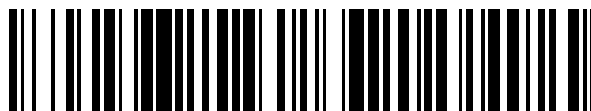


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 374 762**

51 Int. Cl.:
G01N 33/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **08837800 .5**
96 Fecha de presentación: **10.10.2008**
97 Número de publicación de la solicitud: **2210108**
97 Fecha de publicación de la solicitud: **28.07.2010**

54 Título: **PROCEDIMIENTOS PARA DETECTAR COMPLICACIONES CARDIOVASCULARES Y CEREBROVASCULARES ADVERSAS GRAVES.**

30 Prioridad:
10.10.2007 US 998563 P
11.10.2007 US 998756 P

45 Fecha de publicación de la mención BOPI:
21.02.2012

45 Fecha de la publicación del folleto de la patente:
21.02.2012

73 Titular/es:
BG MEDICINE, INC.
601 N. LINCOLN STREET
WALTHAM, MA 02451, US

72 Inventor/es:
MUNTENDAM, Pieter;
BALASUBRAMANIAN, Rajalakshmi y
MYERS, Rene

74 Agente: **de Elzaburu Márquez, Alberto**

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

ES 2 374 762 T3

DESCRIPCIÓN

Procedimientos para detectar complicaciones cardiovasculares y cerebrovasculares adversas graves

Campo de la invención

- 5 Las presentes enseñanzas se refieren a procedimientos para predecir si un ser humano padecerá una complicación cardiovascular o cerebrovascular adversa grave, o MACCE. Más específicamente, las presentes enseñanzas se refieren a procedimientos para cribar un individuo para si está en riesgo de tener o desarrollar una complicación cardiovascular o cerebrovascular adversa grave usando uno o más analitos.

Antecedentes

- 10 El infarto de miocardio es la principal causa de muerte (véase www.americanheart.org). Una de cada 5 muertes en los Estados Unidos se produce de un infarto de miocardio. En 2004 hubo 452.327 muertes en los Estados Unidos debidas a infarto de miocardio, produciendo aproximadamente 1.200.000 infartos cardiovasculares nuevos y recurrentes.

- 15 El accidente cerebrovascular es la tercera causa principal de muerte en los Estados Unidos (véase www.americanheart.org). El accidente cerebrovascular mató a 150.147 personas en 2004, produciendo aproximadamente 700.000 complicaciones cerebrovasculares nuevas y recurrentes. El accidente cerebrovascular es la causa principal de discapacidad grave a largo plazo en los Estados Unidos. Aproximadamente 5.700.000 de supervivientes de accidentes cerebrovasculares viven hoy en día en los Estados Unidos. 2.400.000 son hombres y 3.300.000 son mujeres.

- 20 Tanto el infarto de miocardio como ciertos tipos de accidente cerebrovascular pueden producir la rotura de la placa aterosclerótica vulnerable (Naghavi y col., *Circulation* 108:1664-72 y 108:1772-8, 2003). Actualmente, el riesgo de tener un infarto de miocardio o accidente cerebrovascular se evalúa en la población general considerando ciertos factores de riesgo clínicos y bioquímicos (Wilson y col., *Circulation* 97:1837-47, 1998; ATP HI, *JAMA* 285:2486-97, 2001), pero estas características no explican completamente el riesgo cardiovascular (Khot y col., *JAMA* 290:898-904, 2003; Greenland y col., *JAMA* 290:891-7, 2003).

- 25 Si pudiera mejorarse la capacidad de predecir la aparición futura de infarto de miocardio o accidente cerebrovascular, los individuos con tal riesgo podrían elegirse como diana para medidas preventivas y podría reducirse la incidencia global de esta causa principal de muerte.

- 30 La medición de múltiples proteínas y metabolitos en la sangre de un individuo ofrece la esperanza de una "ventana" en el estado bioquímico del individuo y podría proporcionar una mejor indicación del estado de su sistema cardiovascular y la probabilidad de que el sujeto experimentara un futuro infarto de miocardio o accidente cerebrovascular (Vasan, *Circulation* 113:2335-62, 2006). Con esta exposición razonada, los inventores realizaron un estudio con el objetivo de descubrir un perfil de biomarcadores moleculares (por ejemplo, un conjunto de proteínas, un conjunto de metabolitos, un conjunto de proteínas y metabolitos, o un conjunto de otros analitos que puedan incluir proteínas y/o metabolitos) en sangre o plasma en sangre y un algoritmo asociado que predijera una complicación cardiovascular o cerebrovascular adversa grave (MACCE) a medio plazo.

- 35 El documento US 2007/003981 desvela paneles de biomarcadores para identificar SCA (síndrome coronario agudo) en pacientes en riesgo. Estos paneles pueden incluir, entre otros, glutatión S-transferasa, factor tisular y alfa-fetoproteína, y algunos de estos paneles pueden incluir cinco o más biomarcadores.

- 40 Por tanto, sigue existiendo la necesidad de procedimientos de diagnóstico para predecir complicaciones cardíacas agudas graves. En particular, se necesitan procedimientos y composiciones fidedignas y rentables para permitir el diagnóstico y/o la predicción de complicaciones cardiovasculares y cerebrovasculares adversas graves.

Resumen de la invención

La presente invención proporciona un procedimiento de evaluación de la probabilidad de una complicación cardiovascular o cerebrovascular adversa grave en un ser humano, procedimiento que comprende:

- 45 medir una concentración, en una muestra basada en sangre de un ser humano, de un conjunto de analitos que comprende alfa-fetoproteína, antígeno del cáncer 125, glutatión S-transferasa y factor tisular;
- determinar un índice para el conjunto de analitos; e
- identificar el ser humano que tiene una alta probabilidad de una complicación cardiovascular o cerebrovascular adversa grave si el índice es mayor que cero, o una baja probabilidad de una complicación cardiovascular o cerebrovascular adversa grave si el índice es inferior o igual a cero,
- 50 en el que determinar un índice para el conjunto de analitos comprende:

normalizar la concentración medida de cada analito para obtener una concentración normalizada;

multiplicar la concentración normalizada de cada analito por un constante de analito para obtener un valor de analito; y

sumar el valor de analito de cada analito para obtener el índice.

- 5 Según diversas realizaciones, el conjunto de analitos también puede incluir CD40, fibrinógeno, IL-3, IL-8, SGOT y factor de von Willebrand.

Según diversas realizaciones, normalizar la concentración medida puede incluir restar de la concentración medida un valor promedio de la población para obtener un resultado y luego dividir el resultado entre una desviación estándar del valor promedio de la población.

- 10 Según diversas realizaciones, la muestra basada en sangre puede ser suero o plasma.

En otras realizaciones, el procedimiento también puede incluir transmitir, mostrar, almacenar o proporcionar el índice a un dispositivo de interfaz del usuario, un medio de almacenamiento legible por ordenador o un sistema de ordenador local o remoto.

Breve descripción de los dibujos

- 15 Lo anterior y otros objetos, características y ventajas de las presentes enseñanzas descritas anteriormente se entenderán más completamente a partir de la siguiente descripción de diversas realizaciones ilustrativas cuando se lean junto con los dibujos adjuntos. En los dibujos, números de referencia similares generalmente se refieren a las mismas partes en todas las vistas diferentes. Debe entenderse que los dibujos descritos más adelante son sólo para fines de ilustración. Los dibujos no pretenden limitar de ningún modo el alcance de las presentes enseñanzas.

- 20 La Figura 1 es una curva de las características de operación del receptor (ROC) que representa la predicción de la aparición de una MACCE dentro de dos años usando los niveles de los 20 primeros analitos de la Tabla 5 en una muestra de plasma según una realización a modo de ejemplo de las presentes enseñanzas.

- 25 La Figura 2 es una curva de las características de operación del receptor (ROC) que representa la predicción de la aparición de una MACCE dentro de dos años usando los niveles de los 10 primeros analitos de proteína de la Tabla 6 en una muestra de plasma según una realización a modo de ejemplo de las presentes enseñanzas.

Descripción detallada

- 30 Se han identificado analitos que son pronósticos de complicaciones cardiovasculares o cerebrovasculares adversas graves. Si uno o más de estos analitos están presentes en una muestra de fluido corporal de un individuo en cantidades diferentes de aquellas en un patrón, pueden ser indicativos de que el individuo está en riesgo de tener o desarrollar una complicación cardiovascular o cerebrovascular adversa grave.

- 35 Se identificaron setecientos veintitrés (723) analitos. Ciertos procedimientos según las presentes enseñanzas utilizan dos o más de estos analitos para predecir una complicación cardiovascular o cerebrovascular adversa grave. Específicamente, en estos procedimientos, una muestra de un sujeto puede cribarse para determinar si la muestra contiene niveles de dos o más de cada uno de estos setecientos veintitrés (723) analitos que son diferentes de una muestra patrón. Si la muestra contiene una cantidad de cada uno de dos o más de estos analitos que es diferente de la cantidad de estos analitos en un patrón, el cribado se considera un "cribado positivo" (es decir, el individuo está en riesgo de sufrir una complicación cardiovascular o cerebrovascular adversa grave). Normalmente puede obtenerse mayor sensibilidad y especificidad en la clasificación de muestras de complicación cardiovascular o cerebrovascular adversa grave usando un mayor número de analitos. Las muestras que posiblemente contienen estos analitos
- 40 pueden extraerse de múltiples muestras biológicas (por ejemplo, fluidos corporales, tejido, células) obtenidas de múltiples fuentes (por ejemplo, sangre completa, plasma en sangre, suero en sangre, orina, líquido cefalorraquídeo, células epiteliales y células endoteliales). Debe entenderse que todas las posibles combinaciones de los setecientos veintitrés (723) analitos desvelados en la presente memoria (y no sólo los cincuenta (50) analitos identificados en la Tabla 5) pueden usarse en procedimientos según las presentes enseñanzas.

- 45 Los setecientos veintitrés (723) analitos se identificaron usando la metodología descrita más completamente en el Ejemplo 1. Brevemente, mediante un protocolo de clasificación analítica específica se analizó un conjunto de setecientos veintitrés (723) analitos y se obtuvieron picos espectrales. Los apéndices 1 a 4 proporcionan moléculas específicas que comprenden los setecientos veintitrés (723) analitos analizados. Estos picos caracterizan moléculas específicas. Utilizando estos picos (que incluyen los 723 analitos) se identificaron cincuenta picos (además de otros picos preferidos como se describe más adelante). En tanto que los picos caracterizan moléculas, identificar cincuenta picos de entre un número tan grande de picos significa que los analitos se eligen de un grupo de moléculas que incluye más de las cincuenta moléculas correspondientes a los cincuenta analitos. Los analitos (es decir, las moléculas) pueden ser cualquier tipo de molécula. Los analitos (moléculas) incluyen, pero no se limitan a, proteínas, péptidos, aminoácidos, lípidos, esteroides, ácidos nucleicos, metabolitos y elementos. La Tabla 5
- 50

proporciona moléculas específicas que comprenden los cincuenta picos (es decir, identificar los analitos elegidos). La segunda columna de la Tabla 5 ordena por orden de importancia los picos por peso. Por consiguiente, el mayor pico de peso es el número de la clasificación 1 y el menor pico de peso es el número de la clasificación 50.

Ahora que se conocen los analitos (es decir, cualquiera de los setecientos veintitrés analitos identificados o cualquiera de los cincuenta analitos preferidos) pueden usarse para cribar un individuo para determinar si la cantidad o concentración absoluta de cada uno de dos o más de estos analitos en una muestra del individuo es diferente o no de la cantidad de cada uno de los dos o más analitos de un patrón, para determinar la concentración relativa individual de cada uno de los dos o más analitos en la muestra en comparación con un patrón y clasificar el individuo, a una cierta especificidad y sensibilidad, como que tiene o está en riesgo de desarrollar una complicación cardiovascular o cerebrovascular adversa grave. Por supuesto, la cantidad o concentración medida de un analito puede normalizarse antes de la comparación. La sensibilidad y especificidad deseada del ensayo puede elegirse basándose en el número de analitos examinados. El patrón puede ser una muestra real o datos empíricos previamente generados. El patrón puede obtenerse de una persona normal conocida. La persona normal conocida puede ser una persona sana y puede tener una ingesta alimentaria predeterminada durante un tiempo predeterminado antes del muestreo. Además, la muestra puede obtenerse de una persona normal conocida del mismo sexo que el sujeto. Alternativamente, los analitos podrían compararse con los de un sujeto conocido con complicación cardiovascular o cerebrovascular adversa grave, en cuyo caso se examinaría la similitud entre las dos muestras, o la concentración relativa del analito en comparación con un patrón. Un profesional médico puede usar diversas técnicas y/o kits para cribar muestras de sujetos con el fin de determinar el nivel y/o la cantidad de un analito particular en una muestra del sujeto. Ejemplos de tales ensayos se describen más adelante e incluyen, pero no se limitan a, un inmunoensayo, espectroscopía de masas, cromatografía, un análisis químico, un ensayo colorimétrico, un análisis espectrofotométrico, un análisis electroquímico y resonancia magnética nuclear. Adicionalmente, tales ensayos pueden realizarse en cualquier muestra biológica que incluye sangre completa, plasma en sangre, suero en sangre, líquido cefalorraquídeo, saliva, orina, líquido seminal, aspirado del pezón de la mama, líquido pancreático y combinaciones de los mismos. Estos ensayos se eligen basándose en que son los más aptos para detectar un analito particular, además de que son los más aptos para usar con una muestra biológica particular. Por consiguiente, pueden usarse múltiples ensayos para detectar los analitos deseados, y pueden analizarse muestras de una o más fuentes.

Un analito puede detectarse y/o cuantificarse usando uno o más procedimientos de separación. Por ejemplo, procedimientos de separación adecuados pueden incluir un procedimiento de espectrometría de masas tal como espectrometría de masas de ionización por electropulverización (EM-ESI), EM/EM-ESI, (EM)ⁿ/EM-ESI (n es un número entero mayor de cero), espectrometría de masas de tiempo de vuelo de ionización por desorción láser asistida por matriz (EM-TOF-MALDI), espectrometría de masas de tiempo de vuelo de desorción/ionización por láser potenciada por superficie (EM-TOF-SELDI), desorción/ionización sobre silicio (DIOS), espectrometría de masas de ión secundario (SIMS), cuadrupolo-tiempo de vuelo (Q-TOF), espectrometría de masas de ionización química a presión atmosférica (EM-APCI), EM/EM-APCI, (EM)ⁿ-APCI, espectrometría de masas de fotoionización a presión atmosférica (EM-APPI), EM/EM-APPI y (EM)ⁿ-APPI. Otros procedimientos de espectrometría de masas pueden incluir, entre otros, cuadrupolo, espectrometría de masas por transformada de Fourier (FTMS) y trampa de iones. Las técnicas espectrométricas que también pueden usarse incluyen espectroscopía de resonancia y espectroscopía óptica.

Otros procedimientos de separación adecuados incluyen separación por extracción química, cromatografía en columna, cromatografía de intercambio iónico, cromatografía de líquidos hidrófobos (fase inversa), isoelectroenfoque, electroforesis en gel de poli(acrilamida) (PAGE) unidimensional, electroforesis en gel de poli(acrilamida) bidimensional (2D-PAGE), u otras técnicas cromatográficas tales como cromatografía en capa fina, de gases o líquidos, o cualquier combinación de las mismas. En una realización, la muestra biológica que va a ensayarse puede fraccionarse antes de aplicarse al procedimiento de separación.

La unión en tándem de cromatografía (por ejemplo, cromatografía de líquidos ("CL")) y espectrometría de masas ("EM") puede ser útil para detectar y cuantificar uno o más de los analitos. La CL puede usarse para separar las moléculas, que pueden incluir un analito, en una muestra de un individuo. Una pequeña cantidad de la muestra, disuelta en un disolvente, puede inyectarse en el puerto de inyección del dispositivo de CL, que puede mantenerse a alta temperatura. La columna de CL del dispositivo contiene un sustrato sólido que puede ser tanto polar como apolar. Debido a las diferentes polaridades de las moléculas en la muestra, las moléculas tendrán diferentes afinidades por el sustrato sólido en la columna y eluirán en tiempo diferentes. Cuanto más fuerte sea la afinidad de la molécula por el sustrato, mayor será el tiempo de retención de la molécula en la columna. Cuando las moléculas salen de la columna entran en el espectrómetro de masas. El espectrómetro de masas ioniza las moléculas. En el modo de espectrometría de masas en tándem, si el sistema puede normalizarse apropiadamente, cada compuesto enviado a un espectrómetro de masas se fragmenta en iones de diversas masas y abundancias formando un patrón de firma único para esa sustancia. Comparando el espectrógrafo de masas en tándem de cada pico con una base de datos computerizada, el ordenador puede identificar normalmente las moléculas con un alto grado de certeza. Alternativamente, o adicionalmente, esta comparación puede llevarse a cabo por inspección humana. Una vez se ha establecido una identidad, el ordenador integra el área bajo cada pico y así determina la cantidad relativa de cada molécula en la mezcla. Hasta el punto de que cualquiera de las moléculas se identifica como un analito, la cantidad del analito puede compararse con la cantidad del analito de un patrón para determinar si hay una diferencia.

