(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 111164704 A
(43)申请公布日 2020.05.15

(21)申请号 201880061882.9

(22)申请日 2018.09.20

(30)优先权数据
15/715,548 2017.09.26 US
15/801,714 2017.11.02 US

(85)PCT国际申请进入国家阶段日
2020.03.24

(86)PCT国际申请的申请数据
PCT/IB2018/057236 2018.09.20

(87)PCT国际申请的公布数据
WO2019/064128 EN 2019.04.04

(71)申请人 国际商业机器公司
地址 美国纽约

(72)发明人 J·维尔玛 罗衡 胡建英 张平

(74)专利代理机构 中国国际贸易促进委员会专利商标事务所 11038
代理人 邹丹

(51)Int.Cl.
G16H 50/20(2018.01)
G16H 70/40(2018.01)
G16C 20/10(2019.01)
G06N 20/00(2019.01)
G16C 20/30(2019.01)
G16C 20/70(2019.01)

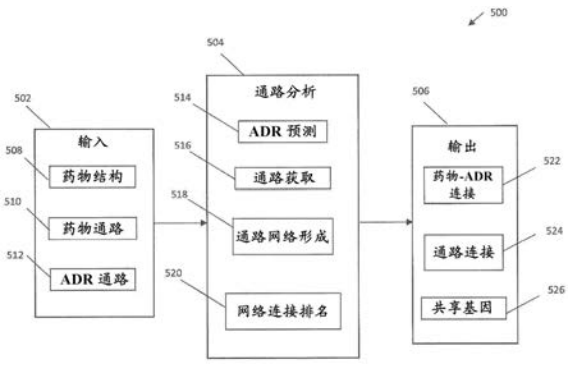
权利要求书2页 说明书12页 附图18页

(54)发明名称

用于候选药物药物不良反应预测的作用机理推导

(57)摘要

实施例包括用于生成作用机理假说的方法、系统和计算机程序产品。各方面包括接收候选药物数据以及与候选药物数据相关联的多个预测的药物不良反应(ADR)。各方面包括接收候选药物的药物通路数据和该多个预测的药物不良反应中的每一个的药物不良反应通路数据。各方面包括构建通路网络,其中通路网络包括多个药物通路节点、多个ADR通路节点和多个通路连接。各方面还包括生成通路输出。



CN 111164704 A

1.一种用于生成针对药物不良反应的作用机理假说的计算机实现的方法,所述方法包括:

由处理器接收识别候选药物的候选药物数据以及与所述候选药物数据相关联的多个预测的药物不良反应;

由所述处理器接收候选药物的药物通路数据;

由所述处理器接收针对所述多个预测的药物不良反应中的每一个的药物不良反应通路数据;

由所述处理器构建通路网络,其中所述通路网络包括多个药物通路节点、多个药物不良反应通路节点和多个通路连接;以及

生成通路输出。

2.根据权利要求1所述的计算机实现的方法,其中所述通路输出包括用于所述通路连接的可视化输出。

3.根据权利要求2所述的计算机实现的方法,其中,所述可视化输出可视地描绘所述通路连接中的每个通路连接的统计显著性。

4.根据前述权利要求中的任一项所述的计算机实现的方法,还包括动态通路输出,所述动态通路输出包括用于药物通路节点与药物不良反应节点之间的连接之一的基因的列表。

5.根据前述权利要求中任一项所述的计算机实现的方法,其中,构建所述通路网络包括:

识别所述药物的药物通路与所述药物不良反应的药物不良反应通路之间的通路连接,并对所述通路连接进行统计分析。

6.根据权利要求5所述的计算机实现的方法,其中统计分析所述通路连接包括将Jaccard指数应用于所述通路连接。

7.根据前述权利要求中任一项所述的计算机实现的方法,还包括将机器学习模型应用于所述候选药物以生成所述多个预测的药物不良反应。

8.一种用于生成针对药物不良反应的作用机理假说的处理系统,包括:

与一种或多种类型的存储器通信的处理器,所述处理器被配置为:

接收候选药物数据,所述候选药物数据识别候选药物以及与所述候选药物数据相关联的多个预测的药物不良反应;

接收所述候选药物的药物通路数据;

接收针对所述多个预测的药物不良反应中的每一个的药物不良反应通路数据;

构建通路网络,所述通路网络包括多个药物通路节点、多个药物不良反应通路节点和多个通路连接;以及

生成通路输出。

9.根据权利要求8所述的处理系统,其中所述通路输出包括用于所述通路连接的可视化输出。

10.如权利要求9所述的处理系统,其中所述可视化输出可视地描绘所述通路连接中的每个通路连接的统计显著性。

11.如权利要求8至10中任一项所述的处理系统,其中构建所述通路网络包括识别所述

药物的药物通路与所述药物不良反应的药物不良反应通路之间的通路连接，并统计分析所述通路连接。

12.根据权利要求8至11中任一项所述的处理系统，其中，所述处理器被配置为将机器学习模型应用于所述候选药物以生成所述多个预测的药物不良反应。

13.一种用于显示针对药物不良反应的作用机理假说的计算机实现的方法，所述方法包括：

构建候选药物与药物不良反应(ADR)之间的通路网络，其中所述通路网络包含用于药物的多个药物通路节点、用于相关联ADR的多个ADR节点以及所述药物通路与ADR通路之间的连接；

在图形用户界面上的药物通路区域中显示所述多个药物通路节点；

在所述图形用户界面上的ADR通路区域中显示多个ADR通路节点；以及

通过由一条或多条线将所述药物通路节点中的一个或多个连接至所述ADR通路节点中的一个或多个来显示多个通路连接，其中，所述线中的每条线的相对粗细反映药物通路-ADR通路连接的统计显著性。

14.根据权利要求13所述的计算机实现的方法，还包括在所述图形用户界面上的共享基因区域中动态地显示作为所述通路连接中的一个或多个通路连接的基础的一组基因。

15.根据权利要求14所述的计算机实现的方法，其中显示多个通路连接包括显示Sankey图。

16.一种用于生成作用机理假说的系统，包括：

输入，包括药物结构输入、药物通路输入和药物不良反应(ADR)输入；

通路分析引擎，包括ADR预测模块、通路获取模块、通路网络形成模块和网络连接排名模块；以及

系统输出接口。

17.根据权利要求16所述的系统，其中所述输出接口包括多个药物-ADR连接和多个通路连接。

18.根据权利要求17所述的系统，其中所述输出接口进一步包括所述通路连接中的一个或多个通路连接的共享基因的列表。

19.一种用于生成针对药物不良反应的作用机理假说的计算机程序产品，所述计算机程序产品包括：一种计算机可读存储介质，其可由处理电路读取并且存储用于由所述处理电路执行以用于执行根据权利要求1至7中任一项所述的方法的指令。

20.一种存储在计算机可读介质上并且可加载到数字计算机的内部存储器中的计算机程序，包括软件代码部分，当所述程序在计算机上运行时，用于执行根据权利要求1至7中任一项所述的方法。

21.一种用于显示针对药物不良反应的作用机理假说的计算机程序产品，所述计算机程序产品包括：一种计算机可读存储介质，其可由处理电路读取并且存储用于由所述处理电路执行以用于执行根据权利要求13至15中任一项所述的方法的指令。

22.一种存储在计算机可读介质上并且可加载到数字计算机的内部存储器中的计算机程序，包括软件代码部分，当所述程序在计算机上运行时，用于执行根据权利要求13至15中任一项所述的方法。

用于候选药物药物不良反应预测的作用机理推导

技术领域

[0001] 本发明总体上涉及药物不良反应预测,更具体地,涉及用于候选药物-药物不良反应的预测的作用机理假说推导。

背景技术

[0002] 药物不良反应(ADR)是由药物的正常使用引起的非预期的和潜在有害的反应。ADR代表全世界的重要公共健康问题。预测ADR对于安全药物开发和精确药物治疗是非常有价值和重要的。已经开发了机器学习模型以帮助开发和评估候选药物,例如帮助预测药物不良反应ADR。尽管一些机器学习模型可以有效地预测ADR,但是这样的模型可能缺乏生物学解释,并且输出可能难以从生物学角度解释。例如,机器学习模型的输出可以包括潜在ADR的识别和表示与ADR相关联的统计可能性的数值。

[0003] 因此,在本领域中需要解决上述问题。

发明内容

[0004] 从第一方面来看,本发明提供了一种用于生成针对药物不良反应的作用机理假说的计算机实现的方法,该方法包括:由处理器接收识别候选药物的候选药物数据以及与候选药物数据相关联的多个预测的药物不良反应;由处理器接收候选药物的药物通路数据;由处理器接收针对该多个预测的药物不良反应中的每一个的药物不良反应通路数据;由处理器构建通路网络,其中通路网络包括多个药物通路节点、多个药物不良反应通路节点和多个通路连接;以及生成通路输出。

[0005] 从另一方面来看,本发明提供了一种用于生成针对药物不良反应的作用机理假说的处理系统,包括:与一种或多种类型的存储器通信的处理器,处理器被配置为:接收识别候选药物的候选药物数据以及与候选药物数据相关联的多个预测的药物不良反应;接收候选药物的药物通路数据;接收针对该多个预测的药物不良反应中的每一个的药物不良反应通路数据;构建通路网络,其中通路网络包括多个药物通路节点、多个药物不良反应通路节点和多个通路连接;以及生成通路输出。

[0006] 从另一方面来看,本发明提供了一种用于显示药物不良反应的作用机理假说的计算机实现的方法,方法包括:构建候选药物与药物不良反应(ADR)之间的通路网络,其中通路网络包含用于药物的多个药物通路节点、用于相关联ADR的多个ADR节点以及药物通路与ADR通路之间的连接;在图形用户界面上的药物通路区域中显示多个药物通路节点;在图形用户界面上的ADR通路区域中显示多个ADR通路节点;以及通过用一条或多条线将药物通路节点中的一个或多个连接到ADR通路节点中的一个或多个来显示多个通路连接,其中每条线的相对粗细反映药物通路-ADR通路连接的统计显著性。

[0007] 从另一方面来看,本发明提供了一种用于生成作用机理假说的系统,包括:输入,包括药物结构输入、药物通路输入和药物不良反应(ADR)输入;通路分析引擎,包括ADR预测模块、通路获取模块、通路网络形成模块和网络连接排名模块;以及系统输出接口。

[0008] 从另一方面来看，本发明提供了一种用于生成针对药物不良反应的作用机理假说的计算机程序产品，该计算机程序产品包括计算机可读存储介质，该计算机可读存储介质可由处理电路读取并且存储由处理电路执行以执行用于执行本发明的步骤的方法的指令。

[0009] 从另一方面来看，本发明提供了一种用于显示针对药物不良反应的作用机理假说的计算机程序产品，所述计算机程序产品包括计算机可读存储介质，该计算机可读存储介质可由处理电路读取并且存储由处理电路执行以执行用于执行本发明的步骤的方法的指令。

[0010] 从另一方面来看，本发明提供了一种存储在计算机可读介质上并且可加载到数字计算机的内部存储器中的计算机程序，包括软件代码部分，当所述程序在计算机上运行时，用于执行本发明的步骤。

[0011] 根据本发明的实施例，提供了一种用于生成针对药物不良反应的作用机理假说的计算机实现的方法。该方法的非限制性实例包括通过处理器接收识别候选药物的候选药物数据以及与候选药物数据相关联的多个预测的药物不良反应。该方法还包括通过处理器接收用于候选药物的药物通路数据。该方法还包括由处理器接收针对多个预测的药物不良反应中的每一个的药物不良反应通路数据。该方法还包括由处理器构建通路网络，其中通路网络包括多个药物通路节点、多个药物不良反应通路节点和多个通路连接。该方法还包括生成通路输出。

[0012] 根据本发明的实施例，提供了一种用于生成针对药物不良反应的作用机理假说的计算机程序产品。计算机程序产品的非限制性示例包括可由处理电路读取的计算机可读存储介质。计算机可读存储介质存储由处理电路执行以执行方法的程序指令。该方法包括接收识别候选药物的候选药物数据以及与候选药物相关联的多个预测的药物不良反应。该方法还包括接收候选药物的药物通路数据。该方法还包括接收多个预测的药物不良反应中的每一个的药物不良反应通路数据。该方法还包括构建通路网络，其中通路网络包括多个药物通路节点、多个药物不良反应通路节点和多个通路连接。该方法还包括生成通路输出。

[0013] 根据本发明的实施例，一种用于生成针对药物不良反应的作用机理假说的处理系统包括与一种或多种类型的存储器通信的处理器。处理器被配置为接收识别候选药物的候选药物数据以及与候选药物数据相关联的多个预测的药物不良反应。处理器还被配置为接收用于候选药物的药物通路数据。处理器还被配置为接收针对所述多个预测的药物不良反应中的每一个的药物不良反应通路数据。处理器还被配置成构建通路网络，其中通路网络包括多个药物通路节点、多个药物不良反应通路节点和多个通路连接。处理器还被配置为生成通路输出。

[0014] 根据本发明的实施例，提供了一种用于显示针对药物不良反应的作用机理假说的计算机实现的方法。该方法的非限制性实例包括在候选药物和药物不良反应(ADR)之间建立通路网络，其中通路网络包括用于药物的多个药物通路节点、用于相关联ADR的多个ADR节点以及药物通路和ADR通路之间的连接。该方法还包括在图形用户界面上的药物通路区域中显示多个药物通路节点。该方法还包括在图形用户界面上的ADR通路区域中显示多个ADR通路节点。该方法还包括通过用一条或多条线将药物通路节点中的一个或多个连接到ADR通路节点中的一个或多个来显示多个通路连接，其中每条线的相对粗细反映药物通路-ADR通路连接的统计显著性。

[0015] 根据本发明的实施例，提供了一种用于生成作用机理假说的系统。该系统的非限制性示例包括输入，该输入包括药物结构输入、药物通路输入和药物不良反应(ADR)输入。该系统还包括通路分析引擎，该通路分析引擎包括ADR预测模块、通路获取模块、通路网络形成模块和网络连接排名模块。该系统还包括系统输出接口。

附图说明

[0016] 本发明的主题在说明书结尾处的权利要求中被特别指出并明确要求保护。从结合附图的以下详细描述中，本文描述的一个或多个实施例的前述和其它特征和优点是清晰的，在附图中：

