+).

Puntuación H: Una indicación de la expresión de proteínas que pondera las células teñidas fuertemente con más peso que las células teñidas débilmente. Por ejemplo, una puntuación H puede indicar el porcentaje de células que se tiñen débilmente (por ejemplo, 1+) más dos veces el porcentaje de células que se tiñen moderadamente (por ejemplo, 2+) más tres veces el porcentaje de células que se tiñen fuertemente (por ejemplo, 3+) (por ejemplo, véase Cuzick *et al.*, *J. Clin. Oncol.* 29:4273-8, 2011, incorporado en el presente documento mediante referencia) (véase también www.pathogenesys.com/html/semi-quantitative_ihc.html). En el presente documento se describen técnicas de cálculo de la puntuación H ejemplares.

Puntuación de heterogeneidad: Una indicación de la cantidad de heterogeneidad de la expresión de proteínas de un biomarcador en una muestra, tal como tinción en cuanto a ER, HER2, Ki-67 o PR en una muestra de cáncer de mama. La puntuación de heterogeneidad proporciona una medida de cuán diferente es un FOV de otro FOV, para el mismo marcador.

Receptor del factor de crecimiento epidérmico humano 2 (HER2): Un miembro de la familia de proteínas ErbB, que es un protooncogén localizado en el brazo largo del cromosoma 17 (17q11.2-q12) humano. Aproximadamente un 25-30 % de los cánceres de mama tienen una amplificación del gen HER2/neu o sobreexpresión de su producto de proteína, denominados "positivos para HER2" (HER2+). Los pacientes HER2+ pueden recibir el anticuerpo monoclonal trastuzumab (Herceptin) como un tratamiento para el cáncer de mama, y, en algunos ejemplos, se usa en combinación con el anticuerpo monoclonal pertuzumab. La sobreexpresión de HER2 en el cáncer de mama se ha asociado a una recidiva de la enfermedad incrementada y a un peor pronóstico.

Puntuación inmunohistoquímica (IHC) combinada: Una puntuación de pronóstico basada en un número de marcadores IHC, en la que el número de marcadores es mayor que uno. IHC4 es una de dichas puntuaciones basada en cuatro marcadores IHC medidos, en concreto, ER, HER2, Ki-67 y PR, en una muestra de cáncer de mama (por ejemplo, véase Cuzick *et al.*, *J. Clin. Oncol.* 29:4273-8, 2011 y Barton *et al.*, *Br. J. Cancer* 1-6, 24 de abril de 2012, ambas incorporadas en el presente documento mediante referencia). En un ejemplo, se calcula una puntuación IHC4 usando, por ejemplo, la siguiente fórmula:

$$IHC4 = 94,7 \times \{-0,100 ER_{10} - 0,079 PR_{10} + 0,586 HER2 + 0,240 \ln(1 + 10 \times Ki67)\}.$$

Un experto en la técnica apreciará que son posibles otras puntuaciones IHC combinadas (por ejemplo, IHC3, IHC5 o similares).

Ki-67: Una proteína nuclear asociada a la proliferación celular y la transcripción de ARN ribosómico. La inactivación del antígeno Ki-67 da lugar a la inhibición de la síntesis de ARN ribosómico. Ki-67 se usa, por ejemplo, como un marcador de proliferación.

Marcador: Un agente con capacidad de detección, por ejemplo, mediante espectrofotometría, citometría de flujo o microscopía (tal como microscopía óptica). Por ejemplo, se pueden acoplar uno o más marcadores a un anticuerpo, permitiendo así la detección de la proteína diana. Los marcadores ejemplares incluyen isótopos radioactivos, fluoróforos, ligandos, agentes quimioluminiscentes, enzimas y combinaciones de los mismos. En un ejemplo, el marcador es un punto cuántico (quantum dot).

Tejido o células normales: Tejido y células no malignos, no tumorales.

Puntuación de pronóstico de salida: El resultado de combinar y ponderar las puntuaciones IHC combinadas (por ejemplo, IHC4) y las puntuaciones de heterogeneidad para un sujeto, a partir de las cuales se puede determinar el pronóstico de cáncer de mama.

Receptor de progesterona (PR o PgR): Un receptor esteroideo intracelular que se une específicamente a progesterona. Los receptores de progesterona se sobreexpresan en algunos casos de cáncer de mama, denominados "positivos para PR" (PR+).

Pronóstico: El procedimiento para determinar el resultado probable de que un sujeto tenga una enfermedad (por ejemplo, cáncer de mama en estadio temprano) en ausencia de tratamiento adicional. En un ejemplo, los procedimientos divulgados permiten el pronóstico de una forma más agresiva de un cáncer de mama en estadio temprano si se detecta una puntuación de pronóstico de salida por encima de un umbral. En cambio, se determina el pronóstico de una forma menos agresiva de un cáncer de mama en estadio temprano si se detecta una puntuación de pronóstico de salida por debajo de un umbral. Por ejemplo, el pronóstico puede estar relacionado con la predicción de acontecimientos futuros, tales como la esperanza de vida (por ejemplo, la probabilidad de supervivencia a 1 año, 3 años o 5 años), predicción de una probable recidiva (local o metastásica) de cáncer de mama (por ejemplo, en 1, 3 o 5 años). El término similar "determinar el pronóstico" también se usa en el presente documento.

Cuantificar: Expresar como una cantidad numérica, ya sea como una cantidad real o una cantidad relativa.

Muestra: Un espécimen biológico que puede contener, por ejemplo, ADN genómico, ARN (por ejemplo, ARNm), proteína o combinaciones de los mismos, obtenida de un sujeto. Los ejemplos incluyen, pero no se limitan a, aspirado de biopsia con aguja fina, biopsia de tejido, espécimen quirúrgico y material de necropsia. En un ejemplo, una muestra incluye tejido mamario, tal como el obtenido durante una biopsia por punción, tumorectomía o mastectomía.

Estadificación del cáncer: Un cáncer, tal como cáncer de mama, se puede estadificar para describir el alcance o gravedad de un cáncer basándose en el alcance del tumor original (primario) y el grado de diseminación en el cuerpo. El cáncer de mama se puede estadificar de acuerdo con el sistema TNM (véase el Manual de estadificación del AJCC), donde T describe el tamaño del tumor y si ha invadido tejido cercano, N describe cualesquiera ganglios linfáticos implicados y M describe metástasis (diseminación del cáncer desde una parte del cuerpo a otra).

Los estadios son como sigue: estadio 0 - carcinoma *in situ*; estadio I - el tumor (T) no implica ningún ganglio linfático axilar (N); estadio IIA - T de 2-5 cm y N negativo o T < 2 cm y N positivo; estadio IIB - T > 5 cm y N negativo o T 2-5 cm y N positivo (< 4 nódulos axilares); estadio IIIA - T > 5 cm y N positivo o T 2-5 cm con 4 o más ganglios axilares; estadio IIIB - el T se ha infiltrado en la pared torácica o piel y se puede haber diseminado a < 10 N axilares; estadio IIIC - T tiene > 10 N axilares, 1 o más N supraclaviculares o infraclaviculares, o N mamario interno; y estadio IV - metástasis (M) a distancia.

Sujeto: Organismos vertebrados multicelulares vivos, una categoría que incluye mamíferos humanos y no humanos, tales como sujetos veterinarios. En un ejemplo particular, un sujeto es uno que tuvo o se sospecha que ha tenido cáncer de mama, tal como un cáncer de mama en estadio temprano.

Molécula diana: Una biomolécula cuya detección o medición se desea, tal como un marcador de cáncer de mama. Los ejemplos de moléculas diana incluyen ER, PR, Ki-67 y HER2.

En condiciones suficientes para: Una frase que se usa para describir cualquier entorno que permita la actividad deseada. Un ejemplo incluye poner en contacto un anticuerpo con una muestra de cáncer de mama suficiente para permitir la detección de una o más moléculas diana (por ejemplo, ER, PR, HER2, Ki-67) en la muestra y puede incluir la cuantificación de una o más moléculas diana en la muestra.

Sistema ejemplar para determinar el pronóstico de cáncer de mama

La FIG. 1 es un diagrama de bloques de un sistema ejemplar para llevar a cabo las tecnologías descritas en el presente documento. En el ejemplo, se hace funcionar un dispositivo de adquisición de imágenes 10, tal como un equipo de obtención de imágenes de preparaciones (por ejemplo, iScan Coreo de Ventana Medical Systems, Inc. o similar) para que acepte una preparación 5 preparada como se describe en el presente documento y generar una imagen de la preparación 5 para su análisis como se describe en el presente documento. En la práctica, se usa una pluralidad de preparaciones como se describe en el presente documento.

El sistema computacional 20 puede incluir uno o más dispositivos de entrada 23, una o más pantallas 24 y uno o más ordenadores 21, que pueden ejecutar una aplicación o plataforma de procesamiento de imágenes como se describe en el presente documento.

Los datos, tales como imágenes obtenidas, se pueden almacenar, por ejemplo, de manera remota en un servidor 12 y/o en un ordenador 21. La adquisición de imágenes se puede realizar por separado (por ejemplo, en un sistema diferente, mediante un actor diferente o similar) del procesamiento de imágenes.

En algunas implementaciones se puede implementar un escenario basado en la nube o de software como servicio, en las que la aplicación de procesamiento de imágenes 22 se halla parcial o totalmente fuera del ordenador 21 (por ejemplo, en el servidor 12 o en uno o más de otros servidores).

Procedimientos ejemplares para determinar el pronóstico de cáncer de mama

En el presente documento se muestra que se puede determinar el pronóstico de los cánceres de mama, por ejemplo, cánceres de mama en estadio temprano, basándose en la obtención de una puntuación inmunohistoquímica (IHC) combinada (tal como una puntuación IHC4) y/o una puntuación de heterogeneidad. Estas dos puntuaciones se combinan y ponderan, dando como resultado una puntuación de pronóstico de salida. La puntuación de pronóstico de salida resultante se puede usar para determinar el pronóstico del paciente con cáncer de mama. Por ejemplo, si la puntuación de pronóstico de salida está por encima de un valor umbral, esto indica que el cáncer de mama es más agresivo o que es más probable que recidive, mientras que una puntuación de pronóstico de salida que está por debajo de un valor umbral indica que el cáncer de mama es menos agresivo y que es menos probable que recidive. Esto proporciona a los médicos y pacientes la capacidad de tomar decisiones de tratamiento y seguimiento más apropiadas. Por ejemplo, el pronóstico de salida puede ser beneficioso en la determinación de un tratamiento apropiado (por ejemplo, elegir entre el seguimiento, una mastectomía, una tumorectomía o una tumorectomía combinada con quimioterapia).

En el presente documento se proporcionan procedimientos para determinar el pronóstico de cáncer de mama, tales como cáncer de mama en estadio temprano, en un sujeto. En algunos ejemplos, se sabe que el cáncer de mama es positivo para el receptor de estrógenos (ER+), positivo para el receptor de progesterona (PR+) y negativo para el receptor del factor de crecimiento epidérmico humano 2 (HER2-) o positivo para el receptor del factor de crecimiento epidérmico humano 2 (HER2+). Por ejemplo, los procedimientos se pueden usar para determinar la probable agresividad del cáncer o la probabilidad de que el cáncer recidive, por ejemplo, la probabilidad de que el cáncer recidive en un plazo de 5 años.

Las tecnologías del presente documento pueden generar una puntuación IHC combinada y una puntuación de heterogeneidad basándose en campos de visión indicados en imágenes de una o más preparaciones que tienen tejido de cáncer de mama tomado de un sujeto. A continuación, las puntuaciones se pueden combinar para generar una puntuación de pronóstico como se describe en el presente documento. Se pueden generar diversas puntuaciones y/o cálculos intermedios antes de llegar a la puntuación IHC combinada o puntuación de heterogeneidad como se describe en el presente documento.

En un procedimiento ejemplar, se toma una muestra de tejido de cáncer de mama de un sujeto. En algunos ejemplos, la muestra de cáncer de mama es una muestra de cáncer de mama en estadio temprano, tal como una que es negativa para HER2 (por ejemplo, tinción IHC de menos de 2+ como se describe en el presente documento) y/o FISH amplificado, positiva para ER y/o con ganglios linfáticos negativos. De esta manera, una muestra negativa para HER2 puede ser una que (1) muestre algo de tinción IHC, tal como tinción que sea de menos de 2+ o una puntuación H de <1 o (2) se amplifica como se indica mediante análisis de FISH, pero es una en la que un médico o patólogo llegaría a la conclusión de que, de todos modos, era negativa para HER2. De manera alternativa, la muestra puede ser positiva para HER2.

Además de FISH, se pueden usar otros tipos de hibridación *in situ* (ISH), tales como hibridación *in situ* cromogénica (CISH), hibridación *in situ* cromogénica de doble color (DISH) o similares.

La FIG. 2 es un diagrama de bloques que muestra cómo se usan las imágenes 32, 34, 36 y 38 de preparaciones teñidas para diferentes proteínas para generar una puntuación de pronóstico de recidiva de cáncer de mama 58. Aunque se muestran diferentes preparaciones, se puede emplear un enfoque de micromatrices de tejidos como se describe en el presente documento, dando como resultado un menor número de preparaciones. Por ejemplo, una

única preparación puede incluir una pluralidad de secciones de tejido diferentes, que se pueden analizar cada una para una proteína diferente. En otro ejemplo, se puede analizar una única sección de tejido para una pluralidad de proteínas. Una muestra, por ejemplo, un bloque de tejido, se divide en secciones, por ejemplo, cuatro secciones, que pueden ser o no ser secciones adyacentes o en serie de la muestra. A cada sección se aplican agentes específicos para los biomarcadores de interés, por ejemplo, sondas o tinciones. En un ejemplo, se aplica un agente específico para el receptor de estrógenos ER a una sección, se aplica un agente específico para el PR a otra sección y similares. Los agentes ejemplares incluyen anticuerpos y aptámeros. De esta manera, la muestra se puede marcar de manera detectable con anticuerpos específicos para cada uno de ER, HER2, Ki-67 y PR. A continuación, las imágenes 32, 24, 36 y 38 de las secciones de muestra se adquieren como se describe en el presente documento. A partir de las imágenes 32, 34, 36 y 38, se calculan una puntuación H para ER 41, una positividad en porcentaje para PR 42, una positividad en porcentaje para Ki-67 43 y una puntuación en intervalos de clase para HER2 44. La puntuación en intervalos de clase 44 se puede convertir en una puntuación binaria (por ejemplo, 0 o 1) 45. A continuación, las puntuaciones se combinan como se describe en el presente documento para generar una puntuación IHC4 55.

La puntuación de heterogeneidad para ER 51 se basa en una determinación de positividad en porcentaje para ER 47 y la puntuación de heterogeneidad para PR 52 se basa en una determinación de positividad en porcentaje para PR 42. Las puntuaciones de heterogeneidad 51 y 52 se pueden combinar como se describe en el presente documento para generar una puntuación de heterogeneidad combinada 54.

A continuación, la puntuación IHC4 55 y la puntuación de heterogeneidad combinada 54 se pueden combinar para generar una puntuación de pronóstico de recidiva de cáncer de mama 58.

La FIG. 3 proporciona una vista general de un procedimiento ejemplar. El procedimiento incluye determinar una puntuación IHC4 por medio de la expresión medida para cada uno de ER, Ki-67, PR y HER2 112, por ejemplo, determinando o midiendo una positividad en porcentaje para cada uno de PR y Ki-67, una puntuación en intervalos de clase para HER2 (por ejemplo, en una escala típica de 0 a 3 para representar la intensidad, en la que 0 se asigna a tinción negativa y 3 se asigna a muestras teñidas muy intensamente) y una puntuación H para ER. El procedimiento también incluye determinar una puntuación de heterogeneidad para al menos ER y PR y en algunos ejemplos también para Ki-67 y HER2 114. La puntuación IHC4 112 y la puntuación de heterogeneidad 114 resultantes se combinan 116 para producir una puntuación de pronóstico de salida 118. Basándose en la puntuación de pronóstico de salida, se determina el pronóstico 120 de cáncer de mama.

En un ejemplo, se obtienen imágenes de cada preparación y se genera una imagen digital de cada preparación usando cualquiera de los sistemas descritos en el presente documento. Se genera una puntuación IHC combinada mediante un procedimiento implementado por ordenador basado en las preparaciones, y se genera una puntuación de heterogeneidad para cada preparación (o sección de la preparación que se analiza, tal como una región que contiene una sección de tejido para el análisis de ER). En un modo de realización ejemplar, la puntuación IHC combinada se denomina "puntuación IHC4", ya que se utilizan cuatro biomarcadores para recabar información. La puntuación IHC4 se puede combinar con las puntuaciones de heterogeneidad como se describe en el presente documento para generar una puntuación de pronóstico de recidiva de cáncer de mama.

En un ejemplo, para calcular la puntuación IHC4, el procedimiento incluye seleccionar, en la imagen de cada sección de tejido en una o más preparaciones de la muestra de cáncer de mama que se ha marcado con un agente (tal como un anticuerpo), al menos dos campos de visión (FOV) diferentes para cada una de una imagen de preparación de secciones de tejido para ER, una imagen de preparación de secciones de tejido para HER2, una imagen de preparación de secciones de tejido para Ki-67 y una imagen de preparación de secciones de tejido para PR. En un modo de realización ejemplar, se seleccionan tres (3) FOV para cada preparación o sección para calcular la puntuación IHC4. Se eligen regiones tumorales como campos de visión.

La FIG. 4A muestra un conjunto ejemplar de imágenes de preparaciones 180 y anotaciones mediante las cuales se seleccionan los campos de visión (FOV) 181A, 181B y 181C (por ejemplo, por un patólogo) en las imágenes de preparaciones respectivas para calcular una puntuación IHC combinada. Como se muestra, las anotaciones pueden ser independientes, sin ninguna relación entre los campos de visión de un biomarcador. En los ejemplos, los campos de visión son rectangulares. Sin embargo, como se describe en el presente documento, se pueden usar otras técnicas (por ejemplo, los campos de visión pueden ser otras formas conocidas o formas irregulares). En algunos ejemplos, el campo de visión es un área de interés, tal como una región anatómica de interés (por ejemplo, glándula). En un ejemplo, el campo de visión es una preparación completa o sección de tejido completa.

Basándose en los campos de visión, se calcula una puntuación IHC combinada. Al calcular los componentes para la puntuación IHC combinada, los campos de visión se pueden tomar en conjunto (por ejemplo, combinados y/o considerados colectivamente) para un biomarcador particular. Por ejemplo, la puntuación H, la positividad en porcentaje y la puntuación en intervalos de clase para HER2 se pueden basar en conjuntos respectivos de campos de visión tomados en conjunto (por ejemplo, la puntuación H para ER se basa en las células observadas en los campos de visión). Son posibles otras técnicas (por ejemplo, promediar puntuaciones, votar por una puntuación, o similares).

Por ejemplo, la puntuación IHC combinada se puede calcular usando la fórmula para IHC4 de Dowsett o una variación de la misma:

5 $IHC4 = 94,7 \times \{-0,100 ER_{10} - 0,079 PR_{10} + 0,586 HER2 + 0,240 \ln(1 + 10 \times Ki67)\}$,
 en la que ER_{10} es la puntuación H/30 para ER, en la que PR_{10} se obtiene dividiendo el porcentaje de células con
 tinción positiva (por ejemplo, con un límite superior de un 10 % impuesto sobre el porcentaje de células) entre 10
 para generar una variable con un intervalo de 0 a 10, en la que HER2 es la puntuación para HER2 binaria (por
 ejemplo, 0 si es menos de 2+, 1 de otro modo) y Ki67 es la positividad en porcentaje para Ki-67. En la práctica, ER_{10}
 10 se ajusta a escala a un intervalo de 0 – 10 dividiendo la puntuación H entre 30 como se muestra en la fórmula.

Para algunos de los campos de visión, los núcleos con tinción positiva se diferencian de los núcleos con tinción
 negativa (por ejemplo, no teñidos) en cada campo de visión. A continuación, para las preparaciones para PR y Ki-67,
 se calcula la positividad en porcentaje (por ejemplo, se suman el número total de núcleos de células (por ejemplo,
 15 células malignas) con tinción positiva en cada campo de visión en la imagen digital de una preparación y se dividen
 entre el número total de núcleos con tinción positiva y negativa de cada uno de los campos de visión de una imagen
 digital) en una única preparación como sigue:

Positividad en porcentaje = número de células con tinción positiva / (número de células con tinción positiva + número
 de células con tinción negativa)

20 En un ejemplo, la positividad en porcentaje se determina manualmente (por ejemplo, sin el uso de una imagen
 digital, con el uso de análisis de imágenes, o diversos otros procedimientos de análisis de imágenes).

25 En un ejemplo, los núcleos se detectan y clasifican como núcleos con tinción positiva o núcleos con tinción negativa
 por medio de las técnicas de identificación de núcleos descritas en el presente documento. Por ejemplo, las células
 con tinción negativa se pueden diferenciar, por ejemplo, de linfocitos y tejido estromal. Dicha técnica puede ser
 superior a usar un umbral de color universal.

30 Adicionalmente, por ejemplo, se calcula la puntuación H (que refleja la intensidad de las células teñidas y el número
 de células teñidas para una preparación individual) para la preparación (o sección de tejido) para ER. Dicha
 puntuación se puede basar en los FOV tomados en conjunto colectivamente. En modos de realización ejemplares,
 se usa una escala numérica (por ejemplo, 0-300) para la puntuación H. Para las células (por ejemplo, en los campos
 de visión) con tinción positiva, se determina la intensidad. Dicha intensidad se puede determinar por medio de un
 algoritmo, en comparación con un umbral, y asignar un número de intervalos de clase. Dicho número de intervalos
 35 de clase puede ser, por ejemplo, 0, 1, 2 o 3. Como resultado, habrá un recuento de núcleos (es decir, células) en
 cada intervalo de clase. Los recuentos de los intervalos de clase se pueden usar para calcular la puntuación H.

40 La etapa intermedia de agrupar los núcleos en intervalos de clase se puede reutilizar más tarde al calcular una
 puntuación de heterogeneidad. Las intensidades positivas (por ejemplo, marrón) (por ejemplo, 1, 2 y 3) se pueden
 agregar para determinar un recuento de núcleos con tinción positiva global según sea necesario (por ejemplo, para
 la puntuación de heterogeneidad para ER).

45 El cálculo de la puntuación binaria para HER2 puede incluir determinar la puntuación en intervalos de clase (que se
 relaciona con la completitud de la tinción de la membrana celular y la intensidad de la membrana celular teñida) para
 la preparación para HER2 (por ejemplo, basándose en los campos de visión). La puntuación en intervalos de clase
 para los FOV combinados es 0, 1, 2 o 3 como se describe en el presente documento. 0 o 1 se consideran negativos
 (0 para la puntuación binaria) y 2 o 3 se consideran positivos (1 para la puntuación binaria). De esta manera, se
 calcula una puntuación binaria para HER2. De manera alternativa, las puntuaciones para los FOV se pueden
 50 combinar (por ejemplo, promediar, promedio y redondeo o similares) para determinar una puntuación global para
 HER2.

A continuación, se puede determinar la puntuación IHC4 utilizando la positividad en porcentaje, puntuación H y
 puntuación binaria descritas anteriormente. La expresión de las proteínas ER, HER2, Ki-67 y PR se puede detectar o
 medir en cada uno de los FOV seleccionados para ayudar a realizar estas determinaciones. Por ejemplo, se puede
 55 examinar o medir la expresión de proteínas para determinar o medir una positividad en porcentaje para cada uno de
 Ki-67 y PR, una puntuación de intensidad de tinción en intervalos de clase para HER2 y una puntuación H para ER.
 Basándose en esta información, se puede determinar o calcular una puntuación inmunohistoquímica (IHC)
 combinada de acuerdo con la fórmula anterior.

60 Aparte de la puntuación IHC combinada colectiva (por ejemplo, IHC4) para la muestra, también se puede calcular la
 heterogeneidad, por ejemplo, la heterogeneidad de la expresión de proteínas para biomarcadores respectivos (por
 ejemplo, ER, PER, HER2 y Ki67). Las determinaciones de la heterogeneidad añaden valor pronóstico al valor ya
 existente de la puntuación IHC combinada (por ejemplo, IHC4). En un modo de realización ejemplar, se determina la
 heterogeneidad, por ejemplo, heterogeneidad regional, para cada biomarcador (por ejemplo, ER, PER, HER2 y
 65 Ki67). La FIG. 4B es un conjunto ejemplar de imágenes de preparaciones 185 que muestran anotaciones mediante
 las cuales se seleccionan los campos de visión 186A, 186B y 186C (por ejemplo, por un patólogo) en las imágenes

de preparaciones respectivas para calcular una puntuación de heterogeneidad. Como se muestra, las anotaciones pueden ser independientes, sin ninguna relación entre los campos de visión de un biomarcador. Se puede usar un conjunto de anotaciones diferente para la heterogeneidad que las usadas para la puntuación IHC combinada. En algunos casos, puede haber un solapamiento parcial entre el FOV usado para la puntuación IHC combinada y el FOV usado para las puntuaciones de heterogeneidad (por ejemplo, se puede compartir un campo de visión entre los mismos para un biomarcador).

En un modo de realización ejemplar, se determina la heterogeneidad para una preparación (o sección de tejido) para medir cuán diferentes son las diversas regiones tumorales de una referencia, por ejemplo, entre sí. En un modo de realización ejemplar, un patólogo selecciona un número de campos de visión para una preparación (o sección de tejido). En un modo de realización, el patólogo selecciona tres campos de visión, incluyendo cada uno regiones tumorales, por ejemplo, con diferentes positivities en porcentaje.

La heterogeneidad de proteínas para ER y PR se mide o detecta usando los FOV indicados. Por ejemplo, medir la heterogeneidad de proteínas puede incluir determinar una métrica de variabilidad (VM) para cada una de ER y PR, en la que $VM = \text{STD}(PP(FS_1), PP(FS_2), \dots, PP(FS_N))$. PP(FS) es la positividad en porcentaje para cada FOV, FS (por ejemplo, los campos de visión de una imagen digital de una muestra de tejido que se ha puesto en contacto con un agente específico para la sonda de ER, un agente específico para PR o similares). Basándose en la métrica de variabilidad, se determina o calcula una puntuación de heterogeneidad para cada uno de ER y PR. Por ejemplo, la puntuación de heterogeneidad para cada uno de ER y PR se puede calcular usando la siguiente fórmula, donde $\alpha = [0,1]$ (por ejemplo, un número que varía desde 0 a 1) es un factor de normalización, y S es una puntuación de preparaciones con positividad en porcentaje promedio como se describe en el presente documento:

$$H = \begin{cases} \alpha * \frac{VM}{0,05}, & S < 10 \% \\ \alpha * \frac{VM}{S}, & \text{de otro modo} \end{cases}$$

Los modos de realización que calculan la heterogeneidad para Ki-67 pueden usar la misma fórmula o similar que la usada para ER y PR. La puntuación de heterogeneidad de proteínas para HER2, que puede ser útil en algunos modos de realización, se puede calcular usando la siguiente fórmula:

$$H = \sum_{\substack{v_{i,j} \\ i \neq j}} \|P(FS_i) - P(FS_j)\|$$

en la que P(FS) es la puntuación para HER2 en intervalos de clase (por ejemplo, determinada como se describe en el presente documento) para cada campo de visión, FS. Si las puntuaciones en intervalos de clase de cada campo de visión son iguales (por ejemplo, $P(FS_1) = P(FS_2) = P(FS_N)$), entonces $H = 0$. De esta manera, H indica cuán diferente es la expresión de proteínas en los campos de visión diferentes para un biomarcador.