Los analitos también pueden detectarse y/o cuantificarse mediante procedimientos que no requieren separación física de los propios analitos. Por ejemplo, la espectroscopia de resonancia magnética nuclear (RMN) puede usarse para resolver un perfil de un analito de una mezcla compleja de moléculas. Un uso análogo de RMN para clasificar tumores se desvela, por ejemplo, en Hagberg, NMR Biomed. 11: 148-56 (1998). Procedimientos adicionales incluyen tecnologías de amplificación de ácidos nucleicos, que pueden usarse para determinar un perfil de analito sin separación física de moléculas individuales (véanse, por ejemplo, Stordeur y col., J. Immunol. Methods 259: 55-64 (2002) y Tan y col., Proc. Nat'l Acad. Sci. USA 99: 11387-11392 (2002)).

Un analito en una muestra también puede detectarse y/o cuantificarse, por ejemplo, combinando el analito con un resto de unión que puede unirse específicamente al analito. El resto de unión puede incluir, por ejemplo, un miembro de un par ligando-receptor, es decir, un par de moléculas que puede tener una interacción de unión específica. El resto de unión también puede incluir, por ejemplo, un miembro de un par de unión específica tal como anticuerpo-antígeno, enzima-sustrato, ácido nucleico-ácido nucleico, proteína-ácido nucleico, proteína-proteína, u otros pares de unión específica conocidos en la materia. Pueden diseñarse proteínas de unión que tienen afinidad potenciada por una diana. Opcionalmente, el resto de unión puede ligarse con una marca detectable tal como una marca de partícula enzimática, fluorescente, radiactiva, fosforescente o coloreada. El complejo marcado puede detectarse, por ejemplo, visualmente o con la ayuda de un espectrofotómetro u otro detector, y/o puede cuantificarse.

Un analito también puede detectarse y/o cuantificarse usando técnicas de electroforesis en gel disponibles en la materia. En la electroforesis en gel bidimensional, las moléculas se separan primero en un gel de gradiente de pH según su punto isoelectrico. Entonces, el gel resultante puede colocarse sobre un segundo gel de poli(acrilamida) y las moléculas separarse según el peso molecular (véase, por ejemplo, O'Farrell J. Biol. Chem. 250: 4007-4021 (1975)). Un analito para complicación cardiovascular o cerebrovascular adversa grave puede detectarse aislando primero moléculas de una muestra obtenida de un individuo del que se sospecha que está en riesgo de una complicación cardiovascular o cerebrovascular adversa grave y luego separando las moléculas por electroforesis en gel bidimensional para producir un patrón de electroforesis en gel bidimensional característico. El patrón puede entonces compararse con un patrón de gel producido separando, bajo las mismas condiciones o similares, moléculas aisladas del patrón (por ejemplo, sujetos sanos o con complicación cardíaca aguda grave). El patrón de gel patrón puede almacenarse en, y recuperarse de, una base de datos electrónica de patrones de electroforesis. Por tanto, puede determinarse si la cantidad del marcador en el sujeto es diferente de la cantidad en el patrón. La presencia de una pluralidad, por ejemplo, de dos a cincuenta, de analitos de complicación cardiovascular o cerebrovascular adversa grave en el gel bidimensional en una cantidad diferente a un patrón normal conocido indica un cribado positivo para una complicación cardiovascular o cerebrovascular adversa grave en el individuo. Por tanto, el ensayo permite la predicción y el tratamiento de complicaciones cardiovasculares o cerebrovasculares adversas graves.

Un analito también puede detectarse y/o cuantificarse usando cualquiera de un amplio intervalo de técnicas de inmunoensayo disponibles en la materia. Por ejemplo, la forma de inmunoensayo de tipo sándwich puede usarse para detectar y/o cuantificar un analito en una muestra de un sujeto. Alternativamente pueden usarse procedimientos inmunohistoquímicos convencionales para detectar y/o cuantificar la presencia de un analito en una muestra usando una o más proteínas de unión marcadas.

En un inmunoensayo de tipo sándwich se usan generalmente dos anticuerpos que pueden unirse a analitos, por ejemplo, uno inmovilizado sobre un soporte sólido y uno libre en disolución y marcado con un compuesto químico detectable. Ejemplos de marcas químicas que pueden usarse para el segundo anticuerpo incluyen radioisótopos, compuestos fluorescentes y enzimas u otras moléculas que generan productos coloreados o electroquímicamente activos cuando se exponen a un sustrato reactivo o enzimático. Si una muestra que contiene el analito se coloca en este sistema, el analito se une tanto al anticuerpo inmovilizado como al anticuerpo marcado para formar un inmunocomplejo de tipo "sándwich" sobre la superficie del soporte. El analito complejo se detecta lavando los componentes de muestra no unidos y el exceso de anticuerpo marcado, y midiendo la cantidad de anticuerpo marcado complejo con el analito sobre la superficie del soporte. Alternativamente, el anticuerpo libre en disolución, que puede marcarse con un resto químico, por ejemplo, un hapteno, puede detectarse por un tercer anticuerpo marcado con un resto detectable que se une al anticuerpo libre o, por ejemplo, el hapteno acoplado al mismo.

Tanto el inmunoensayo de tipo sándwich como los procedimientos inmunohistoquímicos de tejido son altamente específicos y muy sensibles, siempre que se usen marcas con buenos límites de detección. Una revisión detallada del diseño del ensayo inmunológico, teoría y protocolos puede encontrarse en numerosos textos en la materia que incluyen Butt, W.R., Practical Immunology, ed. Marcel Dekker, New York (1984) y Harlow y col. Antibodies, A Laboratory Approach, ed. Cold Spring Harbor Laboratory (1988).

En general, las consideraciones del diseño de inmunoensayos incluyen la preparación de anticuerpos (por ejemplo, anticuerpos monoclonales o policlonales) que tienen especificidad de unión suficientemente alta por la diana para formar un complejo que pueda distinguirse de forma fidedigna de productos de interacciones no específicas. Como se usa en la presente memoria, el término "anticuerpo" se entiende que significa proteínas de unión, por ejemplo, anticuerpos u otras proteínas que comprenden un dominio de unión similar a región variable de la inmunoglobulina que tiene las afinidades y especificidades de unión apropiadas por la diana. Cuanto mayor sea la especificidad de

unión del anticuerpo, menor será la concentración diana que puede detectarse. Como se usa en la presente memoria, los términos “unión específica” o “se une específicamente” se entiende que significan que el resto de unión, por ejemplo, una proteína de unión, tiene una afinidad de unión por la diana de más de aproximadamente 10^5 M^{-1} , más preferentemente más de aproximadamente 10^7 M^{-1} .

- 5 Los anticuerpos para un analito diana aislado que son útiles en ensayos para predecir una complicación cardiovascular o cerebrovascular adversa grave en un individuo pueden generarse usando procedimientos inmunológicos convencionales muy conocidos y descritos en la materia. Véase, por ejemplo, Practical Immunology, arriba. Brevemente, un analito aislado puede usarse para producir anticuerpos en un huésped xenogénico tal como un ratón, cabra u otro mamífero adecuado. El analito puede combinarse con un adyuvante adecuado que puede
10 potenciar la producción de anticuerpos en el huésped, y puede inyectarse en el huésped, por ejemplo, por administración intraperitoneal. Puede usarse cualquier adyuvante adecuado para estimular la respuesta inmunitaria del huésped. Un adyuvante comúnmente usado es adyuvante completo de Freund (una emulsión que comprende células microbianas muertas y secadas y está disponible de, por ejemplo, Calbiochem Corp., San Diego, o Gibco, Grand Island, NY). Si se desean múltiples inyecciones de antígeno, las inyecciones posteriores pueden comprender el antígeno en combinación con un adyuvante incompleto (por ejemplo, emulsión libre de células). Los anticuerpos policlonales pueden aislarse del huésped productor de anticuerpos extrayendo suero que contiene anticuerpos para la proteína de interés. Los anticuerpos monoclonales pueden producirse aislando células huésped que producen el anticuerpo deseado, fusionando estas células con células de mieloma usando procedimientos convencionales conocidos en la técnica de la inmunología y cribando células híbridas (hibridomas) que reaccionan específicamente
20 con la diana y tienen la afinidad de unión deseada.

También pueden producirse biosintéticamente dominios de unión de anticuerpos y manipularse la secuencia de aminoácidos del dominio de unión para potenciar la afinidad de unión con un epítipo preferido sobre la diana. Se entienden bien las metodologías de anticuerpos específicos y se describen en la bibliografía. Una descripción más detallada de su preparación puede encontrarse, por ejemplo, en Practical Immunology, (arriba).

- 25 Además, pueden usarse sitios de unión de anticuerpos biosintéticos genéticamente manipulados, también conocidos en la técnica como BABS o sFv, para determinar si una muestra contiene un analito. Los procedimientos para preparar y usar BABS que comprenden (i) dímeros V_H y V_L sintéticos asociados no covalentemente o unidos por enlace disulfuro, (ii) sitios de unión monocatenarios de V_H - V_L ligados covalentemente, (iii) dominios V_H o V_L individuales, o (iv) sitios de unión de anticuerpos monocatenarios se desvelan, por ejemplo, en las patentes de EE.UU. n°: 5.091.513; 5.132.405; 4.704.692; y 4.946.778. Además, BABS que tienen especificidad necesaria por el
30 analito pueden derivarse por clonación de anticuerpos en fago de bibliotecas de genes combinatorias (véase, por ejemplo, Clackson y col. Nature 352: 624-628 (1991)). Brevemente, fagos, cada uno los cuales expresa sobre sus superficies de la cubierta BABS que tienen regiones variables de inmunoglobulina codificadas por secuencias de genes de la región variable derivadas de ratones preinmunizados con un analito aislado, o un fragmento de las mismas, se criban para actividad de unión contra el analito inmovilizado. Los fagos que se unen al analito inmovilizado se recogen y el gen que codifica BABS puede secuenciarse. Entonces, las secuencias de ácidos nucleicos resultantes que codifican BABS de interés pueden expresarse en sistemas de expresión convencionales para producir la proteína de BABS.

- 40 Un analito aislado también puede usarse para el desarrollo de kits de diagnóstico y de evaluación de otros tejidos y ensayos para monitorizar el nivel de los analitos en una muestra de tejido o de líquido. Por ejemplo, el kit puede incluir anticuerpos u otras proteínas de unión específica que se unen específicamente a uno o más analitos y que permiten la presencia y/o la cantidad del uno o más analitos que van a detectarse y/o cuantificarse en una muestra de tejido o de líquido.

- 45 Se contemplan kits adecuados para detectar uno o más analitos que incluyen, pero no se limitan a, un receptáculo u otros medios para capturar una muestra que va a evaluarse y un medio para detectar la presencia y/o la cantidad en la muestra de uno o más de los analitos descritos en la presente memoria. Los medios para detectar en una realización incluyen, pero no se limitan a, uno o más anticuerpos específicos para estos analitos y medios para detectar la unión de los anticuerpos a estos analitos, por ejemplo, por un inmunoensayo de tipo sándwich convencional como se describe en la presente memoria. Si va a detectarse la presencia de un analito localizado dentro de una célula (por ejemplo, como de una muestra de tejido), el kit también puede comprender medios para
50 romper la estructura celular de manera que se expongan los componentes intracelulares.

- Los analitos de las presentes enseñanzas pueden incluir ácidos nucleicos de una secuencia particular. Uno o más de los analitos pueden detectarse y/o cuantificarse determinando una cantidad o concentración absoluta del analito de ácido nucleico en una muestra usando, por ejemplo, PCR cuantitativa en tiempo real (RT-PCR) y comparando la
55 cantidad medida con un patrón para determinar una concentración relativa del analito ácido nucleico en una muestra. La RT-PCR mide eficazmente la cantidad de un analito ácido nucleico resultante de PCR. Un resultado positivo representa una cantidad medida del analito ácido nucleico que es diferente de la cantidad del analito de un patrón, o una concentración relativa que tiene un valor por encima o por debajo de cero.

- 60 Pueden revelarse cebadores que son complementarios a la secuencia de ácidos nucleicos de un analito de ácido nucleico particular. Estos cebadores dirigen una polimerasa para copiar y amplificar ese ácido nucleico particular. La

RT-PCR detecta la acumulación del analito de ácido nucleico amplificado durante la reacción. Durante la fase exponencial de la reacción de PCR puede medirse el analito de ácido nucleico que se acumula. Un patrón de calibración que tiene una concentración de ácido nucleico conocida puede usarse para preparar una curva patrón de la que puede extrapolarse la cantidad del analito de ácido nucleico en la muestra probada.

- 5 Una vez se conoce la cantidad o concentración absoluta de un analito de ácido nucleico en una muestra, puede compararse con la cantidad del analito de un patrón para determinar una concentración relativa de un analito de ácido nucleico en una muestra. El patrón para la clasificación de sujetos con complicaciones cardiovasculares o cerebrovasculares adversas graves puede determinarse por medios empíricos. Por ejemplo, la cantidad puede determinarse amplificando el analito de ácido nucleico en una muestra de una población de uno o más individuos normales conocidos y analizando cuantitativamente la cantidad de un analito de ácido nucleico en la población.

- 10 Por tanto, pueden realizarse formas adicionales de análisis químico de una muestra. Por ejemplo, pueden llevarse a cabo pruebas cuantitativas que indican las cantidades o concentraciones absolutas de cada analito en una muestra. Un ensayo colorimétrico es un análisis químico cuantitativo que mide la intensidad de color producida haciendo reaccionar una muestra con un reactivo como representante de la cantidad del material ensayado en una muestra. Pueden proporcionarse reactivos que, cuando reaccionan con cualquier analito, producen un color en la muestra del ensayo. La intensidad de ese color puede depender de la cantidad de analito en la muestra. Comparando la intensidad con una carta de color calibrado y/o patrón puede determinarse la cantidad de analito en la muestra. Esta cantidad puede compararse entonces con la cantidad de analito de un patrón (tal como de una persona normal conocida) para determinar una concentración relativa de los analitos en una muestra.

- 20 Adicionalmente pueden usarse análisis de orina para determinar la cantidad o concentración absoluta de un analito en una muestra de orina. Las muestras de orina se prueban con una variedad de instrumentos y técnicas diferentes. Algunas pruebas usan tiras reactivas que son delgadas tiras de plástico que cambian de color en presencia de sustancias específicas. Las tiras reactivas podrían usarse para medir la cantidad de un analito.

- 25 Además, la determinación de una concentración relativa comparando el nivel absoluto o la concentración de cada uno de los al menos dos analitos con el nivel de cada uno de dos analitos de un patrón puede hacerse como una medida de cribado preventiva y no sólo cuando se observe una complicación cardiovascular o cerebrovascular adversa (es decir, después de que haya podido progresar la enfermedad). Por ejemplo, asumiendo que no hay pruebas de una complicación cardiovascular o cerebrovascular adversa, los sujetos podrían monitorizarse después de una cierta edad y a intervalos predeterminados con el fin de obtener un diagnóstico de tener o estar en riesgo de tener una complicación cardiovascular o cerebrovascular adversa grave en el momento más pronto posible. Hasta el punto de que el cribado sea positivo, un profesional médico podría recomendar monitorización adicional de la progresión de la enfermedad, y/o el profesional médico podría empezar tratamientos con un fármaco u otra terapia.

- 30 Los resultados del análisis, que incluyen, por ejemplo, la cantidad o concentración absoluta de cada uno de los analitos, la concentración relativa de cada uno de los analitos con respecto a un patrón y/o una probabilidad de tener o estar en riesgo de tener una complicación cardiovascular o cerebrovascular adversa grave, pueden mostrarse o proporcionarse a un dispositivo de interfaz del usuario, un medio de almacenamiento legible por ordenador o un sistema de ordenador local o remoto. El mostrar o proporcionar un resultado o diagnóstico significa que los resultados de cualquiera de los anteriores análisis se comunican a un usuario usando cualquier medio tal como, por ejemplo, por vía oral, escrita, visualización visual, etc., medio legible por ordenador o sistema informático. Será evidente para un experto en la materia que proporcionar el resultado no se limita a proporcionar a un usuario o a un componente(s) externo(s) ligado(s) tal como un sistema informático o memoria informática, sino que alternativamente o adicionalmente puede proporcionarse a componentes externos tales como cualquier medio legible por ordenador. Los medios legibles por ordenador pueden incluir, pero no se limitan a, discos duros, disquetes, CD-ROM, DVD y DAT. Los medios legibles por ordenador no incluyen ondas portadoras u otras formas de ondas para la transmisión de datos. Será evidente para un experto en la materia que la diversa evaluación de muestras y procedimientos de diagnóstico desvelados y reivindicados en la presente memoria pueden, pero no necesitan, implementarse por ordenador y que, por ejemplo, la etapa de mostrar o proporcionar puede hacerse, por ejemplo, comunicando a una persona oralmente o por escrito (por ejemplo, escrito a mano).

