

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 777 002**

51 Int. Cl.:

G01N 33/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.09.2012 PCT/US2012/058060**

87 Fecha y número de publicación internacional: **04.04.2013 WO13049674**

96 Fecha de presentación y número de la solicitud europea: **28.09.2012 E 12836076 (5)**

97 Fecha y número de publicación de la concesión europea: **12.02.2020 EP 2761289**

54 Título: **Predicción de eventos de riesgo cardiovascular y usos de la misma**

30 Prioridad:

30.09.2011 US 201161541828 P

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

03.08.2020

73 Titular/es:

**SOMALOGIC, INC. (100.0%)
2945 Wilderness Place
Boulder, CO 80301, US**

72 Inventor/es:

**GILL, ROSALYNN DIANNE;
WILLIAMS, STEPHEN ALARIC;
STEWART, ALEX A.E.;
MEHLER, ROBERT;
FOREMAN, TRUDI y
SINGER, BRITTA**

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 2 777 002 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

Predicción de eventos de riesgo cardiovascular y usos de la misma

5 **Campo de la invención**

La presente solicitud se refiere en general a la detección de biomarcadores y a un método de evaluación del riesgo de un evento cardiovascular futuro en un individuo y, más específicamente, a uno o más biomarcadores, métodos, dispositivos, reactivos, sistemas y kits utilizados para evaluar a un individuo para la predicción del riesgo de desarrollar un evento cardiovascular (CV) durante un período de 5 años. Dichos eventos incluyen, pero sin limitación, infarto de miocardio, ictus, insuficiencia cardíaca congestiva o muerte.

Antecedentes

15 La siguiente descripción proporciona un resumen de información relevante para la presente solicitud y no es una admisión que cualquier información proporcionada o publicaciones a las que se hace referencia en el presente documento sea técnica anterior de la presente solicitud.

20 La enfermedad cardiovascular es la principal causa de muerte en los EE.UU. Existen varios predictores existentes e importantes de riesgo de eventos primarios (D'Agostino, R et al., "General Cardiovascular Risk Profile for Use in Primary Care: The Framingham Heart Study" *Circulation* 117: 743-53 (2008); y Ridker, P. et al., "Development and Validation of Improved Algorithms for the Assessment of Global Cardiovascular Risk in Women" *JAMA* 297 (6): 611-619 (2007)) y eventos secundarios (Shlipak, M. et al. "Biomarkers to Predict Recurrent Cardiovascular Disease: The Heart & Soul Study" *Am. J. Med.* 121: 50-57 (2008)) que se usan ampliamente en la práctica clínica y en ensayos terapéuticos. Desafortunadamente, las curvas características de funcionamiento del receptor, las relaciones de riesgos instantáneos y la concordancia muestran que el rendimiento de los factores de riesgo y biomarcadores existentes es modesto (las ABC de -0,75 significan que estos factores están solo a medio camino entre un lanzamiento de moneda y la perfección). Además de la necesidad de mejorar el rendimiento del diagnóstico, existe la necesidad de un producto de riesgo que sea a corto plazo y que responda personalmente en los individuos a las intervenciones beneficiosas (y destructivas) y los cambios en el estilo de vida. La ecuación de Framingham utilizada habitualmente tiene tres problemas principales. En primer lugar, es demasiado a largo plazo: proporciona cálculos de riesgo a 10 años, pero los seres humanos descartan los riesgos futuros y son reacios a realizar modificaciones de comportamiento y estilo de vida basadas en ellos. En segundo lugar, no responde muy bien a las intervenciones: su factor más ponderado es la edad cronológica, que no puede disminuirse. En tercer lugar, dentro de la población de alto riesgo prevista en el presente documento, los factores de Framingham no logran discriminar bien entre riesgo alto y bajo: la relación de riesgos instantáneos entre los cuartiles superior e inferior es de solo 2.

40 Los factores de riesgo de enfermedad cardiovascular se usan ampliamente para impulsar la intensidad y la naturaleza del tratamiento médico y su uso ha contribuido indudablemente a la reducción de la morbilidad y la mortalidad cardiovasculares que se ha observado en las últimas dos décadas. Estos factores se han combinado rutinariamente en algoritmos, pero desafortunadamente no capturan todo el riesgo (la presentación inicial más común para la enfermedad cardíaca sigue siendo la muerte). De hecho, probablemente solo capturan la mitad del riesgo. Un área bajo la curva ROC de -0,76 es típica para dichos factores de riesgo y, nuevamente, está a medio camino entre un lanzamiento de moneda a 0,5 y la perfección a 1,0.

45 La adición de nuevos biomarcadores a las puntuaciones de riesgo clínico ha sido decepcionante. Por ejemplo, en el estudio de Framingham (Wang et al., "Multiple Biomarkers for the Prediction of First Major Cardiovascular Events and Death" *N. Eng. J. Med.* 355: 2631-2637 (2006)) en 3209 personas, la adición de 10 biomarcadores (CRP, BNP, NT-proBNP, aldosterona, renina, fibrinógeno, dímero D, inhibidor del activador de plasminógeno de tipo 1, homocisteína y la relación de albúmina urinaria a creatinina), no mejoró significativamente el ABC cuando se añadieron a factores de riesgo existentes: el ABC para eventos de 0-5 años fue de 0,76 con la edad, el sexo y los factores de riesgo convencionales, y de 0,77 con la mejor combinación de biomarcadores añadidos a la mezcla.

50 La identificación precoz de pacientes con mayor riesgo de un evento cardiovascular dentro de una ventana de 1-10 años es importante porque un tratamiento más agresivo de individuos con riesgo elevado puede mejorar el resultado. Por tanto, el abordaje óptimo requiere una intervención agresiva para reducir el riesgo de un evento cardiovascular en aquellos pacientes que se considera que tienen un riesgo mayor, mientras que a los pacientes con un riesgo menor de un evento cardiovascular se les pueden evitar tratamientos caros y potencialmente invasivos, que probablemente no tendrán ningún efecto beneficioso para el paciente.

60 La selección de biomarcadores para la predicción del riesgo de tener una patología o afección específica dentro de un período de tiempo definido implica primero la identificación de marcadores que tienen una diferencia medible y estadísticamente significativa en las poblaciones en las que el evento ha ocurrido o no durante el período de tiempo para una aplicación médica específica. Los biomarcadores pueden incluir moléculas secretadas o desprendidas que son paralelas al desarrollo o progresión de una enfermedad o afección y difunden fácilmente en el torrente sanguíneo desde el tejido cardiovascular o desde los tejidos circundantes y las células circulantes en respuesta a un evento

cardiovascular. El biomarcador o conjunto de biomarcadores identificados generalmente se validan clínicamente o se demuestra que son un indicador fiable para el uso original previsto para el que se seleccionó. Los biomarcadores pueden incluir moléculas pequeñas, péptidos, proteínas y ácidos nucleicos. Algunos de los problemas clave que afectan a la identificación de biomarcadores incluyen el ajuste excesivo de los datos disponibles y el sesgo en los datos.

Se ha utilizado diversos métodos en un intento de identificar biomarcadores y diagnosticar o predecir el riesgo de tener una enfermedad o afección. Para los marcadores a base de proteínas, estos incluyen electroforesis bidimensional, espectrometría de masas y métodos de inmunoensayo. Para los marcadores de ácido nucleico, estos incluyen perfiles de expresión de ARNm, perfiles de microARN, FISH, análisis en serie de expresión génica (SAGE), matrices de expresión génica a gran escala, secuenciación y genotipado de genes (SNP o análisis de variantes pequeñas).

La utilidad de la electroforesis bidimensional está limitada por la sensibilidad de detección baja; problemas con la solubilidad de proteína, carga e hidrofobia; reproducibilidad en gel; y la posibilidad de un único punto que represente múltiples proteínas. Para la espectrometría de masas, dependiendo del formato utilizado, las limitaciones giran en torno al procesamiento y la separación de la muestra, la sensibilidad a las proteínas de abundancia baja, las consideraciones de señal a ruido y la incapacidad para identificar inmediatamente la proteína detectada. Las limitaciones en los enfoques de inmunoensayo para el descubrimiento de biomarcadores se centran en la incapacidad de los ensayos múltiples basados en anticuerpos para medir un gran número de analitos. Podría simplemente imprimirse una matriz de anticuerpos de alta calidad y, sin sándwiches, medir los analitos unidos a esos anticuerpos. (Esto sería el equivalente formal de usar un genoma completo de secuencias de ácido nucleico para medir mediante hibridación todas las secuencias de ADN o ARN en un organismo o una célula. El experimento de hibridación funciona porque la hibridación puede ser un ensayo riguroso de identidad. Incluso anticuerpos muy buenos no son suficientemente rigurosos en la selección de sus compañeros de unión para trabajar en el contexto de sangre o incluso extractos celulares porque el conjunto de proteínas en esas matrices tiene abundancias extremadamente diferentes). Por tanto, debe usarse un enfoque diferente con enfoques a base de inmunoensayo para el descubrimiento de biomarcadores: sería necesario usar ensayos ELISA multiplexados (es decir, sándwiches) para obtener la rigurosidad suficiente para medir muchos analitos simultáneamente para decidir qué analitos son realmente biomarcadores. Los inmunoensayos de tipo sándwich no se escalan a un contenido alto y, por tanto, el descubrimiento de biomarcadores usando inmunoensayos sándwich rigurosos no es posible usando formatos de matriz convencionales. Por último, los reactivos de anticuerpos se someten a variabilidad de lote e inestabilidad de reactivo sustanciales. La presente plataforma para el descubrimiento de biomarcadores de proteína supera este problema.

Muchos de estos métodos dependen o requieren algún tipo de fraccionamiento de muestra antes del análisis. Por tanto, la preparación de la muestra requerida para ejecutar un estudio suficientemente potente diseñado para identificar y descubrir biomarcadores estadísticamente relevantes en una serie de poblaciones de muestra bien definidas es extremadamente difícil, costoso y lento. Durante el fraccionamiento, puede introducirse un amplio intervalo de variabilidad en las diversas muestras. Por ejemplo, un marcador potencial podría ser inestable para el proceso, la concentración del marcador podría cambiar, podría producirse agregación o desagregación inadecuada y podría producirse una contaminación inadvertida de la muestra y, de este modo, ocultar los cambios sutiles anticipados en la enfermedad precoz.

Es ampliamente aceptado que los métodos de descubrimiento y detección de biomarcadores usando estas tecnologías tienen graves limitaciones para la identificación de biomarcadores de diagnóstico o predictivos. Estas limitaciones incluyen la incapacidad de detectar biomarcadores de abundancia baja, la incapacidad de cubrir uniformemente todo el intervalo dinámico del proteoma, la irreproducibilidad en el procesamiento y fraccionamiento de la muestra y la irreproducibilidad global y la falta de robustez del método. Adicionalmente, estos estudios han introducido sesgos en los datos y no han abordado adecuadamente la complejidad de las poblaciones de muestra, incluyendo los controles adecuados, en términos de distribución y aleatorización requeridas para identificar y validar biomarcadores dentro de una población de enfermedades objetivo.

Aunque los esfuerzos dirigidos al descubrimiento de biomarcadores nuevos y eficaces han continuado durante varias décadas, los esfuerzos han sido en gran parte infructuosos. Se han identificado biomarcadores para diversas enfermedades normalmente en laboratorios académicos, por lo general a través de un descubrimiento accidental mientras se realizaba una investigación básica sobre algunos procesos de enfermedad. Basándose en el descubrimiento y con pequeñas cantidades de datos clínicos, se publicaron artículos que sugerían la identificación de un nuevo biomarcador. La mayoría de estos biomarcadores propuestos, sin embargo, no se han confirmado como biomarcadores reales o útiles, principalmente porque el pequeño número de muestras clínicas sometidas a ensayo proporciona solamente pruebas estadísticas débiles de que se ha descubierto un biomarcador eficaz. Es decir, la identificación inicial no fue rigurosa con respecto a los elementos básicos de la estadística. En cada uno de los años 1994 a 2003, una búsqueda en la bibliografía científica muestra que se publicaron miles de referencias referidas a biomarcadores. Durante ese mismo período de tiempo, sin embargo, la FDA aprobó el uso diagnóstico, como máximo, de tres nuevos biomarcadores proteínicos al año y en varios años no se aprobaron biomarcadores proteínicos nuevos.

Basándose en la historia de esfuerzos fallidos de descubrimiento de biomarcadores, se han propuesto teorías que promueven adicionalmente la comprensión general de que los biomarcadores para el diagnóstico, pronóstico o

predicción del riesgo de desarrollar enfermedades y afecciones son raros y difíciles de encontrar. La investigación de biomarcadores basada en geles 2D o espectrometría de masas respalda estas nociones. Se han identificado muy pocos biomarcadores útiles a través de estos enfoques. Sin embargo, por lo general se pasa por alto que el gel 2D y la espectrometría de masas miden las proteínas que están presentes en la sangre a concentraciones de aproximadamente 1 nM y más altas, y que este conjunto de proteínas puede ser el menos propenso a cambiar con la enfermedad o el desarrollo de una afección particular. Además de la presente plataforma de descubrimiento de biomarcadores, no existen plataformas de descubrimiento de biomarcadores proteómicos que puedan medir con precisión los niveles de expresión de proteínas a concentraciones mucho más bajas.

Se sabe mucho sobre las vías bioquímicas para la biología humana compleja. Muchas vías bioquímicas culminan o son iniciadas por proteínas secretadas que trabajan localmente dentro de la patología; por ejemplo, se secretan factores de crecimiento para estimular la replicación de otras células en la patología y se secretan otros factores para evitar el sistema inmunitario, etc. Aunque muchas de estas proteínas secretadas funcionan de manera paracrina, algunas operan distalmente en el cuerpo. Un experto en la materia con una comprensión básica de las vías bioquímicas comprendería que muchas proteínas específicas de la patología deberían existir en la sangre a concentraciones por debajo (incluso muy por debajo) de los límites de detección de geles 2D y espectrometría de masas. Lo que debe preceder a la identificación de este número relativamente abundante de biomarcadores de enfermedad es una plataforma proteómica que puede analizar proteínas a concentraciones por debajo de las detectables mediante geles 2D o espectrometría de masas.

Como se ha analizado anteriormente, los eventos cardiovasculares pueden prevenirse mediante un tratamiento agresivo si la propensión a dichos eventos puede determinarse con precisión. Los ensayos de marcadores múltiples existentes requieren la recogida de múltiples muestras de un individuo o requieren repartir una muestra entre múltiples ensayos. De manera óptima, un ensayo mejorado requeriría una sola muestra de sangre, orina u otra, y un solo ensayo. En consecuencia, existe la necesidad de biomarcadores, métodos, dispositivos, reactivos, sistemas y kits que permitan la predicción de eventos cardiovasculares dentro de un período de 5 años.

El documento WO2011/059721 describe biomarcadores de proteína y lípidos que proporcionan una mejora constante de la predicción de la diabetes de tipo 2.

El documento US2011/0003707 describe un inmunoensayo de múltiples marcadores, incluyendo biomarcadores de patología cardíaca e inflamación vascular para la detección precoz de la insuficiencia cardíaca congestiva (ICC).

Sumario de la invención

La invención es como se define en las reivindicaciones adjuntas.

La presente solicitud incluye biomarcadores, métodos, reactivos, dispositivos, sistemas y kits para la predicción del riesgo de tener un evento cardiovascular (CV) dentro de un período de 5 años. Los biomarcadores de la presente solicitud se identificaron usando un ensayo basado en SOMAmero múltiple que se describe en detalle en los Ejemplos 1 y 2. Mediante el uso del método de identificación de biomarcadores basado en SOMAmero que se describe en el presente documento, la presente solicitud describe un número sorprendentemente grande de biomarcadores de eventos CV que son útiles para la predicción de eventos CV. La población de muestra utilizada para descubrir biomarcadores asociados al riesgo de un evento CV era del *Heart & Soul Study* (Estudio de Corazón y Alma), un estudio de cohorte prospectivo que examina la progresión de la arteriopatía coronaria en una población con enfermedad CV preexistente, incluyendo infarto de miocardio previo, evidencia de una estenosis superior al 50 % en 1 o más vasos coronarios, isquemia inducida por ejercicio mediante cinta de correr o ensayos nucleares o revascularización coronaria previa. Los participantes fueron reclutados en el Área de la Bahía de San Francisco. El tipo de evento CV y el tiempo para la población de estudio se muestran en la Tabla 4. En la identificación de estos biomarcadores de eventos CV, se midieron más de 1000 proteínas de más de 900 muestras individuales, algunas de las cuales estaban en concentraciones en el intervalo femtomolar bajo. Esto es aproximadamente cuatro órdenes de magnitud inferior a los experimentos de descubrimiento de biomarcadores realizados con geles 2D y/o espectrometría de masas.

Aunque algunos de los biomarcadores de eventos CV descritos son útiles solos para la predicción del riesgo de tener un evento CV, en el presente documento se describen métodos para la agrupación de múltiples subconjuntos de los biomarcadores de eventos CV que son útiles como un panel de biomarcadores. Una vez que se ha identificado un biomarcador individual o un subconjunto de biomarcadores, la predicción del riesgo de un evento CV en un individuo puede lograrse usando cualquier plataforma o formato de ensayo que sea capaz de medir diferencias en los niveles del biomarcador o biomarcadores seleccionados en una muestra biológica.

Sin embargo, solo mediante el uso del método de identificación de biomarcadores basado en SOMAmero que se describe en el presente documento, en donde más de 1000 valores de biomarcadores potenciales separados se seleccionaron individualmente entre un gran número de individuos que habían sido diagnosticados anteriormente como que tenían o no un evento CV dentro de un marco temporal de 5 años, fue posible identificar los biomarcadores de eventos CV que se desvelan en el presente documento. Este enfoque de descubrimiento contrasta con el

descubrimiento de biomarcadores en muestras de tejido, medios acondicionados o células lisadas, ya que cuestiona un sistema más relevante para el paciente que no requiere traducción a la patología humana. Además, esta forma de medición basada en la sangre es mucho más aplicable clínicamente.

- 5 Por tanto, en un aspecto de la presente solicitud, se proporcionan uno o más biomarcadores para su uso solos o en diversas combinaciones para predecir el riesgo de la incidencia de un evento CV dentro de un marco temporal de 5 años. Las realizaciones de ejemplo incluyen los biomarcadores que se proporcionan en la Tabla 1, Col. 7, "NOMBRE PÚBLICO", que como se ha indicado anteriormente, se identificaron usando un ensayo basado en SOMAmérico múltiple, como se describe en general en el Ejemplo 1 y más específicamente en el Ejemplo 2. Los marcadores proporcionados en la Tabla 1 son útiles en la predicción del riesgo de tener un evento CV dentro de un periodo de tiempo de 5 años. Los biomarcadores de la Tabla 2 y la Tabla 3, respectivamente, demuestran la reducción de los 155 biomarcadores de la Tabla 1 a un número menor que realiza la misma tarea con menor complejidad técnica y coste; sin embargo, pueden compilarse otras combinaciones con eficacia similar a partir de la Tabla 1.
- 10
- 15 Aunque algunos de los biomarcadores de riesgo de eventos CV que se describen son útiles solos para la predicción del riesgo de un evento CV en 5 años, en el presente documento también se describen métodos para la agrupación de múltiples subconjuntos de los biomarcadores de riesgo de eventos CV que son útiles como panel de dos o más biomarcadores. Por tanto, diversas realizaciones de la presente solicitud proporcionan combinaciones que comprenden N biomarcadores, en donde N es al menos dos biomarcadores. En otras realizaciones, N se selecciona para que sea cualquier número entre 2-155 biomarcadores.
- 20

En otras realizaciones más, N se selecciona para que sea cualquier número entre 2-7, 2-10, 2-15, 2-20, 2-25, 2-30, 2-35, 2-40, 2-45, 2-50, 2-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 2-155. En otras realizaciones, N se selecciona para que sea cualquier número entre 3-7, 3-10, 3-15, 3-20, 3-25, 3-30, 3-35, 3-40, 3-45, 3-50, 3-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 3-155. En otras realizaciones, N se selecciona para que sea cualquier número entre 4-7, 4-10, 4-15, 4-20, 4-25, 4-30, 4-35, 4-40, 4-45, 4-50, 4-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 4-155. En otras realizaciones, N se selecciona para que sea cualquier número entre 5-7, 5-10, 5-15, 5-20, 5-25, 5-30, 5-35, 5-40, 5-45, 5-50, 5-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 5-155. En otras realizaciones, N se selecciona para que sea cualquier número entre 6-10, 6-15, 6-20, 6-25, 6-30, 6-35, 6-40, 6-45, 6-50, 6-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 6-155. En otras realizaciones, N se selecciona para que sea cualquier número entre 7-10, 7-15, 7-20, 7-25, 7-30, 7-35, 7-40, 7-45, 7-50, 7-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 7-155. En otras realizaciones, N se selecciona para que sea cualquier número entre 8-10, 8-15, 8-20, 8-25, 8-30, 8-35, 8-40, 8-45, 8-50, 8-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 8-155. En otras realizaciones, N se selecciona para que sea cualquier número entre 9-15, 9-20, 9-25, 9-30, 9-35, 9-40, 9-45, 9-50, 9-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 9-155. En otras realizaciones, N se selecciona para que sea cualquier número entre 10-15, 10-20, 10-25, 10-30, 10-35, 10-40, 10-45, 10-50, 10-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 10-155. Se apreciará que N puede seleccionarse para abarcar intervalos similares, pero de orden superior.

25

30

35

40

Como se ha analizado anteriormente, los eventos cardiovasculares pueden evitarse mediante un tratamiento agresivo si la propensión a dichos eventos puede determinarse con precisión. Los ensayos de marcadores múltiples de la técnica anterior requieren la recogida de múltiples muestras de un individuo o requieren que se reparta una muestra entre múltiples ensayos. Se preferiría proporcionar un ensayo de pronóstico que requeriría solamente una única muestra biológica, medida en un único ensayo, en lugar de múltiples muestras para diferentes tipos de analitos (lípidos, proteínas, metabolitos) o paneles de analitos. El beneficio central de un ensayo de muestra única es la simplicidad en el punto de uso, puesto que un ensayo con colecciones de múltiples muestras es más complejo de administrar y esto constituye una barrera para su adopción. Una ventaja adicional deriva de ejecutar esa única muestra en un único ensayo para múltiples proteínas. Un único ensayo debería mitigar la variación no deseada debido a la calibración de múltiples resultados de ensayo juntos. El ensayo que forma la base de la presente solicitud es un ensayo de "muestra única, ensayo único" de este tipo. Esta combinación de muestra única y ensayo único es una característica novedosa de este ensayo de riesgo de eventos cardiovasculares que aborda la complejidad logística de recoger múltiples muestras y los problemas y riesgos biológicos implicados en la división de muestras en múltiples alícuotas para múltiples procedimientos analíticos independientes.

45

50

55

Se sabe que las enfermedades cardiovasculares implican múltiples procesos biológicos y tejidos. Son ejemplos bien conocidos de sistemas y procesos biológicos asociados a enfermedades cardiovasculares la inflamación, trombosis, angiogénesis asociada a enfermedad, activación de plaquetas, activación de macrófagos, respuesta aguda hepática, remodelación de la matriz extracelular y función renal. Estos procesos pueden observarse en función del género, el estado menopáusico y la edad y de acuerdo con el estado de coagulación y la función vascular. Puesto que estos sistemas se comunican parcialmente a través de sistemas de señalización basados en proteínas y pueden medirse múltiples proteínas en una única muestra de sangre, la invención proporciona un ensayo basado en múltiples proteínas de muestra única y ensayo único, enfocado en proteínas de los sistemas biológicos y procesos específicos implicados en enfermedades cardiovasculares.

60

65

Como se analiza en el presente documento, una de las funciones centrales de medir el riesgo de un evento cardiovascular es permitir la evaluación del progreso en respuesta al tratamiento y los cambios conductuales tales como la dieta y el ejercicio. Los métodos actuales de predicción de riesgo, tales como la ecuación de Framingham, incluyen información covariante clínica claramente correlacionada, siendo el factor más importante la edad del sujeto.

5 Esto hace que la ecuación de Framingham sea menos útil para controlar el cambio en el riesgo de un individuo, aunque puede ser precisa para una población. Una característica novedosa del presente ensayo de riesgo de eventos CV es que no necesita la edad como parte del modelo de pronóstico. La presente invención se basa en la premisa de que, dentro de la biología del envejecimiento, existen factores causales que son variables y, por tanto, se usan mejor para evaluar el riesgo. La invención se basa en la creencia de que la edad en sí misma no es un factor causal en la enfermedad y que la edad actúa como un sustituto o representante de la biología subyacente. Aunque la edad es, de hecho, un factor pronóstico de eventos CV, no puede usarse para evaluar la mejoría individual y, presumiblemente, el efecto de la edad está mediado a través de la función biológica. Este efecto puede determinarse mejor a través de la medición de la biología relevante. En la presente invención, las proteínas que son dianas están implicadas en la biología de la enfermedad. Por tanto, la invención captura la información biológica que se refleja en la correlación entre la edad y el riesgo de un evento CV. De hecho, añadir un factor de edad al modelo de riesgo de los presentes inventores basado en proteínas no mejora el rendimiento en la predicción de eventos.

La estrategia para identificar proteínas de múltiples procesos implicados en enfermedades cardiovasculares requería elegir parámetros que proporcionasen una amplia gama/diversidad de pacientes con enfermedad CV que presentasen diversos eventos o síntomas. Los eventos debidos a enfermedades cardiovasculares son heterogéneos e implican dos clases principales de eventos: eventos trombóticos y relacionados con la ICC. Algunos eventos de presentación pueden carecer de información diagnóstica específica (por ejemplo, muerte en el hogar). En vista de estas características de las enfermedades CV, el ensayo de la invención se desarrolló midiendo proteínas implicadas en los procesos biológicos asociados a enfermedades CV, en muestras de sangre de una amplia gama de eventos. Esta estrategia dio como resultado la inclusión de información de múltiples procesos implicados en la enfermedad (por ejemplo, angiogénesis, activación de plaquetas, activación de macrófagos, respuesta aguda hepática, otra inflamación linfocítica, remodelación de la matriz extracelular y función renal). Con el fin de desarrollar un ensayo de muestra única de pronóstico basado en proteínas múltiples para enfermedades CV, la población de estudio elegida era un grupo de sujetos de alto riesgo del estudio "*Heart & Soul*". Mediante la elección de este conjunto de sujetos con una alta tasa de eventos CV, fue posible determinar el riesgo asociado a las mediciones de proteínas con mayor precisión de lo que hubiera sido posible en la población general (dentro de la cual los eventos son más raros). El desarrollo del presente ensayo en este grupo de alto riesgo permitió la identificación de combinaciones de biomarcadores de proteína que podrían generalizarse debido a la biología común. Como resultado, el presente ensayo de la invención y los biomarcadores probablemente sean eficaces más allá de la predicción de eventos en una población más grande que aquellos individuos que coinciden con los criterios de entrada del estudio "*Heart & Soul*".

Como se ha mencionado anteriormente, las enfermedades CV implican el sistema de coagulación de la sangre, los glóbulos blancos inflamatorios y la activación plaquetaria. Las señales de la activación de estos sistemas en el cuerpo pueden ocultarse debido a errores comunes en la preparación de la muestra que conducen a que las plaquetas y los glóbulos blancos se sedimenten parcialmente en las muestras de plasma. Si estas células no se sedimentan por completo, pueden lisarse por congelación-descongelación cuando las muestras se transportan y someten a ensayo. Durante el curso de la identificación de los presentes biomarcadores, se hizo evidente que, al menos en algunos casos, las muestras preparadas convencionalmente contienen células enteras y plaquetas después de la congelación-descongelación. Durante el posterior ensayo proteómico, cualquier célula completa se lisaría e interferiría con la detección de proteínas características de los procesos de enfermedad de activación *in vivo* de plaquetas y monocitos. Por tanto, en una realización de la presente invención, se realiza una etapa adicional de volver a centrifugar las muestras después de la descongelación antes del ensayo. Esta etapa de centrifugado adicional puede retirar las plaquetas y los monocitos que de otro modo evitarían la identificación de biomarcadores relacionados con la activación de plaquetas y monocitos. La etapa adicional para retirar los componentes insolubles y celulares de las muestras (mediante centrifugación o filtración), representa un avance para un ensayo de riesgo de eventos cardiovasculares que se cree que no se describe en la técnica anterior.

Aunque existen proteínas específicas de la bibliografía que se sabe que son factores pronósticos para eventos CV, tales como la apolipoproteína B, la apolipoproteína A-1, BNP y CRP, y que tienen asociaciones conocidas con enfermedades CV, no ha quedado claro que un conjunto particular de mediciones de proteínas pueda combinarse para obtener un rendimiento óptimo en términos de rendimiento de predicción debido a que la información biológica común está representada por mediciones de proteínas múltiples. Un ejemplo específico de cómo las combinaciones de proteínas que se observa que varían durante las enfermedades CV pueden no proporcionar un rendimiento de predicción óptimo, implica las muchas proteínas séricas pequeñas excretadas en la orina que están relacionadas con la función renal según se mide mediante la tasa de filtración glomerular (TFG). La mala función renal, indicada mediante la TFG baja, se relaciona con el riesgo cardiovascular. Por tanto, las muchas proteínas pequeñas asociadas a la TFG baja, parecen estar relacionadas con las enfermedades CV, pero no independientemente. Durante el desarrollo del presente ensayo, se realizaron correcciones en las mediciones de proteínas para la TFG estimada con el fin de determinar qué proteínas proporcionaron un valor pronóstico adicional más allá de la TFG.

65 La medición de la TFG es claramente útil para predecir el riesgo de un evento CV. Sin embargo, la medición clínica

de la TFG implica la recogida de orina durante 24 horas, lo que no cumple con el presente criterio de un ensayo de "muestra única, ensayo único". Otras estimaciones de la TFG son menos onerosas; sin embargo, para cumplir con el objetivo de un ensayo de pronóstico de "muestra única", la estrategia subyacente a la presente invención buscó el uso de las propias mediciones de proteínas para proporcionar información de TFG para el análisis de riesgo. Por ejemplo,

5 en el modelo de diez marcadores de la Tabla 3, la proteína ESAM predice fuertemente el riesgo de evento CV debido a su correlación con la TFG. Después de corregir las mediciones de la proteína ESAM para retirar la correlación con la TFG estimada, ESAM ya no es predictiva de riesgo. Este uso de una proteína tal como ESAM para transmitir la señal biológica relacionada con la TFG en una "muestra única, ensayo único" representa un avance novedoso para el pronóstico de un evento CV.

10 La identificación de los biomarcadores de la Tabla 3 implicó la selección de proteínas que podrían funcionar juntas en el pronóstico de un evento CV. Estudios, tales como Wang, T. et al., (2006) *N. Eng. J. Med.* 355: 2631-9, han demostrado que las combinaciones de biomarcadores con frecuencia no pueden proporcionar un rendimiento adicional sobre cálculos de riesgo simples usando información clínica común tal como la edad y los niveles de lípidos. Con el fin de evitar una combinación *ad hoc* de biomarcadores, la presente invención proporciona un procedimiento de análisis estadístico en el que las proteínas se seleccionaron para determinar su poder pronóstico individual y también, de manera crucial, la capacidad de las proteínas para trabajar juntas sinérgicamente para mejorar el valor pronóstico de la combinación. Se representan múltiples procesos biológicos independientes en el modelo de proteína de diez marcadores de la Tabla 3 que se proporciona en el presente documento.

20 En una realización, la invención comprende un método para evaluar el riesgo de un evento cardiovascular (CV) futuro dentro de un período de 5 años en una población. Este método comprende la detección, en una muestra biológica de un individuo de la población, de valores de biomarcadores que corresponden cada uno a uno de al menos N biomarcadores seleccionados entre la Tabla 1, en donde se evalúa el riesgo para un individuo de un evento CV, basado en los valores de biomarcadores y en donde N = 2-155. En otra realización, los biomarcadores se seleccionan entre la Tabla 2 y N = 2-46. En otra realización, los biomarcadores se seleccionan entre la Tabla 3 y N = 2-10.

25 En una realización, la selección de una población es de manera que la población se caracterice por no tener antecedentes de enfermedad cardiovascular. Como alternativa, la población puede seleccionarse de manera que se caracterice por tener antecedentes de enfermedad cardiovascular.

30 Los antecedentes pueden comprender infarto de miocardio previo, evidencia angiográfica de una estenosis superior al 50 % en 1 o más vasos coronarios, isquemia inducida por ejercicio mediante cinta de correr o ensayos nucleares o revascularización coronaria previa.

35 Adicionalmente, la población puede seleccionarse de manera que se caracterice por factores de riesgo genéticos que comprenden mutaciones, polimorfismos de un único nucleótido e inserción/supresiones. Dichos factores de riesgo genéticos pueden usarse para complementar la evaluación del riesgo.

40 La evaluación del riesgo de un evento CV puede medirse en una escala dinámica que responda al cambio a lo largo del tiempo en respuesta a intervenciones que comprenden terapias, programas nutricionales, suplementos, modificación del estilo de vida, programas para dejar de fumar y protocolos de abordaje de enfermedades.

45 Los métodos anteriores relacionados con la evaluación del riesgo de eventos CV durante un período de 5 años pueden usarse para asignar los individuos a programas de abordaje de enfermedades aumentados o disminuidos, en función de sus valores de biomarcadores. El método también puede usarse para estratificar a los individuos en diferentes bandas de riesgo relacionadas con la cobertura del seguro de vida, dependiendo del valor de dicho biomarcador. Además, puede usarse para la evaluación del riesgo de evento CV con el fin de estratificar a los individuos en diferentes bandas de riesgo relacionadas con la cobertura del seguro de salud, dependiendo del valor de dicho biomarcador.

50 Adicionalmente, puede usarse para evaluar posibles candidatos para la asociación dependiendo de los valores de biomarcadores.

55 Adicionalmente, los métodos anteriores para la predicción del riesgo de eventos CV pueden usarse para: predecir el consumo de recursos médicos de la población basándose en los valores de biomarcadores; usar el valor de biomarcador del individuo como criterio de entrada para ensayos clínicos de terapias CV; predicción de la eficacia de resultados de ensayos clínicos basándose en dicho valor de biomarcador; usar el valor de biomarcador para la vigilancia de seguridad cardiovascular de un producto terapéutico CV o cualquier agente terapéutico; usar el valor de biomarcador como criterio de valoración sustituto de eficacia de productos terapéuticos CV; y/o controlar el cumplimiento de cualquier intervención, protocolo dietético o terapéutico basándose en dicho valor de biomarcador.

60 Con respecto a la vigilancia de seguridad CV de un agente terapéutico CV o de cualquier agente terapéutico, dicha vigilancia es importante en estudios de seguridad CV de fase 3 grandes, costosos, necesarios para fármacos terapéuticos CV y no cardiovasculares para casi todos los usos crónicos.

65 El presente método para evaluar un riesgo de evento CV también puede usarse para seleccionar o derivar al individuo a otros procedimientos de diagnóstico basados en dicho valor de biomarcador. Adicionalmente, el presente método puede usarse para seleccionar un producto terapéutico CV basándose en el valor de biomarcador.

5 En el presente método para la evaluación del riesgo de evento CV, los valores de biomarcadores pueden detectarse realizando un ensayo *in vitro*. El ensayo *in vitro* puede implicar al menos un reactivo de captura correspondiente a cada uno de los biomarcadores y, adicionalmente, puede incluir al menos un reactivo de captura del grupo que consiste en SOMAmeros, anticuerpos y una sonda de ácido nucleico. En una realización preferida, el reactivo de captura es un SOMAmero.

10 En otra realización, el ensayo *in vitro* puede seleccionarse entre el grupo que consiste en un inmunoensayo, un ensayo basado en SOMAmero, un ensayo histológico o citológico y un ensayo de nivel de expresión de ARNm.

15 En los presentes métodos, la muestra biológica puede ser sangre completa, plasma, suero, orina o similares. En una realización preferida, la muestra biológica es suero, plasma u orina.

Adicionalmente, se proporciona que en los presentes métodos el individuo puede ser un mamífero y, en particular, un ser humano.

En realizaciones alternativas de los presentes métodos, N = 3-10; N = 3-15; N = 2-10; N = 4-10; o N = 5-10.

20 En otra realización, la divulgación proporciona un método implementado por ordenador para evaluar el riesgo de un evento cardiovascular (CV). Este método puede incluir recuperar en un ordenador información de biomarcadores para un individuo, en donde la información de biomarcadores comprende valores de biomarcadores que corresponden cada uno a al menos N biomarcadores seleccionados entre la Tabla 1; realizar con el ordenador una clasificación de cada uno de los valores de biomarcadores; e indicar un resultado de la evaluación del riesgo para un evento CV para dicho individuo basándose en una pluralidad de clasificaciones, y en donde N = 2-155. En realizaciones alternativas del presente método, los biomarcadores pueden seleccionarse entre la Tabla 2 (con N = 2-46) o la Tabla 3 (con N = 2-10).

30 El resultado de la evaluación del riesgo de un evento CV para el individuo puede mostrarse en una pantalla de ordenador.

35 En otra realización, la divulgación comprende un producto de programa informático para evaluar el riesgo de un evento CV. El producto de programa informático puede incluir un medio legible por ordenador que incorpore un código de programa ejecutable por un procesador de un dispositivo o sistema informático, donde el código de programa comprende: código que recupera datos atribuidos a una muestra biológica de un individuo, en donde los datos comprenden valores de biomarcadores que corresponden cada uno a uno de al menos N biomarcadores seleccionados entre la Tabla 1, donde los biomarcadores se detectaron en la muestra biológica; y un código que ejecuta un método de clasificación que indica un resultado de la evaluación del riesgo para un evento CV del individuo en función de dichos valores de biomarcadores. Cuando los biomarcadores se seleccionan entre la Tabla 1, col. 7, N = 2-155. En otras realizaciones, los biomarcadores pueden seleccionarse entre la Tabla 2 (con N = 2-46) o entre la Tabla 3 (con N = 2-10).

