



(12) 发明专利申请



(10) 申请公布号 CN 117813402 A

(43) 申请公布日 2024. 04. 02

(21) 申请号 202280036122.9

(22) 申请日 2022.03.17

(30) 优先权数据

63/163,414 2021.03.19 US

63/306,054 2022.02.02 US

(85) PCT国际申请进入国家阶段日

2023.11.17

(86) PCT国际申请的申请数据

PCT/US2022/020815 2022.03.17

(87) PCT国际申请的公布数据

W02022/197968 EN 2022.09.22

(71) 申请人 赛福医药公司

地址 美国马萨诸塞州

(72) 发明人 维亚切斯拉夫·R·埃克麦弗

西奥多·R·梅勒斯

(74) 专利代理机构 北京安信方达知识产权代理有限公司 11262

专利代理师 徐爱文 武晶晶

(51) Int.Cl.

G12Q 1/6883 (2006.01)

G16B 40/00 (2006.01)

G16B 25/00 (2006.01)

G16B 20/00 (2006.01)

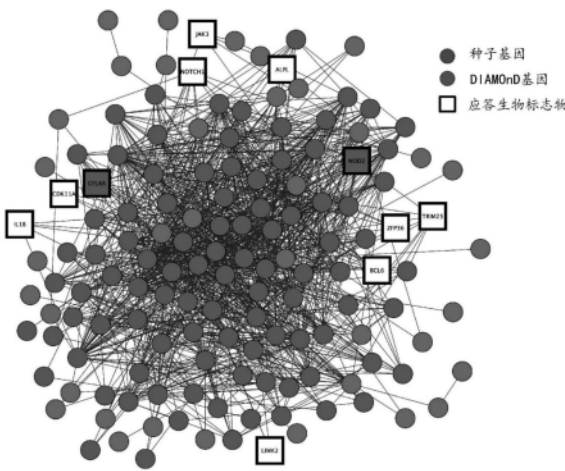
权利要求书4页 说明书49页 附图23页

(54) 发明名称

分类和治疗患者的方法

(57) 摘要

本文提供了用于开发可用于预测对特定治疗的应答的分类器的系统和方法。例如,在一些实施方案中,本公开提供了治疗患有自身免疫性病症的对象的方法,该方法包括:向已通过分类器确定为有应答的对象施用抗TNF疗法,所述分类器被建立用于区分队列中的已接受抗TNF疗法的有应答的先前对象和无应答的先前对象。例如,在一些实施方案中,本公开提供了在治疗性疗法期间治疗患有自身免疫性病症的对象的方法,该方法包括:在从施用抗TNF疗法开始的一段时间内鉴定有应答的先前对象和无应答的先前对象。