- 40 Además, los analitos pueden usarse para validar modelos animales de complicaciones cardiovasculares o cerebrovasculares adversas graves. Por ejemplo, en cualquier modelo particular, una muestra podrían analizarse para determinar si los niveles de los analitos en el animal son los mismos que los niveles de los analitos en un sujeto con complicación cardiovascular o cerebrovascular adversa grave conocido. Esto validaría el modelo, por ejemplo, para probar fármacos candidatos en el modo descrito anteriormente.

Ejemplo 1

- 55 Este ejemplo describe la metodología usada para identificar analitos para complicaciones cardiovasculares o cerebrovasculares adversas graves. Brevemente, ciento treinta y seis (136) sujetos se incluyeron en el estudio. Una clasificación de los sujetos en tanto categorías de enfermedad como de control se logró identificando sujetos que había padecido una complicación cardiovascular o cerebrovascular adversa grave, definida como un infarto de miocardio, intervención coronaria percutánea o muerte, en el plazo de dos años desde una cateterización cardíaca

índice.

Se obtuvieron muestras de plasma congeladas de un conjunto de muestras obtenidas en un estudio a largo plazo de individuos que habían padecido cateterización cardíaca en el Centro médico de la Universidad Duke (el estudio CATHGEN). Las muestras de plasma fueron de 136 sujetos en estudio, 68 casos de enfermedad y 68 controles de correspondencia (como se define más adelante). Los casos de enfermedad y los controles se agruparon según su índice de enfermedad de las arterias coronarias (EAC), edad, sexo y raza. Algunos de los sujetos en estudio en tanto los casos de enfermedad como los grupos de control no presentaron signos de enfermedad de las arterias coronarias tras la cateterización cardíaca y angiografía. Otros sujetos tuvieron diversos grados de gravedad de estenosis de la arteria coronaria como se indica por mayores puntuaciones del índice de EAC.

Los parámetros clínicos que se registraron para todas las muestras fueron aquellos mostrados en la siguiente Tabla 1. Estos parámetros, que no fueron aquellos con los que se correspondieron los casos enfermedad y los controles, se incluyeron como factores clínicos adicionales en algunos de los análisis estadísticos (como se trata más adelante).

Tabla 1: Parámetros clínicos de sujetos en estudio

Característica	Categorías/ Unidad	Descripción
Edad	Años	Edad (número entero) en la cateterización
Raza	• Caucásica • Africana • Americana • Otra	
Índice de EAC		Índice de gravedad de enfermedad de las arterias coronarias
IMC	• <25 • 25 - 29,9 • ≥30	Índice de masa corporal
Dislipidemia	• Sí • No	Historia de dislipidemia
Historia familiar	• Sí • No	Historia familiar de enfermedad coronaria
Hipertensión	• Sí • No	Sí = Historia de hipertensión o BP sistólica > 140 o BP diastólica > 90
Diabetes	• Sí • No	Sí = Diabetes de tipo I o II
Estado de tabaquismo	• Sí • No	Historia de tabaquismo
Tasa de eliminación de creatinina	Medida continua	Aproximación de Cockcroft-Gault basándose en la edad, peso, creatinina en plasma y sexo
Uso diario de aspirina	• Sí • No	Obtenido de la curación experta de la tabla de medicaciones
Uso de estatinas	• Sí • No	Obtenido de la curación experta de la tabla de medicaciones

Se identificó que los casos de enfermedad tenían las siguientes características. Antes de la cateterización índice, los sujetos no tenían historia de cirugía de injerto de derivación de las arterias coronarias (IDAC), intervención coronaria percutánea (ICP) para complicaciones cardíacas o infarto de miocardio (IM). Los sujetos no tuvieron IDAC posterior. Las fracciones expulsadas de los sujetos fueron > 40% (en la cateterización índice) y no estuvieron ausentes. El IM, ICP o muerte se produjo en el plazo de 2 años desde la cateterización índice y el muestreo de sangre. Se excluyeron aquellos sujetos que claramente tuvieron muerte no relacionada con la enfermedad de las arterias coronarias (es decir, hipertensión pulmonar, cáncer, etc.). Se excluyeron aquellos sujetos con una complicación relacionada con ICP en el plazo de 7 días desde la cateterización índice.

- 5 Se identificó que los casos de control tenían las siguientes características. Antes de la cateterización índice, los sujetos no tenían historia de cirugía de injerto de derivación de las arterias coronarias (IDAC), intervención coronaria percutánea (ICP) para complicaciones cardíacas o infarto de miocardio (IM). Los sujetos no tuvieron IDAC posterior. Las fracciones expulsadas de los sujetos fueron > 40% (en la cateterización índice) y no estuvieron ausentes. No se produjo IM, ICP o muerte durante un periodo de al menos 2 años tras la cateterización índice y el muestreo de sangre.

Las muestras se sometieron a bioanálisis exhaustivo usando las plataformas presentadas en la siguiente Tabla 2 y se describen en detalle más adelante.

Tabla 2: Procedimientos de bioanálisis

Proteómica	Número de analitos usados en análisis estadísticos
Proteómica de descubrimiento por EM/EM-MALDI-CL	217
Proteómica elegida como diana (inmunoensayo multiplexado)	66
Metabolómica	
Metabolómica de EM/CL de lípidos	125
Metabolómica de EM/CL polar	164
Metabolómica de EM/CG	130
Ácidos grasos libres	21

- 10 Se midieron satisfactoriamente un total de 723 analitos en la parte de bioanálisis del estudio. La contribución al total de cada una de las plataformas bioanalíticas se muestra en la Tabla 2 anterior. Los resultados de las plataformas bioanalíticas se analizaron basándose en plataforma por plataforma y también se combinaron en un conjunto de datos integrado único para el posterior análisis estadístico (véase la discusión más adelante).

Metabolómica de espectrometría de masas en tándem con cromatografía de líquidos ("CL") ("EM/CL") de lípidos

- 15 Este procedimiento utiliza condiciones de cromatografía de líquidos y espectrometría de masas que se optimizan para la resolución y la detección de moléculas de lípidos.

- 20 Se usaron los siguientes materiales según el procedimiento. Los disolventes y los reactivos, que incluyen agua, metanol e isopropanol y diclorometano (calidad para HPLC) se compraron de VWR (EE.UU.). El ácido fórmico (99%), el acetato de amonio, el diclorometano y la reserpina se compraron de Sigma-Aldrich (Milwaukee, WI). La precolumna de HPLC (Symmetry 300 C4 300 Å 3,5 µ 2,1 x 10 mm) y la columna analítica (C4 300 Å 3,5 µ 2,1 x 150 mm) se compraron de Waters (Milford, Ma). Los viales del inyector automático, viales de recuperación total de 300 µl y los tapones de rosca se compraron de Waters (Milford, Ma). Los tubos Eppendorf de 500 µl para el almacenamiento se compraron de VWR. Los patrones internos de lípidos 17:0 LPC, 24:0 PC y 40:0 PC se compraron de Avanti Polar Lipids (Alabaster, AL) y 30:0 PC y 51:0 TG se compraron de Sigma-Aldrich (Milwaukee, WI).

- 30 Las muestras se extrajeron según el siguiente protocolo. Cinco patrones internos se añadieron en cantidades conocidas a la disolución de extracción para la normalización de datos. Se prepararon disoluciones madre de patrones internos disolviendo cantidades apropiadamente pesadas de cada compuesto de patrón interno en una disolución 1:9 de DCM/IPA para lograr concentraciones finales en el intervalo de 1-3 mg/ml dependiendo del lípido particular. La disolución de extracción se creó tomando alícuotas de 300 µl de cada una de las 5 disoluciones madre en 300 ml de una disolución 1:9 de DCM/IPA (referencia a SOP: BG-QC27 R01). Esta disolución se usó como la disolución de extracción de lípidos con los cinco patrones internos añadidos en cantidades conocidas con relaciones de concentraciones de 2:1:2:3:2 (µg/ml) para 17:0 LPC, 24:0 PC, 30:0 PC, 40:0 PC, 51:0 TG respectivamente.

- 35 Las muestras se prepararon para análisis de EM/CL según el siguiente protocolo. Las muestras de cada lote se sacaron del congelador junto con muestras de control de calidad y se descongelaron sobre hielo. A cada uno de los viales de muestra que contenían 10 µl de plasma se añadió un volumen apropiado de disolución de extracción. El vial se agitó cuidadosamente con vórtex y se centrifugó a 14.000 rpm a 24°C durante 10 minutos. El sobrenadante se transfirió a un vial intermedio y se mezcló adicionalmente por vórtex durante algunos segundos, se pipetearon 2 alícuotas de 60 µl cada una en dos viales de inyector automático con código de barras para el análisis de EM/CL como duplicados. El sobrenadante restante se almacenó a -40°C para su uso en análisis de EM/EM/CL. Se desechó el extracto restante con el sedimento.

El análisis de cromatografía de líquidos de alta resolución (HPLC) se realizó según el siguiente protocolo. El aparato

cromatográfico consistió en un sistema de HPLC Waters 2795 Alliance con un sistema de gradiente cuaternario, inyector automático, un conjunto de calentador externo de la columna a 45°C y una columna de HPLC analítica con una precolumna en línea. Los analitos se separaron y se administraron al espectrómetro de masas a una velocidad de flujo de 350 µl/min. Se usó un gradiente binario (0 min - 20% de B, 2 min - 20% de B, 4 min - 80% de B, 20 min - 100% de B, 25 min - 100% de B, 25,1 min - 20% de B, 35 min - 20% de B) para la separación de analitos con el disolvente A que comprendía 95% de H₂O y 5% de MeOH y el disolvente B que consistía en 99% de MeOH y 1% de H₂O. Ambas corrientes de disolvente contuvieron acetato de amonio 10 mM y 0,1% de ácido fórmico. Se inyectaron 5 µl de muestra en un bucle de muestra de 20 µl y se cargaron sobre la precolumna antes de la separación por la columna analítica. Entre cada muestra, la aguja de inyección, el bucle y la jeringuilla se lavaron con disolución de isopropanol/MeOH.

El análisis de EM/CL se realizó en un instrumento de cuadrupolo-tiempo de vuelo Waters/Micromass (Waters, Q-ToF™) equipado con una fuente de ESI LockSpray. La temperatura del bloque de la fuente de ESI se mantuvo a 120°C y la temperatura de desolvatación se fijó a 320°C. La sensibilidad de la señal de EM se optimizó usando el pico de reserpina a 609,2814. La calibración de masa se realizó en modo de EM usando picos de polialanina protonados. La resolución promedio del instrumento de ~8500 se eligió como diana para los 5 picos de IS con una precisión de la calibración con LockMass de 3 a 5 ppm. Basándose en estudios previos, los datos de EM se adquirieron de 7 a 23 minutos durante la elución de CL. Todos los datos de EM/CL se adquirieron en modo centroide barriendo a m/z 300 -1000 en 0,3 segundos.

El análisis de EM/EM posterior para la identificación de analitos se realizó en el mismo sistema de EM/CL con la condición de CL idéntica. La calibración de EM/EM se realizó usando fragmentos de picos de polialanina de precursor a m/z 1084. Todos los datos de EM/EM se adquirieron en modo centroide barriendo a m/z 100 - 1000 en 1 ó 2 segundos. Cuando fue necesario se aplicó la corrección del tiempo de retención de CL a la lista diana. Los datos de EM/EM se recogen por análisis elegido como diana.

La detección de picos se realizó según el siguiente protocolo. Se captaron los picos y se integraron de los archivos de datos brutos de las ejecuciones aceptadas. Sólo los picos que alcanzaron los umbrales definidos pasaron al procedimiento de alineamiento. Cada pico se caracterizó por un valor de m/z y tiempo de retención, por ejemplo, un pico detectado a m/z 510 y eluido a 7,88 minutos se etiquetaría 510_0788. Cada muestra se procesa independientemente.

Después de completarse la captación de picos, los archivos .MLL generados por el algoritmo de captación de picos se "alinearon" entre sí. El programa fija primero una ventana de tiempo de retención usando los picos de patrón interno de todas las inyecciones y luego "alinean" lógicamente cada analito dentro del lote basándose en el tiempo de retención medio de los picos de referencia. Al final del procedimiento, el tiempo de retención físico puede cambiarse de 510_0785 a 510_0788 y un umbral adicional puede aplicarse a las áreas de picos.

El control de calidad (CC) se evaluó para cada lote por muchas mediciones, incluyendo calcular el porcentaje de desviación estándar relativa (% de DER) de los patrones internos a lo largo de todas las muestras, representar las áreas extraídas (intensidad) de picos del patrón interno en función del tiempo para identificar cualquier tendencia y comprobar las variaciones del tiempo de retención tanto antes como después del alineamiento. La inspección de CC visual de cada inyección también tuvo lugar durante los análisis. Estas revisiones se basaron en cómo de bien se recubrían las inyecciones duplicadas, el comportamiento de las muestras de CC con el tiempo y las tendencias generales de los instrumentos. Pueden producirse cambios menores en la retención y el área del pico durante el estudio y se corrigieron en el procedimiento de normalización de datos (por ejemplo, ligeras tendencias temporales), pero cambios más significativos (como inyecciones parciales) pueden requerir volver a analizarse.

En el procedimiento de medición de EM/CL, cada especie molecular de interés normalmente aparece como múltiples entidades químicas. Esto puede ser debido a varios factores: primero, la presencia de múltiples isótopos estables de elementos constituyentes y, segundo, la formación de iones pseudomoleculares de los analitos mediante la formación de aductos con sodio (Na), potasio (K), amonio (NH₄) y otros analitos produciendo múltiples tipos de iones. Además, el procedimiento de ionización de espectrometría de masas puede ir frecuentemente acompañado de fragmentación de la molécula original. Por consiguiente, los conjuntos de datos adquiridos están exagerados un orden de magnitud en el número de picos que representan un analito. Estos picos adicionales que son redundantes en su contenido de información contribuyen a la introducción de ruido, disminución de la potencia estadística, enmascaramiento de efectos en el análisis bioinformático y, debido al gran número de picos, pueden hacer más difícil extraer conclusiones científicas. Se usó una plataforma computacional automatizada, que agrupa entidades medidas según sus compuestos putativos de origen, siguiendo la captación de picos y el alineamiento de espectros, y antes o paralelamente a la normalización y al análisis estadístico de datos. El algoritmo central barre los datos según el tiempo de retención (RT) con una estrecha ventana de tiempo en movimiento. Los picos candidatos de tiempo de retención similar se agrupan según patrones de diferencias de masas conocidos dependiendo de las características químicas específicas de la plataforma.

Tras la detección y el alineamiento de picos se producen varias etapas adicionales del procesamiento de datos antes de que los datos estén disponibles para los análisis estadísticos unifactoriales y multifactoriales. Estas etapas incluyen la eliminación de picos que se asignan a blancos de muestra/estudio y otros picos de referencia, corrección

y normalización de lotes. La completitud de estas etapas representa el hito denominado en lo sucesivo “conjunto de datos bloqueados”.

La identificación de espectrometría de masas y el seguimiento del procesamiento de datos generó una lista de analitos para el posterior análisis estadístico. Los analitos primarios, normalmente el pico más abundante para cada familia de analitos, se sometieron a EM/EM elegida como diana usando el mismo sistema de EM/CL o similar como experimento de identificación inicial. Los espectros de EM/EM/CL elegidos como diana se adquirieron en modo centroide. Los espectros de EM/EM obtenidos por análisis de EM/EM que elige diana de los analitos se interpretaron por una combinación de tiempo de retención, masa, regla del nitrógeno, iones específicos del fragmento de la cadena lateral de ácidos grasos (AG) y comparación con espectros de referencia de una biblioteca de EM/EM de lípidos.

Metabolómica de espectrometría de masas en tándem con cromatografía de gases (EM/CG)

La CG tiene la ventaja de alta eficiencia de separación y tiempos de retención buenos con detección de masa sensible y selectiva (impacto de electrones). Sin embargo, muchos compuestos contienen grupos funcionales polares que son tanto térmicamente lábiles a las temperaturas requeridas para su separación como son nada volátiles en absoluto. Además, la forma del pico de compuestos con grupos funcionales polares puede ser insatisfactoria debido a la interacción con la columna no deseada tal como adsorción irreversible. Con el fin de tratar estos asuntos puede ser necesaria la derivatización de los compuestos antes de análisis de CG.