[0017] 图1示出了根据本发明实施例的云计算环境。

[0018] 图2示出了根据本发明实施例的抽象模型层。

[0019] 图3示出了根据本发明实施例的计算机系统。

[0020] 图4描绘了用于本发明的实施例中的药物-药物不良反应(ADR)预测的示例性机器学习模型界面。

[0021] 图5示出了根据本发明实施例的示例性系统。

[0022] 图6示出根据本发明实施例的示例性系统界面。

[0023] 图7描述了说明根据本发明实施例的示例性方法的流程图。

[0024] 图8描述了说明根据本发明实施例的示例性方法的流程图。

[0025] 图9示出了根据本发明实施例的示例性系统的各方面。

[0026] 图10示出了根据本发明实施例的示例性系统的各方面。

[0027] 图11示出了根据本发明实施例的示例性系统的各方面。

[0028] 图12示出了根据本发明实施例的示例性系统的各方面。

[0029] 图13示出了根据本发明实施例的示例性系统的各方面。

[0030] 图14示出了根据本发明实施例的示例性系统的各方面。

[0031] 图15示出了根据本发明实施例的示例性系统的各方面。

[0032] 图16示出了根据本发明实施例的示例性系统的各方面。

[0033] 图17示出了根据本发明实施例的示例性系统的各方面。

[0034] 图18示出了根据本发明实施例的示例性系统的各方面。

具体实施方式

[0035] 首先应当理解，尽管本公开包括关于云计算的详细描述，但其中记载的技术方案的实现却不限于云计算环境，而是能够结合现在已知或以后开发的任何其它类型的计算环境而实现。

[0036] 云计算是一种服务交付模式，用于对共享的可配置计算资源池进行方便、按需的网络访问。可配置计算资源是能够以最小的管理成本或与服务提供者进行最少的交互就能快速部署和释放的资源，例如可以是网络、网络带宽、服务器、处理、内存、存储、应用、虚拟机和服务。这种云模式可以包括至少五个特征、至少三个服务模型和至少四个部署模型。

[0037] 特征包括：

[0038] 按需自助式服务：云的消费者在无需与服务提供者进行人为交互的情况下能够单

方面自动地按需部署诸如服务器时间和网络存储等的计算能力。

[0039] 广泛的网络接入:计算能力可以通过标准机制在网络上获取,这种标准机制促进了通过不同种类的瘦客户机平台或厚客户机平台(例如移动电话、膝上型电脑、个人数字助理PDA)对云的使用。

[0040] 资源池:提供者的计算资源被归入资源池并通过多租户(multi-tenant)模式服务于多重消费者,其中按需将不同的实体资源和虚拟资源动态地分配和再分配。一般情况下,消费者不能控制或甚至并不知晓所提供的资源的确切位置,但可以在较高抽象程度上指定位置(例如国家、州或数据中心),因此具有位置无关性。

[0041] 迅速弹性:能够迅速、有弹性地(有时是自动地)部署计算能力,以实现快速扩展,并且能迅速释放来快速缩小。在消费者看来,用于部署的可用计算能力往往显得是无限的,并能在任意时候都能获取任意数量的计算能力。

[0042] 可测量的服务:云系统通过利用适于服务类型(例如存储、处理、带宽和活跃用户帐号)的某种抽象程度的计量能力,自动地控制和优化资源效用。可以监测、控制和报告资源使用情况,为服务提供者和消费者双方提供透明度。

[0043] 服务模型如下:

[0044] 软件即服务(SaaS):向消费者提供的能力是使用提供者在云基础架构上运行的应用。可以通过诸如网络浏览器的瘦客户机接口(例如基于网络的电子邮件)从各种客户机设备访问应用。除了有限的特定于用户的应用配置设置外,消费者既不管理也不控制包括网络、服务器、操作系统、存储、乃至单个应用能力等的底层云基础架构。

[0045] 平台即服务(PaaS):向消费者提供的能力是在云基础架构上部署消费者创建或获得的应用,这些应用利用提供者支持的程序设计语言和工具创建。消费者既不管理也不控制包括网络、服务器、操作系统或存储的底层云基础架构,但对其部署的应用具有控制权,对应用托管环境配置可能也具有控制权。

[0046] 基础架构即服务(IaaS):向消费者提供的能力是消费者能够在其中部署并运行包括操作系统和应用的任意软件的处理、存储、网络和其他基础计算资源。消费者既不管理也不控制底层的云基础架构,但是对操作系统、存储和其部署的应用具有控制权,对选择的网络组件(例如主机防火墙)可能具有有限的控制权。

[0047] 部署模型如下:

[0048] 私有云:云基础架构单独为某个组织运行。云基础架构可以由该组织或第三方管理并且可以存在于该组织内部或外部。

[0049] 共同体云:云基础架构被若干组织共享并支持有共同利害关系(例如任务使命、安全要求、政策和合规考虑)的特定共同体。共同体云可以由共同体内的多个组织或第三方管理并且可以存在于该共同体内部或外部。

[0050] 公共云:云基础架构向公众或大型产业群提供并由出售云服务的组织拥有。

[0051] 混合云:云基础架构由两个或更多部署模型的云(私有云、共同体云或公共云)组成,这些云依然是独特的实体,但是通过使数据和应用能够移植的标准化技术或私有技术(例如用于云之间的负载平衡的云突发流量分担技术)绑定在一起。

[0052] 云计算环境是面向服务的,特点集中在无状态性、低耦合性、模块性和语意的互操作性。云计算的核心是包含互连节点网络的基础架构。

[0053] 现在参考图1，描述了根据本发明的一个或多个实施例的说明性云计算环境50。如图所示，云计算环境50包括云消费者使用的本地计算设备可以与其通信的一个或多个云计算节点10，本地计算设备例如个人数字助理(PDA)或蜂窝电话54A、台式计算机54B、膝上型计算机54C和/或汽车计算机系统54N。节点10可以彼此通信。它们可以被物理地或虚拟地分组(未示出)在一个或多个网络中，诸如如上文描述的私有云、共同体云、公共云或混合云或其组合。这允许云计算环境50提供基础架构、平台和/或软件即服务，云消费者不需要为其维护本地计算设备上的资源。应当理解，图1中所示的计算设备54A-N的类型仅旨在说明，并且计算节点10和云计算环境50可以通过任何类型的网络和/或网络可寻址连接(例如，使用web浏览器)与任何类型的计算机化设备通信。

[0054] 现在参考图2，示出了根据本发明的一个或多个实施例的由云计算环境50(图1)提供的一组功能抽象层。应当预先理解，图2中所示的组件、层和功能仅旨在说明，并且本发明的实施例不限于此。如所描绘的，提供了以下层和相应的功能：

[0055] 硬件和软件层60包括硬件和软件组件。硬件组件的示例包括：主机61；基于RISC(精简指令集计算机)架构的服务器62；服务器63；刀片服务器64；存储装置65；以及网络和网络组件66。在本发明的一些实施例中，软件组件包括网络应用服务器软件67和数据库软件68。

[0056] 虚拟化层70提供抽象层，从该抽象层可以提供虚拟实体的以下示例：虚拟服务器71；虚拟存储器72；虚拟网络73，包括虚拟专用网络；虚拟应用和操作系统74；以及虚拟客户机75。

[0057] 在一个示例中，管理层80可以提供以下描述的功能。资源供应81提供用于在云计算环境内执行任务的计算资源和其它资源的动态采购。计量和定价82提供了在云计算环境内利用资源时的成本跟踪，以及因为消耗这些资源而进行帐单开具或发票开具。在一个示例中，这些资源可以包括应用软件许可。安全性为云消费者和任务提供身份验证，以及为数据和其他资源提供保护。用户门户83为消费者和系统管理员提供对云计算环境的访问。服务级别管理84提供云计算资源分配和管理，使得满足所需的服务级别。服务水平协议(SLA)规划和履行85提供对云计算资源的预安排和采购，其中根据SLA预期未来需求。

[0058] 工作负载层90提供了可以利用云计算环境的功能的示例。可以从该层提供的工作负载和功能的示例包括：绘图和导航91；软件开发和生命周期管理92；虚拟教室教育传送93；数据分析处理94；交易处理95；和药物不良反应分析96。

[0059] 现在参考图3，其中示出了根据本发明的一个或多个实施例的包括在分布式云环境或云服务网络中的云计算节点100的示意图。云计算节点100仅仅是适合的云计算节点的一个示例，不应对本发明实施例的功能和使用范围带来任何限制。总之，云计算节点100能够被用来实现和/或执行以上所述的任何功能。

[0060] 云计算节点100具有计算机系统/服务器12，其可与众多其它通用或专用计算系统环境或配置一起操作。众所周知，适于与计算机系统/服务器12一起操作的计算系统、环境和/或配置的例子包括但不限于：个人计算机系统、服务器计算机系统、瘦客户机、厚客户机、手持或膝上设备、基于微处理器的系统、机顶盒、可编程消费电子产品、网络个人电脑、小型计算机系统、大型计算机系统和包括上述任意系统的分布式云计算技术环境，等等。

[0061] 计算机系统/服务器12可以在由计算机系统执行的计算机系统可执行指令(诸如

程序模块)的一般语境下描述。通常,程序模块可以包括执行特定的任务或者实现特定的抽象数据类型的例程、程序、目标程序、组件、逻辑、数据结构等。计算机系统/服务器12可以在通过通信网络链接的远程处理设备执行任务的分布式云计算环境中实施。在分布式云计算环境中,程序模块可以位于包括存储设备的本地或远程计算系统存储介质上。

[0062] 如图3所示,云计算节点100中的计算机系统/服务器12以通用计算设备的形式表现。计算机系统/服务器12的组件可以包括但不限于:一个或者多个处理器或者处理单元16,系统存储器28,连接不同系统组件(包括系统存储器28和处理单元16)的总线18。

[0063] 总线18表示几类总线结构中的一种或多种,包括存储器总线或者存储器控制器,外围总线,图形加速端口,处理器或者使用多种总线结构中的任意总线结构的局域总线。举例来说,这些体系结构包括但不限于工业标准体系结构(ISA)总线,微通道体系结构(MAC)总线,增强型ISA总线、视频电子标准协会(VESA)局域总线以及外围组件互连(PCI)总线。

[0064] 计算机系统/服务器12典型地包括多种计算机系统可读介质。这些介质可以是能够被计算机系统/服务器12访问的任意可获得的介质,包括易失性和非易失性介质,可移动的和不可移动的介质。

[0065] 系统存储器28可以包括易失性存储器形式的计算机系统可读介质,例如随机存取存储器(RAM)30和/或高速缓存存储器32。计算机系统/服务器12可以进一步包括其它可移动/不可移动的、易失性/非易失性计算机系统存储介质。仅作为举例,存储系统34可以用于读写不可移动的、非易失性磁介质(图1未显示,通常称为“硬盘驱动器”)。尽管图1中未示出,可以提供用于对可移动非易失性磁盘(例如“软盘”)读写的磁盘驱动器,以及对可移动非易失性光盘(例如CD-ROM,DVD-ROM或者其它光介质)读写的光盘驱动器。在这些情况下,每个驱动器可以通过一个或者多个数据介质接口与总线18相连。存储器28可以包括至少一个程序产品,该程序产品具有一组(例如至少一个)程序模块,这些程序模块被配置以执行本发明各实施例的功能。

[0066] 具有一组(至少一个)程序模块42的程序/实用工具40,可以存储在存储器28中,这样的程序模块42包括但不限于操作系统、一个或者多个应用程序、其它程序模块以及程序数据,这些示例中的每一个或某种组合中可能包括网络环境的实现。程序模块42通常执行本发明所描述的实施例中的功能和/或方法。

[0067] 计算机系统/服务器12也可以与一个或多个外部设备14(例如键盘、指向设备、显示器24等)通信,还可与一个或者多个使得用户能与该计算机系统/服务器12交互的设备通信,和/或与使得该计算机系统/服务器12能与一个或多个其它计算设备进行通信的任何设备(例如网卡,调制解调器等等)通信。这种通信可以通过输入/输出(I/O)接口22进行。并且,计算机系统/服务器12还可以通过网络适配器20与一个或者多个网络(例如局域网(LAN),广域网(WAN)和/或公共网络,例如因特网)通信。如图所示,网络适配器20通过总线18与计算机系统/服务器12的其它模块通信。应当明白,尽管图中未示出,其它硬件和/或软件模块可以与计算机系统/服务器12一起操作,包括但不限于:微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、RAID系统、磁带驱动器以及数据备份存储系统等。

[0068] 现在转到与本发明的方面更具体相关的技术的概述,药物不良反应(ADR)是可能由按正常剂量服用药物引起的潜在不良临床状况。药物开发公司和研究人员不断地寻求在药物开发早期识别ADR和潜在ADR,以减少与开发活动相关联的时间、成本和患者风险。

[0069] 为了识别和预测ADR,可以使用计算方法,诸如关于结构描述符的机器学习模型、分子对接分布型的相似性分析、基于网络的方法和电子健康记录的数据挖掘。尽管一些方法可以生成ADR预测,但是没有生成关于潜在的生物学方法的信息。例如,这样的模型可以输出量化药物或候选药物与ADR的关联的得分。如本文所用,"药物"和"候选药物"可互换使用,并且旨在包括任何研究中的化学物质。

[0070] 图4描绘了用于本发明的实施例中的机器学习模型的示例性药物ADR预测界面。在所示的例子中,输入界面402可以包括三维化学结构描绘404和/或化学描述符输入406,例如简化的分子线性输入系统(SMILES)输入。输出404显示机器学习模型药物-ADR预测的结果,其可包括ADR列表408,其基于化学描述符描绘所得的预测的ADR,以及与药物-ADR预测相关联的置信度得分列表410。仍然不清楚的是这些模型的生物学解释,包括特定ADR为何对于一种候选药物比对于另一种具有更高的关联评分的基本原理或原因。对药物-ADR关联的更深入理解和解释是非常需要的。