El método de clasificación puede usar una puntuación o medida o métrica de riesgo continua. El método de clasificación también puede usar dos o más clases.

45 La presente divulgación comprende adicionalmente un método para seleccionar un individuo para la evaluación del riesgo de un evento CV. Este método comprende la detección, en una muestra biológica de un individuo, de valores de biomarcadores que corresponden cada uno a uno de al menos N biomarcadores seleccionados entre la Tabla 1, en donde el individuo se evalúa para determinar el riesgo de un evento CV basándose en dichos valores de biomarcadores y en donde N = 2-155. En otras realizaciones, los biomarcadores pueden seleccionarse entre la Tabla 2 (con N = 2-46) o entre la Tabla 3 (con N = 2-10), respectivamente.

50 En los presentes métodos, la detección de los valores de biomarcadores puede realizarse en un ensayo *in vitro*. Dicho ensayo *in vitro* puede incluir al menos un reactivo de captura correspondiente a cada uno de los biomarcadores y puede comprender adicionalmente la selección del al menos un reactivo de captura entre el grupo que consiste en SOMAmeros, anticuerpos y una sonda de ácido nucleico. Preferentemente, el al menos un reactivo de captura es un SOMAmero. El ensayo *in vitro* puede seleccionarse entre un inmunoensayo, un ensayo basado en SOMAmero, un ensayo histológico o citológico y un ensayo de nivel de expresión de ARNm.

60 La muestra biológica puede seleccionarse entre sangre completa, plasma, suero, orina y similares. Preferentemente, la muestra biológica es suero, plasma u orina. En el presente método, el individuo puede ser un mamífero y preferentemente es un ser humano.

65 En una de las realizaciones de la presente divulgación, se proporciona que el individuo se evalúa para determinar el riesgo de un evento CV, basándose en dichos valores de biomarcadores y al menos un elemento de información biomédica adicional correspondiente a dicho individuo. Al menos un elemento de información biomédica adicional puede incluir, pero sin limitación, cualquiera de los siguientes: (a) información correspondiente a la presencia de

factores de riesgo cardiovasculares incluyendo uno o más de entre un infarto de miocardio previo, evidencia angiográfica de una estenosis superior al 50 % en uno o más vasos coronarios, isquemia inducida por ejercicio mediante cinta de correr o ensayos nucleares o revascularización coronaria previa; (b) información correspondiente a descriptores físicos de dicho individuo; (c) información correspondiente a un cambio en el peso de dicho individuo; (d) información correspondiente a la etnia de dicho individuo; (e) información correspondiente al género de dicho individuo; (f) información correspondiente al historial de tabaquismo de dicho individuo; (g) información correspondiente al historial de consumo de alcohol de dicho individuo; (h) información correspondiente al historial ocupacional de dicho individuo; (i) información correspondiente a antecedentes familiares de enfermedades cardiovasculares u otras afecciones del sistema circulatorio de dicho individuo; (j) información correspondiente a la presencia o ausencia en dicho individuo de al menos un marcador genético que se correlaciona con un mayor riesgo de enfermedad cardiovascular en dicho individuo o un miembro de la familia de dicho individuo; (k) información correspondiente a síntomas clínicos del individuo; (l) información correspondiente a otros ensayos de laboratorio; (m) información correspondiente a valores de expresión génica del individuo; (n) información correspondiente al consumo de factores de riesgo cardiovasculares conocidos de dicho individuo, tales como una dieta rica en grasas saturadas, rica en sal; (o) información correspondiente a los estudios de imagen del individuo, incluyendo electrocardiograma, ecocardiografía, ultrasonido carotídeo para determinar el espesor íntima-media, dilatación mediada por flujo, velocidad de la onda del pulso, índice tobillo-brazo, ecocardiografía con estrés, imágenes de perfusión miocárdica, calcio coronario por TC, angiografía por TC de alta resolución, resonancia magnética y otras modalidades de formación de imágenes; e (p) información sobre medicamentos.

La divulgación comprende adicionalmente un panel de biomarcadores para evaluar el riesgo de un evento CV futuro dentro de un período de tiempo de cinco años, en donde el panel comprende N biomarcadores seleccionados entre el grupo que consiste en los biomarcadores de la Tabla 1, en donde N = 2-155. En realizaciones alternativas, los biomarcadores se seleccionan entre la Tabla 2, en donde N = 2-46, o entre la Tabla 3, en donde N = 2-10.

En otro aspecto, la divulgación comprende un procedimiento para seleccionar un individuo en una población evaluando o pronosticando el riesgo de un evento CV futuro dentro de un período de 5 años, detectando, en una muestra biológica del individuo, un valor de biomarcador para angiopoyetina 2, y determinando el riesgo de un evento CV futuro basándose en el valor de biomarcador de angiopoyetina 2. El valor de biomarcador puede expresarse como una puntuación de medición o una clasificación en una de una pluralidad de clasificaciones.

En una variación del método de selección usando angiopoyetina 2, la presente divulgación también comprende añadir al método, antes de la etapa de determinación, la etapa que comprende proporcionar información sobre el uso de una estatina por el individuo. Por tanto, la determinación del riesgo de un evento CV futuro se basa en el valor de biomarcador de angiopoyetina 2 y la información de estatinas.

La angiopoyetina 2 es sorprendentemente útil en el pronóstico de un evento cardiovascular secundario para individuos que toman estatinas. Se ha informado que las estatinas en la técnica anterior no solo reducen el riesgo de un evento cardiovascular secundario, sino que también provocan un aumento en la angiopoyetina 2. Se esperaba que este aumento en la angiopoyetina 2 negara su uso como biomarcador. Inesperadamente, la angiopoyetina 2 ha demostrado ser un buen marcador para la predicción de eventos cardiovasculares secundarios en individuos de alto riesgo.

El método de detección del valor de biomarcador de angiopoyetina 2 puede comprender la etapa adicional de: detectar en la muestra biológica un valor de biomarcador para uno o más de los biomarcadores MMP7, CHRDL1, MATN2, PSA-ACT o una combinación de los mismos. El método puede implicar adicionalmente la detección, en la muestra biológica, de valores de biomarcadores que corresponden cada uno a N biomarcadores seleccionados entre los biomarcadores de la Tabla 3, en donde N = 2-10.

En otro aspecto, la divulgación proporciona un panel de biomarcadores para seleccionar un individuo en una población evaluando o pronosticando el riesgo de un evento CV futuro dentro de un período de 5 años. El panel incluye al menos el biomarcador de angiopoyetina 2. Este panel puede incluir adicionalmente uno o más de los biomarcadores MMP7, CHRDL1, MATN2, PSA-ACT o cualquier combinación de los mismos. Además, el panel puede incluir uno o más biomarcadores seleccionados entre la Tabla 3, en donde N = 2-10.

En otra realización, la divulgación proporciona un método para seleccionar un individuo mediante la evaluación del riesgo de un evento CV futuro dentro de un período de 5 años, en donde la evaluación comprende un pronóstico diferencial de un evento trombotico o evento de insuficiencia cardíaca congestiva (ICC). Este método comprende: detectar en una muestra biológica del individuo de la población, los valores de biomarcadores que corresponden cada uno al biomarcador GPVI para el pronóstico del evento trombotico y el biomarcador MATN2 para el pronóstico del evento de ICC. El método puede implicar la etapa adicional de detectar en la muestra biológica los valores de biomarcadores para N biomarcadores seleccionados entre el grupo de biomarcadores que se expone en la Tabla 3, en donde N = 3-10. El evento trombotico puede incluir cualquiera de entre un infarto de miocardio (IM), un ataque isquémico transitorio (AIT), un ictus, un síndrome coronario agudo y la necesidad de revascularización coronaria.

Adicionalmente, se proporciona un panel de biomarcadores para seleccionar un individuo en una población evaluando o pronosticando el riesgo de un evento CV futuro dentro de un período de 5 años, en donde el panel comprende un

biomarcador GPVI y un biomarcador MATN2. El panel también puede incluir adicionalmente al menos uno de los N biomarcadores seleccionados entre el grupo que consiste en los biomarcadores que se exponen en la Tabla 3.

5 Hay disponibles múltiples clases de tratamiento para las enfermedades CV, lo que refleja la diversidad de sistemas biológicos implicados. Por ejemplo, se han utilizado antitrombóticos, inhibidores de plaquetas, medicamentos para el equilibrio del metabolismo lipídico, de los líquidos y los electrolitos, y betabloqueantes, en el tratamiento de las enfermedades CV. Con el fin de guiar el tratamiento, es útil identificar no solo el riesgo global, sino también distinguir la clase de evento indicada por la biología. El método anterior usando MATN2 y GPVI permite distinguir las clases de evento probables de eventos tromboticos y eventos de ICC. El GPVI es más específico para el desarrollo del evento trombotico y el MATN2 es más específico para los eventos de ICC. La especificidad de GPVI para eventos tromboticos se demuestra en las Figuras 8A y 8B. La especificidad de MATN2 para el pronóstico de ICC se ilustra en las Figuras 9A y 9B. Estas diferencias pueden interpretarse en términos de los procesos biológicos relacionados. Por tanto, este ensayo basado en múltiples proteínas puede proporcionar al paciente, por tanto, información para distinguir el riesgo de desarrollar ICC frente al riesgo de eventos tromboticos. Esta es una característica significativa e importante de la invención que se cree que no se describe en la técnica anterior.

Además de proporcionar un pronóstico del riesgo de eventos CV basado únicamente en mediciones de proteínas, el presente método también proporciona la ventaja de una imagen más completa derivada de tener en cuenta información simple tal como el género, la medicación, otros marcadores tales como el colesterol LDL, el colesterol HDL, el colesterol total y otras afecciones tales como la diabetes. Dichos modelos pueden construirse sobre el modelo existente de diez proteínas de la Tabla 3 que se presenta en el presente documento.

En el presente documento se proporciona adicionalmente un kit para seleccionar un individuo en una población mediante la evaluación o el pronóstico del riesgo de un evento CV futuro dentro de un período de 5 años. El kit incluye los siguientes componentes: al menos uno de los biomarcadores que se exponen en la Tabla 1; al menos un reactivo de captura correspondiente, en donde cada uno de los reactivos de captura correspondientes es específico de los biomarcadores seleccionados; y un material generador de señal, siendo dicho material específico para los biomarcadores correspondientes seleccionados y/o los reactivos de captura correspondientes, en donde cada señal se activa tras la unión de cada reactivo de captura al biomarcador correspondiente.

En otro aspecto, el kit puede comprender uno o más biomarcadores seleccionados entre el grupo que consiste en: biomarcador de angiopoyetina 2; biomarcador de angiopoyetina 2 y cualquier biomarcador seleccionado entre el grupo que consiste en MMP7, CHRDL1, MATN2, PSA-ACT y cualquier combinación de los mismos; un biomarcador seleccionado entre el grupo que consiste en biomarcador GPVI, biomarcador MATN2 y cualquier combinación de los mismos; N biomarcadores seleccionados entre el grupo de biomarcadores que se expone en la Tabla 3, en donde N = 2-10; y cualquier combinación de los mismos.

Los reactivos de captura de los kits pueden ser uno o más de entre SOMAmeros, anticuerpos y sondas de ácido nucleico o una combinación de los mismos. El kit también puede incluir instrucciones o uno o más productos de software o de programa informático para clasificar al individuo de quien se obtuvo la muestra biológica, ya sea que tenga o no un mayor riesgo de un evento CV.

En otra realización, la presente divulgación comprende un clasificador que comprende los biomarcadores de la Tabla 1, col. 7, la Tabla 2 o la Tabla 3.

Breve descripción de los dibujos

La Figura 1A es un diagrama de flujo para un método de ejemplo para la predicción de un evento CV en una muestra biológica. La Figura 1B es un diagrama de flujo de un método de ejemplo para la predicción de un evento CV en una muestra biológica usando un método de clasificación de Bayes sin exposición previa.

La Figura 2A muestra un análisis del componente Principal para un subgrupo de casos con Eventos y controles sin eventos CV en 6 meses. Los casos con eventos están parcialmente separados de los controles a lo largo del eje vertical.

La Figura 2B muestra un análisis DSGA (análisis genómico específico de enfermedad, por sus siglas en inglés) para un subgrupo de casos con Eventos y controles sin eventos CV en 6 meses. Los casos con eventos están parcialmente separados de los controles a lo largo del eje horizontal.

La Figura 3 proporciona un análisis de puntuación de riesgo para la población de estudio. Esta puntuación se calculó construyendo un modelo de riesgo instantáneo proporcional simple de Cox usando el logaritmo de las mediciones de las diez proteínas de la Tabla 3. La población se dividió en quintiles basándose en esta puntuación. Las gráficas de Kaplan Meier de la Figura 3 demuestran cómo estos quintiles difieren en la proporción del individuo que experimenta un evento cardiovascular o muerte para diversos tipos de eventos.

La Figura 3A muestra las gráficas de Kaplan Meier de todas las muertes y eventos cardiovasculares de la población de estudio, con la población dividida en quintiles basándose en las puntuaciones de proteína de la Tabla 3.

La Figura 3B muestra las gráficas de Kaplan Meier para los casos de eventos CV: muertes sin clasificar, aquellas muertes sin una causa proximal conocida tal como IM o ICC (insuficiencia cardíaca congestiva), con la población dividida en quintiles basándose en las puntuaciones de proteína de la Tabla 3.

La Figura 3C muestra las gráficas de Kaplan Meier para los casos de eventos CV: ICC incidente con la población dividida en quintiles basándose en las puntuaciones de proteína de la Tabla 3.

La Figura 3D muestra las gráficas de Kaplan Meier para los casos de eventos CV: recurrencia de ICC para pacientes con ICC crónica (aquellos con un diagnóstico previo de ICC) con la población dividida en quintiles basándose en las puntuaciones de proteína de la Tabla 3.

La Figura 3E muestra las gráficas de Kaplan Meier para los casos de eventos CV: evento trombótico (IM + ictus) con la población dividida en quintiles basándose en las puntuaciones de proteína de la Tabla 3.

La Figura 3F muestra las gráficas de Kaplan Meier para los casos de eventos CV: todo ICC, con la población dividida en quintiles basándose en las puntuaciones de proteína de la Tabla 3.

La Figura 4 ilustra un sistema informático de ejemplo para su uso con diversos métodos implementados por ordenador que se describen en el presente documento.

La Figura 5 es un diagrama de flujo de un método de indicación y evaluación del riesgo de un evento CV de acuerdo con una realización. La Figura 6 es un diagrama de flujo para un método de evaluación del riesgo de un evento CV de acuerdo con una realización.

La Figura 7 ilustra un ensayo de aptámero de ejemplo que puede usarse para detectar uno o más biomarcadores de eventos CV en una muestra biológica.

La Figura 8 muestra las gráficas de Kaplan Meier basadas en GPVI, una de las diez proteínas de la Tabla 3, lo que demuestra que esta proteína distingue entre eventos trombóticos y de ICC. La población se divide en cuartiles de GPVI. La Figura 8A muestra que el cuartil más alto de GPVI es pronóstico para eventos cardiovasculares trombóticos. La Figura 8B muestra que los cuartiles de GPVI tienen poca o ninguna capacidad discriminadora para pronosticar eventos de ICC.

La Figura 9 muestra las gráficas de Kaplan Meier basadas en MATN2, una de las diez proteínas de la Tabla 3, lo que demuestra que esta proteína distingue entre eventos trombóticos y de ICC. La población se divide en cuartiles de MATN2.

La Figura 9A muestra que los cuartiles de MATN2 no son pronósticos para eventos cardiovasculares trombóticos.

La Figura 9B muestra que los individuos del cuartil más alto de MATN2 tienen una tasa más alta de eventos de ICC.

La Figura 10 muestra las gráficas de Kaplan Meier de todos los 538 sujetos que toman medicamentos con estatinas que muestran que aquellos individuos en el 4to cuartil de la distribución de la población para la angiopoyetina-2 padecen eventos cardiovasculares a una tasa mayor en comparación con aquellos que no están en el 4to cuartil para la angiopoyetina-2. Por tanto, a pesar de los efectos del tratamiento con estatinas, la angiopoyetina-2 es un biomarcador útil del riesgo de eventos CV.

La Figura 11 muestra las gráficas de Kaplan Meier de los 538 sujetos que toman medicamentos con estatinas que demuestran que CHRDL1 se asocia a la supervivencia sin eventos cardiovasculares en individuos tratados con estatinas. Por tanto, a pesar de los efectos del tratamiento con estatinas, CHRDL1 es un biomarcador útil del riesgo de eventos CV.

Descripción detallada

Ahora se hará referencia en detalle a realizaciones representativas de la divulgación. Aunque la divulgación se describirá junto con las realizaciones enumeradas, se comprenderá que no se pretende limitar la invención a esas realizaciones. Por el contrario, se pretende que la invención incluya todas las alternativas, modificaciones y equivalentes que puedan incluirse dentro del alcance de la presente invención como se define en las reivindicaciones.

El experto en la materia reconocerá muchos métodos y materiales similares o equivalentes a los que se describen en el presente documento, que podrían usarse y están dentro del alcance de la puesta en práctica de la presente invención. La presente invención no se limita de ninguna manera a los métodos y materiales que se describen.

A menos que se defina lo contrario, los términos técnicos y científicos utilizados en el presente documento tienen el mismo significado que el que entiende normalmente un experto habitual en la materia a la que pertenece la presente invención. Aunque puede usarse cualquier método, dispositivo y material similar o equivalente a los que se describen en el presente documento en la puesta en práctica o ensayo de la invención, se describen ahora los métodos, dispositivos y materiales preferidos.

Todas las publicaciones, documentos de patente publicados y solicitudes de patente que se citan en la presente solicitud son indicativos del nivel de habilidad en la técnica o técnicas a las que pertenecen la solicitud.

Como se usa en la presente solicitud, incluyendo las reivindicaciones adjuntas, las formas en singular "un", "una", "el" y "la" incluyen referencias en plural, a menos que el contenido imponga claramente lo contrario, y se usan indistintamente con "al menos uno" y "uno o más". Por tanto, la referencia a "un SOMAmero" incluye mezclas de SOMAmeros, la referencia a "una sonda" incluye mezclas de sondas y similares.

Como se usa en el presente documento, el término "aproximadamente" representa una modificación o variación insignificante del valor numérico de manera que la función básica del elemento con el que se relaciona el valor numérico no cambia.

Como se usa en el presente documento, las expresiones "comprende", "que comprende", "incluye", "que incluye", "contiene", "que contiene", y cualquier variación de las mismas, tienen por objeto abarcar una inclusión no exclusiva, tal como un proceso, método, producto definido por su proceso o composición de la materia que comprende, incluye o contiene un elemento o lista de elementos no incluye solo esos elementos, sino que puede incluir otros elementos no enumerados expresamente o inherentes a dicho proceso, método, producto definido por su proceso o composición de la materia.

La presente solicitud incluye biomarcadores, métodos, dispositivos, reactivos, sistemas y kits para la predicción del riesgo de eventos CV dentro de un período de tiempo definido, tal como 5 años.

"Evento cardiovascular" significa una insuficiencia o disfunción de cualquier parte del sistema circulatorio. En una realización, "Evento cardiovascular" significa ictus, ataque isquémico transitorio (AIT), infarto de miocardio (IM), muerte súbita atribuible a la disfunción del sistema circulatorio y/o insuficiencia cardíaca. En otra realización, "Evento cardiovascular" se refiere a cualquiera de las disfunciones anteriores y/o angina inestable, necesidad de stent o angioplastia, o similares.

Los eventos cardiovasculares incluyen "insuficiencia cardíaca congestiva" o "ICC" y "eventos trombóticos". Los eventos trombóticos incluyen IM, ataques isquémicos transitorios (AIT), ictus, síndrome coronario agudo y la necesidad de revascularización coronaria.

En un aspecto, se proporcionan uno o más biomarcadores para su uso ya sea solo o en diversas combinaciones para evaluar el riesgo de un evento CV futuro dentro de un período de 5 años con eventos CV definidos como infarto de miocardio, ictus, muerte e insuficiencia cardíaca congestiva. Los eventos trombóticos (Figura 3e) consisten en infarto de miocardio e ictus combinados. Como se describe en detalle a continuación, las realizaciones de ejemplo incluyen los biomarcadores que se proporcionan en la Tabla 1, Col. 7, que se identificaron usando un ensayo basado en SOMAmero múltiple que se describe en general en el Ejemplo 1 y más específicamente en el Ejemplo 2.

La Tabla 1, Col. 7, expone los resultados obtenidos del análisis de cientos de muestras de sangre individuales de pacientes que han tenido un evento CV dentro de un período de 6 meses a 10 años (Evento Positivo) después de la extracción de sangre inicial (punto temporal 1) y cientos de muestras de sangre individual equivalentes de individuos que no tuvieron un evento CV dentro de ese marco temporal (Evento Negativo). Los biomarcadores potenciales se midieron en muestras individuales en lugar de agrupar las muestras de sangre de Evento Positivo y Evento Negativo; Esto permitió una mejor comprensión de las variaciones individuales y grupales en los fenotipos asociados a la presencia y ausencia de un evento CV. Puesto que se realizaron más de 1000 mediciones de proteínas en cada muestra y se midieron individualmente varios cientos de muestras de cada una de las poblaciones de Evento Positivo y Evento Negativo, los biomarcadores publicados en la Tabla 1, Col. 7 fueron resultado de un análisis de un conjunto inusualmente grande de datos. Las mediciones se analizaron usando los métodos que se describen en la sección "Clasificación de biomarcadores y cálculo de puntuaciones de riesgo" en el presente documento. La Tabla 1, Col. 7 enumera los 155 biomarcadores que se encuentran útiles para estratificar la población de individuos de acuerdo con su propensión a presentar un evento CV futuro en el período de 0-5 años después de la extracción de la muestra de sangre. Las curvas de Kaplan-Meier en las Figuras 3A-3F muestran una fuerte dependencia del riesgo de eventos del quintil de una puntuación determinada por un pequeño subconjunto de dichos biomarcadores, como se enumera en la Tabla 3.

Aunque algunos de los biomarcadores de eventos CV que se describen son útiles solos para evaluar el riesgo de un evento CV, también se describen métodos en el presente documento para la agrupación de múltiples subconjuntos de los biomarcadores de eventos CV, donde cada selección de agrupación o subconjunto es útil como panel de tres o más biomarcadores, denominados indistintamente en el presente documento "panel de biomarcadores" y un panel. Por tanto, diversas realizaciones de la presente solicitud proporcionan combinaciones que comprenden N biomarcadores, en donde N es al menos dos biomarcadores. En otras realizaciones, N se selecciona entre 2-155 biomarcadores.

En otras realizaciones más, N se selecciona para que sea cualquier número entre 2-7, 2-10, 2-15, 2-20, 2-25, 2-30, 2-35, 2-40, 2-45, 2-50, 2-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 2-155. En otras realizaciones, N se selecciona para que sea cualquier número entre 3-7, 3-10, 3-15, 3-20, 3-25, 3-30, 3-35, 3-40, 3-45, 3-50, 3-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 3-155. En otras realizaciones, N se selecciona para que sea cualquier número entre 4-7, 4-10, 4-15, 4-20, 4-25, 4-30, 4-35, 4-40, 4-45, 4-50, 4-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 4-155. En otras realizaciones, N se selecciona para que sea cualquier número entre 5-7, 5-10, 5-15, 5-20, 5-25, 5-30, 5-35, 5-40, 5-45, 5-50, 5-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 5-155. En otras realizaciones, N se selecciona para que sea cualquier número entre 6-10, 6-15, 6-20, 6-25, 6-30, 6-35, 6-40, 6-45, 6-50, 6-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 6-155. En otras realizaciones, N se selecciona para que sea cualquier número entre 7-10, 7-15, 7-20, 7-25, 7-30, 7-35, 7-40, 7-45, 7-50, 7-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 7-155. En otras realizaciones, N se selecciona para que sea cualquier número entre 8-10, 8-15, 8-20, 8-25, 8-30, 8-35, 8-40, 8-45, 8-50, 8-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 8-155. En otras

realizaciones, N se selecciona para que sea cualquier número entre 9-15, 9-20, 9-25, 9-30, 9-35, 9-40, 9-45, 9-50, 9-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 9-155. En otras realizaciones, N se selecciona para que sea cualquier número entre 10-15, 10-20, 10-25, 10-30, 10-35, 10-40, 10-45, 10-50, 10-55 o en incrementos sucesivos de 5 para el límite superior del intervalo, hasta e incluyendo 10-155. Se apreciará que N puede seleccionarse para abarcar intervalos similares, pero de orden superior.

En una realización, el número de biomarcadores útil para un subconjunto o panel de biomarcadores se basa en el valor de sensibilidad y especificidad para la combinación particular de valores de biomarcadores. Los términos "sensibilidad" y "especificidad" se usan en el presente documento con respecto a la capacidad de clasificar correctamente a un individuo, basándose en uno o más valores de biomarcadores detectados en su muestra biológica, como que tienen un mayor riesgo de tener un evento CV dentro de 5 años o de no tener un mayor riesgo de tener un evento CV dentro del mismo período de tiempo. La "sensibilidad" indica el rendimiento del biomarcador o biomarcadores con respecto a clasificar correctamente a los individuos que tienen un mayor riesgo de un evento CV. La "especificidad" indica el rendimiento del biomarcador o biomarcadores con respecto a clasificar correctamente a los individuos que no tienen un mayor riesgo de un evento CV. Por ejemplo, una especificidad del 85 % y una sensibilidad del 90 % para un panel de marcadores utilizados para someter a ensayo un conjunto de muestras de Evento Negativo y muestras de Evento Positivo indican que el panel clasificó correctamente el 85 % de las muestras de control como muestras de Evento Negativo y que el panel clasificó correctamente el 90 % de las muestras de Evento Positivo como muestras de Evento Positivo.

En un método alternativo, las puntuaciones pueden publicarse en un intervalo continuo, con un umbral de riesgo alto, intermedio o bajo de un evento CV, con umbrales determinados basándose en hallazgos clínicos.

Los biomarcadores de riesgo de eventos CV identificados en el presente documento representan un número extremadamente grande de opciones para subconjuntos o paneles de biomarcadores que pueden usarse para predecir el riesgo de un evento CV. La selección del número deseado de dichos biomarcadores depende de la combinación específica de biomarcadores elegidos. Es importante recordar que los paneles de biomarcadores para predecir el riesgo de eventos CV también pueden incluir biomarcadores que no se encuentran en la Tabla 1, Col. 7, y que la inclusión de biomarcadores adicionales que no se encuentran en la Tabla 1, Col. 7 puede reducir el número de biomarcadores en el subconjunto o panel particular que se selecciona entre la Tabla 1, Col. 7. El número de biomarcadores de la Tabla 1, Col. 7 utilizados en un subconjunto o panel también puede reducirse si se usa información biomédica adicional junto con los valores de biomarcadores para establecer valores umbral aceptables para un ensayo dado.

Otro factor que puede afectar a la cantidad de biomarcadores que se usarán en un subconjunto o panel de biomarcadores son los procedimientos utilizados para obtener muestras biológicas de individuos que se están evaluando para determinar el riesgo de un evento CV. En un entorno de obtención de muestras cuidadosamente controlado, el número de biomarcadores necesarios para cumplir con la sensibilidad y especificidad deseadas y/o los valores umbral será más bajo que en una situación en la que puede haber más variación en la recogida, manipulación y almacenamiento de muestras.

Un aspecto de la presente solicitud puede describirse en general con referencia a las Figuras 1A y 1B. Se obtiene una muestra biológica de un individuo o individuos de interés. Después, la muestra biológica se analiza para detectar la presencia de uno o más (N) biomarcadores de interés y para determinar un valor de biomarcador para cada uno de dichos N biomarcadores (referido en la Figura 1B como marcador RFU). Una vez que se ha detectado un biomarcador y se ha asignado un valor de biomarcador, cada marcador se puntúa o clasifica como se describe en detalle en el presente documento. Después, las puntuaciones de los marcadores se combinan para proporcionar una puntuación de diagnóstico total, lo que indica la probabilidad de que el individuo del que se obtuvo la muestra tenga un riesgo alto, medio o bajo de un evento CV, en particular cuando se publica en un intervalo continuo.

"Muestra biológica", "muestra" y "muestra de ensayo" se usan indistintamente en el presente documento para referirse a cualquier material, líquido biológico, tejido o célula obtenido o derivado de otro modo de un individuo. Esto incluye sangre (incluyendo sangre completa, leucocitos, células mononucleares de sangre periférica, capa leucocitaria, plasma y suero), manchas de sangre seca (por ejemplo, obtenida de bebés), esputo, lágrimas, moco, lavados nasales, aspirados nasales, aire espirado, orina, semen, saliva, lavados peritoneales, ascitis, líquido cístico, líquido meníngeo, líquido amniótico, líquido glandular, líquido pancreático, líquido linfático, líquido pleural, aspirado del pezón, aspirado bronquial, cepillado bronquial, líquido sinovial, aspirado articular, secreciones orgánicas, células, un extracto celular y líquido cerebroespinal. Esto también incluye fracciones separadas experimentalmente de todos los anteriores. Por ejemplo, una muestra de sangre puede fraccionarse en suero, plasma o en fracciones que contienen tipos particulares de células sanguíneas, tales como glóbulos rojos o glóbulos blancos (leucocitos). Si se desea, una muestra puede ser una combinación de muestras de un individuo, tal como una combinación de una muestra de tejido y de líquido. La expresión "muestra biológica" también incluye materiales que contienen material sólido homogeneizado, tal como a partir de una muestra de heces, una muestra tisular o una biopsia tisular, por ejemplo. La expresión "muestra biológica" también incluye materiales derivados de un cultivo tisular o un cultivo celular. Puede emplearse cualquier método adecuado para obtener una muestra biológica; los métodos de ejemplo incluyen, por ejemplo, flebotomía, hisopado (por ejemplo, hisopado bucal) y un procedimiento de biopsia de aspirado con agua fija. Los tejidos de ejemplo

susceptibles de aspiración con agua fina incluyen ganglio linfático, pulmón, lavados pulmonares, LBA (lavado broncoalveolar), tiroides, mama, páncreas e hígado. También pueden recogerse muestras, por ejemplo, mediante microdissección (por ejemplo, microdissección de captura láser (MCL) o microdissección láser (MDL)), lavado de vejiga, frotis (por ejemplo, un frotis de PAP) o lavado ductal. Una "muestra biológica" obtenida o derivada de un individuo incluye cualquier muestra de este tipo que ha sido procesada de cualquier manera adecuada después de haberse obtenido del individuo.

Adicionalmente, debe tenerse en cuenta que una muestra biológica puede derivar de tomar muestras biológicas de varios individuos y agruparlas o de agrupar una alícuota de la muestra biológica de cada individuo. La muestra agrupada puede tratarse como una muestra de un único individuo y si se establece un riesgo mayor o menor de un evento CV en la muestra agrupada, entonces cada muestra biológica individual puede volverse a someter a ensayo para determinar qué individuo o individuos tienen un aumento o disminución del riesgo de un evento CV.

Como se ha mencionado anteriormente, la muestra biológica puede ser orina. Las muestras de orina proporcionan determinadas ventajas sobre las muestras de sangre o de suero. La recogida de muestras de sangre o plasma a través de punción venosa es más compleja de lo deseable, puede entregar volúmenes variables, puede ser preocupante para el paciente e implica cierto riesgo (pequeño) de infección. Además, la flebotomía requiere personal cualificado. La simplicidad de la recogida de muestras de orina puede conducir a una aplicación más generalizada de los presentes métodos.

Con el fin de determinar la idoneidad del uso de la orina como muestra, se evaluó la calidad y la cantidad de cada proteína en dichas muestras sanas de sujetos sanos, y esta información se combinó con la calidad de los biomarcadores de las Tablas 1-3 en el pronóstico de riesgo CV. Como la orina es un ultrafiltrado de plasma, la cantidad de una proteína específica excretada en la orina es proporcional a la concentración de proteína en la sangre. Si hay disponibles biomarcadores de calidad de cualquiera de las Tablas 1-3 en cantidad suficiente en la orina, entonces son adecuados para su uso en un método para seleccionar un individuo para la evaluación del riesgo de un evento CV. Los biomarcadores predictivos de un evento CV que se han descubierto en la orina incluyen ESAM, MMP7 y GP6, que muestran una señal fuerte. Estos biomarcadores son más pequeños y el riñón los filtra libremente en la orina. Además, se ha descubierto que el PSA-ACT y el plasminógeno muestran variabilidad entre individuos en la orina. Esto indica que la cuantificación de estos biomarcadores en orina también puede ser útil en el método de selección de un individuo para determinar el riesgo de un evento CV. Por tanto, estas cinco proteínas proporcionan un ensayo simple a base de orina que ha de usarse en los presentes métodos de selección de individuos para determinar el riesgo de un evento CV.

Para los fines de la presente memoria descriptiva, la frase "datos atribuidos a una muestra biológica de un individuo" tiene por objeto significar que los datos en alguna forma derivan de, o se generaron usando, la muestra biológica del individuo. Los datos pueden haber sido reformateados, revisados o alterados matemáticamente hasta cierto punto después de haber sido generados, tal como, por ejemplo, mediante la conversión de unidades en un sistema de medición en unidades en otro sistema de medición; pero, se entiende que los datos derivan de, o se generaron usando, la muestra biológica.

"Diana", "molécula diana" y "analito" se usan indistintamente en el presente documento para referirse a cualquier molécula de interés que pueda estar presente en una muestra biológica. Una "molécula de interés" incluye cualquier variación menor de una molécula particular, tal como, en el caso de una proteína, por ejemplo, variaciones menores en la secuencia de aminoácidos, formación de enlaces disulfuro, glucosilación, lipidación, acetilación, fosforilación o cualquier otra manipulación o modificación, tal como conjugación con un componente marcador, que no altera sustancialmente la identidad de la molécula. Una "molécula diana", "diana" o "analito" es un conjunto de copias de un tipo o especie de molécula o estructura multimolecular. "Moléculas diana", "dianas" y "analitos" se refieren a más de un conjunto de dichas moléculas. Las moléculas diana de ejemplo incluyen proteínas, polipéptidos, ácidos nucleicos, hidratos de carbono, lípidos, polisacáridos, glucoproteínas, hormonas, receptores, antígenos, anticuerpos, afficuerpos, miméticos de anticuerpos, virus, patógenos, sustancias tóxicas, sustratos, metabolitos, análogos del estado de transición, cofactores, inhibidores, fármacos, colorantes, nutrientes, factores de crecimiento, células, tejidos y cualquier fragmento o porción de cualquiera de los anteriores.

Como se usa en el presente documento, "polipéptido", "péptido" y "proteína" se usan indistintamente en el presente documento para referirse a polímeros de aminoácidos de cualquier longitud. El polímero puede ser lineal o ramificado, puede comprender aminoácidos modificados y puede estar interrumpido por no aminoácidos. Los términos también abarcan un polímero de aminoácidos que se ha modificado de manera natural o por intervención; por ejemplo, formación de enlaces disulfuro, glucosilación, lipidación, acetilación, fosforilación o cualquier otra manipulación o modificación, tal como conjugación con un componente marcador. En la definición también se incluyen, por ejemplo, polipéptidos que contienen uno o más análogos de un aminoácido (incluyendo, por ejemplo, aminoácidos no naturales, etc.), así como otras modificaciones conocidas en la técnica. Los polipéptidos pueden ser cadenas únicas o cadenas asociadas. En la definición también se incluyen las preproteínas y las proteínas maduras intactas; péptidos o polipéptidos derivados de una proteína madura; fragmentos de una proteína; variantes de corte y empalme; formas recombinantes de una proteína; variantes proteínicas con modificaciones, supresiones o sustituciones de aminoácidos; digestiones; y modificaciones postraduccionales, tales como glucosilación, acetilación, fosforilación, y similares.

Como se usa en el presente documento, "marcador" y "biomarcador" se usan indistintamente para referirse a una molécula diana que indica o es un signo de un proceso normal o anormal en un individuo o de una enfermedad u otra afección en un individuo. Más específicamente, un "marcador" o "biomarcador" es un parámetro anatómico, fisiológico, bioquímico o molecular asociado a la presencia de un estado o proceso fisiológico específico, ya sea normal o anormal y, si es anormal, ya sea crónico o agudo. Los biomarcadores pueden detectarse y medirse mediante diversos métodos incluyendo ensayos de laboratorio y formación de imágenes médicas. Cuando un biomarcador es una proteína, también es posible usar la expresión del gen correspondiente como una medida sustituta de la cantidad o presencia o ausencia del biomarcador proteínico correspondiente en una muestra biológica o estado de metilación del gen que codifica el biomarcador o las proteínas que controlan la expresión del biomarcador.

Como se usa en el presente documento, "valor de biomarcador", "valor", "nivel de biomarcador" y "nivel" se usan indistintamente para referirse a una medición que se realiza usando cualquier método analítico para detectar el biomarcador en una muestra biológica y que indica la presencia, ausencia, cantidad o concentración absoluta, cantidad o concentración relativa, título, un nivel, un nivel de expresión, una relación de niveles medidos, o similares, de, para, o correspondientes al biomarcador en la muestra biológica. La naturaleza exacta del "valor" o "nivel" depende del diseño específico y los componentes del método analítico particular empleado para detectar el biomarcador.

Cuando un biomarcador indica o es un signo de un proceso anormal o una enfermedad u otra afección en un individuo, ese biomarcador generalmente se describe como sobreexpresado o subexpresado en comparación con un nivel o valor de expresión del biomarcador que indica o es un signo de un proceso normal o la ausencia de una enfermedad u otra afección en un individuo. "Regulación positiva", "regulado positivamente", "sobrexpresión", "sobrexpresado" y cualquier variación de los mismos se usan indistintamente para referirse a un valor o nivel de un biomarcador en una muestra biológica que es superior a un valor o nivel (o intervalo de valores o niveles) del biomarcador que normalmente se detecta en muestras biológicas similares de individuos sanos o normales. Los términos también pueden referirse a un valor o nivel de un biomarcador en una muestra biológica que es superior a un valor o nivel (o intervalo de valores o niveles) del biomarcador que puede detectarse en una etapa diferente de una enfermedad particular.