Los siguientes materiales se usaron según el procedimiento. Los disolventes y reactivos incluyeron metanol; piridina; clorhidrato de etoxiamina; y N-metil-N-trimetilsililtrifluoroacetamida (MSTFA). Los patrones internos incluyeron leucina-D3; fenilalanina-D5; ácido glutámico-D3; glucosa-D7; ácido cólico-D4; alanina-D4; trifluoroacetilanzano (TFAA); difluorobifenilo (DFBP); y ftalato de dicitohexilo (DCHP).

Las muestras se descongelaron a temperatura ambiente y se añadieron 10 µl de un patrón 250 ng/µl de leucina-D3, fenilalanina-D5, ácido glutámico-D3 y glucosa-D7 en H₂O seguido de 400 µl de metanol (para la precipitación de proteína). Las muestras se agitaron con vórtex, se centrifugaron durante 10 minutos a 10000 rpm y el sobrenadante se transfirió a un vial de inyector automático. Después de una etapa de secado con nitrógeno se añadieron 10 µl de un patrón 250 ng/µl de ácido cólico-D4 y alanina-D4 en piridina. Antes de tapar el vial del inyector automático se añadieron 30 µl de disolución de clorhidrato de etoxiamina en piridina y la muestra se incubó a 40°C durante 90 minutos.

Después de la oximación se añadieron los patrones internos restantes: 10 µl de un patrón 250 ng/µl de difluorobifenilo, ftalato de dicitohexilo y trifluoroacetilanzano en piridina. La muestra se sililó añadiendo 100 µl de MSTFA y calentando durante 50 minutos a 40°C. Antes de la inyección en el CG, las muestras se centrifugaron durante 10 minutos a 3500 rpm.

Los análisis de EM/CG emplearon un cromatógrafo de gases Agilent 6890 N equipado con un inyector PTV (vaporizador de temperatura programada) y un inyector automático CTC Analytics Combi-Pal. Para la detección se usó un detector selectivo de masa Agilent 5973. El sistema se controló por el software Enhanced Chemstation G1701CA versión D.01.02.

Todos los compuestos que se eluyen de la columna de EM/CG se detectaron en modo de barrido completo. Cada pico se caracterizó por su tiempo de retención y varios fragmentos (valores de m/z). La cantidad de datos (es decir, el número de variables) se redujo aplicando un procedimiento de procesamiento diana.

El enfoque diana reduce el número de variables al menos un factor de 20. La mayoría de los artefactos del procedimiento podrían eliminarse de los datos no incluyendo estos picos en la tabla diana. Para este estudio se compiló una tabla diana específica del estudio. En este procedimiento se usó la información de proyectos de EM/CG en plasma humano previos. Se seleccionaron varios cromatogramas representativos de las muestras del estudio que incluían muestras extremas para incluir tantos picos observados en este estudio como fuera posible en la tabla diana. Durante la etapa inicial del proyecto, la identidad de los picos no se conocía y se etiquetaron 'desconocido 1' a 'desconocido x', y cada pico se caracterizó por su tiempo de retención y un ión de cuantificación. Los picos encontrados en los blancos de procedimiento se eliminaron de la tabla diana. En los casos de la incertidumbre de origen de un pico, los picos se retuvieron en la lista diana para evitar la posible pérdida de analitos relevantes.

En muestras de estudio, el analito individual/compuestos, distintos de patrones internos, encontrados en los cromatogramas representativos se correspondieron con la base de datos de compuestos de referencia (tiempos de retención y espectros de masas). Después de esta etapa de correspondencia, a varios compuestos se les asignó un ID provisional basándose en la base de datos de referencia.

El alineamiento de tiempos de retención de picos no se requiere cuando se aplica un enfoque de procesamiento elegido como diana. La tabla diana puede ajustarse en los casos en los que se observaron desplazamientos del tiempo de retención (los patrones internos sirvieron de 'puntos de referencia'). Todos los compuestos presentes en la tabla diana se integraron primero usando parámetros de integración estándar. Estos parámetros no son aplicables a

- 5 todos los compuestos y muestras y pueden conducir a resultados de integración incorrectos para algunos compuestos. Se desarrolló un procedimiento para encontrar errores de integración. Para todas las dianas puede calcularse la desviación del área del pico de la media (de todo el estudio y de muestras de CC). Los resultados se representaron y se usaron para detectar picos problema (errores de integración para aquellos picos que abarcan múltiples muestras) y muestras problema (integración para múltiples picos dentro de una muestra). Para las muestras problema, la integración se realizó manualmente. Para los picos problema, la integración automatizada se realizó después de ajustar los parámetros de integración).
- 10 En diversas etapas en la preparación de muestras y análisis, uno o más patrones internos se añadieron con el fin de monitorizar la calidad de los datos después del análisis. Después de cada lote, la respuesta de los patrones internos después de la corrección para DCHP se evaluó en todas las muestras. Si el área del pico corregida de cada patrón interno se desvió menos de 20% de la media del lote, el lote puede aprobarse. Si, para una o más muestras, la desviación es superior al 20%, entonces aquellas muestras volvieron a analizarse en el siguiente lote. En último lugar puede generarse un conjunto de datos bloqueados para el análisis de EM/CG.
- 15 La primera etapa de identificación puede ser hacer corresponder la lista de compuestos diana con la base de datos de patrones de referencia. Varios compuestos de la lista diana pueden identificarse y confirmarse comparando con las bases de datos de referencia y analizando los patrones de referencia. Tras el análisis unifactorial y multifactorial, los desconocidos con prioridad se evaluaron primero comprobando los datos brutos y los resultados de CC. Se dividieron en dos categorías: i) compuestos con muy baja intensidad y/o altas DER en los patrones de CC y ii) compuestos con buena señal en una o más muestras y bajas DER en los patrones de CC. La identificación empieza con la segunda categoría. Para ambas categorías se realizó la comprobación adicional para artefactos.
- 20 Después de este análisis se seleccionaron procedimientos de identificación adicionales dependiendo de los compuestos individuales (que iban a identificarse). Para algunos compuestos, los blancos se encontraron en bibliotecas espectrales comerciales y no se requirieron experimentos de ID adicionales. Para otros compuestos se realizaron ionización química, masa exacta u otros experimentos de derivatización.
- 25 Metabolómica de EM/CL polar
- Los siguientes materiales se usaron según el procedimiento: metanol, Biosolve, calidad para EM-CL, cat. nº 13687801, o equivalente; HCl, Merck, 37%, cat. nº 1.00317, o equivalente; n-butanol, Baker cat. nº 8017, o equivalente; agua Milli-Q, agua purificada con el sistema ELGA, o equivalente; ditiotreitolo DTT, Sigma, 99%, cat. nº D9779, o equivalente; ácido fórmico, Merck, cat. nº 1.00264, o equivalente; acetonitrilo, calidad para gradiente de HPLC-S de Biosolve, cat. nº 12000701, o equivalente; fenilalanina-d5, 99% de átomos de D de CDN Isotopes, cat. nº D-1597, o equivalente; glutamato-d3, 99% de átomos de D de CDN Isotopes, cat. nº D-1196, o equivalente; leucina-d3, 99% de átomos de D de CDN Isotopes, cat. nº D-1973, o equivalente; alanina-d3, 99% de átomos de D de CDN Isotopes, cat. nº D-1462, o equivalente; creatina-d3, 99% de átomos de D de CDN Isotopes, cat. nº D-1462, o equivalente; metionina-d4, 99% de átomos de D de CDN Isotopes, cat. nº D-1462, o equivalente; metil-(d3)-histidina, 99% de átomos de D de CDN Isotopes, cat. nº D-1462, o equivalente; tirosina-d7, 99% de átomos de D de CDN Isotopes, cat. nº D-1462, o equivalente.
- 30 Se usaron los siguientes instrumentos según el procedimiento: agitación con Vortex; centrífuga; horno; equipo de evaporación; ThermoFinnigan Surveyor o Surveyor Plus HPLC; espectrómetro de masas de trampa de iones ThermoFinnigan LTQ equipado con fuente de ESI.
- 40 A una muestra de plasma de 10 µl en un vial Eppendorf pequeño se añadieron 10 µl de disolución de trabajo de IS y la muestra se agitó brevemente con vórtex. Después de la adición de disolución de DTT y desproteinización de la muestra, el sobrenadante se liofilizó. Entonces, las muestras se derivatizaron con HCl-butanol a 65°C. El exceso de reactivo se eliminó por liofilización. La muestra se reconstituyó en una disolución de agua de DTT que contenía tirosina D7 sin derivatizar como patrón interno.
- 45 Se realizó cromatografía de líquidos de alta resolución según las siguientes especificaciones:
- | | |
|---|---|
| Columna: | Varian/Chrompack Inertsil 5 µm ODS-3 100*3 mm |
| Precolumna: | Varian/Chrompack R2 10 x 2 mm de d.i. |
| Fase móvil A: | 0,1% de ácido fórmico |
| Se añade 1 ml de ácido fórmico a 1000 ml de agua, se mezcla y se desgasifica por ultrasonificación durante 5 minutos | |
| Fase móvil B: | 80% de acetonitrilo en 0,1% de ácido fórmico |
| Mezcla de 800 ml de acetonitrilo, agua hasta 1000 ml y 1 ml de ácido fórmico y desgasificación por ultrasonificación durante 5 minutos. | |
- 50

Temperatura de la columna:	25°C-30°C
Temperatura del inyector automático:	10°C
Volumen de inyección:	10 µl

5 La espectrometría de masas se realizó con un divisor con una relación de aproximadamente 4:1 que significa un flujo de ~ 75 µl a la fuente de ESI. Los parámetros de ESI y EM son:

Voltaje de pulverización ESI:	4kV
Capilar calentado:	250-300°C
Gas protector:	50-60
Gas auxiliar:	5
10 Polaridad:	positiva
Intervalo de barridos:	125-1250
Número de microbarridos:	5-6
Tiempo de inyección máximo:	200 ms
CID de la fuente:	0-5 V
15 Distancia de la fuente:	posición C

Se analizaron muestras seleccionadas para la masa exacta. Para este fin se seleccionaron dos muestras de CC de cada lote de derivatización y también varias muestras "extremas". Los experimentos de masa exacta se realizaron en un instrumento ThermoFinnigan LTQ-FTMS usando los mismos parámetros que se han descrito anteriormente. La detección de los iones se realizó en FTMS. La resolución a m/z 400 es ~100.000.

20 Todos los compuestos que eluyen de la columna de EM/CL se detectaron en modo de barrido completo. Cada pico se caracterizó por su tiempo de retención y varios iones (valores de m/z). La cantidad de datos (es decir, número de variables) se redujo aplicando un procedimiento de procesamiento diana en el que cada compuesto en el cromatograma estaba representado en la mayoría de los casos por una sola entrada en la tabla diana.

25 En enfoque diana redujo el número de variables al menos un factor de 20. La mayoría de los artefactos del procedimiento podrían eliminarse de los datos no incluyendo estos picos en la tabla diana.

30 Para este estudio se compiló una tabla diana específica del estudio. Esto se realizó captando picos de todos los QCY y muestras de estudio de los pocos primeros lotes. La tabla de picos resultante se clasificó basándose en la intensidad y se seleccionaron los 1000 picos mayores. De esta tabla parcial se eliminaron los picos de isótopos y, en la medida en la que se reconocieron, los aductos. En esta etapa, la identidad de los picos no se conocía y se etiquetaron como m/z en el tiempo de retención (en minutos).

El alineamiento de tiempos de retención de picos no se requiere cuando se aplica un enfoque de procesamiento elegido como diana. La tabla diana se ajustó en casos en los que se observaron desplazamientos del tiempo de retención (los patrones internos sirvieron de 'puntos de referencia').

35 Después de la completitud del análisis de todas las ejecuciones, todos los compuestos presentes en la tabla diana se integraron primero usando parámetros de integración estándar. Estos parámetros no son aplicables a todos los compuestos y muestras y pueden conducir a resultados de integración incorrectos para algunos compuestos. Se desarrolló un procedimiento para encontrar errores de integración. Para todas las dianas puede calcularse la desviación del área del pico de la media (de todo el estudio y de muestras de CC).

40 El control de calidad del conjunto de datos completo se realizó basándose en la respuesta de IS en todas las muestras, excepto los blancos y la comprobación de las desviaciones estándar relativas de todos los compuestos diana en las muestra de CQY y CTRL. Estos resultados se compararon con aquellos observados en estudios previos. Al completarse estas etapas puede generarse un "conjunto de datos bloqueados" para la plataforma de EM/CL polar. Este conjunto de datos forma la fundación para los posteriores análisis estadísticos.

45 La primera etapa de identificación fue hacer corresponder la lista de compuestos diana con la base de datos de patrones de referencia que contiene varios aminoácidos y compuestos relacionados. Los tiempos de retención y las masas exactas se compararon y, cuando fue necesario, se compararon espectros de EM/EM.

Los compuestos diana restantes se identificaron sólo después de enumerarse en las listas de prioridad (unifactorial/multifactorial). Los desconocidos con prioridad se evaluaron primero comprobando los datos brutos y los

resultados de CC. Se dividieron en dos categorías: i) compuestos con muy baja intensidad y/o altas DER en los patrones de CC y ii) compuestos con buena señal en una o más muestras y bajas DER en los patrones de CC. La identificación empezó con la segunda categoría. Para ambas categorías se realizó la comprobación adicional para artefactos; el análisis de los blancos se comprobó para la presencia de estos compuestos.

- 5 Para más identificación se usaron datos de tiempo de retención, masa exacta y datos de EM/EM. Basándose en la masa exacta y el conocimiento sobre la derivatización usada se buscaron posibles composiciones elementales en las bases de datos KEGG, Merck y Chemfinder. Los posibles blancos se evaluaron y para este fin se usaron los tiempos de retención y los espectros de EM/EM. En varios casos se compraron patrones y se analizaron en disolución y después de la adición convencional a las muestras de estudio.

10 Identificación por EM/CL de ácidos grasos libres

El plasma se desproteinizó y se extrajo con IPA. Los ácidos grasos libres se separaron sobre una columna analítica, se ionizaron por electropulverización en modo de ión negativo y se detectaron en modo de barrido completo. Se analizaron varios ácidos grasos libres como patrones externos y, si necesitaron determinarse las concentraciones en plasma reales, éstas se calcularon (aproximadas) usando los patrones de calibración.

- 15 Las muestras se prepararon y se analizaron usando los siguientes materiales. Los instrumentos incluyeron una centrífuga para tubos Eppendorf, un Thermo Electron LTQ, HPLC Thermo Electron Surveyor con inyector automático y horno de columna y Vortex. Los productos químicos incluyeron agua desmineralizada, (sistema 4 de ELGA o cualquier otro fabricante, por ejemplo, Millipore); isopropanol para análisis (Baker 6068 o igual); metanol, (calidad para HPLC, Cromasolve 34966); diclorometano para análisis, (Baker 7053 o igual); acetato de amonio para análisis (Fluka 09690 o igual); y ácido fórmico para análisis, (Merck 1.00264.1000 o igual).

El material incluyó columna de HPLC Alltech Prosphere C4 300Å (150 x 3,0 mm de d.i., 5 µm), nº de pieza 35548; precolumna Symmetry 300 C4 (10 x 2,1 mm de d.i., 3,5 µm), nº de pieza 186000278; viales de inyector automático (nº de pieza de Alltech AV055201, 32*11,6 mm); tubos Eppendorf; y pipetas (por ejemplo Eppendorf).

- 25 Se usaron los siguientes patrones: ácido heptadecanoico C17:0 FA (Sigma-Aldrich, H3500); C14:0 FA ácido mirístico (Sigma M3128); C16:0 FA ácido palmítico (Sigma P5585); C16:1 FA ácido palmitoleico (Sigma P9417); C18:0 FA ácido esteárico (Sigma S4751); C18:1 FA ácido oleico (Sigma O1008); C18:2 FA ácido linoleico (Sigma L1376); y C20:4 FA ácido araquidónico (Sigma A9673).

- 30 Las disoluciones de patrón interno se prepararon según los siguientes procedimientos. Las disoluciones madre de ácido heptadecanoico C17:0 FA (1 mg/ml) se prepararon pesando 5 mg en un matraz volumétrico de 5 ml, añadiendo 500 µl de DCM, disolviendo por ultrasonificación hasta 5 ml con IPA. La disolución de trabajo de IS (para la precipitación de proteína y extracción de lípidos) se preparó en IPA que contenía C17:0 FA en una concentración de 1 µg/ml.