[0071] 现在转到本发明的方面的概述,本发明的一个或多个实施方案通过提供药物或候选药物和ADR代谢通路的比较,从而提供药物-ADR关联的作用机理的假说,解决了现有技术的上述缺点。本发明的一个或多个实施方案可以提供药物和ADR之间的统计学关联的可视表示,从而提供增强的输出以更深入地理解和解释药物-ADR关联。如本文所用,通路数据和代谢通路数据可互换使用,并且包括与药物(例如,药物通路数据)和/或ADR(例如,ADR通路数据)相关联或链接的基因、基因相互作用、基因富集等,如所示。在本发明的一些实施方案中,可视表示和/或代谢通路比较提供了对与药物相关联的ADR的作用机理的假说。通过显示共享通路和这些通路中的共同基因以及相对显著性,药物开发者可以容易地确定预测的ADR的可能原因,从而使他们的研究工作流线化并更快地开发安全有效的药物。

[0072] 本发明的上述方面通过从基因表达分析和从文献中获取用预测模型生成的药物-ADR对中的药物和ADR的通路信息,并确定和定量药物-ADR对之间的通路连接,解决了现有技术的缺点。药物和ADR之间的通路连接可以例如通过Jaccard指数进行统计分析和排名。排名靠前的连接可作为药物或候选药物如何诱导ADR的候选作用机理假说。本发明的一些实施方案提供了可视图,例如Sankey图、Venn图、Upset可视化等,其可视地捕获并提供从药物到药物通路到ADR通路到ADR的信息流。这种可视图可以提供各种比较的排名的明确显示,例如使用边(即药物和ADR之间的连接物)的粗细,使得代谢比较的结果易于观察和比较,提供优良的候选药物评估。在本发明的一些实施例中,提供了交互式用户界面,其中用户可以与显示的边或连接物交互以获得药物和ADR之间的共享基因的列表以及统计相似性,例如Jaccard相似性。

[0073] 现在转到本发明的方面的更详细描述,图5描绘了根据本发明的一个或多个实施例的示例性系统500的示意图。系统500可以包括如图所示配置和布置的输入502、通路分析引擎504和输出506。输入502可以包括药物结构输入508、药物通路输入510和ADR通路输入512。

[0074] 药物结构输入508可以适于基于其化学结构或子结构,例如SMILES、IUPAC国际化学标识符(InChITM)密钥等,以任何合适的特征向量或格式接收二维或三维化学结构信息。在本发明的一些实施例中,药物结构输入508可以接收药物特征,例如结构描述符(例如,来自PubChem)、例如来自SIDER数据库的标签、以及相关信息。

[0075] 药物通路输入510可以适于从基因表达分析和文献接收药物的通路信息。在本发明的一些实施方式中，药物通路输入510适于接收来自药物效果的基因表达分析的通路数据。根据本发明的实施例，可以访问提供药物通路数据的任何数据库，例如药物路径和/或小分子通路数据库(SMPDB)。

[0076] ADR通路输入512可以适于接收从文献中组织的ADR的通路信息。ADR通路输入可以包括结构化和非结构化数据，并且可以包括但不限于以叙事格式发布在期刊文章中的数据。根据本发明的实施方式，可以访问提供药物通路数据的任何数据库，包括例如京都基因和基因组百科全书(KEGG)、Reactome通路数据库和/或通过进化关系的蛋白质分析(PANTHER)系统。

[0077] 通路分析引擎504可以包括例如ADR预测模块514、通路获取模块516、通路网络形成模块518和网络连接排名模块520。ADR预测模块514可以例如通过使用机器学习模型或用于生成药物的预测的ADR的任何其他预测模型来预测给定药物的一个或多个ADR。通路获取模块516可以从药物通路输入510和从ADR通路输入512获取药物的通路。通路网络形成模块518可以例如使用Jaccard指数、IDF归一化余弦、交集等来识别并统计评估药物通路和ADR通路之间的连接。在本发明的一些实施方案中，使用不受通路中基因数量影响的统计方法，例如Jaccard指数。网络连接排名模块516可以对药物通路和ADR通路之间的连接进行排名。在本发明的一些实施方案中，网络连接排名模块516可生成用于从药物到药物通路到ADR通路到ADR的连接和/或信息流的可视化输出，例如Sankey图，从而提供优良的候选药物评估。

[0078] 输出506可以包括例如药物-ADR连接522，诸如由机器学习模型关联的药物-ADR对、通路连接524以及在每种药物与其具有共同或共享通路的关联ADR之间的共享基因526。药物-ADR连接522、通路连接524和共享基因中的一个或多个可以包括交互式和/或动态组件以增强对候选药物的评估。

[0079] 图6示出了根据本发明实施例的示例性系统输出接口。输出界面可以包括药物602、药物通路区域610、ADR通路区域620和一个或多个ADR 604a、604b、…604n。药物通路区域610可包含多个药物通路节点612a、612b、612c、612d、612e、612f、612g。ADR通路区域620可以包括多个ADR通路节点622a、622b、622c、622d、622e、622f、622g、622h、622i。通路节点代表从相关数据库中获取的用于指定药物和/或ADR的通路。如图所示，数据可以用Sankey图来描述，其中通路节点之间的连接物("边")的粗细可视地表示通路之间的连接的统计显著性。输出界面可以包括动态边，例如，使得当用户点击或触摸边时，出现列出所指示的通路、通路中的基因(基因1、基因2、基因3、…基因n)和诸如Jaccard指数的其他相关信息的图。基因列表可包括所示通路中的全部或部分基因列表，并可任选地通过分析突出显示药物和ADR通路中所示的基因。

[0080] 图7描述了示出根据本发明的实施例的示例性方法700的流程图。方法700可包括接收候选药物和与候选药物相关联的一个或多个预测的ADRS，如框702所示。例如，候选药物和预测的ADR可包括来自机器学习模型的输出。方法700还可以包括接收候选药物的药物通路数据，如在框704所示。方法700还可以包括接收ADR的药物通路数据，如框706所示。如方框708所示，方法700包括构建通路网络，其中通路网络包括药物通路节点和ADR通路节点以及药物通路和ADR通路之间的连接。该方法还包括生成通路输出，如框710所示，其中通路输出包括预测的ADR和药物-ADR通路连接以及相关的统计排名。在本发明的一些实施例中，

作用机理输出包括通路网络的可视化显示和统计排名，例如在具有一个或多个动态组件的Sankey图中。

[0081] 图8示出了根据本发明的另一个实施例的用于生成通路输出的示例性方法800的流程图。方法800包括在候选药物和ADR之间构建通路网络，其中通路网络包括用于药物的多个药物通路节点、用于相关联ADR的ADR通路节点以及药物通路和ADR通路之间的连接，如框802所示。该方法还包括在图形用户界面上的药物通路区域中显示多个药物通路节点，如在块804处所示。该方法还包括在图形用户界面上的ADR通路区域中显示多个ADR通路节点，如在框806处所示。如方框808所示，该方法还包括通过将每个药物通路节点通过一条或多条线连接到一个或多个ADR通路节点来显示多个通路连接，其中每条线的粗细反映药物通路-ADR通路连接的相对统计显著性。如在框810处所示，该方法还包括可选地在图形用户界面上的共享基因区域中动态地显示作为通路连接中的一个或多个的基础的一组基因。

[0082] 图9示出了根据本发明的示例性实施例的基于机器学习的ADR预测的示意图。在本发明的一些实施例中，方法包括使用机器学习模型来预测ADR。机器学习模型可以使用包括已知诱发给定ADR的药物的阳性药物组902和包括对给定ADR未知的药物的阴性药物组904。机器学习模型可以使用诸如SMILES格式的药物结构，以及例如来自PubChem的指纹的结构描述符906，如框910中所示，作为用于给定ADR的机器学习模型908的输入。机器学习模型可以为药物生成预测的ADR的集合。

[0083] 图10描述了根据本发明的示例性实施方式的通路集合1000的示意图。例如从机器学习模型导出的药物1002和ADR 1008可以作为输入被接收。如图所示，对于选择的药物1002，可以从一个或多个已知的通路数据库，例如药物通路Drug-Path，获取多个药物通路1004a、1004b、1004c、…1004n。对于给定的ADR 1008，可以从已知的通路数据源获取多个ADR通路1006a、1006b、1006c、…1006n，例如通过从文献(例如KEGG、Reactome等)中管理通路信息。

[0084] 图11描述了根据本发明的示例性实施例的通路连接和统计分析的示意图。例如，从如图10所示收集的药物通路和ADR通路数据开始，可以通过共享基因分析药物通路和ADR通路的关联。作为这种分析的结果，可以形成药物通路和ADR通路之间的连接，以形成通路网络1102。在和不在ADR通路中的基因可以相对在和不在药物通路中的基因作图，如图11中的1104所示，并且Jaccard指数1106可以通过已知方法确定并输出。

[0085] 图12描绘了根据本发明的实施例的示例性可视化1200和假说生成。通路连接可以例如通过Jaccard指数来统计地排名，并且排名靠前的连接可以被强调，例如通过线粗细来强调通路连接。药物-ADR通路连接通过提供关于药物如何诱导ADR的基本原理和/或原因而用作作用机理假说。药物通路和ADR通路共享的基因可以是湿实验假说检验的起点。

[0086] 例如，在早期药物开发阶段，制药公司可以使用根据本发明实施方案的系统和方法来预测候选药物的潜在ADR并识别潜在的作用机理。公司随后可以修改候选药物以避免ADR并提高安全性。

[0087] 作为另一个例子，在药物开发的售后阶段，制药公司可以使用根据本发明实施方案的系统和方法来识别与其药物产品相关联的ADR的作用机理。通过研究通路和在共享通路中识别的基因，有可能发现对某些ADR敏感的遗传生物标记。因此，医疗提供者和药物提供者可以建议具有所识别的生物标志物的患者调整剂量或处方以避免ADR(称为"精确医

学")。

[0088] 在任何可能的技术细节结合层面,本发明可以是系统、方法和/或计算机程序产品。计算机程序产品可以包括计算机可读存储介质,其上载有用于使处理器实现本发明的各个方面的计算机可读程序指令。

[0089] 计算机可读存储介质可以是可以保持和存储由指令执行设备使用的指令的有形设备。计算机可读存储介质例如可以是——但不限于——电存储设备、磁存储设备、光存储设备、电磁存储设备、半导体存储设备或者上述的任意合适的组合。计算机可读存储介质的更具体的例子(非穷举的列表)包括:便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、静态随机存取存储器(SRAM)、便携式压缩盘只读存储器(CD-ROM)、数字多功能盘(DVD)、记忆棒、软盘、机械编码设备、例如其上存储有指令的打孔卡或凹槽内凸起结构、以及上述的任意合适的组合。这里所使用的计算机可读存储介质不被解释为瞬时信号本身,诸如无线电波或者其他自由传播的电磁波、通过波导或其他传输媒介传播的电磁波(例如,通过光纤电缆的光脉冲)、或者通过电线传输的电信号。

[0090] 这里所描述的计算机可读程序指令可以从计算机可读存储介质下载到各个计算/处理设备,或者通过网络、例如因特网、局域网、广域网和/或无线网下载到外部计算机或外部存储设备。网络可以包括铜传输电缆、光纤传输、无线传输、路由器、防火墙、交换机、网关计算机和/或边缘服务器。每个计算/处理设备中的网络适配卡或者网络接口从网络接收计算机可读程序指令,并转发该计算机可读程序指令,以供存储在各个计算/处理设备中的计算机可读存储介质中。

[0091] 用于执行本发明操作的计算机程序指令可以是汇编指令、指令集架构(ISA)指令、机器指令、机器相关指令、微代码、固件指令、状态设置数据、集成电路配置数据或者以一种或多种编程语言的任意组合编写的源代码或目标代码,所述编程语言包括面向对象的编程语言—诸如Smalltalk、C++等,以及过程式编程语言—诸如"C"语言或类似的编程语言。计算机可读程序指令可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络—包括局域网(LAN)或广域网(WAN)—连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。在一些实施例中,通过利用计算机可读程序指令的状态信息来个性化定制电子电路,例如可编程逻辑电路、现场可编程门阵列(FPGA)或可编程逻辑阵列(PLA),该电子电路可以执行计算机可读程序指令,从而实现本发明的各个方面。

[0092] 这里参照根据本发明实施例的方法、装置(系统)和计算机程序产品的流程图和/或框图描述了本发明的各个方面。应当理解,流程图和/或框图的每个方框以及流程图和/或框图中各方框的组合,都可以由计算机可读程序指令实现。

[0093] 这些计算机可读程序指令可以提供给通用计算机、专用计算机或其它可编程数据处理装置的处理器,从而生产出一种机器,使得这些指令在通过计算机或其它可编程数据处理装置的处理器执行时,产生了实现流程图和/或框图中的一个或多个方框中规定的功能/动作的装置。也可以把这些计算机可读程序指令存储在计算机可读存储介质中,这些指

令使得计算机、可编程数据处理装置和/或其他设备以特定方式工作，从而，存储有指令的计算机可读介质则包括一个制造品，其包括实现流程图和/或框图中的一个或多个方框中规定的功能/动作的各个方面的指令。

[0094] 也可以把计算机可读程序指令加载到计算机、其它可编程数据处理装置、或其它设备上，使得在计算机、其它可编程数据处理装置或其它设备上执行一系列操作步骤，以产生计算机实现的过程，从而使得在计算机、其它可编程数据处理装置、或其它设备上执行的指令实现流程图和/或框图中的一个或多个方框中规定的功能/动作。

[0095] 附图中的流程图和框图显示了根据本发明的多个实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段或指令的一部分，所述模块、程序段或指令的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个连续的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的是，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或动作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

[0096] 本文所用的术语仅是为了描述特定实施例的目的，而不是要限制本发明。如本文所用，单数形式"一"、"一个"和"该"旨在也包括复数形式，除非上下文另有明确指示。还将理解，术语"包括"和/或"包含"在本说明书中使用时，指定所陈述的特征、整数、步骤、操作、元件和/或组件的存在，但不排除一个或多个其它特征、整数、步骤、操作、元件组件和/或其群组的存在或添加。

[0097] 以下权利要求中的所有装置或步骤加功能元素的对应结构、材料、动作和等同物旨在包括用于按照具体要求保护的方式与其它要求保护的元素组合执行功能的任何结构、材料或动作。已经出于说明和描述的目的给出了本发明的描述，但是该描述不是穷举的或者将本发明限制为所描述的形式。在不背离本发明范围的情况下，许多修改和变化对于本领域普通技术人员来说是显而易见的。选择和描述实施例是为了最好地解释本发明的原理和实际应用，并且使本领域的其他普通技术人员能够理解本发明的具有各种修改的各种实施例，这些修改适合于所设想的特定用途。