"Regulación negativa", "regulado negativamente", "subexpresión", "subexpresado" y cualquier variación de los mismos se usan indistintamente para referirse a un valor o nivel de un biomarcador en una muestra biológica que es inferior a un valor o nivel (o intervalo de valores o niveles) del biomarcador que normalmente se detecta en muestras biológicas similares de individuos sanos o normales. Los términos también pueden referirse a un valor o nivel de un biomarcador en una muestra biológica que es inferior a un valor o nivel (o intervalo de valores o niveles) del biomarcador que puede detectarse en una etapa diferente de una enfermedad particular.

Adicionalmente, un biomarcador que está sobreexpresado o subexpresado también puede denominarse "expresado diferencialmente" o que tiene un "nivel diferencial" o "valor diferencial" en comparación con un nivel o valor de expresión "normal" del biomarcador que indica o es un signo de un proceso normal o la ausencia de una enfermedad u otra afección en un individuo. Por tanto, la "expresión diferencial" de un biomarcador también puede denominarse una variación de un nivel de expresión "normal" del biomarcador.

La expresión "expresión génica diferencial" y "expresión diferencial" se usan indistintamente para referirse a un gen (o su producto de expresión proteínica correspondiente) cuya expresión se activa a un nivel más alto o más bajo en un sujeto que padece una enfermedad o afección específica, con respecto a su expresión en un sujeto normal o de control. Los términos también incluyen genes (o los productos de expresión proteínica correspondientes) cuya expresión se activa a un nivel más alto o más bajo en diferentes etapas de la misma enfermedad o afección. También se entiende que un gen expresado diferencialmente puede estar activado o inhibido a nivel de ácido nucleico o a nivel de proteína, o puede someterse a corte y empalme alternativo para dar como resultado un producto polipeptídico diferente. Dichas diferencias pueden evidenciarse por diversos cambios incluyendo niveles de ARNm, expresión superficial, secreción u otra separación de un polipéptido. La expresión génica diferencial puede incluir una comparación de la expresión entre dos o más genes o sus productos génicos; o una comparación de las relaciones de la expresión entre dos o más genes o sus productos génicos; o incluso una comparación de dos productos procesados de manera diferente del mismo gen, que difieren entre sujetos normales y sujetos que padecen una enfermedad; o entre diversas etapas de la misma enfermedad. La expresión diferencial incluye diferencias, tanto cuantitativas como cualitativas, en el patrón de expresión temporal o celular en un gen o sus productos de expresión entre, por ejemplo, células normales y enfermas o entre células que han experimentado diferentes eventos de enfermedad o etapas de enfermedad.

Como se usa en el presente documento, "individuo" se refiere a un sujeto de ensayo o paciente. El individuo puede ser un mamífero o no mamífero. En diversas realizaciones, el individuo es un mamífero. Un individuo mamífero puede ser un ser humano o no humano. En diversas realizaciones, el individuo es un ser humano. Un individuo sano o normal es un individuo en el que la enfermedad o afección de interés (incluyendo, por ejemplo, eventos cardiovasculares tales como infarto de miocardio, ictus e insuficiencia cardíaca congestiva) no es detectable mediante métodos de diagnóstico convencionales.

"Diagnóstico", "diagnosticar", "diagnosis" y variaciones de los mismos se refieren a la detección, determinación o

reconocimiento de un estado o condición de salud de un individuo basándose en uno o más signos, síntomas, datos u otra información pertinente a ese individuo. El estado de salud de un individuo puede diagnosticarse como sano/normal (es decir, un diagnóstico de la ausencia de una enfermedad o afección) o puede diagnosticarse como enfermo/anormal (es decir, un diagnóstico de la presencia o la evaluación de las características, de una enfermedad o afección). Los términos "diagnosticar", "diagnostico", "diagnosis", etc., abarcan, con respecto a una enfermedad o afección particular, la detección inicial de la enfermedad; la caracterización o clasificación de la enfermedad; la detección de la progresión, remisión o recurrencia de la enfermedad; y la detección de la respuesta a la enfermedad después de la administración de un tratamiento o terapia al individuo. La predicción del riesgo de un evento CV incluye distinguir individuos que tienen un mayor riesgo de un evento CV de individuos que no lo tienen.

"Pronóstico", "pronosticar", "prognosis" y variaciones de los mismos se refieren a la predicción de un curso futuro de una enfermedad o afección en un individuo que tiene la enfermedad o afección (por ejemplo, predicción de la supervivencia del paciente) y dichos términos abarcan la evaluación de la respuesta a una enfermedad o afección después de la administración de un tratamiento o terapia al individuo.

"Evaluar", "evaluación", "evaluar" y variaciones de los mismos abarcan tanto "diagnosticar" como "pronosticar" y también abarcan determinaciones o predicciones sobre el curso futuro de una enfermedad o afección en un individuo que no tiene la enfermedad, así como determinaciones o predicciones sobre el riesgo de que una enfermedad o afección reaparezca en un individuo que aparentemente se ha curado de la enfermedad o en el que se ha resuelto la afección. El término "evaluar" también abarca la evaluación de la respuesta de un individuo a una terapia, tal como, por ejemplo, predecir si es probable que un individuo responda favorablemente a un agente terapéutico o si es improbable que responda a un agente terapéutico (o experimentará un efecto tóxico u otros efectos secundarios no deseados, por ejemplo), seleccionar un agente terapéutico para la administración a un individuo o controlar o determinar la respuesta de un individuo a una terapia que se ha administrado al individuo. Por tanto, "evaluar" el riesgo de un evento CV puede incluir, por ejemplo, cualquiera de los siguientes: predecir el riesgo futuro de un evento CV en un individuo; predecir el riesgo de un evento CV en un individuo que aparentemente no tiene problemas CV; o determinar o predecir la respuesta de un individuo a un tratamiento CV o seleccionar un tratamiento CV para administrar a un individuo basándose en una determinación de los valores de biomarcadores derivados de la muestra biológica del individuo. La evaluación del riesgo de un evento CV puede incluir realizaciones tales como la evaluación del riesgo de un evento CV en una escala continua o la clasificación del riesgo de un evento CV en clasificaciones crecientes. La clasificación del riesgo incluye, por ejemplo, la clasificación en dos o más clasificaciones, tales como "Sin riesgo elevado de un evento CV" y "Riesgo elevado de un evento CV". La evaluación del riesgo de un evento CV es durante un período de tiempo definido; dicho período de tiempo puede ser, por ejemplo, 5 años.

Como se usa en el presente documento, "información biomédica adicional" se refiere a una o más evaluaciones de un individuo, distintas del uso de cualquiera de los biomarcadores que se describen en el presente documento, que se asocian al riesgo CV o, más específicamente, el riesgo de evento CV. La "información biomédica adicional" incluye cualquiera de los siguientes: descriptores físicos de un individuo, incluyendo la altura y/o el peso de un individuo; la edad de un individuo; el género de un individuo; el cambio de peso; el origen étnico de un individuo; el historial ocupacional; los antecedentes familiares de enfermedades cardiovasculares (u otros trastornos del sistema circulatorio); la presencia de uno o más marcadores genéticos que se correlacionan con un mayor riesgo de enfermedades cardiovasculares (u otros trastornos del sistema circulatorio) en las alteraciones del individuo o de un miembro de la familia en el espesor de la íntima carótida; los síntomas clínicos tales como dolor en el pecho, el aumento o la pérdida de peso en los valores de expresión génica; descriptores físicos de un individuo, incluyendo los descriptores físicos observados mediante formación de imágenes radiológicas; tabaquismo; historial de consumo de alcohol; historial ocupacional; hábitos alimenticios: consumo de sal, grasas saturadas y colesterol; consumo de cafeína; e información de imágenes tales como electrocardiograma, ecocardiografía, ultrasonido carotídeo para determinar el espesor íntima-media, dilatación mediada por flujo, velocidad de la onda del pulso, índice tobillo-brazo, ecocardiografía con estrés, imágenes de perfusión miocárdica, calcio coronario por TC, angiografía por TC de alta resolución, resonancia magnética y otras modalidades de formación de imágenes; y los medicamentos del individuo. Los ensayos de los niveles de biomarcadores en combinación con una evaluación de cualquier información biomédica adicional, incluyendo otros ensayos de laboratorio (por ejemplo, HDL, ensayos de LDL, niveles de CRP, ensayos de Nt-proBNP, ensayos de albúmina sérica, ensayos de creatina), pueden, por ejemplo, mejorar la sensibilidad, la especificidad y/o el ABC para la predicción de eventos CV en comparación con el ensayo de biomarcadores solo o la evaluación de cualquier elemento particular de información biomédica adicional sola (por ejemplo, formación de imágenes de espesor de la íntima carótida sola). Puede obtenerse información biomédica adicional de un individuo usando técnicas de rutina conocidas en la técnica, tal como del propio individuo mediante el uso de un cuestionario de rutina del paciente o un cuestionario de historial de salud, etc. o de un médico, etc. Los ensayos de niveles de biomarcadores en combinación con una evaluación de cualquier información biomédica adicional puede, por ejemplo, mejorar la sensibilidad, la especificidad y/o los umbrales para la predicción de eventos CV (u otros usos relacionados con enfermedades cardiovasculares) en comparación con el ensayo de biomarcadores solo o la evaluación de cualquier elemento particular de información biomédica adicional sola (por ejemplo, formación de imágenes por tomografía computarizada sola).

Como se usa en el presente documento, "detectar" o "determinar" con respecto a un valor de biomarcador incluye el uso tanto del instrumento necesario para observar y registrar una señal correspondiente a un valor de biomarcador

como del material o materiales necesarios para generar esa señal. En diversas realizaciones, el valor de biomarcador se detecta usando cualquier método adecuado, incluyendo fluorescencia, quimioluminiscencia, resonancia de plasmón superficial, ondas acústicas superficiales, espectrometría de masas, espectroscopía infrarroja, espectroscopía Raman, microscopía de fuerza atómica, microscopía de túnel de barrido, métodos de detección electroquímica, resonancia magnética nuclear, puntos cuánticos y similares.

"Soporte sólido" se refiere en el presente documento a cualquier sustrato que tenga una superficie a la que las moléculas puedan unirse, directa o indirectamente, a través de enlaces covalentes o no covalentes. Un "soporte sólido" puede tener diversos formatos físicos, que pueden incluir, por ejemplo, una membrana; un chip (por ejemplo, un chip de proteína); un portaobjetos (por ejemplo, un portaobjetos o cubreobjetos de vidrio); una columna; una partícula hueca, sólida, semisólida, que contenga poros o cavidades, tal como, por ejemplo, una perla; un gel; una fibra, incluyendo un material de fibra óptica; una matriz; y un receptáculo de muestra. Los receptáculos de muestra de ejemplo incluyen pocillos, tubos, capilares, viales de muestra y cualquier otro recipiente, surco o hendidura capaz de contener una muestra. Un receptáculo de muestra puede estar contenido en una plataforma de múltiples muestras, tal como una placa de microtitulación, un portaobjetos, un dispositivo de microfluidos y similares. Un soporte puede estar compuesto por un material natural o sintético, un material orgánico o inorgánico. La composición del soporte sólido sobre el que se unen los reactivos de captura en general depende del método de unión (por ejemplo, unión covalente). Otros receptáculos de ejemplo incluyen microgotas y emulsiones de aceite/acuosas controladas por microfluidos o a granel dentro de las cuales pueden realizarse ensayos y manipulaciones relacionadas. Los soportes sólidos adecuados incluyen, por ejemplo, plásticos, resinas, polisacáridos, sílice o materiales a base de sílice, vidrio funcionalizado, silicio modificado, carbono, metales, vidrios inorgánicos, membranas, nylon, fibras naturales (tales como, por ejemplo, seda, lana y algodón), polímeros y similares. El material que compone el soporte sólido puede incluir grupos reactivos tales como, por ejemplo, grupos carboxi, amino o hidroxilo, que se usan para la unión de los reactivos de captura. Los soportes sólidos poliméricos pueden incluir, por ejemplo, poliestireno, tetraftalato de polietilenglicol, acetato de polivinilo, cloruro de polivinilo, polivinilpirrolidona, poliacrilonitrilo, metacrilato de polimetilo, politetrafluoroetileno, caucho de butilo, caucho de estirenobutadieno, caucho natural, polietileno, polipropileno, (poli)tetrafluoroetileno, fluoruro de (poli)vinilideno, policarbonato y polimetilpenteno. Las partículas de soporte sólido adecuadas que pueden usarse incluyen, por ejemplo, partículas codificadas, tales como partículas codificadas de tipo Lumindex®, partículas magnéticas y partículas de vidrio.

Usos de ejemplo de biomarcadores

En diversas realizaciones de ejemplo, se proporcionan métodos para evaluar el riesgo de un evento CV en un individuo mediante la detección de uno o más valores de biomarcadores correspondientes a uno o más biomarcadores que están presentes en la circulación de un individuo, tal como en suero o plasma, mediante cualquier número de métodos analíticos, incluyendo cualquiera de los métodos analíticos que se describen en el presente documento. Estos biomarcadores, por ejemplo, se expresan diferencialmente en individuos con mayor riesgo de un evento CV en comparación con individuos sin mayor riesgo de un evento CV. La detección de la expresión diferencial de un biomarcador en un individuo puede usarse, por ejemplo, para permitir la predicción del riesgo de un evento CV en un marco temporal de 5 años.

Además de someter a ensayo los niveles de biomarcadores como ensayo de diagnóstico independiente, los niveles de biomarcadores también pueden hacerse junto con la determinación de SNP u otras lesiones genéticas o la variabilidad que son indicativas de un mayor riesgo de susceptibilidad a enfermedades o afecciones. (Véase, por ejemplo, Amos et al., *Nature Genetics* 40, 616-622 (2009)).

Además de someter a ensayo los niveles de biomarcadores como ensayo de diagnóstico independiente, los niveles de biomarcadores también pueden usarse junto con la detección radiológica. Los niveles de biomarcadores también pueden usarse junto con síntomas relevantes o ensayos genéticos. La detección de cualquiera de los biomarcadores que se describen en el presente documento puede ser útil después de que se haya evaluado el riesgo de evento CV para guiar la atención clínica adecuada del individuo, incluyendo el aumento a niveles de atención más agresivos en individuos de alto riesgo después de que se haya determinado el riesgo de evento CV. Además de someter a ensayo los niveles de biomarcadores junto con síntomas relevantes o factores de riesgo, también puede evaluarse información con respecto a los biomarcadores junto con otros tipos de datos, en particular datos que indican el riesgo de un individuo para eventos cardiovasculares (por ejemplo, historial clínico del paciente, síntomas, antecedentes familiares de enfermedades cardiovasculares, antecedentes de tabaquismo o consumo de alcohol, factores de riesgo tales como la presencia de un marcador o marcadores genéticos y/o el estado de otros biomarcadores, etc.). Estos diversos datos pueden evaluarse mediante métodos automatizados, tales como un programa/software informático, que puede incorporarse en un ordenador u otro aparato/dispositivo.

Además de someter a ensayo los niveles de biomarcadores junto con el barrido radiológico en individuos de alto riesgo (por ejemplo, evaluar los niveles de biomarcadores junto con el bloqueo detectado en un angiograma coronario), también puede evaluarse información con respecto a los biomarcadores junto con otros tipos de datos, en particular datos que indican el riesgo de un individuo de tener un evento CV (por ejemplo, historial clínico del paciente, síntomas, antecedentes familiares de enfermedades cardiovasculares, factores de riesgo tales como si el individuo es fumador o no, si es consumidor frecuente de alcohol y/o el estado de otros biomarcadores, etc.). Estos diversos datos pueden

evaluarse mediante métodos automatizados, tales como un programa/software informático, que puede incorporarse en un ordenador u otro aparato/dispositivo.

Los ensayos de biomarcadores también pueden asociarse a pautas y algoritmos de riesgo cardiovascular actualmente en uso en la práctica clínica. Por ejemplo, la puntuación de riesgo de Framingham usa los siguientes factores de riesgo para dar como resultado una puntuación de riesgo: el tono vascular, los niveles de colesterol LDL y colesterol HDL, los niveles de glucosa alterados, el hábito de fumar, la presión arterial sistólica y la diabetes. La frecuencia de los pacientes de alto riesgo aumenta con la edad y los hombres comprenden una mayor proporción de pacientes de alto riesgo que las mujeres.

También puede usarse cualquiera de los biomarcadores que se describen en ensayos de formación de imágenes. Por ejemplo, puede acoplarse un agente de formación de imágenes a cualquiera de los biomarcadores que se describen, que pueden usarse para ayudar a predecir el riesgo de un evento cardiovascular, para controlar la respuesta a intervenciones terapéuticas, para seleccionar poblaciones diana en un ensayo clínico, entre otros usos.

Detección y determinación de biomarcadores y valores de biomarcadores

Puede detectarse un valor de biomarcador para los biomarcadores que se describen en el presente documento usando cualquiera de diversos métodos analíticos conocidos. En una realización, se detecta un valor de biomarcador usando un reactivo de captura. Como se usa en el presente documento, un "agente de captura" o "reactivo de captura" se refiere a una molécula que es capaz de unirse específicamente a un biomarcador. En diversas realizaciones, el reactivo de captura puede exponerse al biomarcador en solución o puede exponerse al biomarcador mientras el reactivo de captura se inmoviliza sobre un soporte sólido. En otras realizaciones, el reactivo de captura contiene una característica que es reactiva con una característica secundaria sobre un soporte sólido. En estas realizaciones, el reactivo de captura puede exponerse al biomarcador en solución y, después, la característica del reactivo de captura puede usarse junto con la característica secundaria sobre el soporte sólido para inmovilizar el biomarcador sobre el soporte sólido. El reactivo de captura se selecciona basándose en el tipo de análisis que ha de realizarse. Los reactivos de captura incluyen, pero sin limitación, SOMAmeros, anticuerpos, adnectinas, anquirinas, otros miméticos de anticuerpos y otros anticuerpos de proteínas, autoanticuerpos, quimeras, moléculas pequeñas, un fragmento F(ab')₂, un fragmento de anticuerpo monocatenario, un fragmento Fv, un fragmento Fv monocatenario, un ácido nucleico, una lectina, un receptor de unión a ligando, afficuerpos, nanocuerpos, polímeros impresos, avímeros, peptidomiméticos, un receptor hormonal, un receptor de citocinas y receptores sintéticos, y modificaciones y fragmentos de éstos.

En algunas realizaciones, un valor de biomarcador se detecta usando un complejo de biomarcador/reactivo de captura.

En otras realizaciones, el valor de biomarcador deriva del complejo de biomarcador/reactivo de captura y se detecta indirectamente, tal como, por ejemplo, como resultado de una reacción que es posterior a la interacción biomarcador/reactivo de captura, pero que depende de la formación del complejo de biomarcador/reactivo de captura.

En algunas realizaciones, el valor de biomarcador se detecta directamente a partir del biomarcador en una muestra biológica.

En una realización, los biomarcadores se detectan usando un formato multiplexado que permite la detección simultánea de dos o más biomarcadores en una muestra biológica. En una realización del formato multiplexado, los reactivos de captura se inmovilizan, directa o indirectamente, covalentemente o no covalentemente, en ubicaciones discretas sobre un soporte sólido. En otra realización, un formato multiplexado usa soportes sólidos discretos donde cada soporte sólido tiene un reactivo de captura único asociado a ese soporte sólido, tal como, por ejemplo, puntos cuánticos. En otra realización, se usa un dispositivo individual para la detección de cada uno de múltiples biomarcadores que han de detectarse en una muestra biológica. Los dispositivos individuales pueden configurarse para permitir que cada biomarcador en la muestra biológica se procese simultáneamente. Por ejemplo, puede usarse una placa de microtitulación de manera que cada pocillo en la placa se use para analizar de forma única uno de los múltiples biomarcadores que han de detectarse en una muestra biológica.

En una o más de las realizaciones anteriores, puede usarse un marcador fluorescente para marcar un componente del complejo de biomarcador/reactivo de captura para permitir la detección del valor de biomarcador. En diversas realizaciones, el marcador fluorescente puede conjugarse con un reactivo de captura específico para cualquiera de los biomarcadores que se describen en el presente documento usando técnicas conocidas y, después, el marcador fluorescente puede usarse para detectar el valor de biomarcador correspondiente. Los marcadores fluorescentes adecuados incluyen quelatos de tierras raras, fluoresceína y sus derivados, rodamina y sus derivados, dansilo, alofocianina, PBXL-3, Qdot 605, Lisamina, ficoeritrina, rojo Texas y otros compuestos de este tipo.

En una realización, el marcador fluorescente es una molécula de colorante fluorescente. En algunas realizaciones, la molécula de colorante fluorescente incluye al menos un sistema de anillo de indolio sustituido en el que el sustituyente en el carbono 3 del anillo de indolio contiene un grupo químicamente reactivo o una sustancia conjugada. En algunas realizaciones, la molécula de colorante incluye una molécula de AlexaFluor, tal como, por ejemplo, AlexaFluor 488, AlexaFluor 532, AlexaFluor 647, AlexaFluor 680 o AlexaFluor 700. En otras realizaciones, la molécula de colorante

incluye un primer tipo y un segundo tipo de molécula de colorante, tal como, por ejemplo, dos moléculas de AlexaFluor diferentes. En otras realizaciones, la molécula de colorante incluye un primer tipo y un segundo tipo de molécula de colorante y las dos moléculas de colorante tienen espectros de emisión diferentes.

5 La fluorescencia puede medirse con diversos instrumentos compatibles con una amplia gama de formatos de ensayo. Por ejemplo, se han diseñado espectrofluorímetros para analizar placas de microtitulación, portaobjetos de microscopio, matrices impresas, cubetas, etc. Véase *Principles of Fluorescence Spectroscopy*, de J.R. Lakowicz, Springer Science + Business Media, Inc., 2004. Véase *Bioluminescence & Chemiluminescence: Progress & Current Applications*; Philip E. Stanley y Larry J. Kricka editores, World Scientific Publishing Company, enero de 2002.

10 En una o más de las realizaciones anteriores, puede usarse un marcador de quimioluminiscencia opcionalmente para marcar un componente del complejo de biomarcador/reactivo de captura para permitir la detección de un valor de biomarcador. Los materiales quimioluminiscentes adecuados incluyen cualquiera de entre cloruro de oxalilo, Rodamina 6G, Ru(bipi)₃²⁺, TMAE (tetraquis(dimetilamino)etileno), Pirogalol (1,2,3-trihidroxibenceno), Lucigenina, peroxioxalatos, oxalatos de arilo, ésteres de acridinio, dioxetanos y otros.

15 En otras realizaciones más, el método de detección incluye una combinación enzima/sustrato que genera una señal detectable que corresponde al valor de biomarcador. En general, la enzima cataliza una alteración química del sustrato cromógeno que puede medirse usando diversas técnicas, incluyendo espectrofotometría, fluorescencia y quimioluminiscencia. Las enzimas adecuadas incluyen, por ejemplo, luciferasas, luciferina, malato deshidrogenasa, ureasa, peroxidasa de rábano picante (HRPO), fosfatasa alcalina, beta-galactosidasa, glucoamilasa, lisozima, glucosa oxidasa, galactosa oxidasa y glucosa-6-fosfato deshidrogenasa, uricasa, xantina oxidasa, lactoperoxidasa, microperoxidasa y similares.

20 En otras realizaciones más, el método de detección puede ser una combinación de fluorescencia, quimioluminiscencia, radionúclidos o combinaciones de enzima/sustrato que generan una señal medible. La señalización multimodal podría tener características únicas y ventajosas en los formatos de ensayo de biomarcadores.

25 Más específicamente, los valores de biomarcadores para los biomarcadores que se describen en el presente documento pueden detectarse usando métodos analíticos conocidos incluyendo, ensayos de SOMAmeros únicos, ensayos de SOMAmeros múltiples, inmunoensayos únicos o múltiples, realización de perfiles de expresión de ARNm, realización de perfiles de expresión de miRNA, análisis espectrométrico de masas, métodos histológicos/citológicos, etc. como se detalla a continuación.

35 **Determinación de valores de biomarcadores usando ensayos a base de SOMAmero**

Los ensayos dirigidos a la detección y cuantificación de moléculas fisiológicamente significativas en muestras biológicas y otras muestras son herramientas importantes en la investigación científica y en el campo de la atención médica. Una clase de dichos ensayos implica el uso de una micromatriz que incluye uno o más aptámeros inmovilizados sobre un soporte sólido. Los aptámeros son capaces cada uno de unirse a una molécula diana de una manera altamente específica y con una afinidad muy alta. Véase, por ejemplo, la Patente de los EE.UU. N.º 5.475.096 titulada "*Nucleic Acid Ligands*"; véanse también, por ejemplo, la Patente de los EE.UU. N.º 6.242.246, la Patente de los EE.UU. N.º 6.458.543 y la Patente de los EE.UU. N.º 6.503.715, cada una de las cuales se titula "*Nucleic Acid Ligand Diagnostic Biochip*". Una vez que la micromatriz se pone en contacto con una muestra, los aptámeros se unen a sus respectivas moléculas diana presentes en la muestra y, por tanto, permiten la determinación de un valor de biomarcador correspondiente a un biomarcador.

Como se usa en el presente documento, un "aptámero" se refiere a un ácido nucleico que tiene una afinidad de unión específica por una molécula diana. Se reconoce que las interacciones de afinidad son una cuestión de grado; sin embargo, en este contexto, la "afinidad de unión específica" de un aptámero por su diana significa que el aptámero se une a su diana generalmente con un grado mucho mayor de afinidad que si se une a otros componentes en una muestra de ensayo. Un "aptámero" es un conjunto de copias de un tipo o especie de molécula de ácido nucleico que tiene una secuencia de nucleótidos particular. Un aptámero puede incluir cualquier número adecuado de nucleótidos, incluyendo cualquier número de nucleótidos modificados químicamente. "Aptámeros" se refiere a más de uno de dichos conjuntos de moléculas. Diferentes aptámeros pueden tener el mismo o diferente número de nucleótidos. Los aptámeros pueden ser ADN o ARN o ácidos nucleicos modificados químicamente y pueden ser monocatenarios, bicatenarios o contener regiones bicatenarias, y pueden incluir estructuras de orden superior. Un aptámero también puede ser un fotoaptámero, donde se incluye un grupo funcional fotorreactivo o químicamente reactivo en el aptámero para permitir que se una covalentemente a su diana correspondiente. Cualquiera de los métodos de aptámeros que se desvelan en el presente documento puede incluir el uso de dos o más aptámeros que se unen específicamente a la misma molécula diana. Como se describe adicionalmente a continuación, un aptámero puede incluir un marcador. Si un aptámero incluye un marcador, no es necesario que todas las copias del aptámero tengan el mismo marcador. Además, si diferentes aptámeros incluyen cada uno un marcador, estos diferentes aptámeros pueden tener el mismo marcador o un marcador diferente.

65 Puede identificarse un aptámero usando cualquier método conocido, incluyendo el proceso SELEX. Una vez

identificado, un aptámero puede prepararse o sintetizarse de acuerdo con cualquier método conocido, incluyendo métodos de síntesis química y métodos de síntesis enzimática.

5 Como se usa en el presente documento, un "SOMAmerno" o aptámero modificado de tasa de disociación lenta se refiere a un aptámero que tiene características de tasa de disociación mejoradas. Pueden generarse SOMAmernos usando los métodos SELEX mejorados que se describen en la Publicación de los EE.UU. N.º 2009/0004667, titulada "*Method for Generating Aptamers with Improved Off-Rates*".

10 Las expresiones "SELEX" y "proceso SELEX" se usan indistintamente en el presente documento para referirse en general a una combinación de (1) la selección de aptámeros que interactúan con una molécula diana de una manera deseable, por ejemplo, la unión con alta afinidad a una proteína, con (2) la amplificación de los ácidos nucleicos seleccionados. El proceso SELEX puede usarse para identificar aptámeros con alta afinidad por una diana o un biomarcador específicos.

15 En general, SELEX incluye preparar una mezcla candidata de ácidos nucleicos, unir la mezcla candidata a la molécula diana deseada para formar un complejo de afinidad, separar los complejos de afinidad de los ácidos nucleicos candidatos no unidos, separar y aislar el ácido nucleico del complejo de afinidad, purificar el ácido nucleico e identificar una secuencia de aptámero específica. El proceso puede incluir múltiples ciclos para perfeccionar adicionalmente la afinidad del aptámero seleccionado. El proceso puede incluir etapas de amplificación en uno o más puntos en el proceso. Véase, por ejemplo, la Patente de los EE.UU. N.º 5.475.096, titulada "*Nucleic Acid Ligands*". El proceso SELEX puede usarse para generar un aptámero que se une covalentemente a su diana así como un aptámero que se une no covalentemente a su diana. Véase, por ejemplo, la Patente de los EE.UU. N.º 5.705.337 titulada "*Systematic Evolution of Nucleic Acid Ligands by Exponential Enrichment: Chemi-SELEX*".

25 El proceso SELEX puede usarse para identificar aptámeros de alta afinidad que contienen nucleótidos modificados que confieren características mejoradas en el aptámero, tal como, por ejemplo, características de entrega mejoradas o estabilidad *in vivo* mejorada. Los ejemplos de dichas modificaciones incluyen sustituciones químicas en las posiciones de la ribosa y/o el fosfato y/o la base. Se describen aptámeros identificados con el proceso SELEX que contienen nucleótidos modificados en la Patente de los EE.UU. N.º 5.660.985, titulada "*High Affinity Nucleic Acid Ligands Containing Modified Nucleotides*", que describe oligonucleótidos que contienen derivados de nucleótidos modificados químicamente en las posiciones 5' y 2' de pirimidinas. la Patente de los EE.UU. N.º 5.580.737, véase más arriba, describe aptámeros altamente específicos que contienen uno o más nucleótidos modificados con 2'-amino (2'-NH₂), 2'-fluoro (2'-F) y/o 2'-O-metilo (2'-OMe). Véase también la Publicación de Solicitud de Patente de los EE.UU. N.º 20090098549, titulada "*SELEX and PHOTOSELEX*", que describe bibliotecas de ácidos nucleicos que tienen propiedades químicas y físicas expandidas y su uso en SELEX y photoSELEX.

30 También puede usarse SELEX para identificar aptámeros que tienen características de tasa de disociación deseables. Véase la Publicación de Solicitud de Patente de los EE.UU. 20090004667, titulada "*Method for Generating Aptamers with Improved Off-Rates*", que describe métodos SELEX mejorados para generar aptámeros que pueden unirse a moléculas diana. Como se ha mencionado anteriormente, estos aptámeros de tasa de disociación lenta se conocen como "SOMAmerno". Se describen métodos para producir aptámeros o SOMAmernos y fotoaptámeros o SOMAmernos que tienen tasas más lentas de disociación de sus respectivas moléculas diana. Los métodos implican poner en contacto la mezcla candidata con la molécula diana, permitir que se produzca la formación de complejos ácido nucleico-diana y realizar un proceso de enriquecimiento de tasa de disociación lenta en donde los complejos ácido nucleico-diana con tasas de disociación rápidas se disociarán y no se reformarán, mientras que los complejos con tasa de disociación lenta permanecerán intactos. Adicionalmente, los métodos incluyen el uso de nucleótidos modificados en la producción de mezclas de ácidos nucleicos candidatas para generar aptámeros o SOMAmernos con un rendimiento de tasa de disociación mejorado.

45 Una variación de este ensayo emplea aptámeros que incluyen grupo funcionales fotoreactivos que permiten que los aptámeros se unan covalentemente o "fotoreticulen" sus moléculas diana. Véase, por ejemplo, la Patente de los EE.UU. N.º 6.544.776 titulada "*Nucleic Acid Ligand Diagnostic Biochip*". Estos aptámeros fotoreactivos también se denominan fotoaptámeros. Véase, por ejemplo, la Patente de los EE.UU. N.º 5.763.177, la Patente de los EE.UU. N.º 6.001.577 y la Patente de los EE.UU. N.º 6.291.184, cada una de las cuales se titula "*Systematic Evolution of Nucleic Acid Ligands by Exponential Enrichment: Photoselection of Nucleic Acid Ligands and Solution SELEX*"; véanse también, por ejemplo, la Patente de los EE.UU. N.º 6.458.539, titulada "*Photoselection of Nucleic Acid Ligands*". Después de poner la micromatriz en contacto con la muestra y de que los fotoaptámeros hayan tenido la oportunidad de unirse a sus moléculas diana, los fotoaptámeros se fotoactivan y el soporte sólido se lava para retirar cualquier molécula no específicamente unida. Pueden usarse condiciones de lavado duras, puesto que las moléculas diana que están unidas a los fotoaptámeros generalmente no se retiran, debido a los enlaces covalentes creados por el grupo o grupos funcionales fotoactivados sobre los fotoaptámeros. De esta manera, el ensayo permite la detección de un valor de biomarcador correspondiente a un biomarcador en la muestra de ensayo.

65 En ambos de estos formatos de ensayo, los aptámeros o SOMAmernos se inmovilizan sobre el soporte sólido antes de ponerse en contacto con la muestra. En determinadas circunstancias, sin embargo, la inmovilización de los aptámeros o SOMAmernos antes del contacto con la muestra puede no proporcionar un ensayo óptimo. Por ejemplo, la

preinmovilización de los aptámeros o SOMAmeros puede dar como resultado una mezcla ineficiente de los aptámeros o SOMAmeros con las moléculas diana sobre la superficie del soporte sólido, conduciendo, quizás, a tiempo de reacción largos y, por tanto, períodos de incubación prolongados para permitir una unión eficiente de los aptámeros o SOMAmeros a sus moléculas diana. Adicionalmente, cuando se emplean fotoaptámeros o fotoSOMAmeros en el ensayo y dependiendo del material utilizado como soporte sólido, el soporte sólido puede tender a dispersar o absorber la luz utilizada para efectuar la formación de enlaces covalentes entre los fotoaptámeros o fotoSOMAmeros y sus moléculas diana. Además, dependiendo del método empleado, la detección de moléculas diana unidas a sus aptámeros o fotoSOMAmeros puede estar sujeta a imprecisión, puesto que la superficie del soporte sólido también puede estar expuesta y verse afectada por cualquier agente marcador que se use. Por último, la inmovilización de los aptámeros o SOMAmeros sobre el soporte sólido generalmente implica una etapa de preparación de aptámero o SOMAmero (es decir, la inmovilización) antes de la exposición de los aptámeros o SOMAmeros a la muestra, y esta etapa de preparación puede afectar a la actividad o funcionalidad de los aptámeros o SOMAmeros.

También se han descrito ensayos de SOMAmero que permiten a un SOMAmero capturar su diana en solución y después emplear etapas de separación que se diseñan para retirar componentes específicos de la mezcla de SOMAmero-diana antes de la detección (véase la Publicación de Solicitud de Patente de los EE.UU. 20090042206, titulada "*Multiplexed Analyses of Test Samples*"). Los métodos de ensayo de SOMAmero que se describen permiten la detección y cuantificación de una diana no de ácido nucleico (por ejemplo, una diana de proteína) en una muestra de ensayo mediante la detección y cuantificación de un ácido nucleico (es decir, un SOMAmero). Los métodos que se describen crean un sustituto de ácido nucleico (es decir, el SOMAmero) para detectar y cuantificar una diana no de ácido nucleico, permitiendo, de este modo, la amplia diversidad de tecnologías de ácido nucleico, incluyendo la amplificación, que ha de aplicarse a una gama más amplia de dianas deseadas, incluyendo dianas de proteína.

Pueden construirse SOMAmeros para facilitar la separación de los componentes de ensayo de un complejo de biomarcadores de SOMAmero (o complejo covalente de biomarcadores de fotoSOMAmero) y permitir el aislamiento del SOMAmero para la detección y/o cuantificación. En una realización, estas construcciones pueden incluir un elemento escindible o liberable dentro de la secuencia de SOMAmero. En otras realizaciones, puede introducirse una funcionalidad adicional en el SOMAmero, por ejemplo, un componente marcado o detectable, un componente espaciador o un elemento de inmovilización o marcador de unión específico. Por ejemplo, el SOMAmero puede incluir un marcador conectado al SOMAmero a través de un resto escindible, un marcador, un componente espaciador que separe el marcador y el resto escindible. En una realización, un elemento escindible es un enlazador fotoescindible. El enlazador fotoescindible puede unirse a un resto de biotina y una sección espaciadora, puede incluir un grupo NHS para la derivatización de aminas y puede usarse para introducir un grupo biotina a un SOMAmero, permitiendo de este modo la liberación del SOMAmero más adelante en un método de ensayo.

Los ensayos homogéneos, realizados con todos los componentes de ensayo en solución, no requieren la separación de la muestra y los reactivos antes de la detección de señal. Estos métodos son rápidos y fáciles de usar. Estos métodos generan una señal basándose en un reactivo de unión o reactivo de captura molecular que reacciona con su diana específica. Para la predicción de eventos CV, los reactivos de captura molecular serían un SOMAmero o un anticuerpo o similar, y la diana específica sería un biomarcador de eventos CV de la Tabla 1, Col. 7.

En una realización, un método para la generación de señal aprovecha el cambio de señal de anisotropía debido a la interacción de un reactivo de captura marcado con fluoróforo con su diana biomarcadora específica. Cuando el reactivo de captura marcado reacciona con su diana, el aumento del peso molecular provoca que el movimiento de rotación del fluoróforo unido al complejo sea mucho más lento cambiando el valor de anisotropía. Mediante el control del cambio de anisotropía, pueden usarse eventos de unión para medir cuantitativamente los biomarcadores en soluciones. Otros métodos incluyen ensayos de polarización de fluorescencia, métodos de baliza molecular, extinción de fluorescencia resuelta en el tiempo, quimioluminiscencia, transferencia de energía de resonancia de fluorescencia y similares.