- 35 Las muestras se prepararon según el siguiente procedimiento. 300 µl de disolución de trabajo de IS se añadieron a 10 µl de muestra de plasma descongelada en una copa Eppendorf y se agitaron con vórtex. La disolución se sometió a centrifuga en Eppendorf durante 3-5 minutos a 10.000 rpm y se transfirió 2 veces de 50 µl del sobrenadante claro a dos viales del inyector automático con un inserto y el extracto restante se almacenó a <-18°C.

- 40 Se realizó cromatografía de líquidos y espectrometría de masas usando una columna de HPLC Alltech Prosphere C4 300Å (150 x 3,2 mm de d.i., 5 µm) y una precolumna Phenomenex Widespore C4 (4 x 3 mm de d.i., 5 µm). La fase móvil A consistió en 5% de metanol en tampón; la fase móvil B consistió en metanol que contenía NH₄Ac 2 mM; y el lavado necesario consistió en 60% de IPA en MeOH. El gradiente fue del siguiente modo:

Tiempo (min)	Flujo (ml/min)	% de A	% de B
0	0,4	70	30
2	0,4	70	30
6	0,4	30	70
10	0,4	0	100
15	0,4	0	100
15,1	0,4	70	30
20	0,4	70	30

La temperatura de la columna fue 40°C; la temperatura del inyector automático fue 20°C; y el volumen de inyección

fue 20 µl.

Para el procedimiento de espectrometría de masas, el instrumento de EM/CL se afinó usando los siguientes patrones de calibración:

Instrumento	ThermoFinnigan LTQ
Modo	negativo
Capilar calentado	250°C
Voltaje de pulverización	3,5 kV
Gas protector	35 arb
Gas auxiliar	15 arb
Posición de pulverización	C
Intervalo de barridos	180-400
Velocidad de barrido	Automática
Multiplicador	Automático

5 Se usó Thermo XCalibur LCQuan V 2.0 para la integración de analitos diana seleccionados. Los datos se procesaron usando una tabla diana que contenía al menos los siguiente ácidos grasos (basándose en estudios previos): C12:0 FA, C12:1 FA, C14:0 FA, C14:1 FA, C16:0 FA, C16:1 FA, C16:2 FA, C18:0 FA, C18:1 FA, C18:2 FA, C18:3 FA, C20:0 FA, C20:1 FA, C20:2 FA, C20:3 FA, C20:4 FA, C20:5 FA, C22:5 FA, C22:6 FA y C22:7 FA.

Se obtuvieron datos de integración para todos los compuestos de la tabla diana. El resultado fue una tabla de Excel que contenía un valor de integración para cada compuesto diana.

10 Proteómica de descubrimiento por EM/EM-MALDI-CL

El flujo de trabajo de proteómica se basó en una cromatografía de líquidos multidimensional - análisis de EM/EM de péptidos generados a partir de las muestras de plasma. Las ventajas de este enfoque incluyen la insensibilidad a múltiples heterogeneidades de proteínas del plasma que producen dificultades en un enfoque basado en gel 2D, la fácil automatización de técnicas de cromatografía de líquidos y la posibilidad de utilizar marcado con isótopos estables de péptidos.

15 La metodología iTRAQ™ recientemente introducida (iTRAQ es una marca registrada de Applied Biosystems) fue particularmente adecuada para este estudio, ya que la reacción de marcado fue fácil de llevar a cabo hasta completitud, genérica, casi todos los péptidos se marcan con iTRAQ y permitió la cuantificación multiplexada de hasta cuatro muestras por experimento. Esto último fue importante en este estudio en el que la coherente cuantificación (e identificación) de las mismas proteínas de muestra a muestra es un requisito clave. Otra ventaja significativa del marcado con isótopos estables fue la robustez del procedimiento con respecto a variaciones del comportamiento de cromatografía de líquidos de péptidos de muestra a muestra (se combinaron conjuntos marcados diferencialmente una vez).

25 En el procedimiento iTRAQ, los digestos enzimáticos de muestras de proteína se tratan con los reactivos (un éster de N-hidroxisuccinimida (NHS)) que derivatizan los grupos amino primarios libres de péptidos: el extremo N (si no está bloqueado) y los residuos de lisina. Cuatro variedades diferentes del reactivo se caracterizan por la misma masa, pero diferente posicionamiento de las marcas de isótopo estable. Por tanto, en modo de EM, en el que la señal del ión refleja el peso molecular de los péptidos, los cuatro péptidos marcados de forma diferente aparecen como un único componente. En modo de EM/EM, las leves diferencias en la estructura de la marca se vuelven
30 visibles y, mientras que los fragmentos del esqueleto peptídico (los llamados a, b e y) todavía son isóbaros, cuatro fragmentos indicadores diferentes se generan correspondientes a m/z 114, 115, 116 y 117, respectivamente. La cuantificación relativa de los péptidos de cuatro muestras diferentes puede llevarse a cabo por la determinación de intensidades relativas de estos iones indicadores.

35 En este estudio se usó un espectrómetro de masas MALDI-TOF/TOF (ABI 4800) en el que las muestras se depositaron sobre placas de MALDI por un procedimiento de recogida de fracciones automatizado. La ventaja particular de la EM/EM-MALDI-CL está en la naturaleza fuera de línea del acoplamiento EM-CL. Fue posible analizar una separación por CL 2D entera en modo de EM primero, antes de generar identificaciones de péptidos y datos cuantitativos en modo de EM/EM. Esta característica puede ser muy conveniente cuando tiene que medirse un conjunto constante de péptidos en un gran número de muestras.

- Un único aspecto de analizar muestras de plasma es la enorme información representada por las abundantes proteínas en plasma (albúmina, inmunoglobulinas, etc.). Para tratar este reto se usaron varias técnicas de reducción. En este estudio se usó una columna de anticuerpo IgY de pollo para reducir las muestras de doce proteínas abundantes. El conjunto de proteínas no retenidas sobre la columna de anticuerpo se recuperó en una columna de fase inversa. El conjunto de proteínas recuperado se redujo, se alquiló y se digirió con tripsina. La mezcla de péptidos resultante se marcó con el reactivo de iTRAQ y se combinó con las otras tres muestras designadas para la misma mezcla de iTRAQ. Las mezclas cuádruplex combinadas se fraccionaron primero en una cromatografía de intercambio catiónico fuerte. Se recogieron diez fracciones de la elución de SCX. Estas fracciones se analizaron adicionalmente - después de reunir algunas de las fracciones - por HPLC EM/EM.
- El análisis de iTRAQ permite la cuantificación relativa de cuatro muestras diferentes en un único experimento de EM/EM-CL 2D (cuádruplex). Sin embargo, el asignar cuatro muestras primarias a una mezcla cuádruplex de iTRAQ no permite comparar dos muestras primarias designadas a diferentes mezclas de iTRAQ. Esto se trató creando un conjunto de referencia que se tomó en alícuotas según el número de mezclas de iTRAQ analizadas para el proyecto; un componente constante de cada mezcla de iTRAQ; la preparación de muestras para esta alícuota de referencia se llevó a cabo junto con los otros tres miembros de la misma mezcla de iTRAQ; y se marcó con el reactivo 117.
- El conjunto de referencia se creó a partir de una combinación de aproximadamente 25% de cada muestra del estudio primario analizada en el estudio. Las alícuotas del conjunto de referencia se denominaron en lo sucesivo muestras de QCR.
- La adquisición de EM/EM puede guiarse por listas de precursores de EM para tanto inclusión como exclusión. Una lista de inclusión original se derivó del análisis realizado durante la fase de catálogo del estudio. En esta fase, algunas muestras del conjunto de CC de las muestras se procesaron mediante el flujo de trabajo de proteómica completo. Las listas de inclusión y exclusión se compilaron después de identificar péptidos y hacer corresponder proteínas de estas mezclas.
- El contenido de la lista de inclusión incluye: péptidos trípticos marcados completamente alquilados y de iTRAQ; péptidos semitrípticos marcados completamente alquilados y de iTRAQ, con la exclusión de aquellos debidos a actividad quimiotríptica (escisión en F, Y y W); y péptidos manualmente identificados como que tienen modificaciones postraduccionales no rutinariamente encontradas por Mascot.
- El contenido de la lista de exclusión incluye: los 2.000 componentes sin identificar más abundantes, que incluyen péptidos con marcado por iTRAQ incompleto o alquilación; péptidos con más de 3.000 Da de peso molecular que contienen múltiples sitios de escisión trípticos ausentes; péptidos generados por actividad lateral quimiotríptica de tripsina; posible forma $M+H^+$ de péptidos abundantes que estaban en la lista de inclusión en la forma $M+H^+$; y péptidos autolíticos de tripsina.
- Se usó una aplicación de software a medida para considerar los criterios de inclusión/exclusión y generar adquisiciones de EM/EM en AB4800 basándose en archivos de texto separados que desglosan precursores seleccionados para cualquier inclusión o exclusión. Los ajustes automáticos de posibles desplazamientos de los tiempos de retención de HPLC también se acomodaron por el software. Se adquirieron espectros de EM/EM adicionales de un modo oportunista como tiempo de adquisición permitido después de analizar primero los precursores incluidos.
- Las muestras se descongelaron sobre hielo en lotes diarios de 12 que consistían en 9 muestras primarias y 3 muestras del conjunto de referencia y se procesaron seriadamente con una muestra del conjunto de referencia siguiendo tres muestras primarias (es decir, las muestras 1, 2, 3, luego R, 4, 5, 6, R etc.). Una vez las muestras se habían descongelado completamente, se deslipidaron. Las muestras se diluyeron con PBS y se extrajeron con triclorotrifluoroetano (freón 13), seguido de una etapa de centrifugación. La fase acuosa superior se transfirió a un nuevo tubo y se guardó brevemente sobre hielo antes de colocarse en el inyector automático enfriado del sistema de CL para la eliminación de abundante proteína.
- Usando eliminación de abundante proteína (EAP), las muestras deslipidadas se agotaron de doce de las proteínas del plasma más abundantes. Este procedimiento usó un procedimiento de dos columnas (CL 2D) en el que doce proteínas que incluían albúmina, IgG, IgA, transferrina, α -1 antitripsina, la familia de las haptoglobinas, fibrinógeno, α -1 glicoproteína ácida, IgM, α -2 macroglobulina y HDL (apolipoproteínas A-I y AII) se eligieron como diana por una columna de IgY de afinidad mientras que el material agotado fluía su través y se capturaban en una columna de fase inversa (RP). La columna de RP se lavó independientemente y la etapa se eluyó. El pico de elución, que consiste en el material agotado lavado, se recogió en un colector de fracciones y posteriormente se secó en la centrifuga SpeedVac para el procesamiento adicional en el flujo de trabajo. Entonces, la columna de RP se lavó con una disolución de lavado que elimina el remanente a la siguiente ejecución. Entonces, la columna de IgY se eluyó independientemente y se neutralizó antes de la siguiente ejecución. Se procesaron doce muestras deslipidadas (un único lote) en un periodo de 24 horas seguido de una centrifugación en SpeedVac durante la noche de las fracciones reunidas.
- Las muestras de plasma agotadas se redujeron, se alquilaron, se digirieron y se marcaron con iTRAQ. Las muestras

- agotadas secadas se resuspendieron en tampón TEAB 1 M a pH 8,5. Se añadieron TCEP y SDS para desnaturalizar y reducir las muestras. Después de una incubación de una hora a 71°C, las muestras se alquilaron con yodoacetamida durante 30 minutos a temperatura ambiente en la oscuridad. Se añadió tripsina suspendida en N-acetil-1-cisteína y las muestras se dejaron digerir durante la noche a 48°C. Al día siguiente, los reactivos de iTRAQ se descongelaron y se resuspendieron en 100% de etanol. Las muestras se marcaron con reactivos de iTRAQ específicos. Después de una hora de incubación a temperatura ambiente para el marcado, la reacción de acetilación se inactivó con la adición de 4 microlitros de bicarbonato de amonio 1 M. Las muestras marcadas con iTRAQ para la misma mezcla se combinaron y se secaron en un Speedvac antes de la resuspensión para cromatografía de intercambio catiónico.
- El fraccionamiento por intercambio catiónico fuerte de las mezclas de iTRAQ combinadas se llevó a cabo en una estación de cromatografía Vision (Applied Biosystems) usando una columna de intercambio catiónico fuerte de polisulfato de 4,6 x 50 mm. La separación usó un procedimiento de dos columnas en la que los conjuntos de péptidos se cargan primero a una columna de fase inversa Poros R2 con el fin de eliminar la gran cantidad de productos de hidrólisis de iTRAQ.
- Los péptidos se eluyeron sobre la columna SCX usando un tampón de carga de SCX de 35% de contenido de ACN. Se eliminó el material más hidrófobo apilado sobre la columna de RP (es decir, tripsina residual) por lavado con alto contenido orgánico. Los péptidos unidos sobre la columna SCX se eluyeron con un gradiente. Antes de la inyección, el pH de la muestra se ajusta a 3,5 usando HCl 1 N. De la separación de SCX se recogieron diez fracciones de 0,8 ml (A1 a A10). Estas fracciones se secaron en Speedvac y se resuspendieron en tampón de carga de HPLC.
- Proteómica de descubrimiento por EM/EM-MALDI-CL
- El acoplamiento de HPLC y MALDI se llevó a cabo fuera de línea mediante la recogida de fracciones sobre la diana de MALDI. El equipo fue un sistema de cromatografía UltiMate equipado con el dispositivo de aplicación de puntos de MALDI Probot de Dionex-LC Packings (Hercules, CA). Una bomba de jeringuilla integrada sobre el colector de fracciones permitió la mezcla conjunta del eluyente de HPLC y la matriz de MALDI antes de la deposición sobre la placa de MALDI. Para garantizar la buena calibración de la masa de los espectros de EM-MALDI se añadió un patrón interno - péptido ACTH (18-39); $MH^+ = 2465,199$ a la disolución de la matriz.
- La aplicación de puntos a las placas de MALDI se configuró de forma que las tres ejecuciones de HPLC se depositaran sobre una única placa de MALDI 4800. Cada ejecución de HPLC contuvo una matriz de 6 x 35 de fracciones (210 fracciones/HPLC). Las seis ejecuciones de HPLC que representan una mezcla de iTRAQ se aplicaron por puntos sobre dos placas de MALDI.
- Los parámetros de operación de HPLC-MALDI se enumeran más adelante. El agua de calidad para HPLC y el acetonitrilo de calidad analítica fueron de Burdick & Jackson, acetato de amonio, citrato de amonio, TFA y ácido α -ciano-4-hidroxicinámico fueron de Sigma.

Parámetros de HPLC-MALDI:

Tampón A de HPLC:	95% de agua, 5% de acetonitrilo, 0,1% de TFA
Tampón B de HPLC:	10% de agua, 90% de acetonitrilo, 0,1% de TFA
Tampón de carga de HPLC:	95% de agua, 5% de acetonitrilo, 0,1% de TFA, acetato de amonio 2 mM
Disolución de matriz:	10 g/l de ácido α -ciano-4-hidroxicinámico – 0,1 g/l de citrato de amonio dibásico en 15% de agua, 85% de acetonitrilo
Patrón de calibración de la masa interna:	Péptido ACTH (18 – 39) a 200 fmol/l (en la disolución de matriz)
Columna de resolución de HPLC:	0,180 x 150 mm C18 (Dionex-LC Packings)
Columna de trampa de HPLC:	0,3 x 5 mm C18 (Dionex-LC Packings)
Flujo de HPLC:	2 microlitros/min
Flujo de la matriz:	2 microlitros/min
Intervalo de recogida de fracciones:	9 s

El gradiente de HPLC se programó como:

	0 – 5% de B	en 8 min
	5 – 10% de B	en 2 min
	10 – 37,5% de B	en 30 min
5	37,5 – 100% de B	en 2 min.

Las trazas de UV se registraron usando una celda de flujo de 45 nl monitorizando la longitud de onda de 280 nm.

Las cantidades inyectadas de cada fracción de SCX se determinaron para maximizar la carga de muestra, pero evitando sobrecargar la columna de HPLC. Estas cantidades se determinaron en experimentos de catálogo preliminares.

- 10 Antes de cada conjunto de seis inyecciones que representan las muestras de MALDI de una mezcla de iTRAQ, una mezcla de péptidos patrón (de Michrom) se inyectó para monitorizar la consistencia de los tiempos de retención de HPLC. Obsérvese que cambios moderados en los tiempos de elución de HPLC fueron manipulados a la perfección por el programa de selección del precursor alineando automáticamente las ejecuciones de EM-CL de fracciones de SCX adyacentes. Tras la mezcla de CC patrón, las seis fracciones de SCX se inyectaron y se generaron dos placas de MALDI.