En un modo de realización ejemplar se utiliza al menos una de las puntuaciones de heterogeneidad para determinar la puntuación de pronóstico de salida. En un modo de realización, las puntuaciones de heterogeneidad resultantes para las preparaciones respectivas y la puntuación IHC combinada se combinan y ponderan, generando así una puntuación de pronóstico de salida. En un modo de realización ejemplar, la puntuación de pronóstico de salida es un factor o porcentaje (P1) de la puntuación IHC combinada más un factor o porcentaje (P2) de una puntuación de heterogeneidad combinada, donde P1 y P2 son mayores que 0. Por ejemplo, la puntuación de pronóstico de salida (PS) se puede calcular usando la fórmula:

$$PS = 0,03114 * IHC4 + 1,95119 * \text{puntuación de heterogeneidad combinada}$$

donde la puntuación de heterogeneidad combinada es igual a la raíz cuadrada de (puntuación de heterogeneidad para ER + puntuación de heterogeneidad para PR). La puntuación de pronóstico de salida resultante se usa para determinar el pronóstico del paciente que tiene el cáncer de mama. Por ejemplo, el procedimiento puede determinar el pronóstico de que sea probable que el cáncer de mama en el sujeto sea agresivo o recidive (por ejemplo, en el plazo de 5 años) si la puntuación de pronóstico de salida está por encima de un determinado umbral o de que sea improbable que el cáncer de mama en el sujeto sea agresivo o recidive si la puntuación de pronóstico de salida está por debajo del umbral.

En algunos ejemplos, el procedimiento puede incluir obtener una imagen digitalizada de la muestra de cáncer de mama (por ejemplo, una que esté en uno o más portaobjetos de microscopio) que se marca de manera detectable para cada uno de ER, HER2, Ki-67 y PR. Por ejemplo, se pueden obtener una o más imágenes digitalizadas para cada uno de ER, HER2, Ki-67 y PR. En algunos ejemplos, el procedimiento incluye adicionalmente seleccionar un sujeto para el que se ha determinado el pronóstico de que tiene una probabilidad más alta de recidiva, por ejemplo,

seleccionar al paciente para un tratamiento más agresivo.

La FIG. 5 proporciona detalles adicionales sobre un modo de realización particular de un procedimiento de acuerdo con la tecnología. El procedimiento también puede incluir adquirir u obtener imágenes (tales como una imagen digital) de la muestra de cáncer de mama teñida para detectar ER, HER2, Ki-67 y PR 108. Se seleccionan al menos dos FOV para cada una de ER, HER2, Ki-67 y PR 110, por ejemplo, seleccionados y marcados en una imagen digital. Sin embargo, el número de FOV seleccionados puede variar (por ejemplo, puede haber 2, 3, 4, 5, 6, 7, 8, 9 o 10 FOV). Después de seleccionar los FOV 110, el procedimiento incluye detectar o medir la expresión de ER, HER2, Ki-67 y PR en cada FOV 111. Basándose en esta información, se determina una puntuación IHC (por ejemplo, IHC4) por medio de la expresión para cada uno de ER, Ki-67, PR y Her2 112, por ejemplo, determinando o midiendo una positividad en porcentaje para cada uno de PR y Ki-67, una puntuación binaria basada en una puntuación en intervalos de clase para HER2 (por ejemplo, en una escala típica de 0 a 3 para representar la intensidad, en la que 0 se asigna a tinción negativa y 3 se asigna a muestras que se tiñen muy intensamente) y una puntuación H para ER. El procedimiento también incluye detectar o medir la heterogeneidad de proteínas para cada uno de ER y PR y en algunos ejemplos también Ki-67 y HER2 113. Basándose en esta información, se determina 114 una puntuación de heterogeneidad para al menos ER y PR y en algunos ejemplos también Ki-67 y HER2. La puntuación IHC4 112 y puntuación de heterogeneidad 114 resultantes se combinan para producir una puntuación de pronóstico de salida 118. Basándose en la puntuación de pronóstico de salida, se determina el pronóstico 120 de cáncer de mama.

La FIG. 6 proporciona detalles sobre un procedimiento ejemplar para obtener o adquirir las imágenes IHC de la muestra de tejido mamario que se analiza usando los procedimientos proporcionados en el presente documento. Los procedimientos de obtención y tinción de muestras son rutinarios en la técnica. Por ejemplo, el procedimiento puede incluir obtener el tejido mamario de un paciente 210, por ejemplo, de una biopsia de mama. A continuación, el tejido resultante se fija y se somete a inclusión 212, por ejemplo, usando formalina y parafina. A continuación, el tejido fijado y sometido a inclusión se puede cortar o seccionar y montar sobre un sustrato 214, por ejemplo, montar sobre uno o más portaobjetos de microscopio de vidrio. En algunos ejemplos, una preparación incluye una pluralidad de secciones de tejido, tales como al menos 2 o al menos 4 secciones. A continuación, las secciones de tejido se pueden incubar con anticuerpos apropiados (u otros agentes de unión específicos) para marcar las proteínas ER, HER2, Ki-67 y PR. Por ejemplo, al menos una sección de tejido mamario se puede marcar con un anticuerpo específico para ER, otra sección de tejido mamario se puede marcar con un anticuerpo específico para HER2, otra sección de tejido mamario se puede marcar con un anticuerpo específico para PR y otra sección de tejido mamario se puede marcar con un anticuerpo específico para Ki-67 216. Sin embargo, un experto en la técnica apreciará que una única sección de tejido se puede marcar con más de un anticuerpo (u otro agente de unión específico), siempre y cuando dichas proteínas marcadas sean distinguibles, por ejemplo, usando anticuerpos secundarios marcados de manera diferente. Después de marcar las secciones de tejido, se pueden obtener las imágenes. Por ejemplo, se pueden obtener una o más imágenes de la sección de tejido mamario marcada con el anticuerpo específico para ER, se pueden obtener una o más imágenes de la sección de tejido mamario marcada con el anticuerpo específico para HER2, se pueden obtener una o más imágenes de la sección de tejido mamario marcada con el anticuerpo específico para PR y se pueden obtener una o más imágenes de la sección de tejido mamario marcada con el anticuerpo específico para Ki-67 218.

La FIG. 7 proporciona detalles sobre los parámetros que se pueden medir o determinar al detectar la expresión de proteínas en cada FOV. Dichos valores se pueden usar para calcular la puntuación IHC (por ejemplo, IHC4). Como se advierte en la FIG. 5, después de seleccionar los FOV 110, el procedimiento incluye detectar o medir la expresión de ER, HER2, Ki-67 y PR en cada FOV 111. Como se muestra en la FIG. 7, detectar la expresión de proteínas en cada FOV puede incluir determinar o medir una positividad en porcentaje para cada uno de PR y Ki-67 410, determinar o medir una puntuación en intervalos de clase para HER2 412 como se describe en el presente documento, y determinar o medir una puntuación H para ER 414. Las puntuaciones se pueden combinar en una puntuación inmunohistoquímica como se describe en el presente documento. La combinación puede comprender convertir la puntuación en intervalos de clase para HER2 en una puntuación binaria para HER2.

La FIG. 8 proporciona detalles sobre los parámetros que se pueden medir o determinar al medir la heterogeneidad de proteínas en cada FOV. Dichos valores se pueden usar para calcular la puntuación de heterogeneidad. Como se observa en la FIG. 5, el procedimiento puede incluir detectar o medir la heterogeneidad de proteínas para cada uno de ER, HER2, Ki-67 y PR 113. Como se muestra en la FIG. 8, detectar la heterogeneidad de proteínas puede incluir determinar o medir la variabilidad de la positividad en porcentaje en cada FOV y normalizar los valores, para al menos ER y PR (y en algunos ejemplos también para Ki-67) 510. Por ejemplo, se calcula una métrica de variabilidad (VM) para cada uno de ER y PR (y en algunos ejemplos para Ki-67), en la que $VM = \frac{STD(PP(FS_1), PP(FS_2), \dots, PP(FS_N))}{PP(FS)}$, y PP(FS) es la positividad en porcentaje para cada campo de visión, FS. Como se muestra en la FIG. 8, medir la heterogeneidad de proteínas para HER2 puede incluir determinar o medir la puntuación en intervalos de clase para cada FOV para HER2 y la variación entre puntuaciones en intervalos de clase 512 (por ejemplo, usando la puntuación en intervalos de clase determinada anteriormente en 412, pero incorporando al menos otro FOV).

Sistema ejemplar que implementa las tecnologías

La FIG. 9 es un diagrama de bloques de un sistema ejemplar 900 que implementa las tecnologías para el pronóstico

de cáncer de mama descritas en el presente documento. En el ejemplo, una herramienta de procesamiento de imágenes de preparaciones 920 puede aceptar una pluralidad de imágenes de preparaciones 912 y una pluralidad de campos de visión 914 como entrada. La herramienta 920 arroja una indicación 990 de si es probable que el cáncer recidive en el sujeto. Como se describe en el presente documento, dicha indicación 990 se puede basar en una puntuación de pronóstico de salida.

Las preparaciones pueden representar la expresión de proteínas para biomarcadores respectivos en una muestra de cáncer de mama de un sujeto (por ejemplo, marcada de manera detectable con anticuerpos para los biomarcadores como se describe en el presente documento).

La herramienta 920 puede proporcionar interfaces de usuario para recibir selecciones de los campos de visión (FOV) 914 o dicha funcionalidad se puede proporcionar mediante otra herramienta o componente.

La puntuación IHC combinada se puede calcular (por ejemplo, mediante la herramienta 920 u otra herramienta o componente) basándose en las imágenes de preparaciones y campos de visión dentro de las imágenes de preparaciones.

La herramienta 920 se hace funcionar para calcular una o más puntuaciones de heterogeneidad 927 basándose en las imágenes de preparaciones y campos de visión dentro de las imágenes de preparaciones como se describe en el presente documento.

La herramienta de procesamiento de imágenes de preparaciones 920 puede incluir una herramienta de pronóstico 930 que acepte una puntuación IHC combinada calculada y las una o más puntuaciones de heterogeneidad 927 como entrada y arroje una indicación 990 de si el cáncer es más agresivo y, de esta manera, si es más probable que recidive en el sujeto.

En la práctica, los sistemas mostrados en el presente documento, tales como el sistema 900, pueden ser más complicados, con funcionalidad adicional, entradas más complejas y similares. Por ejemplo, una funcionalidad adicional puede calcular la puntuación IHC combinada, la o las puntuaciones de heterogeneidad o ambas, o dichas puntuaciones pueden ser proporcionadas por otro software.

El sistema 900 y cualquiera de los demás sistemas descritos en el presente documento se pueden implementar conjuntamente con cualquiera de los componentes de hardware descritos en el presente documento, tales como los sistemas computacionales descritos a continuación (por ejemplo, unidades de procesamiento, memoria y similares). En cualquiera de los ejemplos del presente documento, las entradas, salidas y herramientas se pueden almacenar en uno o más medios de almacenamiento legibles por ordenador o dispositivos de almacenamiento legibles por ordenador. Las tecnologías descritas en el presente documento pueden ser genéricas para las especificaciones de los sistemas operativos o hardware y se pueden aplicar en cualquier variedad de entornos para aprovechar las características descritas.

Procedimiento implementado por ordenador ejemplar que implementa las tecnologías

La FIG. 10 es un diagrama de flujo de un procedimiento implementado por ordenador ejemplar 1000 que implementa las tecnologías para el pronóstico de cáncer de mama descritas en el presente documento y se puede implementar, por ejemplo, en el sistema mostrado en la FIG. 9.

El procedimiento se puede realizar en una o más imágenes de entrada de preparaciones que se reciben para su procesamiento. Por ejemplo, se puede usar una imagen de preparación que represente una muestra de cáncer de mama de un sujeto marcada de manera detectable con anticuerpos para un biomarcador como se describe en el presente documento.

En 1010 se recibe una pluralidad de campos de visión digitales dentro de una imagen. Como se describe en el presente documento, se puede recibir una indicación de un campo de visión por medio de selecciones de imágenes visualizadas (por ejemplo, trazando un contorno sobre la imagen para un campo de visión) o una indicación de un límite dentro de una imagen.

En 1012, la heterogeneidad de proteínas se mide entre los campos de visión digitales como se describe en el presente documento (por ejemplo, se determina la heterogeneidad interregional para los campos de visión para una imagen de preparación dada). Por ejemplo, se puede medir la expresión de proteínas para un biomarcador para los campos de visión respectivos y comparar las mediciones frente a la variabilidad como se describe en el presente documento. En la práctica, la heterogeneidad se puede medir para una pluralidad de biomarcadores (por ejemplo, como una puntuación de heterogeneidad). Se pueden usar uno o más de ER, PR y similares (por ejemplo, Ki-67, HER2 o similar).

Al medir la heterogeneidad de proteínas para un segundo biomarcador, se puede recibir una segunda pluralidad de campos de visión digitales dentro de una imagen visualizada que representa una muestra de cáncer de mama del

sujeto marcada de manera detectable con anticuerpos para un segundo biomarcador. La heterogeneidad de proteínas para el segundo biomarcador se puede medir (por ejemplo, como una puntuación de heterogeneidad separada de la primera).

5 En 1016 se recibe una puntuación IHC combinada para el sujeto. Se puede soportar cualquiera de las puntuaciones IHC combinadas descritas en el presente documento. La puntuación IHC combinada puede ser para la misma muestra usada para el análisis de heterogeneidad, o se puede usar una muestra diferente del mismo sujeto.

10 En 1018 se genera una puntuación de pronóstico de salida basada en al menos la heterogeneidad de proteínas medida para el biomarcador entre la pluralidad de campos de visión digitales y la puntuación inmunohistoquímica combinada para el sujeto. Como se describe en el presente documento, una puntuación IHC combinada y una o más puntuaciones de heterogeneidad para biomarcadores respectivos se pueden generar y usar para calcular una puntuación de pronóstico de salida.

15 En 1020, basándose en la puntuación de pronóstico de salida, se arroja una indicación de pronóstico (por ejemplo, indicación 990 de si el cáncer es más agresivo y, de esta manera, es probable que recidive en el sujeto). Por ejemplo, los umbrales descritos en el presente documento se pueden usar para seleccionar entre categorías de indicaciones de pronóstico; indicar un resultado positivo o negativo; o similar.

20 El procedimiento 1000 y cualquiera de los demás procedimientos implementados por ordenador descritos en el presente documento se pueden realizar mediante instrucciones ejecutables por ordenador (por ejemplo, provocar que un sistema computacional realice el procedimiento) almacenadas en uno o más medios legibles por ordenador (por ejemplo, medios de almacenamiento u otros medios tangibles) o almacenadas en uno o más dispositivos de almacenamiento legibles por ordenador.

25

Recepción ejemplar del campo de visión

30 En cualquiera de los ejemplos del presente documento, se puede recibir un campo de visión como una indicación almacenada de un área (por ejemplo, un conjunto de píxeles, un límite o similar) dentro de una imagen o como una anotación (por ejemplo, dibujando, trazando o similar) mediante un sistema de ordenador que hace funcionar un patólogo con referencia a (por ejemplo, en) una imagen (por ejemplo, que se almacena entonces para su uso posterior).

35 Por ejemplo, la anotación se puede realizar mediante una herramienta o software separado o incorporar a la herramienta o software que realiza el análisis de heterogeneidad.

Puntuación inmunohistoquímica combinada ejemplar

40 En cualquiera de los ejemplos del presente documento se puede combinar una puntuación inmunohistoquímica combinada con una puntuación de heterogeneidad para dar una puntuación de pronóstico de recidiva de cáncer de mama. Por ejemplo, se puede usar la puntuación IHC (por ejemplo, IHC4) descrita en el presente documento u otra puntuación inmunohistoquímica combinada que combine dos o más biomarcadores, por ejemplo, HER2, ER, PR, Ki-67 o similar. Dicha puntuación se puede determinar por medio de los campos de visión indicados por un patólogo como se describe en el presente documento.

45

La heterogeneidad se puede analizar para los biomarcadores usados para generar la puntuación IHC combinada como se describe en el presente documento y combinar con la puntuación IHC combinada para dar como resultado una puntuación de pronóstico de salida.

50 ***Imagen ejemplar***

55 En cualquiera de los ejemplos del presente documento, una imagen puede ser una imagen digital que represente una muestra de cáncer de mama de un sujeto marcada de manera detectable con anticuerpos para un biomarcador particular. En la práctica, dicha imagen representa una sección de dicha muestra. Se pueden usar diferentes secciones para diferentes biomarcadores.

Las tecnologías pueden soportar una pluralidad de imágenes para los biomarcadores respectivos. Las implementaciones pueden soportar múltiples biomarcadores en una única imagen (por ejemplo, en un escenario multiplex).

60

Al almacenarse, una imagen se puede representar como datos de imagen, píxeles o vóxeles que tienen, por ejemplo, valores de color, valores de intensidad o ambos. Los vóxeles o píxeles de imagen se pueden procesar como se describe en el presente documento. Por ejemplo, se pueden analizar los campos de visión dentro de las imágenes para cuantificar la expresión de proteínas dentro del campo de visión para un biomarcador dado.

65

Campos de visión ejemplares

En cualquiera de los ejemplos del presente documento, un campo de visión digital puede ser un área dentro de una imagen o una indicación de dicha área y se denomina en ocasiones simplemente un “campo de visión” en el presente documento. Dichos campos de visión pueden comprender regiones de interés según indica un patólogo. En algunas implementaciones, las preparaciones se pueden comparar con otras preparaciones, de modo que el campo de visión pueda ser toda la preparación. Sin embargo, puesto que el campo de visión es típicamente más pequeño que la imagen digital completa, típicamente corresponde a un área que es más pequeña que la sección completa que se está formando como imagen. En la práctica, dicho campo de visión digital se puede almacenar como una imagen separada o indicar mediante un límite (por ejemplo, dibujado originalmente por un patólogo y almacenado electrónicamente) con referencia a una imagen de referencia (por ejemplo, la imagen a partir de la que derivó el campo de visión). Las porciones de la imagen en el interior del límite se consideran dentro del campo de visión, y las porciones de la imagen que quedan fuera del límite se consideran fuera del campo de visión. Las porciones en el límite se pueden considerar dentro o fuera del campo de visión según se desee.

Típicamente, los campos de visión digitales son regiones contiguas de píxeles o vóxeles dentro de la imagen seleccionada por un patólogo de acuerdo con un protocolo apropiado. Se puede soportar cualquier forma arbitraria (por ejemplo, forma rectangular, no rectangular, cuadrada, elíptica, circular, trazada o similar) o área de interés, y se pueden proporcionar herramientas para seleccionar formas particulares (por ejemplo, una herramienta para trazado, una herramienta para elipse, una herramienta para círculo, una herramienta para cuadrado, una herramienta para rectángulo o similar). Un campo de visión puede comprender una estructura anatómica de interés (tal como una glándula).

Los campos de visión diferentes se seleccionan típicamente para puntuaciones inmunohistoquímicas combinadas y puntuaciones de heterogeneidad; sin embargo, es posible que haya solapamiento entre las dos. En otras palabras, los campos de visión para una puntuación inmunohistoquímica combinada se pueden reutilizar para una puntuación de heterogeneidad y viceversa.

Al almacenar, recibir o arrojar campos de visión, se puede usar una referencia al campo de visión digital en lugar de los datos del campo de visión real.

Selección de campos de visión ejemplares

En cualquiera de los ejemplos del presente documento, los campos de visión se pueden seleccionar de acuerdo con un protocolo apropiado para el propósito. Por ejemplo, para los campos de visión usados en una puntuación IHC (por ejemplo, IHC4), el protocolo especifica que los campos seleccionados (por ejemplo, 2, 3, 4, 5, 6, 7 o similar) sean regiones tumorales y sean representativos (por ejemplo, con un aspecto similar) de la expresión de proteínas del biomarcador dentro del tejido tumoral. Para los campos de visión usados para una puntuación de heterogeneidad, el protocolo especifica que se seleccionen campos (por ejemplo, 2, 3, 4, 5, 6, 7 o similar) que representan diferentes (por ejemplo, heterogéneos) niveles de expresión de proteínas del biomarcador dentro del tejido tumoral. En algunos casos, se puede usar un único campo para ambos propósitos (por ejemplo, uno de los campos heterogéneos también se usa en una puntuación IHC combinada).

Una interfaz de usuario presentada al usuario que realiza la selección (por ejemplo, un patólogo) puede indicar para qué propósito se seleccionan los campos de visión y también proporcionar orientación en relación con el protocolo según sea apropiado.

Heterogeneidad de proteínas ejemplar

En cualquiera de los ejemplos del presente documento, la heterogeneidad de proteínas se refiere a la variación espacial de los patrones de tinción histoquímica y molecular, tales como los patrones de tinción para los biomarcadores de cáncer de mama ER, HER2, Ki-67 y PR en una muestra de cáncer de mama y en ocasiones se denomina simplemente “heterogeneidad” en el presente documento. Por ejemplo, la heterogeneidad se incrementa con la variabilidad del nivel de biomarcadores en diferentes localizaciones dentro de una única muestra. La heterogeneidad puede ser un indicador de la variación espacial de la agresividad del tumor y/o de los patrones de crecimiento que se pueden correlacionar con un fenotipo clínico agregado (por ejemplo, un tumor que es probable que recidive).

Los procedimientos de la invención implican medir la heterogeneidad de proteínas para ER y PR como se expone en las reivindicaciones adjuntas. Los ejemplos o modos de realización que exceden el alcance reivindicado constituyen una divulgación útil para comprender la invención.

En el presente documento se muestra que la heterogeneidad biológica de la expresión de proteínas para ER y PR se correlaciona con la recidiva impredecible de una fracción de pacientes con cáncer de mama en estadio temprano.

La heterogeneidad se puede medir mediante una métrica de variabilidad que mide cuán diferentes son los niveles de expresión de proteínas entre los campos de visión para el mismo biomarcador (por ejemplo, la variabilidad de las

mediciones de expresión de proteínas dentro de campos de visión para un único biomarcador). De esta manera, se puede usar heterogeneidad interregional (por ejemplo, inter-FOV).

5 Se puede calcular una medida cuantitativa de heterogeneidad interregional (por ejemplo, heterogeneidad entre los campos de visión) basándose en una desviación (por ejemplo, desviación estándar u otro momento de la distribución) de la medición de expresión de proteínas, PE (por ejemplo, positividad en porcentaje), en los campos de visión diferentes para un biomarcador dado (por ejemplo, ER, PR, Ki-67 o similar). Dicha medición puede cuantificar cuán distantes están (por ejemplo, la diseminación de la distribución) las mediciones de expresión de proteínas. Por ejemplo, se puede realizar un cálculo de heterogeneidad ejemplar para un conjunto de campos de visión que tienen mediciones de expresión de proteínas respectivas PE(FS₁), PE(FS₂), ...PE(FS_N) para un biomarcador dado en los campos de visión FS como sigue:

$$\text{Heterogeneidad ejemplar (H)} = \sigma (\text{PE}(\text{FS}_1), \text{PE}(\text{FS}_2), \dots \text{PE}(\text{FS}_N))$$

15 El valor se puede normalizar de acuerdo con una puntuación de preparaciones promedio (S), que puede ser el promedio de las mediciones de expresión de proteínas para una preparación particular (por ejemplo, el promedio de PE(FS₁), PE(FS₂), ...PE(FS_N)). Si la puntuación de preparaciones promedio (S) está por debajo de un umbral (por ejemplo, un 10 % o similar), se puede usar un valor de sustitución (por ejemplo, un 5 %) como puntuación de preparaciones promedio. La normalización se puede lograr dividiendo la desviación observada entre la puntuación de preparaciones promedio.

20 Para los biomarcadores que implican cálculos en intervalos de clase (por ejemplo, HER2), se puede usar una suma de diferencias de intervalos de clase entre permutaciones de los campos de visión como una medida de heterogeneidad. De esta manera, la heterogeneidad se puede calcular agregando diferencias de puntuaciones en intervalos de clase para campos de visión diferentes para el marcador. Por ejemplo, la heterogeneidad para un conjunto de campos de visión (*i* y *j*) que tiene puntuaciones en intervalos de clase de expresión de proteínas respectivas (por ejemplo, 0, 1, 2, 3) P(FS₁), P(FS₂), ...P(FS_N) para un biomarcador en intervalos de clase dado se puede calcular como sigue:

$$H = \sum_{\substack{\forall i,j \\ i \neq j}} \|P(\text{FS}_i) - P(\text{FS}_j)\|$$

30 En el ejemplo, un único campo de visión tiene una única puntuación en intervalos de clase. Si todas las puntuaciones en intervalos de clase son iguales, la puntuación es 0.

35 Como se describe en el presente documento, se puede hacer la normalización de acuerdo con el biomarcador implicado.

Heterogeneidad espacial de proteínas ejemplar

40 En cualquiera de los ejemplos del presente documento se puede medir la heterogeneidad espacial de proteínas. Dicha heterogeneidad se puede clasificar como geográfica, regional, interglandular, intraglandular o similar.

45 Por ejemplo, la heterogeneidad geográfica se puede medir midiendo la variación en la expresión de proteínas a nivel geográfico: dos bloques de tejido separados entre sí más de una distancia umbral (por ejemplo, 2 pulgadas).

La heterogeneidad regional se puede medir midiendo la variación en la expresión de proteínas a nivel regional: en la misma sección de tejido, por ejemplo, separación entre 0,25 y 2 pulgadas, y campos de visión diferentes, por ejemplo, al menos a una distancia de un objetivo 4x entre sí. Los campos de visión como se describen en el presente documento se pueden usar para medir dicha heterogeneidad regional.

50 La heterogeneidad interglandular se puede medir midiendo la variación en la expresión de proteínas a nivel interglandular: por ejemplo, a una separación de menos de 0,25 pulgadas y dentro de un objetivo 4x.

55 La heterogeneidad intraglandular se puede medir midiendo la variación en la expresión de proteínas a nivel intraglandular: por ejemplo, las formaciones en un objetivo 20x para las categorías regional e interglandular.

Puntuaciones alternativas ejemplares de heterogeneidad

60 En cualquiera de los ejemplos del presente documento, se pueden usar métricas de variabilidad (VM) distintas de la desviación estándar, σ , para medir la heterogeneidad de proteínas interregional. Por ejemplo, se pueden usar diferencias interregionales en las mediciones de expresión de proteínas o un máximo de las mismas (por ejemplo, con más de dos campos de visión). Por ejemplo, se puede realizar un cálculo de heterogeneidad ejemplar para un conjunto de campos de visión que tienen mediciones de expresión de proteínas (PE) respectivas para los campos de

visión FS₁, FS₂, FS₃ para un biomarcador dado como sigue usando una función de valor absoluto (ABS):

VM = MÁX.

5 (ABS(PE(FS₁) - PE(FS₂)), ABS(PE(FS₁) - PE(FS₃)), ABS(PE(FS₂) - PE(FS₃)))

Dicho cálculo también puede explicar el intervalo de variabilidad. Por ejemplo, también se puede tener en cuenta la diferencia interregional mínima. Un cálculo ejemplar puede ser de esta manera:

10 VM_{adj} = VM – MÍN.

(ABS(PE(FS₁) - PE(FS₂)), ABS(PE(FS₁) - PE(FS₃)), ABS(PE(FS₂) - PE(FS₃)))

Son posibles variaciones adicionales.