[0098] 这里描述的流程图仅仅是一个例子。在不脱离本发明的实施例的范围的情况下，可以存在对这里描述的该图或步骤(或操作)的许多变化。例如，可以以不同的顺序执行这些步骤，或者可以添加、删除或修改步骤。所有这些变化都被认为是所要求保护的发明的一部分。

[0099] 以上已经描述了本发明的各实施例，上述说明是示例性的，并非穷尽性的，并且也不限于所披露的各实施例。在不偏离所说明的各实施例的范围和精神的情况下，对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。本文中所用术语的选择，旨在最好地解释各实施例的原理、实际应用或对市场中的技术的技术改进，或者使本技术领域的其它普通技术人员能理解本文披露的各实施例。

[0100] 示例

[0101] 示例1

[0102] 根据本发明的实施例，得出卡马西平(carbamazepine)的作用机理假说。如图13所

示,药物结构1302和SMILES数据1304被提供给预测性的机器学习模型的用户界面1300。该模型在输出1306处生成多个预测的ADR和相关的置信度得分。还输出ADR 1308的证据。

[0103] 如图14所示,对预测的ADR和药物进行了通路关联分析。示例性的通路可视化工具在图形用户界面中生成药物通路和ADR通路的列表,以及许多共享的基因和Jaccard指数。按照Jaccard指数对通路进行排名,如图所示。例如,针对免疫性血小板减少症(immune thrombocytopenia)ADR的卡马西平的主要通路连接是通过MAPK信号传导通路(MAPK signaling pathway)和破骨细胞分化通路(Osteoclast differentiation pathway),包括48个共享基因并且具有Jaccard指数0.142。示例性图形用户界面包括动态组件,其中向用户提供了"点击这里"以可视化通路连接的选项。图15描述了卡马西平的可视化药物-ADR通路分析输出。药物通路和ADR通路使用Sankey图可视化,并且MAPK信号传导通路和破骨细胞分化通路共享的基因通过点击Sankey图中连接这些ADR的边来揭示。基因列表出现在图14的右侧。对列表进行进一步分析,并且在文献中报道了与该连接相关的MAP2K1、MAPK9、MAPK11、MAPK13和MAPK10(在基因列表中识别的第四至第八基因)。

[0104] 示例2

[0105] 根据本发明的实施方案,得到氟米龙(fluorometholone)的作用机理假说。如图16所示,药物结构1602和SMILES数据1604被提供给预测性的机器学习模型的用户界面1600。该模型在输出1606处生成多个预测的ADR和相关的置信度得分。还输出ADR 1608的证据。

[0106] 如图17所示,对预测的药物和药物进行了通路关联分析。示例性的通路可视化工具在图形用户界面中生成药物通路和ADR通路的列表,以及许多共享的基因和Jaccard指数。按照Jaccard指数对通路进行排名,如图所示。例如,针对糖尿病(diabetes mellitus)ADR的氟米龙的主要通路连接是通过T细胞受体信号传导通路(T cell receptor signaling pathway),包括105个共享基因并具有Jaccard指数为1.0。示例性图形用户界面包括动态组件,其中向用户提供了"点击这里"以可视化通路连接的选项。图18描述了氟米龙的可视化药物-ADR通路分析输出。药物通路和ADR通路使用Sankey图可视化。在本示例中,确定药物和ADR具有共同的通路,T细胞受体信号传导通路。在点击Sankey图的顶边之后显示的基因列表出现在图18的右侧。

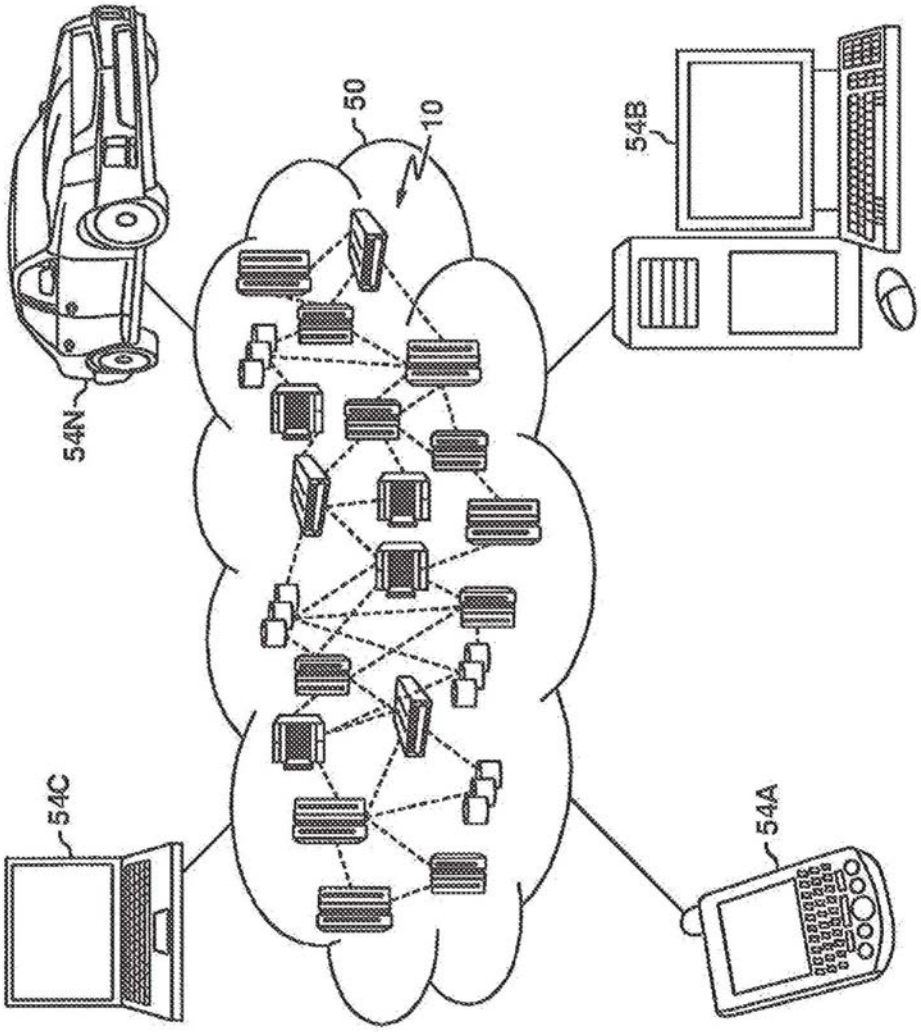


图1

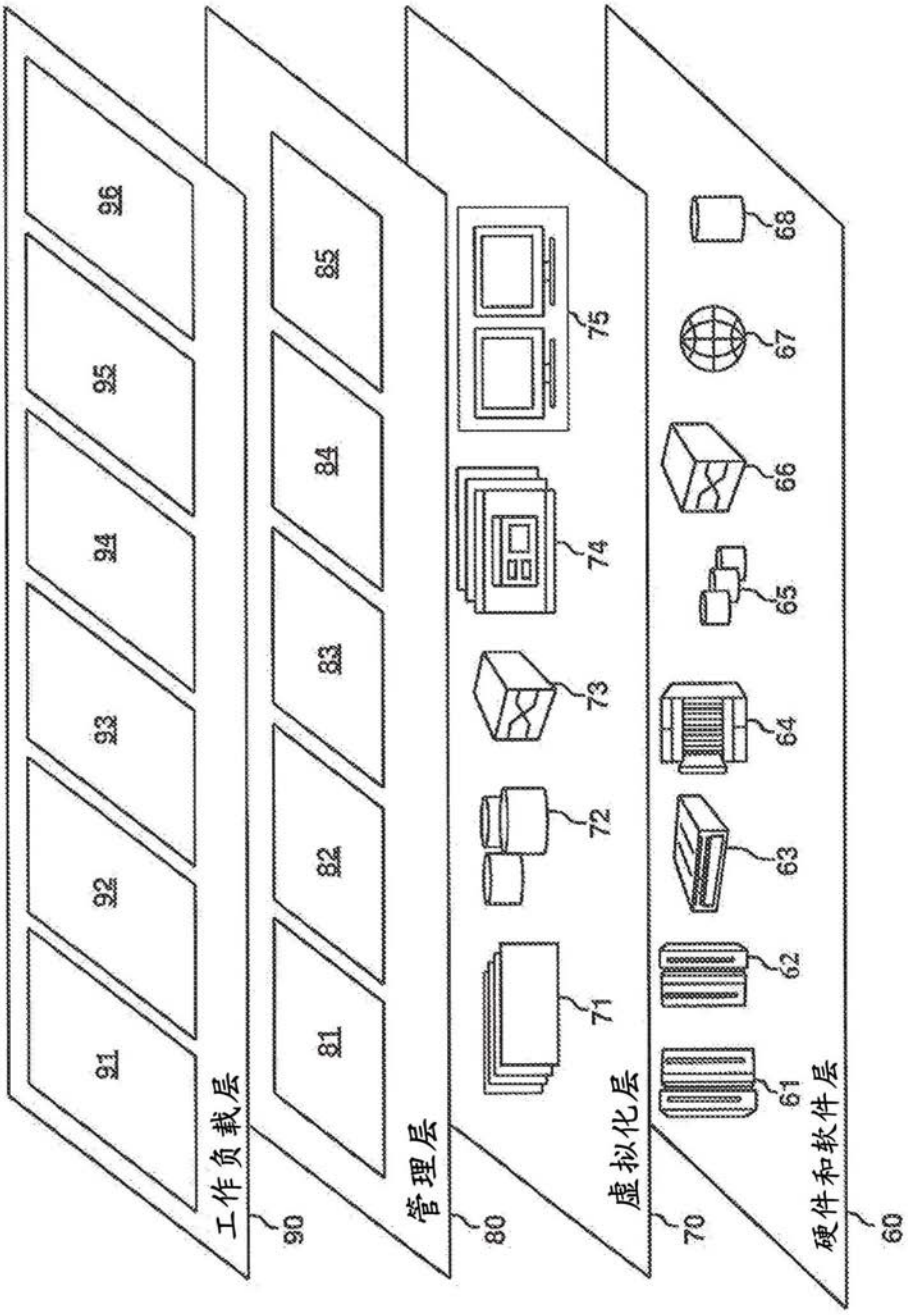

工作负载层
90
91
92
93
94
95
96
管理层
80
81
82
83
84
85
虚拟化层
70
71
72
73
74
75
硬件和软件层
60
61
62
63
64
65
66
67
68

图2

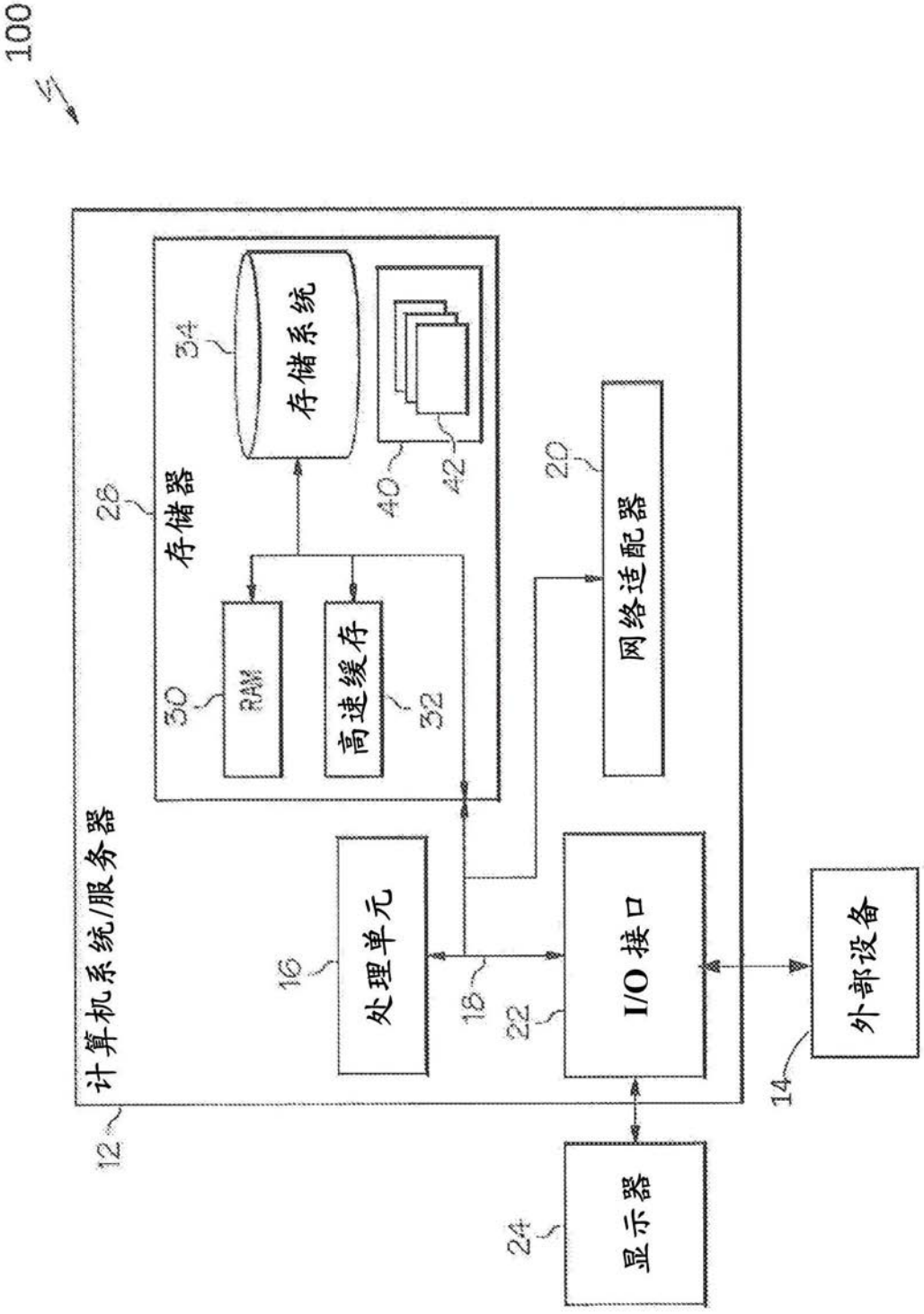

100
计算机系统/服务器
12
存储器
28
存储系统
34
RAM
30
高速缓存
32
40
42
网络适配器
20
处理单元
16
18
I/O 接口
22
外部设备
14
显示器
24