Los conjuntos de dos placas de MALDI se montaron en el cargador automático del instrumento AB4800. Cada placa se ejecutó en modo de EM por la siguiente secuencia:

- Generar el modelo de placa y actualizar la calibración por defecto usando la mezcla patrón de 4800 (Applied Biosystems) aplicada en puntos sobre las ocho posiciones “Cal” sobre la placa
- Ejecutar las fracciones de CL en modo de EM con calibración de la masa interna usando los picos de m/z 2465,199 (patrón de péptido) y m/z 568,136 (trímero de la matriz)

- 25 Los principales objetivos para la selección de precursores tuvieron como objetivo minimizar los tiempos de adquisición (eliminar precursores redundantes), eliminar la interferencia de picos adyacentes en el espectro de EM y seleccionar coherentemente los mismos precursores para todas las mezclas de iTRAQ. La adquisición de EM/EM y los parámetros de procesamiento se optimizaron para producir la mejor medición posible de los picos conductores de iTRAQ a m/z 114, 115, 116, 117. Obsérvese que las adquisiciones de EM/EM se fijaron para acumular 2.100 disparos de láser a menos que la señal de iTRAQ superara los criterios de señal con respecto a ruido después de sólo 1.050 disparos. De esta forma, el tiempo de adquisición podría reducirse sin arriesgar la calidad de las mediciones del área del pico de iTRAQ.

- 30 Los espectros de EM/EM de AB4800 se recogieron de la base de datos Oracle residente en el ordenador de control de la adquisición usando un archivo de registro en Perl a medida accionado por el software de control Pipeline. Los picos de EM/EM se procesaron en dos archivos distintos para registrar por separado los datos de picos de iTRAQ cuantitativos y los picos de fragmentación de péptidos. Cada espectro de EM/EM individual se identificó por una única ID de la tabla de picos de la base de datos y la única ID de la base de datos de la placa de MALDI de la que se adquirieron los datos. Estos dos valores permitieron la posterior asociación con información adicional registrada en tanto los datos de adquisición como el sistema de control de información del laboratorio interno (LIMS). La información cuantitativa de iTRAQ se transfirió a una base de datos llamada PeptideDB del primer archivo. La salida secundaria fue un archivo MGF (formato genérico de Mascot) que contenía los picos de fragmentación de péptidos y líneas de comentarios asociadas adicionales para permitir el seguimiento entre resultados de Mascot y espectros de EM/EM de entrada.

Los picos incluidos en los archivos MGF se filtraron para eliminar picos de alta masa (dentro de 64 Da o el precursor MH^+); eliminar picos de baja masa (< 146 Da); y limitar la “densidad” de picos a 15 masas de fragmento por 200 Da de ventana.

- 45 Se creó un archivo MGF separado por la ejecución de trabajo de EM/EM en AB4800 (un trabajo de EM/EM puede consistir en entre 1 y miles de adquisiciones de EM/EM separadas adquiridas a partir de la misma placa de MALDI). Los archivos MGF individuales correspondientes a la misma mezcla de iTRAQ se concatenaron antes de iniciar una búsqueda con Mascot.

- 50 Cada espectro de EM/EM se asoció a la intensidad del ión de la señal de m/z 114, 115, 116 y 117 en PeptideDB. Como se ha descrito anteriormente, las relaciones de iTRAQ se determinaron con respecto a la intensidad de m/z 117 (ya que esta marca se usó para el conjunto de referencia). Las relaciones se calcularon como: relación de áreas de pico para datos de 4800 (no áreas de grupos como se usa por el software de Applied Biosystems) y relación de recuentos de picos para datos de QToF.

Tras la búsqueda en bases de datos, las secuencias de péptidos correspondientes a los espectros de EM/EM

- “heredaron” los números de cuantificación asignados a los espectros de EM/EM. Las relaciones de proteínas se derivaron de las relaciones de péptidos asignadas a esa proteína como una relación media (o relación promedio en el software de Applied Biosystems). Sin embargo, las asignaciones de péptido-proteína finales en el flujo de trabajo de BGM no se hicieron hasta que se completó la identificación de péptidos a partir de cada mezcla de iTRAQ. En este momento, no sólo se consolidaron las asignaciones de proteínas, sino que las proteínas individuales pudieron descomponerse en múltiples nodos basándose en la información cuantitativa para todo el estudio (es decir, se resolverían isoformas de proteínas, secuencias de péptidos que acomodan sitios polimórficos, variantes alélicas, etc.). Entonces, los nodos de proteínas - que se usan en último lugar para el análisis de redes estadísticas y de correlación - se representaron como la media de las relaciones de péptidos que se corresponden con ellos.
- Si se propagan relaciones de péptidos en relaciones de proteínas o de nodos de proteínas, sólo se consideraron aquellos péptidos que (a) se modificaron completamente al grado esperado; es decir, se marcaron completamente con iTRAQ y se alquilaron completamente si contenían cisteína(s) o (b) no se correspondieron con proteínas elegidas como diana por APR, o con tripsina, con hemoglobina (estas proteínas se consideran contaminantes)
- La identificación de péptidos/proteínas se llevó a cabo usando un enfoque de tres dientes del programa de búsqueda de la base de datos comercial Mascot (MatrixScience Ltd., RU), validación experta por científicos y un procedimiento de correspondencia con espectros de EM/EM.
- Una vez se completó la adquisición de datos para todo el componente de proteómica del estudio y los resultados de búsqueda de Mascot se cargaron en PeptideDB, se ejecutó un procedimiento de correspondencia espectral. El objetivo fue encontrar aquellos espectros de EM/EM que se correspondían con una secuencia de péptidos con los criterios de autovalidación en cualquiera de las mezclas de iTRAQ analizadas en el proyecto de proteómica. La identidad de péptidos se estableció basándose en masa del precursor estrechamente correspondiente; número de fracciones de SCX y tiempo de retención de HPLC; y masas de iones de fragmentos más abundantes. De esta forma fue posible rescatar varios espectros de EM/EM que no cumplieron los criterios de autovalidación por Mascot, pero que tuvieron un caso autovalidado o validado por experto en otra(s) mezcla(s) de iTRAQ. La importancia de tal procedimiento es aumentar la representación de péptidos y proteínas medidos a lo largo de tantas muestras como sea posible en el estudio de proteómica.
- Después de completarse la identificación de péptidos para el estudio completo (que tiene el conjunto de péptidos más completo posible) se ejecutó una herramienta de validación de proteínas para identificar el conjunto mínimo de proteínas que explicaban todas las secuencias de péptidos validadas. En la terminología de la herramienta para la validación de proteínas, las secuencias de péptidos se asignaron a (a) clase de proteína que representa un único gen en el genoma o (b) ejemplar de proteína que representa el caso de secuencias preferidas (normalmente la que tiene la mayor información de anotación) en una clase de proteínas (entre variantes de corte y empalme, formas polimórficas, etc.).
- Se aplicaron varias etapas de procesamiento de datos al conjunto de datos complejo y se influyó el número de proteínas/péptidos y métricas de cuantificación informadas para el estudio. Estas etapas de procesamiento de datos incluyen:
- Consolidación de mediciones de péptidos: Se sumaron las señales de iTRAQ de múltiples mediciones de EM/EM del mismo péptido (misma secuencia y mismas modificaciones) en la misma mezcla de iTRAQ y a partir de ese momento se usó una única medición de relaciones.
- Eliminación de valores atípicos: Un procedimiento estadístico convencional para eliminar 5% de las mediciones de péptidos de valor atípico. Una medición de péptido se consideró un valor atípico si se encontró fuera (más de dos desviaciones estándar) de la distribución de todas las mediciones de péptidos que se correspondían con la misma proteína.
- Asignación de proteínas/nodos finales: Una vez se había completado la herramienta de validación de proteínas, los resultados cuantitativos para todo el proyecto se usaron para finalizar la lista de proteínas para análisis estadístico unifactorial y multifactorial. Las proteínas finales definieron los nodos para el análisis de redes de correlación. Este procedimiento fue necesario para el correcto seguimiento de formas procesadas de proteínas (es decir, C3a y C3b del complemento C3, fibrinopéptido A de la cadena alfa de fibrinógeno, alfa-microglobulina y bikunina de AMBP_HUMAN, etc.) y variantes polimórficas de proteínas.
- El rendimiento cuantitativo se evaluó por la consistencia de las relaciones de péptidos que se corresponden con la misma proteína. Las desviaciones estándares relativas de estas mediciones de péptidos indicaron cómo de exacta era la medición de proteínas (como un promedio de las mediciones de péptidos correspondientes). Estos procedimientos redujeron las desviaciones estándares relativas de péptidos algunos puntos de porcentaje.
- Proteómica elegida como diana: Inmunoensayos multiplexados
- Las concentraciones de las 89 proteínas elegidas como diana mostradas en la siguiente Tabla 3 se midieron en un sistema de inmunoensayo multiplexado (Rules Based Medicine, Austin, TX, EE.UU.) [McDade & Fulton, Device Diagn Ind 19:75-82, 1997; Fulton y col., Clin Chem 43:1749-56, 1997]. El sistema realiza un número muy grande de

- 5 reacciones simultáneas usando un citómetro de flujo y procesador de señales digitales para realizar análisis en tiempo real de múltiples ensayos basados en microesferas. Los tres componentes principales del sistema son un citómetro de flujo de mesa, microesferas y hardware y software informático. El volumen de plasma usado en estas mediciones fue 50 microlitros. La señal de fluorescencia registrada por el citómetro de flujo para cada una de las proteínas elegidas como diana puede compararse con curvas estándar para el cálculo de la concentración de cada una de las proteínas elegidas como diana. En el análisis se incluyeron muestras de control de calidad.

Tabla 3: Las 89 proteínas medidas en inmunoensayo multiplexado

1. Adiponectina	32. GM-CSF	65. MMP-3
2. Alfa-1 Antitripsina	33. Hormona de crecimiento	66. MMP-9
3. Alfa-Fetoproteína	34. Haptoglobina	67. MCP-1
4. Alfa-2 Macroglobulina	35. Inmunoglobulina A	68. Mieloperoxidasa
5. Apolipoproteína A-1	36. Inmunoglobulina E	69. Mioglobina
6. Apolipoproteína C-III	37. Inmunoglobulina M	70. PAI-1
7. Apolipoproteína H	38. Insulina	71. PAPP-A
8. Beta-2 Microglobulina	39. IGF-1	72. PSA, libre
9. BDNF	40. ICAM-1	73. Fosfatasa ácida prostática
10. Proteína reactiva C	41. Interferón-gamma	74. RANTES
11. Calcitonina	42. Interleucina-1 alfa	75. Amiloide P del suero
12. Antígeno del cáncer 19-9	43. Interleucina-1 beta	76. SGOT
13. Antígeno del cáncer 125	44. Interleucina-1 ra	77. Globulina de unión a hormona sexual
14. Antígeno carcinoembrionario	45. Interleucina-2	78. Factor de citoblastos
15. CD40	46. Interleucina-3	79. Trombopoyetina
16. Ligando de CD40	47. Interleucina-4	80. Globulina de unión a tiroides
17. Complemento 3	48. Interleucina-5	81. Hormona estimulante tiroidea
18. CK-MB	49. Interleucina-6	82. Factor tisular
19. Endotelina-1	50. Interleucina-7	83. TIMP-1
20. Eotaxina	51. Interleucina-8	84. Factor de necrosis tumoral-alfa
21. Factor de crecimiento epidérmico	52. Interleucina-10	85. Factor de necrosis tumoral-beta
22. ENA-78	53. Interleucina-12 p40	86. Factor de necrosis tumoral RII
23. Eritropoyetina	54. Interleucina-12 p70	87. VCAM-1
24. ENRAGE	55. Interleucina-13	88. VEGF
25. Factor VII	56. Interleucina-15	89. Factor de von Willebrand
26. Proteína de unión de ácidos grasos	57. Interleucina-16	
27. Ferritina	58. Leptina	
28. Fibrinógeno	59. Lipoproteína (a)	
29. FGF básico	60. Linfotactina	
30. GST	61. MDC	
31. G-CSF	62. MIP-1 alfa	
	63. MIP-1 beta	
	64. MMP-2	

- 10 En análisis de proteínas elegidas como diana dio información útil sobre las concentraciones de 66 proteínas a lo largo del conjunto de muestras de casos y controles. El conjunto de datos para estas 66 proteínas se usó en análisis unifactoriales y multifactoriales para esta plataforma bioanalítica individual y se combinó con los otros conjuntos de datos de plataformas bioanalíticas para generar el conjunto de datos integrado que comprende 723 analitos.

Procedimientos estadísticos

- 15 El análisis estadístico se llevó a cabo para tratar el objetivo primario del estudio, concretamente para descubrir un analito molecular (conjunto), en sangre o plasma en sangre, y algoritmo asociado que predice una MACCE a medio plazo. Los sujetos que tuvieron la aparición de MACCE durante el seguimiento de dos años se denotan como casos de enfermedad y aquellos sin aparición de MACCE durante el seguimiento como controles.

El análisis estadístico incluyó los siguientes componentes:

1. Características de referencia de sujetos en estudio.
2. Análisis de componentes principales (PCA) para visualizar la separación (si la hay) entre los casos de enfermedad y los controles de los datos identificados por cada plataforma.
3. Modelos unifactoriales para detectar analitos individuales que están estadísticamente significativamente asociados a la probabilidad de una MACCE a medio plazo, después de ajustar factores de riesgo convencionales.
4. Modelos multifactoriales: Clasificadores para obtener subconjuntos óptimos de analitos que están estadísticamente significativamente asociados a la probabilidad de una MACCE a medio plazo, con y sin factores de riesgo convencionales. Se compararon curvas de características de operación del receptor (ROC) basándose en resultados de clasificadores con y sin factores clínicos convencionales.

Características de referencia: Las medias y las proporciones para características de referencia se calcularon para los casos de enfermedad y controles. La significancia estadística de las diferencias en características de referencia entre casos de enfermedad y controles se basó en la prueba de Wilcoxon para datos independientes (variables continuas) y las pruebas de la chi al cuadrado (para proporciones y tablas de Rx C).

Análisis de componentes principales: Las representaciones de las puntuaciones de PCA se generaron para datos obtenidos de cada plataforma metabolómica y proteómica en el estudio. Las motivaciones para estos análisis incluyen visualización de la separación entre casos de enfermedad y controles, identificación de muestras de valores atípicos y para la evaluación de la calidad de los conjuntos de datos.

Procedimientos unifactoriales: El análisis unifactorial primario se basó en regresión logística condicional, teniendo en cuenta la correspondencia 1 a 1 de casos de enfermedad con respecto a controles. El análisis unifactorial consistió en modelos de regresión que evaluaban la significancia estadística de la asociación de cada analito en plasma individualmente con el resultado (caso de enfermedad frente a estado de control) después de ajustar covariables clínicas relevantes. Como los casos de enfermedad se correspondieron con los controles en edad, raza, sexo e índice de EAC, estas variables no aparecen en el modelo de regresión. El modelo de regresión logística condicional se ajustó para factores de confusión que no están incluidos en la estrategia de correspondencia. La selección de factores de riesgo convencionales para la inclusión en el modelo se basó en una revisión de artículos recientes en la bibliografía médica que informa sobre los resultados de estudios que investigan los efectos de analitos con respecto a resultados cardiovasculares. Cada analito en plasma se modeló como cuantiles basándose en la distribución del analito en los sujetos de control. Las pruebas de relaciones de probabilidad se usaron para evaluar la significancia estadística de cada analito en la predicción del resultado. El ajuste para múltiples comparaciones se basó en procedimientos de Storey (J. R. Statist. Soc. B 64: 479-498, 2002). Las características con más de 50% de valores ausentes en tanto los casos de enfermedad como los controles se excluyeron de este análisis unifactorial.

Se realizó un segundo conjunto de análisis unifactoriales en el que los valores de analitos se transformaron en variables binarias en las que 0 indica valores ausentes y 1 de otro modo. Como en los análisis unifactoriales previos, las pruebas de las relaciones de probabilidad se usaron para evaluar la significancia estadística de cada analito en predecir el resultado. El ajuste para múltiples comparaciones se basó en procedimientos de Storey (J. R. Statist. Soc. B 64: 479-498, 2002).

También se generaron representaciones gráficas de cajas y bigotes para cada analito estadísticamente significativo, representando casos de enfermedad y controles por separado.

Se generaron representaciones de valor de p ajustado a FDR en comparación con valor de p bruto. La distribución de los valores de p observada en la representación se usó como base para seleccionar un umbral de valores de p para la selección de analitos putativos en cada plataforma. Cuando los datos de todas las plataformas estuvieron disponibles, los valores de p ajustados a FDR se generaron integrando los datos de todas las plataformas. Se generó una lista de analitos unifactorial final a partir del análisis integrado. Los resultados seleccionados se enumeran en la siguiente Tabla 4.

Tabla 4: Resultados del análisis unifactorial (Parte 1).