1. 一种治疗患有自身免疫性病症的对象的方法,所述方法包括:
向所述对象施用抗TNF疗法,其中所述对象已通过分类器被确定为有应答,所述分类器被建立用于区分队列中已接受所述抗TNF疗法的有应答的先前对象和无应答的先前对象;
其中所述分类器是通过评估以下项而开发的:
表达水平与临床应答性 or 无应答性显著相关的一种或多种基因;以及
以下项中的至少一者:
所述一种或多种基因的表达序列中一种或多种单核苷酸多态性 (SNP) 的存在;或
所述有应答的先前对象和所述无应答的先前对象的至少一种临床特征;并且
其中所述分类器由与已接受所述抗TNF疗法的队列不同的独立队列验证;并且
所述一种或多种基因包括:ALPL、ATRAID、BCL6、CDK11A、CFLAR、COMM5、GOLGA1、IL1B、IMPDH2、JAK3、KLHDC3、LIMK2、NOD2、NOTCH1、SPINT2、SPON2、STOML2、TRIM25或ZFP36。
2. 根据权利要求1所述的方法,其中所述对象先前已被施用所述抗TNF疗法。
3. 根据权利要求2所述的方法,其中在所述施用之前至少一个月、至少两个月、至少三个月、至少四个月、至少五个月或至少六个月,所述对象已被施用所述抗TNF疗法。
4. 根据权利要求3所述的方法,其中所述先前施用的抗TNF疗法不同于对所述分类器有应答而施用的所述抗TNF疗法。
5. 根据权利要求1所述的方法,其中所述分类器鉴定初治队列中60%或更多的无应答者。
6. 根据权利要求5所述的方法,其中所述分类器鉴定具有至少350名对象的初治队列中60%或更多的无应答者。
7. 根据权利要求1所述的方法,其中所述一种或多种基因的特征在于当映射到人类相互作用组图上时它们的拓扑性质。
8. 根据权利要求1所述的方法,其中所述SNP是参考人类基因组鉴定的。
9. 根据权利要求1所述的方法,其中所述一种或多种基因包括:ALPL、BCL6、CDK11A、CFLAR、IL1B、JAK3、LIMK2、NOD2、NOTCH1、TRIM25或ZFP36。
10. 根据权利要求1所述的方法,其中所述至少一种临床特征选自:体重指数 (BMI)、性别、年龄、种族、先前疗法治疗、疾病持续时间、C反应蛋白水平、抗环瓜氨酸肽的存在、类风湿因子的存在、患者总体评价、治疗应答率(例如,ACR20、ACR50、ACR70)及其组合。
11. 根据权利要求1所述的方法,其中所述抗TNF疗法包括施用英夫利昔单抗、阿达木单抗、依那西普、赛妥珠单抗、goliluma或其生物仿制药。
12. 根据权利要求1所述的方法,其中所述自身免疫性病症选自类风湿性关节炎、银屑病关节炎、强直性脊柱炎、克罗恩病、溃疡性结肠炎、慢性银屑病、化脓性汗腺炎、多发性硬化和幼年特发性关节炎。
13. 根据权利要求1所述的方法,其中所述分类器是使用源自所述有应答的先前对象和所述无应答的先前对象的微阵列分析来建立的。
14. 根据权利要求1所述的方法,其中所述SNP选自表3。
15. 根据权利要求1所述的方法,其中通过临床特征的统计分析在对象中验证应答。
16. 根据权利要求15所述的方法,其中所述临床特征的统计分析分析在接受抗TNF疗法后临床特征的变化。

17. 根据权利要求15所述的方法, 其中所述临床特征的统计分析分析ACR50、ACR70、CDAI LDA、CDAI缓解、DAS28-CRP LDA或DAS28-CRP缓解中一种或多种的变化。

18. 根据权利要求15所述的方法, 其中所述统计分析是蒙特卡罗分析。

19. 根据权利要求1所述的方法, 其中所述分类器包括以下所有基因和临床特征: ALPL、ATRAID、BCL6、CDK11A、CFLAR、COMMD5、GOLGA1、IL1B、IMPDH2、JAK3、KLHDC3、LIMK2、NOD2、NOTCH1、SPINT2、SPON2、STOML2、TRIM25、ZFP36、BMI、性别、患者总体状况、评估和抗CCP。

20. 根据权利要求1所述的方法, 其中所述方法是自动化的计算机实现的方法。

21. 一种编码有计算机程序的非暂时性计算机可读存储介质, 所述计算机程序包括能够由至少一个处理器执行以创建改进的样品分类应用的指令, 所述非暂时性计算机可读存储介质包括:

存储在所述介质上的至少一个分类器, 其中所述分类器能够区分对抗TNF疗法有应答的对象和无应答的对象;

其中所述分类器是通过评估以下项而开发的:

表达水平与对抗TNF疗法的临床应答性 or 无应答性显著相关的一种或多种基因;

以下项中的至少一者:

所述一种或多种基因的表达序列中一种或多种单核苷酸多态性 (SNP) 的存在; 或

所述有应答的先前对象和所述无应答的先前对象的至少一种临床特征; 并且

其中所述分类器由与已接受所述抗TNF疗法的队列不同的独立队列验证; 并且

所述一种或多种基因包括: ALPL、ATRAID、BCL6、CDK11A、CFLAR、COMMD5、GOLGA1、IL1B、IMPDH2、JAK3、KLHDC3、LIMK2、NOD2、NOTCH1、SPINT2、SPON2、STOML2、TRIM25或ZFP36。