15 **Normalización ejemplar de puntuaciones de heterogeneidad**

En cualquiera de los ejemplos del presente documento, la normalización se puede aplicar a puntuaciones de heterogeneidad (por ejemplo, una puntuación de heterogeneidad para un biomarcador). Por ejemplo, la normalización se puede lograr dividiendo la desviación observada entre la puntuación de preparaciones promedio (por ejemplo, para el biomarcador medido) como se describe anteriormente.

Se puede lograr una normalización adicional basándose en el biomarcador dado implicado. Por ejemplo, basándose en la observación histórica de la heterogeneidad de la expresión de proteínas para un biomarcador dado, se puede usar un coeficiente u otra técnica de normalización.

Un cálculo de heterogeneidad ejemplar con un factor de normalización α (por ejemplo, con un intervalo de cero a uno) para una métrica de variabilidad VM (por ejemplo, desviación estándar o similar) y S (por ejemplo, una puntuación de preparaciones con positividad en porcentaje promedio) puede ser de esta manera:

$$H = \begin{cases} \alpha * \frac{VM}{0,05}, S < 10 \% \\ \alpha * \frac{VM}{S}, \text{de otro modo} \end{cases}$$

Un factor de normalización por biomarcador, α, puede explicar las diferencias en el impacto que tiene la heterogeneidad de un biomarcador específico en la puntuación de pronóstico. Como se observa en los datos para el análisis de heterogeneidad, dichos factores de normalización por biomarcador se pueden adaptar.

En un ejemplo, α era 1,0 para ER y PR, y 0,75 para Ki-67. Dicho factor se puede determinar por medio de procedimientos estadísticos para determinar las ponderaciones apropiadas para el o los biomarcadores implicados (por ejemplo, para separar a los pacientes en categorías de pronóstico basándose en la puntuación de pronóstico resultante).

40 **Mediciones ejemplares de expresión de proteínas**

En cualquiera de los ejemplos del presente documento, la expresión de proteínas para un biomarcador dado se puede cuantificar midiendo el grado en el que se expresa la proteína (por ejemplo, en un campo de visión o colectivamente en campos de visión). Con referencia a las células teñidas en un campo de visión, la expresión de proteínas se puede expresar como un porcentaje de las células en el campo de visión con tinción positiva. Por ejemplo, se puede calcular una cuantificación, PP(FS), de expresión de proteínas para un campo de visión, FS, para determinar la positividad en porcentaje como sigue:

50 PP(FS) = (recuento de células con tinción positiva en FS) / (recuento de células totales en FS)

En dicho caso, el recuento de células totales puede ser una suma del recuento de células con tinción positiva y el recuento de células con tinción negativa. En la práctica, se puede determinar un recuento de células con tinción positiva, se puede determinar un recuento de células con tinción negativa y, basándose en ambos, se puede calcular la expresión de proteínas. Dado que la cuantificación indica el porcentaje de células con tinción positiva, dicha medición se denomina en ocasiones "positividad en porcentaje" para un campo de visión (por ejemplo, para un biomarcador dado).

60 Se pueden usar otras técnicas para medir la expresión de proteínas.

Se puede usar una puntuación en intervalos de clase como se describe en el presente documento para medir la

expresión de proteínas (por ejemplo, para HER2).

Determinación ejemplar de la positividad en porcentaje

5 En cualquiera de los ejemplos del presente documento, la positividad en porcentaje para un campo de visión dado se puede determinar por medio de un cierto número de técnicas. Las tecnologías descritas en el presente documento pueden hacer uso de técnicas actuales y futuras para determinar la positividad en porcentaje.

10 Algunas de dichas técnicas realizan recuentos de núcleos celulares de células con tinción positiva y realizan recuentos de núcleos celulares de células con tinción negativa. Dicha tecnología puede filtrar las regiones estromales y linfocitarias del campo de visión. Si se desea, la información de una preparación completa se puede usar al realizar recuentos de núcleos.

15 De esta manera, la heterogeneidad de proteínas se puede determinar determinando las mediciones de positividad en porcentaje para los campos de visión respectivos para una imagen asociada a un biomarcador, donde las mediciones de positividad en porcentaje indican un porcentaje de células en los campos de visión respectivos con tinción positiva para el biomarcador. A continuación, las mediciones de positividad en porcentaje se pueden comparar como se describe en el presente documento (por ejemplo, para generar una métrica de variabilidad).

20 Además, una determinación de positividad en porcentaje puede incluir determinar la intensidad de la tinción y determinar un recuento de núcleos para una pluralidad de intervalos de clase (por ejemplo, 0, 1, 2, 3). Los núcleos en el intervalo de clase 0 se pueden considerar negativos, y los demás se pueden agregar como las células con tinción positiva. Disponer de dicha información de intervalos de clase puede ser útil para otros propósitos, tales como calcular una puntuación H o similar.

25 **Determinación ejemplar de la implementación de la positividad en porcentaje: identificación de núcleos**

30 La FIG. 11 es un diagrama de flujo de un procedimiento implementado por ordenador ejemplar 1100 para identificar núcleos en una preparación. En el ejemplo, el procedimiento 1100 puede usar información sobre la preparación fuera de los campos de visión para el análisis.

35 En 1110, se recibe una imagen obtenida de preparaciones de una sección de tejido mamario teñida como se describe en el presente documento. Dicha obtención de imagen se puede realizar en una primera ampliación (por ejemplo, 20x, 40x, o similar) y/o resolución.

En 1120, se identifica la región de tejido de la preparación. La imagen obtenida de preparaciones se puede separar en fondo de vidrio y región de tejido. Dicho análisis se puede realizar en una segunda ampliación (por ejemplo, 1x, 2x, o similar) y/o resolución.

40 En 1130, se puede realizar un análisis de componentes de colores dominantes en la región de tejido. Por ejemplo, los colores dominantes se pueden extraer de la región de tejido en la preparación. Dicho análisis se puede realizar en una tercera ampliación (por ejemplo, 5x o similar) y/o resolución.

45 En 1140, se puede realizar una segmentación para segmentar la región de tejido en diferentes regiones marcadas usando características de color y textura a partir de múltiples ampliaciones.

En 1150, los núcleos en la preparación y/o campos de visión se detectan y se realiza un recuento (por ejemplo, de acuerdo con si tienen tinción positiva o negativa). De esta manera, se puede lograr una puntuación.

50 Se pueden usar otras técnicas para identificar núcleos. Por ejemplo, para identificar núcleos se puede usar cualquier técnica que use diferenciación de colores (por ejemplo, entre marrón positivo y azul negativo) y forma (por ejemplo, análisis morfológico) de núcleos candidatos. En dichas técnicas, el posprocesamiento morfológico puede filtrar células estromales y masas desviadas.

55 **Determinación ejemplar de la implementación de la positividad en porcentaje: puntuación del FOV**

60 La FIG. 12 es un diagrama de bloques de un sistema ejemplar 1200 para la puntuación del campo de visión. En el ejemplo, se muestra un campo de visión de entrada 1210. Sin embargo, dicho análisis se puede realizar para múltiples campos de visión, una o más placas de la preparación, toda la preparación o cualquier otra imagen de entrada.

La segmentación de núcleos (por ejemplo, identificación de núcleos en una imagen) 1220 se puede realizar para identificar núcleos en el FOV. Se puede usar cualquier número de técnicas para identificar los núcleos.

65 A continuación, se puede realizar la clasificación de núcleos 1230 para clasificar los núcleos identificados. De nuevo, se puede usar cualquier número de técnicas para clasificar los núcleos como con tinción positiva o negativa.

En 1240, se realiza la puntuación en los núcleos clasificados, dando como resultado la positividad en porcentaje para el biomarcador. Al identificar los núcleos, se identifican las células correspondientes. Se puede realizar la puntuación para un campo de visión particular, incluso si el análisis se realiza en áreas fuera del campo de visión.

5

Determinación ejemplar de la implementación de la positividad en porcentaje: puntuación del FOV

La FIG. 13 es un diagrama de bloques de otro sistema ejemplar 1300 para la puntuación del campo de visión. En el ejemplo, se muestra un campo de visión de entrada 1310. Sin embargo, dicho análisis se puede realizar para múltiples campos de visión, una o más placas de la preparación, toda la preparación o cualquier otra imagen de entrada.

10

Se puede realizar la segmentación de núcleos 1320 (por ejemplo, identificación de núcleos en una imagen). En 1322, se realiza la detección de semillas para la imagen de entrada para localizar las semillas procesadas por el sistema. Una semilla es un punto que se encuentra dentro de un núcleo candidato y sirve como punto de partida para localizar núcleos celulares. La detección de semillas puede usar técnicas tales como funcionar en una imagen de gradiente usando un procedimiento de votación basado en núcleos. La segmentación en primer plano se realiza en 1324. Se puede calcular un enmascaramiento en primer plano asociado a regiones de núcleos. Se realiza la segmentación de masas en 1326. La segmentación en primer plano y la segmentación de masas pueden hacer uso de las variaciones de intensidad global de una imagen a otra. Se puede extraer una representación tipo masa para cada núcleo.

15

20

Se puede realizar la clasificación de núcleos 1330. En 1332, se realiza la construcción de características. En 1336, se realiza la clasificación, con referencia a un modelo 1334 de entrenamiento. Dicho modelo puede ser específico para el biomarcador implicado.

25

En 1340, se realiza la puntuación, dando como resultado una positividad en porcentaje para el biomarcador. Al identificar los núcleos, se identifican las células correspondientes. Se puede realizar la puntuación para un campo de visión particular, incluso si el análisis se realiza en áreas fuera del campo de visión.

30

Determinación ejemplar de la implementación de la positividad en porcentaje: clasificación

La FIG. 14 es un diagrama de bloques de un clasificador ejemplar 1400 para determinar núcleos teñidos verdaderos positivos. En el ejemplo, los núcleos candidatos (por ejemplo, candidatos entrantes) se procesan mediante un clasificador 1410, que los separa en dos clases: positivos 1451 y negativos 1452. Es posible que algunos candidatos entrantes no se clasifiquen en ninguna clase, que se filtren por adelantado o similar. El clasificador puede determinar como negativas 1452 las células (por ejemplo, asociadas con núcleos) que se determinan que son células que no se tiñen. Algunos candidatos (por ejemplo, tejido estromal, linfocitos, etc.) pueden ser distintos de células que no se incluyen en (por ejemplo, se filtra de) la clasificación de negativas 1452 o la clasificación de positivas 1451.

35

40

El clasificador puede usar características tales como características de color (por ejemplo, media, varianza y similares), color de fondo de tejido y contexto, forma (por ejemplo, tamaño, excentricidad, alargamiento), morfología, densidad celular y similares.

45

Determinación ejemplar de la implementación de la positividad en porcentaje: clasificación

La FIG. 15 es un diagrama de bloques de un clasificador binario lineal multietapa 1500 ejemplar para determinar núcleos teñidos verdaderos positivos. En el ejemplo, los núcleos candidatos (por ejemplo, masas entrantes) se procesan mediante un clasificador 1510, que los separa en dos clases preliminares:

50

1) marrón (verdaderos positivos) y masas descartables débiles; y

2) negativos (por ejemplo, núcleos no teñidos), linfocitos y estroma.

55

Un clasificador 1520 separa los candidatos entrantes en estroma 1554 y negativos y linfocitos.

Otro clasificador 1530 separa los candidatos entrantes en negativos 1552 y linfocitos 1553.

60

Todavía otro clasificador 1540 separa los candidatos entrantes en marrón (verdaderos positivos) 1551 y masas descartables marrones débiles 1555.

El recuento de verdaderos positivos 1551 y negativos 1552 se puede usar para determinar la positividad en porcentaje para un campo de visión.

65

Los clasificadores pueden usar características tales como características de color (por ejemplo, media, varianza y similares), color de fondo de tejido y contexto, forma de masa (por ejemplo, tamaño, excentricidad, alargamiento),

morfología, densidad celular y similares.

Determinación ejemplar de la puntuación para HER2 en intervalos de clase

5 En cualquiera de los ejemplos del presente documento se puede generar una puntuación en intervalos de clase para el biomarcador HER2. Para una puntuación IHC combinada, los campos de visión para HER2 se pueden considerar colectivamente (por ejemplo, se puede realizar colectivamente el análisis en los campos de visión digitales). Para las puntuaciones de heterogeneidad, se puede generar una puntuación en intervalos de clase para los campos de visión respectivos dentro de una imagen de una preparación de tejido para el biomarcador HER2. En la práctica, dicha puntuación determina la completitud de la tinción para la membrana celular que rodea un núcleo.

10 Una puntuación en intervalos de clase puede comprender un único número (por ejemplo, 0, 1, 2 o 3) para HER2 (por ejemplo, para los FOV colectivamente o bien FOV respectivos) determinado como se describe en el presente documento.

15 **Procedimiento ejemplar de puntuación para HER2 en intervalos de clase**

La FIG. 16 es un diagrama de flujo de un procedimiento ejemplar 1600 para determinar una puntuación para HER2 en intervalos de clase y se puede usar en cualquiera de los ejemplos del presente documento en los que se usa una puntuación para HER2 en intervalos de clase. Dicha puntuación se puede generar colectivamente para campos de visión o para campos de visión separados como se describe en el presente documento.

20 El procedimiento 1600 se puede realizar para una imagen dada de una preparación preparada y anotada como se describe en el presente documento para HER2. En 1610, el tejido en la imagen se segmenta. Por ejemplo, se pueden determinar áreas estromales y no estromales.

25 En 1620 se pueden detectar los núcleos como se describe en el presente documento, por ejemplo, en las regiones no estromales (por ejemplo, glándula).

30 En 1630 se clasifican los núcleos como teñidos o contrateñidos.

En 1640 se detecta la membrana celular teñida en la imagen (por ejemplo, imagen completa, FOV o similar).

35 En 1650 se puntúan las células asociando los núcleos respectivos con la membrana teñida alrededor de los mismos. Basándose en la presencia de la membrana teñida alrededor de los núcleos, una célula se clasifica en uno de los siguientes tipos: no teñida (por ejemplo, no se encuentra membrana teñida alrededor del núcleo), parcialmente teñida (por ejemplo, el núcleo de la célula está rodeado parcialmente por la membrana teñida), completamente teñida (por ejemplo, el núcleo está completamente rodeado por la membrana teñida). Los parámetros se pueden usar como umbrales para determinar la tinción parcial y completa. Si se desea, un usuario puede ajustar dichos parámetros (por ejemplo, un porcentaje alrededor requerido para calificar como parcial, un porcentaje alrededor requerido para calificar como completa). Por ejemplo, se puede determinar que las células con más de una cantidad umbral (por ejemplo, un 90 %) de membrana teñida están teñidas “completamente”.

40 En la práctica, para las células detectadas (por ejemplo, correspondientes a los núcleos detectados), a una célula se le puede asignar un indicador de completitud de tinción (por ejemplo, que indica si la tinción está completa o no, el grado de completitud o similar) y un valor de intensidad de tinción basándose en la intensidad de los píxeles para el componente de color de tinción (por ejemplo, marrón) (por ejemplo, que varía desde 0 a un valor máximo, tal como 100, 255 o similar). A continuación, dicha información se puede analizar para determinar la puntuación para HER2 como se describe en el presente documento.

45 De esta manera, para una pluralidad de células que aparecen en los campos de visión digitales para HER2, se pueden determinar los valores de intensidad de tinción y los indicadores de completitud de tinción respectivos. A continuación, se pueden analizar los valores de intensidad de tinción y los indicadores de completitud de tinción por medio de las condiciones como se describe en el presente documento.

50 Para una pluralidad de núcleos que aparecen en los campos de visión digitales para HER2, se puede determinar la intensidad de tinción de la membrana celular alrededor de los núcleos respectivos.

55 En 1660, los campos de visión (por ejemplo, colectiva o respectivamente) se puntúan basándose en las puntuaciones de células en el campo de visión. Por ejemplo, se pueden usar las siguientes condiciones (por ejemplo, si el campo de visión califica una puntuación más alta, no se necesita hacer un procesamiento para las puntuaciones más bajas):

Condición	Puntuación asignada
-----------	---------------------

Respuesta a la determinación del porcentaje de células completamente teñidas > un umbral (por ejemplo, un 30 %)	3+
Respuesta a la determinación (porcentaje de células completamente teñidas > un umbral (por ejemplo, un 10 %)) O (porcentaje de células completamente teñidas es > 0 % Y la mediana de la intensidad de membrana es más baja que el umbral de intensidad fuerte)	2+
Respuesta a la determinación (porcentaje de células parcialmente teñidas > 0 %) O (porcentaje de células completamente teñidas > 0% Y porcentaje de células completamente teñidas < un umbral (por ejemplo, un 10 %) Y mediana de la intensidad de membrana >= umbral de intensidad fuerte Y mediana de la intensidad de membrana < umbral de intensidad débil)	1+
Respuesta a la determinación de las condiciones anteriores, que no se alcanzan (por ejemplo, de otro modo)	0

Como se describe, la mediana de la intensidad de membrana se puede comparar con un umbral de intensidad fuerte. También se puede determinar, para una pluralidad de núcleos que aparecen en campos de visión digitales para HER2, si la membrana celular alrededor de los núcleos se tiñe completamente.

5 De esta manera, se puede lograr la puntuación en intervalos de clase para HER2 del campo de visión. Si se desea una puntuación binaria, se puede omitir o combinar la etapa intermedia de determinar la puntuación en intervalos de clase en el procedimiento.

10 **Intensidad de tinción ejemplar**

En cualquiera de los ejemplos del presente documento, se puede cuantificar la tinción positiva (por ejemplo, de la membrana celular) mediante una coloración marrón promedio. Por ejemplo, se puede determinar una coloración marrón promedio en la membrana mediante el valor escalar de luminancia (por ejemplo, 0-255), que se puede calcular a partir de los valores de píxeles RGB de la región de membrana. Los valores RGB se pueden convertir en valor de luminancia (L) por medio de una técnica de conversión RGB a $L^*a^*b^*$. El valor L se puede convertir desde una escala de 0 a 100 a una escala de 0 a 255.

20 Se puede calcular una coloración marrón promedio para una célula promediando L sobre los píxeles de membrana. La coloración marrón promedio se puede usar como la mediana de la intensidad (por ejemplo, de la membrana celular) como se describe anteriormente.

Clasificación de núcleos ejemplar

25 En cualquiera de los ejemplos del presente documento se pueden clasificar los núcleos con el propósito de calcular la puntuación para HER2 en intervalos de clase. Por ejemplo, las células se pueden clasificar como teñidas y contratañidas, basándose en la información de saturación, intensidad y de colores rojo y azul en cada célula a partir de la imagen. La información de saturación e intensidad se puede usar para distinguir las células gris oscuro, de modo que se puedan clasificar como teñidas en lugar de no teñidas.

30 Por ejemplo, se pueden usar las siguientes reglas:
 A. Porcentaje de tinción 1 = $100 * (R - G) / (R + 1)$
 B. Porcentaje de tinción 2 = $100 * (R - B) / (R + 1)$
 C. Si el porcentaje de tinción 1 > % de tinción a nivel de píxel y porcentaje de tinción 2 > % de tinción a nivel de píxel, entonces píxel = teñido
 D. Si saturación < 45 e intensidad < 128, entonces píxel = teñido
 E. Si saturación < 128 e intensidad < 55, entonces píxel = teñido
 F. Si saturación < 255 e intensidad <= 30, entonces píxel = teñido
 G. Si intensidad <= 20, entonces píxel = teñido

40 Usando la información de clasificación de píxeles, cada núcleo se puede clasificar como teñido o no teñido basándose en el porcentaje de los píxeles teñidos dentro de cada núcleo. Por ejemplo, se pueden usar las siguientes reglas:

45 A. Si el umbral a nivel celular > % de los píxeles teñidos dentro de cada núcleo, entonces clasificar la célula como teñida
 B. Si no, clasificarla como no teñida
 Los objetos de los núcleos se pueden filtrar entonces basándose en el tamaño.

50 **Controles de puntuación para HER2 en intervalos de clase ejemplares**

Se pueden usar cuatro controles de líneas celulares para determinar si las puntuaciones para HER2 en intervalos de clase se determinan como se describe en el presente documento. Cuando se procesan y tiñen como se describe en el presente documento, las líneas celulares deben teñirse como se describe en la siguiente tabla. Las líneas celulares están disponibles de Ventana Medical Systems, Inc. como en el catálogo n.º 781-2991.

Puntuación IHC para HER2	Línea celular	N.º de copias del gen HER2
0	MCF-7	1,7
1+	T47D	2,9
2+	MDA-MB-453	5,2
3+	BT-474	18,9

El n.º de copias del gen HER2 se determinó como un promedio de tres lotes de preparaciones de control PATHWAY HER-2 4 en 1 que se determinaron usando la sonda de HER2 PathVysion®.

Se puede encontrar información adicional sobre la puntuación para HER2 en la documentación para "PATHWAY® anti-HER-2/neu (4B5) Rabbit Monoclonal Primary Antibody", con número de catálogo 790-2991, disponible de Ventana Medical Systems, Inc., que se incorpora por referencia en el presente documento.

Cálculo de la puntuación H ejemplar

En cualquiera de los ejemplos del presente documento se puede calcular una puntuación H a partir de un campo de visión o campos de visión colectivamente por medio de técnicas automatizadas. Por ejemplo, se pueden identificar los núcleos, y entonces las células (por ejemplo, correspondientes a los núcleos) se pueden categorizar (por ejemplo, en intervalos de clase) en cuatro intervalos de clase: 0, 1 (con tinción débil), 2 (con tinción moderada) o 3 (con tinción fuerte). Dicha técnica se puede lograr determinando la coloración marrón promedio de una célula o agrupamiento de células (por ejemplo, masa).

Al examinar los componentes de color (por ejemplo, RGB o similar), se puede comparar la coloración marrón promedio con los umbrales para determinar si una célula no se tiñe (0), se tiñe débilmente (1+), se tiñe moderadamente (2+) o se tiñe fuertemente (3+). Se pueden mantener los recuentos de las células en los intervalos de clase, y al concluir el recuento, se puede calcular una puntuación H basándose en el porcentaje de células en cada intervalo de clase como se describe en el presente documento.

Por ejemplo, dados los recuentos de células en una pluralidad de intervalos de clase asociados a intensidades de tinción respectivas (por ejemplo, incluyendo un intervalo de clase para intensidad cero), se puede calcular la puntuación H sumando, para los intervalos de clase, el producto del porcentaje (por ejemplo, porcentaje de recuento de células, tal como el número de células en un intervalo de clase dividido entre el número total de células) de células en un intervalo de clase por el nivel de intensidad respectivo asociado al intervalo de clase. En una disposición que tiene cuatro intervalos de clase, 0-3, la puntuación H oscila desde 0-300. La puntuación máxima en dicha disposición se logra si todas (100 %) las células son de intensidad 3 (por ejemplo, $0 \times 0 + 0 \times 1 + 0 \times 2 + 100 \times 3 = 300$). De esta manera, el porcentaje se puede tratar como un número entero en lugar de una fracción (por ejemplo, $50 \% = 50$). El nivel de intensidad asociado a un intervalo de clase puede ser la porción numérica de un designador de intensidad (por ejemplo, 2+ se trata como 2).

Puntuación de pronóstico de recidiva de cáncer de mama ejemplar

En cualquiera de los ejemplos del presente documento, se puede generar una puntuación de pronóstico de recidiva de cáncer de mama basándose en una puntuación IHC combinada y puntuaciones de heterogeneidad para uno o más biomarcadores. Dicha puntuación de pronóstico se denomina a veces simplemente "puntuación de pronóstico de salida".

La puntuación se puede generar combinando las puntuaciones IHC combinadas y las puntuaciones de heterogeneidad. Por ejemplo, las puntuaciones de heterogeneidad se pueden añadir, multiplicar o de otro modo combinar y, entonces, aplicar operaciones (por ejemplo, raíz cuadrada, exponente, coeficientes, etc.) para formar un componente de heterogeneidad con respecto a la puntuación.

La puntuación inmunohistoquímica combinada también se puede ajustar (por ejemplo, raíz cuadrada, exponente, coeficientes, etc.) para formar un componente inmunohistoquímico combinado con respecto a la puntuación.

Entonces, los dos componentes se pueden combinar (por ejemplo, añadir, multiplicar o similar).

Una técnica particular para calcular la puntuación de pronóstico de salida es como sigue:

$$F_{IHC} * IHC4 + F_{Het} * \sqrt{Het_{ER} + Het_{PR}}$$

5 donde F_{IHC} es un coeficiente inmunohistoquímico combinado, F_{Het} es un coeficiente de heterogeneidad y Het_{ER} es una puntuación de heterogeneidad para ER y Het_{PR} es una puntuación de heterogeneidad para PR. En una implementación, $F_{IHC} = 0,03114$ y $F_{HET} = 1,95119$. Sin embargo, las tecnologías pueden soportar el ajuste de los coeficientes según se desee. Por ejemplo, el análisis estadístico puede indicar coeficientes apropiados que se pueden cambiar en función de los datos acumulados (por ejemplo, para dividir a los pacientes en categorías como se describe en el presente documento).

Indicación de pronóstico ejemplar

15 En cualquiera de los ejemplos del presente documento, una indicación de pronóstico puede comprender una puntuación de pronóstico de salida por sí misma, una categoría (por ejemplo, rojo, amarillo, verde) basándose en umbrales o similar. La indicación puede ser de una forma que indique si se predice o no (o cuán probable es) que el cáncer recidive, basándose en las entradas.

20 La indicación de la recidiva de cáncer de mama se puede elegir (por ejemplo, entre una pluralidad de categorías) basándose en uno o más umbrales almacenados para la puntuación de pronóstico de recidiva de cáncer de mama. Por ejemplo, si la puntuación alcanza un umbral, se puede elegir una primera indicación (por ejemplo, pronóstico de recidiva); si la puntuación no alcanza el umbral, se puede elegir una segunda indicación (por ejemplo, pronóstico de no recidiva).

25 Dichas categorías también pueden comprender un componente de tiempo (por ejemplo, se espera que recidive en un plazo de 5 años) o similar.

Se pueden soportar múltiples indicaciones (por ejemplo, una categoría para si se espera que el cáncer recidive en un plazo de 1 año, otra categoría para si se espera que el cáncer recidive en un plazo de 5 años, etc.).

Implementación ejemplar en la plataforma de formación de imágenes

35 Aunque las tecnologías se pueden implementar de forma independiente, las tecnologías descritas en el presente documento se pueden implementar en una plataforma de patología digital multifuncional que ayuda a los usuarios en el análisis de los datos de imágenes de preparaciones. Por ejemplo, las tecnologías se pueden integrar en el software de gestión de imágenes Virtuoso de Ventana Medical Systems, Inc. o soluciones comparables.

40 Dicha plataforma puede servir como una parte visible para las tecnologías. Las imágenes de preparaciones y la información recopilada sobre ellas (por ejemplo, campos de visión) por medio de la plataforma de patología digital se pueden usar como entradas para las tecnologías descritas en el presente documento.

45 Dicha plataforma puede proporcionar a los patólogos la flexibilidad de trabajar en casos, especímenes e imágenes en cualquier orden deseado. Las imágenes de preparaciones se pueden anotar, y puede estar disponible una variedad de otras funcionalidades.

La plataforma se puede implementar en un escenario de cliente liviano (por ejemplo, software como servicio, informática en la nube o similar) según se desee.

