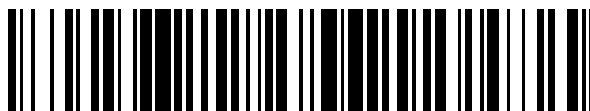


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 742 826**

51 Int. Cl.:

G16H 10/60 (2008.01)

G16H 50/70 (2008.01)

G16H 10/40 (2008.01)

G06F 19/00 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **05.05.2010 PCT/US2010/033695**

87 Fecha y número de publicación internacional: **11.11.2010 WO10129653**

96 Fecha de presentación y número de la solicitud europea: **05.05.2010 E 10772754 (7)**

97 Fecha y número de publicación de la concesión europea: **01.05.2019 EP 2427103**

54 Título: **Sistema y procedimiento para evaluación rápida de distribuciones de valores de laboratorio**

30 Prioridad:

05.05.2009 US 175538 P

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

17.02.2020

73 Titular/es:

**OPTUMINSIGHT, INC. (100.0%)
13625 Technology Drive, MN-051-A130
Eden Prairie, MN 55344, US**

72 Inventor/es:

ANDERSON, DAVID R.

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 742 826 T3

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

DESCRIPCIÓN

Sistema y procedimiento para evaluación rápida de distribuciones de valores de laboratorio

Antecedentes de la invención

Campo de la invención

- 5 La presente invención se refiere al análisis de datos relacionados con la salud y, más particularmente, se refiere a un sistema y a un procedimiento para la evaluación rápida de distribuciones de valores de laboratorio entre grupos de comparación.

Descripción de la técnica relacionada

- 10 La mayoría de las corporaciones, incluyendo las corporaciones de seguros de salud, mantienen un alto volumen de datos. Dichos datos pueden analizarse y explotarse para obtener información valiosa sobre las tendencias comerciales y otras estadísticas importantes. La minería de datos es una estrategia común para identificar y analizar dichos datos.

- 15 Hay muchas formas diferentes de minería de datos. Se pueden desarrollar operaciones analíticas personalizadas para satisfacer necesidades específicas. Como alternativa, herramientas de análisis estadístico disponibles comercialmente, como el software de análisis estadístico (SAS) se puede utilizar para identificar tendencias estadísticas en los datos.

- 20 Las compañías de seguros de salud generalmente mantienen bases de datos de información de reclamaciones de seguros de salud, información demográfica y otros datos sobre los miembros del plan de seguro de salud. Dicha información se puede utilizar para obtener información valiosa sobre el diagnóstico temprano de enfermedades, relación entre pruebas de laboratorio y enfermedades o tratamientos farmacológicos, y la gravedad de las enfermedades. Desafortunadamente, los procedimientos típicos para analizar tales datos son a menudo engorrosos, costosos y requieren tiempos y recursos de procesamiento inmanejables.

- 25 Las deficiencias a las que se hace referencia no pretenden ser exhaustivas, sino que se encuentran entre muchos que tienden a perjudicar la efectividad de las técnicas previamente conocidas en el manejo de enfermedades, diagnóstico y tratamiento; sin embargo, los mencionados aquí son suficientes para demostrar que los procedimientos que aparecen en la técnica no han sido satisfactorios y que existe una necesidad significativa de las técnicas descritas y reivindicadas en la presente divulgación.

- 30 El documento US 2009/076845 A1 desvela sistemas y procedimientos para facilitar la gestión de la atención médica en instalaciones médicas tales como hospitales, clínicas, salas de emergencia, etc., por ejemplo, monitorizando y evaluando la calidad de la atención al paciente. El documento US 2005/004895 A1 desvela sistemas informáticos separados que pueden participar en una red de datos de salud (HDN) de modo que los sistemas informáticos estén vinculados para compartir varios tipos de información relacionada con la atención médica. El documento US 6018713 A desvela un sistema y un procedimiento basado en la red para ordenar e informar resultados acumulativos de pruebas médicas que incluye un ordenador operado en una ubicación médica (tal como un hospital o consultorio
- 35 médico) para ordenar pruebas, recuperar y almacenar datos estadísticos o el estado del progreso de las pruebas ordenadas previamente. El documento US 2009/076841 A1 desvela una aplicación y un entorno de interfaz de usuario de software basados en reglas, que comprende: a) un esquema de base de datos de biblioteca de medidas de calidad clínica, b) un esquema de base de datos de especificaciones de medidas de calidad clínica, c) una biblioteca de reglas, d) un editor de reglas comerciales clínicas (CBR), e) un indicador clave de rendimiento (KPI) o un editor de especificaciones de medidas de calidad clínica, f) un módulo de configuración de tiempo de ejecución de calidad clínica, g) un módulo de configuración de tiempo de ejecución de costo de cuidado y h) un módulo de administrador de tiempo de ejecución y, en algunas realizaciones, i) un módulo de prueba de medidas. Siadaty Mir S y col. desvela la localización de patrones previamente desconocidos en los resultados de minería de datos.

Sumario de la invención

- 45 A partir de la descripción anterior, debería ser evidente que existe la necesidad de un sistema y procedimiento para analizar rápidamente, descubrir y resumir diferencias en la distribución de valores de laboratorio entre grupos de comparación, particularmente en cualquier contexto temporal de cualquier atributo temporal.

El ámbito de la invención se define mediante las reivindicaciones adjuntas.

- 50 Se presenta un sistema para la evaluación rápida de las distribuciones de valor de laboratorio. En un ejemplo, el sistema incluye un dispositivo de almacenamiento de datos configurado para almacenar una base de datos que comprende uno o más registros, en el que los registros están identificados por uno o más códigos de prueba. El sistema también puede incluir un servidor en comunicación de datos con el dispositivo de almacenamiento de datos. El servidor puede programarse adecuadamente para recibir un código médico (donde los valores de laboratorio pueden incluir un rango de interés), buscar en la base de datos para obtener un primer grupo de registros asociados

5 con personas que tienen el código médico, en el que cada registro del primer grupo de registros se identifica por un código de prueba, buscar en la base de datos para obtener un segundo grupo de registros asociados con una población de control, en el que cada registro del segundo grupo de registros se identifica mediante el código de prueba y genera una salida que comprende un gráfico de distribución en una escala compartida del primer y segundo grupo de registros.

10 En un ejemplo, el servidor puede comparar un primer valor que representa el primer grupo y un segundo valor que representa el segundo grupo de registros. En otro ejemplo, el servidor puede interpolar el primer grupo de registros para obtener el primer valor e interpolar el segundo grupo de registros para obtener el segundo valor, en el que el primer y segundo valores están asociados con un percentil compartido, tal como el 5 %, 32 %, 50 %, 95 %, o cualquier percentil intermedio. Por ejemplo, se pueden comparar los valores de la mediana en ambos grupos.

15 En cierto ejemplo, el servidor puede contar registros distintos en el primer grupo de registros y el segundo grupo de registros. El servidor también puede agregar registros del primer grupo y del segundo grupo de acuerdo con un atributo seleccionado, tal como un rango seleccionado de valores de prueba, y/o calcular una probabilidad en respuesta a una relación de un número de registros en el primer grupo que tiene el atributo seleccionado y un número de registros en el segundo grupo que tiene el atributo seleccionado.

El servidor puede seleccionar además el primer grupo de registros dentro de un intervalo de tiempo, como un lapso de tiempo (por ejemplo, aproximadamente 30 días) antes o después del diagnóstico inicial o inicio de la enfermedad. En una realización adicional, el servidor puede seleccionar registros en el primer grupo y en el segundo grupo de acuerdo con un criterio limitante, tal como un parámetro relacionado con la enfermedad.

20 También se presenta un procedimiento para comparar dos grupos de registros para analizar su discrepancia de distribución. En una realización, el procedimiento incluye recibir un código médico, buscar en una base de datos almacenada en un dispositivo de almacenamiento de datos para obtener un primer grupo de registros asociados con personas que tienen el código médico, en el que cada registro del primer grupo de registros se identifica por un código de prueba, buscar en la base de datos para obtener un segundo grupo de registros asociados con una población de control, en el que cada registro del segundo grupo de registros se identifica mediante el código de prueba y genera una salida que comprende un gráfico de distribución en una escala compartida del primer y segundo grupo de registros. El procedimiento en las realizaciones desveladas incluye sustancialmente las etapas necesarias para llevar a cabo las funciones presentadas anteriormente con respecto a la operación del sistema de la presente invención.

30 El procedimiento también puede incluir comparar un primer valor que representa el primer grupo y un segundo valor que representa el segundo grupo de registros. De manera adicional, el procedimiento también puede incluir la interpolación del primer grupo de registros para obtener el primer valor y la interpolación del segundo grupo de registros para obtener el segundo valor, en el que el primer y segundo valores están asociados con un percentil compartido, tal como cualquier percentil intermedio de 5 %-95 %, o. En una realización particular, se pueden comparar los valores medios asociados.

40 En cierta realización, el procedimiento puede incluir contar registros distintos en el primer grupo de registros y el segundo grupo de registros. El procedimiento también puede incluir la agregación de registros del primer grupo y del segundo grupo de acuerdo con un atributo seleccionado, tal valor de prueba o rango. En otra realización más, el procedimiento también puede comprender calcular una probabilidad en respuesta a una relación de una cantidad de registros en el primer grupo que tiene el atributo seleccionado y una cantidad de registros en el segundo grupo que tiene el atributo seleccionado.

45 Para procesar registros en cualquier intervalo temporal relacionado con una enfermedad, el procedimiento puede comprender además seleccionar el primer grupo de registros dentro de un intervalo de tiempo. Tal intervalo de tiempo puede ser relativo a la fecha o la hora de la primera aparición de la enfermedad asociada con el código médico o la fecha o la hora en que se diagnostica por primera vez. En otra realización adicional, el procedimiento puede comprender seleccionar registros en el primer grupo y en el segundo grupo de acuerdo con un criterio limitante, tal como un parámetro relacionado con la enfermedad, edad, género o similar.

50 También se puede proporcionar un producto de programa de ordenador tangible que comprende un medio legible por ordenador que tiene un código de programa utilizable por ordenador ejecutable para realizar operaciones que comprenden: recibir un código médico; buscar en una base de datos almacenada en un dispositivo de almacenamiento de datos para obtener un primer grupo de registros asociados con personas que tienen el código médico, en el que cada registro del primer grupo de registros se identifica por un código de prueba; buscar en la base de datos para obtener un segundo grupo de registros asociados con una población de control, en el que cada registro del segundo grupo de registros se identifica por el código de prueba; y generar una salida que comprende un gráfico de distribución en una escala compartida del primer y segundo grupo de registros.

55 Las operaciones también pueden incluir comparar un primer valor que representa el primer grupo y un segundo valor que representa el segundo grupo de registros. Más aún, las operaciones pueden incluir la interpolación del primer grupo de registros para obtener el primer valor y la interpolación del segundo grupo de registros para obtener el

segundo valor, en el que el primer y segundo valores están asociados con un percentil compartido.

5 En cierta realización, las operaciones pueden incluir contar registros distintos en el primer grupo de registros y el segundo grupo de registros. Las operaciones también pueden incluir la agregación de registros del primer grupo y del segundo grupo de acuerdo con un atributo seleccionado. En otra realización más, las operaciones también pueden incluir calcular una probabilidad en respuesta a una relación de una cantidad de registros en el primer grupo que tiene el atributo seleccionado y una cantidad de registros en el segundo grupo que tiene el atributo seleccionado.

10 Para un análisis temporal, las operaciones pueden comprender además seleccionar el primer grupo de registros dentro de un intervalo de tiempo. En otra realización adicional, las operaciones también pueden incluir la selección de registros en el primer grupo y en el segundo grupo de acuerdo con un criterio limitante, tal como un parámetro relacionado con la enfermedad, edad, género o similar.

15 Para un registro con múltiples valores de pruebas de laboratorio, el registro podría comprender un promedio, una primera lectura o una última lectura de una prueba. Para analizar más a fondo los parámetros asociados a la enfermedad, cada registro en el primer grupo de registros puede compartir un criterio limitante con cada registro en el segundo grupo de registros, por ejemplo, esos registros en el primer y segundo grupo pueden tener el mismo rango de edad, género o residencia geográfica. Asimismo, la población de control podría ser un grupo normal o un grupo de cohorte. En ciertos aspectos, el gráfico de distribución puede ser un histograma o cualquier formato adecuado para presentar una distribución.

20 El término "asociado" se denomina como conectado o relacionado. El término "acoplado" se define como conectado, aunque no necesariamente directamente, y no necesariamente mecánicamente.

Los términos "un" y "una" se definen como uno o más a menos que esta divulgación explícitamente requiera lo contrario.

25 El término "sustancialmente" y sus variaciones se definen en gran medida, pero no necesariamente en su totalidad, como lo entiende un experto en la materia, y en una realización no limitante "sustancialmente" se refiere a intervalos dentro del 10 %, preferentemente dentro del 5 %, más preferentemente dentro del 1 %, y lo más preferiblemente dentro del 0,5 % de lo especificado.

30 Los términos "comprende" (y cualquier forma de comprender, como "comprenden" y "que comprende"), "tiene" (y cualquier forma de tener, como "tienen" y "que tiene"), "incluye" (y cualquier forma de incluir, como "incluyen" y "que incluye") y "contiene" (y cualquier forma de contener, como "contienen" y "que contiene") son verbos de enlace abiertos. Como resultado, un procedimiento o dispositivo que "comprende" "tiene", "incluye" o "contiene" una o más etapas o elementos posee esas una o más etapas o elementos, pero no se limita a poseer solo esos uno o más elementos. Igualmente, una etapa de un procedimiento o un elemento de un dispositivo que "comprende" "tiene", "incluye" o "contiene" una o más características posee esas una o más características, pero no se limita a poseer solo esas una o más características. Asimismo, un dispositivo o estructura que se configura de cierta manera se configura al menos de esa manera, pero también puede configurarse de formas que no están en la lista.

35 Otras características y ventajas asociadas serán evidentes con referencia a la siguiente descripción detallada de realizaciones específicas en relación con los dibujos adjuntos.

Breve descripción de los dibujos

40 Los siguientes dibujos forman parte de la presente memoria descriptiva y se incluyen para demostrar adicionalmente ciertos aspectos de la presente invención. La invención puede entenderse mejor por referencia a uno o más de estos dibujos en combinación con la descripción detallada de realizaciones específicas presentadas en el presente documento.