图3

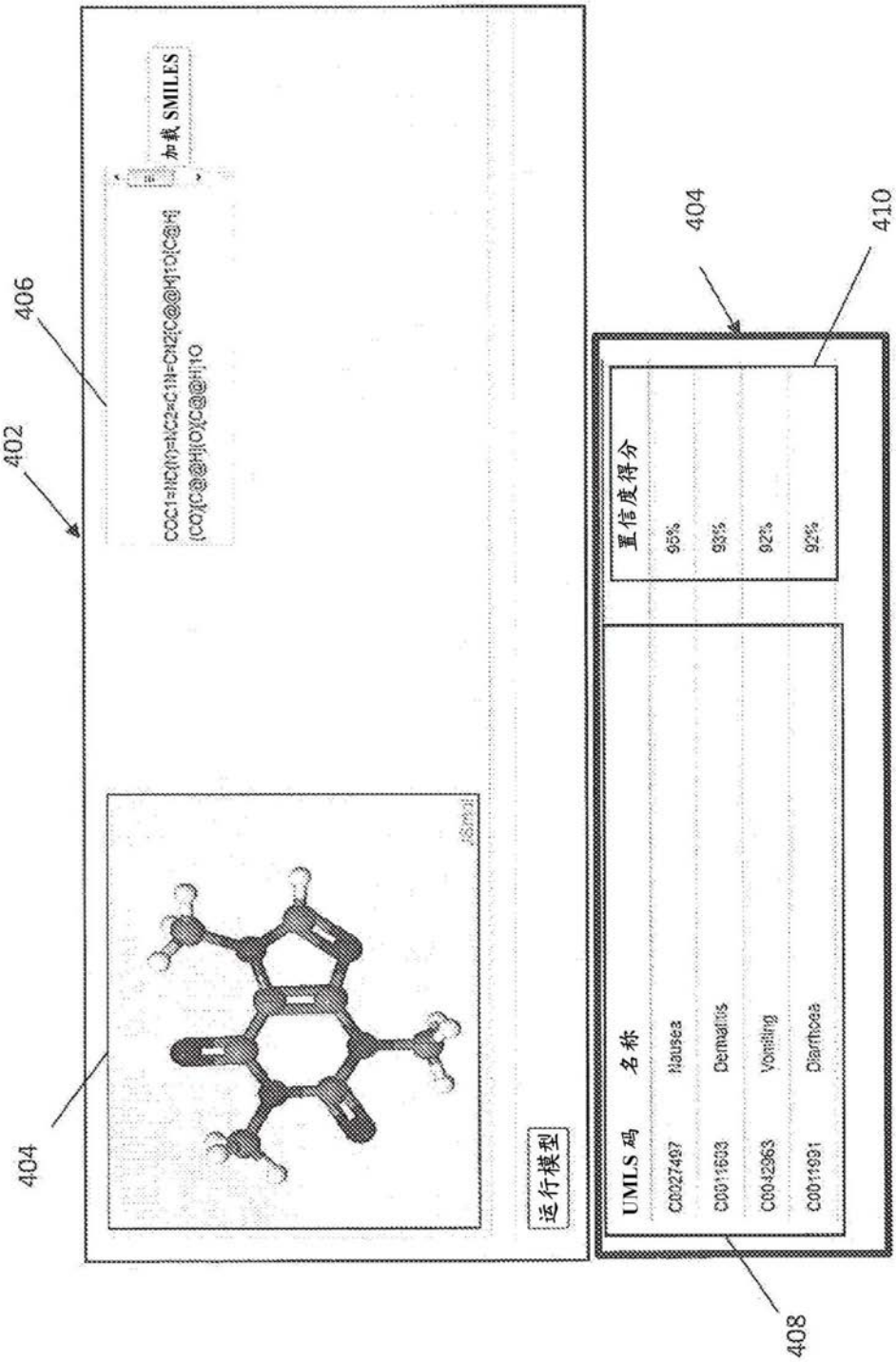

406
402
404
410
404
408
加载 SMILES
运行模型
UMLS 码
名称
置信度得分
C0027497
Nausea
95%
C0011633
Dermatitis
93%
C0042963
Vomiting
92%
C0011991
Diarrhoea
92%

图4

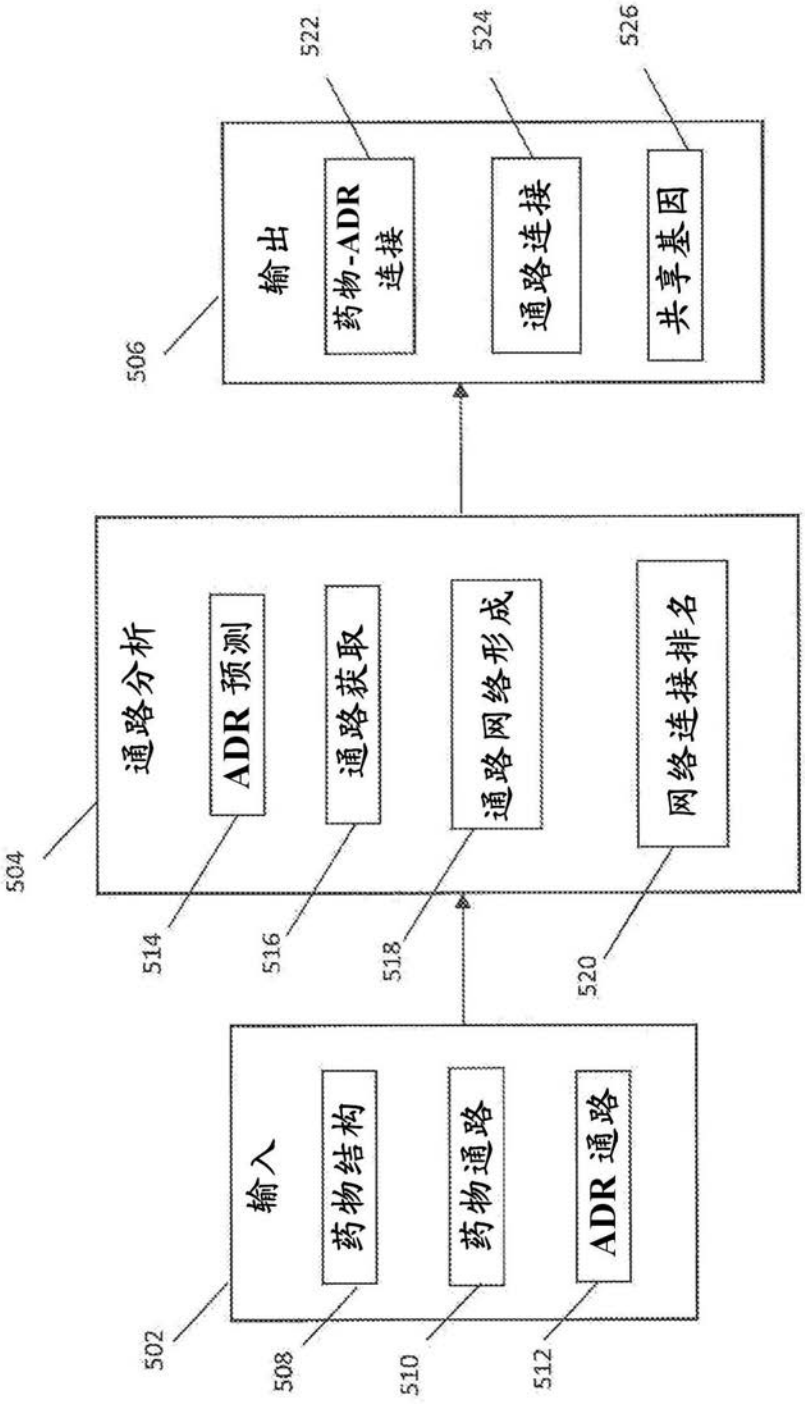

500
输出
药物-ADR 连接
通路连接
共享基因
506
522
524
526
通路分析
ADR 预测
通路获取
通路网络形成
网络连接排名
504
514
516
518
520
输入
药物结构
药物通路
ADR 通路
502
508
510
512

图5

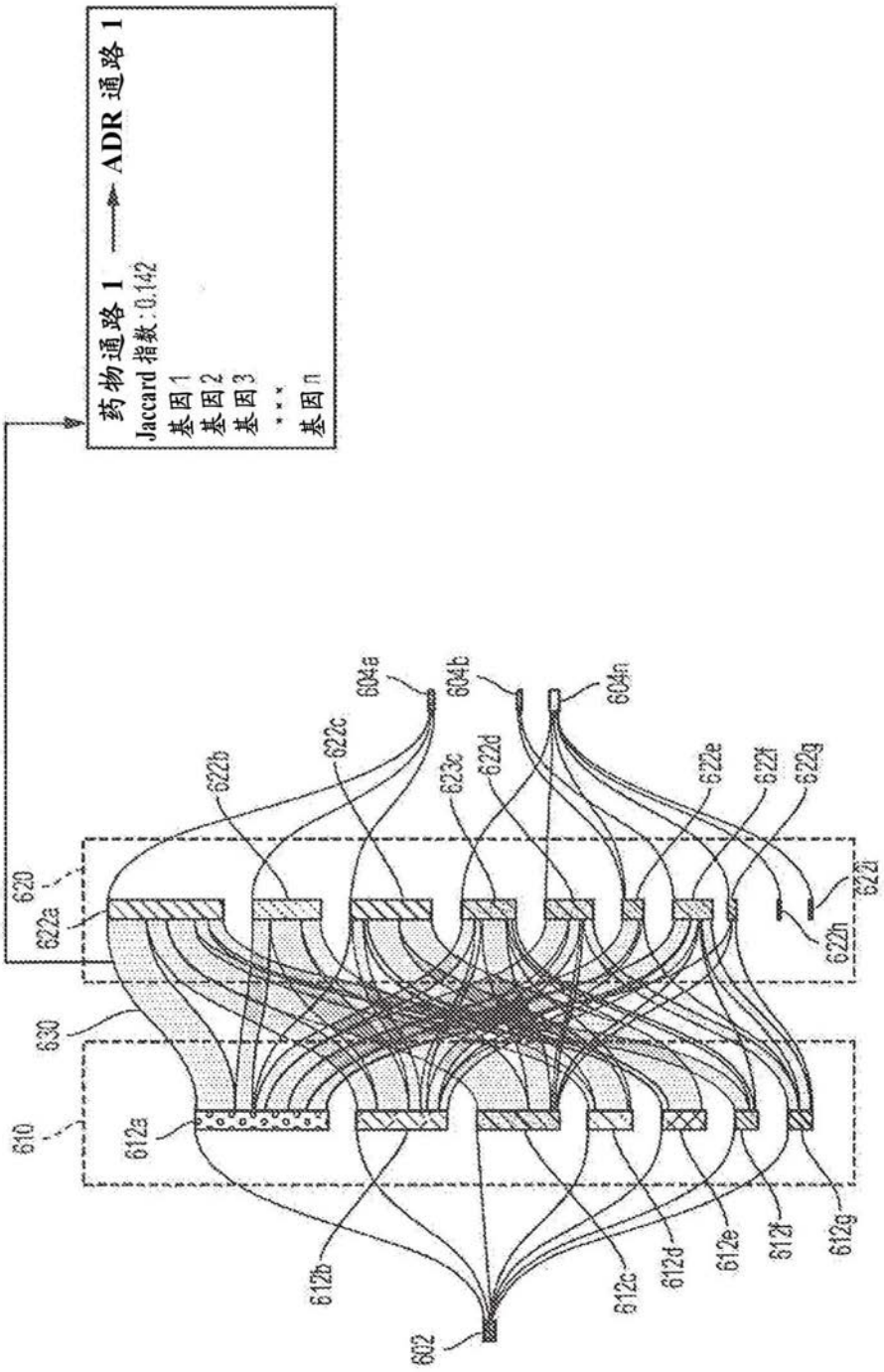

药物通路 1 ⟶ ADR 通路 1
Jaccard 指数：0.142
基因 1
基因 2
基因 3
基因 n
604a
604b
604n
622b
622c
623c
622d
622e
622f
622g
622i
622h
620
622a
630
610
612a
612b
612c
612d
612e
612f
612g
602

图6

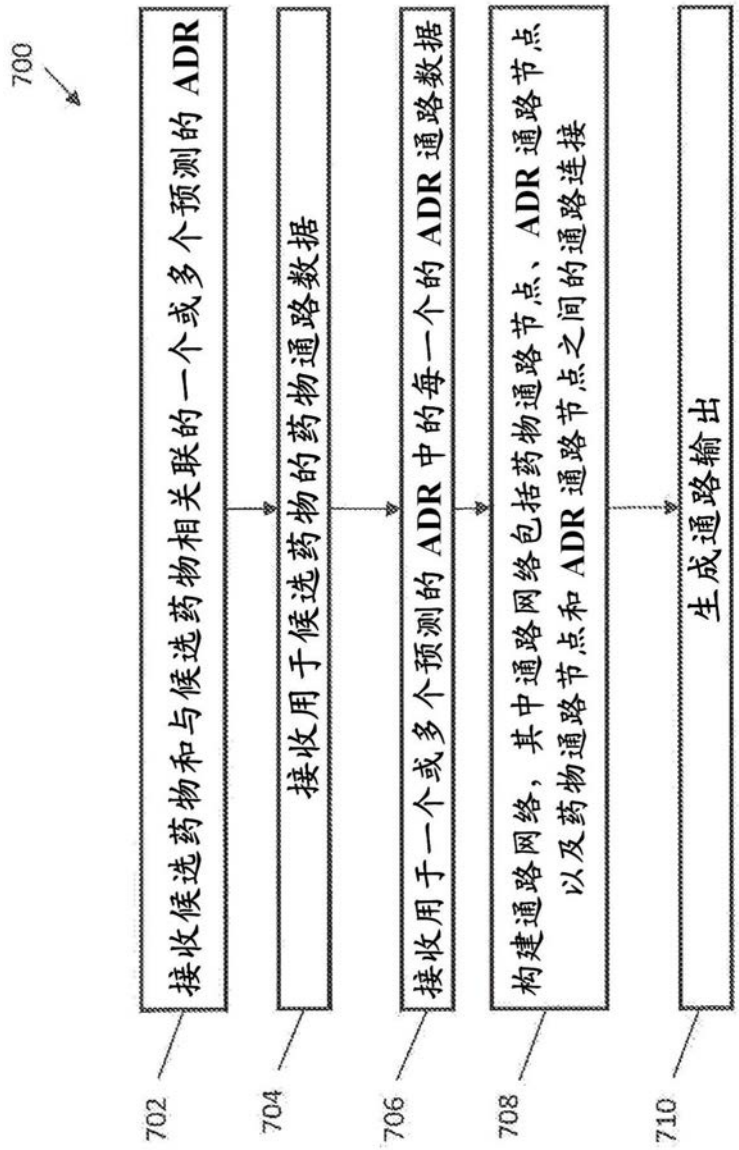

700
接收候选药物和与候选药物相关联的一个或多个预测的 ADR
接收用于候选药物的药物通路数据
接收用于一个或多个预测的 ADR 中的每一个的 ADR 通路数据
构建通路网络，其中通路网络包括药物通路节点、ADR 通路节点以及药物通路节点和 ADR 通路节点之间的通路连接
生成通路输出
702
704
706
708
710