22. 根据权利要求21所述的非暂时性计算机可读存储介质, 其还包括被配置为接收从来自对象的血液样品衍生的基因表达数据的软件模块。

23. 根据权利要求22所述的非暂时性计算机可读存储介质, 其还包括用于将所述分类器应用于所述基因表达数据的软件模块。

24. 根据权利要求23所述的非暂时性计算机可读存储介质, 其还包括用于使用所述分类器输出对所述样品的分类的软件模块, 其中所述分类将所述血液样品分类为是来自对抗TNF疗法有应答的对象, 还是来自对抗TNF疗法无应答的对象。

25. 根据权利要求21所述的非暂时性计算机可读存储介质, 其中所述对象先前已被施用所述抗TNF疗法。

26. 根据权利要求25所述的非暂时性计算机可读存储介质, 其中在所述施用之前至少一个月、至少两个月、至少三个月、至少四个月、至少五个月或至少六个月, 所述对象已被施用所述抗TNF疗法。

27. 根据权利要求21所述的非暂时性计算机可读存储介质, 其中所述分类器鉴定初治队列中60%或更多的无应答者。

28. 根据权利要求27所述的非暂时性计算机可读存储介质, 其中所述分类器鉴定具有至少350名对象的初治队列中60%或更多的无应答者。

29. 根据权利要求21所述的非暂时性计算机可读存储介质, 其中所述一种或多种基因的特征在于当映射到人类相互作用组图上时它们的拓扑性质。

30. 根据权利要求21所述的非暂时性计算机可读存储介质, 其中所述SNP是参考人类基

基因组鉴定的。

31. 根据权利要求21所述的非暂时性计算机可读存储介质,其中所述一种或多种基因包括:ALPL、BCL6、CDK11A、CFLAR、IL1B、JAK3、LIMK2、NOD2、NOTCH1、TRIM25或ZFP36。

32. 根据权利要求21所述的非暂时性计算机可读存储介质,其中所述至少一种临床特征选自:体重指数(BMI)、性别、年龄、种族、先前疗法治疗、疾病持续时间、C反应蛋白水平、抗环瓜氨酸肽的存在、类风湿因子的存在、患者总体评价、治疗应答率(例如,ACR20、ACR50、ACR70)及其组合。

33. 根据权利要求21所述的非暂时性计算机可读存储介质,其中所述抗TNF疗法包括施用英夫利昔单抗、阿达木单抗、依那西普、赛妥珠单抗、goliluma或其生物仿制药。

34. 根据权利要求21所述的非暂时性计算机可读存储介质,其中所述分类器是使用源自所述有应答的先前对象和所述无应答的先前对象的微阵列分析来建立的。

35. 根据权利要求21所述的非暂时性计算机可读存储介质,其中所述SNP选自表3。

36. 根据权利要求21所述的非暂时性计算机可读存储介质,其中通过临床特征的统计分析在对象中验证应答。

37. 根据权利要求36所述的非暂时性计算机可读存储介质,其中所述临床特征的统计分析分析在接受抗TNF疗法后临床特征的变化。

38. 根据权利要求36所述的非暂时性计算机可读存储介质,其中所述临床特征的统计分析分析ACR20、ACR50、ACR70、CDAILDA、CDAI缓解、DAS28-CRP LDA或DAS28-CRP缓解中一种或多种的变化。

39. 根据权利要求36所述的非暂时性计算机可读存储介质,其中所述统计分析是蒙特卡罗分析。

40. 根据权利要求21所述的非暂时性计算机可读存储介质,其中所述分类器包括任何以下基因或临床特征:ALPL、ATRAID、BCL6、CDK11A、CFLAR、COMMD5、GOLGA1、IL1B、IMPDH2、JAK3、KLHDC3、LIMK2、NOD2、NOTCH1、SPINT2、SPON2、STOML2、TRIM25、ZFP36、BMI、性别、患者总体状况、评估或抗CCP。

41. 一种治疗患有自身免疫性病症的对象的方法,所述方法包括:

通过分类器将所述对象确定为有应答,所述分类器被建立用于区分队列中已接受抗TNF疗法的有应答的先前对象和无应答的先前对象;

其中所述分类器是通过评估以下项而开发的:

表达水平与临床应答性或无应答性显著相关的一种或多种基因;以及

以下项中的至少一者:

所述一种或多种基因的表达序列中一种或多种单核苷酸多态性(SNP)的存在;或

所述有应答的先前对象和所述无应答的先前对象的至少一种临床特征;并且

其中所述分类器由与已接受所述抗TNF疗法的队列不同的独立队列验证;并且

所述一种或多种基因包括:ALPL、ATRAID、BCL6、CDK11A、CFLAR、COMMD5、GOLGA1、IL1B、IMPDH2、JAK3、KLHDC3、LIMK2、NOD2、NOTCH1、SPINT2、SPON2、STOML2、TRIM25或ZFP36。

42. 根据权利要求41所述的方法,其中所述对象先前已被施用所述抗TNF疗法。

43. 根据权利要求42所述的方法,其中在所述施用之前至少一个月、至少两个月、至少三个月、至少四个月、至少五个月或至少六个月,所述对象已被施用所述抗TNF疗法。

-
44. 一种试剂盒,其包含根据权利要求21所述的非暂时性计算机可读存储介质。
45. 根据权利要求44所述的试剂盒,还包括描述如何执行所述分类器的指令。

分类和治疗患者的方法

[0001] 相关申请的交叉引用

[0002] 本申请要求2021年3月19日提交的美国临时申请第63/163,414号和2022年2月2日提交的美国临时申请第63/306,054号的权益,这些申请的全部内容通过引用整合到本文中。

背景技术

[0003] 类风湿性关节炎(RA)等自身免疫性疾病影响着数以百万计的患者,其治疗费用在整个医疗保健支出中占很大的比重。自身免疫性疾病可以分为两类,即器官特异性和系统性自身免疫。包括RA在内的类风湿疾病属于系统性自身免疫疾病,主要表现在滑膜关节中,最终导致肌腱、软骨和骨骼的不可逆破坏。虽然目前还没有治愈RA的方法,但主要通过开发用于中和这种细胞因子的促炎信号传导的抗TNF(肿瘤坏死因子)药物,这些患者的治疗管理已经取得了显著的进展。这类生物疗法(例如,

Humira[®]、Enbrel[®]、Remicade[®]、Simponi[®]和Cimzia[®])显著改善了一些RA患者的治疗结果。

[0004] 大约34%的RA患者(低百分比)表现出对抗TNF疗法的临床应答,达到低疾病活动度(LDA),有时甚至达到缓解。在这些所谓“应答者”患者中,疾病进展可能是不适当的TNF驱动的促炎应答的结果。对于抗TNF无应答的患者,可以采用抗CD20、共刺激阻断剂、JAK和抗IL6疗法等已获批准的替代疗法。然而,患者可能会在第一次循环使用不同的抗TNF后才切换到这种替代疗法,该循环可能需要一年多的时间,同时症状持续存在且疾病进一步发展,使其更难达到治疗目标。除了延误治疗的问题以外,伴随抗TNF疗法的严重感染和恶性肿瘤的风险非常大,以至于产品批准时可能需要在标签上包含所谓的“黑框警告”。这种疗法的其他潜在副作用包括例如充血性心力衰竭、脱髓鞘疾病和其他全身性副作用。

发明内容

[0005] 抗TNF疗法的一个重要问题是应答率不一致。无论用于定义应答的方法如何,一部分RA患者可能会对TNFi治疗产生足够的应答,50-70%达到ACR20,30-40%达到ACR50,15-25%达到ACR70应答,且10-25%达到缓解。许多研究试图确定生物标志物并开发模型,以便在治疗开始前预测针对TNFi疗法的应答。在新的患者群体和临床试验中,未能验证和再现这些预测生物标志物的性能是一个典型的结果。患者群体之间的不同特征、生成分子数据的实验室方法和程序以及单队列回顾性血液研究固有的其他偏差不仅阻碍了风湿病学的精准医学进展,也阻碍了其他医学专业的精准医学进展。