La figura 1 es un diagrama de bloques esquemático que ilustra una realización de un sistema para la evaluación rápida de distribuciones de valores de laboratorio;

45 La figura 2 es un diagrama de bloques esquemático que ilustra una realización de un sistema de base de datos para la evaluación rápida de distribuciones de valores de laboratorio;

La figura 3 es un diagrama de bloques esquemático que ilustra una realización de un sistema informático que puede usarse de acuerdo con ciertas realizaciones del sistema para una evaluación rápida de las distribuciones de valores de laboratorio;

50 La figura 4 es un diagrama lógico esquemático que ilustra una realización de capas de operación de abstracción en un sistema para la evaluación rápida de distribuciones de valores de laboratorio;

La figura 5 es un diagrama de bloques esquemático que ilustra una realización de un sistema para la evaluación rápida de distribuciones de valores de laboratorio;

La figura 6 es un diagrama de bloques esquemático que ilustra una realización de un sistema para la evaluación rápida de distribuciones de valores de laboratorio;

La figura 7 es un diagrama de flujo esquemático que ilustra una realización de un procedimiento para la evaluación rápida de distribuciones de valores de laboratorio;

5 La figura 8 es un diagrama de flujo esquemático que ilustra una realización de un procedimiento para la evaluación rápida de distribuciones de valores de laboratorio;

La figura 9 es un diagrama de flujo esquemático que ilustra una realización de un procedimiento para la evaluación rápida de distribuciones de valores de laboratorio;

10 La figura 10 es una salida que ilustra una realización de resultados de descubrimiento generados por una evaluación rápida de distribuciones de valor de laboratorio;

La figura 11 es una salida que ilustra una realización de resultados de distribución generados por una evaluación rápida de distribuciones de valor de laboratorio;

La figura 12 es un ejemplo de interpolación lineal que puede usarse para una evaluación rápida de las distribuciones de valores de laboratorio;

15 La figura 13 es una salida que ilustra una realización de resultados de descubrimiento generados por una evaluación rápida de distribuciones de valor de laboratorio;

La figura 14 es una salida que ilustra una realización de resultados de distribución generados por una evaluación rápida de distribuciones de valor de laboratorio;

20 La figura 15 es un diagrama de captura de pantalla que ilustra una realización de una interfaz gráfica de usuario para una evaluación rápida de las distribuciones de valores de laboratorio.

Descripción detallada

Diversas características y detalles ventajosos se explican más detalladamente con referencia a las realizaciones no limitativas que se ilustran en los dibujos adjuntos y detallados en la siguiente descripción. Descripciones de materiales de partida bien conocidos, técnicas de procesamiento, componentes y equipos se omiten para no oscurecer innecesariamente la invención en detalle. Debe entenderse, sin embargo, que la descripción detallada y los ejemplos específicos, aunque indican realizaciones de la invención, se dan solo a modo de ilustración, y no a modo de limitación. Varias sustituciones, modificaciones, adiciones y/o reorganizaciones dentro del ámbito del concepto inventivo subyacente resultarán evidentes para los expertos en la materia a partir de la presente divulgación.

30 Ciertas unidades descritas en la presente memoria descriptiva han sido etiquetadas como módulos, para enfatizar más particularmente su independencia de implementación. Un módulo es "[un] componente de hardware o software autónomo que interactúa con un sistema más grande". Alan Freedman, "The Computer Glossary" 268 (8ª ed. 1998). Un módulo comprende un componente de una máquina, una máquina o una pluralidad de máquinas que están adecuadamente programadas para operar de acuerdo con instrucciones ejecutables. Por ejemplo, un módulo puede implementarse como un circuito de hardware que comprende circuitos VLSI personalizados o conjuntos de puertas, semiconductores estándar como chips lógicos, transistores u otros componentes discretos. También se puede implementar un módulo en dispositivos de hardware programables, como disposiciones de puertas programables en campo, lógica de matriz programable, dispositivos lógicos programables, un controlador o similar.

40 Los módulos también pueden incluir unidades definidas por software o instrucciones que, cuando lo ejecuta una máquina o dispositivo de procesamiento, recuperan y transforman datos almacenados en un dispositivo de almacenamiento de datos de un primer estado a un segundo estado. Un módulo identificado de código ejecutable puede, por ejemplo, comprender uno o más bloques físicos de instrucciones del ordenador que pueden organizarse como un objeto, procedimiento o función. No obstante, los ejecutables de un módulo identificado no necesitan estar físicamente ubicados juntos, pero puede incluir instrucciones dispares almacenadas en diferentes ubicaciones que, cuando se unen lógicamente, comprende el módulo y, cuando lo ejecuta el procesador, logran la transformación de datos establecida.

50 De hecho, un módulo de código ejecutable puede ser una sola instrucción, o muchas instrucciones, e incluso puede distribuirse en varios segmentos de código diferentes, entre diferentes programas y en varios dispositivos de memoria. De manera similar, los datos operativos pueden identificarse e ilustrarse en el presente documento dentro de los módulos, y pueden incorporarse en cualquier forma adecuada y organizarse dentro de cualquier tipo adecuado de estructura de datos. Los datos operativos pueden recopilarse como un solo conjunto de datos, o pueden distribuirse en diferentes ubicaciones, incluyendo diferentes dispositivos de almacenamiento.

En la descripción que sigue, se proporcionan numerosos detalles específicos, tal como ejemplos de programación, módulos de software, selecciones de usuario, transacciones de red, consultas de bases de datos, estructuras de

bases de datos, módulos de hardware, circuitos de hardware, chips de hardware, etc., para proporcionar una comprensión profunda de las presentes realizaciones. Un experto en la materia relevante reconocerá, sin embargo, que la invención se puede practicar sin uno o más de los detalles específicos, o con otros procedimientos, componentes, materiales, etc. En otros casos, estructuras bien conocidas, materiales u operaciones no se muestran o describen en detalle para evitar ocultar aspectos de la invención.

La figura 1 ilustra una realización de un sistema 100 para la evaluación rápida de distribuciones de valores de laboratorio. El sistema 100 puede incluir un servidor 102, un dispositivo 104 de almacenamiento de datos, una red 108 y un dispositivo 110 de interfaz de usuario. En una realización adicional, el sistema 100 puede incluir un controlador 106 de almacenamiento, o un servidor de almacenamiento, configurado para gestionar comunicaciones de datos entre el dispositivo 104 de almacenamiento de datos y el servidor 102 u otros componentes en comunicación con la red 108. En una realización alternativa, el controlador 106 de almacenamiento puede estar acoplado a la red 108. En una realización general, el sistema 100 puede almacenar bases de datos que comprenden registros, realizan búsquedas de esos registros y generan resultados en respuesta a la información contenida en estos registros. De manera específica, el sistema 100 puede recibir un código médico (donde los valores de laboratorio pueden incluir un rango de interés) y generar una salida basada en un primer grupo de registros asociados con una persona que tiene un código médico en comparación con un segundo grupo de registros de una población de control.

En una realización, el dispositivo 110 de interfaz de usuario se conoce ampliamente y está destinado a abarcar un dispositivo basado en un procesador adecuado, tal como un ordenador de escritorio, un ordenador portátil, un Asistente Digital Personal (PDA), un dispositivo de comunicación móvil o dispositivo organizador que tiene acceso a la red 108. En una realización adicional, el dispositivo 110 de interfaz de usuario puede acceder a Internet para acceder a una aplicación web o servicio web alojado por el servidor 102 y proporcionar una interfaz de usuario para permitir que un usuario ingrese o reciba información. Por ejemplo, el usuario puede ingresar un código médico, un código de prueba, un intervalo de tiempo, un criterio limitante, un atributo seleccionado para generar una salida, o similar.

La red 108 puede facilitar las comunicaciones de datos entre el servidor 102 y el dispositivo 110 de interfaz de usuario. La red 108 puede incluir cualquier tipo de red de comunicaciones, incluyendo, pero sin limitación, una conexión directa de PC a PC, una red de área local (LAN), una red de área amplia (WAN), una conexión de módem a módem, Internet, una combinación de lo anterior, o cualquier otra red de comunicaciones ahora conocida o desarrollada más tarde dentro de las técnicas de redes que permiten la comunicación de dos o más ordenadores, entre sí.

En una realización, el servidor 102 está programado adecuadamente para recibir un código médico (donde los valores de laboratorio pueden incluir un rango de interés), buscar en la base de datos para obtener un primer grupo de registros asociados con personas que tienen el código médico, en el que cada registro del primer grupo de registros se identifica por un código de prueba, buscar en la base de datos para obtener un segundo grupo de registros asociados con una población de control, en el que cada registro del segundo grupo de registros se identifica mediante el código de prueba y genera una salida que comprende un gráfico de distribución en una escala compartida del primer y segundo grupo de registros. De manera adicional, el servidor puede acceder a los datos almacenados en el dispositivo 104 de almacenamiento de datos a través de una conexión de red de área de almacenamiento (SAN), una LAN, un bus de datos o similar.

El dispositivo 104 de almacenamiento de datos puede incluir un disco duro, incluyendo discos duros dispuestos en una matriz redundante de matriz de discos independientes (RAID), una unidad de almacenamiento de cinta que comprende un dispositivo de almacenamiento de datos de cinta magnética, un dispositivo de almacenamiento óptico, o similar. En una realización, el dispositivo 104 de almacenamiento de datos puede almacenar datos relacionados con la salud, tales como datos de reclamaciones de seguros, datos del consumidor, o similares. Los datos se pueden organizar en una base de datos y acceder a los mismos mediante consultas de lenguaje de consulta estructurado (SQL) u otros lenguajes u operaciones de consulta de base de datos.

La figura 2 ilustra una realización de un sistema 200 de gestión de datos configurado para almacenar y gestionar datos para una evaluación rápida de distribuciones de valores de laboratorio. En una realización, el sistema 200 puede incluir un servidor 102. El servidor 102 puede estar acoplado a un bus 202 de datos. En una realización, el sistema 200 también puede incluir un primer dispositivo 204 de almacenamiento de datos, un segundo dispositivo 206 de almacenamiento de datos y/o un tercer dispositivo 208 de almacenamiento de datos. En realizaciones adicionales, el sistema 200 puede incluir dispositivos de almacenamiento de datos adicionales (no mostrados). En una tal realización, cada dispositivo 204-208 de almacenamiento de datos puede alojar una base de datos separada de datos de reclamaciones de atención médica, datos de laboratorio, datos de pruebas físicas, datos de progresión de enfermedades, datos demográficos, datos socioeconómicos, datos administrativos, datos clínicos, o similares. La información del cliente en cada base de datos puede estar asociada a un campo o identificador común, tal como el nombre de una persona, número de seguridad social, número de cliente o similar. Como alternativa, los dispositivos 204-208 de almacenamiento pueden disponerse en una configuración RAID para almacenar copias redundantes de la base de datos o bases de datos a través de actualizaciones de redundancia síncronas o asíncronas.

5 En una realización, el servidor 102 puede enviar una consulta a los dispositivos 204-208 de almacenamiento de datos seleccionados para recopilar un conjunto consolidado de elementos de datos asociados con una persona o grupo de personas. El servidor 102 puede almacenar el conjunto de datos consolidado en un dispositivo 210 de almacenamiento de datos consolidado. En una tal realización, el servidor 102 puede referirse nuevamente al dispositivo 210 de almacenamiento de datos consolidado para obtener un conjunto de elementos de datos asociados con una persona o grupo específico. Como alternativa, el servidor 102 puede consultar cada uno de los dispositivos 204-208 de almacenamiento de datos de forma independiente o en una consulta distribuida para obtener el conjunto de elementos de datos asociados con una persona o grupo específico. En otra realización alternativa, se pueden almacenar múltiples bases de datos en un único dispositivo 210 de almacenamiento de datos consolidado.

10 En diversas realizaciones, el servidor 102 puede comunicarse con los dispositivos 204-210 de almacenamiento de datos a través del bus 202 de datos. El bus 202 de datos puede comprender una SAN, una LAN, o similar. La infraestructura de comunicación puede incluir Ethernet, bucle arbitrado de canal de fibra (FC-AL), interfaz de sistema informático pequeño (SCSI) y/u otros esquemas similares de comunicación de datos asociados con el almacenamiento y la comunicación de datos. Por ejemplo, el servidor 102 puede comunicarse indirectamente con los dispositivos 204-210 de almacenamiento de datos; el servidor puede comunicarse primero con un servidor de almacenamiento o controlador 106 de almacenamiento.

20 En un ejemplo del sistema 200, el primer dispositivo 204 de almacenamiento de datos puede almacenar datos asociados con datos clínicos que pueden estar comprendidos en reclamaciones de seguro hechas por una o más personas. Los datos clínicos pueden incluir datos asociados con servicios médicos, procedimientos y recetas utilizadas por la persona. En una realización, el segundo dispositivo 206 de almacenamiento de datos puede almacenar datos de diagnóstico asociados con la persona. Los datos de diagnóstico pueden incluir uno o más diagnósticos de afecciones que la persona padece o está en riesgo. El tercer dispositivo 208 de almacenamiento de datos puede almacenar datos de pruebas de laboratorio asociadas con la persona. Por ejemplo, el tercer dispositivo 208 de almacenamiento de datos puede incluir datos asociados con los resultados de las pruebas de laboratorio de la persona y/u observaciones clínicas. Un cuarto dispositivo de almacenamiento de datos (no mostrado) puede almacenar datos demográficos. Por ejemplo, los datos demográficos pueden incluir información relacionada con la demografía de la persona, incluyendo el género, raza o etnia, edad, ingresos, discapacidades, movilidad, logros educativos, propiedad de la vivienda, estado de empleo, ubicación o similar.

30 El servidor 102 puede alojar una aplicación de software configurada para una evaluación rápida de las distribuciones de valor de laboratorio. La aplicación de software puede incluir además módulos para interactuar con los dispositivos 204-210 de almacenamiento de datos, interconectando una red 108, interactuando con un usuario, y similares. En una realización adicional, el servidor 102 puede alojar un motor, complemento de aplicación o interfaz de programación de aplicaciones (API). En otra realización, el servidor 102 puede alojar un servicio web o una aplicación de software accesible desde la web.

35 La figura 3 ilustra un sistema informático 300 adaptado de acuerdo con ciertas realizaciones del servidor 102 y/o el dispositivo 110 de interfaz de usuario. La unidad 302 central de procesamiento (CPU) está acoplada al bus 304 del sistema. La CPU 302 puede ser una CPU de propósito general, un procesador o un microprocesador. Las presentes realizaciones no están restringidas por la arquitectura de la CPU 302, siempre que la CPU 302 admita los módulos y las operaciones como se describe en el presente documento. La CPU 302 puede ejecutar las diversas instrucciones lógicas de acuerdo con las presentes realizaciones. Por ejemplo, la CPU 302 puede ejecutar instrucciones a nivel de máquina de acuerdo con las operaciones ejemplares descritas a continuación con referencia a las figuras 7-9.

45 El sistema 300 informático también puede incluir una memoria 308 de acceso aleatorio (RAM), que puede ser SRAM, DRAM, SDRAM, o similar. El sistema 300 informático puede utilizar RAM 308 para almacenar las diversas estructuras de datos utilizadas por una aplicación de software adecuadamente programada para una evaluación rápida de las distribuciones de valores de laboratorio. El sistema 300 informático también puede incluir una memoria 306 de solo lectura (ROM) que puede ser PROM, EPROM, EEPROM, almacenamiento óptico, o similar. La ROM puede almacenar información de configuración para arrancar el sistema 300 informático. La RAM 308 y la ROM 306 contienen datos del usuario y del sistema 100.

50 El sistema 300 informático también puede incluir un adaptador 310 de entrada/salida (E/S), un adaptador 314 de comunicaciones, un adaptador 316 de interfaz de usuario y un adaptador 322 de pantalla. El adaptador 310 de E/S y/o el adaptador 316 de interfaz de usuario pueden, en ciertas realizaciones, permitir a un usuario interactuar con el sistema 300 informático para ingresar información para autenticificar a un usuario, identificar a una persona o grupo, recibir información de perfil de salud o ingresar información como un código médico, un código de prueba, un rango temporal, un percentil, un criterio limitante o un rango de valor de prueba seleccionado. En una realización adicional, el adaptador 322 de pantalla puede mostrar una interfaz gráfica de usuario asociada con un software o una aplicación basada en la web para generar una salida que comprende un gráfico de distribución.