Un ensayo de SOMAmero basado en solución de ejemplo que puede usarse para detectar un valor de biomarcador correspondiente a un biomarcador en una muestra biológica incluye lo siguiente: (a) preparar una mezcla poniendo en contacto la muestra biológica con un SOMAmero que incluye un primer marcador y tiene un afinidad específica por el biomarcador, en donde se forma un complejo de afinidad de SOMAmero cuando el biomarcador está presente en la muestra; (b) exponer la mezcla a un primer soporte sólido que incluye un primer elemento de captura y permitir que el primer marcador se asocie al primer elemento de captura; (c) retirar cualquier componente de la mezcla no asociado al primer soporte sólido; (d) unir un segundo marcador al componente biomarcador del complejo de afinidad de SOMAmero; (e) liberar el complejo de afinidad de SOMAmero del primer soporte sólido; (f) exponer el complejo de afinidad de SOMAmero liberado a un segundo soporte sólido que incluye un segundo elemento de captura y permitir que el segundo marcador se asocie al segundo elemento de captura; (g) retirar cualquier SOMAmero no complejado de la mezcla separando el SOMAmero no complejado del complejo de afinidad de SOMAmero; (h) eluir el SOMAmero del soporte sólido; y (i) detectar el biomarcador detectando el componente SOMAmero del complejo de afinidad de SOMAmero.

Puede usarse cualquier medio conocido en la técnica para detectar un valor de biomarcador detectando el componente SOMAmero de un complejo de afinidad de SOMAmero. Pueden usarse varios métodos de detección diferentes para detectar el componente SOMAmero de un complejo de afinidad, tal como, por ejemplo, ensayos de hibridación,

espectroscopía de masas o QPCR. En algunas realizaciones, pueden usarse métodos de secuenciación de ácido nucleico para detectar el componente SOMAmero de un complejo de afinidad de SOMAmero y, por tanto, detectar un valor de biomarcador. Brevemente, una muestra de ensayo puede someterse a cualquier tipo de método de secuenciación de ácido nucleico para identificar y cuantificar la secuencia o secuencias de uno o más SOMAmeros presentes en la muestra de ensayo. En algunas realizaciones, la secuencia incluye la molécula de SOMAmero completa o cualquier porción de la molécula que pueda usarse para identificar unívocamente la molécula. En otras realizaciones, la secuencia de identificación es una secuencia específica añadida al SOMAmero; dichas secuencias con frecuencia se denominan "marcadores", "códigos de barras", o "códigos postales". En algunas realizaciones, el método de secuenciación incluye etapas enzimáticas para amplificar la secuencia de SOMAmero o para convertir cualquier tipo de ácido nucleico, incluyendo ARN y ADN que contienen modificaciones químicas en cualquier posición, en cualquier otro tipo de ácido nucleico adecuado para secuenciación.

En algunas realizaciones, el método de secuenciación incluye una o más etapas de clonación. En otras realizaciones, el método de secuenciación incluye un método de secuenciación directa sin clonación.

En algunas realizaciones, el método de secuenciación incluye un enfoque dirigido con cebadores específicos que se dirigen a uno o más SOMAmeros en la muestra de ensayo. En otras realizaciones, el método de secuenciación incluye un enfoque aleatorio que se dirige a todos los SOMAmeros en la muestra de ensayo.

En algunas realizaciones, el método de secuenciación incluye etapas enzimáticas para amplificar la molécula diana para la secuenciación. En otras realizaciones, el método de secuenciación secuencia directamente moléculas individuales. Un método basado en secuenciación de ácido nucleico de ejemplo que puede usarse para detectar un valor de biomarcador correspondiente a un biomarcador en una muestra biológica incluye lo siguiente: (a) convertir una mezcla de SOMAmeros que contienen nucleótidos modificados químicamente en ácidos nucleicos no modificados con una etapa enzimática; (b) secuenciar aleatoriamente los ácidos nucleicos no modificados resultantes con una plataforma de secuenciación masivamente paralela tal como, por ejemplo, el Sistema de Secuenciación 454 (454 Life Sciences/Roche), el Sistema de Secuenciación Illumina (Illumina), el Sistema de Secuenciación ABI SOLiD (Applied Biosystems), el Secuenciador de Moléculas Individuales HeliScope (Helicos Biosciences) o el Sistema de Secuenciación de Moléculas Individuales en Tiempo Real Pacific Biosciences (Pacific BioSciences) o el Sistema de Secuenciación Polonator G (Dover Systems); e (c) identificar y cuantificar los SOMAmeros presentes en la mezcla por secuencia específica y recuento de secuencia.

Determinación de valores de biomarcadores usando inmunoensayos

Los métodos de inmunoensayo se basan en la reacción de un anticuerpo con su diana o analito correspondiente y pueden detectar el analito en una muestra dependiendo del formato de ensayo específico. Para mejorar la especificidad y la sensibilidad de un método de ensayo basado en inmunorreactividad, con frecuencia se usan anticuerpos monoclonales debido a su reconocimiento específico de epítipo. También se han utilizado con éxito anticuerpos policlonales en diversos inmunoensayos debido a su mayor afinidad por la diana en comparación con los anticuerpos monoclonales. Se han diseñado inmunoensayos para su uso con una amplia gama de matrices de muestras biológicas. Se han diseñado formatos de inmunoensayo para proporcionar resultados cualitativos, semicuantitativos y cuantitativos.

Se generan resultados cuantitativos a través del uso de una curva patrón creada con concentraciones conocidas del analito específico que ha de detectarse. La respuesta o señal de una muestra desconocida se traza en la curva patrón y se establece una cantidad o valor correspondiente a la diana en la muestra desconocida.

Se han diseñado numerosos formatos de inmunoensayo. ELISA o EIA pueden ser cuantitativos para la detección de un analito. Este método se basa en la unión de un marcador al analito o al anticuerpo y el componente marcador incluye, tanto directa como indirectamente, una enzima. Los ensayos ELISA pueden formatearse para la detección directa, indirecta, competitiva o sándwich del analito. Otros métodos se basan en marcadores tales como, por ejemplo, radioisótopos (¹¹²⁵) o fluorescencia. Las técnicas adicionales incluyen, por ejemplo, aglutinación, nefelometría, turbidimetría, transferencia Western, inmunoprecipitación, inmunocitoquímica, inmunohistoquímica, citometría de flujo, ensayo Luminex y otras (véase *ImmunoAssay: A Practical Guide*, editado por Brian Law, publicado por Taylor & Francis, Ltd., edición de 2005).

Los formatos de ensayo de ejemplo incluyen ensayo inmunosorbente ligado a enzimas (ELISA), radioinmunoensayo, inmunoensayos fluorescentes, de quimioluminiscencia y de transferencia de energía de resonancia de fluorescencia (FRET) o de FRET (TR-FRET) resueltos en el tiempo. Los ejemplos de procedimientos para detectar biomarcadores incluyen la inmunoprecipitación de biomarcadores seguida de métodos cuantitativos que permiten la discriminación a nivel de tamaño y de péptido, tales como electroforesis en gel, electroforesis capilar, electrocromatografía plana y similares.

Los métodos para detectar y/o cuantificar un marcador detectable o material generador de señal dependen de la naturaleza del marcador. Los productos de las reacciones catalizadas por enzimas adecuadas (donde el marcador detectable es una enzima; véase anteriormente) pueden ser, sin limitación, fluorescentes, luminiscentes o radioactivos,

o pueden absorber luz visible o ultravioleta. Los ejemplos de detectores adecuados para detectar dichos marcadores detectables incluyen, sin limitación, película de rayos X, contadores de radiactividad, contadores de centelleo, espectrofotómetros, colorímetros, fluorómetros, luminómetros y densitómetros.

- 5 Cualquiera de los métodos de detección puede realizarse en cualquier formato que permita cualquier preparación, procesamiento y análisis adecuados de las reacciones. Esto puede ser, por ejemplo, en placas de ensayo de múltiples pocillos (por ejemplo, 96 pocillos o 384 pocillos) o usando cualquier matriz o micromatriz adecuada. Las soluciones de stock para diversos agentes pueden hacerse de forma manual o robótica y todos los procesos posteriores de pipeteo, dilución, mezcla, distribución, lavado, incubación, lectura de muestras, recopilación y análisis de datos pueden realizarse de forma robótica usando un software de análisis, robótica e instrumentos de detección disponibles en el mercado capaces de detectar un marcador detectable.

Determinación de valores de biomarcadores usando perfiles de expresión génica

- 15 Puede usarse medición de ARNm en una muestra biológica como sustituto para la detección del nivel de la proteína correspondiente en la muestra biológica. Por tanto, también puede detectarse cualquiera de los biomarcadores o paneles de biomarcadores que se describen en el presente documento detectando el ARN adecuado.

20 Los niveles de expresión de ARNm se miden mediante reacción en cadena de la polimerasa cuantitativa de transcripción inversa (RT-PCR seguida de qPCR). Se usa RT-PCR para crear un ADNc a partir del ARNm. El ADNc puede usarse en un ensayo de qPCR para producir fluorescencia a medida que avanza el proceso de amplificación de ADN. En comparación con una curva patrón, la qPCR puede producir una medición absoluta, tal como el número de copias de ARNm por célula. Se han utilizado transferencias Northern, micromatrices, ensayos Invader y RT-PCR combinados con electroforesis capilar para medir los niveles de expresión de ARNm en una muestra. Véase *Gene Expression Profiling: Methods and Protocols*, Richard A. Shimkets, editor, Humana Press, 2004.

30 Las moléculas de miARN son ARN pequeños que no son codificantes pero que pueden regular la expresión génica. También puede usarse cualquiera de los métodos adecuados para la medición de niveles de expresión de ARNm para el miARN correspondiente. Recientemente, muchos laboratorios han investigado el uso de miRNA como biomarcadores de enfermedades. Muchas enfermedades implican una regulación transcripcional generalizada y no es sorprendente que los miRNA puedan encontrar un papel como biomarcadores. La conexión entre las concentraciones de miARN y la enfermedad con frecuencia es aún menos clara que las conexiones entre los niveles de proteínas y la enfermedad, sin embargo, el valor de los biomarcadores de miARN podría ser sustancial. Por supuesto, como con cualquier ARN expresado diferencialmente durante la enfermedad, los problemas que enfrenta el desarrollo de un producto de diagnóstico *in vitro* incluirán el requisito de que los miARN sobrevivan en la célula enferma y se extraigan fácilmente para su análisis o que los miARN se liberen en la sangre u otras matrices donde deben sobrevivir el tiempo suficiente para ser medidos. Los biomarcadores de proteína tienen requisitos similares, aunque muchos biomarcadores de proteína potenciales se secretan intencionadamente en el sitio de la patología y la función, durante la enfermedad, de forma paracrina. Muchos biomarcadores de proteína potenciales se diseñan para que funcionen fuera de las células dentro de las cuales se sintetizan esas proteínas.

Detección de biomarcadores usando tecnologías de formación de imágenes moleculares *in vivo*

45 También puede usarse cualquiera de los biomarcadores que se describen (véase la Tabla 1, Col. 7) en ensayos de formación de imágenes moleculares. Por ejemplo, puede acoplarse un agente de formación de imágenes a cualquiera de los biomarcadores que se describen, que pueden usarse para ayudar a predecir el riesgo de eventos cardiovasculares en 5 años, para controlar la respuesta a intervenciones terapéuticas, para seleccionar una población para ensayos clínicos, entre otros usos.

50 Las tecnologías de formación de imágenes *in vivo* proporcionan métodos no invasivos para determinar el estado de una enfermedad o afección particular en el cuerpo de un individuo. Por ejemplo, pueden verse partes enteras del cuerpo o incluso todo el cuerpo, como una imagen tridimensional, proporcionando de este modo información valiosa sobre la morfología y las estructuras del cuerpo. Dichas tecnologías pueden combinarse con la detección de los biomarcadores que se describen en el presente documento para proporcionar información sobre el estado cardiovascular de un individuo.

60 El uso de tecnologías de formación de imágenes moleculares *in vivo* se está expandiendo debido a diversos avances tecnológicos. Estos avances incluyen el desarrollo de nuevos agentes de contraste o marcadores, tales como radiomarcadores y/o marcadores fluorescentes, que pueden proporcionar señales fuertes dentro del cuerpo; y el desarrollo de una tecnología de formación de imágenes nueva y potente, que puede detectar y analizar estas señales desde el exterior del cuerpo, con suficiente sensibilidad y precisión para proporcionar información útil. El agente de contraste puede visualizarse en un sistema de formación de imágenes adecuado, proporcionando de este modo una imagen de la porción o porciones del cuerpo en las que se encuentra el agente de contraste. El agente de contraste puede unirse o asociarse a un reactivo de captura, tal como un SOMAmo o un anticuerpo, por ejemplo, y/o con un péptido o proteína, o un oligonucleótido (por ejemplo, para la detección de la expresión génica) o un complejo que contiene cualquiera de estos con una o más macromoléculas y/u otras formas en forma de partículas.

El agente de contraste también puede presentar un átomo radiactivo que es útil en la formación de imágenes. Los átomos radiactivos adecuados incluyen tecnecio-99m o yodo-123 para estudios de centellografía. Otros restos fácilmente detectables incluyen, por ejemplo, marcadores de spin para resonancia magnética (RM) tales como, por ejemplo, yodo-123 nuevamente, yodo-131, indio-113, flúor-19, carbono-13, nitrógeno-15, oxígeno-17, gadolinio, manganeso o hierro. Dichos marcadores son bien conocidas en la técnica y podrían ser seleccionados fácilmente por un experto en la materia.

Las técnicas de formación de imágenes convencionales incluyen, pero sin limitación, resonancia magnética, tomografía computarizada (puntuación de calcio coronario), tomografía de emisión de positrones (PET), tomografía computarizada por emisión de fotón único (SPECT), angiografía por tomografía computarizada y similares. Para el diagnóstico por formación de imágenes *in vivo*, el tipo de instrumento de detección disponible es un factor importante en la selección de un agente de contraste dado, tal como un radionúclido dado y el biomarcador particular que se usa para dirigir (proteína, ARNm y similares). El radionúclido elegido normalmente tiene un tipo de descomposición que es detectable por un tipo dado de instrumento. Además, cuando se selecciona un radionúclido para el diagnóstico *in vivo*, su semivida debe ser lo suficientemente larga como para permitir la detección en el momento de la captación máxima por parte del tejido diana, pero lo suficientemente corta como para minimizar la radiación perjudicial del hospedador.

Las técnicas de formación de imágenes de ejemplo incluyen, pero sin limitación, PET y SPECT, que son técnicas de formación de imágenes en las que un radionúclido se administra de forma sintética o local a un individuo. La captación posterior del radioindicador se mide a lo largo del tiempo y se usa para obtener información sobre el tejido diana y el biomarcador. Debido a las emisiones de alta energía (rayos gamma) de los isótopos específicos empleados y la sensibilidad y sofisticación de los instrumentos utilizados para detectarlos, la distribución bidimensional de la radiactividad puede inferirse desde el exterior del cuerpo.

Los nucleidos emisores de positrones utilizados habitualmente en PET incluyen, por ejemplo, carbono-11, nitrógeno-13, oxígeno-15 y flúor-18. Se usan isótopos que se descomponen por captura de electrones y/o emisión gamma en SPECT e incluyen, por ejemplo, yodo-123 y tecnecio-99m. Un método de ejemplo para marcar aminoácidos con tecnecio-99m es la reducción del ion pertecnetato en presencia de un precursor quelante para formar el complejo de tecnecio-99m-precursor lábil, que, a su vez, reacciona con el grupo de unión al metal de un péptido quimiotáctico modificado bifuncionalmente para formar un conjugado de tecnecio-99m-péptido quimiotáctico.

Con frecuencia se usan anticuerpos para dichos métodos de diagnóstico por formación de imágenes *in vivo*. La preparación y el uso de anticuerpos para el diagnóstico *in vivo* son bien conocidos en la técnica. Pueden inyectarse anticuerpos marcados que se unen específicamente a cualquiera de los biomarcadores de la Tabla 1, Col. 7 en un individuo que se sospecha que tiene un mayor riesgo de un evento CV, detectable de acuerdo con el biomarcador particular utilizado, con el fin de diagnosticar o evaluar la patología o el estado del individuo. El marcador utilizado se seleccionará de acuerdo con la modalidad de formación de imágenes que ha de usarse, como se ha descrito anteriormente. La localización del marcador permite la determinación del daño tisular u otras indicaciones relacionadas con el riesgo de un evento CV. La cantidad de marcador dentro de un órgano o tejido también permite la determinación de la implicación de los biomarcadores de eventos CV debido al riesgo de un evento CV en ese órgano o tejido.

De manera similar, pueden usarse SOMAmeros para dichos métodos de diagnóstico por formación de imágenes *in vivo*. Por ejemplo, un SOMAmero que se usó para identificar un biomarcador particular que se describe en la Tabla 1, Col. 7 (y, por tanto, se une específicamente a ese biomarcador particular) puede marcarse e inyectarse adecuadamente en un individuo que se sospecha que ha tenido un evento CV, detectable de acuerdo con el biomarcador particular, con el fin de diagnosticar o evaluar los niveles de daño tisular, placas ateroscleróticas, componentes de respuesta inflamatoria y otros factores asociados al riesgo de un evento CV en el individuo. El marcador utilizado se seleccionará de acuerdo con la modalidad de formación de imágenes que ha de usarse, como se ha descrito anteriormente. La localización del marcador permite determinar el sitio de los procesos que conducen a un mayor riesgo. La cantidad de marcador dentro de un órgano o tejido también permite la determinación de la infiltración del proceso patológico en ese órgano o tejido. Los agentes de formación de imágenes dirigidos por SOMAmero podrían tener características únicas y ventajosas relacionadas con la penetración de tejidos, la distribución tisular, la cinética, la eliminación, la potencia y la selectividad en comparación con otros agentes de formación de imágenes.

Dichas técnicas también pueden realizarse opcionalmente con oligonucleótidos marcados, por ejemplo, para la detección de la expresión génica a través de formación de imágenes con oligonucleótidos antisentido. Estos métodos se usan para la hibridación *in situ*, por ejemplo, con moléculas fluorescentes o radionucleidos como marcador. Otros métodos para la detección de la expresión génica incluyen, por ejemplo, la detección de la actividad de un gen indicador.

Otro tipo general de tecnología de formación de imágenes es la imagen óptica, en la que las señales fluorescentes dentro del sujeto son detectadas por un dispositivo óptico que es externo al sujeto. Estas señales pueden deberse a fluorescencia real y/o a bioluminiscencia. Las mejoras en la sensibilidad de los dispositivos de detección ópticos han aumentado la utilidad de la formación de imágenes ópticas para ensayos de diagnóstico *in vivo*.

El uso de formación de imágenes de biomarcadores moleculares *in vivo* está aumentando, incluso para ensayos clínicos, por ejemplo, para medir más rápidamente la eficacia clínica en ensayos para nuevas terapias de enfermedades o afecciones y/o para evitar el tratamiento prolongado con un placebo para esas enfermedades, tales como las múltiples esclerosis, en la que dicho tratamiento prolongado puede considerarse éticamente cuestionable.

Para una revisión de otras técnicas, véase N. Blow, *Nature Methods*, 6, 465-469, 2009.

Determinación de valores de biomarcadores usando métodos de espectrometría de masas

Puede usarse diversas configuraciones de espectrómetros de masas para detectar valores de biomarcadores. Existen varios tipos de espectrómetros de masas disponibles o pueden producirse con diversas configuraciones. En general, un espectrómetro de masas tiene los siguientes componentes principales: una entrada de muestra, una fuente de iones, un analizador de masas, un detector, un sistema de vacío y un sistema de control del instrumento, y un sistema de datos. Las diferencias en la entrada de muestra, en la fuente de iones y en el analizador de masas generalmente definen el tipo de instrumento y sus capacidades. Por ejemplo, una entrada puede ser una fuente de cromatografía líquida de columna capilar o puede ser una etapa o una sonda directa tal como la utilizada en la desorción láser asistida por matriz. Son fuentes de iones comunes, por ejemplo, la electronebulización, incluyendo la nanonebulización y la micronebulización, o la desorción láser asistida por matriz. Los analizadores de masas comunes incluyen un filtro de masas de cuadrupolo, un analizador de masas de trampa de iones y un analizador de masas de tiempo de vuelo. Se conocen bien en la técnica métodos adicionales de espectrometría de masas (véase Burlingame et al. *Anal. Chem.* 70: 647 R-716R (1998); Kinter y Sherman, Nueva York (2000)).

Los biomarcadores de proteína y los valores de biomarcadores pueden detectarse y medir mediante cualquiera de los siguientes: espectrometría de masas de ionización por electronebulización (IEN-EM), IEN-EM/EM, IEN-EM/(EM)_n, espectrometría de masas de tiempo de vuelo de ionización por desorción láser asistida por matriz (MALDI-TOF-EM), espectrometría de masas de tiempo de vuelo de desorción/ionización láser de superficie potenciada (SELDI-TOF-EM), desorción/ionización en silicio (DIOS), espectrometría de masas de iones secundarios (SIMS), tiempo de vuelo de cuadrupolo (Q-TOF), tecnología de tiempo de vuelo en tándem (TOF/TOF), denominada ultraflex III TOF/TOF, espectrometría de masas de ionización química a presión atmosférica (APCI-EM), APCI-EM/EM, APCI-(EM)_N, espectrometría de masas de fotoionización a presión atmosférica (APPI-EM), APPI-EM/EM y APPI-(EM)_N, espectrometría de masas de cuadrupolo, espectrometría de masas de transformada de Fourier (EMTF), espectrometría de masas cuantitativa y espectrometría de masas de trampa de iones.

Se usan estrategias de preparación de muestras para marcar y enriquecer las muestras antes de la caracterización espectroscópica de masas de biomarcadores de proteína y la determinación de los valores de biomarcadores. Los métodos de marcado incluyen, pero sin limitación, un marcador isobárico para la cuantificación relativa y absoluta (iTRAQ) y el marcado con isótopos estable con aminoácidos en cultivo celular (SI LAC). Los reactivos de captura utilizados para enriquecer selectivamente muestras para proteínas biomarcadoras candidatas antes del análisis espectroscópico de masas incluyen, pero sin limitación, SOMAmeros, anticuerpos, sondas de ácido nucleico, quimeras, moléculas pequeñas, un fragmento F(ab')₂, un fragmento de anticuerpo monocatenario, un fragmento Fv, un fragmento Fv monocatenario, un ácido nucleico, una lectina, un receptor de unión a ligando, anticuerpos, nanocuerpos, anquirinas, anticuerpos de dominio, armazones de anticuerpos alternativos (por ejemplo, diacuerpos, etc.), polímeros impresos, avímeros, peptidomiméticos, peptoides, ácidos peptidonucleicos, ácido nucleico de treosa, un receptor hormonal, un receptor de citocinas y receptores sintéticos, y modificaciones y fragmentos de éstos.

Determinación de valores de biomarcadores usando un ensayo de ligadura de proximidad

Puede usarse un ensayo de ligadura de proximidad para determinar valores de biomarcadores. Brevemente, una muestra de ensayo se pone en contacto con un par de sondas de afinidad que puede ser un par de anticuerpos o un par de SOMAmeros, con cada miembro del par prolongado con un oligonucleótido. Las dianas para el par de sondas de afinidad pueden ser dos determinantes distintos en una proteína o un determinante en cada una de las dos proteínas diferentes, que pueden existir como complejos homo o hetero multiméricos. Cuando las sondas se unen a los determinantes diana, los extremos libres de las prolongaciones de oligonucleótidos se acercan lo suficiente como para hibridarse entre sí. La hibridación de las prolongaciones de oligonucleótidos se ve facilitada por un oligonucleótido conector común que sirve para unir las prolongaciones de oligonucleótidos cuando se colocan suficientemente cerca. Una vez que las prolongaciones de oligonucleótidos de las sondas se hibridan, los extremos de las prolongaciones se unen entre sí mediante ligadura enzimática de ADN.

Cada prolongación de oligonucleótido comprende un sitio cebador para la amplificación por PCR. Una vez que las prolongaciones de oligonucleótidos se ligan entre sí, los oligonucleótidos forman una secuencia de ADN continua que, a través de la amplificación por PCR, revela información sobre la identidad y la cantidad de la proteína diana, así como, información sobre las interacciones proteína-proteína en las que los determinantes diana están en dos proteínas diferentes. La ligadura por cercanía puede proporcionar un ensayo altamente sensible y específico para la concentración de proteínas en tiempo real y la información de interacción a través del uso de PCR en tiempo real. Las sondas que no se unen a los determinantes de interés no tienen las prolongaciones de oligonucleótidos

correspondientes acercadas y no puede producirse la ligadura ni la amplificación por PCR, dando como resultado que no se produce ninguna señal.

5 Los ensayos anteriores permiten la detección de valores de biomarcadores que son útiles en métodos para predecir el riesgo de eventos CV, donde los métodos comprenden detectar, en una muestra biológica de un individuo, al menos N valores de biomarcadores que corresponden cada uno a un biomarcador seleccionado entre el grupo que consiste en los biomarcadores proporcionados en la Tabla 1, Col. 7, en donde una clasificación, como se describe en detalle a continuación, usando los valores de biomarcadores indica si el individuo tiene un riesgo elevado de que se produzca un evento CV dentro de un período de 5 años. Aunque algunos de los biomarcadores de eventos CV que se describen son útiles solos para predecir el riesgo de un evento CV, en el presente documento también se describen métodos para la agrupación de múltiples subconjuntos de los biomarcadores de eventos CV que son útiles como un panel de tres o más biomarcadores. Por tanto, diversas realizaciones de la presente solicitud proporcionan combinaciones que comprenden N biomarcadores, en donde N es al menos tres biomarcadores. En otras realizaciones, N se selecciona para que sea cualquier número entre 2-155 biomarcadores. Se apreciará que N puede seleccionarse para que sea cualquier número entre cualquiera de los intervalos descritos anteriormente, así como intervalos similares, pero de orden superior. De acuerdo con cualquiera de los métodos que se describen en el presente documento, pueden detectarse y clasificarse valores de biomarcadores individualmente o pueden detectarse y clasificarse colectivamente, como, por ejemplo, en un formato de ensayo múltiple.

20 Un "distintivo" de biomarcador para un ensayo de diagnóstico o predictivo dado contiene un conjunto de marcadores, teniendo cada marcador diferentes niveles en las poblaciones de interés. Diferentes niveles, en este contexto, puede referirse a diferentes medios de los niveles de marcador para los individuos en dos o más grupos, o diferentes variaciones en los dos o más grupos, o una combinación de ambos. Para la forma más simple de un ensayo de diagnóstico, estos marcadores pueden usarse para asignar una muestra desconocida de un individuo a uno de dos grupos, ya sea con un mayor riesgo de un evento CV o no. La asignación de una muestra en uno de dos o más grupos se conoce como clasificación y el procedimiento utilizado para lograr esta asignación se conoce como clasificador o método de clasificación. Los métodos de clasificación también pueden denominarse métodos de puntuación. Existen muchos métodos de clasificación que pueden usarse para construir un clasificador de diagnóstico a partir de un conjunto de valores de biomarcadores. En general, los métodos de clasificación se realizan con mayor facilidad usando técnicas de aprendizaje supervisado donde se recopila un conjunto de datos usando muestras obtenidas de individuos de dos (o más, para múltiples estados de clasificación) grupos distintos que se desean distinguir. Puesto que la clase (grupo o población) a la que pertenece cada muestra se conoce de antemano para cada muestra, el método de clasificación puede entrenarse para proporcionar la respuesta de clasificación deseada. También es posible usar técnicas de aprendizaje sin supervisión para producir un clasificador de diagnóstico.

35 Los enfoques comunes para desarrollar clasificadores de diagnóstico incluyen árboles de decisión; compartimentación, refuerzo, bosques y bosques aleatorios; aprendizaje basado en inferencia de reglas; Ventanas Parzen; modelos lineales; logística; métodos de redes neuronales; agrupamiento sin supervisión; medios K; ascendente/descendente jerárquico; aprendizaje semisupervisado; métodos prototipo; vecino más cercano; estimación de densidad de kernel; máquinas de vectores de soporte; modelos ocultos de Markov; Aprendizaje Boltzmann; y los clasificadores pueden combinarse de manera simple o de manera que minimicen funciones objetivas particulares. Para una revisión, véase, por ejemplo, *Pattern Classification*, R.O. Duda, et al., editores, John Wiley & Sons, 2ª edición, 2001; véanse también, *The Elements of Statistical Learning - Data Mining, Inference, and Prediction*, T. Hastie, et al., editores, Springer Science+Business Media, LLC, 2ª edición, 2009.

45 Para producir un clasificador usando técnicas de aprendizaje supervisado, se obtiene un conjunto de muestras denominado datos de entrenamiento. En el contexto de los ensayos de diagnóstico, los datos de entrenamiento incluyen muestras de los distintos grupos (clases) a los que se asignarán muestras desconocidas más adelante. Por ejemplo, las muestras recogidas de individuos en una población de control y de individuos en una población particular de enfermedad, afección o eventos, pueden constituir datos de entrenamiento para desarrollar un clasificador que pueda clasificar muestras desconocidas (o, más en particular, los individuos de quienes se obtuvieron las muestras) como que tienen la enfermedad, afección o riesgo elevado de un evento, o como que no tienen la enfermedad, afección o riesgo elevado de un evento. El desarrollo del clasificador a partir de los datos de entrenamiento se conoce como entrenamiento del clasificador. Los detalles específicos sobre el entrenamiento del clasificador dependen de la naturaleza de la técnica de aprendizaje supervisado (véase, por ejemplo, *Pattern Classification*, R.O. Duda, et al., editores, John Wiley & Sons, 2ª edición, 2001; véanse también, *The Elements of Statistical Learning - Data Mining, Inference, and Prediction*, T. Hastie, et al., editores, Springer Science+Business Media, LLC, 2ª edición, 2009).

60 Puesto que normalmente hay muchos más valores de biomarcadores potenciales que las muestras en un conjunto de entrenamiento, debe tenerse cuidado para evitar el sobreajuste. Se produce sobreajuste cuando un modelo estadístico describe un error aleatorio o ruido en lugar de la relación subyacente. El sobreajuste puede evitarse de diversas maneras, incluyendo, por ejemplo, limitando el número de marcadores utilizados en el desarrollo del clasificador, suponiendo que las respuestas de marcadores son independientes entre sí, limitando la complejidad del modelo estadístico subyacente empleado y garantizando que el modelo estadístico subyacente se ajuste a los datos.

65 Con el fin de identificar un conjunto de biomarcadores asociados a la incidencia de eventos, el conjunto combinado de

muestras de control y eventos precoces se analizó usando análisis de componentes principales (PCA). El PCA muestra las muestras con respecto a los ejes definidos por las variaciones más fuertes entre todas las muestras, sin tener en cuenta el resultado del caso o del control, mitigando de este modo el riesgo de sobreajustar la distinción entre el caso y el control. Puesto que la incidencia de eventos trombóticos graves tiene un fuerte componente de probabilidad implicado, que requiere la notificación de ruptura de placa inestable en vasos vitales, no se esperaría observar una separación clara entre los conjuntos de muestras de control y de eventos. Aunque la separación observada entre el caso y el control no es grande, se produce en el segundo componente principal, que corresponde a aproximadamente el 10 % de la variación total en este conjunto de muestras, lo que indica que la variación biológica subyacente es relativamente simple de cuantificar (Figura 2A).

En el siguiente conjunto de análisis, pueden analizarse biomarcadores para aquellos componentes de diferencia entre muestras que eran específicas para la separación entre las muestras de control y las muestras de eventos precoces. Un método que puede emplearse es el uso de DSGA (Bair, E. y Tibshirani, R. (2004) *Semi-supervised methods to predict patient survival from gene expression data*. *PLOS Biol.*, 2, 511-522) para retirar (desinflar) las primeras tres direcciones de variación de componentes principales entre las muestras en el conjunto de control. Aunque la reducción de dimensionalidad se realiza en el conjunto de control para descubrir, tanto las muestras en el control como las muestras de las muestras de eventos precoces se ejecutan a través de PCA. Puede observarse separación de casos de eventos precoces a lo largo del eje horizontal (Figura 2B).

Selección validada cruzada de proteínas relevantes para el riesgo cardiovascular

Con el fin de evitar el sobreajuste del poder predictivo de proteínas a las características idiosincrásicas de una selección particular de muestras, se adoptó un enfoque de validación cruzada y reducción dimensional. La validación cruzada implica la selección múltiple de conjuntos de muestras para determinar la asociación de riesgo por proteína combinada con el uso de las muestras no seleccionadas para controlar la susceptibilidad del método para aplicarse a muestras que no se usaron para producir el modelo de riesgo (*The Elements of Statistical Learning - Data Mining, Inference, and Prediction*, T. Hastie, et al., editores, Springer Science+Business Media, LLC, 2a edición, 2009). Los inventores aplicaron el método PCA supervisado de Tibshirani et al (Bair, E. y Tibshirani, R. (2004) *Semi-supervised methods to predict patient survival from gene expression data*. *PLOS Biol.*, 2, 511-522) que es aplicable a conjuntos de datos de alta dimensión en el modelado de riesgo de eventos. El método de PCA supervisado (SPCA) implica la selección univariante de un conjunto de proteínas estadísticamente asociado al evento de peligro instantáneo observado en los datos y la determinación del componente correlacionado que combina información de todas estas proteínas. Esta determinación del componente correlacionado es una etapa de reducción de dimensionalidad que no solo combina información entre proteínas, sino que también mitiga la probabilidad de sobreajuste mediante la reducción del número de variables independientes del menú completo de proteínas de más de 1000 proteínas a unos pocos componentes principales (en este trabajo, los inventores solo examinaron el primer componente principal). El método SPCA de Tibshirani et al. aplica la validación cruzada a la selección de proteínas con el fin de determinar el número de proteínas que son realmente predictivas en los conjuntos de ensayos de validación cruzada retenidos. Estas proteínas se usan para crear el componente correlacionado de variación de proteínas asociado al riesgo de evento. Usando SPCA, los inventores encontraron una lista de 155 proteínas que se asociaron estadísticamente al riesgo de evento usando esta técnica de reducción dimensional con validación cruzada. En la presente aplicación de SPCA, los inventores también permitieron someter a ensayo señales de proteínas correspondientes a secuencias aleatorias de SOMAmoro, así como señales correspondientes a proteínas no humanas, que no estaban presentes en las muestras. Ninguna de estas 10-20 señales no biológicas conocidas se seleccionó en las 155 proteínas mediante SPCA (Tabla 1). Esta etapa que usa el enfoque de SPCA con validación cruzada es importante para detectar asociaciones de marcadores de proteína falsos positivos. El enfoque en Tibshirani et al está especialmente protegido contra el falso descubrimiento mediante el uso del método de prevalidación de validación cruzada y la reducción dimensional inherente en PCA. La lista de 155 proteínas de SPCA se usó para verificar análisis posteriores usando diferentes técnicas para detectar el falso descubrimiento de marcadores de proteína, no contenidos en la lista de 155 proteínas de SPCA.

Análisis univariante y análisis multivariante de la relación de proteínas individuales con el tiempo hasta el evento

El modelo de riesgo instantáneo proporcional de Cox (Cox, David R (1972). "*Regression Models and Life-Tables*". *Journal of the Royal Statistical Society. Serie B (Metodológica)* 34 (2): 187-220.) se usa ampliamente en estadística médica. La regresión de Cox evita ajustar una función específica de tiempo a la supervivencia acumulada y en su lugar emplea un modelo de riesgo relativo referido a una función de riesgo instantáneo de referencia (que puede variar con el tiempo). La función de riesgo instantáneo de referencia describe la forma común de la distribución del tiempo de supervivencia para todos los individuos, mientras que el riesgo relativo proporciona el nivel del riesgo instantáneo para un conjunto de valores covariantes (tales como un único individuo o grupo), como múltiplo del riesgo instantáneo basal. El riesgo relativo es constante con el tiempo en el modelo de Cox.

Los inventores ajustaron 1092 modelos de Cox univariantes simples a todas las señales. 46 proteínas (Tabla 2) tienen valores de P (ensayo de Wald, (Wald, Abraham. (1943). *A Method of Estimating Plane Vulnerability Based on Damage of Survivors*. *Statistical Research Group*, Universidad de Columbia) mejor que 10^{-14} . Todas estas 46 proteínas se incluyeron en la lista de 155 proteínas seleccionadas usando SPCA (anteriormente, Tabla 1). La gran cantidad de

proteínas altamente significativas es sorprendente al principio, sin embargo, la participación del riñón en la enfermedad cardiovascular implica cambios en la TFG (tasa de filtración glomerular). La disminución de la TFG aumentará todas las proteínas con aclaramiento renal distinto de cero, la concentración de una proteína en la sangre se reduce a través de la pérdida de la proteína en la orina a través del riñón (aclaramiento), la reducción de la filtración renal medida mediante la TFG se asocia, por tanto, a un aumento de la concentración de estas proteínas en la sangre que son parcialmente filtradas por el riñón.

Un modelo útil sería más parsimonioso que la lista completa de 46 proteínas. Como también se observa en el PCA es probable que muchas proteínas estén altamente correlacionadas, un modelo eficaz tendrá esto en cuenta. Los inventores filtraron la lista de 46 proteínas altamente significativas hasta 10 proteínas en dos etapas. En primer lugar, los inventores restringieron la lista a las 20 proteínas que proporcionaron un coeficiente con una magnitud mayor de 0,37 (equivalente a un cambio de riesgo instantáneo del 30 % para duplicar la señal de la proteína), esta etapa se tomó en un único modelo de Cox multivariante usando las 46 proteínas. (El logaritmo natural de las mediciones de proteínas se tomó antes de ajustar los modelos de Cox, por tanto, la exponencial del coeficiente de Cox corresponde a la relación de riesgos instantáneos de un cambio de e-veces (2,71) en la medición de proteínas).