Sistema ejemplar con herramienta de heterogeneidad que implementa las tecnologías

50 La FIG. 17 es un diagrama de bloques de un sistema ejemplar 1700 que incluye una herramienta de heterogeneidad 1720 que implementa las tecnologías para el pronóstico de cáncer de mama descritas en el presente documento. En el ejemplo, una herramienta de heterogeneidad 1720 acepta una pluralidad de campos de visión digitales 1711A-N, 1714A-N para biomarcadores respectivos como entrada y, con la ayuda de un normalizador opcional 1725, arroja puntuaciones de heterogeneidad 1731, 1734 para los biomarcadores respectivos.

Las características de la FIG. 9 se pueden entremezclar o incorporar según se desee.

60 Una herramienta de pronóstico 1740 acepta las puntuaciones de heterogeneidad 1731, 1734 y la puntuación IHC combinada 1750 como entrada y, con referencia a un umbral de pronóstico 1745, arroja una indicación 1790 de si es probable que el cáncer recidive en el sujeto.

Procedimiento implementado por ordenador ejemplar que implementa las tecnologías por medio de la puntuación de heterogeneidad para biomarcadores respectivos

65

La FIG. 18 es un diagrama de flujo de un procedimiento implementado por ordenador ejemplar 1800 que implementa las tecnologías para el pronóstico de cáncer de mama descritas en el presente documento por medio de la determinación de una puntuación de heterogeneidad para biomarcadores respectivos y se puede implementar, por ejemplo, en el sistema mostrado en la FIG. 17.

5 En 1810 se recibe una pluralidad de campos de visión digitales para biomarcadores respectivos.

10 En 1812, basándose en los campos de visión digitales, se determina una puntuación de heterogeneidad para los biomarcadores respectivos.

15 En 1816, una puntuación IHC combinada y las puntuaciones de heterogeneidad se combinan en una puntuación de pronóstico.

20 En 1818 se arroja una indicación del pronóstico (por ejemplo, una indicación 1790 de si es probable que el cáncer recidive en el sujeto) basándose en la puntuación de pronóstico.

Cálculo de la puntuación IHC combinada ejemplar

25 Un procedimiento implementado por ordenador puede calcular una puntuación inmunohistoquímica combinada para un sujeto recibiendo respectivas pluralidades de campos de visión digitales dentro de imágenes respectivas que representan una muestra de cáncer de mama marcada de manera detectable con anticuerpos para biomarcadores respectivos (por ejemplo, para una pluralidad de biomarcadores como se describe en el presente documento).

30 A continuación, se puede determinar la positividad en porcentaje para una pluralidad de los biomarcadores como se describe en el presente documento.

35 A continuación, se puede calcular la puntuación inmunohistoquímica combinada. Dicho cálculo puede comprender combinar la positividad en porcentaje para un biomarcador con la positividad en porcentaje para un segundo biomarcador.

A continuación, se puede arrojar una puntuación inmunohistoquímica combinada.

Los biomarcadores pueden comprender HER2, y calcular la puntuación IHC combinada puede comprender calcular una puntuación para HER2 (por ejemplo, en intervalos de clase, binaria o similar) como se describe en el presente documento.

Biomarcadores ejemplares

40 En el ejemplo de la FIG. 17 se usan los biomarcadores ER y PR para determinar la heterogeneidad regional. En la práctica, la heterogeneidad de diferentes combinaciones de biomarcadores se puede calcular y combinar con una puntuación IHC combinada para generar una puntuación de pronóstico.

45 Los biomarcadores que se pueden usar pueden incluir combinaciones de uno o más de los biomarcadores usados como parte de la puntuación IHC combinada (por ejemplo, ER, PR, Ki-67 y HER2). Cualquiera de los demás ejemplos del presente documento se puede modificar en consecuencia (por ejemplo, para obtener campos de visión, calcular puntuaciones de heterogeneidad o similar).

Multiplexación ejemplar

50 En cualquiera de los ejemplos del presente documento se pueden medir biomarcadores en un escenario multiplex. Por ejemplo, como se describe en el presente documento, una única sección de tejido se puede marcar con más de un anticuerpo (tal como dos anticuerpos específicos para dos biomarcadores diferentes, etc.), siempre y cuando dichas proteínas marcadas sean distinguibles, por ejemplo, usando anticuerpos secundarios marcados de manera diferente.

Técnicas de micromatrices de tejidos ejemplares

60 En cualquiera de los ejemplos del presente documento se pueden aplicar técnicas de micromatrices de tejidos para conseguir la tinción de tejidos. Por ejemplo, se puede aplicar una pluralidad de muestras de tejido a un único sustrato, y se puede aplicar una pluralidad de agentes de tinción (por ejemplo, anticuerpos de marcaje) a las muestras (por ejemplo, un agente diferente por cada muestra de tejido). A continuación, se pueden usar las técnicas descritas en el presente documento para determinar una o más puntuaciones como se describe en el presente documento (por ejemplo, basándose en imágenes de las muestras de tejido marcadas de manera detectable).

Umbrales ejemplares

En cualquiera de los ejemplos del presente documento se pueden derivar umbrales mediante evaluación estadística. Por ejemplo, se pueden determinar puntos de corte usando entrenamiento y pruebas para validar un punto de corte óptimo que sea el que mejor defina a los pacientes con riesgo alto frente a bajo.

5 **Sistema ejemplar con máquina diferencial que implementa las tecnologías**

La FIG. 19 es un diagrama de bloques de un sistema ejemplar 1900 que incluye una máquina diferencial 1940 para calcular una puntuación de heterogeneidad para un biomarcador. Las características de la FIG. 9 y/o la FIG. 17 se pueden incorporar o entremezclar según se desee.

10 En el ejemplo, un analizador de imágenes de campo de visión 1920 acepta campos de visión digitales 1711A-N (por ejemplo, para un único biomarcador) como entrada y arroja mediciones de positividad en porcentaje 1931A-N para los campos de visión respectivos.

15 La máquina diferencial 1940 acepta las mediciones de positividad en porcentaje 1931A-N como entrada y, con ayuda de un normalizador opcional 1945, arroja una puntuación de heterogeneidad 1731 para el biomarcador.

Procedimiento implementado por ordenador ejemplar para calcular la puntuación de heterogeneidad

20 La FIG. 20 es un diagrama de flujo de un procedimiento implementado por ordenador ejemplar 2000 para calcular una puntuación de heterogeneidad para su uso en las tecnologías para el pronóstico de cáncer de mama descritas en el presente documento y se puede implementar, por ejemplo, en el sistema mostrado en la FIG. 19.

25 En 2010 se recibe una pluralidad de campos de visión digitales para un biomarcador dado. Dicha imagen puede representar una muestra de cáncer de mama de un sujeto marcada de manera detectable con anticuerpos para un biomarcador como se describe en el presente documento.

30 En 2012 se mide la expresión de proteínas para el biomarcador en los campos de visión digitales. Por ejemplo, se reciben o determinan mediciones de positividad en porcentaje para los biomarcadores respectivos.

35 En 2016 se mide la heterogeneidad de la expresión de proteínas medida para el biomarcador entre la pluralidad de campos de visión digitales. Por ejemplo, se calcula una puntuación de heterogeneidad como la variabilidad entre las mediciones de positividad en porcentaje que aparecen en los campos de visión diferentes (por ejemplo, para un único biomarcador).

La puntuación de heterogeneidad (por ejemplo, métrica de variabilidad o puntuación de heterogeneidad preliminar) se puede normalizar como se describe en el presente documento para ajustar la puntuación de heterogeneidad.

40 En 2020 se arroja la puntuación de heterogeneidad para el biomarcador (por ejemplo, para su uso en las tecnologías descritas en el presente documento).

Interfaz de usuario ejemplar

45 La FIG. 21 es una captura de pantalla de una interfaz de usuario ejemplar 2100 para indicar campos de visión digitales dentro de una imagen. Una interfaz de usuario puede adoptar la forma de una ventana, pantalla, panel o similar presentado a un usuario para aceptar entradas (por ejemplo, indicaciones de campos de visión en una imagen de preparación visualizada).

50 En el ejemplo, un patólogo ha seleccionado tres campos de visión 2120A, 2120B y 2120C en una preparación visualizada para un biomarcador particular (por ejemplo, ER) con propósitos de heterogeneidad. Se pueden seleccionar campos de visión independientes para diferentes biomarcadores. Se pueden seleccionar campos de visión independientes adicionales con propósitos de heterogeneidad y propósitos de puntuación inmunohistoquímica combinada.

55 **Flujo de trabajo del patólogo ejemplar**

La FIG. 22 es un diagrama de flujo de un procedimiento implementado por ordenador 2200 del flujo de trabajo digital del patólogo ejemplar. Se puede reforzar un flujo de trabajo digital del patólogo para facilitar la selección de campos de visión en (por ejemplo, dentro de) las imágenes.

60 En 2210, por medio de una interfaz de usuario que requiere campos de visión para una puntuación IHC combinada, se recibe una pluralidad de campos de visión para la puntuación IHC combinada dentro de una pluralidad de imágenes de preparaciones (o tejidos) visualizadas para determinar la puntuación IHC combinada. Dicha interfaz de usuario puede comprender una pluralidad de imágenes de preparaciones (o tejidos) que representan una muestra de cáncer de mama de un sujeto marcada de manera detectable con anticuerpos para cada uno de ER, HER2, Ki-67 y PR. Recibir la pluralidad de campos de visión para la puntuación IHC combinada puede comprender recibir una

65

5 pluralidad de campos de visión para cada uno de ER, HER2, Ki-67 y PR. Por ejemplo, se puede presentar una serie de preparaciones (o secciones de tejido) y el patólogo puede indicar una pluralidad de campos en las preparaciones respectivas (por ejemplo, FOV para una primera preparación, FOV para una segunda preparación, etc.) (o una pluralidad de campos para cada sección de tejido). Dicha indicación se puede conseguir recibiendo anotaciones (por ejemplo, áreas dibujadas) en una imagen visualizada.

10 En 2212, por medio de una interfaz de usuario que requiere campos de visión para una puntuación de heterogeneidad, se recibe una pluralidad de campos de visión para una puntuación de heterogeneidad dentro de una pluralidad de imágenes de preparaciones (o tejidos) visualizadas para determinar la puntuación de heterogeneidad. Dicha interfaz de usuario puede comprender una pluralidad de imágenes de preparaciones (o tejidos) que representan una muestra de cáncer de mama del sujeto marcada de manera detectable con anticuerpos para cada uno de ER y PR. Recibir la pluralidad de campos de visión para la puntuación IHC combinada puede comprender recibir una pluralidad de biomarcadores de campos de visión, tales como para cada uno de ER y PR. Dicha recepción se puede conseguir recibiendo anotaciones (por ejemplo, áreas dibujadas o identificadas) en una imagen visualizada.

15 En 2216 se miden los valores de expresión de proteínas para los campos de visión para la puntuación IHC combinada y se calcula una puntuación IHC combinada a partir de los mismos.

20 En 2218 se miden los valores de expresión de proteínas para los campos de visión para la puntuación de heterogeneidad y se calcula una puntuación de heterogeneidad a partir de los mismos. Se pueden obtener de manera similar una o más de otras puntuaciones de heterogeneidad para otros biomarcadores respectivos.

25 En 2220, la puntuación IHC combinada y las una o más puntuaciones de heterogeneidad se combinan en una puntuación de pronóstico de recidiva de cáncer de mama.

30 En 2222 se visualiza a una indicación del pronóstico de recidiva de cáncer de mama basándose en la puntuación de pronóstico de recidiva de cáncer de mama (por ejemplo, para su consideración por parte de un patólogo responsable).

Selección de campos de visión ejemplares

35 En cualquiera de los ejemplos del presente documento, los campos de visión para una puntuación IHC combinada se pueden seleccionar para que sean representativos de toda la preparación. Una interfaz de usuario presentada al patólogo puede indicarlo así.

40 En cualquiera de los ejemplos del presente documento, los campos de visión para una puntuación de heterogeneidad se pueden seleccionar para que sean representativos de una región en la preparación (por ejemplo, y de esta manera elegidos para enfatizar diferencias entre regiones). Una interfaz de usuario presentada al patólogo puede indicarlo así.

Descripción adicional ejemplar

A. Muestras biológicas

45 Los procedimientos para obtener una muestra biológica de un sujeto se conocen en la técnica. Por ejemplo, los procedimientos para obtener tejido mamario o células mamarías son rutinarios. Por ejemplo, se puede obtener una muestra de un tumor que contenga material celular mediante extirpación quirúrgica de todo o parte del tumor, recogiendo un aspirado con aguja fina del tumor, así como otros procedimientos conocidos en la técnica.

50 Las muestras pueden ser frescas o procesadas después de su obtención. En algunos ejemplos, se pueden fijar las muestras procesadas (por ejemplo, fijar en formalina) y/o someter a inclusión en cera (por ejemplo, parafina). Los fijadores para las preparaciones de células y tejidos montados se conocen bien en la técnica e incluyen, sin limitación, fijador de Bouin alcohólico al 95 %; fijador de alcohol al 95 %; fijador B5, fijador de Bouin, fijador de formalina, fijador de Karnovsky (glutaraldehído), fijador de Hartman, fijador de Hollande, solución de Orth (fijador de dicromato) y fijador de Zenker (véase, por ejemplo, Carson, *Histotechnology: A Self-Instructional Text*, Chicago:ASCP Press, 1997). De esta manera, la muestra puede ser una muestra de tejido de cáncer de mama, fijada, sometida a inclusión en cera, tal como una muestra de tejido de cáncer de mama en estadio temprano, fijada, sometida a inclusión en cera. En algunos ejemplos, la muestra es una sección de tejido mamario teñida con hematoxilina y eosina (H&E). En algunos ejemplos, la muestra es una sección de tejido mamario marcada con anticuerpos primarios específicos para ER, HER2, Ki-67 y PR, que se puede marcar directa o indirectamente (por ejemplo, con un anticuerpo secundario marcado), que en algunos ejemplos se tiñe adicionalmente con H&E.

65 En algunos ejemplos, la muestra (o una fracción de la misma) está presente sobre un soporte sólido. Los soportes sólidos portan la muestra biológica y permiten la detección apropiada de los componentes, por ejemplo, proteínas) en la muestra. Los soportes o sustratos ejemplares incluyen portaobjetos de microscopio (por ejemplo, portaobjetos

de microscopio de vidrio o portaobjetos de microscopio de plástico), cubreobjetos (por ejemplo, cubreobjetos de vidrio o cubreobjetos de plástico), placas de cultivo de tejidos, placas de pocillos múltiples, membranas (por ejemplo, nitrocelulosa o poli(fluoruro de vinilideno) (PVDF)) o chips BIACORE™.

5 **B. Biomarcadores de cáncer de mama**

1. **Receptor de estrógenos (ER) (OMIM: 133430)**

10 El receptor de estrógenos humano (ER o ESR) tiene dos formas diferentes, denominadas α (ESR1) y β (ESR2). ER α es la forma encontrada en las células de cáncer de mama. Las secuencias de ER α están disponibles públicamente, por ejemplo, de GenBank® (por ejemplo, los números de acceso NP_001116214.1, NP_001116213.1 y P03372.2 (proteínas) y NM_000125.3, NM_001122741.1 y NM_178850.2 (ácidos nucleicos)).

15 El receptor de estrógenos es un factor de transcripción activado por ligando compuesto de varios dominios importantes para la unión a hormonas, unión a ADN y activación de la transcripción. El empalme alternativo da como resultado varios transcritos de ARNm de ESR1, que difieren principalmente en sus regiones no traducidas en 5'.

20 Los anticuerpos para detectar la expresión de ER están disponibles públicamente, tal como de Santa Cruz Biotechnology (Santa Cruz, CA), Thermo Scientific Pierce Antibodies (Rockford, IL), GeneTex (Irvine, CA), ARP American Research Products, Ventana Medical Systems, Inc. (Tucson, AZ) o Abcam (Cambridge, MA), por ejemplo, los n.º de cat. sc-71064, sc-73562 y sc-7207 de Santa Cruz Biotechnology, MA1-310 de Pierce Antibodies, 790-4325 (clon SP1) de Ventana o ab2746, ab27614 o ab37438 de Abcam.

2. **Receptor del factor de crecimiento epidérmico humano 2 (HER2) (OMIM: 164870)**

25 El gen HER2 humano se localiza en el cromosoma 17 (17q21-q22). Las secuencias de HER2 están disponibles públicamente, por ejemplo, de GenBank® (por ejemplo, números de acceso NP_001005862.1, P04626.1 y NP_004439.2 (proteínas) y NM_001005862.1 y NM_001982.3 (ácidos nucleicos)).

30 La amplificación o sobreexpresión de HER2 desempeña un papel en la patogenia y evolución de determinados tipos agresivos de cáncer de mama y es un importante biomarcador y diana de tratamiento para la enfermedad.

35 Los anticuerpos para medir la expresión de HER2 están disponibles públicamente, tal como de Santa Cruz Biotechnology (Santa Cruz, CA), Thermo Scientific Pierce Antibodies (Rockford, IL), GeneTex (Irvine, CA), ARP American Research Products, Ventana Medical Systems, Inc. (Tucson, AZ) o Abcam (Cambridge, MA), por ejemplo, los n.º de cat. 790-2991 (clon 4B5) de Ventana, sc-08, sc-81528 y sc-136294 de Santa Cruz Biotechnology y ab2428, ab16901 o ab36728 de Abcam.

3. **Ki-67 (OMIM: 176741)**

40 El gen Ki-67 humano está localizado en el cromosoma 10. Las secuencias de Ki-67 están disponibles públicamente, por ejemplo, de GenBank® (por ejemplo, números de acceso NP_001139438.1, P46013.2 y NP_002408.3 (proteínas) y NM_001040058.1 y NM_001145966.1 (ácidos nucleicos)).

45 Ki-67 es un marcador que se puede usar para determinar la fracción de crecimiento de una población celular dada. La fracción de células tumorales positivas para Ki-67 (el índice de marcaje de Ki-67) se correlaciona a menudo con la evolución clínica del cáncer, tal como cáncer de mama. De esta manera, se puede usar Ki-67 para determinar el pronóstico de supervivencia del paciente y de recidiva del tumor.

50 Los anticuerpos para medir la expresión de Ki-67 están disponibles públicamente, tal como de Santa Cruz Biotechnology (Santa Cruz, CA), Thermo Scientific Pierce Antibodies (Rockford, IL), GeneTex (Irvine, CA), ARP American Research Products, Ventana Medical Systems, Inc. (Tucson, AZ) o Abcam (Cambridge, MA), por ejemplo, los n.º de cat. 790-4286 (clon 30-9) de Ventana, sc-101861 (clon MIB-1), sc-15402 y sc-23900 de Santa Cruz Biotechnology y ab15580, ab16667 o ab8191 de Abcam.

4. **Receptor de progesterona (PR) (OMIM: 607311)**

60 El gen PR humano (también conocido como subfamilia 3 de receptores nucleares, grupo C, miembro 3, NR3C3) se localiza en el cromosoma 11q22. Las secuencias de PR están disponibles públicamente, por ejemplo, de GenBank® (por ejemplo, números de acceso NP_000917.3, AAA60081.1 y AAS00096.1 (proteínas) y NM_001202474.1 y NM_000926.4 (ácidos nucleicos)).

Se ha observado amplificación o sobreexpresión de PR en algunos cánceres de mama.

65 Los anticuerpos para medir la expresión de PR están disponibles públicamente, tal como de Santa Cruz Biotechnology (Santa Cruz, CA), Thermo Scientific Pierce Antibodies (Rockford, IL), GeneTex (Irvine, CA), ARP

American Research Products, Ventana Medical Systems, Inc. (Tucson, AZ) o Abcam (Cambridge, MA), por ejemplo, los n.º de cat. 790-2223 (clon IE2) de Ventana, sc-810, sc-811 y sc-539 de Santa Cruz Biotechnology y ab2765, ab2764 o ab68195 de Abcam.

5 **5. Secuencias variantes**

Además de las secuencias disponibles públicamente de ER, HER2, Ki-67 y PR específicas proporcionadas en el presente documento, un experto en la técnica apreciará que se pueden presentar variantes de dichas secuencias en un sujeto particular. Por ejemplo, se pueden presentar polimorfismos para un gen o proteína particular. Además, una secuencia puede variar entre diferentes organismos. En ejemplos particulares, una secuencia variante retiene la actividad biológica de su secuencia natural correspondiente. Por ejemplo, una secuencia de ER, HER2, Ki-67 o PR presente en un sujeto particular puede tener cambios conservadores de aminoácidos (tales como sustituciones muy altamente conservadoras, sustituciones altamente conservadoras o sustituciones conservadoras), tales como de 1 a 5 o de 1 a 10 sustituciones conservadoras de aminoácidos. Las sustituciones conservadoras de aminoácidos ejemplares se muestran en la tabla 1.

Tabla 1: Sustituciones conservadoras de aminoácidos ejemplares

Residuo original	Sustituciones muy altamente conservadoras	Sustituciones altamente conservadoras (de la matriz Blosum90)	Sustituciones conservadoras (de la matriz Blosum65)
Ala	Ser	Gly, Ser, Thr	Cys, Gly, Ser, Thr, Val
Arg	Lys	Gln, His, Lys	Asn, Gln, Glu, His, Lys
Asn	Gln; His	Asp, Gln, His, Lys, Ser, Thr	Arg, Asp, Gln, Glu, His, Lys, Ser, Thr
Asp	Glu	Asn, Glu	Asn, Gln, Glu, Ser
Cys	Ser	Ninguna	Ala
Gln	Asn	Arg, Asn, Glu, His, Lys, Met	Arg, Asn, Asp, Glu, His, Lys, Met, Ser
Glu	Asp	Asp, Gln, Lys	Arg, Asn, Asp, Gln, His, Lys, Ser
Gly	Pro	Ala	Ala, Ser
His	Asn; Gln	Arg, Asn, Gln, Tyr	Arg, Asn, Gln, Glu, Tyr
Ile	Leu; Val	Leu, Met, Val	Leu, Met, Phe, Val
Leu	Ile; Val	Ile, Met, Phe, Val	Ile, Met, Phe, Val
Lys	Arg; Gln; Glu	Arg, Asn, Gln, Glu	Arg, Asn, Gln, Glu, Ser
Met	Leu; Ile	Gln, Ile, Leu, Val	Gln, Ile, Leu, Phe, Val
Phe	Met; Leu; Tyr	Leu, Trp, Tyr	Ile, Leu, Met, Trp, Tyr
Ser	Thr	Ala, Asn, Thr	Ala, Asn, Asp, Gln, Glu, Gly, Lys, Thr
Thr	Ser	Ala, Asn, Ser	Ala, Asn, Ser, Val
Trp	Tyr	Phe, Tyr	Phe, Tyr
Tyr	Trp; Phe	His, Phe, Trp	His, Phe, Trp
Val	Ile; Leu	Ile, Leu, Met	Ala, Ile, Leu, Met, Thr

20 En algunos modos de realización, una secuencia de ER, HER2, Ki-67 o PR es una variante de secuencia de una secuencia de ER, HER2, Ki-67 o PR natural, respectivamente, tal como una secuencia de proteína o ácido nucleico que tiene al menos un 99 %, al menos un 98 %, al menos un 95 %, al menos un 92 %, al menos un 90 %, al menos un 85 %, al menos un 80 %, al menos un 75 %, al menos un 70 %, al menos un 65 % o al menos un 60 % de identidad de secuencia con respecto a las secuencias expuestas en un número de acceso de GenBank® al que se hace referencia en el presente documento, en la que la variante resultante retiene la actividad biológica de ER, HER2, Ki-67 o PR. La "identidad de secuencia" es un término usado comúnmente para describir la similitud entre dos secuencias de aminoácidos (o entre dos secuencias de ácidos nucleicos). La identidad de la secuencia se expresa típicamente en términos de identidad en porcentaje; cuanto más alto sea el porcentaje, más similares son las dos secuencias.

30 Los procedimientos para alinear secuencias para su comparación y para determinar la identidad de secuencia se conocen bien en la técnica. Diversos programas y algoritmos de alineación se describen en: Smith y Waterman, *Adv.*

Appl. Math., 2:482, 1981; Needleman y Wunsch, *J. Mol. Biol.*, 48:443, 1970; Pearson y Lipman, *Proc. Natl. Acad. Sci. USA*, 85:2444, 1988; Higgins y Sharp, *Gene*, 73:237-244, 1988; Higgins y Sharp, *CABIOS*, 5:151-153, 1989; Corpet *et al.*, *Nucleic Acids Research*, 16:10881-10890, 1988; Huang, *et al.*, *Computer Applications in the Biosciences*, 8:155-165, 1992; Pearson *et al.*, *Methods in Molecular Biology*, 24:307-331, 1994; Tatiana *et al.*, *FEMS Microbiol. Lett.*, 174:247-250, 1999. Altschul *et al.* presentan una consideración detallada de procedimientos de alineación de secuencias y cálculos de homología (*J. Mol. Biol.*, 215:403-410, 1990).

La herramienta de búsqueda de alineaciones locales básicas (BLAST™, Altschul *et al.*, *J. Mol. Biol.*, 215:403-410, 1990) del National Center for Biotechnology Information (NCBI) está disponible públicamente de varias fuentes, incluyendo el National Center for Biotechnology Information (NCBI, Bethesda, MD) y en Internet, para su uso en relación con los programas de análisis de secuencias blastp, blastn, blastx, tblastn y tblastx. Una descripción de cómo determinar la identidad de secuencia usando este programa está disponible en Internet en la sección de ayuda para BLAST™.

Para las comparaciones de secuencias de aminoácidos de más de aproximadamente 15 aminoácidos, se emplea la función "secuencias de Blast 2" del programa BLAST™ (Blastp) usando la matriz BLOSUM62 predeterminada ajustada a los parámetros predeterminados (coste por abrir un hueco [predeterminado = 5] coste por extender un hueco [predeterminado = 2], penalización por un emparejamiento erróneo [predeterminado = 3], recompensa por un emparejamiento [predeterminado = 1], valor de expectación (E) [predeterminado = 10,0], tamaño de palabra [predeterminado = 3] y número de descripciones en una línea (V) [predeterminado = 100]. Al alinear los péptidos cortos (menos de alrededor de 15 aminoácidos), la alineación se debe realizar usando la función secuencias de Blast 2 "búsqueda de emparejamientos prácticamente exactos y cortos" que emplea la matriz PAM30 ajustada a los parámetros predeterminados (umbral esperado = 20000, tamaño de palabra = 2, costes por hueco: existencia = 9 y extensión = 1) usando la estadística basada en composición.

C. Detección de proteínas

En ejemplos particulares, se analiza una muestra obtenida del sujeto para determinar si contiene niveles detectables de las proteínas ER, HER2, Ki-67 y PR, tal como una muestra de cáncer de mama. De esta manera, la muestra se puede analizar para detectar o medir la presencia de las proteínas ER, HER2, Ki-67 y PR en la muestra, por ejemplo, una medición cualitativa o semicuantitativa. Los patrones de expresión de las proteínas ER, HER2, Ki-67 y PR también se pueden usar para determinar la heterogeneidad de la expresión de proteínas.

Los procedimientos de detección de proteínas son rutinarios. En algunos ejemplos se usan inmunoensayos para detectar la presencia de las proteínas ER, HER2, Ki-67 y PR en la muestra. Generalmente, los inmunoensayos incluyen el uso de uno o más agentes de unión específicos (tales como anticuerpos) que únicamente se pueden unir sustancialmente al péptido diana, tal como ER, HER2, Ki-67 y PR. Dichos agentes de unión pueden incluir un marcador detectable (tal como un radiomarcador, fluoróforo o enzima) que permita la detección de la unión a la proteína. Los inmunoensayos ejemplares que se pueden usar incluyen, pero no se limitan a: inmunolectrotransferencia (Western blot), ELISA, microscopia de fluorescencia y citometría de flujo. Un inmunoensayo particular es inmunohistoquímica.