60 El adaptador 310 de E/S puede conectarse a uno o más dispositivos 312 de almacenamiento, como uno o más de un disco duro, una unidad de disco compacto (CD), una disquetera, una unidad de cinta, al sistema 300 informático. El adaptador 314 de comunicaciones puede adaptarse para acoplar el sistema 300 informático a la red 106, que puede ser una o más de una LAN y/o WAN, y/o Internet. El adaptador 316 de interfaz de usuario puede acoplar

dispositivos de entrada de usuario, tal como un teclado 320 y un dispositivo 318 señalador, al sistema 300 informático. El adaptador 322 de pantalla puede ser accionado por la CPU 302 para controlar la pantalla en el dispositivo 324 de pantalla.

5 Las realizaciones actuales no se limitan a la arquitectura del sistema 300. Por el contrario, el sistema 300 informático se proporciona como un ejemplo de un tipo de dispositivo informático que puede adaptarse para realizar las funciones de un servidor 102 y/o el dispositivo 110 de interfaz de usuario. Por ejemplo, se puede utilizar cualquier dispositivo basado en procesador adecuado, incluyendo, sin limitación, asistentes de datos personales (PDA), consolas de videojuegos y servidores multiprocesador. Más aún, las presentes realizaciones pueden implementarse en circuitos integrados específicos de aplicación (ASIC) o circuitos integrados de muy gran escala (VLSI). De hecho, las personas con conocimientos ordinarios en la técnica pueden utilizar cualquier cantidad de estructuras adecuadas capaces de ejecutar operaciones lógicas de acuerdo con las realizaciones descritas.

10 La figura 4 ilustra una realización de un sistema 400 basado en red para la evaluación rápida de distribuciones de valores de laboratorio. En una realización, el sistema 400 basado en red incluye un servidor 102. De manera adicional, el sistema 400 basado en red puede incluir un dispositivo 110 de interfaz de usuario. En otra realización más, el sistema 400 basado en red puede incluir una o más aplicaciones 402 de cliente basadas en red configuradas para operar en una red 108 que incluye una intranet, Internet, o similares. En otra realización más, el sistema 400 basado en red puede incluir uno o más dispositivos 104 de almacenamiento de datos.

15 El sistema 400 basado en red puede incluir componentes o dispositivos configurados para operar en varias capas de red. Por ejemplo, el servidor 102 puede incluir módulos configurados para trabajar dentro de una capa 404 de aplicación, una capa 406 de presentación, una capa 408 de acceso a datos y una capa 410 de metadatos. En una realización adicional, el servidor 102 puede acceder a uno o más conjuntos 418-422 de datos que comprenden una capa de datos o nivel 412 de datos. Por ejemplo, un primer conjunto 418 de datos, un segundo conjunto 420 de datos y un tercer conjunto 422 de datos pueden comprender un nivel 412 de datos que se almacena en uno o más dispositivos 204-208 de almacenamiento de datos.

20 Una o más aplicaciones 412 web pueden operar en la capa 404 de aplicación. Por ejemplo, un usuario puede interactuar con la aplicación 412 web a través de una o más interfaces 318 y 320 de E/S configuradas para interactuar con la aplicación 412 web a través de un adaptador 310 de E/S que opera en la capa de aplicación. En una realización particular, se puede proporcionar una aplicación 412 web para una evaluación rápida de las distribuciones de valor de laboratorio que incluye módulos de software configurados para realizar las etapas de recibir un código médico, buscar en una base de datos almacenada en un dispositivo de almacenamiento de datos para obtener un primer grupo de registros asociados con personas que tienen el código médico, en el que cada registro del primer grupo de registros se identifica por un código de prueba, buscar en la base de datos para obtener un segundo grupo de registros asociados con una población de control, en el que cada registro del segundo grupo de registros se identifica mediante el código de prueba y genera una salida que comprende un gráfico de distribución en una escala compartida del primer y segundo grupo de registros.

25 En una realización adicional, el servidor 102 puede incluir componentes, dispositivos, módulos de hardware o módulos de software configurados para operar en la capa 406 de presentación para soportar uno o más servicios 414 web. Por ejemplo, una aplicación 412 web puede acceder o proporcionar acceso a un servicio 414 web para realizar una o más funciones basadas en la web para la aplicación 412 web. En una realización, una aplicación 412 web puede operar en un primer servidor 102 y acceder a uno o más servicios 414 web alojados en un segundo servidor (no mostrado) durante la operación.

30 Por ejemplo, una aplicación 412 web para identificar registros de diagnóstico y/o analizar datos de diagnóstico, u otra información puede acceder a un primer servicio 414 web para identificar registros de un primer grupo de personas asociadas con un código de diagnóstico y un segundo servicio 414 web para identificar una población de control como un grupo normal o un grupo de cohorte. Los servicios 414 web pueden recibir un código médico. En respuesta, el servicio 414 web puede devolver datos asociados con personas que tienen el código médico, estadísticas, distribuciones, gráficos, o similares. Un experto en la materia reconocerá varias arquitecturas basadas en web que emplean servicios 414 web para la operación modular de una aplicación 412 web.

35 En una realización, una aplicación 412 web o un servicio 414 web pueden acceder a uno o más de los conjuntos de datos 418-422 a través de la capa 408 de acceso a datos. En ciertas realizaciones, la capa 408 de acceso a datos puede dividirse en una o más capas 416 de acceso a datos independientes (DAL) para acceder a conjuntos 418-422 de datos individuales en el nivel 412 de datos. Estas capas 416 individuales de acceso a datos pueden denominarse tomas de datos o adaptadores. Las capas 416 de acceso a datos pueden utilizar metadatos de la capa 410 de metadatos para proporcionar a la aplicación 412 web o al servicio 414 web acceso específico al conjunto 412 de datos.

40 Por ejemplo, la capa 416 de acceso a datos puede incluir operaciones para realizar una consulta de los conjuntos 418-422 de datos para recuperar información específica para la aplicación 412 web o el servicio 414 web. En un ejemplo más específico, la capa 416 de acceso a datos puede incluir una consulta de registros asociados con personas que han sido diagnosticadas con una enfermedad particular, tal como diabetes o quienes están asociados

con un código médico (donde los valores de laboratorio pueden incluir un rango de interés), tal como un código ICD-9, asociado con un diagnóstico de diabetes.

La figura 5 ilustra una cierta realización de un sistema 500 para procesar y visualizar rápidamente distribuciones de valores de laboratorio. En una realización, el sistema 500 comprende un dispositivo 104 de almacenamiento de datos y un servidor 102 configurado para cargar y operar módulos 502-508 de software configurados para una evaluación rápida de las distribuciones de valor de laboratorio. Como alternativa, el sistema 500 puede incluir módulos 502-508 de hardware configurados con lógica analógica o digital, firmware que ejecuta FPGA o similar configurado para recibir un código médico, buscar en una base de datos almacenada en un dispositivo de almacenamiento de datos para obtener un primer grupo de registros asociados con personas que tienen el código médico, en el que cada registro del primer grupo se identifica por un código de prueba, buscar en la base de datos para obtener un segundo grupo de registros asociados con una población de control, en el que cada registro del segundo grupo se identifica mediante el código de prueba y genera una salida que comprende un gráfico de distribución en una escala compartida del primer y segundo grupo de registros. En tales realizaciones, el sistema 500 también puede incluir una interfaz 502, tal como un adaptador 310 de E/S, un adaptador 314 de comunicación, un adaptador 316 de interfaz de usuario, o similar.

En una realización, el servidor 102 puede incluir uno o más módulos definidos por software configurados para buscar un conjunto 418-422 de datos en un dispositivo 204-208 de almacenamiento de datos para obtener un primer grupo de registros asociados con personas que tienen el código médico, en el que cada registro del primer grupo se identifica por un código de prueba, buscar en la base de datos para obtener un segundo grupo de registros asociados con una población de control, en el que cada registro del segundo grupo se identifica mediante el código de prueba y genera una salida que comprende un gráfico de distribución en una escala compartida del primer y segundo grupo de registros. En una realización, estos módulos pueden incluir un módulo 502 de interfaz, un primer módulo 504 de búsqueda, un segundo módulo 506 de búsqueda y un módulo 510 de distribución. Opcionalmente, el servidor 102 también puede incluir un módulo 508 de descubrimiento como se describe con más detalle a continuación en relación con la figura 6.

Generalmente, el módulo 502 de interfaz puede recibir entradas de usuario y mostrar salidas de usuario. Por ejemplo, el módulo 502 de interfaz puede recibir un código médico. El módulo de interfaz puede recibir además uno o más códigos de prueba, intervalo temporal, criterio limitante, percentil, rangos de valores de prueba seleccionados y/u otras entradas de usuario. En una realización adicional, el módulo 502 de interfaz puede mostrar resultados de análisis de descubrimiento o distribución detallada. Dichos resultados de análisis pueden incluir estadísticas, tablas, diagramas, gráficos, recomendaciones y similares.

Estructuralmente, el módulo 502 de interfaz puede incluir uno o más de un adaptador 310 de E/S, un adaptador 314 de comunicaciones, un adaptador 316 de interfaz de usuario y/o un adaptador 322 de pantalla. El módulo 502 de interfaz puede incluir además puertos de E/S, clavijas, almohadillas, cableados, buses y similares para facilitar las comunicaciones entre la CPU 302 y los diversos adaptadores y componentes 310-324 de interfaz. El módulo de interfaz también puede incluir componentes definidos por software para interactuar con otros módulos de software en el servidor 102.

En una realización específica, el servidor 102 puede cargar y ejecutar software informático configurado para generar, recuperar, enviar u operar de otra manera instrucciones SQL. Por ejemplo, el primer módulo 504 de búsqueda puede comunicar una primera consulta SQL al dispositivo 104 de almacenamiento de datos, que está configurado para buscar en la base de datos un primer grupo de registros asociados con personas que tienen el código médico e identificados por el código de prueba.

El código médico puede incluir un código médico utilizado para agrupar e identificar enfermedades, trastornos, síntomas o signos médicos y almacenados en una base de datos de información de seguro de salud, un código médico indicativo de un riesgo de desarrollar una condición patológica o de enfermedad, o un código médico que identifica un tratamiento o procedimiento. En una realización específica, la primera operación de búsqueda puede identificar un primer grupo de personas que tienen registros que incluyen un código médico ICD-9 especificado. Por ejemplo, la primera búsqueda puede identificar un primer grupo de registros en la base de datos asociados con personas que han sido diagnosticados con diabetes. En otra realización, el código médico puede representar un riesgo, tal como un síndrome metabólico, que se caracteriza por un grupo de factores de riesgo metabólico o una combinación de trastornos médicos que aumentan el riesgo de desarrollar enfermedades cardiovasculares y diabetes. En una realización adicional, el código médico puede incluir una identificación o clasificación de personas después del tratamiento con uno o más medicamentos o procedimientos para el diagnóstico posterior al tratamiento o posterior al procedimiento.

En una realización adicional, el código de prueba puede ser un código o clasificación para identificar una prueba de laboratorio o una prueba clínica u observación, tal como un código LOINC® (identificadores de observaciones lógicas, nombres, códigos), un sistema de código universal para reportar observaciones de laboratorio y otras observaciones clínicas. Para cada observación, la base de datos puede incluir un código (de los cuales 25000 son observaciones de pruebas de laboratorio), un nombre formal largo, un nombre "corto" de 30 caracteres y sinónimos. Fue identificado por la Organización de desarrollo de estándares de salud de nivel siete (HL7) como un conjunto de

códigos preferido para los nombres de pruebas de laboratorio en transacciones entre centros de atención médica, laboratorios, dispositivos de prueba de laboratorio y autoridades de salud pública.

5 Un registro puede comprender múltiples valores de prueba de laboratorio identificados por un código de prueba común, en el que un promedio, primera lectura o última lectura, o similares de los valores múltiples se pueden usar para representar un punto de datos para este registro.

10 La primera o segunda búsqueda también puede implicar un componente temporal para agregar registros. Si un código de prueba tiene un aspecto de tiempo, las búsquedas pueden obtener múltiples registros asociados con la misma persona, pero identificada por un conjunto de códigos de prueba en un punto de tiempo diferente dentro de un intervalo de tiempo común. Esos registros pueden ser procesados, por ejemplo, promediados, para producir un valor de datos que representa el intervalo de tiempo asociado con la persona.

15 En ciertos aspectos, el primer módulo 504 de búsqueda puede configurarse para buscar en la base de datos un primer grupo de registros dentro de un intervalo temporal seleccionado, mientras coincide con el código de prueba y el código médico. De manera específica, el primer módulo 504 de búsqueda puede generar una primera consulta de búsqueda configurada para recuperar un primer grupo de registros dentro de un intervalo temporal seleccionado, mientras coincide con el código de prueba y el código médico. En una realización adicional, el primer grupo de registros puede incluir un componente temporal, que puede especificar un período de tiempo antes, durante o después de un evento, tal como un diagnóstico, una enfermedad o trastorno o inicio, un procedimiento o un tratamiento farmacológico. En un aspecto particular, el primer grupo de registros puede ser de un período limpio o un período libre de enfermedad antes de que las personas sean diagnosticadas con una enfermedad o trastorno, por ejemplo, dentro de aproximadamente 30, 60, 120 o 180 días antes de un diagnóstico de diabetes.

25 El segundo módulo 506 de búsqueda puede generar y/o comunicar una segunda consulta SQL a la base de datos en respuesta o independiente de los resultados de la primera consulta SQL. La segunda consulta puede configurarse para buscar en la base de datos un segundo grupo de registros, cada registro en el segundo grupo de registros identificado con el mismo código de prueba que identifica los registros en el primer grupo, pero asociado con una población de control. El segundo grupo de registros puede ser cohortes del primer grupo de registros, o registros de un grupo normal, o todos los registros identificados por el código de prueba en la base de datos, o registros filtrados por un criterio limitante que es compartido por el primer grupo de registros. Tal criterio limitante puede incluir un valor de campo que indica ciertas características específicas de las personas asociadas con los registros, tal como la edad, género, pruebas de laboratorio, resultados de laboratorio, otras enfermedades o diagnósticos, uso de medicamentos y similares. En cierta realización, el segundo grupo de registros también puede tener un componente temporal o estar asociado con un intervalo temporal, por ejemplo, cada registro en el segundo grupo puede estar asociado con personas que tienen al menos tres años de registro en la base de datos.

35 A modo de ejemplo, el primer módulo 504 de búsqueda puede identificar un primer grupo de personas que han sido diagnosticadas con diabetes o alguna otra enfermedad basado en un código médico ICD-9 asociado con dicho diagnóstico médico. El segundo módulo 506 de búsqueda puede entonces identificar cohortes para cada persona o registro identificado por la primera búsqueda. Las cohortes pueden compartir uno o más atributos comunes de segundo índice con las personas o registros identificados en el primer grupo, tal como la edad, género o similar; sin embargo, las cohortes no tendrían el código médico. En este ejemplo, las cohortes pueden tener la misma edad y sexo que las personas del primer grupo, pero no han sido diagnosticadas con diabetes.

40 En otra realización, el segundo módulo 506 de búsqueda puede generar un grupo de registros asociados con un grupo normal, tal como un grupo relativamente normal de personas no asociadas con el código médico utilizado en la primera búsqueda o un grupo normal de personas libres de enfermedad, un grupo maestro de registros que incluye valores de prueba asociados con cada prueba de laboratorio y unidad de prueba de laboratorio almacenados en la base de datos, o un grupo de registros filtrados asociados con un criterio limitante seleccionado, que se puede crear antes, durante o después de la primera búsqueda, o dinámicamente en tiempo de ejecución aplicando cualquier criterio limitante deseado, tal como edad o sexo.