La siguiente etapa filtró las 20 proteínas hasta diez requiriendo que el valor de p sea más significativo que 0,01. Esta etapa suprime las proteínas covariantes y permite que las proteínas independientes contribuyan. Se realizó un ajuste final a la selección de biomarcadores en que se consideró que C9, un miembro del complejo de ataque de membrana en la vía común final del sistema del complemento, era demasiado inespecífico en su señalización, un asunto que no puede decidirse a partir de este estudio solo, puesto que el estudio se creó para demostrar claramente el riesgo cardiovascular. Se retiró C9 y se evaluaron todas las proteínas restantes en su lugar. Las proteínas sustitutas se clasificaron basándose en la mejora en la puntuación del ensayo de Wald y KLK3.SerpinaA3 fue casi tan eficaz como C9.

Las curvas de supervivencia de Kaplan Meier se muestran en las Figuras 3A-3E para este modelo de diez marcadores de riesgo cardiovascular.

La Tabla 1 identifica 155 biomarcadores que son útiles para la evaluación del riesgo de un evento CV futuro en un individuo. Este es un número sorprendentemente mayor de lo esperado en comparación con lo que se encuentra normalmente durante los esfuerzos de descubrimiento de biomarcadores y puede atribuirse a la escala del estudio que se describe, que abarcó más de 1000 proteínas medidas en cientos de muestras individuales, en algunos casos a concentraciones en el intervalo femtomolar bajo. Presumiblemente, la gran cantidad de biomarcadores descubiertos refleja las diversas vías bioquímicas implicadas en la biología que conducen a un evento cardiovascular y la respuesta del cuerpo al evento CV; Cada vía y proceso implica muchas proteínas. Los resultados muestran que ninguna proteína individual de un pequeño grupo de proteínas es únicamente informativa sobre procesos tan complejos; más bien, que múltiples proteínas están implicadas en procesos relevantes, tales como TFG, aterosclerosis, inflamación y regulación CV hormonal, por ejemplo.

Los resultados del Ejemplo 2 sugieren determinadas conclusiones posibles: En primer lugar, la identificación de una gran cantidad de biomarcadores permite su agregación en una gran cantidad de clasificadores que ofrecen un rendimiento similarmente alto. En segundo lugar, los clasificadores pueden construirse de manera que los biomarcadores particulares puedan sustituirse por otros biomarcadores de una manera que refleje las redundancias que indudablemente impregnan las complejidades de los procesos subyacentes de enfermedades, afecciones o eventos. Es decir, la información sobre la enfermedad, afección o evento aportada por cualquier biomarcador individual identificado en la Tabla 1 se superpone con la información aportada por otros biomarcadores, de manera que puede ser que ningún biomarcador particular o pequeño grupo de biomarcadores en la Tabla 1 deba incluirse en ningún clasificador.

50 Kits

Cualquier combinación de los biomarcadores de la Tabla 1, Col. 7 puede detectarse usando un kit adecuado, tal como para su uso en la realización de los métodos que se desvelan en el presente documento. Además, cualquier kit puede contener uno o más marcadores detectables como se describen en el presente documento, tales como un resto fluorescente, etc.

En una realización, un kit incluye (a) uno o más reactivos de captura (tales como, por ejemplo, al menos un SOMAmero o anticuerpo) para detectar uno o más biomarcadores en una muestra biológica, en donde los biomarcadores incluyen cualquiera de los biomarcadores que se exponen en la Tabla 1, Col. 7 y, opcionalmente, (b) uno o más productos de software o programa informático para clasificar al individuo de quien se obtuvo la muestra biológica como que tiene o que no tiene un mayor riesgo de un evento CV o para determinar la probabilidad de que el individuo tenga un mayor riesgo de un evento CV, como se describe adicionalmente en el presente documento. Como alternativa, en lugar de uno o más productos de programa informático, pueden proporcionarse una o más instrucciones para que un ser humano realice manualmente las etapas anteriores.

La combinación de un soporte sólido con un reactivo de captura correspondiente que tiene un material generador de

señal se denomina en el presente documento "dispositivo de detección" o "kit". El kit también puede incluir instrucciones para usar los dispositivos y reactivos, manipular la muestra y analizar los datos. Adicionalmente, el kit puede usarse con un sistema informático o software para analizar y publicar el resultado del análisis de la muestra biológica.

5 Los kits también pueden contener uno o más reactivos (por ejemplo, tampones de solubilización, detergentes, lavados o tampones) para procesar una muestra biológica. Cualquiera de los kits que se describen en el presente documento también puede incluir, por ejemplo, tampones, agentes de bloqueo, materiales de matriz de espectrometría de masas, agentes de captura de anticuerpos, muestras de control positivo, muestras de control negativo, software e información
10 tales como protocolos, directrices y datos de referencia.

En un aspecto, la divulgación proporciona kits para el análisis del estado de riesgo de eventos CV. Los kits incluyen cebadores de PCR para uno o más SOMAmeros específicos para biomarcadores seleccionados entre la Tabla 1, Col. 7. Adicionalmente, el kit puede incluir instrucciones para el uso y la correlación de los biomarcadores con la predicción
15 de riesgo de un evento CV. El kit también puede incluir una matriz de ADN que contiene el complemento de uno o más de los Somameros específicos para los biomarcadores seleccionados entre la Tabla 1, Col. 7, reactivos y/o enzimas para amplificar o aislar ADN de muestra. Los kits pueden incluir reactivos para PCR en tiempo real, por ejemplo, cebadores y/o sondas TaqMan y enzimas.

20 Por ejemplo, un kit puede comprender (a) reactivos que comprendan al menos reactivo de captura para cuantificar uno o más biomarcadores en una muestra de ensayo, en donde dichos biomarcadores comprenden los biomarcadores que se exponen en la Tabla 1, Col. 7 o cualquier otro biomarcador o panel de biomarcadores que se describe en el presente documento y, opcionalmente, (b) uno o más algoritmos o programas informáticos para realizar las etapas de
25 comparar la cantidad de cada biomarcador cuantificado en la muestra de ensayo con uno o más valores de corte predeterminados y asignar una puntuación para cada biomarcador cuantificado basándose en dicha comparación, combinar las puntuaciones asignadas para cada biomarcador cuantificado para obtener una puntuación total, comparar la puntuación total con una puntuación predeterminada y usar dicha comparación para determinar si un individuo tiene un mayor riesgo de un evento CV. Como alternativa, en lugar de uno o más algoritmos o programas informáticos, pueden proporcionarse una o más instrucciones para que un ser humano realice manualmente las etapas
30 anteriores.

Métodos informáticos y software

Una vez que se selecciona un biomarcador o panel de biomarcadores, un método para diagnosticar a un individuo
35 puede comprender lo siguiente: 1) recoger u obtener de otro modo una muestra biológica; 2) realizar un método analítico para detectar y medir el biomarcador o biomarcadores en el panel en la muestra biológica; 3) realizar cualquier normalización o estandarización de datos requerida para el método utilizado para recoger valores de biomarcadores; 4) calcular la puntuación del marcador; 5) combinar las puntuaciones de los marcadores para obtener una puntuación diagnóstica o predictiva total; y 6) publicar la puntuación diagnóstica o predictiva del individuo. En este enfoque, la
40 puntuación diagnóstica o predictiva puede ser un número único determinado a partir de la suma de todos los cálculos de marcadores que se compara con un valor umbral preestablecido que es una indicación de la presencia o ausencia de enfermedad. O la puntuación diagnóstica o predictiva puede ser una serie de barras que representan cada una un valor de biomarcador y el patrón de las respuestas puede compararse con un patrón preestablecido para determinar la presencia o ausencia de enfermedad, afección o el aumento del riesgo (o no) de un evento.

45 Al menos algunas realizaciones de los métodos que se describen en el presente documento pueden implementarse con el uso de un ordenador. Un ejemplo de un sistema informático 100 se muestra en la Figura 4. Con referencia a la Figura 4, se muestra el sistema 100 compuesto por elementos de hardware que se acoplan eléctricamente a través del bus 108, incluyendo un procesador 101, dispositivo de entrada 102, dispositivo de salida 103, dispositivo de almacenamiento 104, lector de medios de almacenamiento legibles por ordenador 105a, sistema de comunicaciones 106, aceleración de procesamiento (por ejemplo, DSP o procesadores de fin especial) 107 y memoria 109. El lector de medios de almacenamiento legibles por ordenador 105a se acopla adicionalmente a medios de almacenamiento legibles por ordenador 105b, representando exhaustivamente la combinación dispositivos de almacenamiento remotos, locales, y/o extraíbles además de medios de almacenamiento, memoria, etc. para contener temporalmente
50 y/o más permanentemente información legible por ordenador, que pueden incluir un dispositivo de almacenamiento 104, una memoria 109 y/o cualquier otro de dicho recurso de sistema accesible 100. El sistema 100 también comprende elementos de software (que se muestran ubicados actualmente dentro de la memoria de trabajo 191) que incluyen un sistema operativo 192 y otro código 193, tal como programas, datos y similares.

60 Con respecto a la Figura 4, el sistema 100 tiene una amplia flexibilidad y capacidad de configuración. Por tanto, por ejemplo, podría utilizarse una arquitectura única para implementar uno o más servidores que se puedan configurar adicionalmente de acuerdo con los protocolos deseables actualmente, variaciones de protocolo, extensiones, etc. Sin embargo, será evidente para los expertos en la materia que pueden utilizarse realizaciones de acuerdo con requisitos de solicitud más específicos. Por ejemplo, podrían implementarse uno o más elementos del sistema como
65 subelementos dentro de un componente del sistema 100 (por ejemplo, dentro del sistema de comunicaciones 106). También podría utilizarse hardware personalizado y/o pueden implementarse elementos particulares en hardware,

software o en ambos. Adicionalmente, aunque puede emplearse la conexión a otros dispositivos informáticos, tales como dispositivos de entrada/salida de red (que no se muestran), ha de comprenderse que también podría utilizarse módem con cable, inalámbrico y/u otra conexión o conexiones a otros dispositivos informáticos.

5 En un aspecto, el sistema puede comprender una base de datos que contenga características de biomarcadores característicos de la predicción del riesgo de un evento CV. Los datos de biomarcadores (o información de biomarcadores) pueden utilizarse como una entrada al ordenador para su uso como parte de un método implementado por ordenador. Los datos de biomarcadores pueden incluir los datos como se describe en el presente documento.

10 En un aspecto, el sistema comprende adicionalmente uno o más dispositivos para proporcionar datos de entrada al uno o más procesadores.

El sistema comprende adicionalmente una memoria para almacenar un conjunto de datos de elementos de datos clasificados.

15 En otro aspecto, el dispositivo para proporcionar datos de entrada comprende un detector para detectar la característica del elemento de dato, por ejemplo, tal como un espectrómetro de masas o un lector de chips genéticos.

20 El sistema adicionalmente puede comprender un sistema de gestión de base de datos. Las solicitudes o consultas de los usuarios pueden formatearse en un lenguaje adecuado comprendido por el sistema de gestión de base de datos que procesa la consulta para extraer la información relevante de la base de datos de los conjuntos de entrenamiento.

25 El sistema puede conectarse a una red a la que están conectados un servidor de red y uno o más clientes. La red puede ser una red de área local (LAN) o una red de área amplia (WAN), como es sabido en la materia. Preferentemente, el servidor incluye el hardware necesario para ejecutar productos de programas informáticos (por ejemplo, software) para acceder a los datos de la base de datos para procesar las solicitudes de los usuarios.

30 El sistema puede incluir un sistema operativo (por ejemplo, UNIX o Linux) para ejecutar instrucciones a partir de un sistema de gestión de base de datos. En un aspecto, el sistema operativo puede operar en una red de comunicaciones global, tal como Internet, y utilizar un servidor de red de comunicaciones global para conectarse a una red de este tipo.

35 El sistema puede incluir uno o más dispositivos que comprendan una interfaz de representación gráfica que comprenda elementos de interfaz tales como botones, menús desplegados, barras de desplazamiento, campos para ingresar texto y similares, como se encuentran habitualmente en las interfaces gráficas de usuario conocidas en la técnica. Las solicitudes ingresadas en una interfaz de usuario pueden transmitirse a un programa de aplicación en el sistema para formatear y buscar información relevante en una o más de las bases de datos del sistema. Las solicitudes o consultas ingresadas por un usuario pueden construirse en cualquier lenguaje de base de datos adecuado.

40 La interfaz gráfica de usuario puede generarse mediante un código de interfaz gráfica de usuario como parte del sistema operativo y puede usarse para ingresar datos y/o representar datos ingresados. El resultado de los datos procesados puede representarse en la interfaz, imprimirse en una impresora en comunicación con el sistema, guardarse en un dispositivo de memoria y/o transmitirse a través de la red o puede proporcionarse en forma de medio legible por ordenador.

45 El sistema puede estar en comunicación con un dispositivo de entrada para proporcionar datos sobre elementos de datos al sistema (por ejemplo, valores de expresión). En un aspecto, el dispositivo de entrada puede incluir un sistema de realización de perfiles de expresión génica que incluya, por ejemplo, un espectrómetro de masas, lector de chip genético o matriz y similares.

50 Los métodos y aparatos para analizar información de biomarcadores de predicción de riesgo de eventos CV de acuerdo con diversas realizaciones pueden implementarse de cualquier manera adecuada, por ejemplo, usando un programa informático que opere en un sistema informático. Puede usarse un sistema informático convencional que comprenda un procesador y una memoria de acceso aleatorio, tal como un servidor de aplicación accesible remotamente, servidor de red, ordenador personal o estación de trabajo. Los componentes adicionales del sistema informático pueden incluir dispositivos de memoria o sistemas de almacenamiento de información, tales como un sistema de almacenamiento masivo y una interfaz de usuario, por ejemplo, un monitor convencional, teclado y dispositivo de rastreo. El sistema informático puede ser un sistema independiente o parte de una red de ordenadores que incluya un servidor y una o más bases de datos.

60 El sistema de análisis de biomarcadores de predicción de riesgo de eventos CV puede proporcionar funciones y operaciones para completar el análisis de datos, tales como la recopilación, el procesamiento, el análisis, la publicación y/o el diagnóstico de datos. Por ejemplo, en una realización, el sistema informático puede ejecutar el programa informático que puede recibir, almacenar, buscar, analizar y publicar información relacionada con los biomarcadores de predicción de riesgo de eventos CV. El programa informático puede comprender múltiples módulos que realicen diversas funciones u operaciones, tales como un módulo de procesamiento para procesar datos en bruto y generar datos suplementarios y un módulo de análisis para analizar datos en bruto y datos complementarios para generar un

cálculo de riesgo o diagnóstico y/o estado de predicción de riesgo de eventos CV. El cálculo de estado de riesgo para un evento CV puede incluir opcionalmente generar o recopilar cualquier otra información, incluir información biomédica adicional, sobre la condición del individuo con respecto a la enfermedad, afección o evento, identificar si pueden ser deseables ensayos adicionales o evaluar de otro modo el estado de salud del individuo.

5 Con referencia ahora a la Figura 5, puede observarse un ejemplo de un método de utilización de un ordenador de acuerdo con los principios de una realización desvelada. En la Figura 5, se muestra un diagrama de flujo 3000. En el bloque 3004, puede recuperarse información de biomarcadores para un individuo. La información de biomarcadores puede recuperarse de una base de datos informática, por ejemplo, después de realizar los ensayos de la muestra biológica del individuo. La información de biomarcadores puede comprender valores de biomarcadores que corresponden cada uno a uno de al menos N biomarcadores seleccionados entre un grupo que consiste en los biomarcadores proporcionados en la Tabla 1, Col. 7, en donde $N = 2-155$. En el bloque 3008, puede usarse un ordenador para clasificar cada uno de los valores de biomarcadores. Y, en el bloque 3012, puede realizarse una determinación en cuanto a la probabilidad de que un individuo tenga un riesgo aumentado de un evento CV basándose en una pluralidad de clasificaciones. La indicación puede enviarse a una pantalla u otro dispositivo indicador de manera que una persona pueda verla. Por tanto, por ejemplo, puede representarse en la pantalla de un ordenador u otro dispositivo de salida.

20 Con referencia ahora a la Figura 6, puede ilustrarse un método alternativo de utilización de un ordenador de acuerdo con otra realización a través del diagrama de flujo 3200. En el bloque 3204, puede utilizarse un ordenador para recuperar información de biomarcadores para un individuo. La información de biomarcadores comprende un valor de biomarcador correspondiente a un biomarcador seleccionado entre el grupo de biomarcadores proporcionado en la Tabla 1, Col.7. En el bloque 3208, puede realizarse una clasificación del valor de biomarcador con el ordenador. Y, en el bloque 3212, puede hacerse una indicación sobre la probabilidad de que el individuo tenga un riesgo aumentado de un evento CV basándose en la clasificación. La indicación puede enviarse a una pantalla u otro dispositivo indicador de manera que una persona pueda verla. Por tanto, por ejemplo, puede representarse en la pantalla de un ordenador u otro dispositivo de salida.

30 Algunas realizaciones que se describen en el presente documento pueden implementarse de manera de incluir un producto de programa informático. Un producto de programa informático puede incluir un medio legible por ordenador que tenga un código de programa legible por ordenador incorporado en el medio para provocar que un programa de aplicación se ejecute en un ordenador con una base de datos.

35 Como se usa en el presente documento, un "producto de programa informático" se refiere a un conjunto organizado de instrucciones en forma de declaraciones en lenguaje natural o de programación que están contenidas en un medio físico de cualquier naturaleza (por ejemplo, escrito, electrónico, magnético, óptico o de otro tipo) y que puede usarse con un ordenador u otro sistema automatizado de procesamiento de datos. Dichas declaraciones de lenguaje de programación, cuando son ejecutadas por un ordenador o sistema de procesamiento de datos, provocan que el ordenador o sistema de procesamiento de datos actúe de acuerdo con el contenido particular de las declaraciones. Los productos de programa informático incluyen, sin limitación: programas en código fuente y objeto y/o bibliotecas de ensayo o de datos integrados en un medio legible por ordenador. Además, el producto de programa informático que permite que un sistema informático o dispositivo de equipo de procesamiento de datos actúe de formas preseleccionadas puede proporcionarse en varias formas, incluyendo, pero sin limitación, código fuente original, código de ensamblaje, código objeto, lenguaje de máquina, versiones cifradas o comprimidas de lo anterior y cualquiera y todos los equivalentes.

50 En un aspecto, se proporciona un producto de programa informático para la evaluación del riesgo de un evento CV. El producto de programa informático incluye un medio legible por ordenador que incorpora un código de programa ejecutable por un procesador de un dispositivo o sistema informático, comprendiendo el código de programa: código que recupera datos atribuidos a una muestra biológica de un individuo, en donde los datos comprenden valores de biomarcadores que cada uno corresponde a uno de al menos N biomarcadores en la muestra biológica seleccionada entre el grupo de biomarcadores proporcionados en la Tabla 1, Col. 7, en donde $N = 2-155$; y un código que ejecuta un método de clasificación que indica un estado de riesgo de eventos CV del individuo en función de los valores de biomarcadores.

55 En otro aspecto adicional, se proporciona un producto de programa informático para indicar una probabilidad de riesgo de un evento CV. El producto de programa informático incluye un medio legible por ordenador que incorpora un código de programa ejecutable por un procesador de un dispositivo o sistema informático, comprendiendo el código de programa: código que recupera datos atribuidos a una muestra biológica de un individuo, en donde los datos comprenden un valor de biomarcador correspondiente a un biomarcador en la muestra biológica seleccionado entre el grupo de biomarcadores proporcionados en la Tabla 1, Col. 7; y código que ejecuta un método de clasificación que indica un estado de riesgo de eventos CV del individuo en función del valor de biomarcador.

65 Aunque se han descrito diversas realizaciones como métodos o aparatos, debe comprenderse que las realizaciones pueden implementarse a través de código acoplado a un ordenador, por ejemplo, código residente en un ordenador o accesible por el ordenador. Por ejemplo, podrían utilizarse software y bases de datos para implementar muchos de

los métodos analizados anteriormente. Por tanto, además de las realizaciones realizadas por hardware, también se observa que estas realizaciones pueden realizarse a través del uso de un artículo de fabricación compuesto por un medio utilizable por ordenador que tiene un código de programa legible por ordenador incorporado en el mismo, lo que provoca la habilitación de las funciones que se desvelan en la presente descripción. Por tanto, se desea que las realizaciones también se consideren protegidas por la presente patente en sus medios de código de programa también. Además, las realizaciones pueden realizarse como código almacenado en una memoria legible por ordenador de prácticamente cualquier tipo, incluyendo, sin limitación, RAM, ROM, medios magnéticos, medios ópticos o medios magneto-ópticos. Incluso más generalmente, las realizaciones podrían implementarse en software o en hardware o cualquier combinación de los mismos, incluyendo, pero sin limitación, software que se ejecuta en un procesador de uso general, microcódigo, PLA o ASIC.

También se prevé que las realizaciones podrían realizarse como señales informáticas incorporadas en una onda transportadora, así como señales (por ejemplo, eléctricas y ópticas) propagadas a través de un medio de transmisión. Por tanto, los diversos tipos de información analizados anteriormente podrían formatearse en una estructura, tal como una estructura de datos, y transmitirse como una señal eléctrica a través de un medio de transmisión, o almacenarse en un medio legible por ordenador.

También se observa que muchas de las estructuras, materiales y actos que se citan en el presente documento pueden citarse como medios para realizar una función o etapa para realizar una función. Por tanto, debe comprenderse que dicho lenguaje tiene por objeto cubrir todas esas estructuras, materiales o actos que se desvelan dentro de la presente memoria descriptiva y sus equivalentes.

El proceso de identificación de biomarcadores, la utilización de los biomarcadores que se desvelan en el presente documento y los diversos métodos para determinar valores de biomarcadores se han descrito en detalle anteriormente con respecto a la evaluación del riesgo de un evento CV. Sin embargo, la aplicación del proceso, el uso de biomarcadores identificados y los métodos para determinar valores de biomarcadores son completamente aplicables a otros tipos específicos de afecciones cardiovasculares, a cualquier otra enfermedad o afección médica o a la identificación de individuos que pueden o pueden no beneficiarse con un tratamiento médico auxiliar.

30 Ejemplos

Los siguientes ejemplos se proporcionan solamente con fines ilustrativos y no tienen por objeto limitar el alcance de la solicitud como se define por las reivindicaciones adjuntas. Todos los ejemplos que se describen en el presente documento se realizaron usando técnicas convencionales, que son bien conocidas y rutinarias para los expertos en la materia. Las técnicas de biología molecular rutinarias que se describen en los siguientes ejemplos pueden realizarse como se describe en los manuales de laboratorio convencionales, tales como Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 3ª ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., (2001).

40 Ejemplo 1. Análisis de aptámeros multiplexados de muestras

Este ejemplo describe el ensayo de aptámeros múltiples utilizado para analizar las muestras y los controles para la identificación de los biomarcadores que se exponen en la Tabla 1. El protocolo general para el análisis de una muestra se ilustra en las Figuras 1A y 1B. Habitualmente en estudios médicos de datos de supervivencia, el modelo de riesgo instantáneo proporcional de Cox se emplea para producir una puntuación de riesgo de covariantes múltiples de estado patológico. En este trabajo, los inventores han empleado este enfoque simple y bien conocido para diseñar un modelo a partir de los datos de población en el estudio *Heart and Soul*, por ejemplo, adecuado para la aplicación a muestras individuales de acuerdo con este Formalismo de Riesgo Instantáneo Proporcional de Cox flexible y ampliamente utilizado. Los valores de biomarcadores se combinan como se muestra en la Figura 1B tomando la relación logarítmica de las mediciones de biomarcadores con respecto a los niveles normales. El modelo de Cox usa la exponencial de la suma ponderada de estas relaciones logarítmicas para producir una estimación de la relación de riesgos instantáneos para la población normal.

En este método, se cambiaron las puntas de pipeta para cada adición de solución.

Además, a menos que se indique lo contrario, la mayoría de las transferencias de soluciones y adiciones de lavado usaron el cabezal de 96 pocillos de un Beckman Biomek FxP. Las etapas del método pipeteadas manualmente usaron un Pipeteman P200 de doce canales (Rainin Instruments, LLC, Oakland, CA), a menos que se indique lo contrario. Un tampón personalizado denominado SB 17 se preparó en el laboratorio, que comprendía HEPES 40 mM, NaCl 100 mM, KCl 5 mM, MgCl₂ 5 mM, EDTA 1 mM a pH 7,5. Todas las etapas se realizaron a temperatura ambiente a menos que se indique lo contrario.

65 1. Preparación de la solución madre de aptámero

Se prepararon soluciones madre personalizadas de aptámero para suero al 5 %, al 0,316 % y al 0,01 % a una concentración 2x en SB17 1x, Tween-20 al 0,05 %. Estas soluciones se almacenan a -20 °C hasta su uso. El día del ensayo, cada mezcla de aptámero se descongeló a

37 °C durante 10 minutos, se colocó en un baño de agua hirviendo durante 10 minutos y se dejó enfriar a 25 °C durante 20 minutos con mezcla vigorosa entre cada etapa de calentamiento. Después del calor-frío, 55 µl de cada mezcla de aptámero 2x se pipetearon manualmente en una placa de 96 pocillos Hybaid y la placa se selló con aluminio. El resultado final fue tres placas de 96 pocillos Hybaid selladas con papel de aluminio con mezclas de aptámero al 5 %, al 0,316 % o al 0,01 %. La concentración de aptámero individual fue 2x final o 1 nM.

2. Preparación de la muestra de ensayo

Se colocaron alícuotas congeladas de suero al 100 %, almacenadas a -80 °C, en un baño de agua a 25 °C durante 10 minutos. Se colocaron muestras descongeladas en hielo, se agitaron suavemente con formación de vórtice (fijado en 4) durante 8 segundos y después se reemplazaron en hielo.

Se preparó una solución de muestra al 10 % (final 2x) transfiriendo 8 µl de muestra usando un pipeteador de 8 canales de 50 µl en placas de 96 pocillos Hybaid, conteniendo cada pocillo 72 µl del diluyente de muestra adecuado a 4 °C (SB17 1x, Tween-20 al 0,06 %, Z-block_2 11,1 µM, MgCl₂ 0,44 mM, AEBSF 2,2 mM, EGTA 1,1 mM, EDTA 55,6 µM para suero). Esta placa se almacenó en hielo hasta que se iniciaron las siguientes etapas de dilución de muestra en el robot Biomek FxP.

Para comenzar el equilibrio de la muestra y el aptámero, la placa de muestra al 10 % se centrifugó brevemente y se colocó en el Biomek FxP donde se mezcló pipeteando hacia arriba y hacia abajo con el pipeteador de 96 pocillos. Después se preparó una placa de muestra al -0,632 % (final 2x) transfiriendo 6 µl de la placa de muestra al 10 % a 89 µl de 1xSB17, Tween-20 al 0,05 % con AEBSF 2 mM. A continuación, la dilución de 6 µl de la muestra resultante al 0,632 % en 184 µl de 1xSB17, Tween-20 al 0,05 % hizo una placa de muestra al 0,02 % (final 2x). Se hicieron diluciones en el Beckman Biomek FxP. Después de cada transferencia, las soluciones se mezclaron pipeteando hacia arriba y hacia abajo. Después, las 3 placas de dilución de muestra se transfirieron a sus respectivas soluciones de aptámero mediante la adición de 55 µl de la muestra a 55 µl de la mezcla adecuada de aptámero 2x. Las soluciones de muestra y aptámero se mezclaron en el robot pipeteando hacia arriba y hacia abajo.

3. Unión de equilibrio de muestra

Las placas de muestra/aptámero se sellaron con alfombrillas de tapones de silicona y se colocaron en una incubadora a 37 °C durante 3,5 horas antes de continuar con la etapa Captura 1.

4. Preparación de la placa de perlas Captura 2

Una alícuota de 11 ml de perlas de estreptavidina C1 MyOne (Invitrogen Corp., Carlsbad, CA) se lavó 2 veces con volúmenes iguales de NaOH 20 mM (5 minutos de incubación para cada lavado), 3 veces con volúmenes iguales de SB17 1x, Tween-20 al 0,05 % y se resuspendió en 11 ml de SB17 1x, Tween-20 al 0,05 %. Usando una pipeta de 12 canales, se pipetearon 50 µl de esta solución manualmente en cada pocillo de una placa de 96 pocillos Hybaid. Después, la placa se cubrió con papel de aluminio y se almacenó a 4 °C para su uso en el ensayo.

5. Preparación de placas de perlas Captura 1

Se equilibraron tres placas Millipore HV de 0,45 µm (membrana Durapore, n.º de Cat. MAHVN4550) con 100 µl de SB17 1x, Tween-20 al 0,05 % durante al menos 10 minutos. Después, el tampón de equilibrio se filtró a través de la placa y se añadieron 133,3 µl de una suspensión de perlas de estreptavidina-agarosa al 7,5 % (en SB17 1x, Tween-20 al 0,05 %) en cada pocillo. Para mantener las perlas de estreptavidina-agarosa suspendidas mientras se transfieren a la placa de filtro, la solución de perlas se mezcló manualmente con un pipeteador de 12 canales y 200 µl, al menos 6 veces entre los eventos de pipeteo. Después de que las perlas se distribuyesen a través de las 3 placas de filtro, se aplicó un vacío para retirar el sobrenadante de perlas. Por último, las perlas se lavaron en las placas de filtro con 200 µl de SB17 1x, Tween-20 al 0,05 % y después se resuspendieron en 200 µl de SB17 1x, Tween-20 al 0,05 %. Los fondos de las placas de filtro se transfirieron y las placas se almacenaron para su uso en el ensayo.

6. Carga del Cytomat

El cytomat se cargó con todas las puntas, placas, todos los reactivos en cubetas (excepto el reactivo NHS-biotina que se preparó nuevo justo antes de la adición a las placas), 3 placas de filtro captura 1 preparadas y 1 placa MyOne preparada.

7. Captura 1

Después de un tiempo de equilibrio de 3,5 horas, las placas de muestra/aptámero se retiraron de la incubadora, se centrifugaron durante aproximadamente 1 minuto, se retiraron las cubiertas de alfombrilla de tapones y se colocaron en la plataforma del Beckman Biomek FxP. Se inició el programa de Beckman Biomek FxP. Todas las etapas posteriores en Captura 1 fueron realizados por el robot Beckman Biomek FxP a menos que se indique lo contrario. Dentro del programa, el vacío se aplicó a las placas de filtro Captura 1 para retirar el sobrenadante de perlas. Se

añadieron cien microlitros de cada una de las reacciones de unión de equilibrio al 5 %, al 0,316 % y al 0,01 % a sus respectivas placas de filtración Captura 1 y cada placa se mezcló usando un agitador orbital en la plataforma a 800 rpm durante 10 minutos.

- 5 La solución no unida se retiró a través de filtración al vacío. Las perlas Captura 1 se lavaron con 190 µl de biotina 100 µM en SB17 1x, Tween-20 al 0,05 % seguido de 5 veces 190 µl de SB17 1x, Tween-20 al 0,05 % mediante dispensación de la solución y aspirando inmediatamente para filtrar la solución a través de la placa.

8. Marcado

- 10 Se descongeló una alícuota de NHS-PEO4-biotina 100 mM en DMSO anhidro (almacenado a -20 °C) a 37 °C durante 6 minutos y después se diluyó 1:100 con tampón de marcado (SB17 a pH = 7,25, Tween-20 al 0,05 %), inmediatamente antes de la adición manual a un canal en la plataforma por lo que el robot dispensó 100 µl de NHS-PEO4-biotina en cada pocillo de cada placa de filtro Captura 1. Se permitió que esta solución se incubase con perlas Captura 1 agitando a 800 rpm durante 5 minutos en los agitadores orbitales.

9. Exposición cinética y fotoescisión

- 20 La reacción de marcado se retiró por filtración al vacío y la reacción se interrumpió mediante la adición de 150 µl de glicina 20 mM en SB17 1x, Tween-20 al 0,05 % a las placas Captura 1. La solución de glicina se retiró a través de filtración al vacío y se añadieron otros 1500 µl de glicina 20 mM (en SB17 1x, Tween-20 al 0,05 %) a cada placa y se incubaron durante 1 minuto en agitadores orbitales a 800 rpm antes de retirarlas mediante filtración al vacío.

- 25 Los pocillos de las placas Captura 1 se lavaron posteriormente añadiendo 190 µl de SB17 1x, Tween-20 al 0,05 %, seguido inmediatamente de filtración al vacío y, después, añadiendo 190 µl de SB17 1x, Tween-20 al 0,05 % con agitación durante 1 minuto a 800 rpm antes de la filtración al vacío. Estas dos etapas de lavado se repitieron dos veces más con la excepción de que el último lavado no se retiró por filtración al vacío. Después del último lavado, las placas se colocaron sobre una placa de pocillos profundos de 1 ml y se retiraron de la plataforma para centrifugación a 1000 rpm durante 1 minuto para retirar tanta cantidad de volumen extraño de las perlas de agarosa antes de la elución como fuera posible.

Las placas se volvieron a colocar en el Beckman Biomek FxP y se añadieron 85 µl de DxSO4 10 mM en SB 17 1x, Tween-20 al 0,05 % a cada pocillo de las placas de filtro.

- 35 Las placas de filtro se retiraron de la plataforma, se colocaron en un termoagitador Variomag (Thermo Fisher Scientific, Inc., Waltham, MA) bajo las fuentes de luz BlackRay (Ted Pella, Inc., Redding, CA) y se irradiaron durante 5 minutos mientras se agitaban a 800 rpm. Después de la incubación de 5 minutos, las placas se giraron 180 grados y se irradiaron con agitación durante 5 minutos más.

- 40 Las soluciones fotoescindidas se eluyeron secuencialmente de cada placa Captura 1 en una placa común de pocillo profundo colocando en primer lugar la placa de filtro Captura 1 al 5 % en la parte superior de una placa de pocillo profundo de 1 ml y centrifugando a 1000 rpm durante 1 minuto. Después, las placas Captura 1 al 0,316 % y al 0,01 % se centrifugaron secuencialmente en la misma placa de pocillo profundo.

45 10. Captura de perlas Captura 2

El bloque de pocillos profundos de 1 ml que contenía los eluatos combinados de Captura 1 se colocó sobre la plataforma del Beckman Biomek FxP para Captura 2.

- 50 El robot transfirió todo el eluato fotoescindido del pocillo profundo de 1 ml placa sobre la placa Hybaid que contenía las perlas magnéticas MyOne Captura 2 preparadas anteriormente (después de la retirada del tampón MyOne a través de separación magnética).

- 55 La solución se incubó mientras se agitaba a 1350 rpm durante 5 minutos a 25 °C en un termoagitador Variomag (Thermo Fisher Scientific, Inc., Waltham, MA).

El robot transfirió la placa a la estación separadora magnética sobre la plataforma. La placa se incubó sobre el imán durante 90 segundos antes de retirar y desechar el sobrenadante.

60 11. Lavados de glicerol al 30 % a 37 °C

- 65 La placa Captura 2 se trasladó al agitador térmico sobre la plataforma y se transfirieron 75 µl de SB17 1x, Tween-20 al 0,05 % a cada pocillo. La placa se mezcló durante 1 minuto a 1350 rpm y 37 °C para resuspender y calentar las perlas. A cada pocillo de la placa captura 2, se le transfirieron 75 µl de glicerol al 60 % a 37 °C y la placa continuó mezclándose durante otro minuto a 1350 rpm y 3 °C. El robot transfirió la placa al separador magnético a 37 °C, donde se incubó en el imán durante 2 minutos y después el robot retiró y desechó el sobrenadante. Estos lavados se repitieron

dos veces más.

Después de la retirada del tercer lavado de glicerol al 30 % de las perlas Captura 2, se añadieron 150 µl de SB17 1x, Tween-20 al 0,05 % a cada pocillo y se incubaron a 37 °C, agitando a 1350 rpm durante 1 minuto, antes de la retirada mediante separación magnética en el imán a 37 °C.

Las perlas Captura 2 se lavaron por última vez usando 150 µl de SB19 1x, Tween-20 al 0,05 % con incubación durante 1 minuto mientras se agitaba a 1350 rpm, antes de la separación magnética.

12. Elución y neutralización de perlas Captura 2

Se eluyeron los aptámeros de las perlas Captura 2 mediante la adición de 105 µl de CAPSO 100 mM con NaCl 1 M, Tween-20 al 0,05 % a cada pocillo. Las perlas se incubaron con esta solución con agitación a 1300 rpm durante 5 minutos.

Después, la placa Captura 2 se colocó sobre el separador magnético durante 90 segundos antes de transferir 63 µl del eluato a una nueva placa de 96 pocillos que contenía 7 µl de HCl 500 mM, HEPES 500 mM, Tween-20 al 0,05 % en cada pocillo. Después de la transferencia, la solución se mezcló robóticamente pipeteando 60 µl hacia arriba y hacia abajo cinco veces.

13. Hibridación

El Beckman Biomek FxP transfirió 20 µl del eluato Captura 2 neutralizado a una placa Hybaid nueva y se añadieron 6 µl de Agilent Block 10x, que contenía una adición de 10x de controles de hibridación, a cada pocillo. A continuación, se pipetearon manualmente 30 µl de tampón de hibridación Agilent 2x a cada pocillo de la placa que contenía las muestras neutralizadas y tampón de bloqueo y la solución se mezcló pipeteando manualmente 25 µl hacia arriba y hacia abajo 15 veces lentamente para evitar la formación masiva de burbujas. La placa se hizo girar a 1000 rpm durante 1 minuto.

Se diseñaron portaobjetos de micromatrices Agilent personalizados (Agilent Technologies, Inc., Santa Clara, CA) para que contuviesen sondas complementarias a la región aleatoria del aptámero más alguna región cebadora. Para la mayoría de los aptámeros, la longitud óptima de la secuencia complementaria se determinó empíricamente y osciló entre 40-50 nucleótidos. Para los aptámeros posteriores, se eligió una región complementaria de 46meros por defecto. Las sondas se unieron a la superficie del portaobjetos con un enlazador de poli-T para una longitud total de la sonda de 60 nucleótidos.