[0006] 在一些方面,本文所述的方法和组合物允许护理提供者区分不同类别的对象,例如,可能在特定治疗(例如,抗TNF疗法)中受益的对象与没有受益的对象、更有可能实现或遭受特定结果或副作用的对象等。在一些实施方案中,所提供的这类技术由此降低了患者的风险,增加了对无应答者患者群体的护理时机和质量,增加了药物开发的效率,或避免了对无应答者患者施用无效疗法或治疗这些患者在接受相关疗法(例如,抗TNF疗法)时出现

的副作用所产生的成本。

[0007] 在一些方面,本公开提供了用特定疗法(例如,抗TNF疗法)治疗对象的方法,在一些实施方案中,该方法包括:向已通过分类器确定为有应答的对象施用疗法,所述分类器被建立用于区分预期对该疗法有应答的对象和无应答的对象。在一些实施方案中,分类器鉴定初治队列中60%或更多的无应答者。在一些实施方案中,分类器鉴定具有至少350名对象的初治队列中60%或更多的无应答者。

[0008] 分类器可以是源自队列内已知的应答者和无应答者之间的基因表达差异的分子特征应答分类器。在一些实施方案中,包括在应答者和无应答者之间的表达具有统计学显著差异的一种或多种基因作为分子特征应答分类器的一部分。在一些实施方案中,将与在应答者和无应答者之间的表达具有统计学显著差异的基因相关的蛋白质映射到人类相互作用组上,以验证所选基因和疾病生物学之间的关系。

[0009] 所提供的分类器还整合了其他元素,例如,可用于对给定患者中的应答或非应答进行分类的临床特征或单核苷酸多态性。

[0010] 在一些实施方案中,本公开提供了治疗患有自身免疫性病症的对象的方法,在一些实施方案中,该方法包括:向已通过分类器确定为有应答的对象施用抗TNF疗法,所述分类器被建立用于区分队列中的已接受抗TNF疗法的有应答的先前对象和无应答的先前对象;其中分类器是通过评估以下项而开发的:表达水平与临床应答性或无应答性显著相关(例如,以线性或非线性方式)的一种或多种基因;以下项中的至少一者:所述一种或多种基因的表达序列中一种或多种单核苷酸多态性(SNP)的存在;或所述有应答的先前对象和所述无应答的先前对象的至少一种临床特征;并且其中所述分类器由与已接受所述抗TNF疗法的队列不同的独立队列验证。

[0011] 在一些实施方案中,所述对象先前已被施用抗TNF疗法。在一些实施方案中,在所述施用之前至少一个月、至少两个月、至少三个月、至少四个月、至少五个月或至少六个月,所述对象已被施用抗TNF疗法。

[0012] 在一些实施方案中,分类器鉴定初治队列中60%或更多的无应答者。在一些实施方案中,分类器鉴定具有至少350名对象的初治队列中60%或更多的无应答者。

[0013] 在一些实施方案中,一种或多种基因的特征在于当映射到人类相互作用组图上时它们的拓扑性质。在一些实施方案中,参考人类基因组来鉴定SNP。在一些实施方案中,通过评估以下每一种来开发分类器:表达水平与临床应答性或无应答性显著相关(例如,以线性或非线性方式)的一种或多种基因;一种或多种SNP的存在;以及至少一种临床特征。

[0014] 在一些实施方案中,一种或多种基因包括:ALPL、ATRAID、BCL6、CDK11A、CFLAR、COMMD5、GOLGA1、IL1B、IMPDH2、JAK3、KLHDC3、LIMK2、NOD2、NOTCH1、SPINT2、SPON2、STOML2、TRIM25或ZFP36。

[0015] 在一些实施方案中,一种或多种基因包括:ALPL、BCL6、CDK11A、CFLAR、IL1B、JAK3、LIMK2、NOD2、NOTCH1、TRIM25或ZFP36。

[0016] 在一些实施方案中,至少一种临床特征选自:体重指数(BMI)、性别、年龄、种族、先前疗法治疗、疾病持续时间、C反应蛋白水平、抗环瓜氨酸肽的存在、类风湿因子的存在、患者总体评价、治疗应答率(例如,ACR20、ACR50、ACR70)及其组合。