En un ejemplo, el agente de unión específico es un anticuerpo, tal como un anticuerpo policlonal o monoclonal, o fragmento del mismo. En algunos ejemplos, el anticuerpo es un anticuerpo humanizado. En algunos ejemplos, el anticuerpo es un anticuerpo quimérico. Si se desea, el anticuerpo puede incluir un marcador detectable para permitir la detección y, en algunos casos, la cuantificación del complejo proteína/anticuerpo diana. En otros ejemplos, el anticuerpo se detecta con un anticuerpo secundario marcado apropiado.

En algunos ejemplos, los anticuerpos para ER, PR, HER2 y Ki-67 se obtienen de Ventana Medical Systems, Inc. (Tucson, AZ). Sin embargo, un experto en la técnica apreciará que otros anticuerpos que se pueden usar en los procedimientos y kits proporcionados en el presente documento están disponibles comercialmente de otras fuentes, tales como: Novus Biologicals (Littleton, CO), Santa Cruz biotechnology, Inc. (Santa Cruz, CA), Abeam (Cambridge, MA) e Invitrogen (Carlsbad, CA).

La presencia de una señal detectable por encima de los niveles de fondo o control indica la presencia de un péptido diana (por ejemplo, proteína ER, HER2, Ki-67 y PR) en la muestra. El valor obtenido para la muestra de cáncer de mama de prueba se puede comparar con un valor de referencia, tal como un valor de referencia que representa un valor o intervalo de valores esperados. En algunos ejemplos, la referencia es una muestra que posee una cantidad conocida o esperada de proteína ER, HER2, Ki-67 y PR.

En algunos ejemplos, se obtiene una muestra de cáncer de mama y se procesa para IHC. Por ejemplo, la muestra se puede fijar y someter a inclusión, por ejemplo, con formalina y parafina. A continuación, la muestra se puede montar en un soporte, tal como un portaobjetos de microscopio de vidrio. Por ejemplo, se puede realizar microtomos de la muestra en una serie de secciones delgadas y las secciones se pueden montar en uno o más portaobjetos de microscopio. En algunos ejemplos, una única preparación incluye múltiples secciones de tejido. A continuación, diferentes secciones de la muestra de cáncer de mama se pueden marcar individualmente con anticuerpos

específicos para proteína ER, HER2, Ki-67 o PR. Es decir, una sección se puede marcar con anticuerpos específicos para ER y otra sección se puede marcar con anticuerpos específicos para HER2, y así sucesivamente. En algunos ejemplos, una única sección de la muestra de cáncer de mama se puede marcar con anticuerpos específicos para dos o más proteínas de ER, HER2, Ki-67 o PR. Es decir, una sección se puede marcar con anticuerpos específicos para ER y con anticuerpos específicos para HER2, y los anticuerpos se pueden distinguir usando diferentes marcadores (en los que cada marcador es específico para ER, HER2, Ki-67 o PR). Por ejemplo, se puede usar el BenchMark ULTRA de Ventana Medical Systems, Inc. para teñir y procesar las preparaciones.

En algunos ejemplos se obtienen imágenes de las preparaciones que contienen la muestra marcada y se digitalizan, por ejemplo, usando el iScan Coreo (Ventana). De esta manera, para cada una de las proteínas ER, HER2, Ki-67 y PR (que se marcan de manera detectable en la muestra), se puede obtener al menos una imagen digital. Posteriormente, se seleccionan y anotan dos o más campos de visión (FOV) para cada proteína (proteína ER, HER2, Ki-67 y PR), por ejemplo, por un patólogo.

D. Calcular una puntuación IHC combinada

Para calcular la puntuación IHC combinada (por ejemplo, IHC4) se usa una muestra de cáncer de mama con ER, HER2, Ki-67 y PR marcadas de manera detectable (por ejemplo, una o más preparaciones, tales como 1, 2, 3 o 4 preparaciones). En un ejemplo, la muestra de cáncer de mama se puede marcar con anticuerpos primarios específicos para cada uno de ER, HER2, Ki-67 y PR y con anticuerpos secundarios marcados apropiadamente (tales como los marcados con un fluoróforo o enzima, tal como DIG). En un ejemplo se usa el kit de detección ultraView Red ISH DIG según las instrucciones del fabricante (Ventana Medical Systems, Inc., n.º de catálogo 760-505).

Se pueden obtener una o más imágenes digitales de una o más preparaciones que contienen la muestra de cáncer de mama marcada, por ejemplo, usando un software de obtención de imágenes de microscopio. La una o más preparaciones que contienen la una o más muestras de cáncer de mama marcadas (o una imagen digital de las mismas) se evalúa en cuanto a su tinción global, por ejemplo, por un patólogo. Por ejemplo, se pueden detectar regiones de heterogeneidad en la tinción para cada uno de ER, HER2, Ki-67 y PR. Si se observa heterogeneidad regional, se selecciona al menos un FOV que tenga heterogeneidad regional (tal como al menos 2, al menos 3, al menos 4 o al menos 5 áreas de FOV diferentes de heterogeneidad regional en la preparación, tal como 2, 3, 4, 5, 6, 7, 8, 9 o 10 FOV), junto con al menos otros dos FOV que sean similares entre sí (tal como al menos 3, al menos 4 o al menos 5 áreas de FOV diferentes en la preparación que sean similares entre sí, tal como 2, 3, 4, 5, 6, 7, 8, 9 o 10 FOV). Si no se observa heterogeneidad regional, se puede seleccionar un menor número de FOV, tal como al menos 2, al menos 3, al menos 4 o al menos 5 FOV diferentes en la preparación, tal como 2, 3, 4 o 5 FOV en la preparación.

Para cada FOV seleccionado para la proteína PR y Ki-67, se determina la positividad en porcentaje. De esta manera, si se seleccionan cuatro FOV para PR, entonces la positividad en % se determina para cada uno de los cuatro FOV. Para calcular la positividad en % se determina el número total de células tumorales en el FOV, así como el número total de células tumorales teñidas en el FOV (por ejemplo, que sean PR+). La positividad en % es el número total de células tumorales teñidas en el FOV dividido entre el número total de células tumorales en el FOV. La positividad en porcentaje para cada FOV para un marcador particular (por ejemplo, PR) se promedia, dando como resultado una positividad en % para PR y Ki-67.

Para cada FOV para HER2 se determina una puntuación en intervalos de clase (en una escala de 0-3). Por ejemplo, para cada FOV se puede asignar una puntuación en intervalos de clase entre 0 y 3 a la tinción basándose en la evaluación del FOV. Se puede usar una escala de cuatro puntos para describir la inmunotinción de la membrana alrededor de un núcleo para HER2, como sigue: 0, negativo; 1, positividad débil; 2, positividad moderada; y 3+, positividad fuerte, por ejemplo, como se muestra en la tabla 2. La tinción citoplásmica todavía puede estar presente, pero no es necesario incluir dicha tinción en la determinación.

Tabla 2: Puntuación en intervalos de clase para HER2

Patrón de tinción	Puntuación	Evaluación de tinción para HER2
No se observa tinción de la membrana	0	Negativa
Tinción parcial débil de la membrana	1+	Negativa
Tinción completa débil de la membrana, más de un 10 % de células cancerosas	2+	Positiva
Tinción completa intensa de la membrana, más de un 10 % de células cancerosas	3+	Positiva

De manera alternativa, se puede determinar una puntuación binaria (0 o 1) para HER2 (por ejemplo, basándose en FOV individuales o en los FOV colectivamente), en la que un resultado positivo (1) corresponde a una puntuación de tinción de 2 o 3 y un resultado negativo (0) corresponde a una puntuación de tinción de 0 o 1. Se puede usar mayor o menor resolución. Por ejemplo, se puede usar un sistema que usa 0, 1, 2 y 3. Internamente, la puntuación para HER2 se puede representar como una puntuación en intervalos de clase (por ejemplo, 0, 1, 2 o 3) o como un valor binario (por ejemplo, 0 o 1) (por ejemplo, el valor binario se determina basándose en la puntuación en intervalos de clase).

Para los FOV seleccionados para ER, se determina una puntuación H (por ejemplo, para los FOV individuales o los FOV colectivamente). La puntuación H para ER es el porcentaje de células que muestran tinción débil, añadido a dos veces el porcentaje de células que se tiñen moderadamente, añadido a tres veces el porcentaje de células de que se tiñen intensamente, dividido entre 30 para llegar a una variable entre 0-10 (ER_{10}). Una puntuación H de más de 1 es positiva. El porcentaje de células con tinción positiva para PR (por ejemplo, con límite a un 10 %) se dividió entre 10 para obtener una variable entre 0 y 10 (PR_{10}).

De esta manera, después de que un patólogo identifique los FOV para cada uno de los cuatro marcadores, un algoritmo de formación de imágenes puede extraer información digital de la preparación (tal como una muestra de cáncer de mama marcada para detectar la proteína ER, HER2, Ki-67 o PR), transformando la imagen digital en componentes de puntuación (por ejemplo, mediciones de positividad en %, puntuación en intervalos de clase (o binaria) y puntuación H). Estas puntuaciones para cada marcador descritas anteriormente se usan para generar una puntuación IHC combinada, que se calcula usando la fórmula:

$$IHC4 = 94,7 \times \{-0,100 ER_{10} - 0,079 PR_{10} + 0,586 HER2 + 0,240 \ln(1 + 10 \times Ki67)\}$$

E. Calcular una puntuación de heterogeneidad

La heterogeneidad se refiere a la variación espacial de los patrones de tinción histoquímica y molecular, tales como los patrones de tinción para los biomarcadores de cáncer de mama ER, HER2, Ki-67 y PR en una muestra de cáncer de mama. La heterogeneidad puede ser un indicador de la variación espacial de la agresividad del tumor y/o de los patrones de crecimiento que se pueden correlacionar con un fenotipo clínico agregado (por ejemplo, un tumor que es probable que recidive). En el presente documento se muestra que la heterogeneidad biológica de la expresión de proteínas para ER, PR, HER2 y Ki-67 se correlaciona con la recidiva impredecible de una fracción de pacientes con cáncer de mama en estadio temprano.

El procedimiento incluye cuantificar la heterogeneidad regional en la muestra para cada uno de los cuatro marcadores, que se calcula a partir de las puntuaciones del FOV seleccionado. De esta manera, para cada marcador se determina una medida cuantitativa de la heterogeneidad regional. Si el marcador es homogéneo o no detectable (por ejemplo, HER2-), entonces ese valor estará fuera del cálculo de la puntuación de heterogeneidad. Se seleccionan y anotan dos o más campos de visión (FOV) para cada proteína (proteína ER, HER2, Ki-67 o PR), por ejemplo, por un patólogo. El FOV seleccionado puede ser igual o diferente del FOV seleccionado para calcular la puntuación IHC4.

En algunos ejemplos, se selecciona una pluralidad de FOV (por ejemplo, al menos 3, al menos 4 o al menos 5 FOV o similar), por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9 o 10 FOV. Como se describe anteriormente, en respuesta a la observación de la heterogeneidad regional, se pueden seleccionar más FOV, tal como al menos un FOV que tenga heterogeneidad regional (tal como al menos 2, al menos 3, al menos 4 o al menos 5 áreas de FOV diferentes de heterogeneidad regional en la preparación, tal como 2, 3, 4, 5, 6, 7, 8, 9 o 10 FOV), junto con al menos otros dos FOV que sean similares entre sí (tal como al menos 3, al menos 4 o al menos 5 áreas de FOV diferentes en la preparación que sean similares entre sí, tal como 2, 3, 4, 5, 6, 7, 8, 9 o 10 FOV). En respuesta a la observación de no heterogeneidad regional, se puede seleccionar un menor número de FOV, tal como al menos 2, al menos 3, al menos 4 o al menos 5 FOV diferentes en la preparación, tal como 2, 3, 4 o 5 FOV en la preparación.

En algunos casos, puede ser deseable seleccionar FOV adicionales para establecer de manera concluyente (por ejemplo, confirmar o demostrar) que la heterogeneidad regional no está presente. Sin embargo, si es evidente de manera concluyente que no hay heterogeneidad regional, seleccionar FOV adicionales puede consumir recursos de manera innecesaria; por lo tanto, se puede seleccionar un menor número de FOV.

En un ejemplo, se seleccionan FOV con diferente positividad en %, puesto que estas áreas indican heterogeneidad, por ejemplo, heterogeneidad regional.

Para cada una de las proteínas ER, Ki-67, HER2 y PR se determina una puntuación de heterogeneidad. En algunos ejemplos no se determina HER2 (o el valor está fuera, puesto que no hay heterogeneidad para HER2), por ejemplo, si la muestra es negativa para HER2. En algunos ejemplos no se determina Ki-67 (o el valor está fuera, puesto que no hay heterogeneidad para Ki-67).

5 Para los marcadores ER y PR (y en algunos ejemplos también HER2 y Ki-67), la puntuación de heterogeneidad se calcula como sigue. La variabilidad de las puntuaciones de positividad en porcentaje se calcula a partir de regiones heterogéneas en el FOV seleccionado. Por ejemplo, se puede determinar una métrica de variabilidad (VM) para cada marcador, en la que la VM es una indicación de cuán diferente es un FOV seleccionado de otro FOV para la misma proteína. Hay muchas maneras de que esto se pueda calcular, tal como determinar una desviación estándar. En un ejemplo, la VM para cada uno de ER, PR y Ki-67 (y, en algunos ejemplos, HER2) se calcula como sigue:

$$VM = \text{STD}(PPFS_1), PP(FS_2), \dots, PP(FS_N))$$

10 en la que PP(FS) es la positividad en porcentaje para cada campo de visión, FS. La positividad en porcentaje se puede calcular como se describe en el presente documento. A continuación, la puntuación de heterogeneidad para el marcador (por ejemplo, ER y PR) se calcula usando la fórmula (donde $\alpha = [0,1]$):

$$H = \begin{cases} \alpha * \frac{VM}{0,05} \cdot S < 10 \% \\ \alpha * \frac{VM}{S}, \text{ de otro modo} \end{cases}$$

15 Para HER2, se calcula una puntuación en intervalos de clase, P(FS) (escala de 0 a 3, véase anteriormente) para cada FOV, FS. A continuación, se puede calcular la puntuación de heterogeneidad para HER2 usando la fórmula:

$$H = \sum_{\substack{\forall i,j \\ i \neq j}} \|P(FS_i) - P(FS_j)\|$$

20 Si las puntuaciones en intervalos de clase son las mismas para cada FOV, la puntuación de heterogeneidad para HER2 es 0.

25 La siguiente tabla indica diferentes tipos de heterogeneidad, incluyendo heterogeneidad regional como se describe en el presente documento.

Tabla 3: Tipos de heterogeneidad

	Genérica	Proteína ER	Proteína PR	Proteína Ki67	Proteína HER2	Gen HER2
Geográfica 2 bloques separados, > 2 pulg. entre sí	Difusa-focal	No aplicable	n.a.	n.a.	n.a.	n.a.
Regional 1 4x-FOV entre sí, misma sección/bloque, 0,25-2 pulg.	Difusa-focal	Sí	Sí	Sí	Sí	Sí
Interglandular Dentro de 4x-FOV, < 0,25 pulg. Para ambas categorías regionales	> o < 50 % de formaciones similares	No	No	No	Sí	Sí
Intraglandular = célula-célula	> o < 50 % de células similares	Sí	Sí	Sí	Sí	Sí
Todas las formaciones en 1 20x-FOV	Si está en capas múltiples, todas o capas específicas	No	No	No	No	No

	Genérica	Proteína ER	Proteína PR	Proteína Ki67	Proteína HER2	Gen HER2
Para todas las categorías regionales e interglandulares	Difusa-polarizada	No	No	No	No	No
Subcelular	Nuclear, citoplásmica o membranosa Para todas las categorías regionales, inter e intraglandulares	Nuclear	Nuclear	Nuclear	Membranosa	Nuclear
Intensidad	Negativa, 3 categorías positivas	Sí	Sí	Sí	Sí	Sí

F. Calcular una puntuación de pronóstico de salida

5 La puntuación de pronóstico de salida se calcula a partir de la puntuación IHC combinada y la puntuación de heterogeneidad descritas anteriormente. Una puntuación de heterogeneidad combinada se genera como sigue y se puede ajustar para que incluya biomarcadores adicionales:

10 Puntuación de heterogeneidad combinada = raíz cuadrada (puntuación de heterogeneidad para ER + puntuación de heterogeneidad para PR)

15 La puntuación IHC combinada y la puntuación de heterogeneidad combinada se incluyen en un procedimiento de análisis estadístico (por ejemplo, modelo de riesgos proporcionales de Cox), para maximizar las capacidades predictivas combinadas de ambas medidas. Por ejemplo, el modelo se puede usar para determinar la supervivencia libre de evolución (PFS) como resultado, tal como una PFS a 1, 3 o 5 años o de otro modo separar a los pacientes en categorías.

El resultado es una fórmula matemática que genera una puntuación para predecir el riesgo de PFS a 5 años

20 puntuación de pronóstico de salida = $0,03114 \cdot IHC4 + 1,95119 \cdot \text{puntuación de heterogeneidad combinada}$

25 De esta manera, la puntuación de pronóstico de salida es una puntuación de riesgo alto (lo que indica que es probable que el cáncer de mama en estadio temprano recidive, por ejemplo, en un plazo de 5 años, tal como una recidiva local o una metástasis a distancia) o bien una puntuación de riesgo bajo (lo que indica que es probable que el cáncer de mama en estadio temprano no recidive).

G. Arrojar un valor de salida y pronóstico

30 Tras la determinación de la puntuación IHC combinada, la puntuación de heterogeneidad y la puntuación de pronóstico de salida, los resultados del ensayo (tales como la puntuación de pronóstico de salida), los hallazgos, el pronóstico, las predicciones y/o las recomendaciones de tratamiento típicamente se registran y comunican a los técnicos, médicos y/o pacientes, por ejemplo. En determinados modos de realización, se usarán ordenadores para comunicar dicha información a las partes interesadas, tales como, pacientes y/o los médicos que les atienden. Basándose en el pronóstico del cáncer de mama, se puede modificar el tratamiento administrado a un sujeto.

35 En un modo de realización, un pronóstico, predicción y/o recomendación de tratamiento basándose en el valor de salida se comunica a las partes interesadas tan pronto como sea posible después de que se haya completado el ensayo y se haya generado el pronóstico. Los resultados y/o información relacionada se pueden comunicar al sujeto por parte del médico responsable del sujeto. De manera alternativa, los resultados se pueden comunicar directamente a las partes interesadas mediante cualquier medio de comunicación, que incluye por escrito, tal como, por ejemplo, proporcionando un informe escrito, formas electrónicas de comunicación, tal como correo electrónico, o por teléfono. La comunicación se puede facilitar mediante el uso de un ordenador programado adecuadamente, tal como en el caso de comunicaciones por correo electrónico. En determinados modos de realización, la comunicación, que contiene los resultados de una prueba pronóstica y/o conclusiones extraídas de y/o recomendaciones de tratamiento basadas en la prueba, se puede generar y entregar automáticamente a las partes interesadas usando una combinación de hardware y software de ordenador que será familiar para los expertos en la técnica de las telecomunicaciones. Un ejemplo de un sistema de comunicaciones orientado a la asistencia sanitaria se describe en la patente de EE. UU. n.º 6.283.761; sin embargo, la presente divulgación no se limita a los procedimientos que utilizan este sistema de comunicaciones particular.

En determinados modos de realización de los procedimientos de la divulgación, todas o algunas de las etapas del procedimiento, incluyendo el pronóstico de cáncer de mama, y la comunicación de los resultados del ensayo o pronóstico, se pueden llevar a cabo en diversas jurisdicciones (por ejemplo, extranjeras).

5 **H. Tratamientos de seguimiento**

Los procedimientos divulgados pueden incluir adicionalmente seleccionar sujetos para el tratamiento del cáncer de mama, por ejemplo, si se determina el pronóstico de la muestra como un cáncer de mama en estadio temprano que es probable que recidive. De manera alternativa, los procedimientos divulgados pueden incluir adicionalmente seleccionar sujetos para no tratamiento (por ejemplo, solo seguimiento), si la muestra se diagnostica como un cáncer de mama en estadio temprano que no es probable que recidive.

En algunos escenarios ejemplares se puede realizar una o más de las siguientes operaciones dependiendo del pronóstico del paciente: a) prescribir un régimen de tratamiento para el sujeto si el pronóstico determinado del sujeto es que el cáncer de mama en estadio temprano es probable que recidive (tal como el tratamiento con una o más radioterapias y/o agentes quimioterapéuticos, cirugía adicional, seguimiento más frecuente o combinaciones de los mismos); b) no prescribir un régimen de tratamiento para el sujeto si el pronóstico determinado del sujeto es que el cáncer de mama en estadio temprano no es probable que recidive; c) recomendar la administración de un tratamiento (tal como un tratamiento con una o más radioterapias y/o recomendar la administración de agentes quimioterapéuticos, cirugía adicional o combinaciones de los mismos) al sujeto si el pronóstico determinado del sujeto es que el cáncer de mama en estadio temprano es probable que recidive; y d) recomendar no administrar un régimen de tratamiento al sujeto si el pronóstico determinado del sujeto es que el cáncer de mama en estadio temprano no es probable que recidive

25 **Componentes de puntuación ejemplares**

En cualquiera de los ejemplos del presente documento, los componentes de puntuación se pueden referir a la medición de la expresión génica (por ejemplo, positividad en porcentaje o similar), puntuación H, puntuación en intervalos de clase, puntuación binaria o similar para uno o más biomarcadores.

Como los componentes se combinan, los componentes resultantes pueden representar biomarcadores plurales.

35 **Descripción adicional ejemplar**

Para las medidas de la heterogeneidad y cálculo IHC, en lugar de puntuaciones manuales de microscopia, las preparaciones se pueden digitalizar y los algoritmos de análisis de imágenes automatizado pueden calcular la positividad en porcentaje, la puntuación en intervalos de clase (por ejemplo, y convertirla a puntuación binaria) y la puntuación H para los campos (FOV) seleccionados por el patólogo.

En cada preparación digitalizada asociada a los cuatro marcadores mencionados anteriormente, el patólogo anota las regiones específicas para cuantificar la intensidad interregional y la heterogeneidad regional. Usando el análisis de imágenes automatizado para cada región, se calculan los recuentos de células con tinción positiva y negativa, la positividad en porcentaje y la puntuación en intervalos de clase (0, 1+, 2+, 3+).

Para cuantificar la heterogeneidad interregional se pueden usar métricas de heterogeneidad basadas en estas puntuaciones regionales. La puntuación de heterogeneidad se basa en los recuentos de células negativas y positivas (1+, 2+ y 3+) en los FOV seleccionados por el patólogo. La puntuación de heterogeneidad capta la variación de las puntuaciones dentro de los FOV seleccionados. Usando los FOV anotados por el patólogo, la puntuación es una función normalizada de la desviación estándar de las puntuaciones de FOV diferentes. El parámetro de normalización elegido es dependiente del marcador, lo que refleja el hecho de que, en general, la puntuación de positividad en porcentaje interregional tiene una variabilidad más alta para una puntuación de positividad de preparación global más alta.

55 **Sistema computacional ejemplar**

La FIG. 23 ilustra un ejemplo generalizado de un sistema computacional adecuado 2300 en el que se pueden implementar varias de las innovaciones descritas. El sistema computacional 2300 no pretende sugerir ninguna limitación en cuanto al alcance de uso o funcionalidad, puesto que las innovaciones se pueden implementar en diversos sistemas computacionales para uso general o para uso especial. Los sistemas computacionales, como se describen en el presente documento, se pueden usar para implementar una funcionalidad automatizada (por ejemplo, se llevan a cabo procedimientos, acciones y similares mediante un sistema computacional como se describe en el presente documento).

Con referencia a la FIG. 23, el sistema computacional 2300 incluye una o más unidades de procesamiento 2310, 2315 y memoria 2320, 2325. En la FIG. 23 se incluye esta configuración básica 2330 dentro de una línea

discontinua. Las unidades de procesamiento 2310, 2315 ejecutan instrucciones ejecutables por ordenador. Una unidad de procesamiento puede ser una unidad de procesamiento central (CPU) para uso general, un procesador en un circuito integrado de aplicación específica (ASIC) o cualquier otro tipo de procesador. En un sistema de procesamiento múltiple, múltiples unidades de procesamiento ejecutan instrucciones ejecutables por ordenador para incrementar la potencia de procesamiento. Por ejemplo, la FIG. 23 muestra una unidad de procesamiento central 2310, así como una unidad de procesamiento de gráficos o unidad de coprocesamiento 2315. La memoria tangible 2320, 2325 puede ser una memoria volátil (por ejemplo, registros, caché, RAM), memoria no volátil (por ejemplo, ROM, EEPROM, memoria rápida, etc.) o alguna combinación de las dos, accesible para la o las unidades de procesamiento. La memoria 2320, 2325 almacena el software 2380 que implementa una o más de las innovaciones descritas en el presente documento, en forma de instrucciones ejecutables por ordenador adecuadas para su ejecución por la o las unidades de procesamiento.

Un sistema computacional puede tener características adicionales. Por ejemplo, el sistema computacional 2300 incluye un almacenamiento 2340, uno o más dispositivos de entrada 2350, uno o más dispositivos de salida 2360 y una o más conexiones de comunicación 2370. Un mecanismo de interconexión (no mostrado), tal como un bus, controlador o red, interconecta los componentes del sistema computacional 2300. Típicamente, el software del sistema operativo (no mostrado) proporciona un entorno operativo para otro software que se ejecuta en el sistema computacional 2300 y coordina las actividades de los componentes del sistema computacional 2300.

El almacenamiento tangible 2340 puede ser extraíble o no extraíble e incluye discos magnéticos, cintas magnéticas o casetes, CD-ROM, DVD o cualquier otro medio que se pueda usar para almacenar información de una manera no transitoria y al que se pueda acceder dentro del sistema computacional 2300. El almacenamiento 2340 almacena instrucciones para el software 2380 que implementa una o más de las innovaciones descritas en el presente documento.

El o los dispositivos de entrada 2350 pueden ser un dispositivo de entrada táctil, tal como un teclado, ratón, *pen* o bola rastreadora, un dispositivo de entrada de voz, un dispositivo de obtención de imágenes u otro dispositivo que proporcione una entrada al sistema computacional 2300. Para la codificación de vídeo, el/los dispositivo(s) de entrada (2350) puede(n) ser una cámara, tarjeta de vídeo, tarjeta sintonizadora de TV o dispositivo similar que acepte la entrada de vídeo en forma analógica o digital, o un CD-ROM o CD-RW que lea muestras de vídeo en el sistema computacional 2300. El o los dispositivos de salida (2360) pueden ser una pantalla, impresora, altavoz, grabadora de CD u otro dispositivo que proporcione una salida desde el sistema computacional 2300.

La o las conexiones de comunicación (2370) posibilitan la comunicación a través de un medio de comunicación a otra entidad computacional. El medio de comunicación transmite información, tal como instrucciones ejecutables por ordenador, entrada o salida de audio o vídeo u otros datos en una señal de datos modulada. Una señal de datos modulada es una señal que tiene una o más de sus características ajustadas o cambiadas de tal manera que se codifica información en la señal. A modo de ejemplo, y sin limitación, los medios de comunicación pueden usar un vehículo eléctrico, óptico, de RF u otro.