50 En una realización adicional, el primer módulo 504 de búsqueda y el segundo módulo 506 de búsqueda pueden integrarse en un único módulo de búsqueda. De manera específica, se puede usar un solo conjunto de instrucciones SQL para identificar el primer grupo de registros e identificar el segundo grupo de registros. Los beneficios de esta realización pueden incluir una sobrecarga del sistema reducida, tiempo reducido de búsqueda y análisis, mano de obra reducida para la configuración y la generación de consultas, etc. Por ejemplo, con una única consulta SQL integrada, un usuario puede obtener resultados para el análisis en mucho menos tiempo de lo que el usuario podría esperar. Tal realización puede no requerir un análisis separado y la generación de consultas separadas para el primer grupo y el segundo grupo. En consecuencia, se puede lograr un ahorro de tiempo significativo.

55 En una realización, el módulo 508 de descubrimiento puede descubrir una o más pruebas de laboratorio relacionadas con el diagnóstico de enfermedad o la evaluación del tratamiento. Con el fin de hacerlo, el módulo 508 de descubrimiento puede extraer la diferencia entre valores representativos del primer y segundo grupo de registros identificados por el mismo código de prueba, pero asociados con diferentes estados de diagnóstico. Por ejemplo, el módulo de descubrimiento puede incluir lógica analógica o digital, firmware o software configurado para llevar a cabo

uno o más descubrimientos de acuerdo con una o más funciones lógicas predefinidas. En una realización adicional, el servidor 102 puede incluir un módulo 508 de descubrimiento definido por software configurado para realizar análisis y comparación de la información y los datos recuperados de la base de datos para el primer grupo de registros y el segundo grupo de registros asociado.

- 5 En una realización específica, el primer módulo 504 de búsqueda y el segundo módulo 506 de búsqueda pueden alimentar datos recuperados en una hoja de cálculo configurada para realizar uno o más cálculos sobre los datos. Por ejemplo, una hoja de cálculo Excel® puede incluir una o más funciones incrustadas u operaciones configuradas para calcular estadísticas tales como porcentajes, promedios, relaciones de posibilidades y otras probabilidades, recuentos, sumas y similares. Los datos pueden importarse automáticamente a una hoja de cálculo usando una
10 macro, un script basado en software o similar. En una realización alternativa, el módulo 508 de descubrimiento puede incluir funciones de software codificadas o dinámicamente variables para calcular tales estadísticas y generar resultados para un usuario. En una realización adicional, el módulo 508 de descubrimiento también puede crear salidas tales como estadísticas, tablas, diagramas, gráficos, recomendaciones y similares, e identificar particularmente uno o más códigos de prueba que indican una diferencia significativa entre dos grupos de registros
15 en función de un umbral o criterio de descubrimiento seleccionado. La "diferencia significativa" puede denominarse una diferencia estadísticamente significativa, tal como un nivel de significancia del 5 %, 1 % o 0,1 %.

La figura 6 ilustra otra realización de un sistema 500 para la evaluación rápida de distribuciones de valores de laboratorio. El sistema 500 puede incluir un servidor 102 como se describe en la figura 5. En una realización
20 adicional, el servidor 102 puede incluir módulos adicionales definidos por software. Por ejemplo, el servidor 102 puede incluir un módulo 602 de filtro. El módulo 508 de descubrimiento puede incluir además un módulo 602 de interpolación y un módulo 604 de comparación. El módulo 510 de distribución también puede incluir un módulo 606 de recuento, un módulo 608 de agregación, un módulo 610 de cálculo y un módulo 612 gráfico.

En una realización adicional, el módulo 602 de filtro puede reducir el primer grupo o el segundo grupo de registros de acuerdo con un criterio limitante. El módulo 602 de filtro puede reducir el primer o segundo grupo de registros al restringir los parámetros de búsqueda antes de realizar la primera o segunda búsqueda. Como alternativa, el módulo
25 602 de filtro puede filtrar, eliminar o borrar de otra manera los resultados de búsqueda de acuerdo con el criterio de limitación. En cierta realización, se pueden usar múltiples criterios limitantes para restringir el ámbito de los resultados de búsqueda devueltos. En una realización, un criterio limitante puede incluir un valor de campo, tal como la fecha de registro, edad, género o similar.

30 En una realización alternativa, un comando SQL generado por el servidor 102 o almacenado en la RAM 308 o en el dispositivo 312 de almacenamiento de datos puede incluir instrucciones, que cuando se ejecutan mediante un controlador 104 de almacenamiento o la CPU 302 en el servidor 102, pueden realizar una búsqueda de registros que coincidan con el criterio de limitación antes o en combinación con la primera o segunda búsqueda. En una realización específica, una operación SQL incrustada con la primera o segunda búsqueda puede activar el filtro.

35 En cierta realización, el módulo 508 de descubrimiento puede incluir un módulo 602 de interpolación configurado para derivar un valor de percentil del primer grupo o el segundo grupo de registros para comparación. El percentil puede ser cualquiera del 1 % al 99 %, tal como una mediana, 5 %, 32 % o 95 %. La función de interpolación puede implicar ajustar una función a los datos comprendidos en los registros y evaluar esa función en el punto deseado, o calcular los datos interpolados por procedimientos estadísticos. Por ejemplo, los procedimientos de interpolación
40 pueden incluir una interpolación lineal, una interpolación constante por partes, una interpolación polinómica, una interpolación spline, o una interpolación no lineal como una interpolación a través de procesos gaussianos. En una realización específica, el servidor 102 o el módulo 602 de interpolación pueden generar un comando SQL para calcular un valor de percentil a partir del primer o segundo grupo de registros. Por ejemplo, dicho comando SQL puede diseñarse de manera similar al ejemplificado en la figura 12 para la interpolación lineal de un grupo de puntos
45 de datos.

En una realización adicional, los valores representativos del percentil correspondiente de ambos grupos pueden procesarse mediante un módulo 604 de comparación. Por ejemplo, dicho módulo 604 de comparación puede generar una salida que comprende un gráfico que muestra la diferencia entre los dos valores correspondientes, como un cambio porcentual del primer valor asociado con el código médico desde el segundo valor asociado con un
50 grupo de control. Para un código médico particular, el módulo 604 de comparación puede ensamblar una pluralidad de códigos de prueba para mostrar el desplazamiento del grupo diagnosticado del grupo de control, como se ejemplifica en la figura 10. El grado relativo del cambio asociado con un código de prueba puede representar su relevancia con el código médico y su utilidad en el diagnóstico.

En una realización, el módulo 510 de distribución puede incluir un módulo 606 de recuento configurado para contar registros distintos en el primer grupo de registros y el segundo grupo de registros. La función de recuento puede implementarse usando un contador basado en hardware. Como alternativa, la función de recuento puede implementarse en un software. En una realización específica, el servidor 102 o la CPU 302 pueden ejecutar instrucciones SQL configuradas para proporcionar el recuento de registros en respuesta a resultados de búsqueda o consulta. En una tal realización, la función de recuento puede integrarse con las instrucciones de búsqueda y filtro en
55 un solo conjunto de comandos o instrucciones SQL.

En una realización adicional, el módulo 510 de distribución puede incluir un módulo 608 de agregación configurado para agregar registros del primer grupo y del segundo grupo de acuerdo con un atributo seleccionado, como una serie de rangos de depósito creados manualmente o derivados empíricamente. El servidor 102 también puede incluir otros módulos para cálculo, formateado y producción de estadísticas, incluyendo un módulo 610 de cálculo y un módulo 612 de salida. El módulo 610 de cálculo puede calcular una probabilidad en respuesta a una relación de una cantidad de registros en el primer grupo que tiene el atributo seleccionado y una cantidad de registros en el segundo grupo que tiene el atributo seleccionado. El módulo 612 de salida puede generar, formatear y proporcionar una representación gráfica de las estadísticas. Estos módulos 606-612 pueden ser módulos independientes implementados en hardware, firmware o software. Como alternativa, las funciones pueden realizarse a través de productos de cálculo comerciales u hojas de cálculo, instrucciones de software o SQL que están integradas con las otras funciones del servidor 102. En una realización específica, el módulo 510 de distribución, incluyendo algunos o todos sus módulos 606-612 de componentes, puede comunicar las estadísticas con el módulo 502 de interfaz para visualización o comunicación a un usuario.

Aunque las diversas funciones del servidor 102 y la CPU o el procesador 302 se describen en el contexto de los módulos, los procedimientos, procesos y el software descritos en el presente documento no se limitan a una estructura modular. En su lugar, algunas o todas las funciones descritas en relación con los módulos de las figuras 5-6 pueden implementarse en varios formatos, incluyendo, pero sin limitación, un solo conjunto de instrucciones integradas, órdenes, código, consultas, etc. En una realización, las funciones pueden implementarse en instrucciones de consulta de la base de datos, incluyendo SQL, PLSQL, o similares. Como alternativa, las funciones pueden implementarse en software codificado en C, C++, C#, php, Java o similar. En otra realización más, las funciones pueden implementarse en instrucciones basadas en web, incluyendo HTML, XML, etc.

Los diagramas de flujo esquemáticos que siguen se presentan generalmente como diagramas de flujo lógico. En consecuencia, el orden representado y las etapas etiquetadas son indicativas de una realización del procedimiento presentado. Pueden concebirse otras etapas y procedimientos que sean equivalentes en función, lógica, o efecto a una o más etapas, o porciones de las mismas, del procedimiento ilustrado. De manera adicional, el formato y los símbolos empleados se proporcionan para explicar las etapas lógicas del procedimiento y se entiende que no limitan el ámbito del procedimiento. Aunque se pueden emplear varios tipos de flecha y tipos de línea en los diagramas de flujo, se entiende que no limitan el ámbito del procedimiento correspondiente. De hecho, algunas flechas u otros conectores pueden usarse para indicar solo el flujo lógico del procedimiento. Por ejemplo, una flecha puede indicar un período de espera o monitorización de la duración no especificada entre las etapas enumeradas del procedimiento representado. De manera adicional, el orden en que se produce un procedimiento particular puede o no adherirse estrictamente al orden de las etapas correspondientes que se muestran.

La figura 7 ilustra una realización de un procedimiento 700 para la evaluación rápida de distribuciones de valores de laboratorio. En una realización, el procedimiento 700 comienza cuando el módulo 502 de interfaz recibe 702 un código médico. El código médico también se conoce como código de diagnóstico o atributo clínico de interés. El procedimiento 700 puede continuar cuando el servidor 102 emite un comando para buscar 704 una base de datos almacenada en el dispositivo 104 de almacenamiento de datos para un primer grupo de registros. El primer grupo de registros puede asociarse con las personas que tienen el código médico e identificarse mediante un código de prueba. Por ejemplo, el servidor 102 puede enviar una consulta SQL a la base de datos para recuperar los registros de atención médica asociados con personas que han sido diagnosticadas con diabetes como lo indica la presencia de un código ICD-9 asociado con el diagnóstico de diabetes en los registros de la persona, mientras que también almacena información sobre valores de pruebas de laboratorio identificados por uno o más códigos de prueba.

El servidor 102 puede entonces emitir un comando para buscar 706 la base de datos para un segundo grupo de registros. Cada registro en el segundo grupo de registros puede compartir un código de prueba común con un registro en el primer grupo de registros, pero está asociado con una población de control, la mayoría de los cuales no tiene el código médico. Por ejemplo, la consulta SQL emitida por el servidor 102 también puede incluir una declaración de consulta para buscar un segundo grupo de personas, cada persona tiene la misma edad y/o género que una persona identificada en el primer grupo, pero no tener el código médico de diabetes en su registro. Estas personas en el segundo grupo pueden considerarse cohortes de las personas en el primer grupo de registros. En otra realización, el segundo grupo de registros puede estar asociado con un grupo normal que puede incluir a todas las personas sanas o libres de enfermedad o un grupo que incluya a todas las personas con registros disponibles en la base de datos de todos los puntos de tiempo o un período de tiempo específico.

El servidor 102 puede entonces recibir los resultados de las búsquedas 704 y 706 en la base de datos. El módulo 510 de distribución puede generar 708 una salida que comprende un gráfico de distribución en una escala compartida del primer grupo de registros y el segundo grupo de registros. Por ejemplo, un programa de hoja de cálculo puede calcular la distribución de registros identificados por el código de prueba en presencia del código médico recibido en comparación con los de un grupo de control. La distribución puede incluir promedios, probabilidades y otros productos computacionales, incluyendo la identificación de tendencias y puntos en común entre los registros.

La figura 8 ilustra otra realización de un procedimiento 800 para la evaluación rápida de distribuciones de valores de laboratorio. En una realización, el procedimiento 800 comienza cuando el módulo 502 de interfaz recibe 802 un

código de prueba, un código médico, un percentil, y uno o más criterios limitantes o intervalos de tiempo. Por ejemplo, el módulo 502 de interfaz puede incluir una interfaz gráfica de usuario. El módulo 502 de interfaz puede recibir entradas de usuario que consisten en identificadores o indicadores de los códigos de prueba, códigos médicos o criterios limitantes. Dichos indicadores pueden incluir una selección de un valor de campo, tal como un valor de código ICD-9, un valor de edad, un valor de género o similar.

Los criterios limitantes pueden incluir valores de ventanas configurados para limitar o restringir los marcos de tiempo desde los cuales se buscarán los registros, restricciones en el tiempo mínimo de inscripción, número mínimo de registros, restricciones de género, restricciones de edad, restricciones de peso, condiciones físicas, patrones de salud, otros diagnósticos o valores de laboratorio identificados, y otros valores de umbral y límite similares. El módulo 602 de filtro puede incorporar el criterio de limitación en una consulta utilizada para filtrar los registros por el criterio de limitación anterior, durante o después de la búsqueda 806 o la búsqueda 808. Por ejemplo, la consulta puede buscar todos los registros asociados con personas que han sido diagnosticadas con diabetes, pero la consulta puede estar restringida para devolver solo resultados asociados con personas que tienen registros de al menos dos años en la base de datos. En una realización adicional, el módulo 602 de filtro puede usar el mismo criterio limitante para ambas búsquedas 806-808 para identificar una cohorte para el primer grupo de registros. El intervalo de tiempo específico relativo al diagnóstico asociado con el código médico también se puede utilizar para filtrar los registros 804 para evaluar las distribuciones de valores de laboratorio en el intervalo de tiempo específico.

Para el módulo 508 de descubrimiento, el módulo 602 de interpolación puede interpolar valores de percentil de ambos grupos de registros de acuerdo con un percentil recibido o seleccionado. Un ejemplo de interpolación se ilustra en la figura 12. Los registros de precios en la figura 12 pueden ser reemplazados por registros de atención médica. El servidor puede enviar dicha instrucción SQL basada en la figura 12 para interpolar diferentes valores de percentiles de los registros como se ejemplifica en la figura 10 para registros en presencia de una pluralidad de códigos de prueba (por ejemplo, código LOINC) en la primera columna en presencia de un código médico común, 292.4 que representa talasemia. En una realización adicional, el módulo 604 de comparación puede comparar los valores de percentiles correspondientes de ambos grupos y generar una salida que comprende un gráfico que muestra la comparación, tal como un cambio porcentual de la mediana del primer grupo en presencia de talasemia de la mediana del segundo grupo asociado con una población normal para múltiples códigos LOINC en la figura 10. El descubrimiento 810-814 puede ayudar a identificar o confirmar códigos de prueba que pueden ser potencialmente relevantes para un código médico en particular y pueden usarse para el diagnóstico o la predicción de enfermedades, así como la evaluación del tratamiento. Sin embargo, las etapas del descubrimiento de esos códigos de prueba pueden ser opcionales; por ejemplo, estas etapas pueden no ser necesarias para los códigos de prueba que se sabe que están asociados con el código médico.