Plataforma	ID	Unidades	Valor de p	Razón de posibilidades (OR ₄)	Ratio de concentración promedio en casos con respecto a concentración promedio en controles
EM/CL polar	Cisteína	unidades arbitrarias (UA)	<0,01	15,70	1,17
Proteómica elegida como diana	Factor de von Willebrand	Microgramos/ml	<0,01	17,67	1,44
Proteómica elegida como diana	IL-8	pg/ml	<0,01	15,57	1,25
EM/CL de lípidos	16:0/18:1 PC	UA	<0,01	17,67	1,12
EM/CG	N-carboxi-alanina	UA	<0,01	14,31	1,16
Proteómica elegida como diana	Fibrinógeno	mg/ml	<0,01	7,55	1,28
Proteómica elegida como diana	MMP-2	ng/ml	<0,01	17,85	1,47
EM/CL de lípidos	18:0/20:4 PE	UA	<0,01	27,53	1,15
Proteómica elegida como diana	Apolipoproteína A1	mg/ml	<0,01	0,09	0,86
EM/CL de lípidos	16:0/22:6 PE	UA	0,01	14,87	1,08
EM/CL de lípidos	18:1/18:0/18:0 TG	UA	0,01	11,43	1,20
Proteómica elegida como diana	Alfa-1 Antitripsina	mg/ml	0,01	5,59	1,12
EM/CL de lípidos	18:2/18:1/17:0 TG	UA	0,01	69,36	1,11
EM/CL de lípidos	20:1/18:1/18:1 TG	UA	0,02	12,64	1,27
EM/CL de lípidos	16:0/16:0 PC	UA	0,02	6,93	1,12
EM/CL de lípidos	20:4 LPC	UA	0,02	0,22	0,87
EM/CL de lípidos	16:0 SM	UA	0,02	7,91	1,08
Proteómica elegida como diana	SHBG	nmol/ml	0,03	6,81	1,25
EM/CL de lípidos	18:1/17:1/16:0 TG	UA	0,03	11,84	1,11
EM/CG	Arabinosa	UA	0,03	0,14	0,87
EM/CL de lípidos	18:1/18:1/17:0 TG	UA	0,03	10,68	1,16

Tabla 4B: Resultados del análisis unifactorial (Parte 2).

Plataforma	ID	Cuartil 1	Cuartil 2	Cuartil 3	Cuartil 4
EM/CL polar	Cisteína	< 2041,1	2041,2-2699,2	2699,3-3675,9	>3675,9
Proteómica elegida como diana	Factor de von Willebrand	<35,2	35,2-47,7	47,8-69,0	>69,0
Proteómica elegida como diana	IL-8	< 14,1	14,1-17,4	17,5-21,4	>21,4
EM/CL de lípidos	16:0/18:1 PC	<30075,8	30075,9-35320,5	35320,5-61392,4	>61392,4
EM/CG	N-carboxi- alanina	<3431,4	3431,4-3853,5	3853,6-4324,6	>4324,6
Proteómica elegida como diana	Fibrinógeno	<3,5	3,5-4,3	4,3-5,2	>5,2
Proteómica elegida como diana	MMP-2	<1037,5	1037,5-1520,0	1520,1-2157,5	>2157,5
EM/CL de lípidos	18:0/20:4 PE	<1861,8	1861,8-2138,4	2138,5-2429,0	>2429,0
Proteómica elegida como diana	Apolipoproteína A1	<0,30	0,30-0,36	0,37-0,46	>0,46
EM/CL de lípidos	16:0/22:6 PE	<1811,8	1811,8-2066,1	2066,2-2476,7	>2476,7
EM/CL de lípidos	18:1/18:0/18:0 TG	<15038,4	15038,4-17619,3	17619,4-22441,1	>22441,1
Proteómica elegida como diana	Alfa-1 Antitripsina	<1,50	1,50-1,69	1,70-1,96	>1,96
EM/CL de lípidos	18:2/18:1/17:0 TG	<1963,7	1963,7-4369,5	4369,6-12667,6	>12667,7
EM/CL de lípidos	20:1/18:1/18:1 TG	<44210,6	44210,6-54354,2	54354,3-68063,3	>68063,3
EM/CL de lípidos	16:0/16:0 PC	<4859,1	4859,1-6019,0	6019,1-8158,0	>8158,0
EM/CL de lípidos	20:4 LPC	<4203,5	4203,5-5082,6	5082,7-6311,4	>6311,4
EM/CL de lípidos	16:0 SM	<4289,8	4289,8-5239,6	5239,7-6554,6	>6554,6
Proteómica elegida como diana	SHBG	<25,8	25,8-38,8	38,9-62,9	>62,9
EM/CL de lípidos	18:1/17:1/16:0 TG	<1649,8	1649,8-1861,9	1862,0-2189,9	>2189,9
EM/CG	Arabinosa	<26944,7	26944,7-33111,6	33111,7-43695,4	>43695,4
EM/CL de lípidos	18:1/18:1/17:0 TG	<1827,5	1827,5-2473,7	2473,8-3245,7	>3245,7

En la Tabla 4 anterior, PC significa fosfatidilcolina, PE significa fosfatidiletanolamina, TG significa triacilglicerol, SM significa esfingomielina, LPC significa lisofosfatidilcolina, IL significa interleucina y x:y denota cadenas de ácido graso que contienen x átomos de carbono e y dobles enlaces. Las entradas en la columna indicada 'Razón de posibilidades' son las razones de posibilidades correspondientes a una comparación de sujetos con concentraciones de analitos en el cuarto cuartil con respecto a sujetos con concentraciones de analitos en el primer cuartil de la

distribución global de concentraciones para el analito a lo largo de todos los sujetos en el estudio. Para analitos medidos por espectrometría de masas, las unidades de medición se indican como 'unidades arbitrarias' o 'UA'. Las mediciones de concentración derivadas del espectrómetro de masas representan recuentos de iones procesados como se detecta por el detector del instrumento de espectrómetro de masas. Para analitos medidos por espectrometría de masas no se realizó cuantificación absoluta.

Los siguientes analitos de ácidos grasos están asociados con el riesgo de MACCE, en los que, respectivamente dentro de la lista de analitos identificados, el primer analito tiene la asociación más fuerte con riesgo de MACCE y el último analito tiene la asociación más débil con riesgo de MACCE: 18:2/18:1/17:0 TG, 18:0/20:4 PE, 16:0/18:1 PC, 16:0/22:6 PE, 20:4 LPC, 20:1/18:1/18:1 TG, 18:1/17:1/16:0 TG, 18:1/18:0/18:0 TG, 18:1/18:1/17:0 TG, 16:0 SM y 16:0/16:0 PC.

Para cada uno de los analitos de ácidos grasos identificados en la Tabla 4, excepto para 20:4 LPC, una mayor concentración del ácido graso puede estar más fuertemente asociada a riesgo de MACCE (por ejemplo, es más probable que los sujetos en el 4º cuartil de la distribución de población de 18:0/20:4 PE sufran una MACCE que aquellos en el cuartil más bajo, es decir, el 1º). Para 20:4 LPC puede ser lo opuesto: los sujetos en el menor cuartil, o 1º, de la distribución de población de 20:4 LPC tienen más probabilidad de sufrir una MACCE que aquellos en el mayor cuartil, es decir, el 4º.

Según algunas realizaciones, un médico puede identificar un sujeto con una medición de analito en el intervalo del 4º cuartil de la Tabla 4 anterior como un individuo posiblemente con alto riesgo de una MACCE en dos años (con la excepción de apolipoproteína A1, 20:4 LPC o arabinosa, que tienen razones de posibilidades inferiores a 1, significando que los sujetos con valores en el 1º cuartil tienen el mayor riesgo). Según otra realización, un médico puede identificar un sujeto con una medición de analito dentro de los intervalos del 3º o 4º cuartil de la Tabla 4 anterior como un individuo posiblemente con alto riesgo de una MACCE en dos años. Según otra realización, un médico puede identificar un sujeto con una medición de analito dentro de los intervalos del 2º, 3º o 4º cuartil de la Tabla 4 anterior como un individuo posiblemente con alto riesgo de una MACCE en dos años.

Procedimientos de clasificación multifactorial

Modelos de predicción multifactorial supervisados: Se construyeron modelos pronósticos multifactoriales en los que el resultado que iba a predecirse era la aparición de MACCE en el plazo de dos años desde la cateterización índice. Las variables de entrada al modelo fueron (i) analitos en plasma específicos de la plataforma bioanalítica y (ii) analitos en plasma como se identifican por todas las plataformas bioanalíticas en conjunto. Cada modelo multivariable también se construyó con y sin ajustar factores clínicos convencionales.

El procedimiento de clasificación multifactorial Random Forest (Breiman, Machine Learning 45:5-32, 2001) se aplicó a los datos para obtener huellas dactilares multifactoriales pronósticas de MACCE. La validación cruzada externa y la prueba de permutación de marcas se implementó con el fin de evaluar la significancia estadística de los modelos resultantes (Ambroise & McLachlan, Proc. Natl. Acad. Sci. 99:6562-2, 2002). Las siguientes salidas estuvieron disponibles de cada clasificador:

1. Los candidatos a **analito** en plasma multifactoriales pronósticos para la predicción de MACE en los grupos de 'Casos de enfermedad' y 'Controles' se informaron con analitos clasificados por importancia en la exactitud de clasificación.
2. Las curvas ROC de sensibilidad frente a especificidad para cada clasificador (véanse, por ejemplo, las Figuras 1 y 2).

Las características con más de un 50% de valores ausentes en tanto los casos de enfermedad como los controles se excluyeron del análisis multifactorial. Las características con menos de 50% de valores ausentes en casos de enfermedad y controles tuvieron valores ausentes atribuidos usando procedimientos descritos por Troyanskaya y col. (Bioinformatics 17(6) 520 - 525, 2001).

Analitos identificados según el procedimiento

Usando un procedimiento de análisis de clasificadores estadísticos multifactoriales (Random Forests - Breiman, Machine Learning 45:5-32, 2001) con el conjunto de datos bioanalítico combinado de 723 analitos, los inventores descubrieron conjuntos de analitos y algoritmos que podían usarse para predecir la aparición de una complicación cardiovascular o cerebrovascular adversa grave en un individuo dentro de dos años desde el momento en el que se obtiene una muestra de sangre. La siguiente Tabla 5 presenta los 50 analitos más importantes en su orden de importancia en la lista del total de 723 analitos en el clasificador de Random Forests basándose en una semilla de partida (Análisis nº 1). La lista de importancia de analitos parciales de un análisis de Random Forests y un enfoque de eliminación de características recurrentes para obtener un clasificador con exactamente 50 componentes de analito usando la misma semilla de partida que el Análisis nº 1 se muestra en el marcador de columna "Orden de importancia del Análisis nº 2". Las listas de importancia de analitos parciales de 3 análisis de Random Forests adicionales y enfoques de eliminación de características recurrentes para obtener clasificadores con exactamente 50 componentes de analito derivados de diferentes semillas de partida se muestran en la columnas para los Análisis nº

3, nº 4 y nº 5. Independientemente de la semilla de partida para los análisis de Random Forests o el uso de eliminación de características recurrentes para obtener un clasificador de componentes de 50 analitos fijo, los 20 analitos más importantes están casi siempre dentro de los 50 clasificados principales por importancia del Análisis nº 1.

5

Tabla 5: 50 analitos más importantes para predecir MACCE

Analito	Orden de importancia del Análisis nº 1	Orden de importancia del Análisis nº 2	Orden de importancia del Análisis nº 3	Orden de importancia del Análisis nº 4	Orden de importancia del Análisis nº 5
Factor tisular	1	3	3	2	3
Antígeno del cáncer 125	2	2	1	3	2
Glutación S-transferasa	3	1	2	1	1
Alfa-fetoproteína	4	4	4	4	4
IPI00028413 (ITIH3)	5	6	12	9	11
IL-3	6	9	8	7	8
103_0at233 (fragmento relacionado con arabinosa (M))	7	7	7	10	10
Factor de von Willebrand	8	5	6	5	7
IPI00011264(CFHR1)	9	12	15	12	12
IL-8	10		10	11	14
178_lat107 (cisteína M+H (M))	11	11	*	6	5
114_0at120 (fragmento relacionado con acetoacetato (M))	12	13	11	18	17
Factor VII	13		27	23	*
IPI00019755 (GSTO1)	14	25	17	35	21
292_0at200 (fragmento relacionado con ácido eritrónico (M))	15	16	25	*	20
IPI00022417 (LRG1)	16	30	37	15	15
MMP-2	17	14	9	14	6
Fibrinógeno	18	8	5	8	9
TNF RII	19	19	13	*	*
TNF-beta	20	10	16	21	16
IPI00550991 (SERPINA3)	21	20	29	19	18
IPI00298971 (VTN)	22	21	23	27	*
IPI00641737 (HP)	23	*	34	20	27
IPI00296176 (F9)	24	46	*	*	32
202_0at311 (L-Triptófano M+H (M))	25	*	30	36	*

Analito	Orden de importancia del Análisis nº 1	Orden de importancia del Análisis nº 2	Orden de importancia del Análisis nº 3	Orden de importancia del Análisis nº 4	Orden de importancia del Análisis nº 5
ICAM-1	26	28		25	19
217_Oat294 (fragmento relacionado con mioinositol (M))	27	*	31	*	*
103_Oat114 (fragmento relacionado con acetoacetato (M))	28		36	28	33
259_Oat375 (fragmento relacionado con P7478_uk15 desconocido (M))	29	*	*	*	43
IPI00167093 (CFHR1)	30	48			*
IPI00021885 (FGA)	31				*
IPI00022429 (ORM1)	32				*
IPI00019943 (AFM)	33	17	35	34	30
226_2at018 (metilhistidina M+H (M))	34	44	39	44	31
IPI00647915 (TAGLN2)	35				
191_Oat110 (fragmento relacionado con ácido 3-hidroxibutanoico (M))	36	*	*	*	*
361_Oat372 (fragmento relacionado con sacarosa (M))	37			*	*
230_Oat194 (fragmento relacionado con L-4-hidroxiprolina (M))	38	*		*	*
188_Oat097 (fragmento relacionado con ácido pirúvico (M))	39	*	*	*	*
IPI00290283 (MASP1)	40		*	*	*
IPI00654888 (KLKB1)	41	35	*	*	*
IPI00020996 (IGFALS)	42	*	*		
IPI00029658 (EFEMP1)	43	24	20	30	28
764_1213 (16:0/22:6 PE M+H (M))	44	33	*	37	*
IPI00745872 (ALB)	45	*	18	17	26
SGOT	46	15	19	13	13
278_2at143	47	*	*	*	
IPI00745933	48	*	*	*	

* El analito no era un miembro de los 50 analitos principales en el análisis específico

La lista completa de los analitos de análisis de Random Forests usando cuatro semillas de partida se muestra en los Apéndices 1-4 adjuntos junto con los 50 analitos principales de Random Forests más el enfoque de eliminación de

características recurrentes para las mismas cuatro semillas de partida en los Apéndices 5-8 adjuntos.

Usando los niveles de los 20 analitos principales del Análisis nº 1 en una muestra de plasma de un individuo junto con un algoritmo, la predicción para ese individuo de la aparición de una MACCE dentro de dos años se haría con 87% de sensibilidad y 87% de especificidad. Esto se ilustra por la curva de características de operación del receptor (ROC) mostrada en la Figura 1.

Usando un procedimiento de análisis de clasificadores estadísticos multifactoriales (Random Forests - Breiman, Machine Learning 45:5-32, 2001) con el conjunto de datos bioanalíticos de proteómica elegida como diana individual de 66 analitos, los inventores descubrieron conjuntos de analitos y algoritmos que podían usarse para predecir la aparición de un infarto de miocardio en un individuo en el plazo de dos años desde el momento en el que se obtiene una muestra de sangre. La siguiente Tabla 6 presenta los 10 analitos más importantes en un conjunto de analitos tal y su orden o importancia en un clasificador derivado del análisis multifactorial.

Tabla 6: 10 analitos más importantes para predecir MACCE

Analito de proteína	Importancia variable en el clasificador
Glutatión S-transferasa	0,048
Antígeno del cáncer 125	0,045
Factor tisular	0,04
Alfa-fetoproteína	0,028
Factor de von Willebrand	0,023
Fibrinógeno	0,02
SGOT	0,015
IL-3	0,014
IL-8	0,013
CD40	0,008

Usando los niveles de estos 10 analitos de proteína principales en una muestra de plasma de un individuo junto con un algoritmo, la predicción para ese individuo de la aparición de una MACCE en el plazo de dos años podría hacerse con 82% de sensibilidad y 87% de especificidad. Esto se ilustra por la curva de características de operación del receptor (ROC) mostrada en la Figura 2.