Las innovaciones se pueden describir en el contexto general de los medios legibles por ordenador. Los medios legibles por ordenador son cualquier medio tangible al que se pueda acceder dentro un entorno computacional. A modo de ejemplo, y sin limitación, con el sistema computacional 2300, los medios legibles por ordenador incluyen la memoria 2320, 2325, el almacenamiento 2340 y combinaciones de cualquiera de los anteriores.

Las innovaciones se pueden describir en el contexto general de las instrucciones ejecutables por ordenador, tal como las incluidas en los módulos de programa, que se ejecutan en un sistema computacional en un procesador real o virtual de destino. En general, los módulos de programa incluyen rutinas, programas, bibliotecas, objetos, clases, componentes, estructuras de datos, etc. que realizan tareas particulares o implementan tipos de datos abstractos particulares. La funcionalidad de los módulos de programa se puede combinar o dividir entre módulos de programa según se desee en diversos modos de realización. Las instrucciones ejecutables por ordenador para los módulos de programa se pueden ejecutar dentro de un sistema computacional local o distribuido.

Los términos “sistema” y “dispositivo” se usan indistintamente en el presente documento. A menos que el contexto indique claramente de otro modo, ningún término implica ninguna limitación sobre un tipo de sistema computacional o dispositivo computacional. En general, un sistema computacional o dispositivo computacional puede ser local o distribuido, y puede incluir cualquier combinación de hardware para uso especial y/o hardware para uso general con software que implementa la funcionalidad descrita en el presente documento.

Por razones de presentación, la descripción detallada usa términos como “determinar” y “usar” para describir las operaciones de ordenador en un sistema computacional. Estos términos son abstracciones de alto nivel para las operaciones realizadas por un ordenador y no se deben confundir con actos realizados por un ser humano. Las operaciones de ordenador reales correspondientes a estos términos varían dependiendo de la implementación.

Medios legibles por ordenador

Cualquiera de los medios legibles por ordenador del presente documento puede ser no transitorio (por ejemplo, memoria, almacenamiento magnético, almacenamiento óptico o similar).

5 Cualquiera de las acciones de almacenamiento descritas en el presente documento se puede implementar almacenando en uno o más medios legibles por ordenador (por ejemplo, medios de almacenamiento legibles por ordenador u otros medios tangibles).

10 Cualquiera de las cosas descritas como almacenadas se puede almacenar en uno o más medios legibles por ordenador (por ejemplo, medios de almacenamiento legibles por ordenador u otros medios tangibles).

15 Cualquiera de los procedimientos descritos en el presente documento se puede implementar mediante instrucciones ejecutables por ordenador en (por ejemplo, codificadas en) uno o más medios legibles por ordenador (por ejemplo, medios de almacenamiento legibles por ordenador u otros medios tangibles). Dichas instrucciones pueden hacer que un ordenador realice el procedimiento. Las tecnologías descritas en el presente documento se pueden implementar en una variedad de lenguajes de programación.

Procedimientos en dispositivos de almacenamiento legibles por ordenador

20 Cualquiera de los procedimientos descritos en el presente documento se puede implementar mediante instrucciones ejecutables por ordenador almacenadas en uno o más dispositivos de almacenamiento legibles por ordenador (por ejemplo, memoria, almacenamiento magnético, almacenamiento óptico o similar) Dichas instrucciones pueden hacer que un ordenador realice el procedimiento.

Implementación del riesgo de recidiva de cáncer de mama en estadio temprano ejemplar y resultados

25 Este ejemplo describe procedimientos usados para predecir el riesgo de recidiva de cáncer de mama en estadio temprano.

30 Muestras de cáncer de mama de aproximadamente 20 pacientes se marcaron con anticuerpos específicos para ER, PR, HER2 y Ki-67.

35 Se formaron imágenes de estas muestras marcadas resultantes en portaobjetos de microscopio usando patología digital. Se seleccionaron tres FOV para cada uno de ER, PR, HER2 y Ki-67. Se usó un software para cuantificar la tinción celular en porcentaje (por ejemplo, positividad en porcentaje) y la intensidad para cada FOV. De esta manera, se determinó la positividad en porcentaje. Además, se determinaron tres medidas diferentes de heterogeneidad basándose en la variabilidad de las intensidades y la tinción celular en porcentaje para los tres FOV.

40 La tinción celular en porcentaje e intensidad resultantes para cada FOV para cada uno de ER, PR, HER2 y Ki-67 se promediaron (por ejemplo, se promediaron los FOV para ER). Esto da como resultado un valor de tinción celular en porcentaje y un valor de intensidad para cada uno de ER, PR, HER2 y Ki-67. Estos valores se usaron para calcular la puntuación IHC4 usando esta fórmula:

$$\text{IHC4} = 94,7 \times \{-0,100 \text{ER}10 - 0,079 \text{PR}10 + 0,586 \text{HER}2 + 0,240 \ln(1 + 10 \times \text{Ki}67)\}$$

45 Usando la puntuación IHC4 y los 12 valores de heterogeneidad (uno para cada uno de los tres FOV para cada uno de ER, PR, HER2 y Ki-67), se modelizaron cada una de las medidas de heterogeneidad y la puntuación IHC4 usando el modelo de riesgos proporcionales de Cox, que modeliza el tiempo hasta la metástasis a distancia. Las tres medidas de heterogeneidad fueron muy similares, por lo que se usó la primera medida de heterogeneidad para cada uno de los cuatro ensayos.

50 Usando las cinco variables, la puntuación IHC4 y una puntuación de heterogeneidad para cada ensayo (ER, PR, Ki-67 y HER2), se modelizó la puntuación IHC4 con cada una de las puntuaciones de heterogeneidad por separado (usando el modelo Cox).

55 Basándose en los resultados, se determinó que la medida de la puntuación de heterogeneidad para PR fue la más útil para determinar el tiempo hasta la metástasis a distancia, con ER añadiendo un cierto poder predictivo.

60 Se determinó si la relación entre la puntuación de heterogeneidad y el riesgo era lineal o no; puesto que hay una suposición de que cada cambio de unidad en la puntuación de heterogeneidad da como resultado el mismo incremento en el riesgo de metástasis a distancia. Los datos se transformaron (tomando la raíz cuadrada, log natural, etc.) y la variable transformada se introdujo en el modelo de Cox. Se seleccionó la raíz cuadrada, puesto que fue la que más incrementó la capacidad predictiva.

65 La modificación final de la medición de heterogeneidad implicó tomar la puntuación de heterogeneidad (HET) para ER y PR, sumarlas y tomar la raíz cuadrada. Puntuación de heterogeneidad = raíz cuadrada (HET^{PR} + HET^{ER}). La puntuación de HET resultante y la puntuación IHC4 se introdujeron en un modelo de riesgos proporcionales de Cox

que toma las dos variables lineales y encuentra la mejor combinación lineal de las dos para predecir el tiempo hasta la metástasis a distancia. La fórmula resultante es:

$$\text{Puntuación de riesgo de salida} = 1,95119 * (\text{raíz cuadrada} (\text{HET}^{\text{PR}} + \text{HET}^{\text{ER}})) + 0,03114 * \text{IHC4}$$

5 Como se muestra en la FIG. 24, este procedimiento clasificó con precisión una muestra de cáncer de mama como una que es más (evolución) o menos (sin evolución) agresiva. Los casos mostrados en la FIG. 24 son los mismos casos usados para crear el algoritmo.

REIVINDICACIONES

- 5 1. Un procedimiento para determinar el pronóstico de cáncer de mama en un sujeto, que comprende: seleccionar en una muestra de cáncer de mama obtenida del sujeto al menos dos campos de visión (FOV) diferentes para cada uno de receptor de estrógenos (ER), receptor del factor de crecimiento epidérmico humano 2 (HER2), Ki-67 y receptor de progesterona (PR), en el que la muestra se marca de manera detectable con anticuerpos para cada uno de ER, HER2, Ki-67 y PR;
- 10 medir la expresión de las proteínas ER, HER2, Ki-67 y PR en cada uno de los FOV seleccionados;
- determinar una puntuación inmunohistoquímica (IHC) combinada IHC4, siendo la puntuación inmunohistoquímica combinada una puntuación pronóstica basada en dichos cuatro marcadores medidos;
- 15 medir la heterogeneidad de proteínas para ER y PR en cada uno de los FOV seleccionados;
- determinar una puntuación de heterogeneidad de proteínas para cada uno de ER y PR, siendo la puntuación de heterogeneidad de proteínas una indicación de la cantidad de heterogeneidad de la expresión de proteínas de un biomarcador particular en FOV diferentes;
- 20 combinar las puntuaciones de heterogeneidad de proteínas y la puntuación IHC combinada, generando así una puntuación de pronóstico de salida; y
- determinar que el cáncer de mama en el sujeto es probable que sea agresivo si la puntuación de pronóstico de salida alcanza un valor umbral o determinar que el cáncer de mama en el sujeto es improbable que sea agresivo si la puntuación de pronóstico de salida no alcanza el valor umbral.
- 25 2. El procedimiento de la reivindicación 1, en el que la muestra de cáncer de mama es una muestra de cáncer de mama en estadio temprano.
- 30 3. El procedimiento de la reivindicación 2, en el que la muestra de cáncer de mama en estadio temprano es negativa para HER2, positiva para ER y con ganglios linfáticos negativos.
- 35 4. El procedimiento de cualquiera de las reivindicaciones 1 a 3, en el que el procedimiento determina un riesgo de recidiva de cáncer de mama a 5 años.
- 40 5. El procedimiento de cualquiera de las reivindicaciones 1 a 4, que comprende adicionalmente obtener una imagen digitalizada de la muestra de cáncer de mama que se marca de manera detectable para cada uno de ER, HER2, Ki-67 y PR.
- 45 6. El procedimiento de cualquiera de las reivindicaciones 1 a 5, en el que medir la expresión de proteínas ER, HER2, Ki-67 y PR comprende determinar:
- la positividad en porcentaje para cada uno de ER, Ki-67 y PR;
- la puntuación en intervalos de clase para HER2; y
- la puntuación H para ER.
- 50 7. El procedimiento de cualquiera de las reivindicaciones 1 a 6, en el que determinar una puntuación IHC combinada comprende usar la fórmula $IHC4 = 94,7 \times \{-0,100 ER_{10} - 0,079 PR_{10} + 0,586 HER2 + 0,240 \ln(1 + 10 \times Ki67)\}$, en la que ER_{10} es la puntuación H para ER, en la que PR_{10} es una puntuación de positividad en porcentaje ajustada para PR, en la que HER2 es la puntuación en intervalos de clase para HER2 y Ki67 es la positividad en porcentaje para Ki-67.
- 55 8. El procedimiento de cualquiera de las reivindicaciones 1 a 7, en el que medir la heterogeneidad de proteínas para ER, HER2, Ki-67 y PR comprende determinar una métrica de variabilidad (VM) para cada uno de ER, HER2, Ki-67 y PR, en la que $VM = STD(PP(FS_1), PP(FS_2), \dots, PP(FS_N))$, y PP(FS) es una positividad en porcentaje para cada campo de visión, FS.
- 60 9. El procedimiento de cualquiera de las reivindicaciones 1 a 8, en el que determinar una puntuación de heterogeneidad de proteínas para cada uno de ER y PR comprende usar la fórmula

$$\begin{cases} \alpha + \frac{VM}{0,05}, S < 10 \% \\ \alpha + \frac{VM}{S}, \text{ de otro modo} \end{cases}$$

en la que

- 5 10. El procedimiento de cualquiera de las reivindicaciones 1 a 9, en el que determinar una puntuación de heterogeneidad de proteínas para HER2 comprende usar la fórmula

$$H = \sum_{V_i, J} \|P(FS_i) - P(FS_j)\|$$

- 10 en la que P es una puntuación en intervalos de clase para cada campo de visión,

11. El procedimiento de cualquiera de las reivindicaciones 1 a 10, en el que generar una puntuación de pronóstico de salida comprende usar la fórmula:

- 15 puntuación de pronóstico de salida = (coeficiente inmunohistoquímico combinado)*IHC4 + (coeficiente de heterogeneidad)*puntuación de heterogeneidad combinada, en la que la puntuación de heterogeneidad combinada es igual a la raíz cuadrada (puntuación de heterogeneidad para ER + puntuación de heterogeneidad para PR).

- 20 12. El procedimiento de cualquiera de las reivindicaciones 1 a 11, en el que:
el coeficiente inmunohistoquímico combinado es igual a 0,03114; y el coeficiente de heterogeneidad es igual a 1,95119.

- 25 13. El procedimiento de cualquiera de las reivindicaciones 1 a 12, en el que una o más de las etapas se realizan mediante un ordenador programado adecuadamente.

14. El procedimiento de cualquiera de las reivindicaciones 1 a 13, en el que la selección de al menos dos FOV diferentes en una muestra de cáncer de mama es realizada por un patólogo.