En el módulo 510 de distribución, los registros también pueden recuperarse mediante la búsqueda 806 y 808, y opcionalmente el filtrado 804. En una realización, el módulo 606 de recuento puede contar registros distintos en ambos grupos para un número total de registros y también registros en cada categoría agregada ya que el módulo 608 de agregación puede agregar registros del primer grupo y del segundo grupo de acuerdo con un atributo seleccionado, tal como una serie de rangos agrupados de valores de prueba para distribución. Un ejemplo de agregación 906 se ilustra en la figura 11. El módulo 610 de cálculo puede calcular una relación de posibilidades, un promedio, o alguna otra estadística matemática y el módulo 612 de salida puede generar una salida que comprende uno o más gráficos de distribución para mostrar una o más estadísticas calculadas por el servidor 102. Por ejemplo, el gráfico puede ser un histograma como se ilustra en la figura 11.

En un ejemplo específico, el servidor 102 puede enviar un único conjunto de instrucciones SQL para realizar la primera búsqueda 806, realizar la segunda búsqueda 808 y otras funciones de búsqueda, tal como el filtrado 804.

En una realización adicional, las instrucciones SQL pueden incluir instrucciones para interpolar valores de percentil de ambos grupos. En otra realización adicional, las instrucciones SQL también pueden incluir comparar los valores de percentiles correspondientes y generar una salida que comprende un gráfico que muestra la comparación, por ejemplo, en el formato de cambio porcentual que también puede clasificarse según el grado de cambio. Por ejemplo, Una realización de las instrucciones SQL que pueden usarse para realizar el procedimiento de la figura 8 pueden incluir:

```

select c.loinc_code||' '||loinc_code_desc,a.*,b.*,-
(a.interpolated50th-
  b.interpolated50th)/(a.interpolated50th+b.interpolated50t
h)/2 percent_shift_of_median from
foo_loinc_distribution_sigmas a,
loinc c,
(
select loinc_code,result_units_name,u_mem,max(case when .50
between decile and lead_decile then (lead_delta-
delta_from_index)/(lead_decile-decile) *0.50
+(delta_from_index-(lead_deltadelta-
from_index)/(lead_decile-decile)*decile) else -
99999 end ) interpolated50th, max(case when .32 between
decile and lead_decile then (lead_delta-

```

ES 2 742 826 T3

```

delta_from_index)/(lead_decile-decile)*0.32 +(delta_from_index-
(lead_delta- delta_from_index)/(lead_deciledecile)*
decile) else -99999 end) interpolated32th, max(case
when .95 between decile and lead_decile then (lead_delta-
5   delta_from_index)/(lead_decile-decile)*0.95
+(delta_from_index-(lead_delta-
delta_from_index)/(lead_decile-decile)*decile) else -
99999 end ) interpolated95th,max(case when .05 between
10  decile and lead_decile then (lead_delta-
delta_from_index)/(lead_decile-decile)*0.05
+(delta_from_index-(lead_delta-
delta_from_index)/(lead_decile-decile)*decile) else -
15  99999 end ) interpolated5th, max(case when .68 between decile
and lead_decile then (lead_delta-
delta_from_index)/(lead_decile-decile)*0.68
+(delta_from_index-(lead_delta-
delta_from_index)/(lead_decile-decile)*decile) else -
20  99999 end ) interpolated68th
from
(
select iv4.*, cast(result_value_nbr as float)
delta_from_index,cast(rn as float)/u_mem decile,lead(cast(rn
as float)/u_mem) over (partition by
25  loinc_code,result_units_name order by rn)
lead_decile,lead(result_value_nbr) over (partition by
loinc_code,result_units_name order by rn) lead_delta
from
(
select iv3.*, max(rn) over( partition by
30  loinc_code,result_units_name ) u_mem
from
(
select iv2.*, row_number() over ( partition by
35  loinc_code,result_units_name order by result_value_nbr asc)
rn
from
(
select
40  a.loinc_code,a.result_units_name,a.individual_id,avg(result_val
ue_nbr) result_value_nbr
from
(
select individual_id,min(service_from_date) min_dos from
45  diagnosis a, foo_members_with_condition6 b where
decn_code like'250%' and b.dx=a.diagnosis_key group by
individual_id
) iv,
foo_members_with_lab6 a
where a.individual_id=iv.individual_id and service_from_date
50  between min_dos+-1400 and min_dos+-365
group by a.loinc_code,a.result_units_name,a.individual_id
) iv2
) iv3
) iv4 where u_mem>19
55  ) iv5 group by loinc_code,result_units_name,u_mem
) b where a.loinc_code=b.loinc_code and
a.result_units_name=b.result_units_name and a.loinc_code not
in ('0','1') and a.result_units_name <> ' '
and a.loinc_code=c.loinc_code and ( b.interpolated50th not
60  between a.interpolated32th and a.interpolated68th or
a.interpolated50th not between b.interpolated32th and
b.interpolated68th)
order by abs((a.interpolated50thb.
65  interpolated50th)/(a.interpolated50th+b.interpolated50th)/2)
desc

```

En una realización adicional, las instrucciones SQL pueden incluir instrucciones para contar 904 registros distintos y/o agregar 906 registros. En otra realización más, el mismo conjunto de instrucciones SQL puede incluir funciones para generar 908 una salida que comprende un gráfico de distribución de probabilidad calculando estadísticas, como la razón de posibilidades de un valor de prueba específico o rango dentro de los registros. Una realización particular de las instrucciones SQL que pueden usarse para realizar el procedimiento de la figura 9 puede incluir:

```

select iv4.*,u_mems/tot1 percent_all,members_with_disease/tot2
percent_with_disease from (
select iv3.*,sum(u_mems) over ( ) tot1,sum(members_with_disease)

```

```

over () tot2 from (
select
b.loinc_code,b.result_units_name,rng,u_mems,rn,count(distinct
individual_id) members_with_disease from
5 (select * from foo_normal_distributions where loinc_code='785-
6' and result_units_name='pg' ) b
left outer join
(
10 select
a.individual_id,min_dos,decm_code,loinc_code,result_units_name,
avg(result_value_nbr) result_value_nbr from
(
15 SELECT decm_code, b.individual_id, MIN(service_from_date)
min_dos
FROM diagnosis a, foo_members_with_condition6 b
WHERE decm_code like '282.4'
AND b.dx=a.diagnosis_key
GROUP BY decm_code, b.individual_id
) iv2
20 ,
foo_members_with_lab6 a
where a.individual_id=iv2.individual_id and service_from_date
between min_dos+-180 and min_dos+-2
group by
25 a.individual_id,min_dos,decm_code,loinc_code,result_units_name
) iv3 on ( iv3.loinc_code=b.loinc_code and
iv3.result_units_name=b.result_units_name and
result_value_nbr >bucketstart and result_value_nbr<=
30 bucketend)
group by b.loinc_code,b.result_units_name,rng,u_mems,rn
) iv3
) iv4 order by rn

```

La figura 10 ilustra una realización de una salida 1000 del módulo 508 de descubrimiento, que puede incluir una tabla 1002 de interpolación o un gráfico 1004 de comparación. En cierta realización, cada fila en la tabla 1002 puede representar un primer grupo de registros en presencia de un código 282.4 médico de talasemia que se identifica mediante un código de prueba específico en la primera columna. Por ejemplo, el código de prueba puede ser el código LOINC 785-6 que representa la hemoglobina corpuscular media de los eritrocitos, y el primer grupo de registros asociado (no mostrado) puede interpolarse para obtener valores que representan el 50 % (mediana), 32 %, 95 %, y 5 %. En una realización particular, la mediana del primer grupo de registros identificados por el código de prueba específico en presencia del código médico se compara con la mediana de un grupo de registros de control, en particular, un grupo normal, que es la línea media correspondiente al 0,0 % en el gráfico 1004. La diferencia entre las dos medianas para cada código de prueba se representa como un cambio porcentual (= (mediana del grupo de registros enfermos-mediana del grupo normal de registros)/mediana del grupo normal de registros x 100 %) desde la línea media, que también se ha utilizado para clasificar el orden de la lista de códigos de prueba de menor a mayor en términos de valor absoluto de diferencia.

La figura 11 ilustra una realización de resultados o resultados que pueden derivarse del módulo 510 de distribución. Un código de prueba específico, por ejemplo, código LOINC 785-6, se recibe o se selecciona para la primera búsqueda 806 para identificar un primer grupo de registros asociados con el código de prueba y en presencia del código 282.4 médico. Una tabla 1102 puede incluir uno o más campos para mostrar el recuento, agregando y calculando resultados. El primer y segundo grupo de registros se pueden agregar de acuerdo con los atributos seleccionados, en particular, un conjunto de rangos de valores de prueba (campo RNG), y contados en cada rango con resultados presentados en el campo U_MEM para el segundo grupo (control normal) y en MIEMBROS_CON_ENFERMEDAD para el primer grupo (enfermos). El porcentaje del número de registros en cada rango de valores de prueba sobre el número total de registros en el primer o segundo grupo también se puede calcular para obtener resultados que se muestran en los campos PORCENTAJE_ENFERMOS o PORCENTAJE_TODOS, respectivamente. Los gráficos 1104 y 1106 muestran una distribución de números de registros en los rangos de cubos para el grupo normal y el grupo enfermo, respectivamente. El histograma 1108 superpone los porcentajes de registros en cada rango de valores para ambos grupos y muestra una diferencia en su patrón de distribución.

La figura 12 ilustra un ejemplo para la interpolación de un valor de percentil de un grupo de registros. El protocolo 1202 explica la interpolación de una fuente bruta de precios aleatorios mediante el uso de una función lineal para extraer un valor al 80 % mediante instrucciones SQL. La tabla 1204 presenta una fuente tan cruda de precios aleatorios y la tabla 1206 muestra el resultado de la interpolación del 80 %.

La figura 13 ilustra una realización del módulo de descubrimiento utilizado para el diagnóstico previo a la enfermedad o para predecir el riesgo de una enfermedad. La lista 1302 proporciona ejemplos de factores de riesgo para identificar candidatos a síndrome metabólico, uno o más de los cuales pueden usarse como código médico para

5 buscar el primer grupo de registros en la base de datos. Como alternativa, se puede usar una ventana temporal en combinación con el código médico para diabéticos para buscar un primer grupo de registros recopilados tres años a treinta días antes del diagnóstico de diabetes o en presencia de un código médico para diabéticos. El gráfico 1304 muestra la salida que comprende el cambio porcentual clasificado de las medianas del primer grupo de registros de las medianas normales para una pluralidad de códigos de prueba. Un código de prueba con un cambio porcentual significativo puede ser un posible marcador prediabético o un predictor temprano de diabetes útil, como el código LOINC 2086-7, asociado con un cambio de casi -12 %.

10 La figura 14 ilustra una realización del módulo de distribución utilizado para el diagnóstico previo a la enfermedad o para predecir el riesgo de una enfermedad. La distribución del porcentaje de registros en cada rango de histograma para el grupo preenfermedad (uno a tres años antes del diagnóstico de diabetes) y el grupo control se comparan en el gráfico 1400.

15 La figura 15 ilustra una realización de una interfaz 1500 gráfica de usuario (GUI) para su uso de acuerdo con el módulo 502 de interfaz. Como se ilustra, la GUI 1500 puede incluir uno o más campos para recibir entradas del usuario, uno o más botones para emitir un comando o controlar el servidor, aplicación o código, y uno o más nombres o elementos descriptivos o informativos para proporcionar instrucciones e indicaciones para el usuario. Como se ilustra en el diagrama informativo en la figura 15, ciertos criterios limitantes o intervalos temporales pueden incluir un valor de período limpio o un período de tiempo anterior a la aparición del código médico. Por ejemplo, si el código médico es un código médico para la diabetes, los criterios del período de limpieza pueden ser un marco de tiempo, tal como dos años, antes de la aparición del código médico. En una tal realización, el primer módulo 504 de búsqueda solo puede devolver registros de personas que no solo han sido diagnosticadas con diabetes, pero también tiene un mínimo de dos años de registros antes del diagnóstico.

20 Los criterios limitantes también pueden incluir una ventana de agregación o un marco de tiempo para recopilar registros asociados con la persona. La variable de ventana de agregación puede incluir un marco de tiempo antes o después de la aparición del código médico desde el cual recopilar registros. Por ejemplo, el módulo 602 de filtro solo puede recopilar registros durante dos años antes de la aparición del código médico y durante noventa días después de la aparición del código médico.

25 Todos los procedimientos desvelados y reivindicados en el presente documento pueden realizarse y ejecutarse sin experimentación indebida a la vista de la presente divulgación. Aunque los sistemas y procedimientos de la presente invención se han descrito en términos de realizaciones preferidas, Será evidente para los expertos en la materia que se pueden aplicar variaciones a los procedimientos y en las etapas o en la secuencia de etapas del procedimiento descrito en el presente documento sin apartarse del concepto y del ámbito de la invención. Además, se pueden hacer modificaciones al aparato desvelado y los componentes se pueden eliminar o sustituir por los componentes descritos en el presente documento donde se obtendrían los mismos resultados o similares. Se considera que todos los sustitutos y modificaciones similares evidentes para los expertos en la materia están dentro del ámbito y el concepto de la invención tal como se define en las reivindicaciones adjuntas.