Se colocó un portaobjetos de junta en una cámara de hibridación Agilent y se pipetearon manualmente 40 µl de cada una de las muestras que contenían solución de hibridación y bloqueo en cada junta. Se usó un pipeteador de expansión variable de 8 canales de una manera que tiene por objeto minimizar la formación de burbujas. Los portaobjetos personalizados Agilent, con el código de barras hacia arriba, se bajaron lentamente sobre los portaobjetos de junta (véase el manual de Agilent para obtener una descripción detallada).

La parte superior de las cámaras de hibridación se colocó sobre el sándwich portaobjetos/protector y los soportes de sujeción se deslizaron sobre todo el conjunto. Estos conjuntos se sujetaron firmemente girando los tornillos de forma segura.

Cada sándwich de portaobjetos/portaobjetos protector fue inspeccionado visualmente para garantizar que las burbujas de la solución puedan moverse libremente dentro de la muestra. Si las burbujas no se movían libremente, el conjunto de la cámara de hibridación se golpeó suavemente para desencajar las burbujas alojadas cerca de la junta.

Las cámaras de hibridación ensambladas se incubaron en un horno de hibridación Agilent durante 19 horas a 60 °C girando a 20 rpm.

14. Lavado posterior a la hibridación

Se colocaron aproximadamente 400 ml de tampón de lavado Agilent 1 en cada una de las dos placas de tinción de vidrio separadas. Una de las placas de tinción se colocó sobre una placa de agitación magnética y se colocaron una rejilla del portaobjetos y una barra de agitación en el tampón.

Se preparó una placa de tinción para Lavado de Agilent 2 colocando una barra de agitación en una placa de tinción de vidrio vacía.

Se colocó una cuarta placa de tinción de vidrio para el lavado final con acetonitrilo.

Cada una de las seis cámaras de hibridación se desmontó. Uno por uno, el sándwich de portaobjetos/protector se retiró de su cámara de hibridación y se sumergió en la placa de tinción que contenía Lavado 1. El sándwich de

portaobjetos/respaldo se separó con un par de pinzas, mientras aún se sumergía el portaobjetos de micromatrices. El portaobjetos se transfirió rápidamente a la rejilla de portaobjetos en el plato de tinción Lavado 1 sobre la placa de agitación magnética.

- 5 La rejilla de portaobjetos se levantó y se bajó suavemente 5 veces. El agitador magnético se encendió a un ajuste bajo y los portaobjetos se incubaron durante 5 minutos.

- 10 Cuando quedaba un minuto para el Lavado 1, se añadió Tampón de Lavado 2 precalentado a 37 °C en una incubadora a la segunda placa de tinción preparada. La rejilla de portaobjetos se transfirió rápidamente al Tampón de Lavado 2 y se retiró cualquier exceso de tampón en la parte inferior de la rejilla raspándolo en la parte superior de la placa de tinción. La rejilla de portaobjetos se levantó y se bajó suavemente 5 veces. El agitador magnético se encendió a un ajuste bajo y los portaobjetos se incubaron durante 5 minutos. La rejilla de portaobjetos se retiró lentamente del Lavado 2, tomando aproximadamente 15 segundos para retirar los portaobjetos de la solución.

- 15 Con un minuto restante en el Lavado 2 se añadió acetonitrilo (ACN) a la cuarta placa de tinción. La rejilla de portaobjetos se transfirió a la placa de tinción de acetonitrilo. La rejilla de portaobjetos se levantó y se bajó suavemente 5 veces. El agitador magnético se encendió a un ajuste bajo y los portaobjetos se incubaron durante 5 minutos.

- 20 La rejilla de portaobjetos se sacó lentamente de la placa de tinción de ACN y se colocó sobre una toalla absorbente. Los bordes inferiores de los portaobjetos se secaron rápidamente y el portaobjetos se colocó en una caja de portaobjetos limpia.

15. Formación de imágenes de micromatrices

- 25 Los portaobjetos de micromatrices se colocaron en los soportes de portaobjetos del escáner Agilent y se cargaron en el escáner Micromatriz Agilent de acuerdo con las instrucciones del fabricante.

- 30 Los portaobjetos se fotografiaron en el canal Cy3 a una resolución de 5 µm en la configuración de PMT al 100 % y la opción de XRD habilitada a 0,05. Las imágenes tiff resultantes se procesaron usando el software de extracción de características Agilent versión 10.5.

Ejemplo 2. Identificación de biomarcadores

- 35 La identificación de posibles biomarcadores de eventos CV se realizó para predecir el riesgo de un evento CV en una población de individuos en el área de la Bahía de San Francisco. Los participantes tenían que cumplir con uno de los siguientes criterios de inscripción para este estudio: infarto de miocardio previo, evidencia angiográfica de una estenosis superior al 50 % en 1 o más vasos coronarios, isquemia inducida por ejercicio mediante cinta de correr o ensayos nucleares, o revascularización coronaria previa. Los criterios de exclusión incluían infarto de miocardio reciente, incapacidad para caminar aproximadamente 1 manzana y planes de reubicación. Se recogieron muestras de sangre en ayunas y se almacenaron alícuotas de suero y plasma a -70 °C. El ensayo de afinidad de SOMAmero multiplexado como se describe en el Ejemplo 1 se usó para medir y publicar el valor de RFU para 1034 analitos en cada una de estas 987 muestras.

- 45 Con el fin de identificar un conjunto de biomarcadores asociados a la incidencia de eventos, el conjunto combinado de muestras de control y eventos precoces se analizó mediante PCA. El PCA muestra las muestras con respecto a los ejes definidos por las variaciones más fuertes entre todas las muestras, sin tener en cuenta el resultado del caso o del control, mitigando de este modo el riesgo de sobreajustar la distinción entre el caso y el control. ("Control" se refiere a individuos que cumplían al menos uno de los criterios de ingreso, pero que no tenían un evento CV durante el curso del estudio; "caso" se refiere a individuos que cumplían al menos uno de los criterios de ingreso, pero que tenía un evento CV durante el curso del estudio). Aunque la separación observada entre el caso y el control no es grande, se produce en el segundo componente principal, que corresponde a aproximadamente el 10 % de la variación total en este conjunto de muestras, lo que indica que la variación biológica subyacente es relativamente simple de cuantificar (Figura 2A).

- 55 En el siguiente conjunto de análisis, pueden analizarse biomarcadores para aquellos componentes de diferencia entre muestras que eran específicas para la separación entre las muestras de control y las muestras de eventos precoces. Aunque la reducción de la dimensionalidad se realiza en el conjunto de control solo, para determinar el espacio de variación multidimensional multivariante abarcado por las diferencias entre las muestras de control, tanto las muestras en el conjunto de control como el conjunto de muestras de eventos precoces se desinflan para el espacio de variación determinado entre muestras de control, la variación residual se enriquece en aquellos componentes que separan el caso del control. Esto se conoce como método DSGA. La separación de casos de eventos precoces puede observarse a lo largo del eje horizontal (Figura 2B) (Nicolau M, Tibshirani R, Børresen-Dale AL, Jeffrey SS. *Disease-specific genomic analysis: Identifying the signature of pathologic biology*. *Bioinformatics*. 2007; 23: 957-965).

- 65 Con el fin de evitar el sobreajuste del poder predictivo de proteínas a las características idiosincrásicas de una selección particular de muestras, se adoptó un enfoque de validación cruzada y reducción dimensional. Esta

determinación del componente correlacionado es una etapa de reducción de dimensionalidad que no solo combina información entre proteínas, sino que también mitiga la probabilidad de sobreajuste mediante la reducción del número de variables independientes del menú completo de proteínas de más de 1000 proteínas a unos pocos componentes principales (en este trabajo, los inventores solo examinaron el primer componente principal). Estas proteínas se usan para crear el componente correlacionado de variación de proteínas asociado al riesgo de evento. Usando SPCA (Análisis de Componente Principal Supervisado), los inventores encontraron una lista de 155 proteínas que se asociaron estadísticamente al riesgo de evento usando esta técnica de reducción dimensional con validación cruzada. En la presente aplicación de SPCA, también se permitieron señales de proteínas de ensayo correspondientes a secuencias aleatorias de SOMAmero, así como señales correspondientes a proteínas no humanas, que no estaban presentes en las muestras. Ninguna de estas 10-20 señales no biológicas conocidas se seleccionó en las 155 proteínas por SPCA (Tabla 1). Esta etapa que usa el enfoque de SPCA con validación cruzada es importante para detectar asociaciones de marcadores de proteína falsos positivos. El enfoque en Tibshirani et al. (Bair, E. y Tibshirani, R. (2004) *Semi-supervised methods to predict patient survival from gene expression data. PLOS Biol.*, 2, 511-522) está especialmente protegido contra el falso descubrimiento mediante el uso del método de prevalidación de la validación cruzada y la reducción dimensional inherente en PCA. La lista de 155 proteínas de SPCA se usó para verificar análisis posteriores usando diferentes técnicas para detectar el falso descubrimiento de marcadores de proteína, no contenidos en la lista de 155 proteínas de SPCA.

Ejemplo 3. Análisis univariante de la relación de proteínas individuales con el tiempo con el evento CV

El modelo de riesgo instantáneo proporcional de Cox (Cox, David R (1972). "*Regression Models and Life-Tables*". *Journal of the Royal Statistical Society. Serie B (Metodológica)* 34 (2): 187-220)) es ampliamente utilizado en estadística médica. La regresión de Cox evita ajustar una función específica de tiempo a la supervivencia acumulada y en su lugar emplea un modelo de riesgo relativo referido a una función de riesgo instantáneo de referencia (que puede variar con el tiempo). La función de riesgo instantáneo de referencia describe la forma común de la distribución del tiempo de supervivencia para todos los individuos, mientras que el riesgo relativo proporciona el nivel del riesgo instantáneo para un conjunto de valores covariantes (tales como un único individuo o grupo), como múltiplo del riesgo instantáneo basal. El riesgo relativo es constante con el tiempo en el modelo de Cox.

El método consistió en ajustar 1092 modelos de Cox univariantes simples a todas las señales. Cuarenta y seis proteínas tienen valores de P (Wald, Abraham. (1943). *A Method of Estimating Plane Vulnerability Based on Damage of Survivors. Statistical Research Group*, Universidad de Columbia) mejor que 10^{-14} . La gran cantidad de proteínas altamente significativas es sorprendente al principio, sin embargo, la implicación del riñón en la enfermedad cardiovascular implica cambios en la tasa de filtración glomerular (TFG). La disminución de la TFG aumentará todas las proteínas con un aclaramiento renal distinto de cero.

Un modelo útil (en términos de complejidad técnica y coste en el laboratorio) sería más parsimonioso que la lista completa de 46 proteínas que se muestra en la Tabla 2. Además, como se observa en el PCA, muchas proteínas probablemente estén altamente correlacionadas; un modelo eficaz tendrá esto en cuenta. La lista de 46 proteínas altamente significativas se filtró a 10 proteínas como se muestra en la Tabla 3 en dos etapas. En primer lugar, la lista se restringió a las 20 proteínas que proporcionaron un coeficiente con una magnitud superior a 0,37 (equivalente a un cambio de riesgo instantáneo del 30 % para duplicar la señal de proteína). Esta etapa se tomó en un único modelo de Cox multivariante usando las 46 proteínas. (El logaritmo natural de las mediciones de proteínas se tomó antes de ajustar los modelos de Cox. Por tanto, el exponencial del coeficiente de Cox corresponde a la relación de riesgos instantáneos de un cambio de e-veces (2,71) en la medición de proteínas).

La siguiente etapa filtró las 20 proteínas hasta nueve requiriendo que el valor P sea más significativo que 0,01. Esta etapa suprime las proteínas covariantes y permite que las proteínas independientes contribuyan. Se realizó un ajuste final a la selección de biomarcadores en que se consideró que C9, un miembro del complejo de ataque de membrana en la vía común final del sistema del complemento, era demasiado inespecífico en su señalización, un asunto que no puede decidirse a partir de este estudio solo, puesto que el estudio se creó para demostrar claramente el riesgo de evento CV. Se retiró C9 y se evaluaron todas las proteínas restantes en su lugar. Las proteínas sustitutas se clasificaron basándose en la mejora en la puntuación del ensayo de Wald y KLK3.SerpinaA3 fue casi tan eficaz como C9.

Las curvas de supervivencia de Kaplan Meier se muestran en las Figuras 3A-4F para este modelo de diez marcadores de riesgo cardiovascular (Tabla 3).

Ejemplo 4. Análisis univariante de la relación de proteínas individuales con el tipo específico de evento CV

Los eventos cardiovasculares se dividen en dos clases: trombóticos y de ICC. Distinguir entre riesgo trombótico y de ICC tiene utilidad médica para guiar la terapia, eligiendo entre medicamentos antitrombóticos y diuréticos, por ejemplo. Aunque gran parte de la biología se comparte entre las clases de eventos trombóticos y de ICC, los eventos trombóticos implican específicamente la biología de la coagulación de la sangre (como implica el nombre trombótico). Usando las diez proteínas de la Tabla 3 identificadas en el modelo de riesgo instantáneo proporcional de Cox (Ejemplo 3), fue posible buscar las señales vinculadas a la coagulación y las señales vinculadas a la remodelación del tejido.

Para determinar cualquier señal diferencial entre la ICC y los eventos trombóticos, las curvas de Kaplan Meier relevantes se trazaron por separado para los eventos de ICC y trombóticos.

Las plaquetas o trombocitos tienen una función clave en la biología de la coagulación. GPVI es una glucoproteína de membrana plaquetaria y, para esta proteína, se realizó un análisis de la asociación con la supervivencia sin eventos tanto para eventos de ICC como para eventos trombóticos en la Figura 8. La Figura 8A muestra una fuerte asociación de supervivencia sin eventos trombóticos con el nivel de GPVI, trazada como cuartiles de la distribución de la población. La Figura 8B muestra que los cuartiles de GPVI no se asocian a la supervivencia sin eventos para eventos de ICC.

A diferencia de GPVI, MATN2 (matrilina 2) es una proteína asociada a la matriz extracelular. La Figura 9A muestra que los cuartiles de MATN2 no se asocian al riesgo de eventos trombóticos, mientras que la Figura 9B muestra una fuerte asociación entre MATN2 y los eventos de ICC. La supervivencia sin eventos para aquellos individuos con una MATN2 en el 4to cuartil de la población es notablemente peor que los primeros tres cuartiles.

Tomados en conjunto, los resultados de este ejemplo demuestran que los diez marcadores de proteínas de los inventores pueden discriminar entre eventos trombóticos y eventos de ICC en términos del riesgo durante algunos años después del muestreo de sangre.

Ejemplo 5. Utilidad de angiopoyetina 2 y CHRDL1 en individuos medicados con estatinas.

Muchos individuos están medicados con estatinas, con enfermedad cardiovascular conocida, y muchos sin afecciones cardiovasculares específicas, tales como aquellos con colesterol LDL alto. Estos fármacos potentes alteran la capacidad de muchos biomarcadores para discriminar aquellos en riesgo de aquellos que no lo están. Es valioso que los biomarcadores funcionen en esta población.

La Figura 10 muestra que, en aquellos sujetos con estatinas, la angiopoyetina 2 sigue siendo muy útil para la predicción de un evento CV en individuos de alto riesgo. La Figura 10 muestra gráficas de Kaplan Meier de los 538 sujetos que toman medicamentos con estatinas e ilustra que aquellos individuos en el 4to cuartil de la distribución de la población para angiopoyetina-2, padecen eventos cardiovasculares a una tasa mayor en comparación con aquellos que no están en el 4to cuartil para angiopoyetina-2. Por tanto, a pesar de los efectos del tratamiento con estatinas, la angiopoyetina-2 es un biomarcador útil del riesgo de eventos cardiovasculares.

La Figura 11 muestra que, en aquellos sujetos con estatinas, la CHRDL1 también se asocia al riesgo de eventos cardiovasculares en esta población de alto riesgo. La Figura 11 muestra las gráficas de Kaplan Meier de los 538 sujetos que toman medicamentos con estatinas. Ilustra que CHRDL1 se asocia a la supervivencia sin eventos cardiovasculares en individuos tratados con medicamentos con estatinas. Por tanto, a pesar de los efectos del tratamiento con estatinas, CHRDL1 es un biomarcador útil del riesgo de eventos cardiovasculares.

Tabla 1: Biomarcadores de eventos CV

	Designación de SOMAmero	Nombre del gen	Diana	Id de Prot de Swiss	Id de gen de Entrez	NOMBRE_PÚBLICO	Ascendente o descendente
1	CRP.4337.49.2	CRP	CRP	P02741	1401	Proteína C reactiva	ascendente
2	IGFBP1.2771.35.2	IGFBP1	IGFBP-1	P08833	3484	Proteína de unión 1 al factor de crecimiento similar a insulina	ascendente
3	IGFBP2.2570.72.5	IGFBP2	IGFBP-2	P18065	3485	Proteína de unión 2 al factor de crecimiento similar a insulina	ascendente
4	B2M.3485.28.2	B2M	Microglobulina b2	P61769	567	Microglobulina β 2	ascendente
5	TNFRSF1B.3152.57.1	TNFRSF 1B	R-II de TNF	P20333	7133	Miembro 1B de la superfamilia de receptores del factor de necrosis tumoral	ascendente
6	CFD.2946.52.2	CFD	Factor D	P00746	1675	Factor del complemento D	ascendente
7	FCER2.3291.30.2	FCER2	CD23	P06734	2208	CD23	ascendente
8	LYZ.4920.10.1	LYZ	Lisozima	P61626	4069	Lisozima	ascendente
9	FSTL3.3438.10.2	FSTL3	FSTL3	095633	10272	Similar a follistatina 3	ascendente
10	ANGPT2.2602.2.2	ANGPT2	Angiopoyetina-2	015123	285	Angiopoyetina-2	ascendente
11	PI3.4982.54.1	PI3	Elafina	P19957	5266	Elafina	ascendente
12	MMP7.2789.26.2	MMP7	MMP-7	P09237	4316	Metaloproteinasa de matriz 7/Matrilisina	ascendente
13	PLA2G2A.2692.74.2	PLA2G2A	NPS-PLA2	P14555	5320	Fosfolipasa A2, Grupo IIA	ascendente
14	LCN2.2836.68.2	LCN2	Lipocalina 2	P80188	3934	Lipocalina 2	ascendente
15	PAPPA.4148.49.2	PAPPA	PAPP-A	Q13219	5069	Proteína plasmática asociada al embarazo A	ascendente
16	RETN.3046.31.1	RETN	resistina	Q9HD8 9	56729	Resistina	ascendente
17	CCL14.2900.53.3	CCL14	HCC-1	Q16627	6358	Quimiocina 1 CC de hemofiltrado /CCL14	ascendente
18	PRSS2.5034.79.1	PRSS2	Tripsina 2	P07478	5645	Tripsina-2	ascendente
19	CAPG.4968.50.1	CAPG	CAPG	P40121	822	Proteína de protección de macrófagos	ascendente
20	TFF3.4721.54.2	TFF3	TFF3	Q07654	7033	Factor de trébol 3	ascendente

(continuación)

	Designación de SOMAmero	Nombre del gen	Diana	Id de Prot de Swiss	Id de gen de Entrez	NOMBRE_PÚBLICO	Ascendente o descendente
21	CHRD1.3362.61.2	CHRD1	CRDL1	Q9BU4 0	91851	Similar a cordina 1	ascendente
22	PPY.4588.1.2	PPY	PH	P01298	5539	Hormona pancreática	ascendente
23	GP6.3194.36.2	GP6	GPVI	Q9HCN 6	51206	GPVI/Gluco proteina plaquetaria VI	ascendente
24	CNDP1.3604.6.4	CNDP1	CNDP1	Q96KN 2	84735	Carnosina dipeptidasa 1	descendente
25	CA3.3799.11.2	CA3	Anhidrasa carbónica III	P07451	761	Anhidrasa carbónica III	ascendente
26	CCL15.3509.1.1	CCL15	MIP-5	Q16663	6359	Proteina inflamatoria de macrófagos 5/CCL15	ascendente
27	TIMP1.2211.9.6	TIMP1	TIMP-1	P01033	7076	Inhibidor tisular de metaloproteinasas 1	ascendente
28	CTSH.3737.6.3	CTSH	Catepsina H	P09668	1512	Catepsina H	ascendente
29	FCGR2B.3310.62.1	FCGR2B	FCG2B	P31994	2213	Receptor II-b de la región Fc de inmunoglobulina G, baja afinidad	ascendente
30	AKR1A1.4192.10.2	AKR1A1	AK1A1	P14550	10327	Miembro A1 de la familia 1 de Alcohol deshidrogenasa (NADP+)/Aldo-ceto reductasa	ascendente
31	PGLYRP1.3329.14.2	PGLYRP1	PGRP-S	075594	8993	Proteína de reconocimiento peptidoglucano, corta	ascendente
32	ROR1.2590.69.4	ROR1	ROR1	Q01973	4919	Receptor transmembrana de proteina tirosina cinasa ROR1	ascendente
33	ESAM.2981.9.3	ESAM	ESAM	Q96AP 7	90952	Molécula de adhesión selectiva de células endoteliales	ascendente
34	VCAM1.2967.8.1	VCAM1	VCAM-1	P19320	7412	Proteina 1 de adhesión de células vasculares/CAM 1	ascendente
35	CD5L.3293.2.3	CD5L	CD5L	043866	922	Antígeno similar a CD5	ascendente
36	COL18A1.2201.17.6	COL18A1	Endostatina	P39060	80781	Endostatina	ascendente
37	CTSZ.4971.1.1	CTSZ	CATZ	Q9UBR2	1522	Catepsina Z	ascendente
38	CCL18.3044.3.2	CCL18	PARC	P55774	6362	Proteina inflamatoria de macrófagos 4/Quimiocina pulmonar y regulada por activación/CCL18	ascendente

(continuación)

	Designación de SOMAmero	Nombre del gen	Diana	Id de Prot de Swiss	Id de gen de Entrez	NOMBRE_PÚBLICO	Ascendente o descendente
39	IGFBP4.2950.57.2	IGFBP4	IGFBP-4	P22692	3487	Proteína de unión 4 al factor de crecimiento similar a insulina	ascendente
40	PLAUR.2652.15.1	PLAUR	suPAR	Q03405	5329	Receptor de superficie de activador de plasminógeno de urocina	ascendente
41	IL16.2774.10.3	IL16	IL-16	Q14005	3603	Interleucina-16	ascendente
42	THBS2.3339.33.1	THBS2	TSP2	P35442	7058	Trombospondina-2	ascendente
43	IGFBP6.2686.67.2	IGFBP6	IGFBP-6	P24592	3489	Proteína de unión 6 al factor de crecimiento similar a insulina	ascendente
44	TNFRSF1A.2654.19.1	TNFRSF1A	R-I de TNF	P19438	7132	Miembro 1A de la superfamilia del receptor del factor de necrosis tumoral	ascendente
45	MATN2.3325.2.2	MATN2	MATN2	000339	4147	Matrilina-2	ascendente
46	MMP1.4924.32.1	MMP1	MMP-1	P03956	4312	Metaloproteinasa de matriz 1/colagenasa 1	ascendente
47	IGF1.2952.75.2	IGF1	IGF-I	P05019	3479	Factor de crecimiento similar a la insulina I	descendente
48	CRK.4976.57.1	CRK	CRK	P46108	1398	Proteína adaptadora Crk-1	ascendente
49	MB.3042.7.2	MB	Mioglobina	P02144	4151	Mioglobina	ascendente
50	SLPI.4413.3.2	SLPI	SLPI	P03973	6590	Inhibidor de la proteasa leucocitaria secretora	ascendente
51	IL18BP.3073.51.2	IL18BP	IL-18 BPa	095998	10068	Proteína de unión de interleucina 18	ascendente
52	IL1RL1.4234.8.2	IL1RL1	IL-1 R4	Q01638	9173	Receptor 4 de interleucina-1	ascendente
53	F3.4931.59.1	F3	TF	P13726	2152	Factor tisular	ascendente
54	STC1.4930.21.1	STC1	Estafocalcina-1	P52823	6781	Estafocalcina-1	ascendente
55	ADIPOQ.3554.24.1	ADIPOQ	Adiponectina	Q15848	9370	Adiponectina	ascendente
56	PROC.2961.1.2	PROC	Proteína C	P04070	5624	Proteína C	descendente
57	REN.3396.54.2	REN	Renina	P00797	5972	Renina	ascendente
58	CCL23.2913.1.2	CCL23	MPIF-1	P55773	6368	Factor 1 inhibidor de progenitoras mieloides/CCL23	ascendente

(continuación)

	Designación de SOMAmero	Nombre del gen	Diana	Id de Prot de Swiss	Id de gen de Entrez	NOMBRE_PÚBLICO	Ascendente o descendente
59	LBP.3074.6.2	LBP	LBP	P18428	3929	Proteína de unión a lipopolisacáridos	ascendente
60	GCG.4891.50.1	GCG	Glucagón	P01275	2641	Glucagón	ascendente
61	YWHAG.4179.57.3	YWHAG	Proteína gamma 14-3-3	P61981	7532	Proteína Y 14-3-3	ascendente
62	CCDC80.3234.23.2	CCDC80	URB	Q76M96	15188 7	Proteína 80 que contiene dominio de enrollamiento enrollado	ascendente
63	CNTFR.2711.6.2	CNTFR	CNTFR alfa	P26992	1271	Receptor α del factor neurotrófico ciliar	ascendente
64	EFNA5.2615.60.2	EFNA5	Efrina-A5	P52803	1946	Efrina-A5	ascendente
65	CST3.2609.59.2	CST3	Cistatina C	P01034	1471	Cistatina C	ascendente
66	FUT5.4549.78.2	FUT5	FUT5	Q11128	2527	Fucosiltransferasa 5	ascendente
67	TNFRSF17.2665.26.2	TNFRSF 17	BCMA	Q02223	608	Proteína de maduración de células B	ascendente
68	ERP29.4983.6.1	ERP29	ERP29	P30040	10961	Proteína 29 residente del retículo endoplasmático	ascendente
69	RARRES2.3079.62.2	RARRES 2	TIG2	Q99969	5919	Quemerina	ascendente
70	MAP2K2.3628.3.4	MAP2K2	MP2K2	P36507	5605	MAPK cinasa 2	ascendente
71	EPHA1.3431.54.2	EPHA1	EphA1	P21709	2041	Receptor 1 de efrina de tipo A	ascendente
72	CLEC11A.4500.50.2	CLEC11A	SCGF-alfa	Q9Y240	6320	Factor α de crecimiento de células madre	ascendente
73	F10.4878.34.1	F10	Factor de coagulación X	P00742	2159	Factor de coagulación X	descendente
74	CHIT1.3600.2.3	CHIT1	Quitotriosidasa -1	Q13231	1118	Quitotriosidasa-1	ascendente
75	ITGA1.ITGB1..3503.4.2	ITGA1 ITGB 1	Integrina α 1b1	P56199, P05556	3672, 3688	Integrina α -1: complejo β -1	ascendente
76	CKB.CKM..3714.49.2	CKB CKM	CK-MB	P12277, P06732	1152, 1158	Creatina cinasa-MB	descendente
77	CTSB.3061.61.2	CTSB	Catepsina B	P07858	1508	Catepsina B	ascendente

(continuación)

	Designación de SOMAmero	Nombre del gen	Diana	Id de Prot de Swiss	Id de gen de Entrez	NOMBRE_PÚBLICO	Ascendente o descendente
78	CD163.5028.59.1	CD163	SCD163	Q86VB 7	9332	Cadena M130 de proteína de tipo 1 rica en cisteína de receptor eliminador/CD163 soluble	ascendente
79	PTN.3045.72.2	PTN	PTN	P21246	5764	Pleiotrofina	ascendente
80	CSF1R.2638.12.2	CSF1R	M-CSFR	P07333	1436	receptor del factor 1 estimulante de colonias de macrófagos	ascendente
81	IGFBP3.2571.12.3	IGFBP3	IGFBP-3	P17936	3486	Proteína de unión 3 al factor de crecimiento similar a insulina	descendente
82	CXCL12.3516.60.2	CXCL12	SDF-1b	P48061	6387	Factor derivado de células estromales 1β	ascendente
83	PEBP1.4276.10.2	PEBP1	Proteína de unión prostática	P30086	5037	Proteína 1 de unión a fostatidiletanolamina	ascendente
84	CCL22.3508.78.3	CCL22	MDC	000626	6367	Quimiocina derivada de macrófagos	ascendente
85	CST2.4324.33.2	CST2	CYTT	P09228	1470	Cistatina SA	ascendente
86	CCL23.3028.36.2	CCL23	Ck-b-8-1	P55773	6368	Ck-β-8-1/Corte y empalme de proteína 3 inflamatoria de macrófagos	ascendente
87	CADM1..3326.58.2	CADM1	Proteína 2 similar a la nectina	Q9BY6 7	23705	Proteína 2 similar a la nectina	ascendente
88	VWF.3050.7.2	VWF	vWF	P04275	7450	Factor de von Willebrand	ascendente
89	CST5.3803.10.2	CST5	CYTD	P28325	1473	Cistatina D	ascendente
90	CHST15.4469.78.2	CHST15	ST4S6	Q7LFX 5	51363	Sulfotransferasa de hidrato de carbono 15	ascendente
91	IL1R1.2991.9.2	IL1R1	RI de IL-1	P14778	3554	Receptor 1 de interleucina-1	ascendente
92	C9.3060.43.2	C9	C9	P02748	735	Complemento C9	ascendente
93	AFM.4763.31.3	AFM	Afamina	P43652	173	Afamina	descendente
94	TGFBR3.3009.3.2	TGFBR3	R III de TGF-b	Q03167	7049	Receptor β de tipo III de factor de crecimiento transformante	ascendente

(continuación)

	Designación de SOMAmero	Nombre del gen	Diana	Id de Prot de Swiss	Id de gen de Entrez	NOMBRE_PÚBLICO	Ascendente o descendente
95	FRZB.2841.13.2	FRZB	sFRP-3	Q92765	2487	Proteína 3 relacionada con Frizzled, secretada	ascendente
96	MMP12.4496.60.2	MMP12	MMP-12	P39900	4321	Metaloproteinasa de matriz 12/Metaloelastasa de macrófagos	ascendente
97	CD33.3166.92.1	CD33	Siglec-3	P20138	945	Siglec-3	ascendente
98	EPHA2.4834.61.2	EPHA2	Cinasa de células epiteliales	P29317	1969	Receptor 2 de efrina de tipo A	ascendente
99	CXCL16.2436.49.4	CXCL16	CXCL16, soluble	Q9H2A 7	58191	Receptor eliminador para fosfatidilserina y lipoproteína de baja densidad oxidada/CXCL16	ascendente
100	IGFBP7.3320.49.2	IGFBP7	IGFBP-7	Q16270	3490	Proteína de unión 7 al factor de crecimiento similar a insulina	ascendente
101	KLK3.SERPINA 3.4153.11.2	KLK3 SERPINA A3	PSA-ACT	P07288, P01011	354 12	PSA: complejo de α -1-antiquimiotripsina	ascendente
102	PRTN3.3514.49.2	PRTN3	Proteinasa-3	P24158	5657	Proteinasa-3	ascendente
103	GNLY.3195.50.2	GNLY	Granulinsina	P22749	10578	Granulinsina	ascendente
104	LY9.3324.51.1	LY9	LY9	Q9HBG 7	4063	Antígeno de superficie de linfocitos T Ly-9/CD229	ascendente
105	NRP1.3214.3.2	NRP1	NRP1	014786	8829	Neuropilina-1	ascendente
106	C7.2888.49.2	C7	C7	P10643	730	Complemento C7	ascendente
107	KLKB1.4152.58.2	KLKB1	Precallicreina	P03952	3818	Precallicreina	descendente
108	CCL21.2516.57.3	CCL21	6Ckine	000585	6366	6Ckina/CCL21	ascendente
109	RGMB.3331.8.1	RGMB	RGMB	Q6NW4 0	285704	Miembro B de la familia de dominio RGM	ascendente
110	ESM1.3805.16.2	ESM1	Endocan	Q9NQ3 0	11082	Endocan	ascendente
111	DKK3.3607.71.1	DKK3	DKK3	Q9UBP4	27122	Proteína 3 relacionada con Dickkopf	ascendente
112	CDNF.4962.52.1	CDNF	ARMEL	Q49AH 0	441549	Factor neurotrófico de dopamina conservado	ascendente
113	SPON1.4297.62.3	SPON1	Espondina-1	Q9HCB 6	10418	Espondina-1	ascendente
114	LSAMP.2999.6.2	LSAMP	LSAMP	Q13449	4045	Proteína de membrana asociada al sistema límbico	ascendente

(continuación)

	Designación de SOMAmero	Nombre del gen	Diana	Id de Prot de Swiss	Id de gen de Entrez	NOMBRE_PÚBLICO	Ascendente o descendente
115	DCTPP1.4314.12.2	DCTPP1	XTP3A	Q9H773	79077	dCTP pirofosfatasa 1	ascendente
116	NID1.3213.65.2	NIDI	Nidogén	P14543	4811	Nidogén	ascendente
117	A2M.3708.62.1	A2M	Macroglobulina a2	P01023	2	Macroglobulina a2	ascendente
118	FCGR3B.3311.27.1	FCGR3B	FCG3B	075015	2215	Receptor III-B de la región Fc de inmunoglobulina G, baja afinidad	ascendente
119	PROC.3758.63.3	PROC	Proteína C activada	P04070	5624	Proteína C activada	descendente
120	ADAMTS13.3175.51.5	ADAMT S13	ATS 13	Q76LX 8	11093	Metalopectidasa ADAM con motivos de trombospondina 13	descendente
121	ANG.4874.3.1	ANG	Angiogenina	P03950	283	Angiogenina	ascendente
122	GRN.4992.49.1	GRN	GRN	P28799	2896	Progranulina	ascendente
123	CD48.3292.75.1	CD48	CD48	P09326	962	CD48	ascendente
124	FGA.FGB.FGG.4907.56.1	FGA FGB FGG	Dímero D	P02671 P02675 P02679	2243 2244 2266	Dímero D	ascendente
125	IGHG1.IGHG2.1 IGHG3.IGHG4.I GK..IGL. 3700.1 5.4	IGHG1 IGHG2 IGHG3 IGHG4 IGK@ IGL@	IgG	NA	NA	IgG	
126	NAGK.3894.15.2	NAGK	NAGK	Q9UJ70	55577	N-acetil-D-glucosamina cinasa	ascendente
127	TNC.4155.3.2	TNC	Tenascina	P24821	3371	Tenascina	ascendente
128	RET.3220.40.2	RET	RET	P07949	5979	Receptor Ret de tirosina-proteína cinasa de proto-oncogén	descendente
129	MDK.2911.27.2	MDK	Midquina	P21741	4192	Midquina	ascendente
130	TNFRSF10D.3129.73.2	TNFRSF 10D	TRAIL R4	Q9UBN 6	8793	Miembro 10D de la superfamilia de receptores del factor de necrosis tumoral	ascendente
131	CD84.3642.4.1	CD84	SLAF5	Q9UIB8	8832	Molécula 5 de activación linfocítica de señalización	ascendente

(continuación)

	Designación de SOMAmero	Nombre del gen	Diana	Id de Prot de Swiss	Id de gen de Entrez	NOMBRE_PÚBLICO	Ascendente o descendente
132	EGFR.2677.1.1	EGFR	ERBB1	P00533	1956	erbB1/HER1	descendente
133	SERPINA4.3449.58.2	SERPINA A4	Kalostatina	P29622	5267	Kalostatina	descendente
134	MRC2.3041.55.2	MRC2	MRC2	Q9UBG 0	9902	Receptor 2 de manosa de macrófagos	ascendente
135	GHR.2948.58.2	GHR	Receptor de hormona de crecimiento	P10912	2690	Receptor de hormona de crecimiento	descendente
136	CXCL12.2330.2.1	CXCL12	SDF-1a	P48061	6387	Factor derivado de células estromales 1α	ascendente
137	SERPINF2.3024.18.2	SERPINF 2	Antiplasmina α2	P08697	5345	Antiplasmina α2	descendente
138	RUNX2.3457.57.1	RUNX2	Factor de transcr 2 específico de osteoblastos	Q13950	860	Factor de transcripción 2 específico de osteoblastos	ascendente
139	CLEC11 A.2966.65.2	CLEC11 A	SCGF-beta	Q9Y240	6320	Factor β de crecimiento de células madre	ascendente
140	SIGLEC7.2742.68.2	SIGLEC7	Siglec-7	Q9Y286	27036	Siglec-7	ascendente
141	VEGFA.2597.8.3	VEGFA	VEGF	P15692	7422	Factor A de crecimiento endotelial vascular	ascendente
142	BMPER.3654.27.1	BMPER	BMPER	Q8N8U 9	168667	Proteína reguladora endotelial de unión a proteína morfogenética ósea	descendente
143	SOD2.5008.51.1	SOD2	Mn SOD	P04179	6648	Superóxido dismutasa [Mn]	descendente
144	OPCML.3634.5.4	OPCML	OBCAM	Q14982	4978	Molécula de adhesión celular de unión a opioides	descendente
145	CTSS.3181.50.2	CTSS	Catepsina S	P25774	1520	Catepsina S	ascendente
146	PLG.4151.6.2	PLG	Plasminógeno	P00747	5340	Plasminógeno	descendente
147	PAFAH1B2.2642.4.1	PAFAH1 B2	Subunidad beta de PFAH	P68402	5049	Subunidad β de IB de acetilhidrolasa de factor activador de plaquetas/subunidad de β de PFAH	ascendente
148	IGF1R.4232.19.2	IGF1R	R de IGF-I	P08069	3480	Receptor I de factor de crecimiento similar a la insulina	ascendente
149	AGT.3484.60.2	AGT	Angiotensinógeno	P01019	183	Angiotensinógeno	descendente
150	GDF11.2765.4.3	GDF11	GDF-11	O95390	10220	Factor 11 de diferenciación del crecimiento	descendente