[0017] 在一些实施方案中,抗TNF疗法包括施用英夫利昔单抗(infliximab)、阿达木单抗

(infliximab)、依那西普(etanercept)、赛妥珠单抗(cirtolizumab pegol)、golilumab或其生物仿制药。在一些实施方案中,疾病、病症或病况选自类风湿性关节炎、银屑病关节炎、强直性脊柱炎、克罗恩病、溃疡性结肠炎、慢性银屑病、化脓性汗腺炎、多发性硬化和幼年特发性关节炎。在一些实施方案中,利用源自应答性的先前对象和无应答性的先前对象的微阵列分析来建立分类器。在一些实施方案中,使用源自独立队列的RNAseq数据来验证分类器。在一些实施方案中,SNP选自表3。

[0018] 在一些实施方案中,本公开提供了一种系统,该系统用于将患有自身免疫性疾病的对象分类为对抗TNF疗法可能有应答或可能无应答,之后对所述对象进行所述抗TNF疗法的任何施用,所述系统包括:处理器;以及存储器,所述存储器上具有指令,所述指令在由所述处理器执行时使所述处理器:a)接收一组数据,所述一组数据包括对象的一种或多种基因中每种基因的表达水平,这些基因包括ALPL、ATRAID、BCL6、CDK11A、CFLAR、COMMD5、GOLGA1、IL1B、IMPDH2、JAK3、KLHDC3、LIMK2、NOD2、NOTCH1、SPINT2、SPON2、STOML2、TRIM25或ZFP36。

[0019] 对于本领域的技术人员来说,本公开的其他方面和优点通过以下的详细描述将变得显而易见,其中仅示出和描述了本公开的说明性实施方案。正如所意识到的,本公开能够有其他和不同的实施方案,且其若干细节能够在各种明显的方面进行修改,所有这些都不会脱离本公开。因此,以下附图和描述在本质上来说应被视为是说明性的、而不是限制性的。

[0020] 引用并入

[0021] 本说明书中提及的所有出版物、专利和专利申请均以引用的方式并入本说明书中,其引用程度等同于每份出版物、专利或专利申请均以引用的方式被具体和单独地并入本说明书中。如果通过引用并入的出版物和专利或专利申请与本说明书中包含的公开内容相矛盾,本说明书应当取代或优先于任何此类的矛盾材料。

附图说明

[0022] 图1是将预测应答的转录物编码的蛋白质映射到人类相互作用组的示例性实施方案。蛋白质用圆圈表示,成对的物理蛋白质-蛋白质相互作用用线条表示。RA疾病模块由种子基因(红色)和DIAMOND基因(茶色)组成。11个转录物特征(方块)编码的蛋白质与RA疾病模块显著相关(p 值 <0.05)。

[0023] 图2A、图2B、图2C和图2D展示了分子特征应答分类器(“MSRC”)在来自Corrona CERTAIN研究的245名患者中进行的交叉验证。图2A展示了基于CDAI、DAS28-CRP、ACR70和ACR50临床结果进行的患者分层的对象工作曲线。图2B展示了具有或不具有无应答分子特征的患者模型评分的比较。方框和交叉线分别表示四分位间距和中位数。平分彩色线表示平均值的变化。在CDAI缓解、LDA、中度或高度疾病活动中具有或不具有无应答分子特征的患者所占的百分比比率如图2C(CDAI)和图2D(DAS28-CRP)所示。柱形表示高于1.0时具有分子特征或低于1.0时不具有分子特征的患者的较大比例。NA:不适用,类别中没有患者。

[0024] 图3A、图3B、图3C、图3D、图3E和图3F展示了MSRC的验证,以鉴定对TNFi疗法不太可能有应答的未经靶向疗法的患者。基于在3个月(图3A)和6个月(图3B)时CDAI、DAS28-CRP、ACR70和ACR50临床结果进行的患者分层的对象工作曲线图。具有或不具有无应答分子特征