REIVINDICACIONES

1. Un procedimiento (700) implementado por ordenador para comparar dos grupos de registros para identificar pruebas de laboratorio para su uso en diagnóstico, predicción o evaluación de tratamiento de una enfermedad, comprendiendo el procedimiento:
- 5 recibir (702) un código médico asociado con un diagnóstico de la enfermedad;
 buscar (704) en una base de datos almacenada en un dispositivo de almacenamiento de datos para obtener un primer grupo de registros asociados con personas que tienen el código médico, en el que cada registro del primer grupo de registros comprende uno o más primeros valores de prueba de laboratorio, en el que cada primer valor de prueba de laboratorio se identifica por un código de prueba de uno o más códigos de prueba;
- 10 buscar (706) en la base de datos para obtener un segundo grupo de registros asociados con una población de control, en el que cada registro del segundo grupo de registros comprende uno o más segundos valores de prueba de laboratorio, en el que cada segundo valor de prueba de laboratorio se identifica por un código de prueba respectivo del uno o más códigos de prueba;
 derivar valores de percentiles del primer grupo y el segundo grupo de registros para comparación;
- 15 comparar valores representativos de los valores de los percentiles correspondientes tanto del primer grupo como del segundo grupo y determinar, desde un grado relativo de cambio entre los dos valores correspondientes, una relevancia de cada código de prueba con el código médico;
 identificar uno o más códigos de prueba en los que la diferencia relativa indica una diferencia estadísticamente significativa entre los dos grupos de registros en función de un umbral o criterio de descubrimiento seleccionado;
- 20 y
 generar (708) una salida que comprende un gráfico de distribución en una escala compartida del primer y segundo grupo de registros, en el que la salida representa la distribución de registros identificados por el uno o más códigos de prueba identificados en presencia del código médico recibido en comparación con los de la población de control.
- 25 2. Un producto de programa de ordenador tangible que comprende un medio legible por ordenador que tiene un código de programa utilizable por ordenador ejecutable para realizar el procedimiento de la reivindicación 1.
3. El procedimiento (700) de la reivindicación 1, en el que la etapa de derivación comprende interpolar cada uno de los uno o más primeros valores de prueba de laboratorio del primer grupo de registros para obtener un primer valor de percentil e interpolar cada uno de los uno o más segundos valores de prueba de laboratorio del segundo grupo de registros para obtener un valor del segundo valor de percentil, en el que el primer y segundo valores de percentil están asociados con un valor de percentil compartido.
- 30 4. El procedimiento (700) de la reivindicación 1, en el que generar la salida comprende además contar registros distintos en el primer grupo de registros y el segundo grupo de registros.
5. El procedimiento (700) de la reivindicación 1, en el que generar la salida comprende además agregar registros del primer grupo y del segundo grupo de acuerdo con un atributo seleccionado.
- 35 6. El procedimiento (700) de la reivindicación 5, en el que la generación de la salida también comprende calcular una probabilidad en respuesta a una relación de un número de registros en el primer grupo que tiene el atributo seleccionado y un número de registros en el segundo grupo que tiene el atributo seleccionado.
7. El procedimiento (700) de la reivindicación 1, que comprende además seleccionar el primer grupo de registros dentro de un intervalo de tiempo.
- 40 8. El procedimiento (700) de la reivindicación 1, en el que el registro comprende un promedio, una primera lectura o una última lectura de una prueba.
9. El procedimiento (700) de la reivindicación 1, que comprende además seleccionar registros en el primer grupo y en el segundo grupo de acuerdo con un criterio limitante.
- 45 10. Un sistema (100) que comprende un dispositivo de almacenamiento de datos y medios de servidor configurados para realizar el procedimiento de la reivindicación 1.
11. El sistema (100) de la reivindicación 10, en el que los medios están configurados adicionalmente para interpolar cada uno de los uno o más primeros valores de prueba de laboratorio del primer grupo de registros para obtener un primer valor de percentil e interpolar cada uno de los uno o más segundos valores de prueba de laboratorio del segundo grupo de registros para obtener un segundo valor de percentil, en el que el primer y segundo valores de percentil están asociados con un valor de percentil compartido.
- 50 12. El sistema (100) de la reivindicación 10, en el que los medios están configurados adicionalmente para contar registros distintos en el primer grupo de registros y el segundo grupo de registros.
13. El sistema (100) de la reivindicación 10, en el que los medios de servidor están configurados además para

agregar registros del primer grupo y del segundo grupo de acuerdo con un atributo seleccionado.

14. El sistema (100) de la reivindicación 13, en el que los medios están configurados además para calcular una probabilidad en respuesta a una relación de un número de registros en el primer grupo que tiene el atributo seleccionado y un número de registros en el segundo grupo que tiene el atributo seleccionado.

5 15. El sistema (100) de la reivindicación 10, en el que los medios están configurados además para seleccionar el primer grupo de registros dentro de un intervalo de tiempo.

16. El sistema (100) de la reivindicación 10, en el que los medios están configurados adicionalmente para seleccionar registros en el primer grupo y en el segundo grupo de acuerdo con un criterio limitante.