FIG. 1

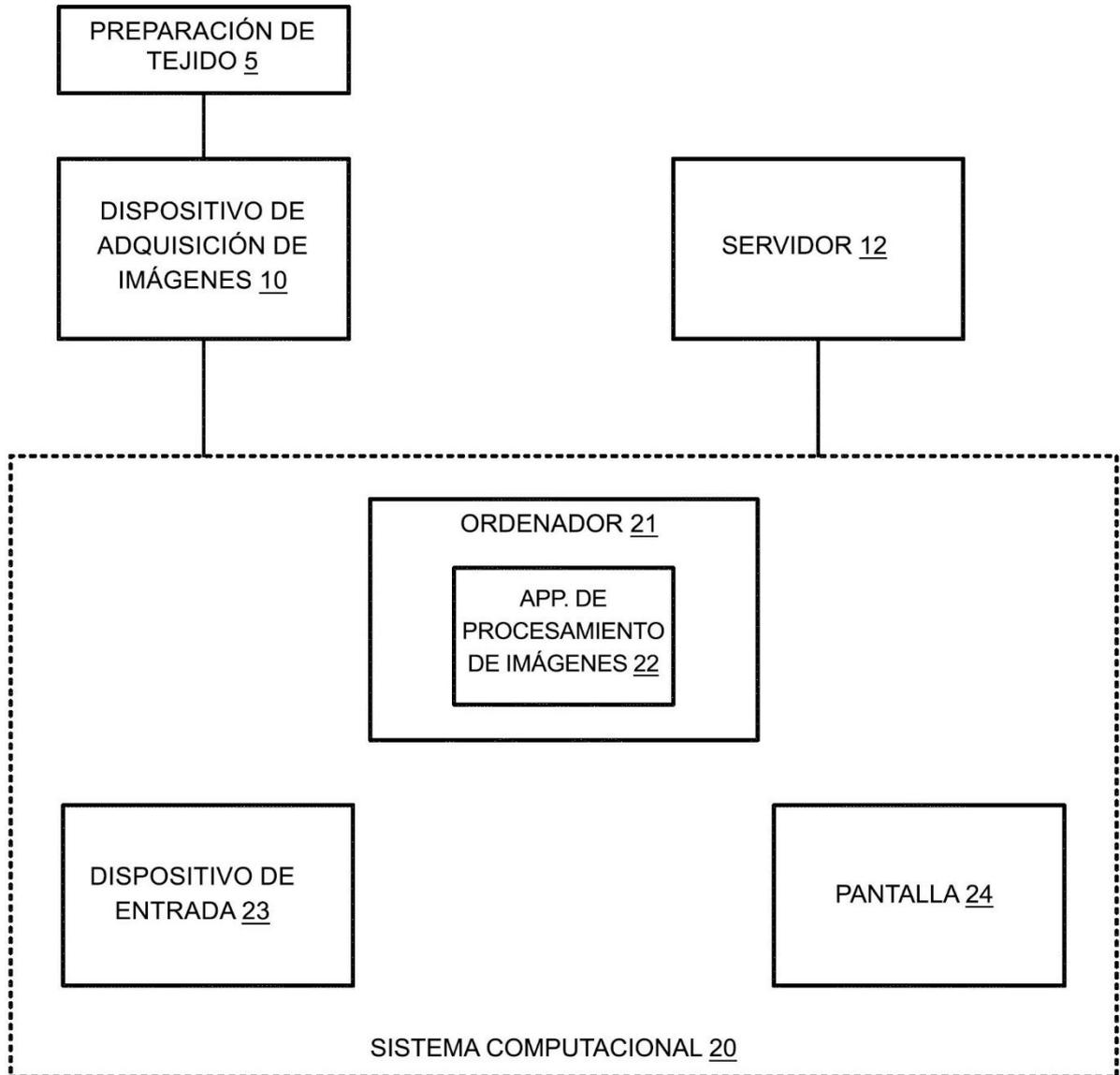


FIG. 2

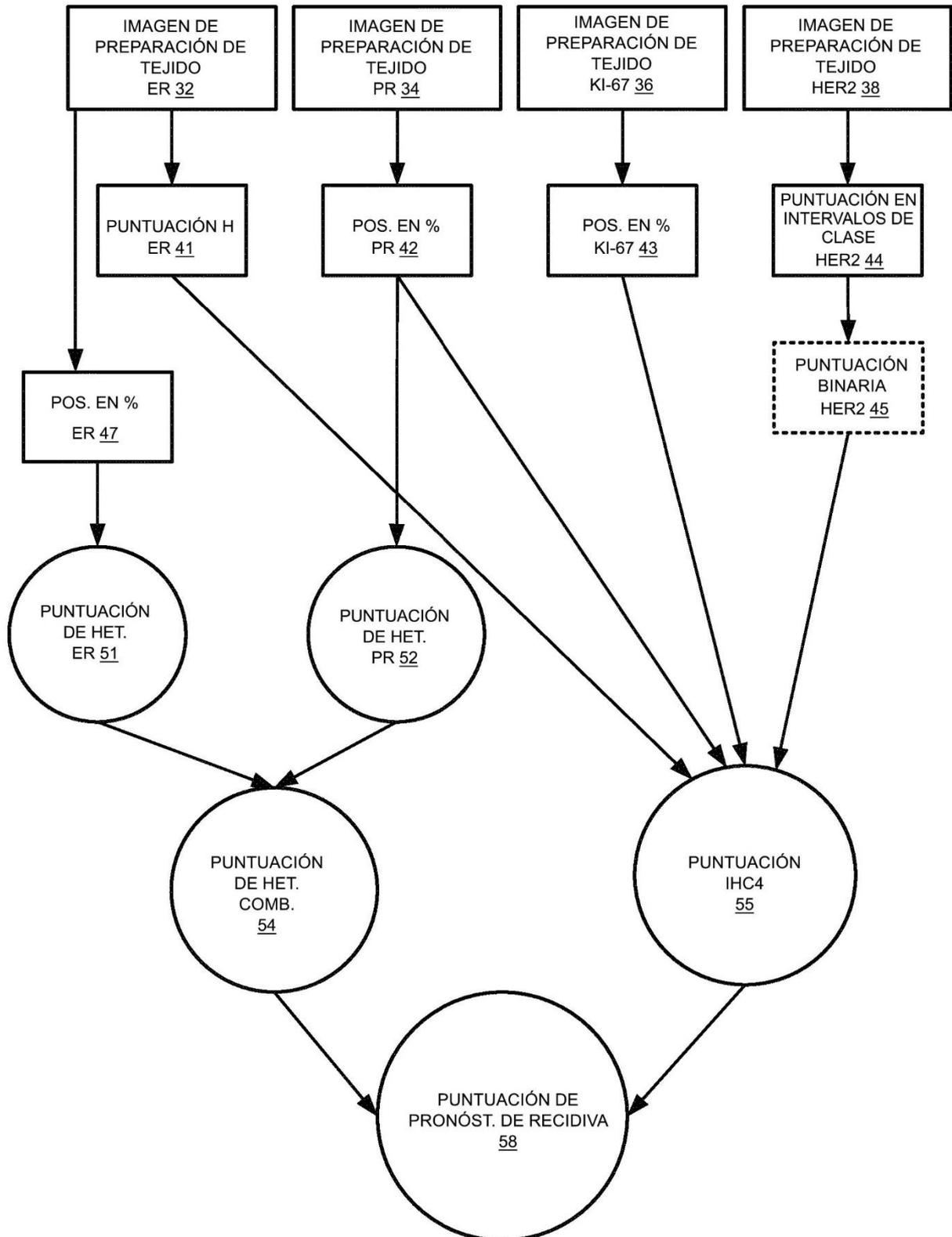


FIG. 3

VISTA GENERAL DEL PROCEDIMIENTO

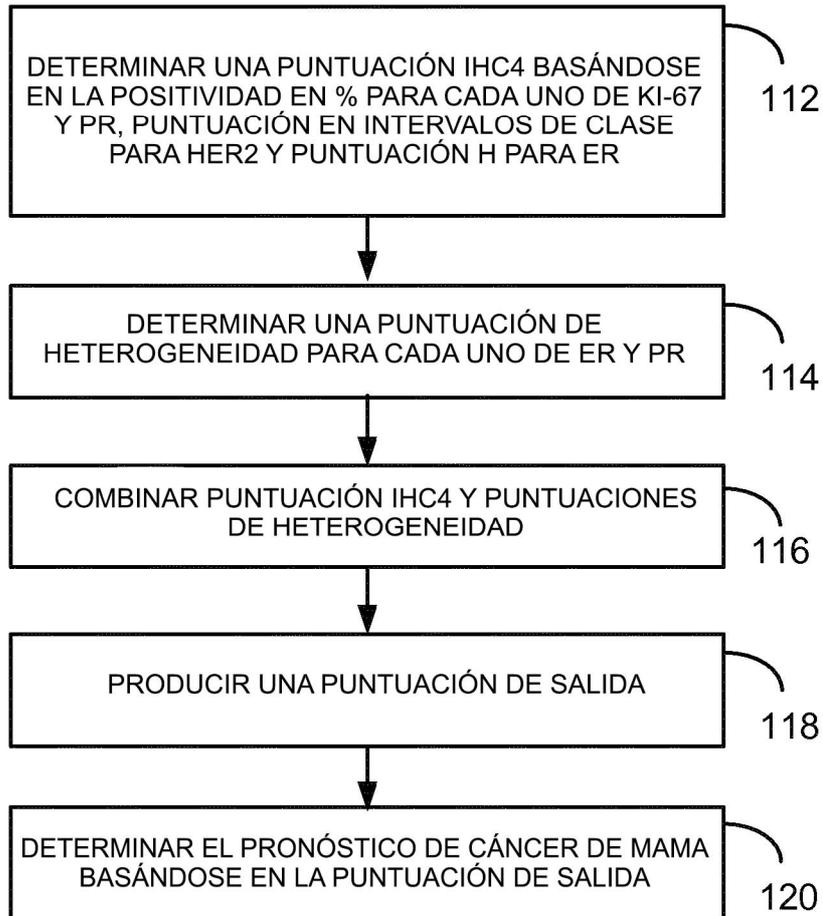


FIG. 4A

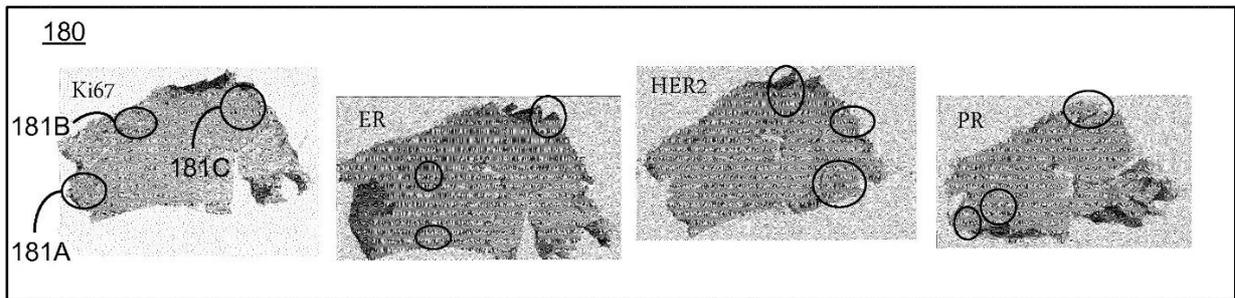


FIG. 4B

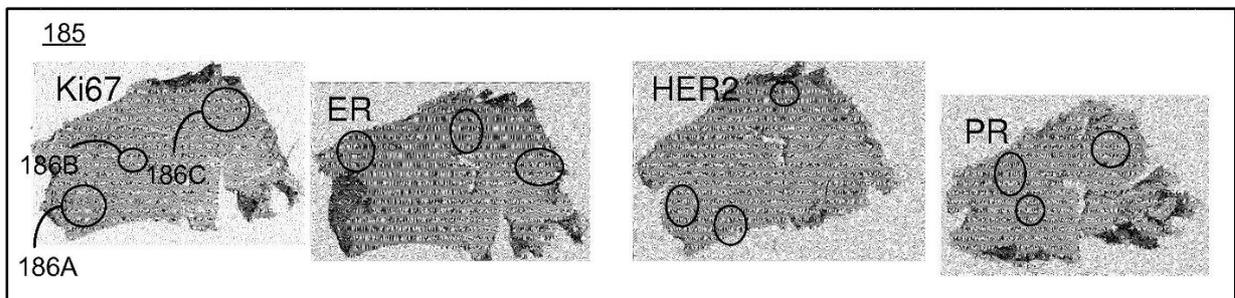


FIG. 5

UN MODO DE REALIZACIÓN DEL PROCEDIMIENTO

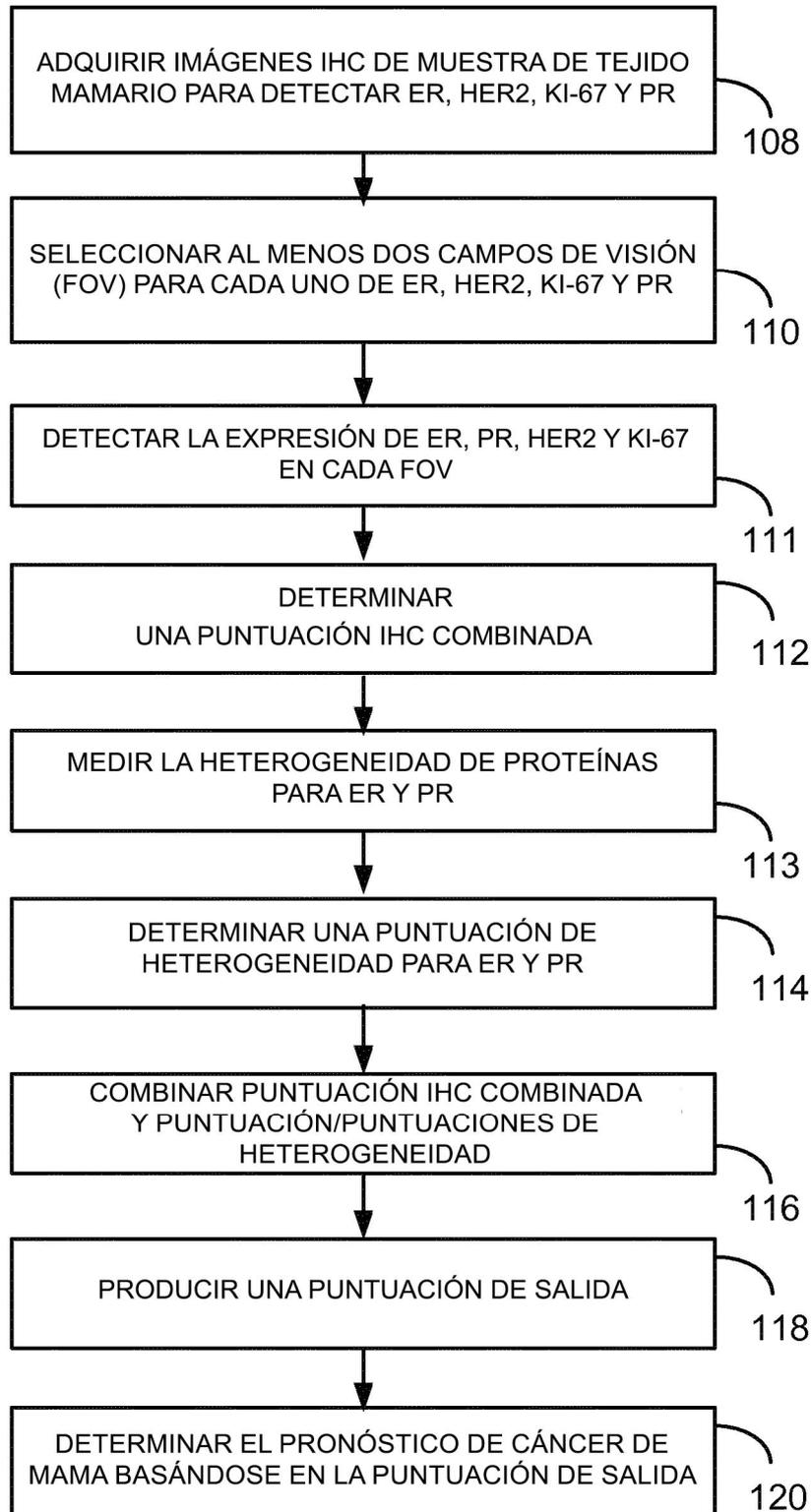


FIG. 6

ADQUIRIR IMÁGENES IHC DE MUESTRA DE
TEJIDO MAMARIO

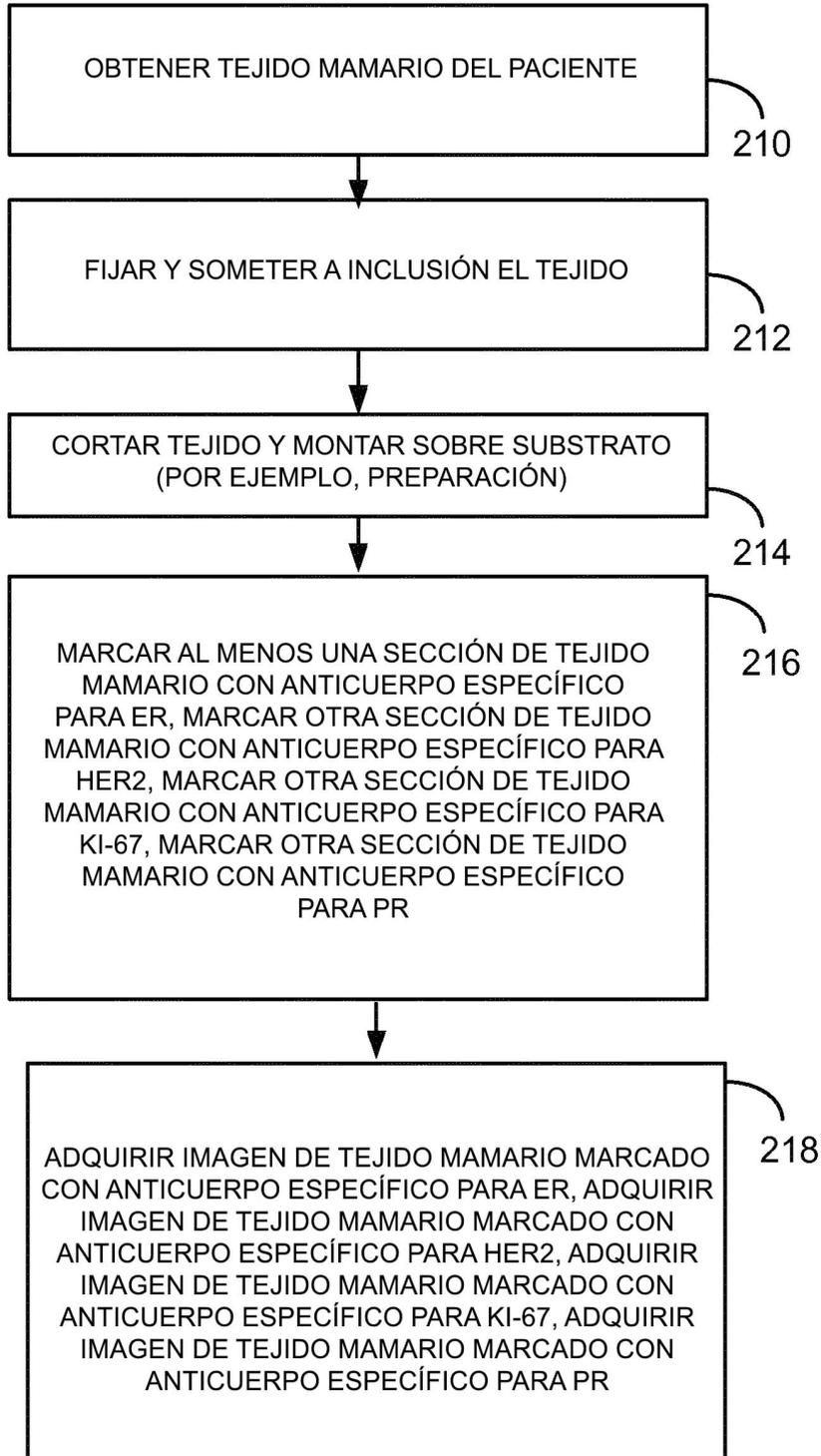


FIG. 7

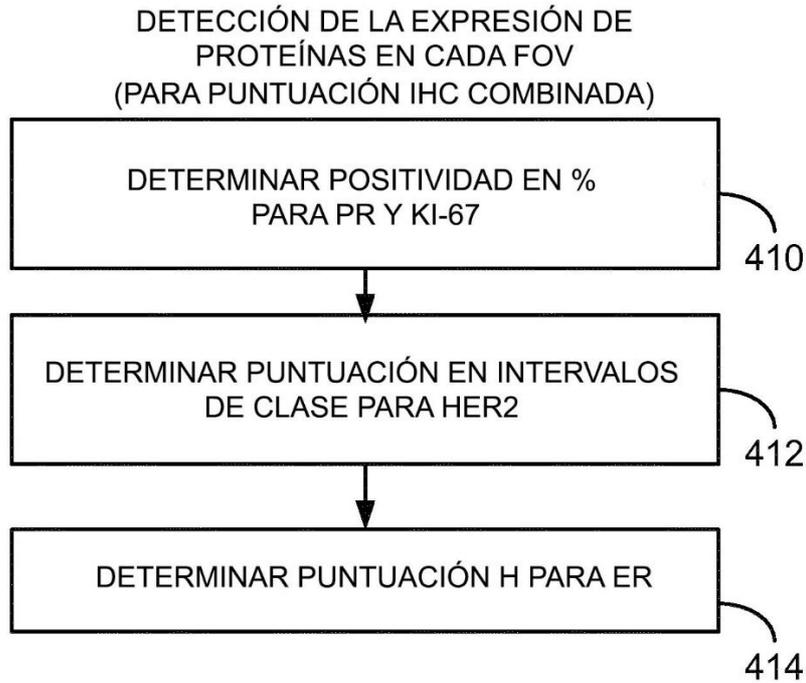


FIG. 8

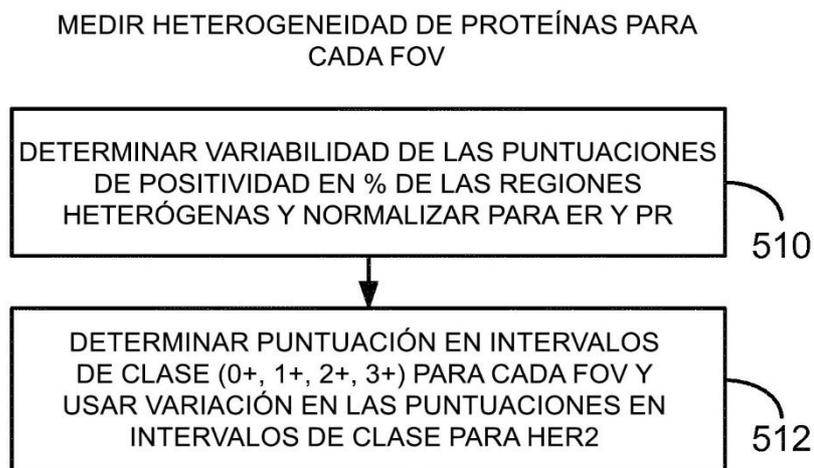


FIG. 9

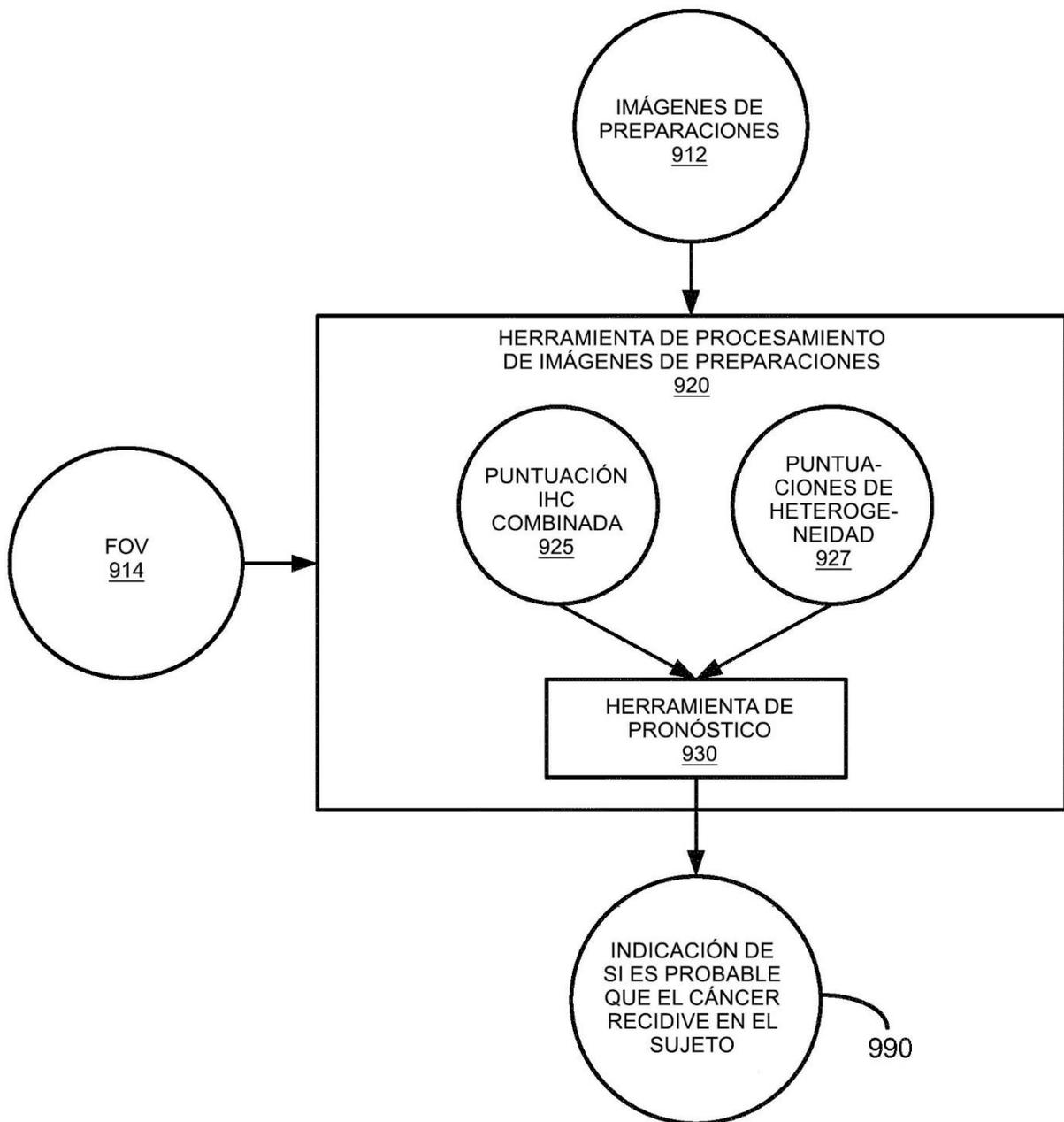
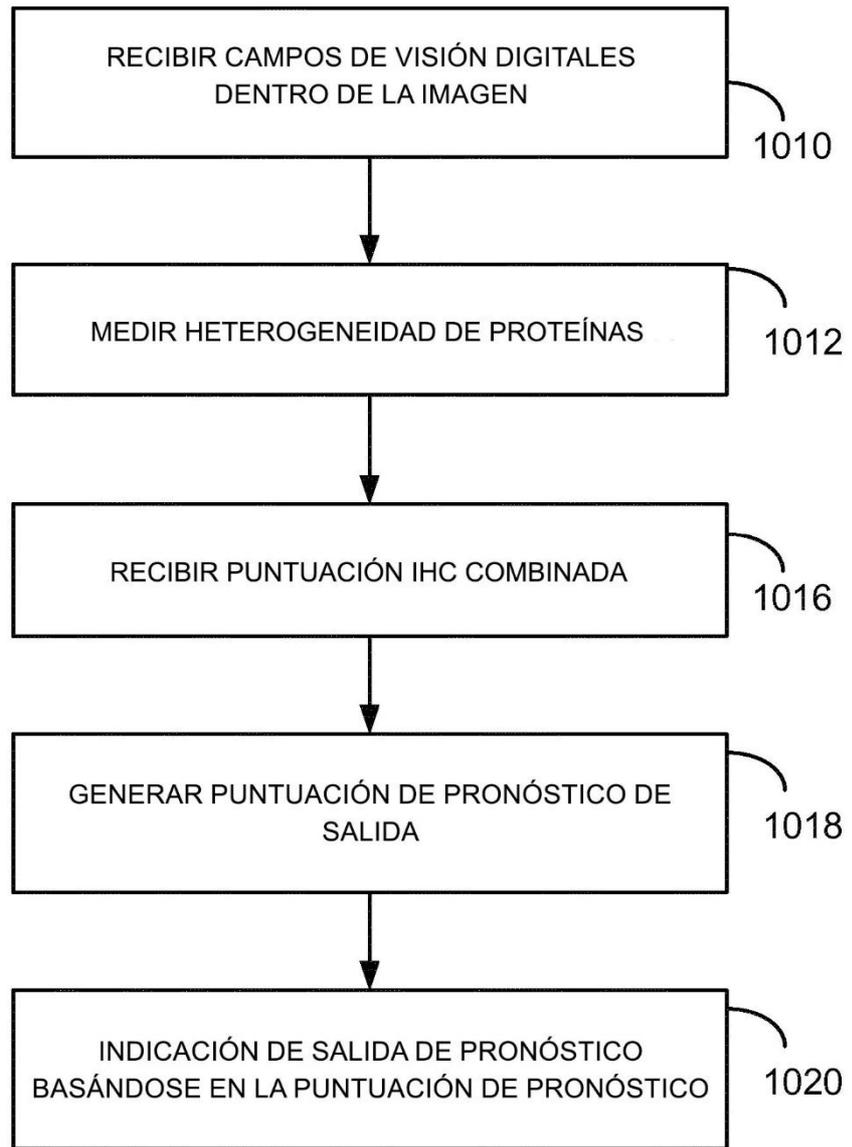
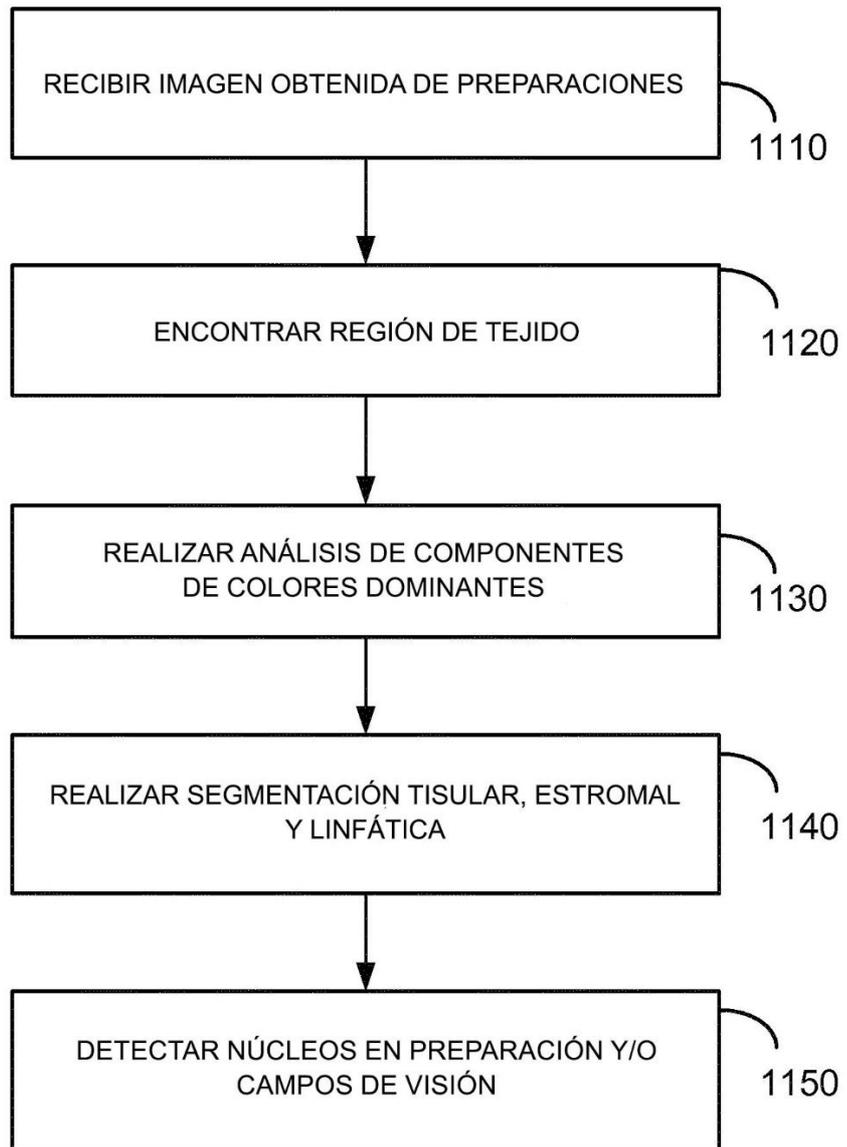


FIG. 10



1000 ↗

FIG. 11



1100 ↗

FIG. 12

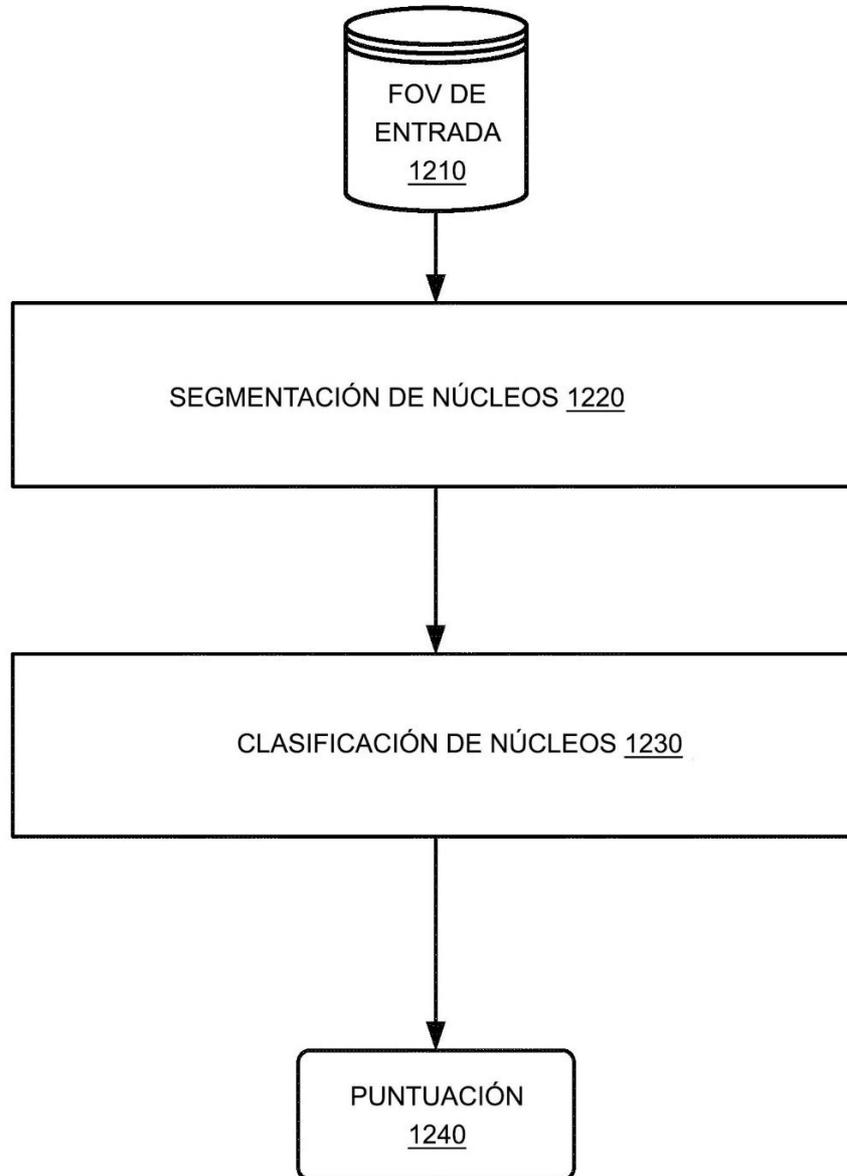
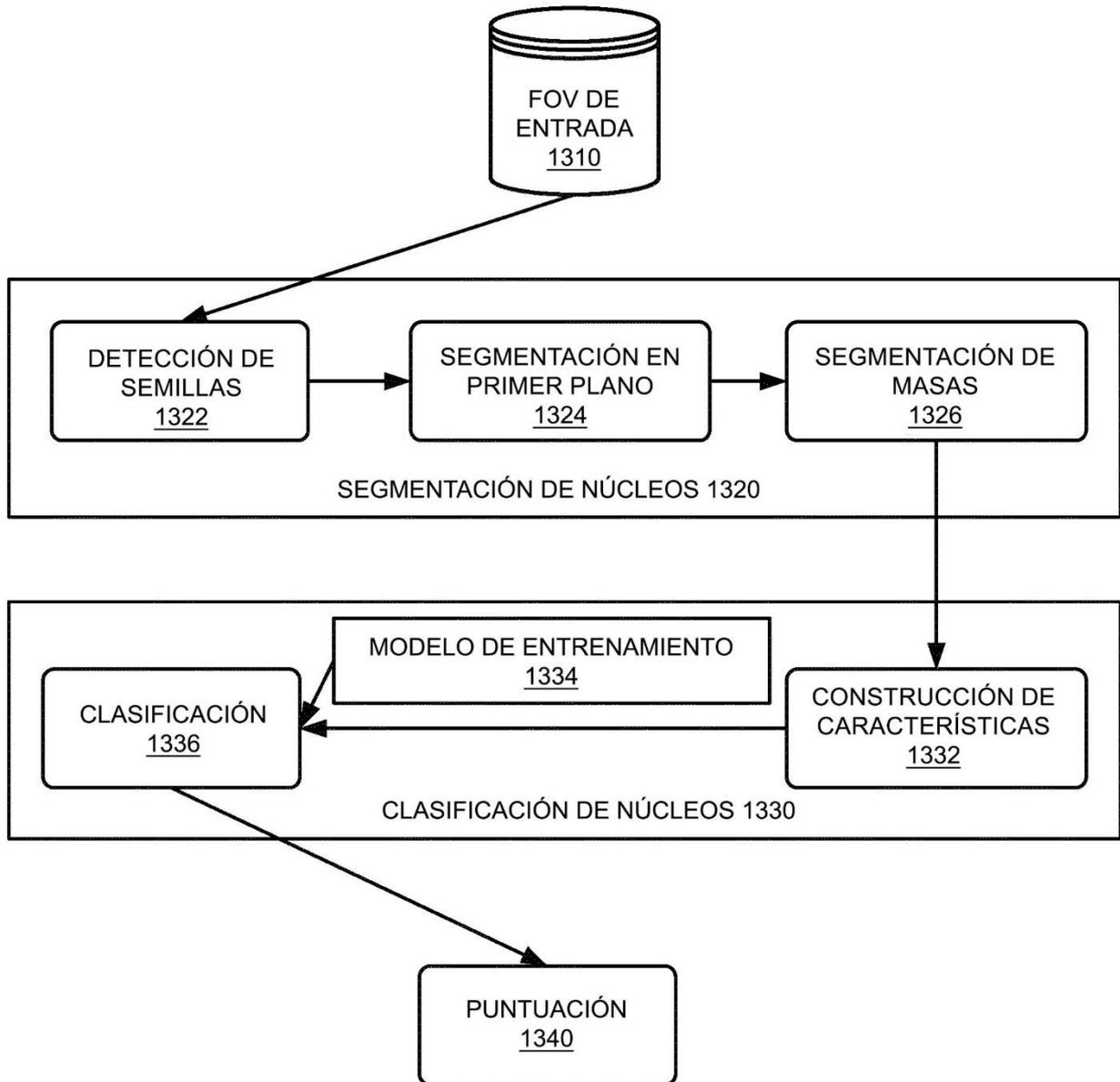