(continuación)

	Designación de SOMAmero	Nombre del gen	Diana	Id de Prot de Swiss	Id de gen de Entrez	NOMBRE_PÚBLICO	Ascendente o descendente
151	MMP14.5002.76.1	MMP14	MMP-14	P50281	4323	Metaloproteinasa de matriz 14 /Metaloproteinasa de matriz 1 de tipo membrana	ascendente
152	SELP.4154.57.2	SELP	P-Selectina	P16109	6403	P-Selectina	ascendente
153	BGN.3284.75.1	BGN	BGN	P21810	633	Biglucano	ascendente
154	IL6ST.2620.4.2	IL6ST	gp130, soluble	P40189	3572	Subunidad β de receptor de interleucina 6/gp130	ascendente
155	PSMA2.4280.47.2	PSMA2	PSA2	P25787	5683	Subunidad de proteasoma $\alpha 2$	ascendente

Tabla 2: Lista de 46 biomarcadores

	Designación de SOMAmero	Nombre del gen	Diana	ID de Prot de Swiss	Id de gen de Entrez	NOMBRE_PÚBLICO	Ascendente o descendente
1	ANGPT2.2602.2.2	ANGPT2	Angiopoyetina-2	015123	285	Angiopoyetina-2	ascendente
2	CTSH.3737.6.3	CTSH	Catepsina H	P09668	1512	Catepsina H	ascendente
3	TIMP1.2211.9.6	TIMP1	TIMP-1	P01033	7076	Inhibidor tisular de metaloproteinasas 1	ascendente
4	B2M.3485.28.2	B2M	Microglobulina b2	P61769	567	Microglobulina β2	ascendente
5	TNFRSF1B.3152.57.1	TNFRSF1 B	R-II de TNF	P20333	7133	Miembro 1B de la superfamilia de receptores del factor de necrosis tumoral	ascendente
6	FSTL3.3438.10.2	FSTL3	FSTL3	095633	10272	Similar a follistatina 3	ascendente
7	MMP7.2789.26.2	MMP7	MM P-7	P09237	4316	Metaloproteinasa de matriz 7/Matrilisina	ascendente
8	IGFBP4.2950.57.2	IGFBP4	IGFBP-4	P22692	3487	Proteína de unión 4 al factor de crecimiento similar a insulina	ascendente
9	SLPI.4413.3.2	SLPI	SLPI	P03973	6590	Inhibidor de la proteasa leucocitaria secretora	ascendente
10	THBS2.3339.33.1	THBS2	TSP2	P35442	7058	Trombospondina-2	ascendente
11	CHRDL1.3362.61.2	CHRDL1	CRDL1	Q9BU40	91851	Similar a cordina 1	ascendente
12	CFD.2946.52.2	CFD	Factor D	P00746	1675	Factor del complemento D	ascendente
13	SPON1.4297.62.3	SPON1	Espondina-1	Q9HCB6	10418	Espondina-1	ascendente
14	COL18A1.2201.17.6	COL18A1	Endostatina	P39060	80781	Endostatina	ascendente
15	IGFBP2.2570.72.5	IGFBP2	IGFBP-2	P18065	3485	Proteína de unión 2 al factor de crecimiento similar a insulina	ascendente
16	MMP12.4496.60.2	MMP12	MMP-12	P39900	4321	Metaloproteinasa de matriz 12/Metalolasiada de macrófagos	ascendente
17	CCL14.2900.53.3	CCL14	HCC-1	Q16627	6358	Quimiocina 1 CC de hemofiltrado /CCL14	ascendente
18	EFNA5.2615.60.2	EFNA5	Efrina-A5	P52803	1946	Efrina-A5	ascendente
19	PLAUR.2652.15.1	PLAUR	suPAR	Q03405	5329	Receptor de superficie de activador de plasminógeno de urocinasa	ascendente
20	C7.2888.49.2	C7	C7	P10643	730	Complemento C7	ascendente
21	STC1.4930.21.1	STC1	Estañocalcina -1	P52823	6781	Estañocalcina-1	ascendente
22	KLK3.SERPINA 3.4153.11.2	KLK3 SERPINA 3	PSA-ACT	P07288, P01011	354 12	Complejo PSA:α-1-antiquimotripsina	ascendente
23	CST3.2609.59.2	CST3	Cistatina C	P01034	1471	Cistatina C	ascendente
24	CTSZ.4971.1.1	CTSZ	CATZ	Q9UBR2	1522	Catepsina Z	ascendente
25	SERPINF2.3024. 18.2	SERPINF 2	Antiplasmina a2	P08697	5345	Antiplasmina α2	descendente

(continuación)

	Designación de SOMAmero	Nombre del gen	Diana	ID de Prot de Swiss	Id de gen de Entrez	NOMBRE_PÚBLICO	Ascendente o descendente
26	PROC.2961.1.2	PROC	Proteína C	P04070	5624	Proteína C	descendente
27	CCL15.3509.1.1	CCL15	MI P-5	Q16663	6359	Proteína inflamatoria de macrófagos 5/CCL15	ascendente
28	CD84.3642.4.1	CD84	SLAF5	Q9UJB8	8832	Molécula 5 de activación linfocítica de señalización	ascendente
29	VCAM1.2967.8.1	VCAM1	VCAM-1	P19320	7412	Proteína 1 de adhesión celular Vasculár/VCAM 1	ascendente
30	TNFRSF1A.2654.19.1	TNFRSF1A	R-I deTNF	P19438	7132	Miembro 1A de la superfamilia del receptor del factor de necrosis tumoral	ascendente
31	CCDC80.3234.23.2	CCDC80	URB	Q76M96	151887	Proteína 80 que contiene dominio de enrollamiento enrollado	ascendente
32	IL18BP.3073.51.2	IL18BP	IL-18 BPa	095998	10068	Proteína de unión de interleucina 18	ascendente
33	IL1R1.2991.9.2	IL1R1	RI de IL-1	P14778	3554	Receptor 1 de Interleucina-1	ascendente
34	CXCL12.2330.2.1	CXCL12	SDF-1a	P48061	6387	Factor derivado de células estromales 1a	ascendente
35	PLA2G2A.2692.74.2	PLA2G2A	NPS-PLA2	P14555	5320	Fosfolipasa A2, Grupo IIA	ascendente
36	CAPG.4968.50.1	CAPG	CAPG	P40121	822	Proteína de protección de macrófagos	ascendente
37	CNDP1.3604.6.4	CNDP1	CNDP1	Q96KN2	84735	Carnosina dipeptidasa 1	descendente
38	ESAM.2981.9.3	ESAM	ESAM	Q96AP7	90952	Molécula de adhesión selectiva de células endoteliales	ascendente
39	MATN2.3325.2.2	MATN2	MATN2	000339	4147	Matrilina-2	ascendente
40	LYZ.4920.10.1	LYZ	Lisozima	P61626	4069	Lisozima	ascendente
41	PLG.4151.6.2	PLG	Plasminógeno	P00747	5340	Plasminógeno	descendente
42	C9.3060.43.2	C9	C9	P02748	735	Complemento C9	ascendente
43	LCN2.2836.68.2	LCN2	Lipocalina 2	P80188	3934	Lipocalina 2	ascendente
44	NID1.3213.65.2	NID1	Nidogén	P14543	4811	Nidogén	ascendente
45	CHST15.4469.78.2	CHST15	ST4S6	Q7LFX5	51363	Sulfotransferasa de hidrato de carbono 15	ascendente
46	ROR1.2590.69.4	ROR1	ROR1	Q01973	4919	Receptor transmembrana de tirosina-proteína cinasa ROR1	ascendente

Tabla 3: Lista de 10 biomarcadores

	Designación de SOMAmero	Nombre del gen	Diana	ID de Prot Swiss	Id de gen de Entrez	NOMBRE_PÚBLICO	ascendente o descendente en alto riesgo
1	ANGPT2.2602.2.2	ANGPT2	Angiopoyetina-2	015123	285	Angiopoyetina-2	ascendente
2	MMP7.2789.26.2	MMP7	MMP-7	P09237	4316	Metaloproteinasa de matriz 7/Matrilisina	ascendente
3	THBS2.3339.33.1	THBS2	TSP2	P35442	7058	Trombospondina-2	ascendente
4	CHRDL1.3362.61.2	CHRDL1	CRDL1	Q9BU40	91851	Similar a cordina 1	ascendente
5	C7.2888.49.2	C7	C7	P10643	730	Complemento C7	ascendente
6	KLK3.SERPINA3.4153.11.2	KLK3 SERPINA3	PSA-ACT	P07288, P01011	354 12	Complejo PSA:α-1-antiquimotripsina	ascendente
7	ESAM.2981.9.3	ESAM	ESAM	Q96AP7	90952	Molécula de adhesión selectiva de células endoteliales	ascendente
8	MATN2.3325.2.2	MATN2	MATN2	000339	4147	Matrilina-2	ascendente
9	PLG.4151.6.2	PLG	Plasminógeno	P00747	5340	Plasminógeno	descendente
10	GP6.3194.36.2	GP6	GPVI	Q9HCN6	51206	GPVI/Gluco proteina plaquetaria VI	ascendente

Tabla 4: Población, Evento, Tiempo y Tipo de estudio *Heart & Soul*

TIEMPO	ICC	ICC crónica ICC	MUERTE	IM	N.º de Evento	Crónico. ICC Sin evento	Ictus
0-6 meses	8	5	8	13	0	0	5
6-12 meses	3	6	8	11	0	0	2
12-24 meses	16	8	15	32	0	0	8
2-3 años	8	6	22	16	0	0	9
3-4 años	12	6	17	9	0	0	1
4-5 años	13	5	39	14	0	0	5
5 años más	4	3	48	5	0	0	4
Sin eventos durante el estudio	0	0	0	0	515	77	0
Total	64	39	157	100	515	77	34
986							

REIVINDICACIONES

1. Un método para seleccionar un individuo para la evaluación del riesgo de un evento CV, comprendiendo el método: proporcionar un panel de biomarcadores que comprende N de las proteínas biomarcadoras enumeradas en la Tabla 1; detectar proteínas biomarcadoras, en una muestra biológica de un individuo, para proporcionar valores de biomarcadores que correspondan a los biomarcadores en el panel de biomarcadores con el fin de predecir el riesgo de un evento CV, basándose en dichos valores de biomarcadores, y en donde N = 3-155 y en donde el panel de biomarcadores incluye PSA-ACT.
2. El método de la reivindicación 1, en donde los biomarcadores se seleccionan entre la Tabla 2 y N = 3-46 o los biomarcadores se seleccionan entre la Tabla 3 y N = 3-10.
3. El método de la reivindicación 1 o de la reivindicación 2, en donde el panel de biomarcadores incluye angiopoyetina 2, MMP7, CHRDL1 o MATN2.
4. El método de la reivindicación 1 en donde el panel de biomarcadores comprende los diez biomarcadores de la Tabla 3.
5. El método de una cualquiera de las reivindicaciones anteriores, en donde la detección de los valores de biomarcadores comprende realizar un ensayo *in vitro* que comprende al menos un reactivo de captura correspondiente a cada uno de dichos biomarcadores, y que comprende adicionalmente seleccionar dicho al menos un reactivo de captura entre el grupo que consiste en SOMAmeros y anticuerpos.
6. El método de una cualquiera de las reivindicaciones anteriores, en donde la muestra biológica se selecciona entre el grupo que consiste en sangre completa, manchas de sangre seca, plasma, suero y orina.
7. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en donde dicho individuo se evalúa para determinar el riesgo de un evento CV, basándose en dichos valores de biomarcadores y al menos un elemento de información biomédica adicional correspondiente a dicho individuo, en donde dicho al menos un elemento de información biomédica adicional se selecciona independientemente entre el grupo que consiste en:
- (a) información correspondiente a la presencia de factores de riesgo cardiovasculares seleccionados entre el grupo que consiste en infarto de miocardio previo, evidencia angiográfica de una estenosis superior al 50 % en uno o más vasos coronarios, isquemia inducida por ejercicio mediante cinta de correr o ensayos nucleares o revascularización coronaria previa,
 - (b) información correspondiente a descriptores físicos de dicho individuo,
 - (c) información correspondiente a un cambio en el peso de dicho individuo,
 - (d) información correspondiente a la etnia de dicho individuo,
 - (e) información correspondiente al género de dicho individuo,
 - (f) información correspondiente al historial de tabaquismo de dicho individuo,
 - (g) información correspondiente al historial de consumo de alcohol de dicho individuo,
 - (h) información correspondiente al historial ocupacional de dicho individuo,
 - (i) información correspondiente al historial familiar de enfermedades cardiovasculares u otras afecciones del sistema circulatorio de dicho individuo,
 - (j) información correspondiente a la presencia o la ausencia en dicho individuo de al menos un marcador genético que se correlacione con un mayor riesgo de enfermedades cardiovasculares en dicho individuo o en un miembro de la familia de dicho individuo,
 - (k) información correspondiente a síntomas clínicos de dicho individuo,
 - (l) información correspondiente a otros ensayos de laboratorio,
 - (m) información correspondiente a valores de expresión génica de dicho individuo, y
 - (n) información correspondiente al consumo de factores de riesgo cardiovasculares conocidos de dicho individuo, tales como una dieta rica en grasas saturadas, rica en sal, rica en colesterol,
 - (o) información correspondiente a los resultados de formación de imágenes del individuo obtenidos mediante técnicas seleccionadas entre el grupo que consiste en electrocardiograma, ecocardiografía, ultrasonido carotídeo para determinar el espesor íntima-media, dilatación mediada por flujo, velocidad de la onda del pulso, índice tobillo-brazo, ecocardiografía con estrés, imágenes de perfusión miocárdica, calcio coronario por TC, angiografía por TC de alta resolución, RM y otras modalidades de formación de imágenes, y
 - (p) información sobre los medicamentos del individuo.
8. El método de acuerdo con una cualquiera de las reivindicaciones anteriores en donde se evalúa el riesgo de un evento cardiovascular (CV) futuro en un período de 5 años.
9. El método de la reivindicación 8, que comprende adicionalmente: seleccionar la población sobre la base de una característica seleccionada entre el grupo que consiste en:
- a) una población que tiene o no antecedentes de enfermedades cardiovasculares; y

b) una población que tiene o no tiene factores de riesgo genéticos que comprenden mutaciones, polimorfismos de nucleótido único y/o inserción/supresiones.

5 10. El método de la reivindicación 8, en donde el riesgo evaluado de evento CV para individuos permite su estratificación en categorías seleccionadas entre el grupo que consiste en:

- 10 a) diferentes categorías para programas de abordaje de enfermedades aumentados o disminuidos;
b) diferentes categorías de riesgo relacionadas con la cobertura del seguro de vida;
c) diferentes categorías de riesgo relacionadas con la cobertura del seguro de salud;
d) diferentes categorías de candidatura para la asociación;
e) diferentes categorías de elegibilidad para ingresar en ensayos clínicos de terapéutica CV;
f) diferentes categorías de riesgo relacionadas con la seguridad cardiovascular para un producto terapéutico CV o cualquier producto terapéutico.

15 11. Un método implementado por ordenador para evaluar el riesgo de un evento cardiovascular (CV), comprendiendo el método:

- 20 proporcionar un panel de biomarcadores que comprende N de los biomarcadores de proteína de la Tabla 1;
recuperar en un ordenador información de biomarcadores para un individuo, en donde la información de biomarcadores comprende valores de biomarcadores que corresponden a uno de al menos N biomarcadores seleccionados entre la Tabla 1;
realizar con el ordenador una clasificación de cada uno de dichos valores de biomarcadores;
indicar un resultado de la evaluación del riesgo para un evento CV para dicho individuo basándose en una pluralidad de clasificaciones, en donde indicar el resultado de la evaluación del riesgo de un evento CV para el
25 individuo opcionalmente comprende representar el resultado en una pantalla de ordenador; y
en donde N = 3-155, y en donde el panel de biomarcadores incluye PSA-ACT.