的患者在3个月(图3C)和6个月(图3D)时的模型评分的比较。方框和交叉线分别表示四分位间距和中位数。平分彩色线表示平均值的变化。根据CDAI(图3E)和DAS28-CRP(图3F)的CDAI缓解、LDA、中度或高度疾病活动中具有或不具有无应答分子特征的患者所占的百分比比率。柱形表示高于1.0时具有分子特征或低于1.0时不具有分子特征的患者的较大比例。NA:不适用,类别中没有患者;NS:不显著。

[0025] 图4A和图4B展示了MSRC的验证,以鉴定对TNFi疗法不太可能有应答的暴露于TNFi的患者。图4A示出了基于测试结果后3个月的CDAI缓解或DAS28-CRP缓解的实现情况,对接受TNFi疗法的患者进行分层的对象工作曲线。图4B示出了具有或不具有无应答分子特征的患者模型评分的比较。方框和交叉线分别表示四分位间距和中位数。平分彩色线表示平均值的变化。

[0026] 图5展示了对TNFi疗法不充分应答的生物学。MSRC包括编码参与RA病理生理学的众多方面的蛋白质的转录物:先天免疫应答、细胞因子生物合成、T和B细胞稳态、骨稳态、解折叠蛋白应答、自噬、凋亡和促炎信号传导。

[0027] 图6是研究设计的流程图。对来自CERTAIN研究的345名患者的子集进行了分析:100名用于鉴定对TNFi疗法无应答的转录物生物标志物,245名用于交叉验证。273名入组了NETWORK-004前瞻性观察研究;244名通过了初步入组筛选,194名完成了3个月的随访,168名完成了6个月的随访。87%(146/168)完成研究的患者具有进行验证分析所需的完整分子和临床数据。

[0028] 图7是维恩图,显示了在暴露于TNF疗法3个月、6个月以及在3个月和6个月两者时提供样品的患者的分类情况。

[0029] 图8A、图8B、图8C和图8D提供了ROC曲线,显示了在TNF开始后3个月和6个月收集的患者样品中的PrismRA性能。图8A显示了使用+3个月结果的3个月样品。图8B显示了使用+6个月结果的3个月样品。图8C显示了使用+3个月结果的6个月样品。图8D显示了使用+6个月结果的6个月样品。

[0030] 图9A和图9B提供了ROC曲线,显示了提供3个月和6个月样品的122名患者中的模型性能。图9A显示了使用+6个月终点的3个月样品。图9B显示了使用+3个月终点的6个月样品。

[0031] 图10提供了示例计算机系统,用于执行根据本公开的一些方面或实施方案的方法。

具体实施方式

[0032] 各种疗法(例如,抗TNF疗法)的一个重要问题是应答率不一致。事实上,最近召开的国际会议旨在将免疫学和风湿病学领域的顶尖科学家和临床医生聚集在一起,以确定这些领域尚未满足的需求,会议几乎普遍认为应答率的不确定性是一个持续性的挑战。例如,第19届国际靶向疗法年会举行了关于多种疾病治疗挑战的分组会议,包括类风湿性关节炎、银屑病关节炎、中轴性脊柱关节炎、系统性红斑狼疮和结缔组织疾病(例如干燥综合征、系统性硬化、血管炎,包括Bechet和IgG4相关疾病),确定了所有这些疾病共同面临的某些问题,尤其是“需要更好地了解每种疾病的异质性…以便可以开发治疗应答的预测工具”。参见Winthrop,等人,“The unmet need in rheumatology:Reports from the targeted therapies meeting 2017,”Clin.Immunol.pii:S1521-6616(17)30543-0,Aug.12,2017,其

内容通过引用并入本文中以备用于所有目的。类似地,与抗肿瘤坏死因子疗法治疗克罗恩病相关的大量文献也一直在抱怨应答率不稳定,以及无法预测哪些患者将从中受益。例如,参见M.T.Abreu,“Anti-TNF Failures in Crohn’s Disease,”*Gastroenterol Hepatol* (NY), 7 (1):37-39 (Jan.2011),另参见Ding等人,“Systematic review:predicting and optimising response to anti-TNF therapy in Crohn’s disease-algorithm for practical management,”*Aliment Pharmacol.Ther.*, 43(1):30-51 (Jan.2016),这些文献的内容通过引用并入本文中以备用于所有目的(报告称“抗TNF疗法的最初无应答影响着13-40%的患者”)。