图7

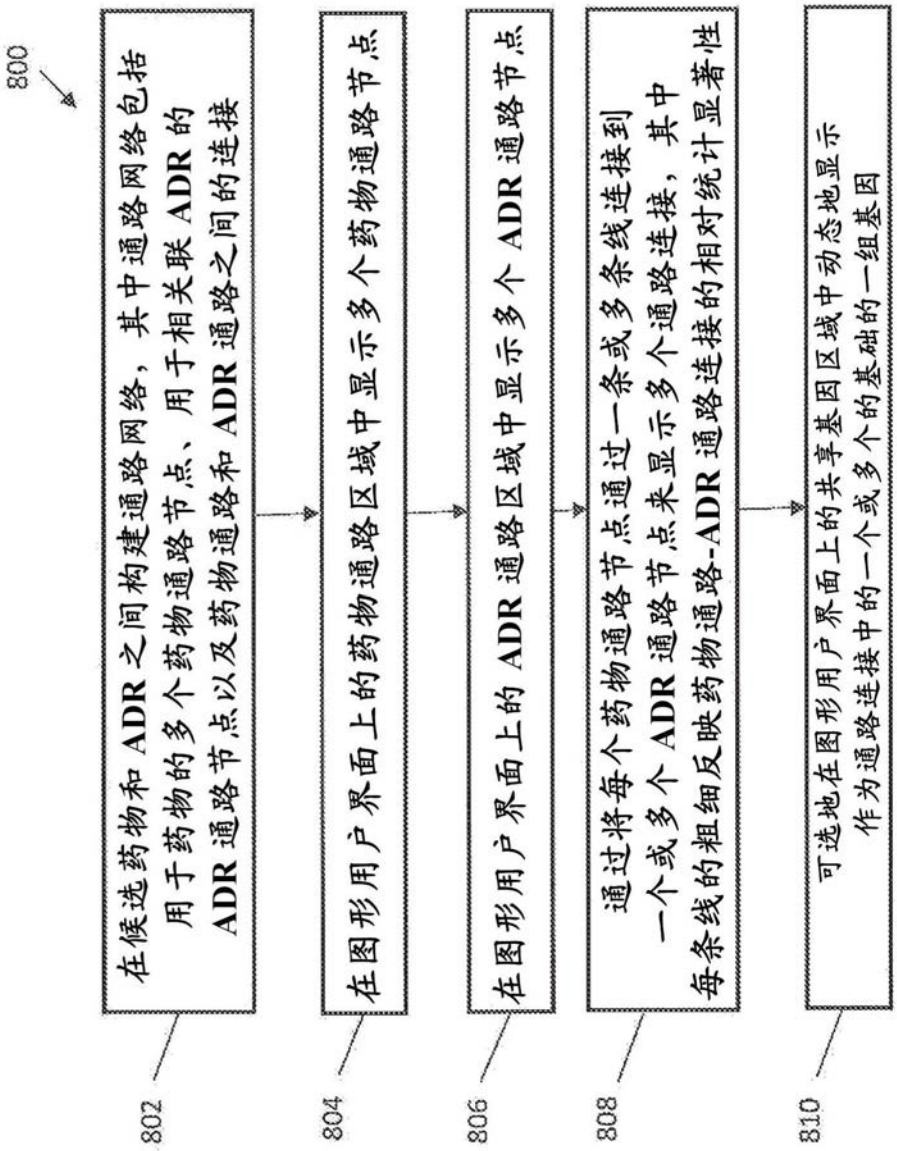

800
在候选药物和 ADR 之间构建通路网络，其中通路网络包括用于药物的多个药物通路节点、用于相关联 ADR 的 ADR 通路节点以及药物通路和 ADR 通路之间的连接
802
在图形用户界面上的药物通路区域中显示多个药物通路节点
804
在图形用户界面上的 ADR 通路区域中显示多个 ADR 通路节点
806
通过将每个药物通路节点通过一条或多条线连接到一个或多个 ADR 通路节点来显示多个通路连接，其中每条线的粗细反映药物通路-ADR 通路连接的相对统计显著性
808
可选地在图形用户界面上的共享基因区域中动态地显示作为通路连接中的一个或多个的基础的一组基因
810

图8

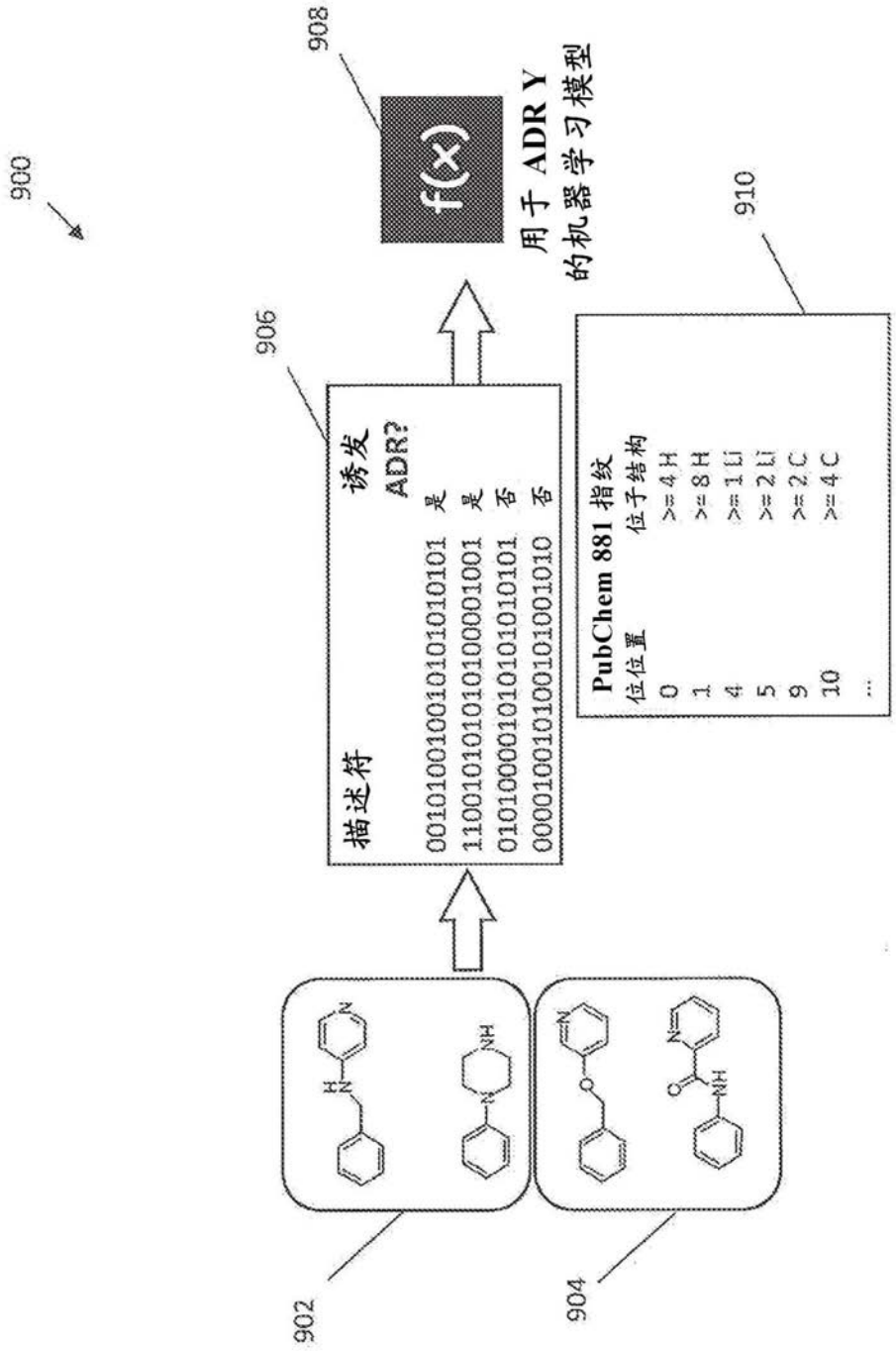

900
908
f(x)
用于 ADR Y
的机器学习模型
910
906
描述符
诱发
ADR?
00101001001010101010101 是
11001010101010100001001 是
01010000101010101010101 否
00001001010010101001010 否
PubChem 881 指纹
位位置 位子结构
0 >= 4 H
1 >= 8 H
4 >= 1 Li
5 >= 2 Li
9 >= 2 C
10 >= 4 C
...
902
904

图9

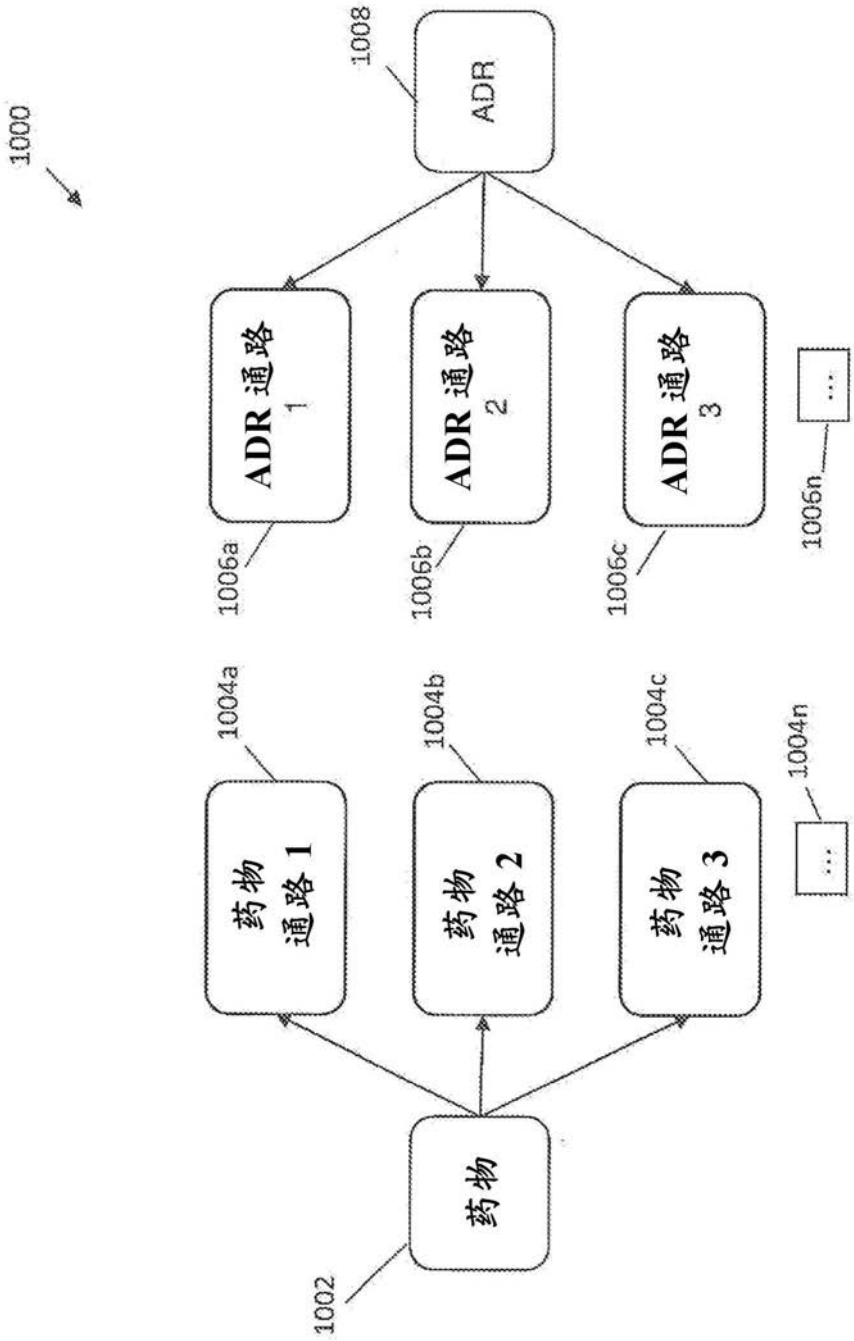

1000
1008
ADR
ADR 通路 1
ADR 通路 2
ADR 通路 3
...
1006a
1006b
1006c
1006n
1004a
1004b
1004c
1004n
药物 通路 1
药物 通路 2
药物 通路 3
...
药物
1002