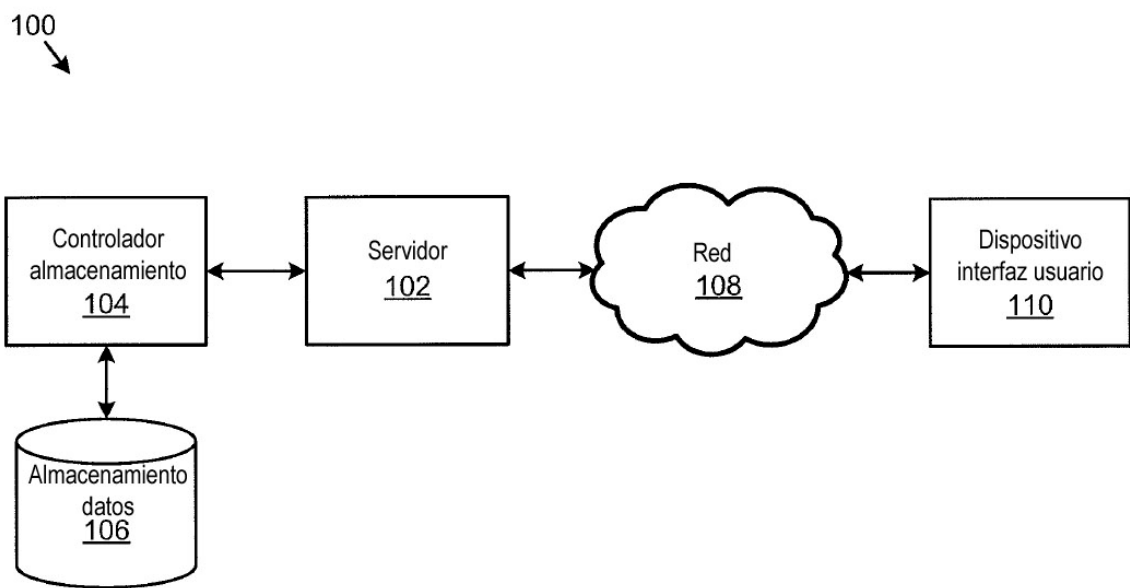


FIG. 1

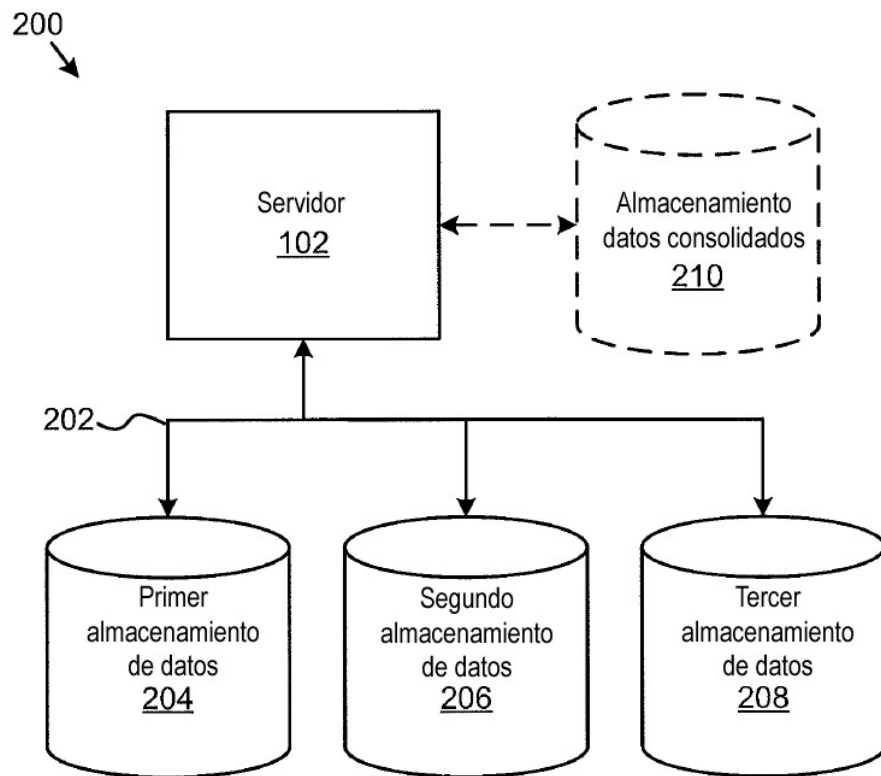


FIG. 2

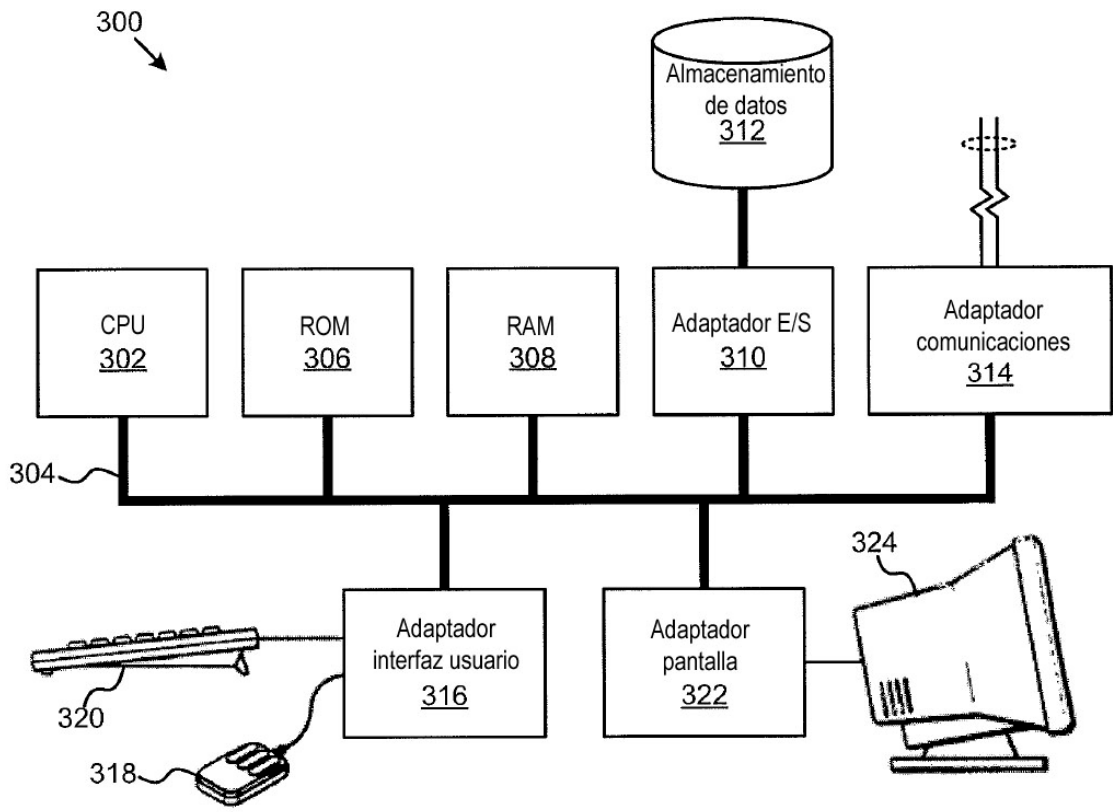


FIG. 3

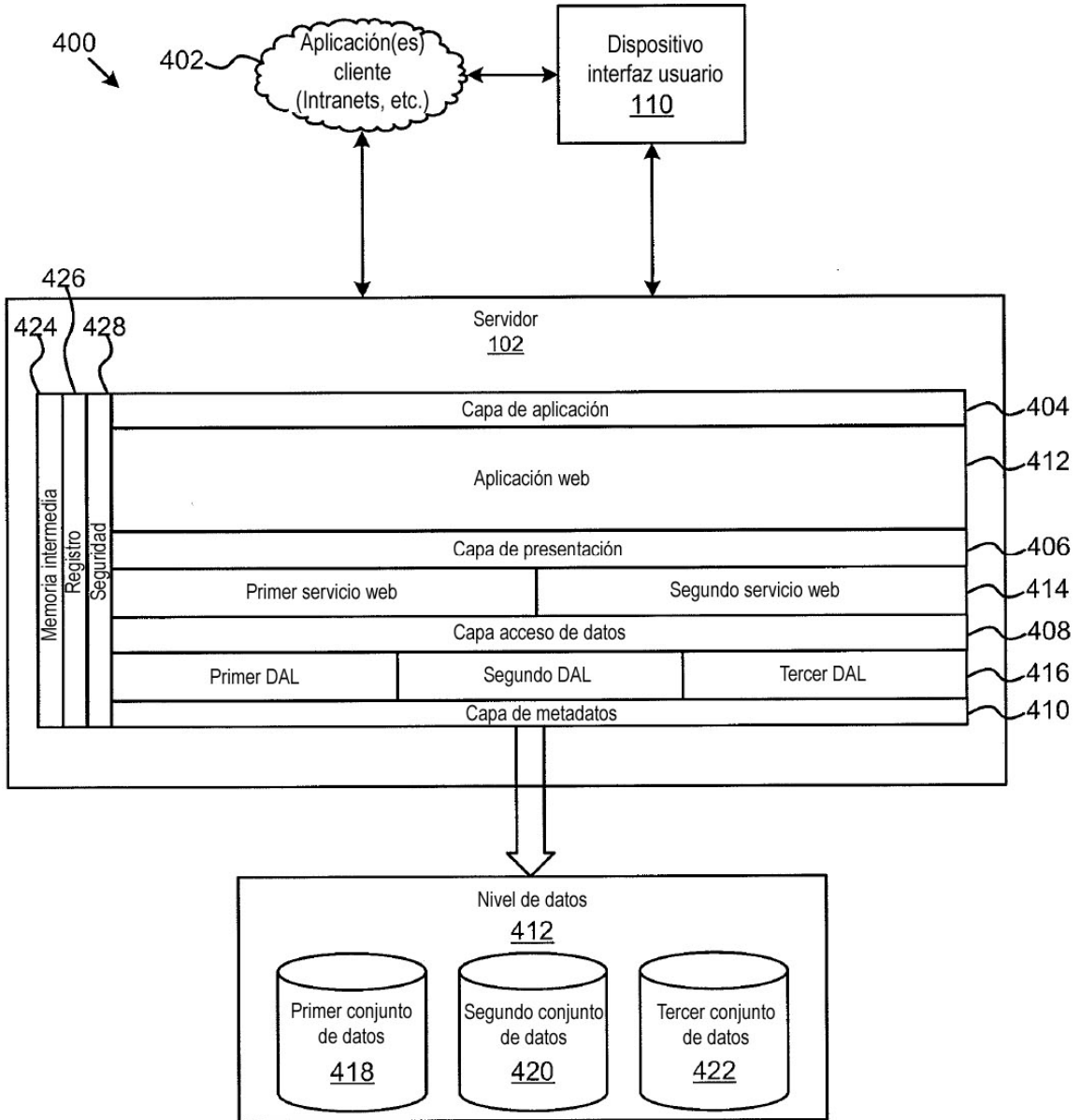


FIG. 4

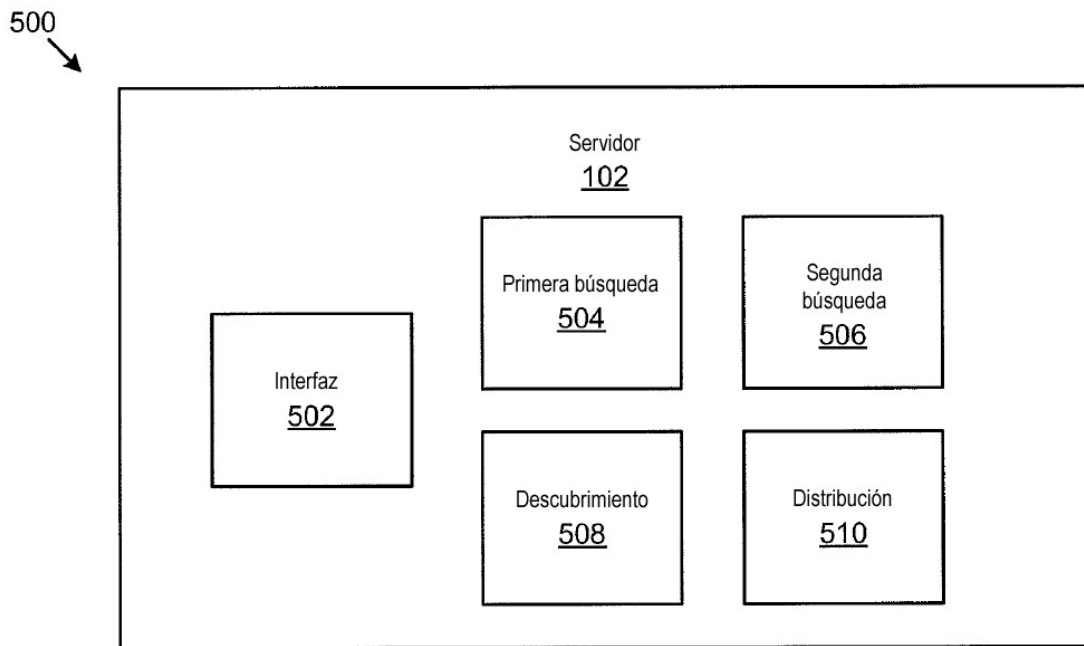


FIG. 5

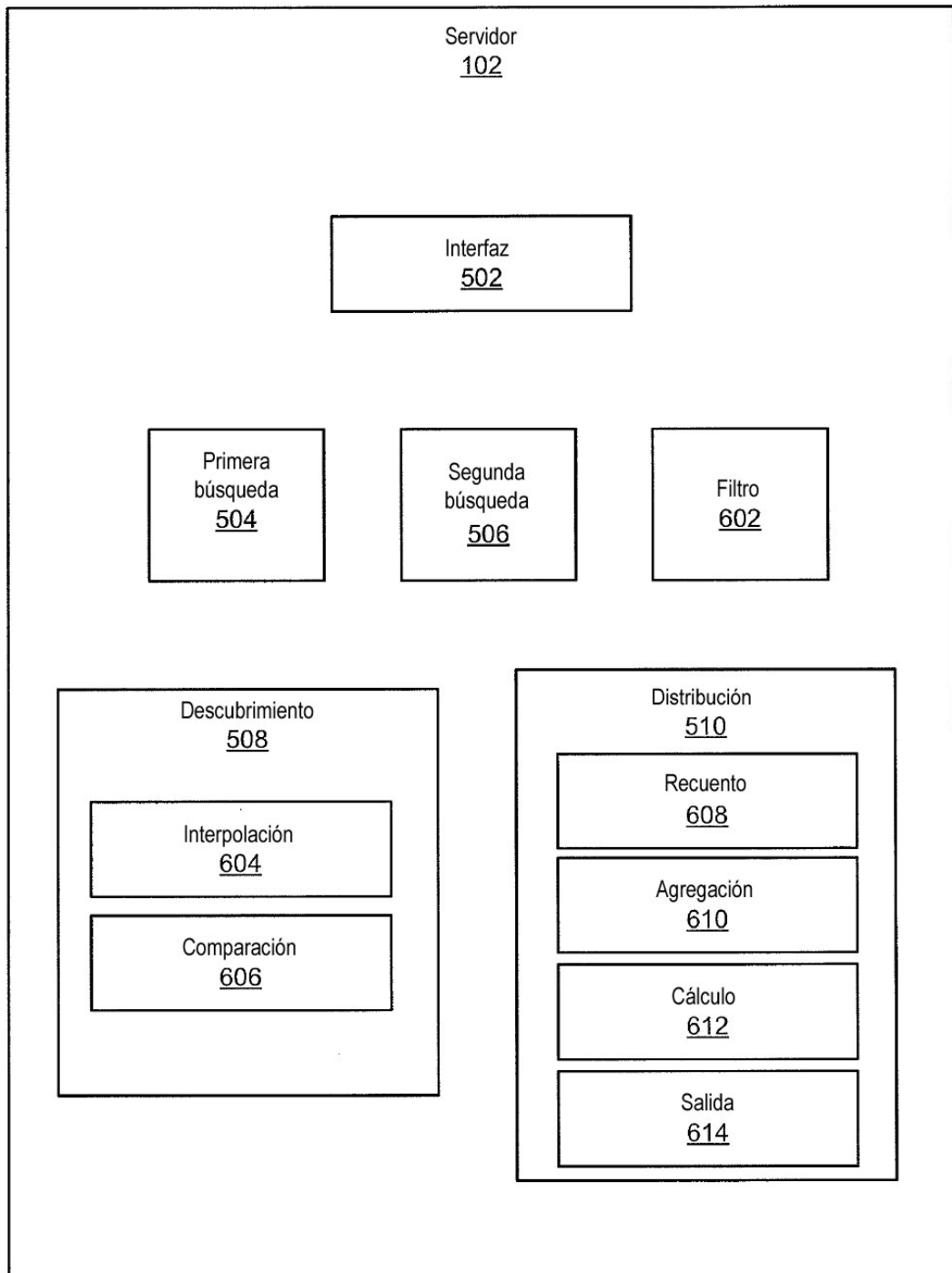


FIG. 6

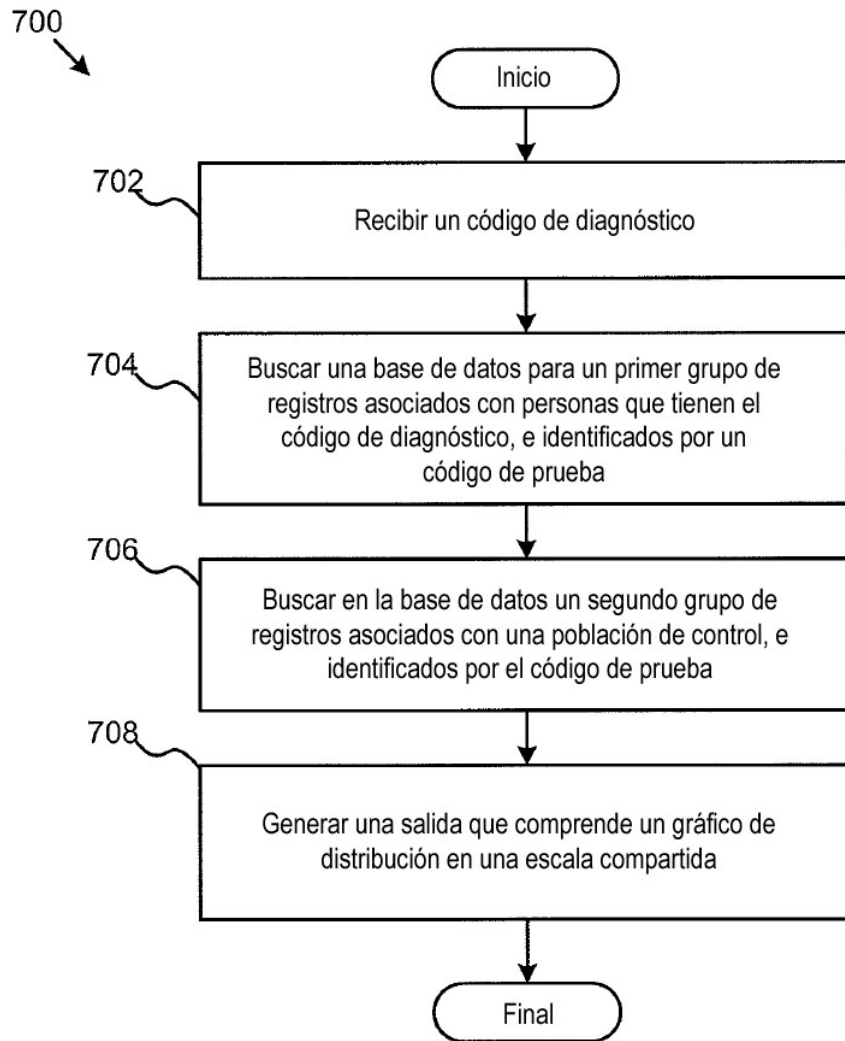


FIG. 7

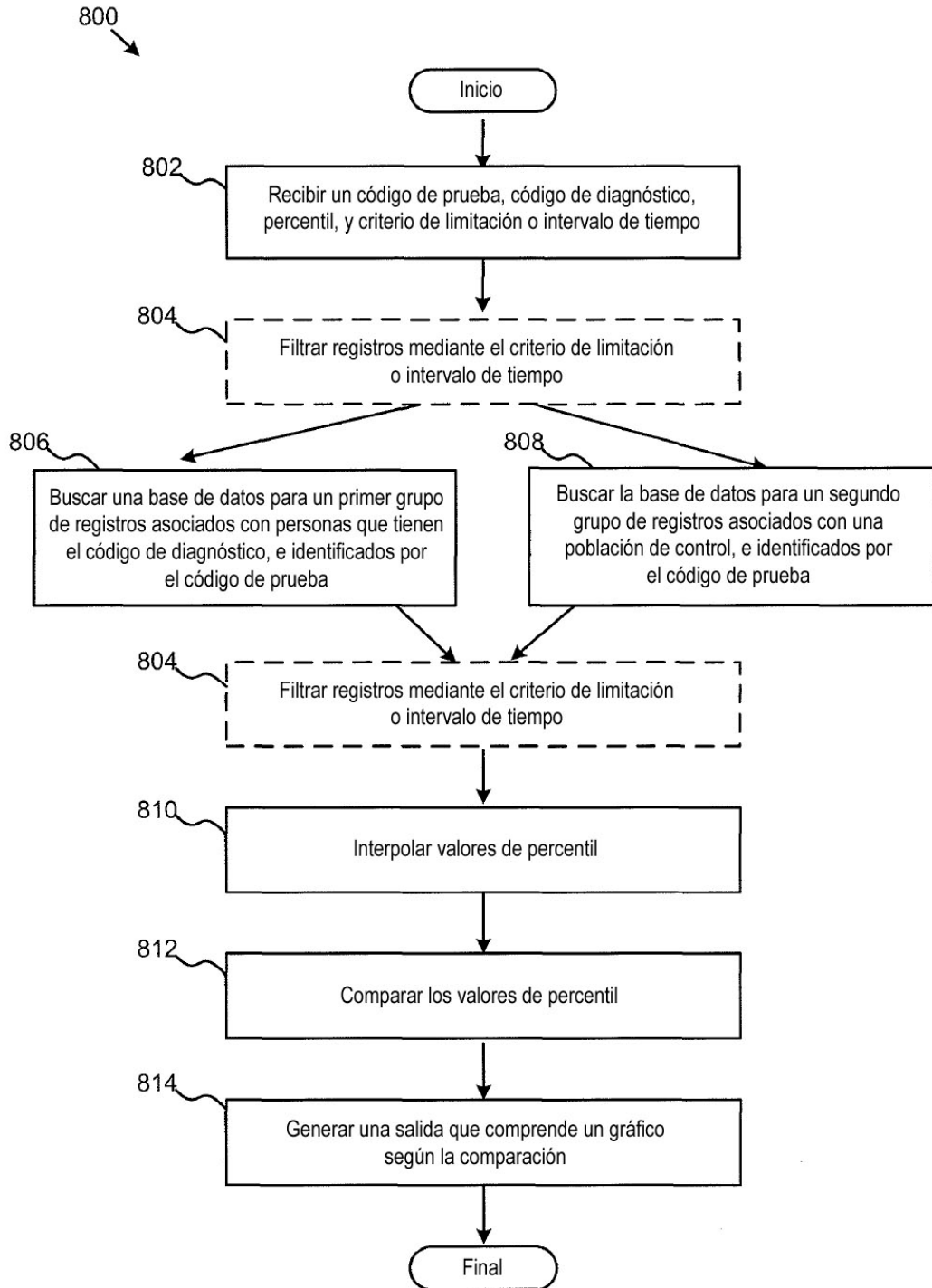


FIG. 8

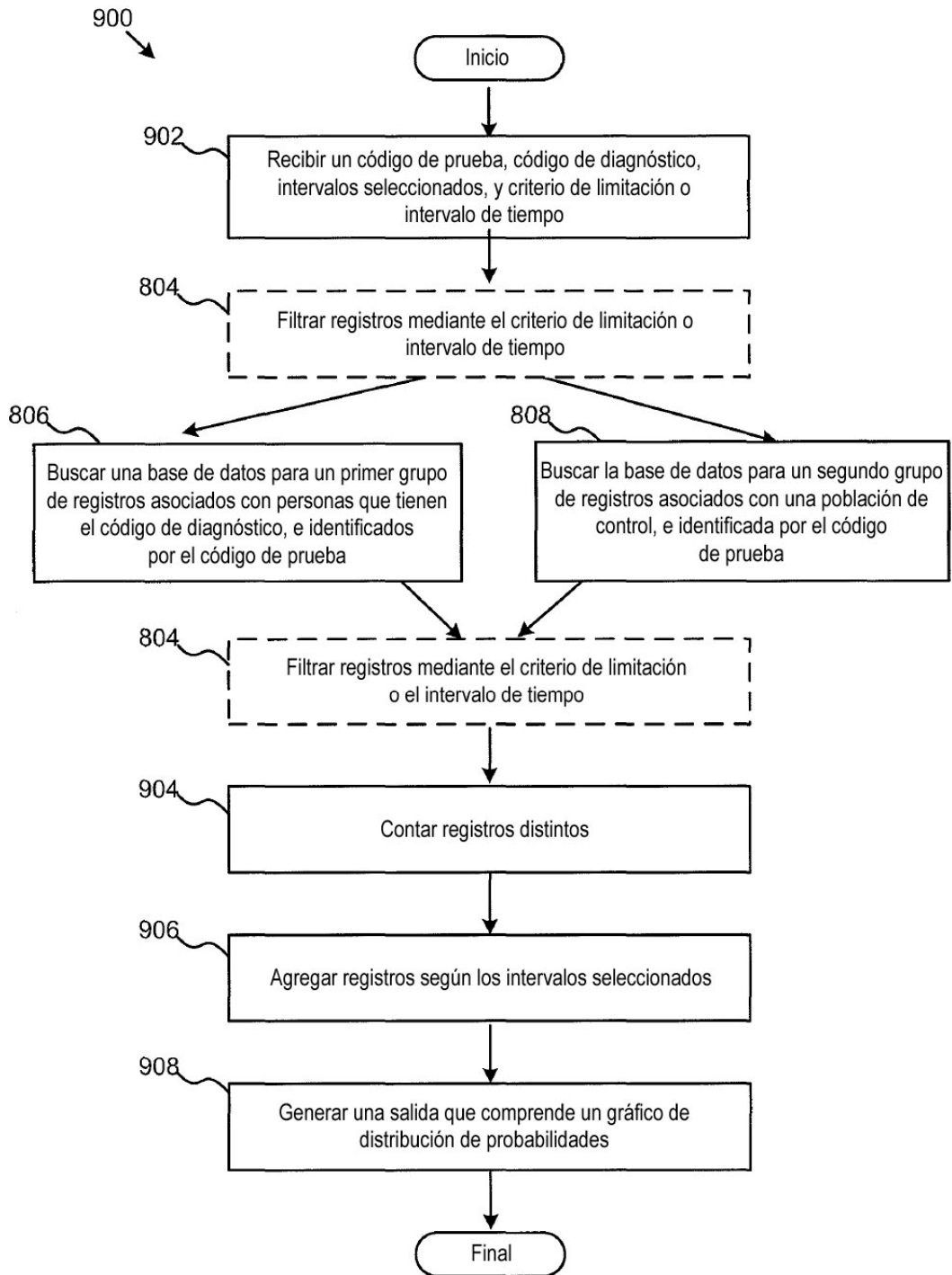


FIG. 9

1000

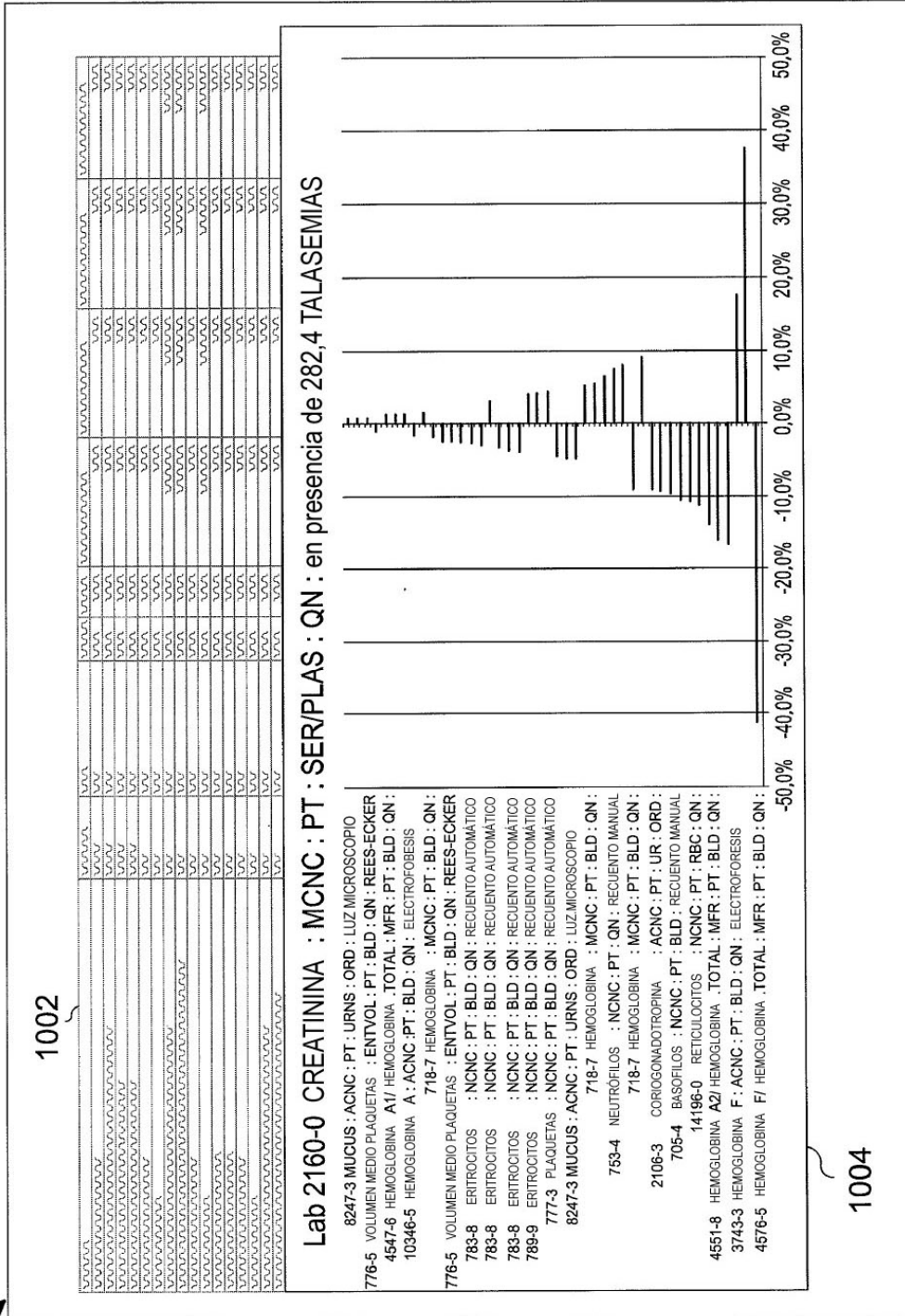


FIG. 10

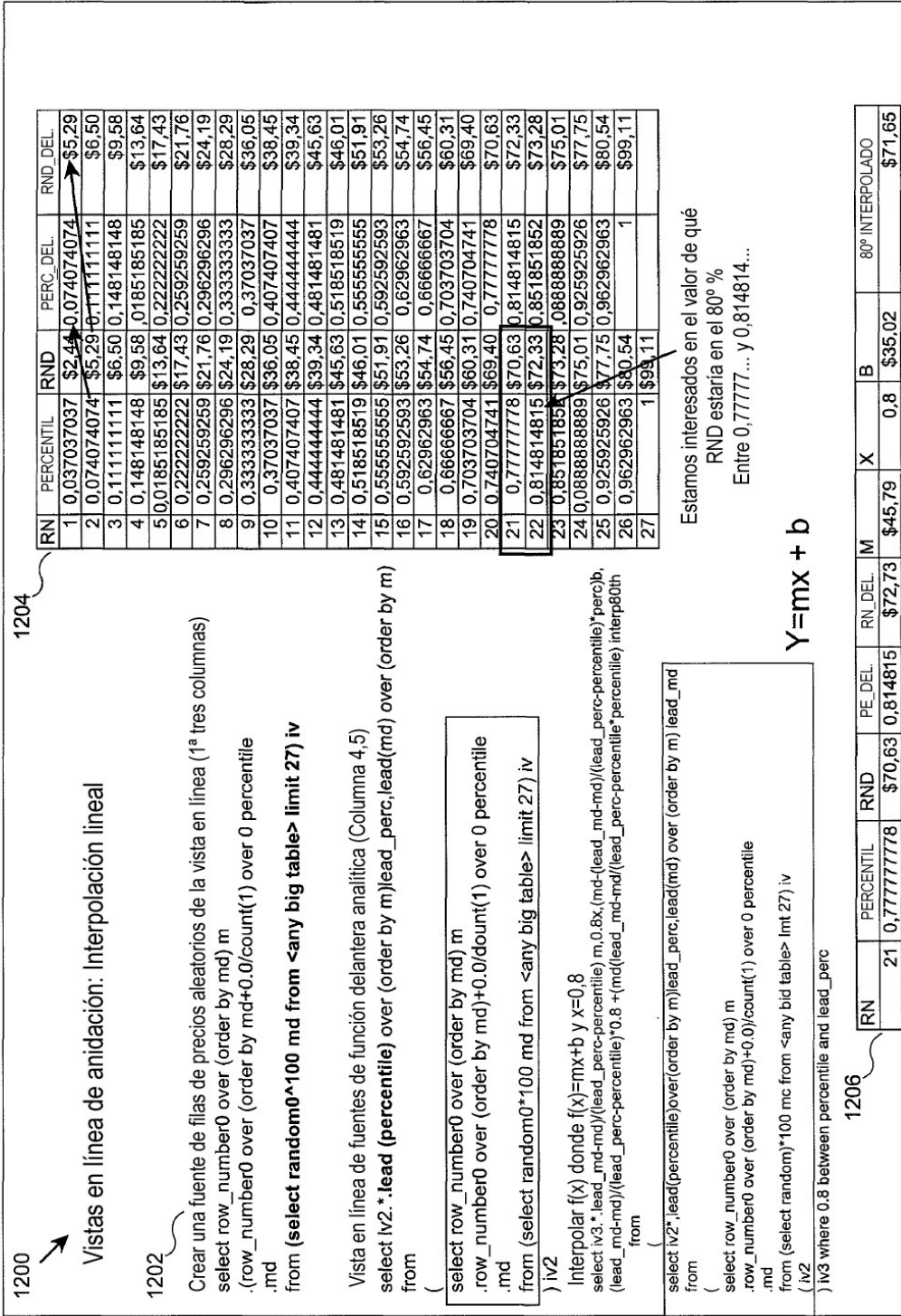


FIG. 12

Intervalos de valores de laboratorio prediabéticos como se ven en datos de cuidado de salud unidos para candidatos de síndrome metabólico de 3 años a 30 días antes de diagnóstico diabético

1300



1302

* Elevada circunferencia de cintura:

- * Hombres -- Igual o mayor que 40 pulgadas (102 cm)
- * Mujeres -- Igual o mayor que 35 pulgadas (88 cm)

* Triglicéridos elevados: Igual o mayor que 150 mg/dL

- * Colesterol HDL ("bueno") reducido
- * Hombres -- Menos de 50 mg/dL
- * Mujeres: Menos de 50 mg/dL

* Presión sanguínea elevada: Igual o mayor que 130/85 mm Hg o uso de medicación para hipertensión

* Glucosa rápida elevada: Igual o mayor que 100 mg/dL (5,6 mmol/L) o uso de medicación para hiperglicemia

1304

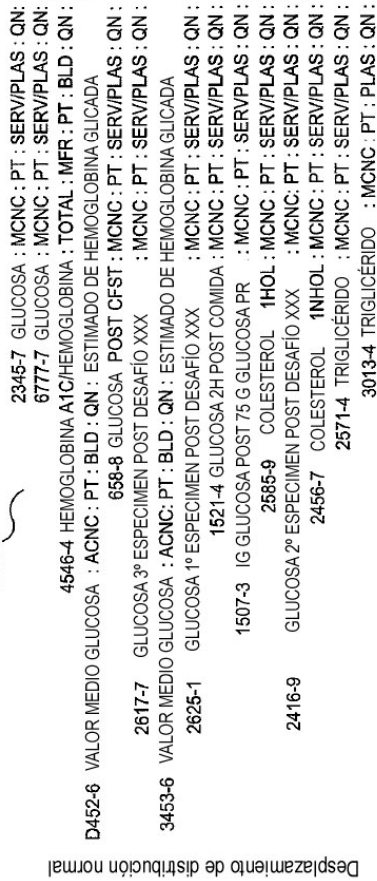


FIG. 13

1400

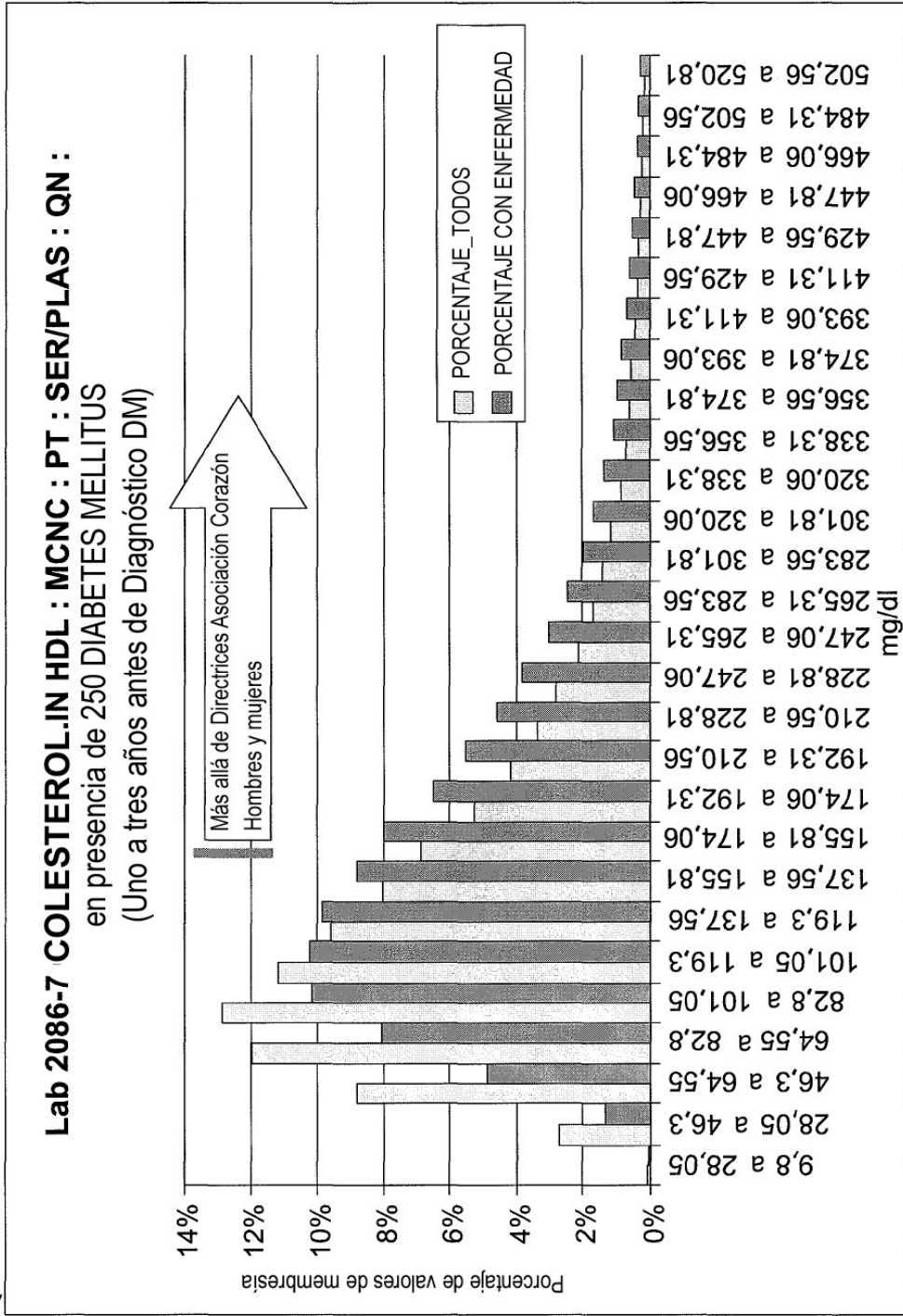


FIG. 14

1500

A	B	C	D	E	F	G	H	I	J
1	CÓDIGO_TITULO						1	Confirmaciones crónicas requeridas	
2	282.4 TALASEMIAS	Código índice (añadir tipo de código)			Títulos código cruzado			Prevalencia histórica de código	
3	<input checked="" type="radio"/> Índice de Dx	282.4	ICD9		282.4 TALASEMIAS			(Miembro contado solamente si código de año de calendario está presente)	