Según algunas realizaciones, una índice de MACCE puede definirse usando diez (10) de los principales analitos por la siguiente ecuación lineal:

$$Y = C_1X_1 + C_2X_2 + C_3X_3 + C_4X_4 + C_5X_5 + C_6X_6 + C_7X_7 + C_8X_8 + C_9X_9 + C_{10}X_{10}$$

[Ecuación 1]

en la que los valores de c_1 a c_{10} se corresponden con los valores enumerados en la Tabla 9 más adelante y las variables x_1 a x_{10} representan las mediciones de los analitos enumerados en la Tabla 7 más adelante. Los analitos se miden como se describe en la sección Proteómica elegida como diana: Inmunoensayos multiplejados anterior. Las unidades de medición para cada analito se enumeran en la tabla más adelante. Antes de entrar los valores de medición en la Ecuación 1, los valores deben normalizarse. Es decir, de cada valor de medición se restó el promedio de todos los valores derivados de la población en estudio entera (es decir, todos los casos y sujetos de control), y posteriormente el resultado se dividió entre la desviación estándar de la medición del analito derivada de la población en estudio entera (es decir, todos los casos y sujetos de control). Los valores correspondientes, concretamente los promedios de población y las desviaciones estándares, para los 10 analitos de la Ecuación 1 se enumeran en la Tabla 8.

Usando una ecuación tal como la Ecuación 1, los sujetos con mediciones de estos diez analitos, cada medición normalizada como se ha descrito anteriormente usando los valores apropiados de la Tabla 8, para los que el valor resultante y en la ecuación anterior, el índice de MACCE, es mayor que cero se clasificarían como que están en riesgo de MACCE en el plazo de dos años, y de otro modo se clasificarían como que no están en riesgo de MACCE en el plazo de dos años.

Tabla 7: Ecuación 1, variables x.

Variable en la Ecuación 1	Analito	Unidades de medición
x1	Alfa-fetoproteína	nanogramos / mililitro (ng/ml)
x2	Antígeno del cáncer 125	unidades / mililitro (U/ml)
x3	CD40	ng/ml
x4	Fibrinógeno	mg/ml
x5	Glutación S-transferasa	ng/ml
x6	IL-3	ng/ml
x7	IL-8	picogramos / mililitro (pg/ml)
x8	SGOT	microgramos por mililitro
x9	Factor tisular	ng/ml
x10	Factor de von Willebrand	microgramos por mililitro

Tabla 8: Valores estándar.

Analito representado en la Ecuación 1	Valor promedio de la población (en unidades de medición de la Tabla 7)	Desviación estándar (en unidades de medición de la Tabla 7)
Alfa-fetoproteína	1,729	1,138
Antígeno del cáncer 125	20,180	45,201
CD40	1,195	2,916
Fibrinógeno	4,515	1,603
Glutación S-transferasa	1,877	0,971
IL-3	0,555	0,139
IL-8	19,795	12,619
SGOT	26,108	7,069
Factor tisular	3,392	2,135
Factor de von Willebrand	54,738	29,320

Tabla 9: Ecuación 1, variables c

c1	-0,26243
c2	0,0947013
c3	0,00672228
c4	-0,102292
c5	0,138434
c6	-0,0191226
c7	-0,121236
c8	0,0659886
c9	0,181278
c10	-0,0900255

Todavía otro índice de MACCE puede definirse usando cuatro (4) de los principales analitos por la siguiente ecuación lineal:

$$y = c_1X_1 + c_2X_2 + c_3X_3 + c_4X_4$$

[Ecuación 2]

- 5 en la que los valores de c_1 a c_4 toman aquellos enumerados en la Tabla 11 más adelante y las variables x_1 a x_{10} representan las mediciones de los analitos enumerados en la Tabla 10 más adelante. Los analitos se miden como se describe en la sección *Proteómica elegida como diana: Inmunoensayos multiplexados* anterior. Las unidades de medición para cada analito se enumeran en la tabla más adelante. Antes de entrar los valores de medición en la Ecuación 2, los valores deben normalizarse. Es decir, de cada valor de medición se restó el promedio de todos los valores derivados de la población en estudio entera (es decir, todos los casos y sujetos de control), y posteriormente el resultado se dividió entre la desviación estándar de la medición del analito derivada de la población en estudio entera (es decir, todos los casos y sujetos de control). Los valores correspondientes, concretamente los promedios de población y las desviaciones estándares, para los 4 analitos de la Ecuación 2 se enumeran en la Tabla 8 anterior.

Tabla 10: Ecuación 2, variables x

Variable en la Ecuación 2	Analito	Unidades de medición
x1	Alfa-fetoproteína	nanogramos / mililitro (ng/ml)
x2	Antígeno del cáncer 125	unidades / mililitro (U/ml)
x3	Glutación S-transferasa	ng/ml
x4	Factor tisular	ng/ml

Tabla 11: Ecuación 2, variables c

c1	-0,240085
c2	0,0331361
c3	0,17824
c4	0,207721

Usando una ecuación tal como la Ecuación 2, los sujetos con mediciones de estos cuatro (4) analitos, cada medición normalizada como se ha descrito anteriormente usando los valores apropiados de la Tabla 8, para los que el valor resultante y en la ecuación anterior, el índice de MACCE, es mayor que cero se clasificarían como que están en riesgo de MACCE en el plazo de dos años, y de otro modo se clasificarían como que no están en riesgo de MACCE en el plazo de dos años.

Un procedimiento de análisis de clasificación multifactorial adicional, el análisis de predicción de micromatrices (PAM, Tibshirani y col., PNAS 99:6567-6572, 2002), se aplicó al conjunto de datos de plataformas combinadas de 723 analitos. Este análisis dio conjuntos de analitos mínimos óptimos para la clasificación de casos o controles. Los conjuntos de analitos para cuatro (4) análisis de PAM usando las mismas cuatro semillas que aquellas para los Análisis n° 2, n° 3, n° 4 y n° 5, en los casos de Random Forests anteriores, se muestran en las siguientes tablas:

Tabla 12: Semilla n° 1

Analito	Valores delta escalados	
	1. Caso	2. Control
Glutación S-transferasa	0,079	-0,079
Antígeno del cáncer 125	0,062	-0,062
Factor de von Willebrand	-0,047	0,047
Factor tisular	0,041	-0,041
Alfa-fetoproteína	0,036	-0,036
178_1at107 (cisteína M+H (M))	-0,03	0,03
MMP-2	-0,023	0,023
IL-8	-0,013	0,013
TNF-beta	0,006	-0,006
TNF RII	-0,006	0,006

Sensibilidad = 83,8%; especificidad = 83,8%

Tabla 13: Semilla n° 2

Analito	Valores delta escalados	
	1. Caso	2. Control
Glutación S-transferasa	0,079	-0,079
Antígeno del cáncer 125	0,062	-0,062
Factor de von Willebrand	-0,047	0,047
Factor tisular	0,041	-0,041
Alfa-fetoproteína	0,036	-0,036
178_1at107 (cisteína M+H (M))	-0,03	0,03
MMP-2	-0,023	0,023
IL-8	-0,013	0,013
TNF RII	-0,006	0,006
TNF-beta	0,006	-0,006

Sensibilidad = 82,4%; especificidad = 80,9%

Tabla 14: Semilla n° 3

Analito	Valores delta escalados	
	1. Caso	2. Control
Glutación S-transferasa	-0,095	0,095
Antígeno del cáncer 125	-0,078	0,078
Factor de von Willebrand	0,063	-0,063
Factor tisular	-0,057	0,057
Alfa-fetoproteína	-0,052	0,052
178_1at107 (Cisteína M+H (M))	0,046	-0,046
MMP-2	0,039	-0,039
IL-8	0,029	-0,029
TNF-beta	-0,022	0,022
TNF RII	0,021	-0,021
103_0at233 (fragmento relacionado con arabinosa (M))	0,008	-0,008
IPI00022417 (LRG1)	0,006	-0,006
IPI00745872 (ALB)	-0,005	0,005

Sensibilidad = 80,9%; especificidad = 80,9%

Tabla 15: Semilla n° 4

Analito	Valores delta escalados	
	1. Caso	2. Control
Glutación S-transferasa	-0,087	0,087
Antígeno del cáncer 125	-0,07	0,07
Factor de von Willebrand	0,055	-0,055
Factor tisular	-0,049	0,049
Alfa-fetoproteína	-0,044	0,044
178_1at107 (cisteína M+H (M))	0,038	-0,038
MMP-2	0,031	-0,031
IL-8	0,021	-0,021
TNF RII	0,014	-0,014
TNF-beta	-0,014	0,014
103_0at233 (fragmento relacionado con arabinosa (M))	0	0

Sensibilidad = 80,9%; especificidad = 82,4%

- 5 Este análisis de clasificadores multifactoriales alternativos por PAM dio esencialmente el mismo conjunto de analitos para cada una de las diferentes semillas de partida y los analitos se clasificaron en orden de importancia de acuerdo con, aunque idéntico a, los resultados de los análisis de Random Forests.

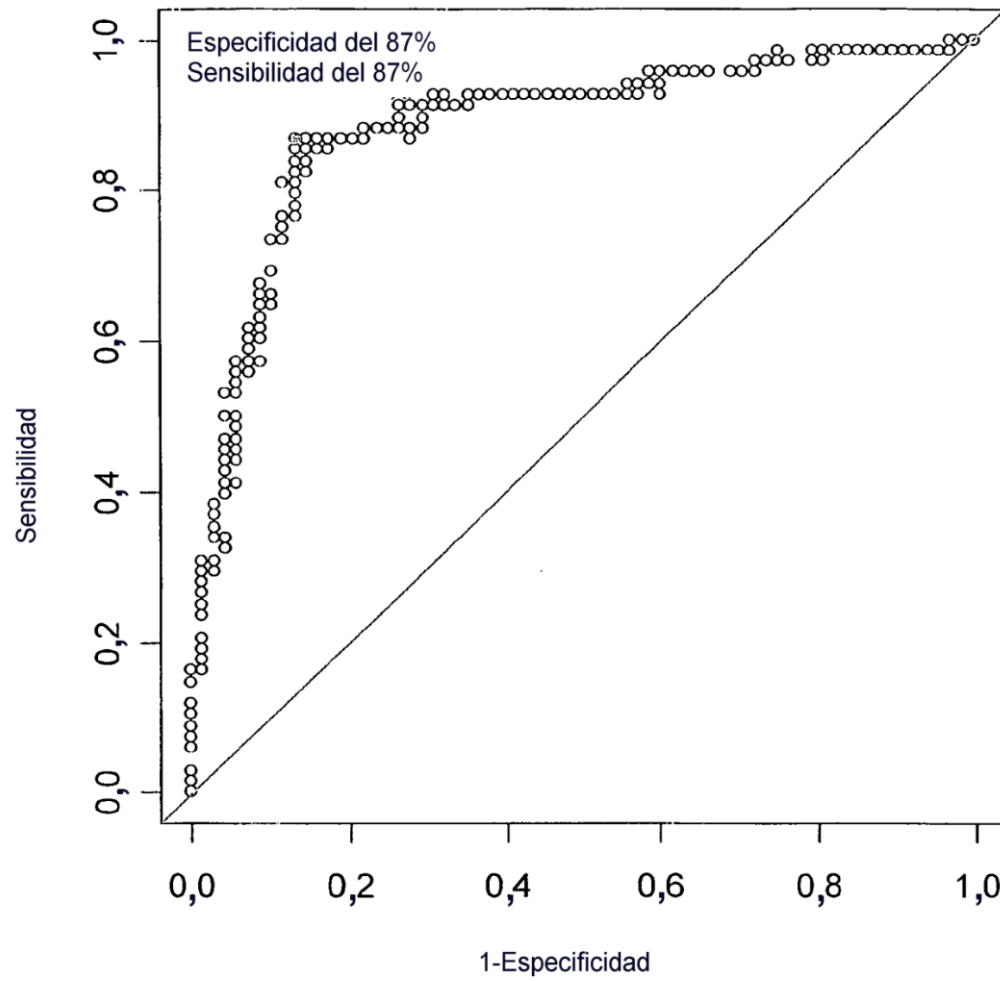
Las presentes enseñanzas incluyen las identidades de los conjuntos de analitos que pueden medirse en muestras

de sangre de un individuo y usarse conjuntamente con algoritmos para calcular un índice de MACCE y para predecir la aparición de una MACCE en el plazo de dos años desde el momento del muestreo de sangre para ese individuo.

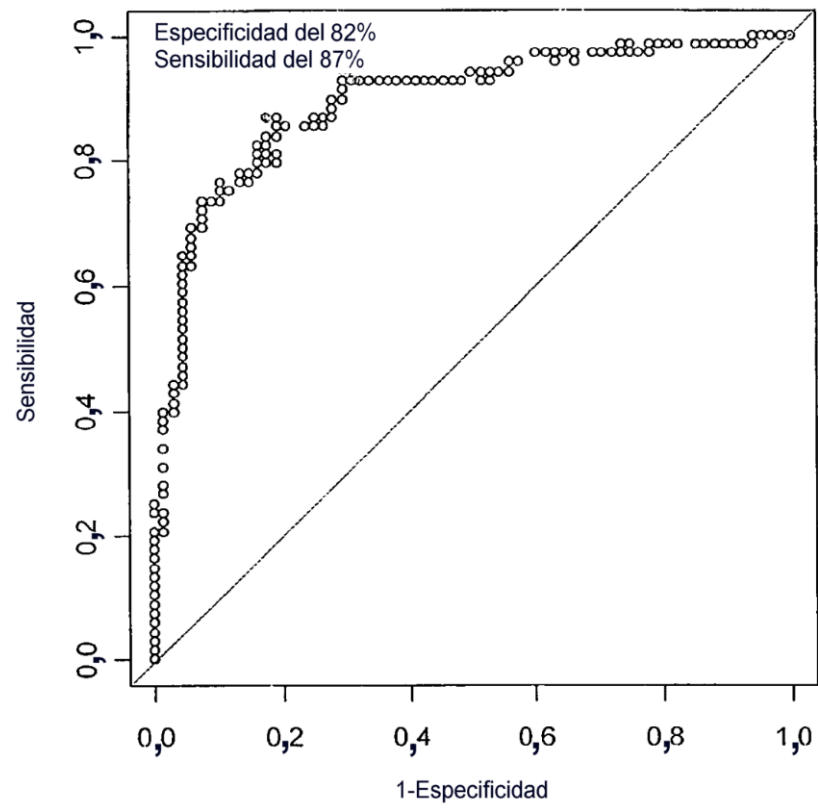
- 5 Las presentes enseñanzas también incluyen los algoritmos que pueden usarse con información sobre niveles de sangre, plasma en sangre o suero en sangre de conjuntos de analitos para calcular un índice de MACCE y para predecir la aparición de una MACCE en el plazo de dos años desde el momento del muestreo de sangre para ese individuo como se describe para los conjuntos de analitos tales como este por Breiman (Machine Learning 45:5-32, 2001).

REIVINDICACIONES

1. Un procedimiento de evaluación de la probabilidad de una complicación cardiovascular o cerebrovascular adversa grave en un ser humano, procedimiento que comprende:
- 5 medir una concentración, en una muestra basada en sangre de un ser humano, de un conjunto de analitos que comprende alfa-fetoproteína, antígeno del cáncer 125, glutatión S-transferasa y factor tisular;
- determinar un índice para el conjunto de analitos; e
- identificar el ser humano que tiene una alta probabilidad de una complicación cardiovascular o cerebrovascular adversa grave si el índice es mayor que cero, o una baja probabilidad de una complicación cardiovascular o cerebrovascular adversa grave si el índice es inferior o igual a cero,
- 10 en el que determinar un índice para el conjunto de analitos comprende:
- normalizar la concentración medida de cada analito para obtener una concentración normalizada;
- multiplicar la concentración normalizada de cada analito por una constante de analito para obtener un valor de analito; y
- sumar el valor de analito de cada analito para obtener el índice.
- 15 2. El procedimiento de cualquiera de la reivindicación 1, en el que el conjunto de analitos comprende además CD40, fibrinógeno, IL-3, IL-8, SGOT y factor de von Willebrand.
3. El procedimiento de la reivindicación 1 ó 2, en el que normalizar la concentración medida comprende restar de la concentración medida un valor promedio de la población para obtener un resultado y luego dividir el resultado entre una desviación estándar del valor promedio de la población.
- 20 4. El procedimiento de una cualquiera de las reivindicaciones 1-3, en el que la muestra basada en sangre comprende suero o plasma.
5. El procedimiento de una cualquiera de las reivindicaciones 1-4 que comprende además transmitir, mostrar, almacenar o proporcionar el índice a un dispositivo de interfaz del usuario, un medio de almacenamiento legible por ordenador o un sistema de ordenador local o remoto.
- 25



El área bajo esta curva de ROC es 0,89.



El área bajo esta curva de ROC es 0,90.