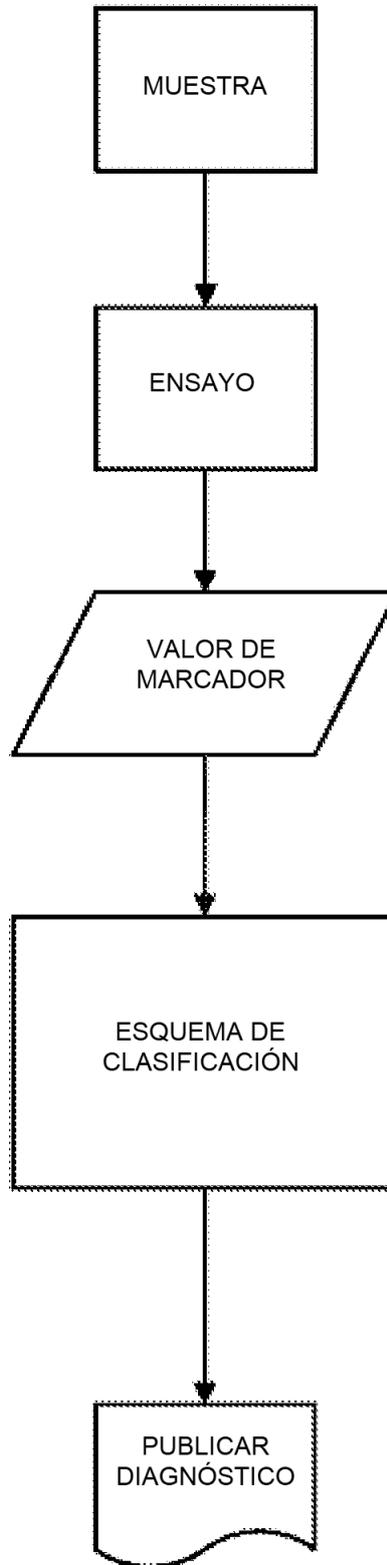


FIG. 1A

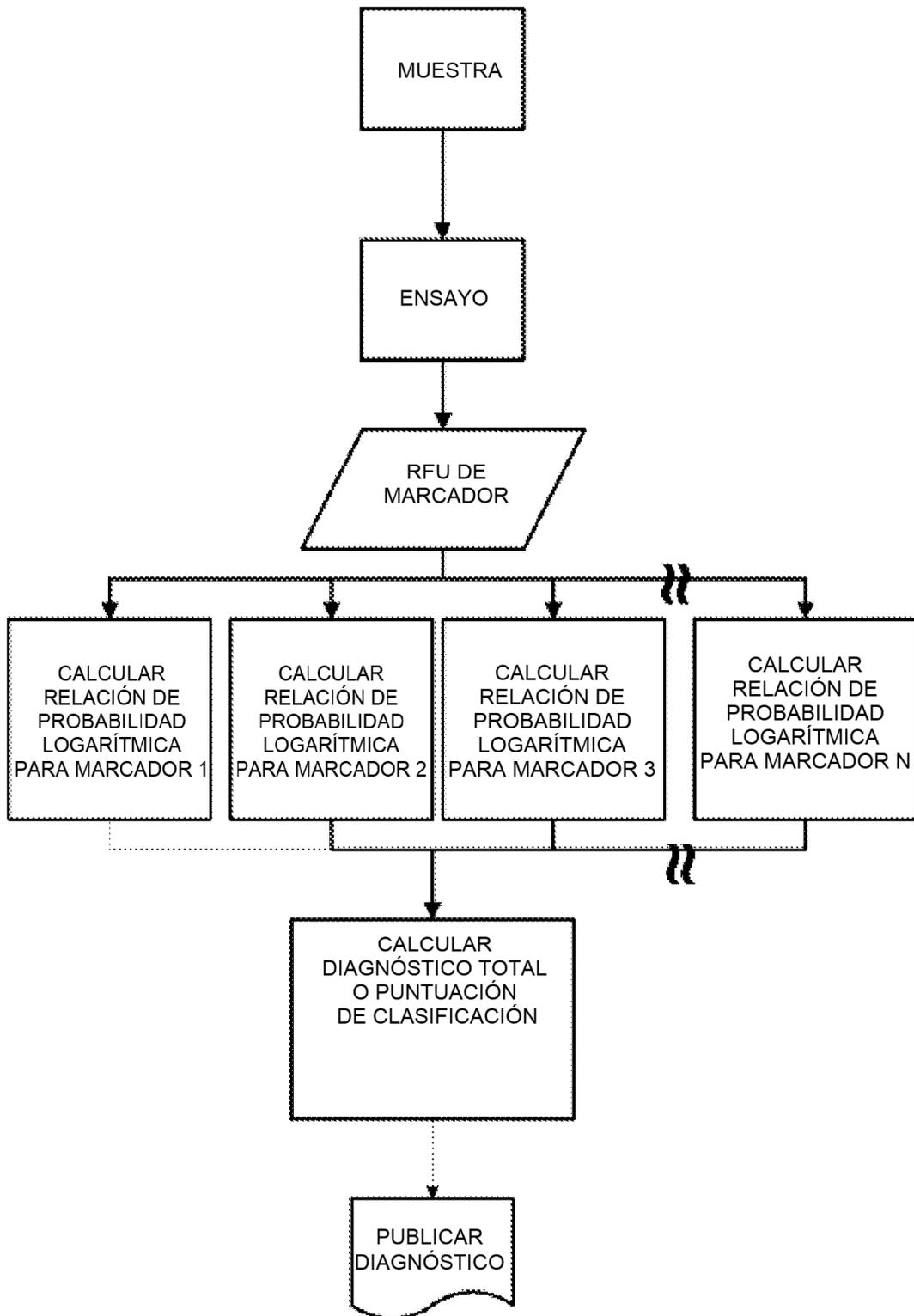


FIG. 1B

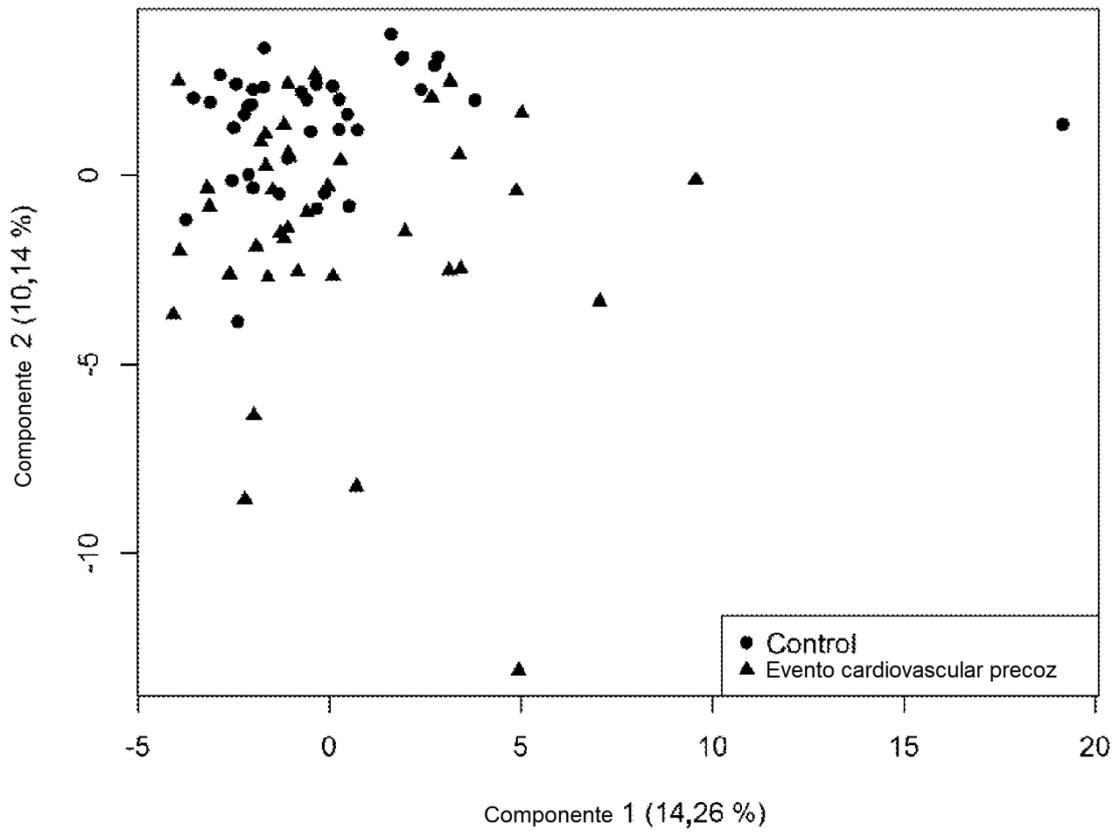


FIG. 2A

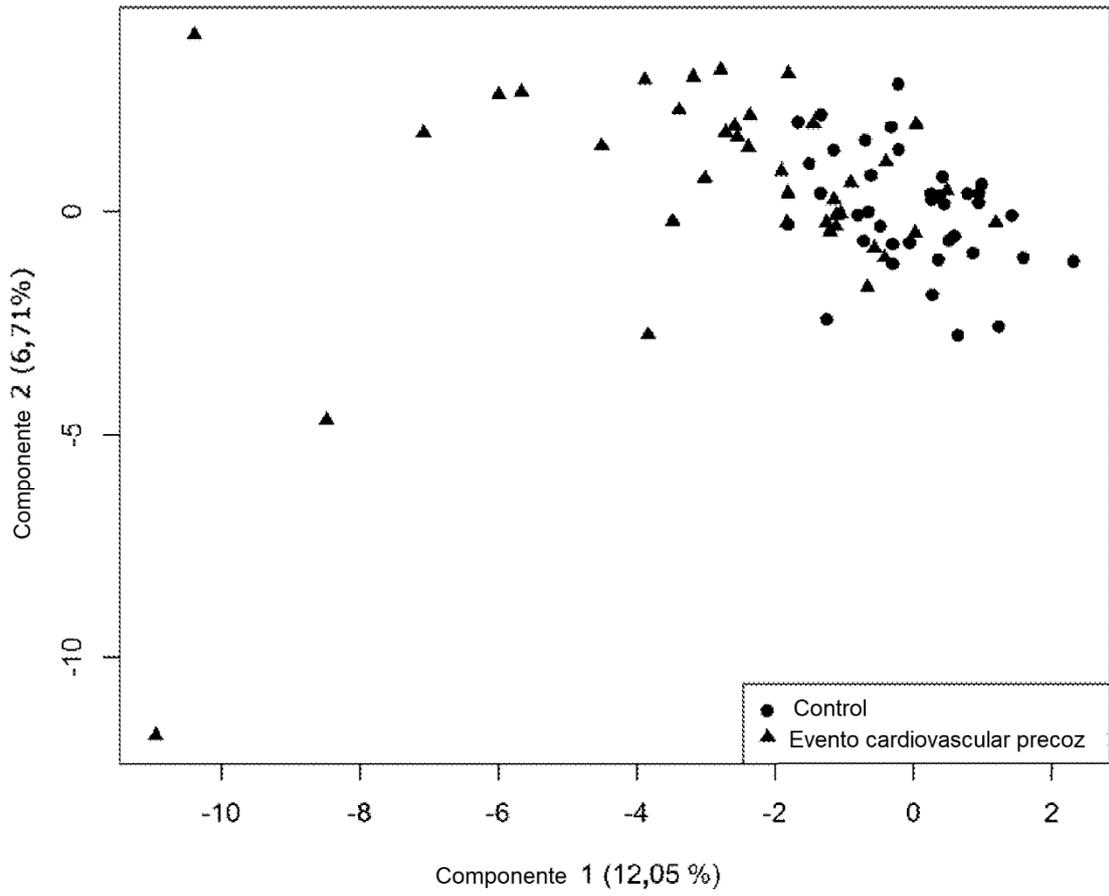


FIG. 2B

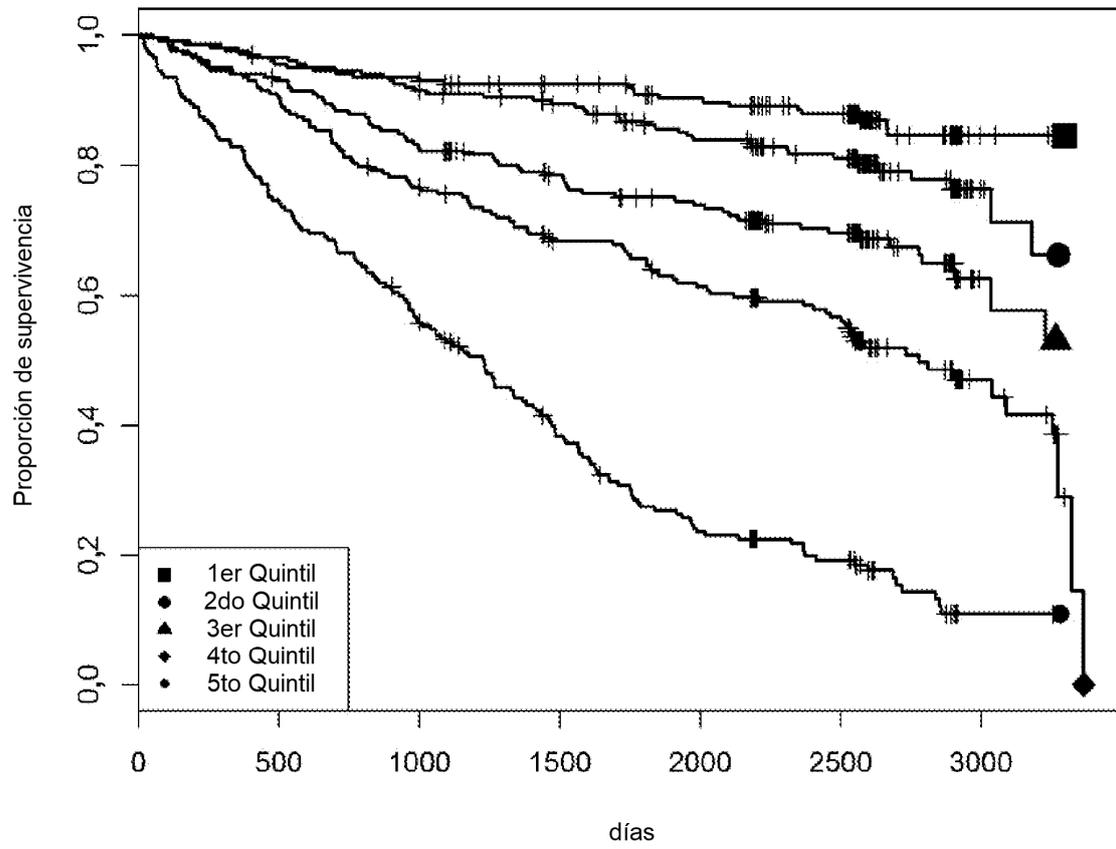


FIG. 3A

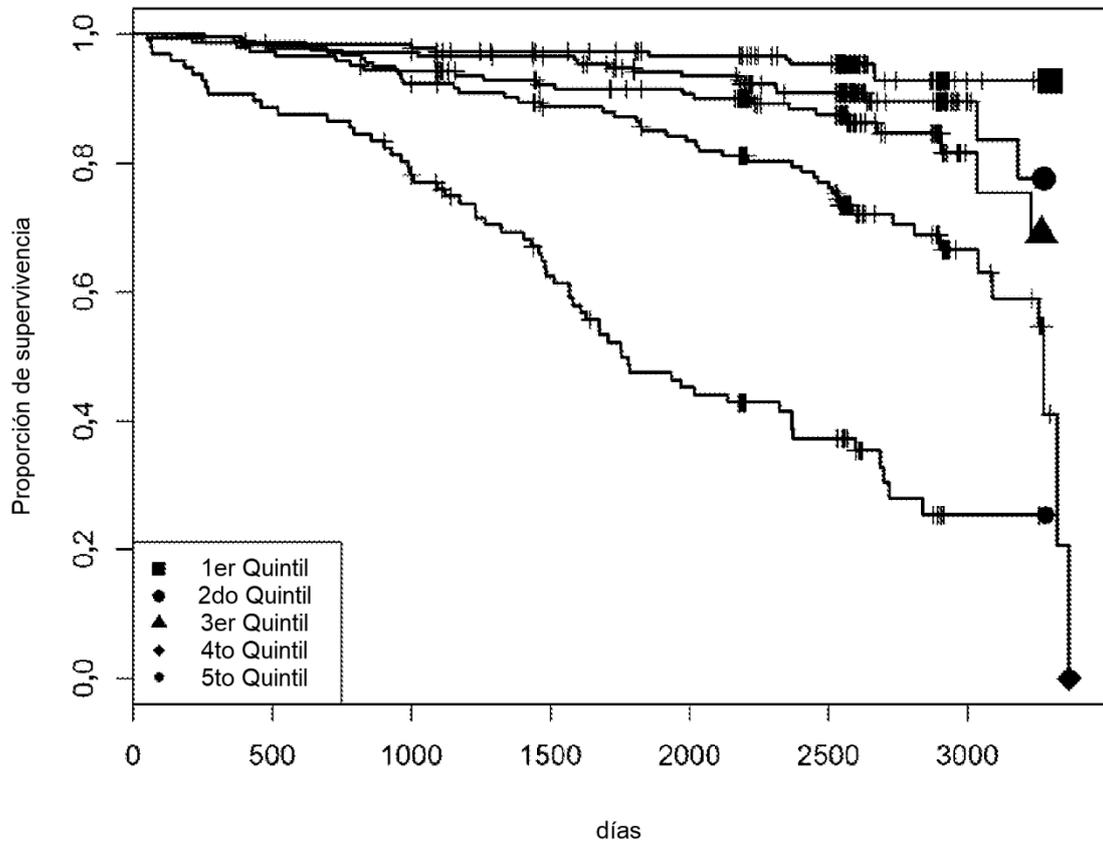


FIG. 3B

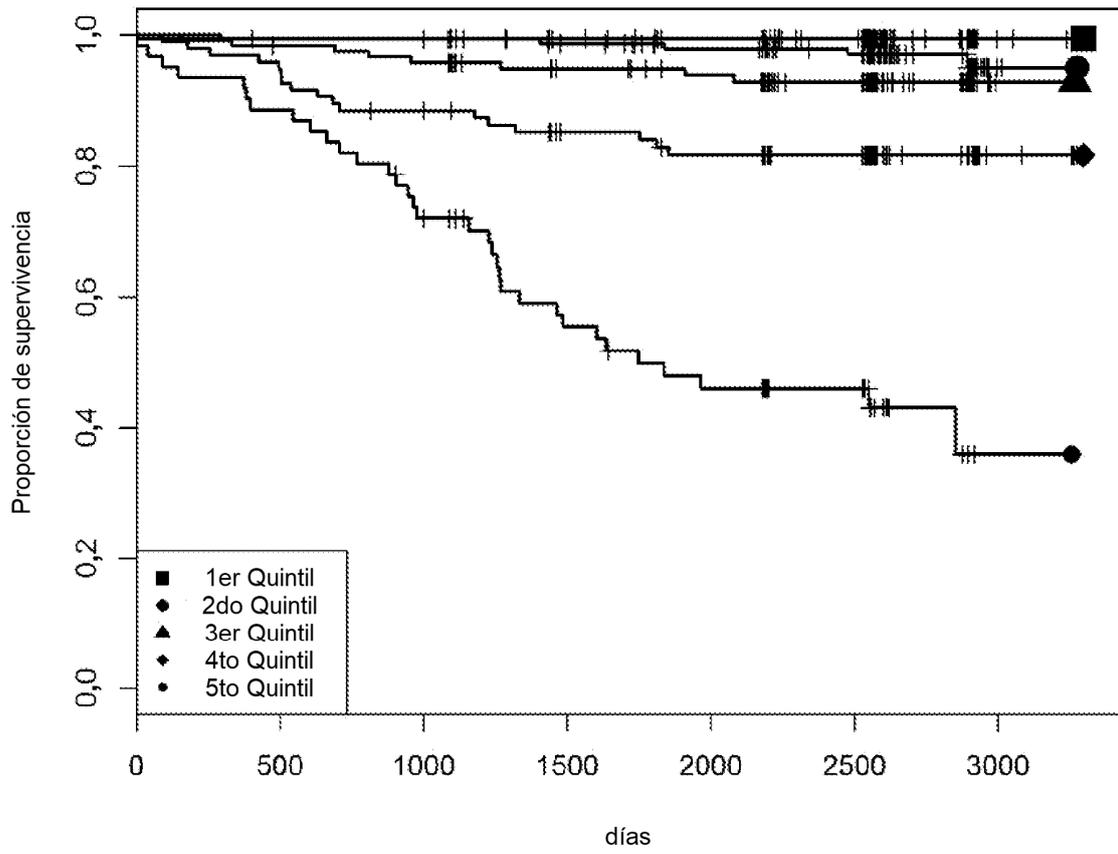


FIG. 3C

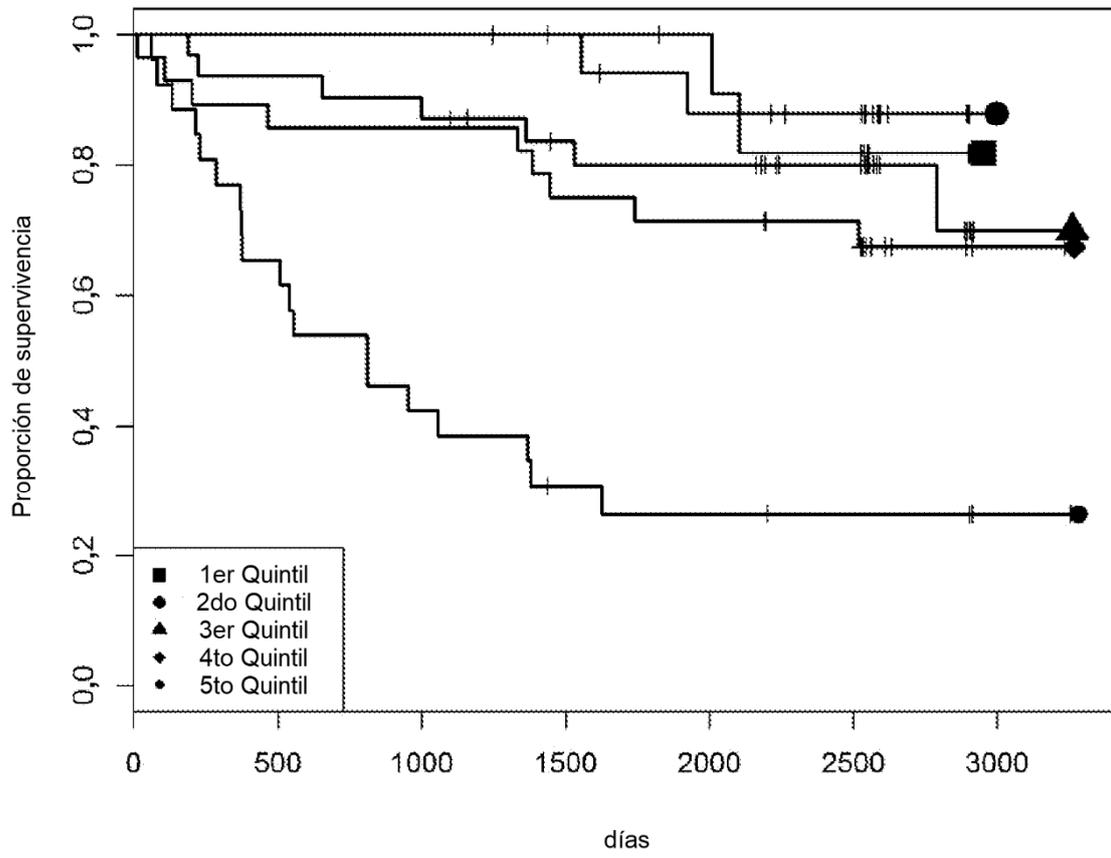


FIG. 3D

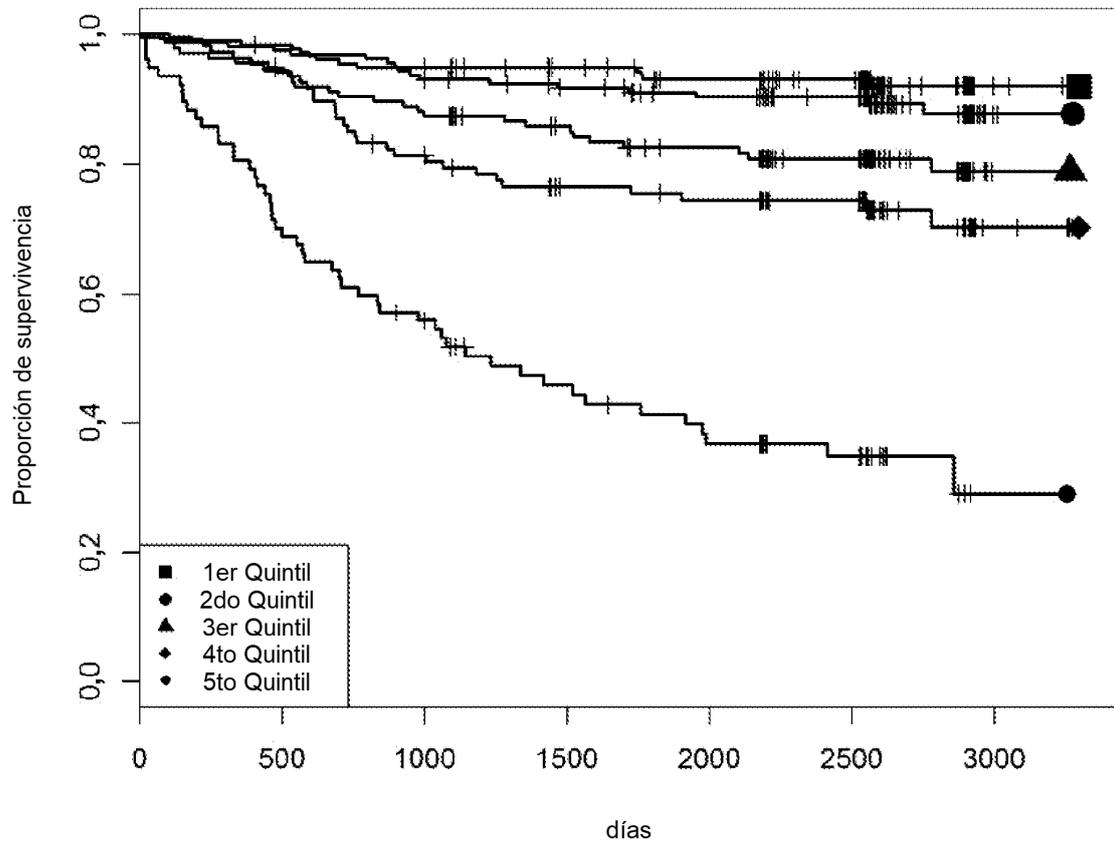


FIG. 3E

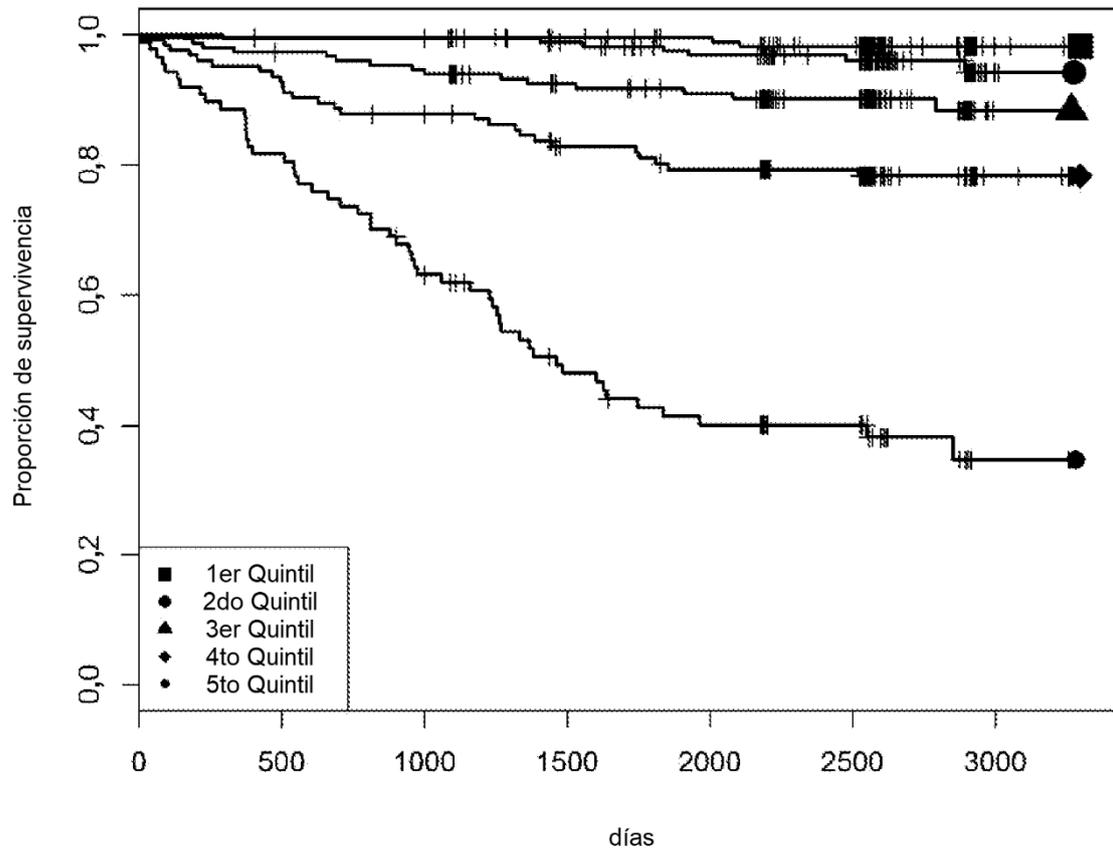


FIG. 3F

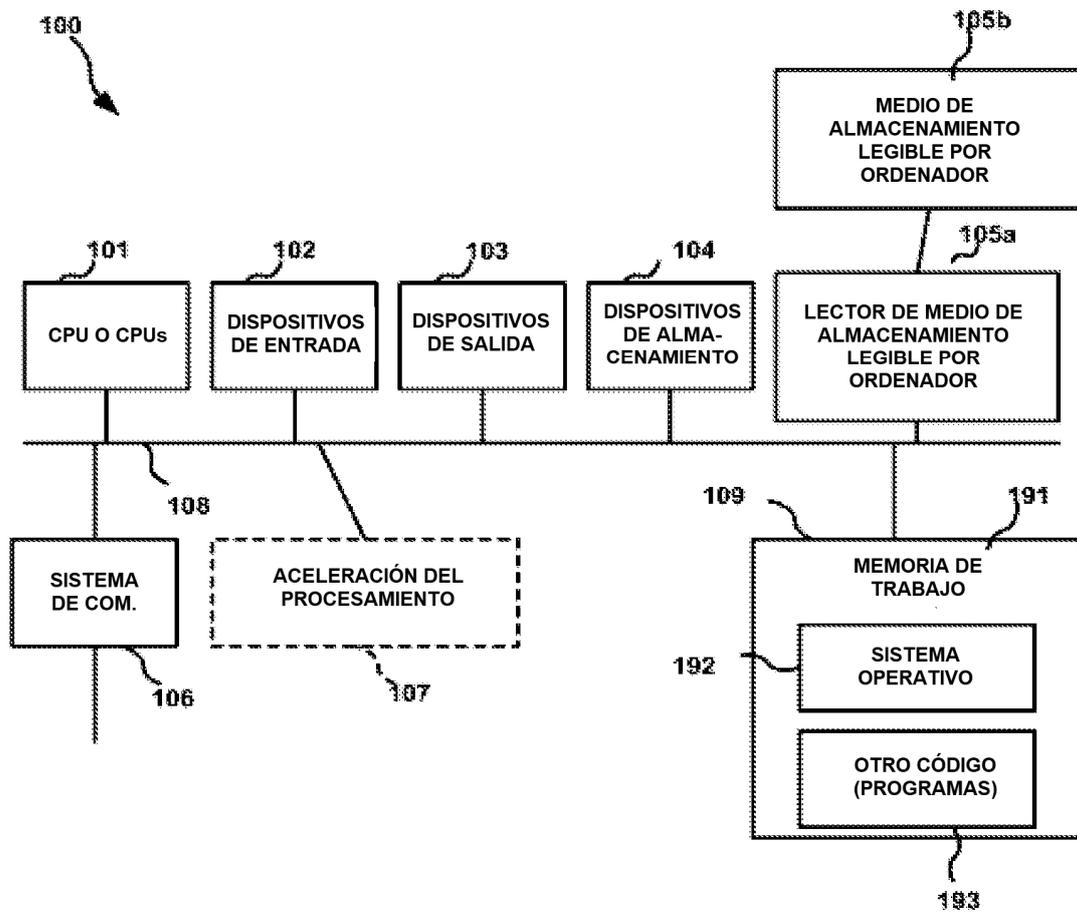


FIG. 4

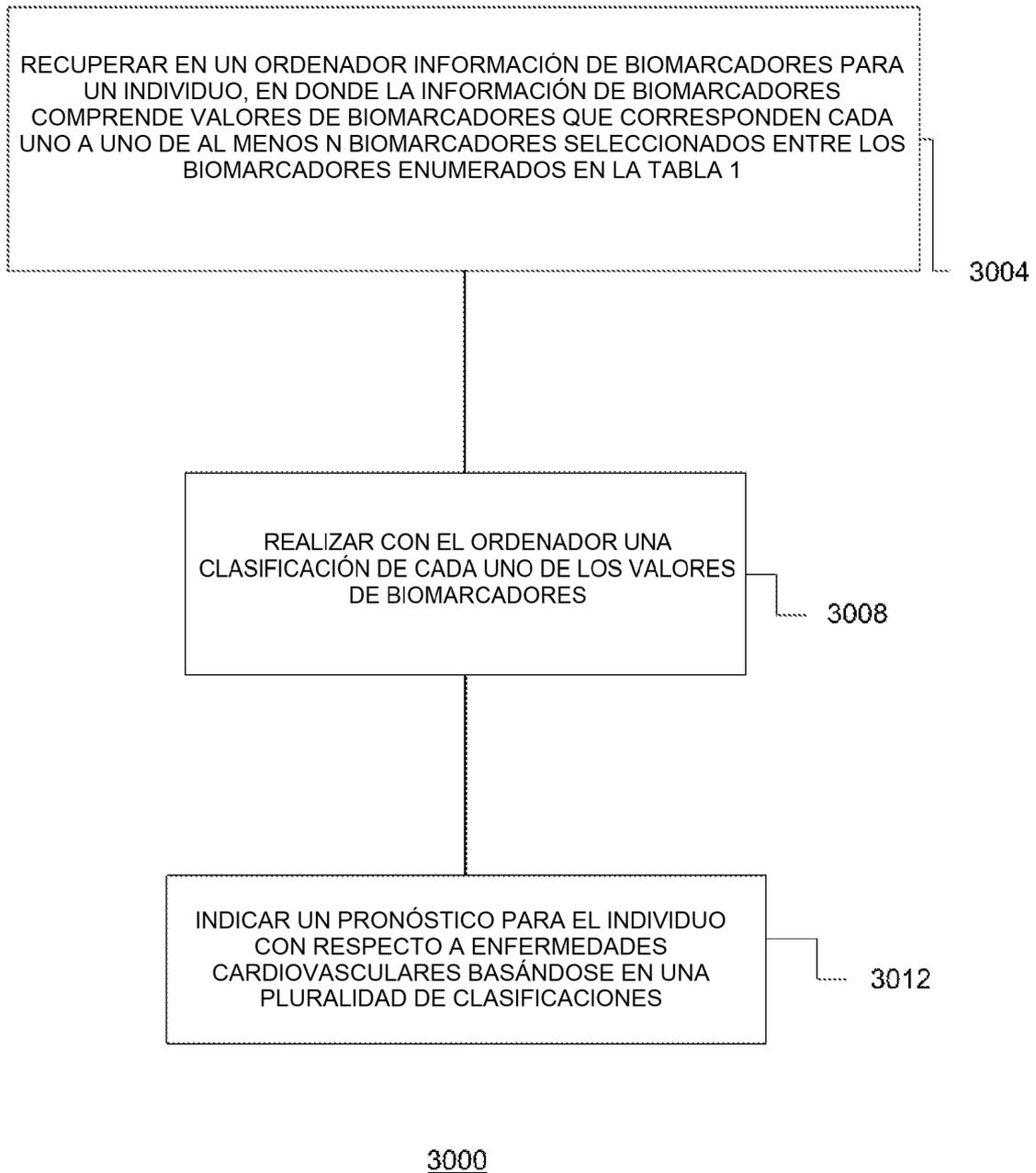


FIG. 5

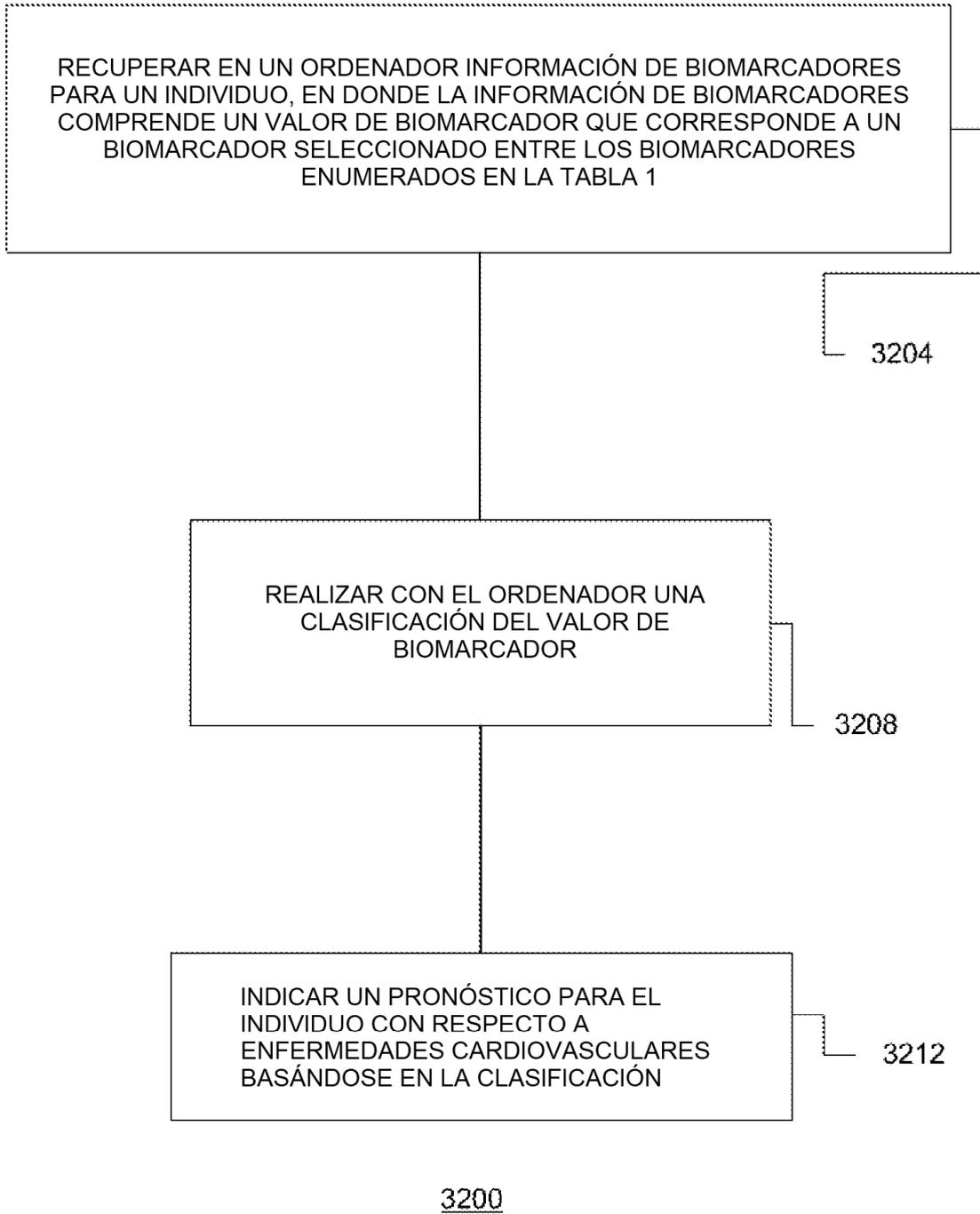


FIG. 6

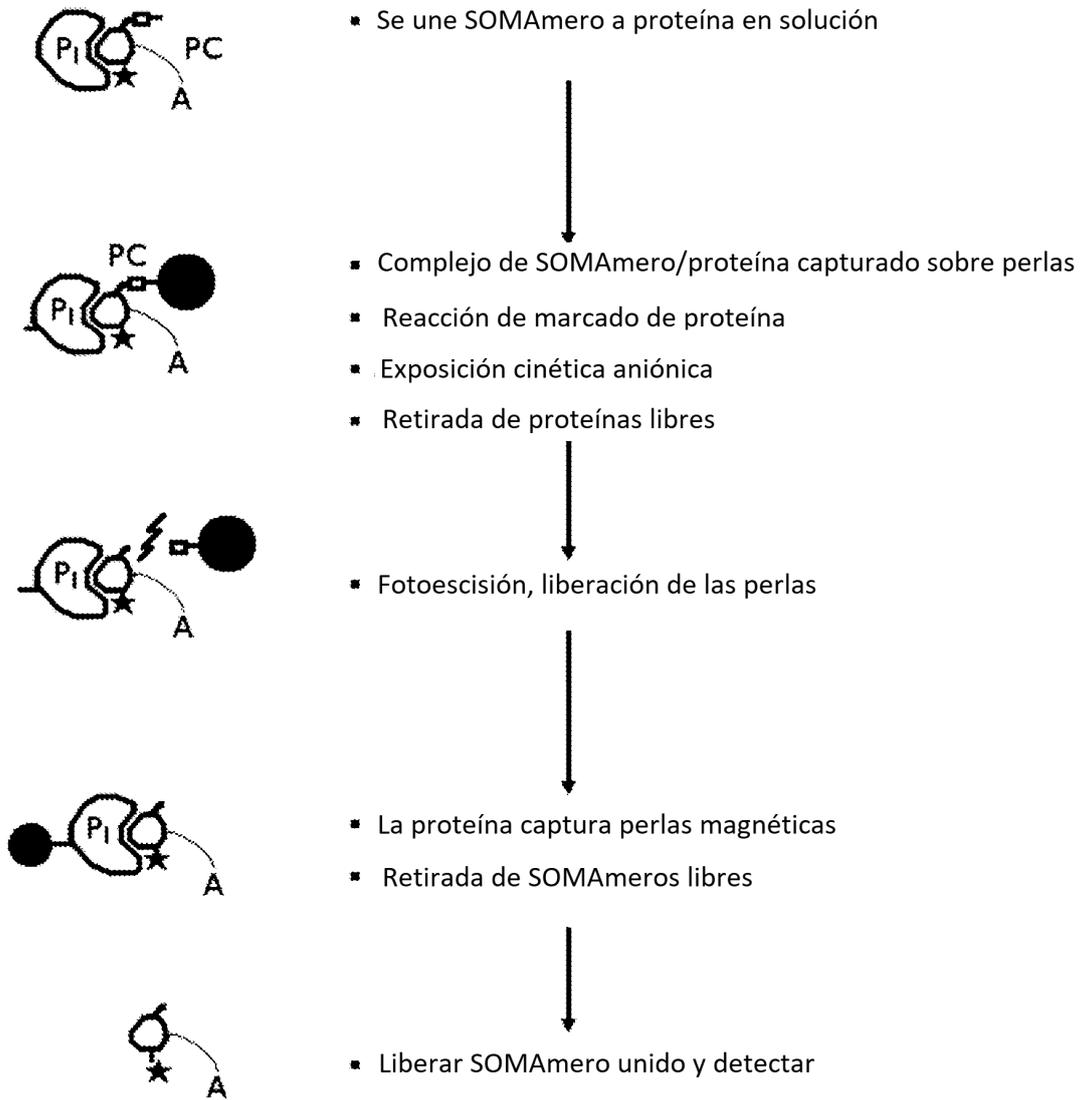


FIG. 7

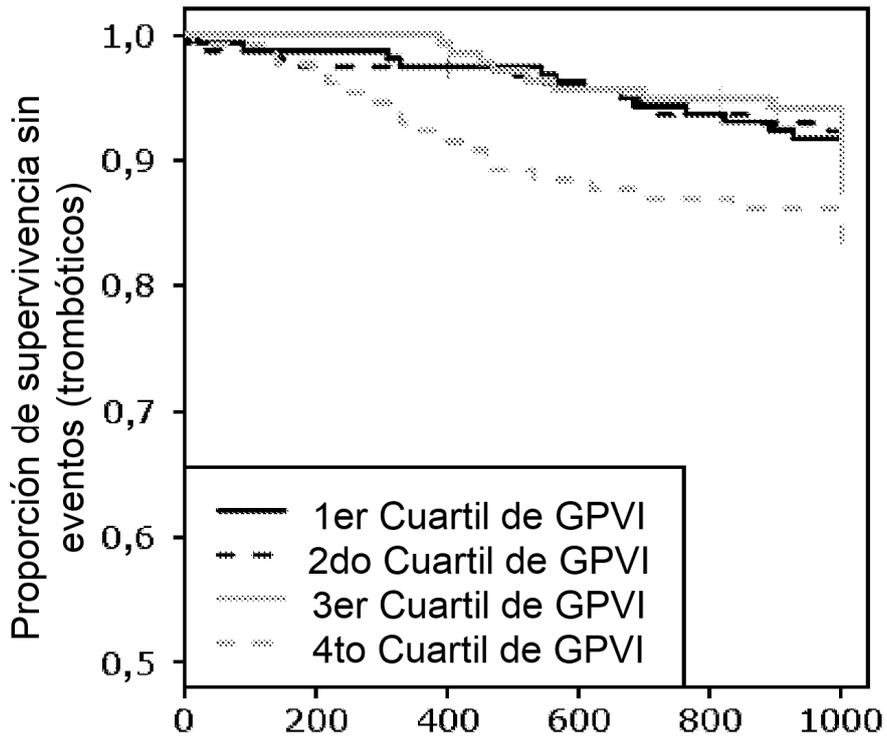


FIG. 8A

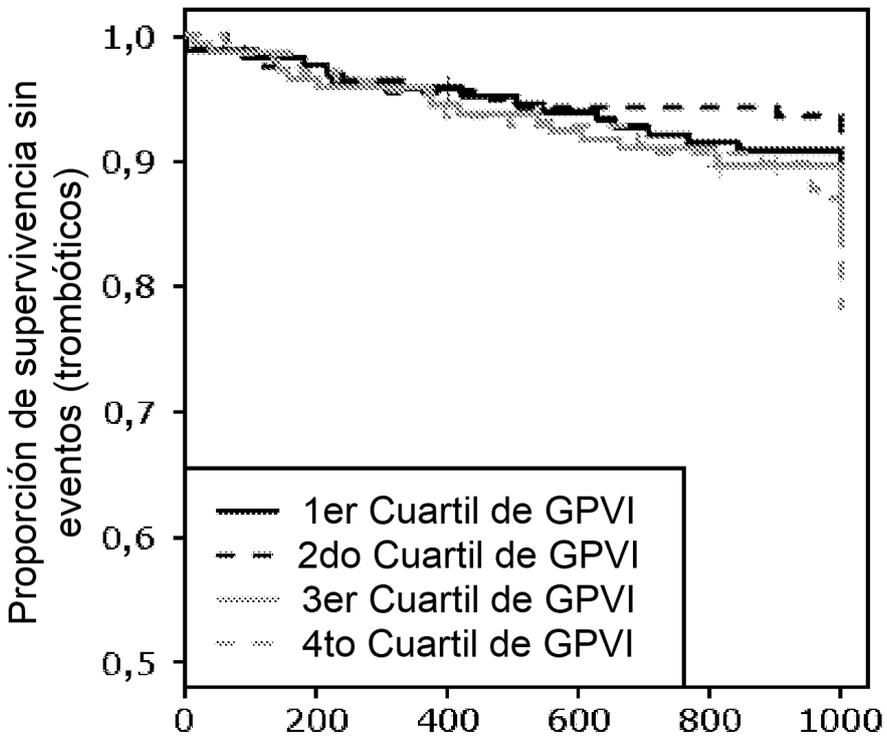


FIG. 8B

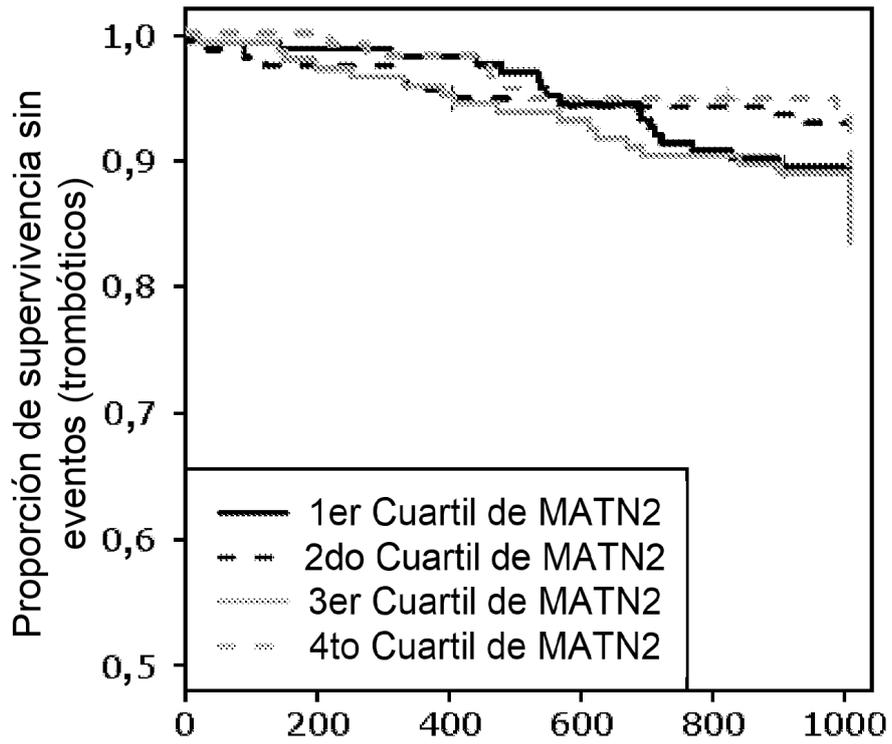


FIG. 9A

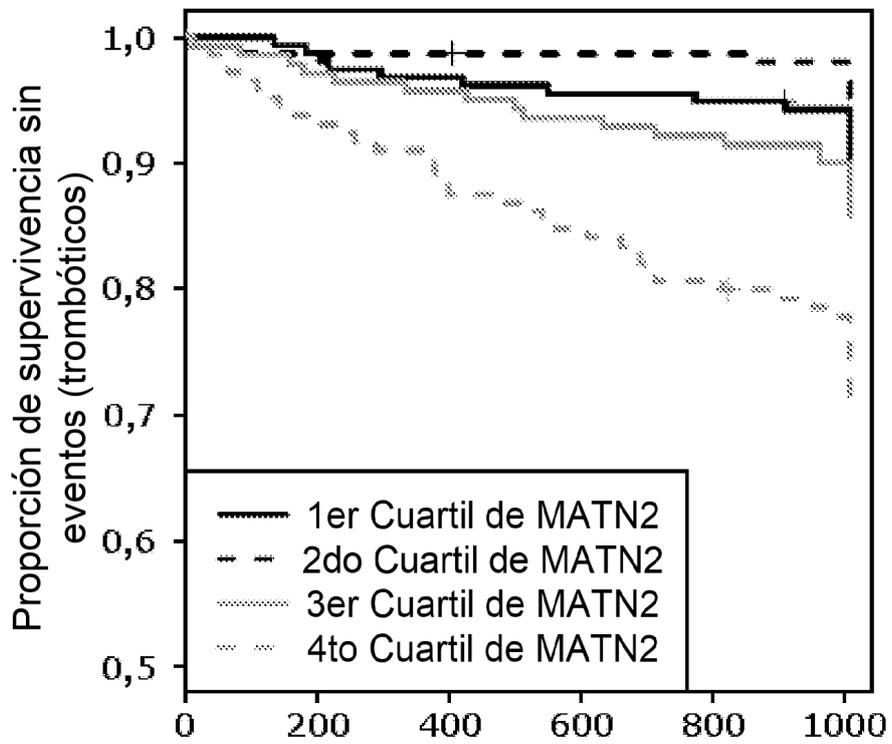


FIG. 9B

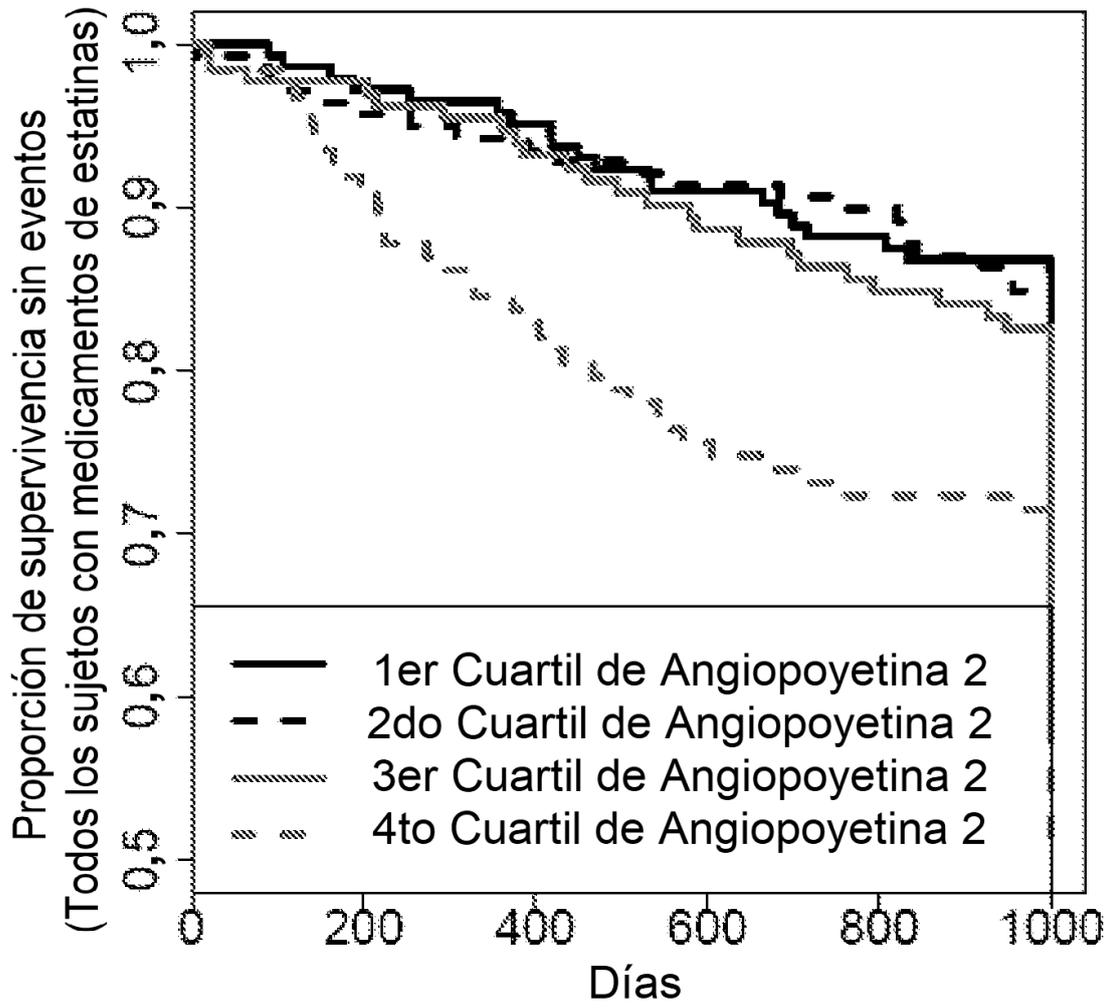


FIG. 10

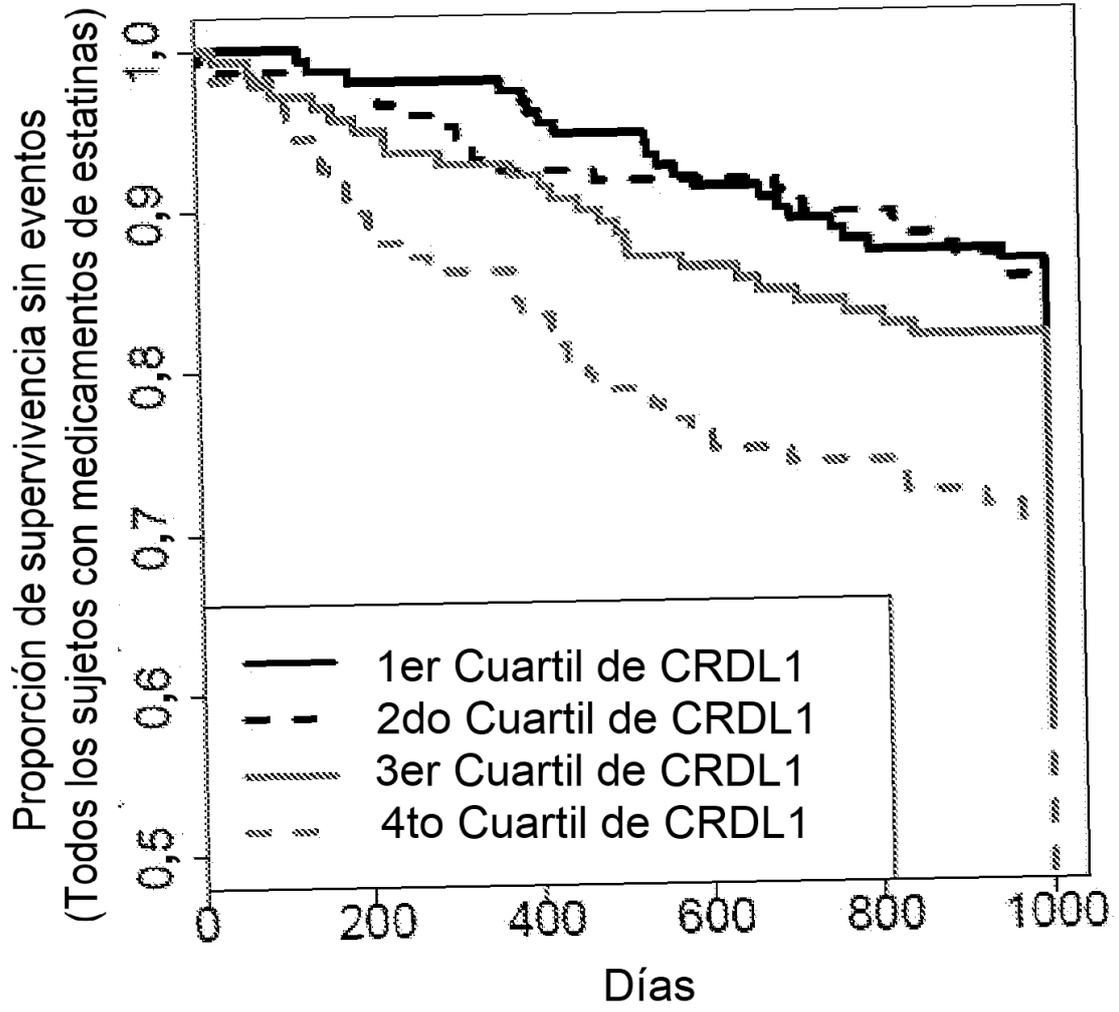


FIG. 11