4	<input type="radio"/> Índice de Px								
5									
6									
7	<input type="radio"/> Índice de Rx								
8									
9	<input type="radio"/> Índice de LX	785.6	LOINC		2160-0 CREATININA : MC			Prevalencia histórica (crónica) de código	
10		valor entre 1.6			9999.0 Valor	pg		(El miembro se contará al presentarse y en años elegibles posteriores)	
11									
12		periodo limpio	180 días	ok					
13		periodo agregación ventana (w)							
14									
15									
16									
17									
18									
19									
20									
21									
22									
23									
24									
25									
26									
27									
28									
29									
30									
31									
32									
33									
34									
35									
36									
37									

Mediciones de laboratorio cohorte clínica pre/post índice

Índice demográfico (inicio)

Valores de laboratorio específicos (LOINC) en presencia de enfermedad (por WAP)

Valores de laboratorio (LOINC) que muestran desplazamiento en distribución en presencia de atributo (por WAP)

Prej. para Ventanas pre-índice -730 a -365 días

Ventana agregación

Periodo limpio

Prej. +1095 días

Prej. Ventanas post-índice -90 a +365 días

Ventana agregación

Sólomente reivindicar buscando miembros como cohortes

Edad de muerte en presencia de código

Inicio a muerte en presencia de código

Mortalidad longitudinal actualmente... Período fijado de 2001 a 2007 Inicio en 2001

Agrupación de diagnóstico temporal por atributo de índice (valor)

Agrupación de procedimiento temporal por atributo de índice (valor)

Agrupación de fármacos temporal por atributo de índice (valor)

FIG. 15