FIG. 13



1300 ↗

FIG. 14

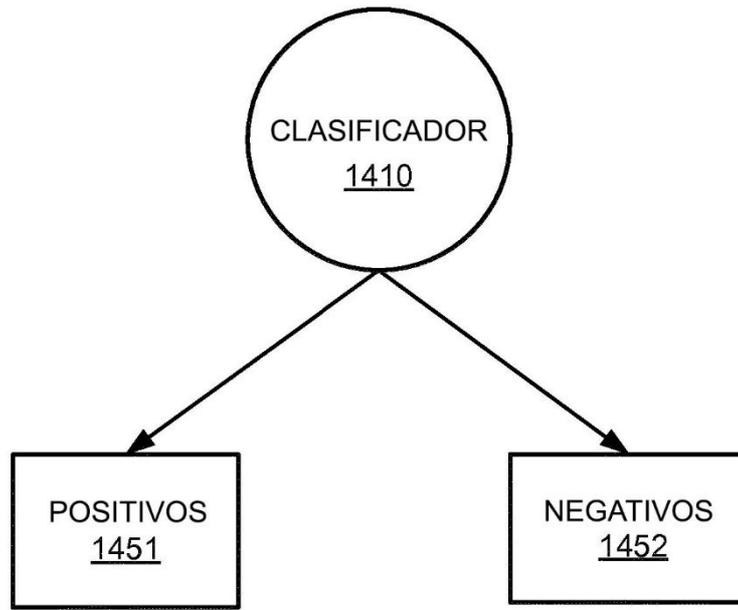
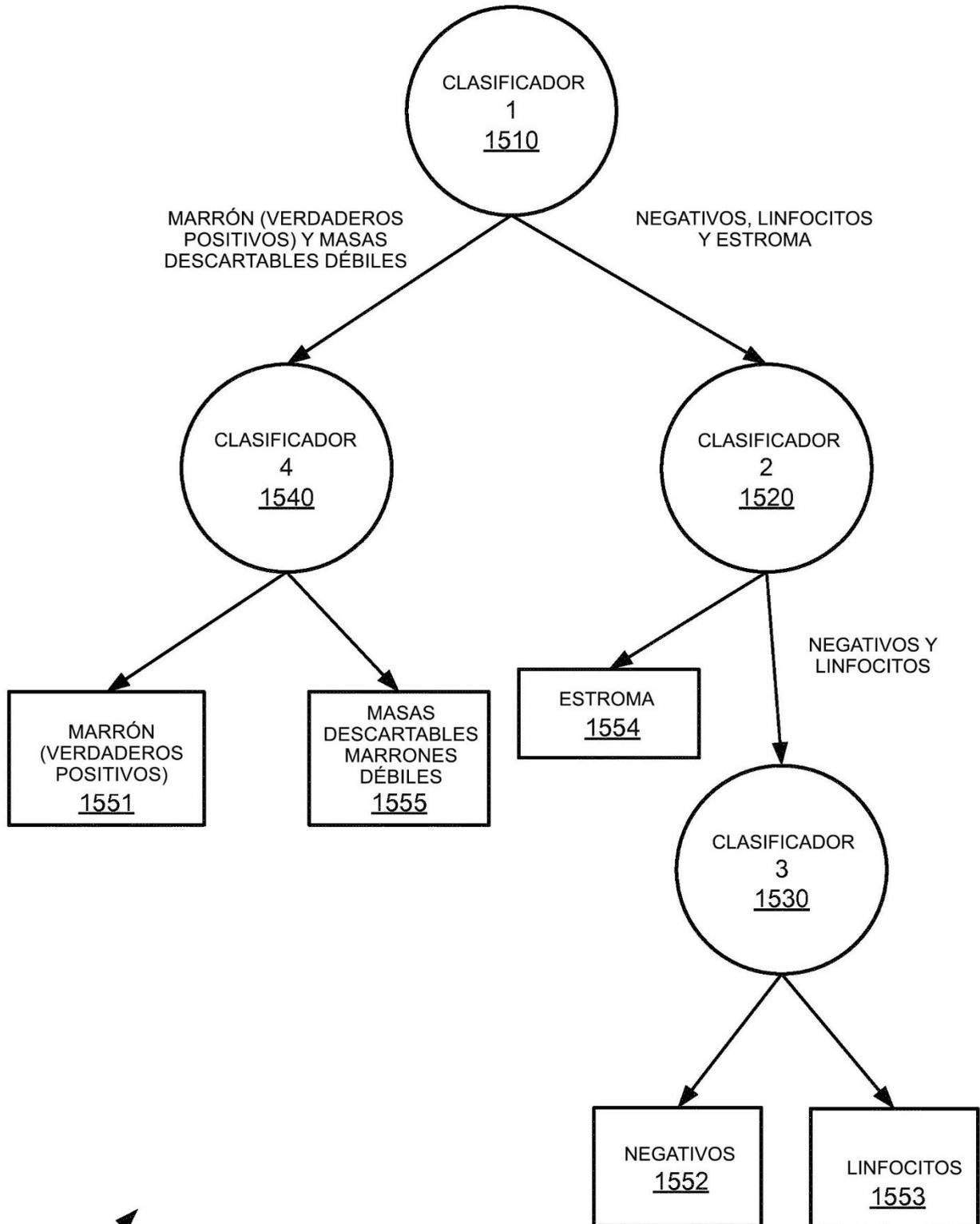
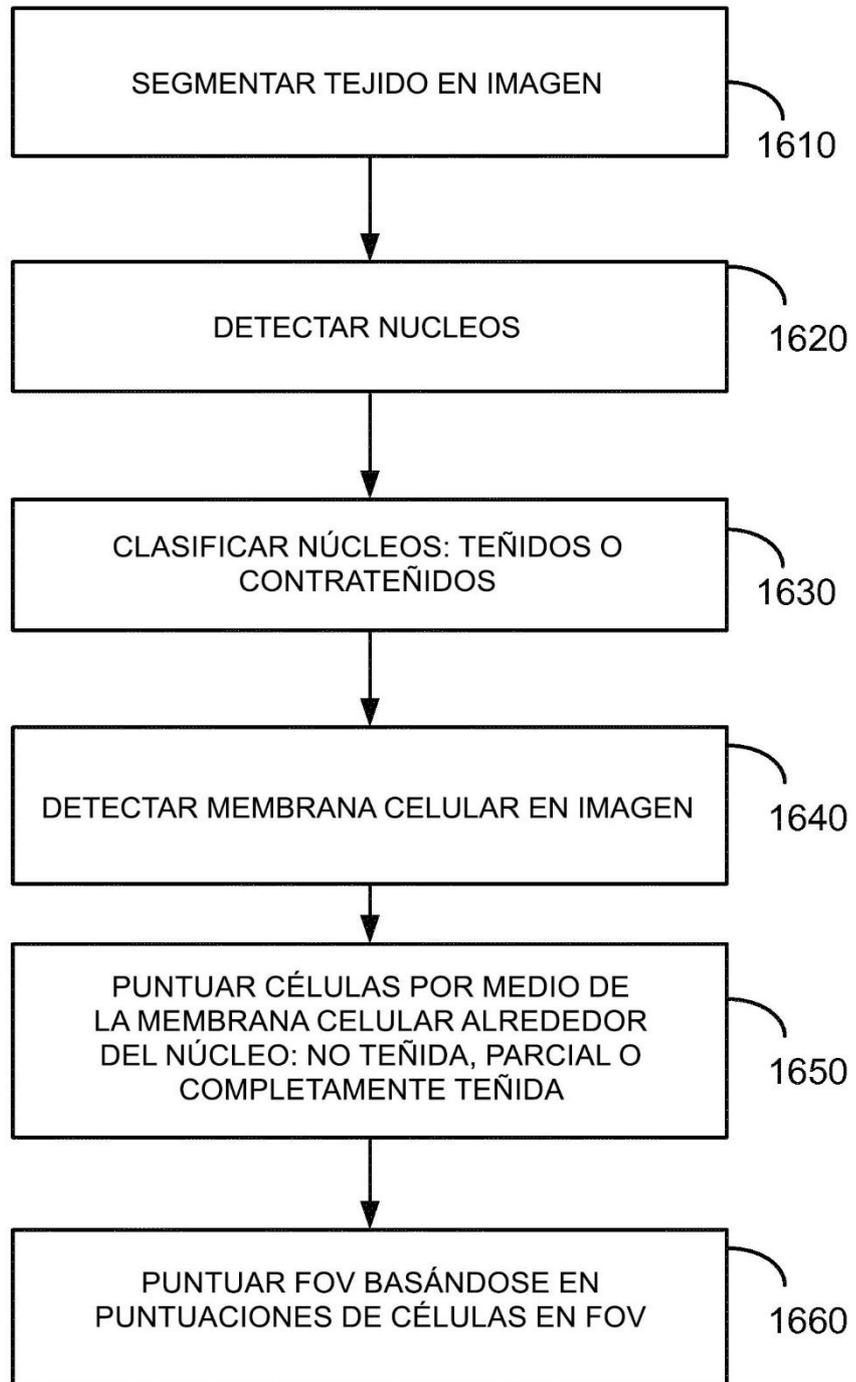


FIG. 15



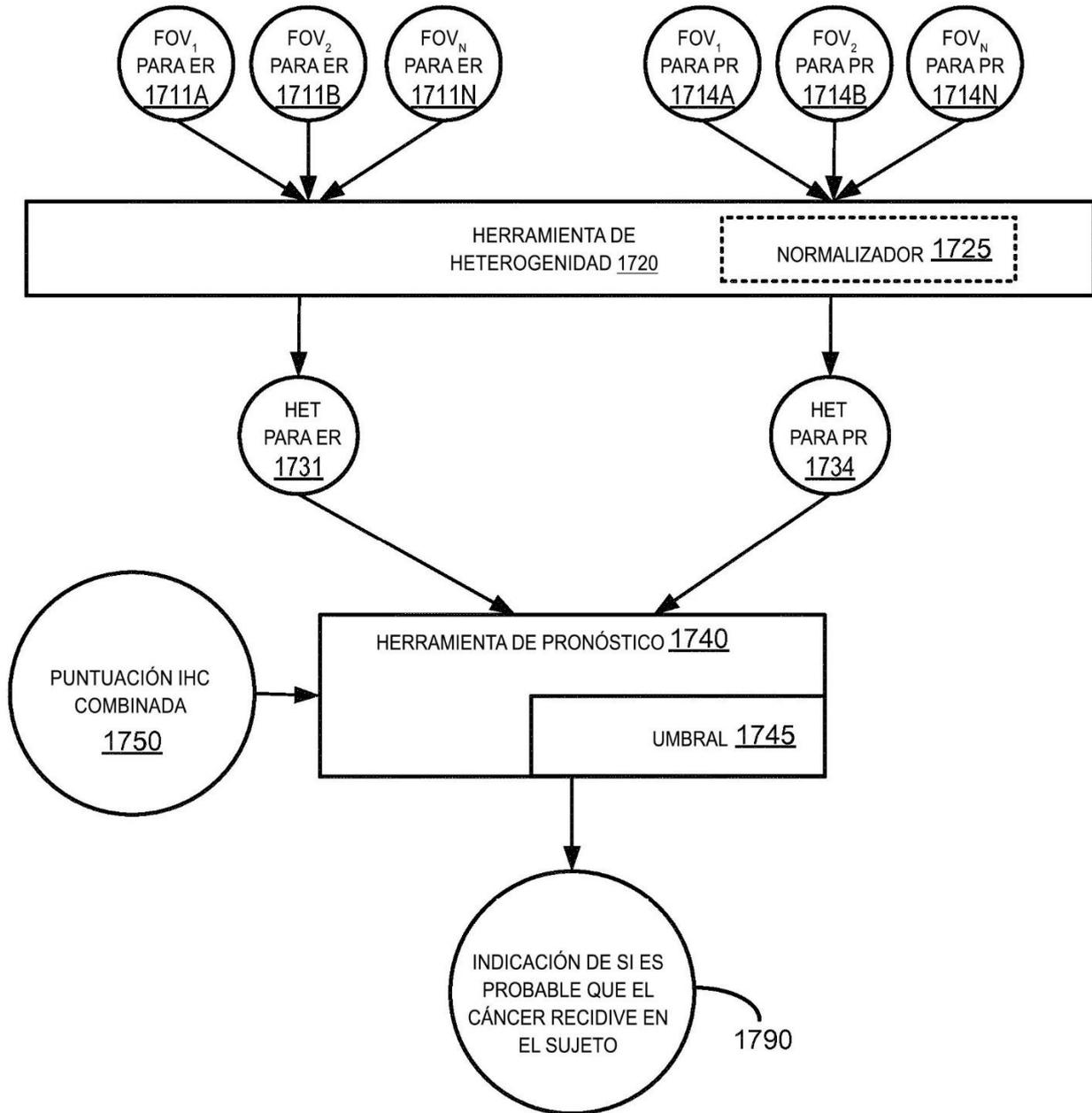
1500 ↗

FIG. 16



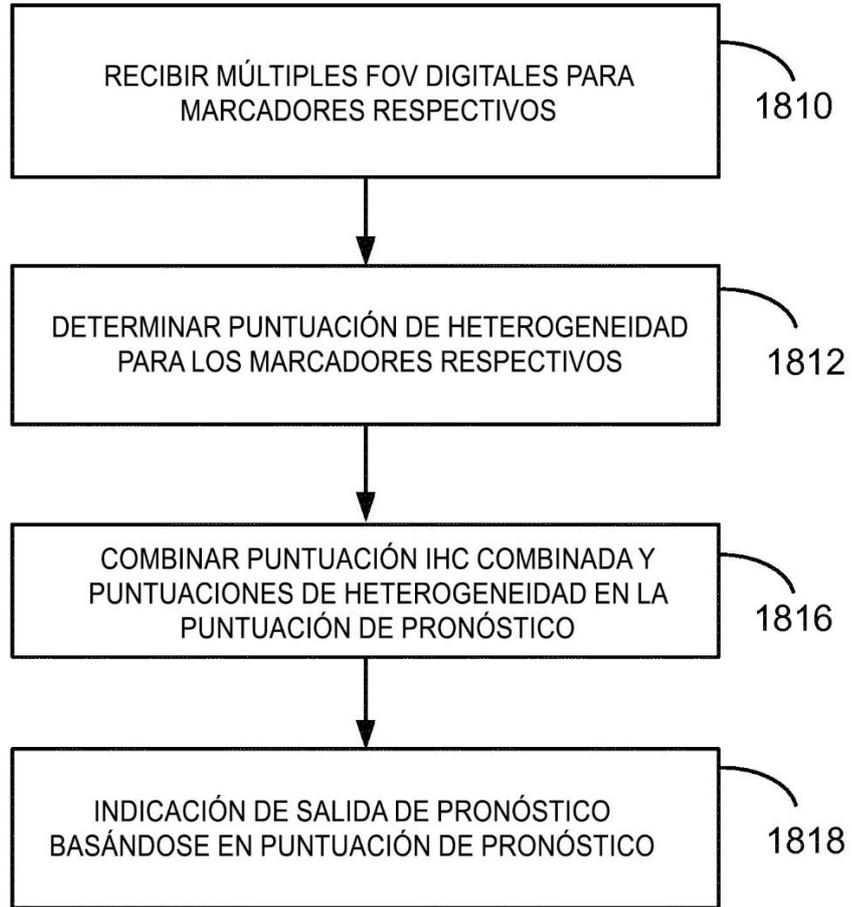
1600 ↗

FIG. 17



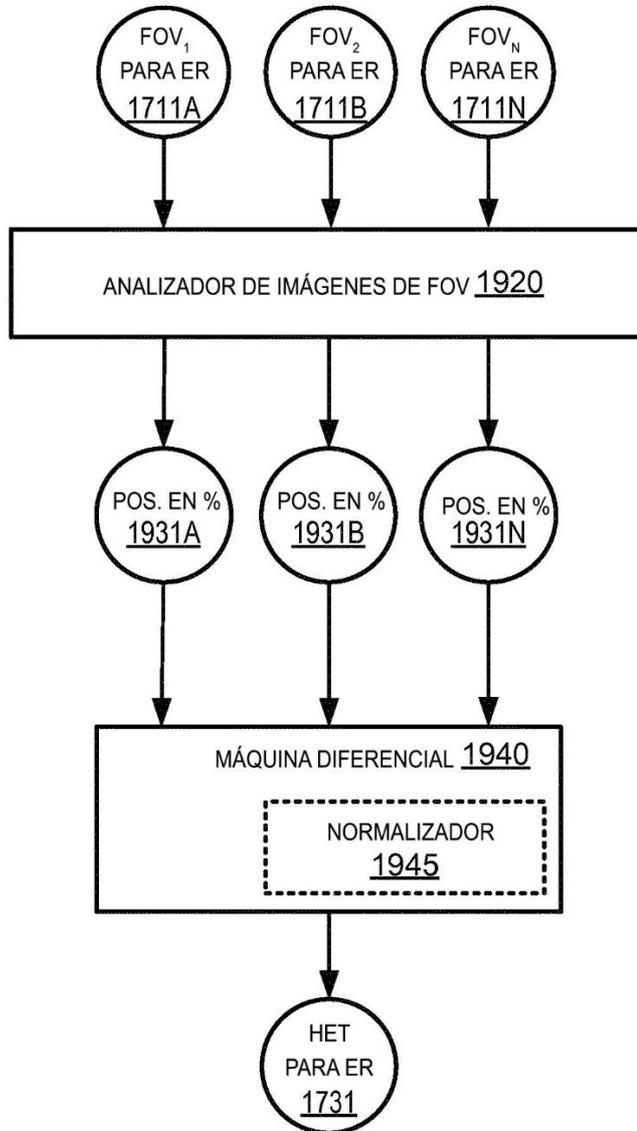
1700

FIG. 18



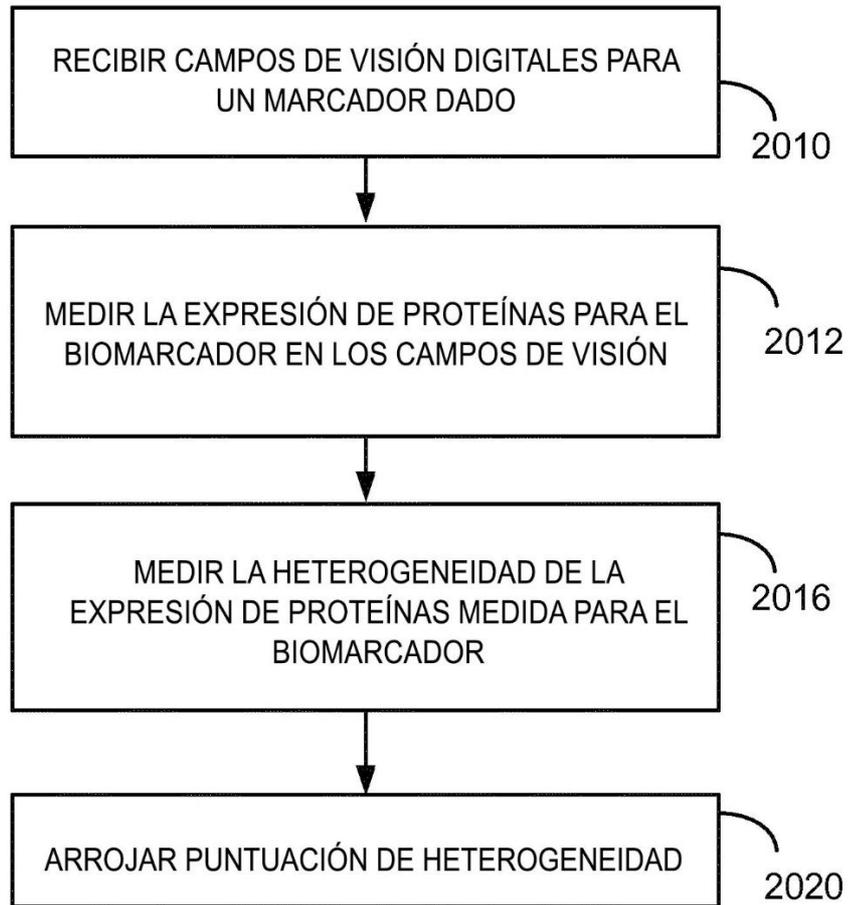
1800 ↗

FIG. 19



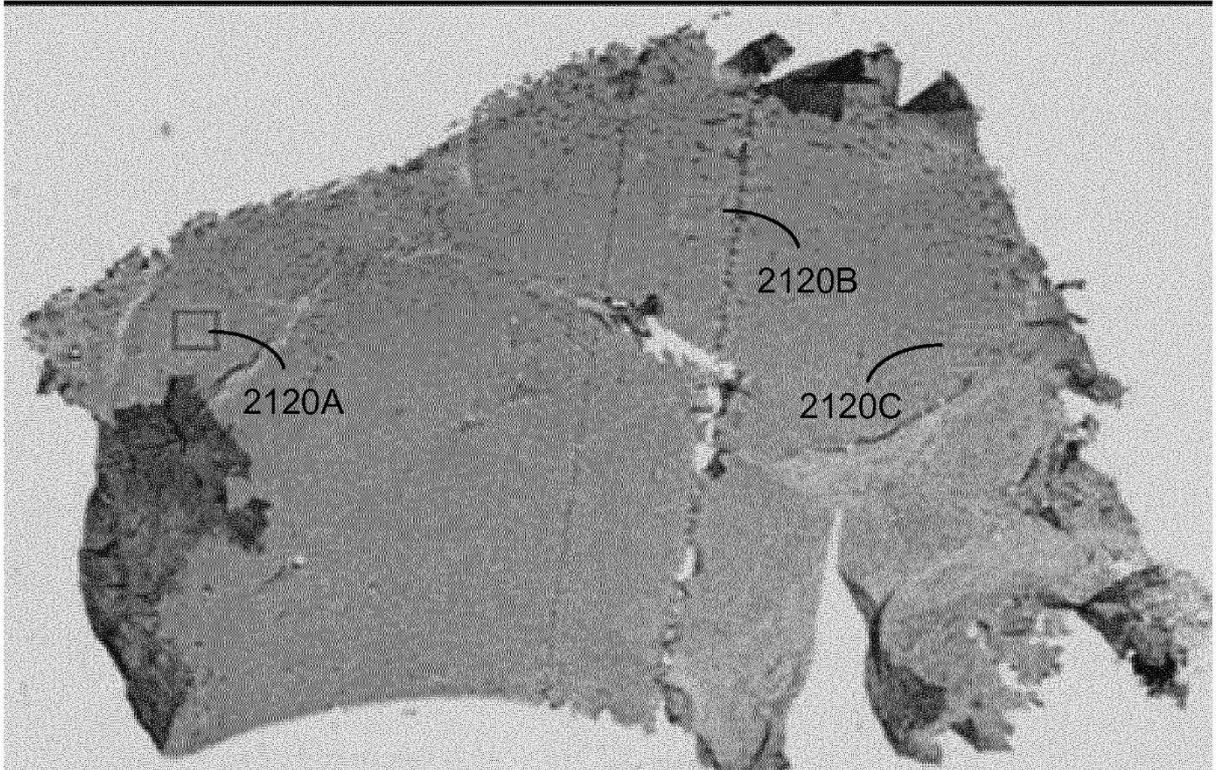
1900 ↗

FIG. 20



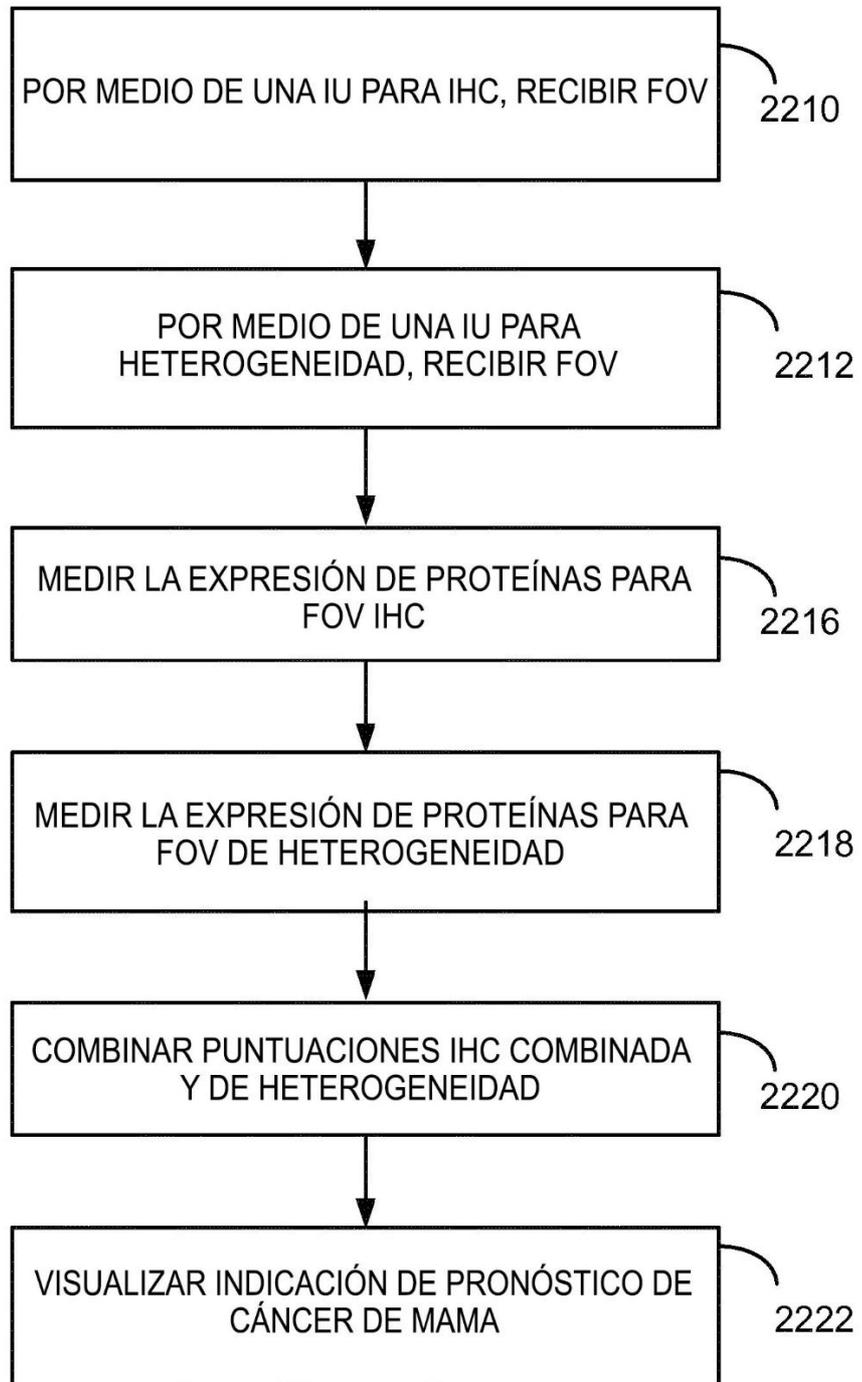
2000 

FIG. 21



2100 

FIG. 22



2200 ↗

FIG. 23

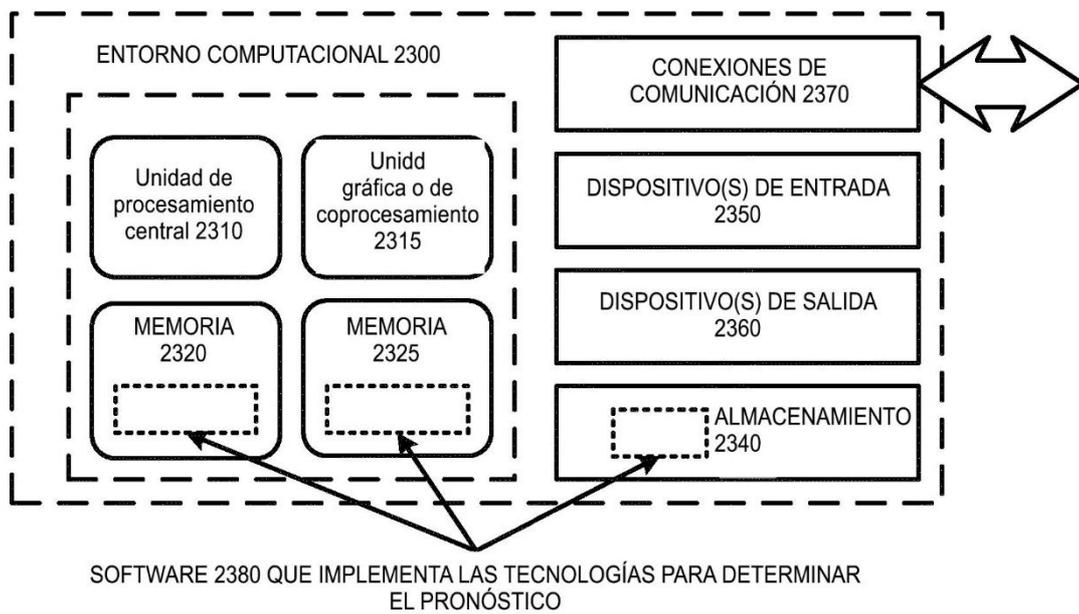


FIG. 24