图10

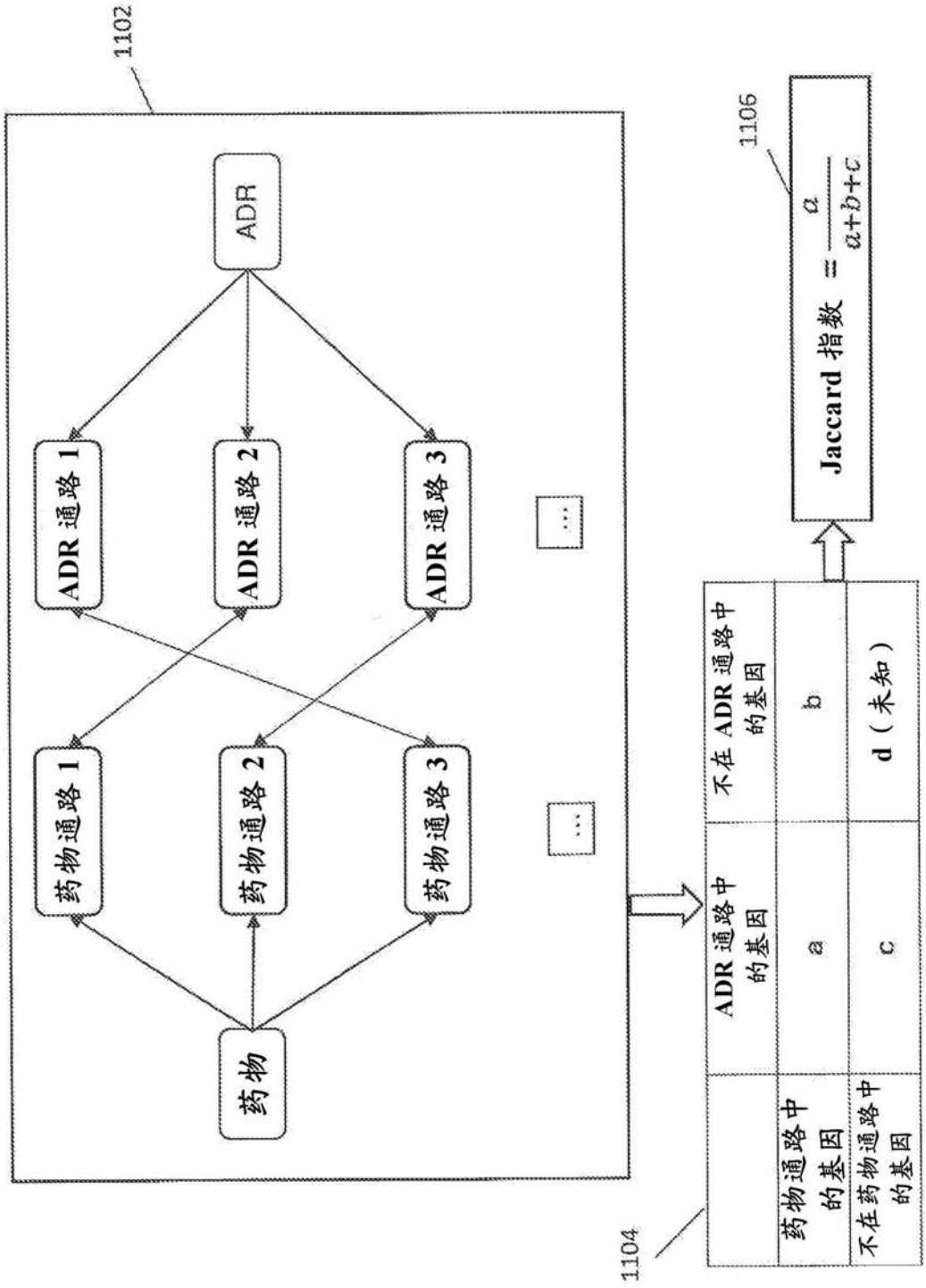

1102
药物
药物通路 1
药物通路 2
药物通路 3
ADR
ADR 通路 1
ADR 通路 2
ADR 通路 3
1104
ADR 通路中的基因
不在 ADR 通路中的基因
药物通路中的基因
a
b
c
d（未知）
1106
Jaccard 指数 = a/(a+b+c)

图11

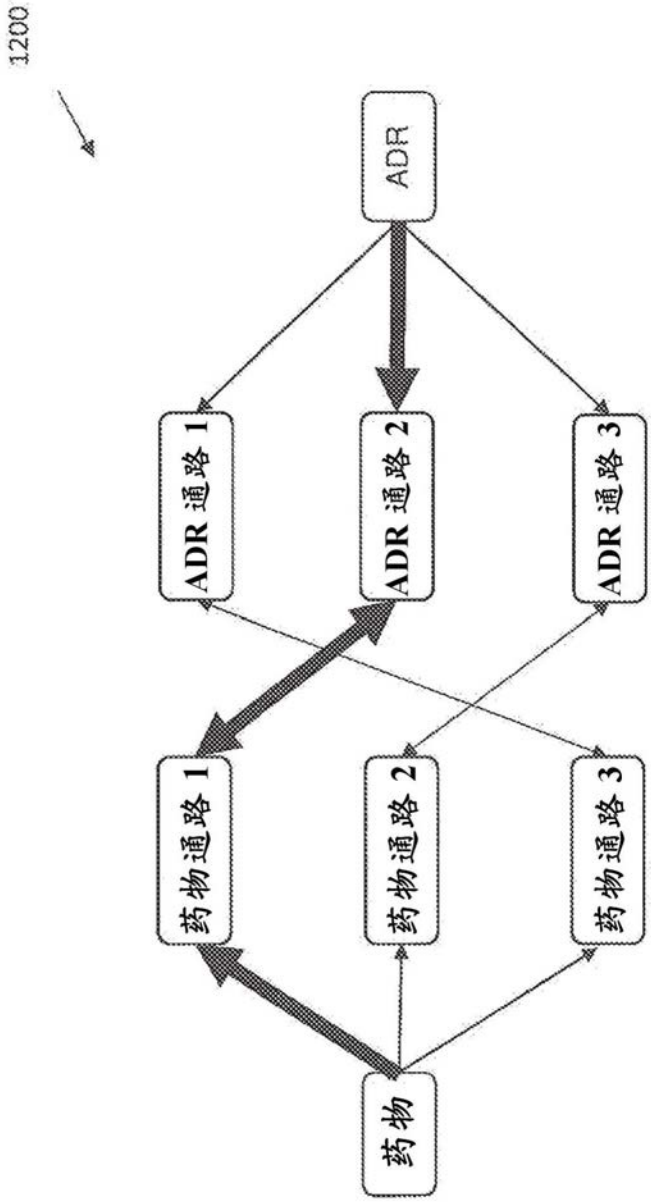


图12

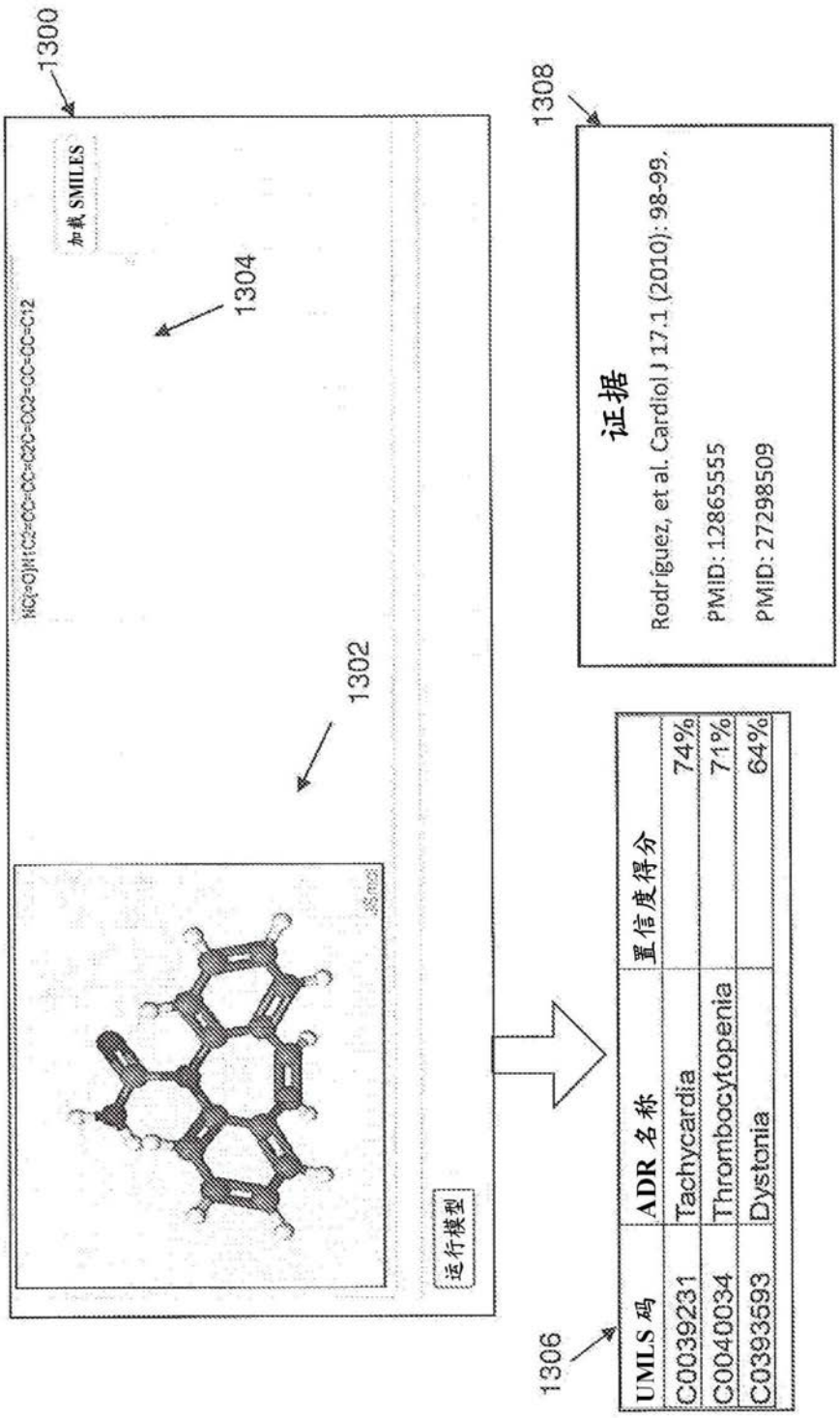

1300
加载 SMILES
1304
1302
运行模型
1308
证据
Rodríguez, et al. Cardiol J 17.1 (2010): 98-99.
PMID: 12865555
PMID: 27298509
1306
UMLS 码
ADR 名称
置信度得分
C0039231
Tachycardia
74%
C0040034
Thrombocytopenia
71%
C0393593
Dystonia
64%

图13

药物通路（基因数量）	ADR	ADR 通路（基因数量）	共享基因	Jacard
MAPK signaling pathway (255)	Immune thrombocytopenia	Osteoclast differentiation (132)	48	0.142
Graft-versus-host disease (41)	Immune thrombocytopenia	Phagosome (154)	20	0.114
Type I diabetes mellitus (43)	Immune thrombocytopenia	Phagosome (154)	20	0.113
Focal adhesion (203)	Immune thrombocytopenia	Fc gamma R-mediated phagocytosis (93)	29	0.109
Small cell lung cancer (86)	Immune thrombocytopenia	Osteoclast differentiation (132)	19	0.095
Small cell lung cancer (86)	Primary dystonia	Adipocytokine signaling pathway (70)	13	0.091
MAPK signaling pathway (255)	Primary dystonia	Dopaminergic synapse (120)	28	0.078
Focal adhesion (203)	Immune thrombocytopenia	Osteoclast differentiation (132)	22	0.07
Small cell lung cancer (86)	Immune thrombocytopenia	Fc gamma R-mediated phagocytosis (93)	11	0.065
MAPK signaling pathway (255)	Catecholaminergic polymorphic ventricular tachycardia	Cardiac muscle contraction (78)	20	0.064
MAPK signaling pathway (255)	Primary dystonia	Cardiac muscle contraction (78)	20	0.064
MAPK signaling pathway (255)	Immune thrombocytopenia	Fc gamma R-mediated phagocytosis (93)	21	0.064
MAPK signaling pathway (255)	Catecholaminergic polymorphic ventricular tachycardia	Calcium signaling pathway (132)	23	0.058
Cytokine-cytokine receptor interaction (265)	Immune thrombocytopenia	Osteoclast differentiation (132)	20	0.053
MAPK signaling pathway (255)	Primary dystonia	Adipocytokine signaling pathway (70)	14	0.045
Focal adhesion (203)	Immune thrombocytopenia	Phagosome (154)	14	0.041

图14

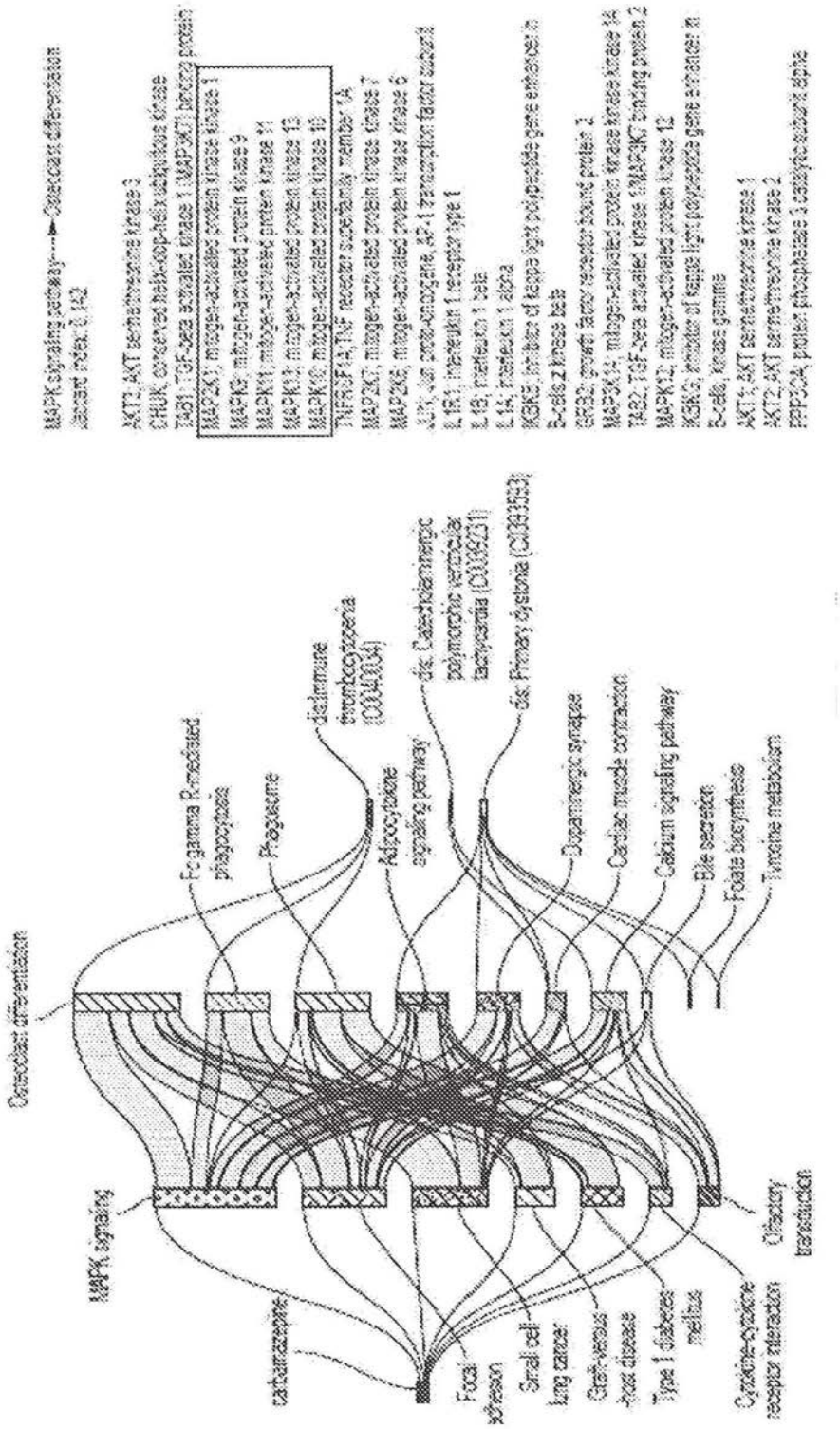

MAPK signaling pathway
Osteoclast differentiation
Fc gamma R-mediated phagocytosis
Phagosome
Adipocytokine signaling pathway
Dopaminergic synapse
Cardiac muscle contraction
Calcium signaling pathway
Bile secretion
Folate biosynthesis
Tyrosine metabolism
Olfactory transduction
Cytokine-cytokine receptor interaction
Type I diabetes mellitus
Graft-versus-host disease
Small cell lung cancer
Focal adhesion
pathways in cancer

图15

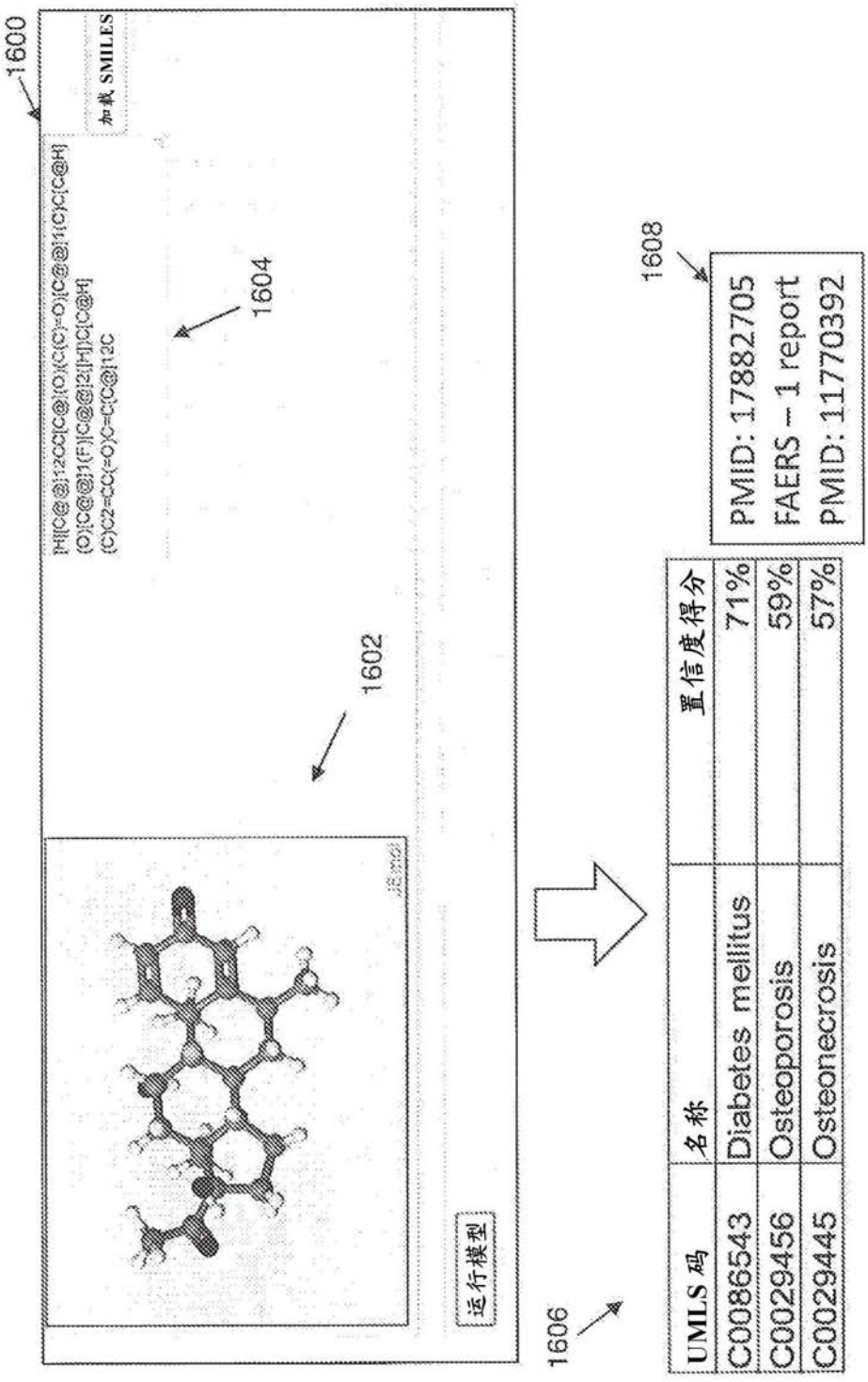

1600
加载 SMILES
1604
1602
运行模型
1606
1608
UMLS 码
名称
置信度得分
C0086543
Diabetes mellitus
71%
C0029456
Osteoporosis
59%
C0029445
Osteonecrosis
57%
PMID: 17882705
FAERS – 1 report
PMID: 11770392

图16

通路可视化工具
药物名称：
氟米龙
例如：卡马西平、氟米龙
选择的 ADR 数量：
3
提交
输出：
映射的疾病：
C0011849 Diabetes mellitus -> Diabetes mellitus
C0029456 Osteoporosis -> Osteoporosis
C0029445 Osteonecrosis -> Glucocorticoid-induced osteonecrosis
点击这里，使通路连接可视化：
结果表格：
药物通路（基因数量） ADR ADR 通路（基因数量） 共享基因 Jaccard
T cell receptor signaling pathway (105) Diabetes mellitus T cell receptor signaling pathway (103) 105 1.0
Cytokine-cytokine receptor interaction (265) Diabetes mellitus Cytokine-cytokine receptor interaction (265) 265 1.0
B cell receptor signaling pathway (73) Diabetes mellitus T cell receptor signaling pathway (105) 48 0.369
Cytokine-cytokine receptor interaction (265) Diabetes mellitus Jak-STAT signaling pathway (158) 93 0.282
T cell receptor signaling pathway (105) Diabetes mellitus Neurotrophin signaling pathway (121) 48 0.215
GnRH signaling pathway (92) Diabetes mellitus Neurotrophin signaling pathway (121) 37 0.21
B cell receptor signaling pathway (73) Diabetes mellitus Neurotrophin signaling pathway (121) 32 0.198
B cell receptor signaling pathway (73) Glucocorticoid-induced osteonecrosis HIF-1 signaling pathway (103) 19 0.121
Systemic lupus erythematosus (135) Diabetes mellitus Type I diabetes mellitus (53) 19 0.119
GnRH signaling pathway (92) Diabetes mellitus T cell receptor signaling pathway (105) 19 0.107
T cell receptor signaling pathway (105) Glucocorticoid-induced osteonecrosis HIF-1 signaling pathway (103) 19 0.101
T cell receptor signaling pathway (105) Diabetes mellitus Jak-STAT signaling pathway (158) 23 0.096
B cell receptor signaling pathway (73) Diabetes mellitus Jak-STAT signaling pathway (158) 17 0.079
GnRH signaling pathway (92) Osteoporosis Wnt signaling pathway (143) 17 0.078
B cell receptor signaling pathway (73) Osteoporosis Wnt signaling pathway (143) 14 0.069
GnRH signaling pathway (92) Glucocorticoid-induced osteonecrosis HIF-1 signaling pathway (103) 11 0.06

图17

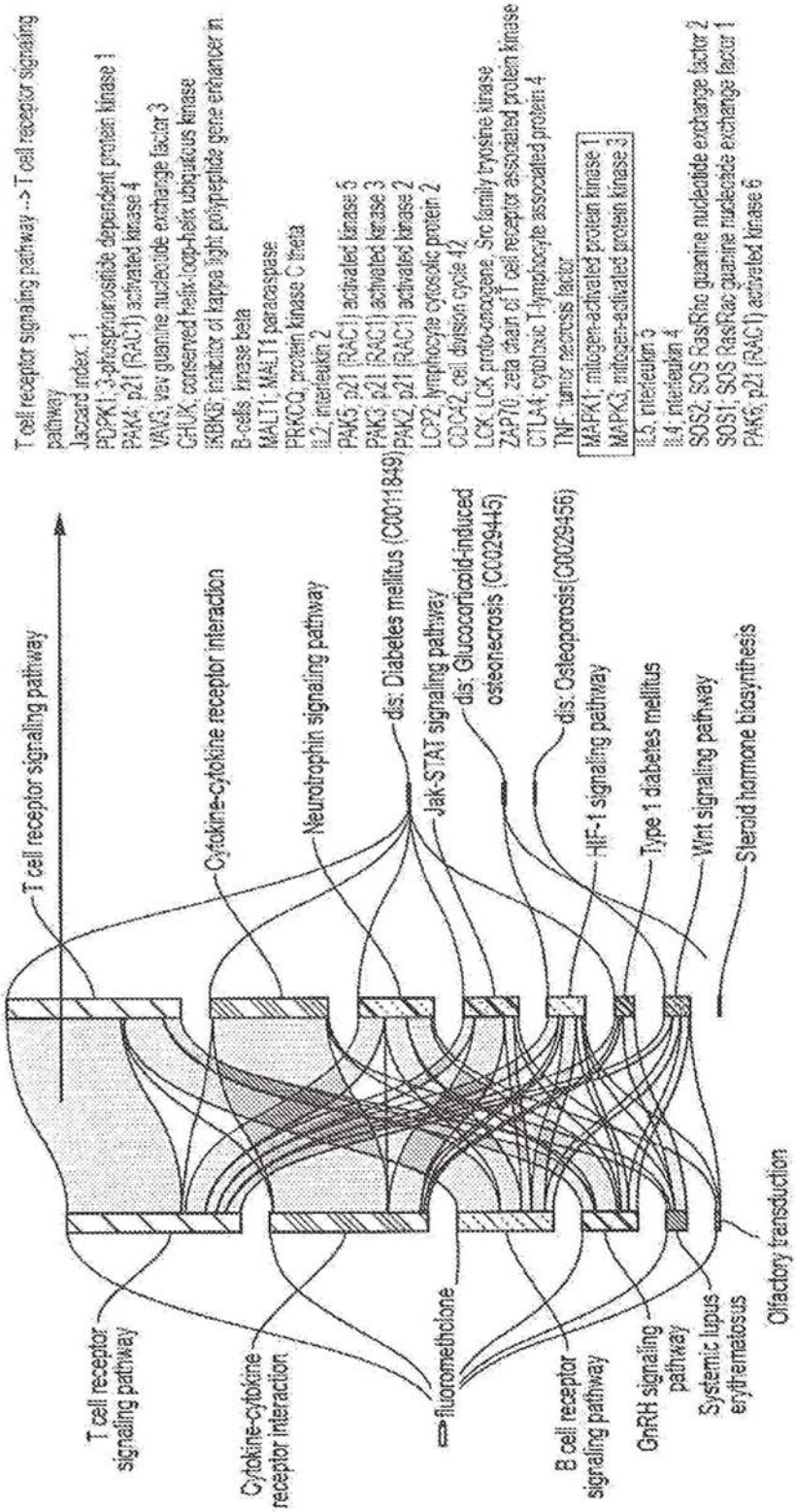

T cell receptor signaling pathway -> T cell receptor signaling pathway
Jaccard index: 1
PDPK1; 3-phosphoinositide dependent protein kinase 1
PAK4; p21 (RAC1) activated kinase 4
VAV3; vav guanine nucleotide exchange factor 3
CHUK; conserved helix-loop-helix ubiquitous kinase
IKBKG; inhibitor of kappa light polypeptide gene enhancer in
B-cells; kinase beta
MALT1; MALT1 paracaspase
PRKCQ; protein kinase C theta
IL2; interleukin 2
PAK5; p21 (RAC1) activated kinase 5
PAK3; p21 (RAC1) activated kinase 3
PAK2; p21 (RAC1) activated kinase 2
LCP2; lymphocyte cytosolic protein 2
CDC42; cell division cycle 42
LCK; LCK proto-oncogene, Src family tyrosine kinase
ZAP70; zeta chain of T cell receptor associated protein kinase
CTLA4; cytotoxic T-lymphocyte associated protein 4
TNF; tumor necrosis factor
MAPK1; mitogen-activated protein kinase 1
MAPK3; mitogen-activated protein kinase 3
IL5; interleukin 5
IL4; interleukin 4
SOS2; SOS Ras/Rho guanine nucleotide exchange factor 2
SOS1; SOS Ras/Rac guanine nucleotide exchange factor 1
PAK6; p21 (RAC1) activated kinase 6
T cell receptor signaling pathway
Cytokine-cytokine receptor interaction
Neurotrophin signaling pathway
dis: Diabetes mellitus (C0011849)
Jak-STAT signaling pathway
dis: Glucocorticoid-induced osteonecrosis (C0029445)
dis: Osteoporosis(C0029456)
HIF-1 signaling pathway
Type 1 diabetes mellitus
Wnt signaling pathway
Steroid hormone biosynthesis
Olfactory transduction
T cell receptor signaling pathway
Cytokine-cytokine receptor interaction
fluorometholone
B cell receptor signaling pathway
GnRH signaling pathway
Systemic lupus erythematosus

图18