[0033] 因此,目前正在接受抗TNF疗法的大量患者并没有从治疗中获益,甚至可能受到伤害。伴随抗TNF疗法的严重感染和恶性肿瘤的风险非常大,以至于产品批准时可能需要在标签上注明所谓的“黑框警告”。这类治疗的其他潜在副作用包括例如充血性心力衰竭、脱髓鞘疾病和其他全身性副作用。此外,鉴于患者在被鉴定为对抗TNF疗法无应答(例如,为抗TNF疗法的无应答者)之前需要数周至数月的治疗,由于目前无法鉴定有应答与无应答的对象,对这类患者的适当治疗可能会大大延迟。例如,参见Roda等人,“Loss of Response to Anti-TNFs:Definition,Epidemiology,and Management,”*Clin.Trani.Gastroenterol.*, 7 (1):e135 (Jan.2016),其内容通过引用并入本文中以备用于所有目的(引用了Hanauer等人,“ACCENT I Study group.Maintenance Infliximab for Crohn’s disease:the ACCENT I randomized trial,”*Lancet* 59:1541-1549 (2002);Sands等人,“Infliximab maintenance therapy for fistulizing Crohn’s disease,”*N.Engl.J.Med.* 350:876-885 (2004))。

[0034] 因此,在一些实施方案中,本公开提供了用抗TNF疗法治疗对象的方法,该方法包括:向通过分类器确定为有应答的患者施用抗TNF疗法,所述分类器被建立用于区分队列中的已接受抗TNF疗法的有应答的先前对象和无应答的先前对象,其中所述分类器是通过评估以下项而开发的:表达水平与临床应答性 or 无应答性显著相关(例如,以线性或非线性方式)的一种或多种基因;以及以下项中的至少一者:所述一种或多种基因的表达序列中一种或多种单核苷酸多态性(SNP)的存在;或所述有应答的先前对象和所述无应答的先前对象的至少一种临床特征。

[0035] 本文提供了用于自动预测对象对抗TNF疗法的应答的系统和方法。本文还提供了用于自动解读基因组或多组学数据的模块化系统。

[0036] 如本文所用,术语“施用”通常是指将组合物施用到对象或系统,例如,以实现作为组合物或包含在组合物中或以其他方式通过组合物递送的药剂的递送。

[0037] 如本文所用,术语“药剂”是通常指实体(例如,脂质、金属、核酸、多肽、多糖、小分子等,或其复合物、组合、混合物或系统[例如,细胞、组织、生物体])或现象(例如,热、电流或电场、磁力或磁场等)。

[0038] 如本文所用,术语“氨基酸”通常是指可以例如通过形成一个或多个肽键而掺入多肽链的任何化合物或物质。在一些实施方案中,氨基酸具有一般结构 $H_2N-C(H)(R)-COOH$ 。在一些实施方案中,氨基酸是天然存在的氨基酸。在一些实施方案中,氨基酸是非天然氨基酸;在一些实施方案中,氨基酸是D-氨基酸;在一些实施方案中,氨基酸是L-氨基酸。如本文所用,术语“标准氨基酸”是指在天然存在的肽中常见的二十种L-氨基酸中的任一种。“非标

准氨基酸”是指除标准氨基酸以外的任何氨基酸,无论其是否可以存在于天然来源中。在一些实施方案中,与以上一般结构相比,包括多肽中的羧基或氨基末端氨基酸在内的氨基酸可以包含结构修饰。例如,在一些实施方案中,与一般结构相比,氨基酸可通过甲基化、酰胺化、乙酰化、聚乙二醇化、糖基化、磷酸化或取代(例如,氨基、羧基基团、一个或多个质子或羟基)进行修饰。在一些实施方案中,与含有原本相同的未修饰氨基酸的多肽相比,此类修饰可例如改变含有修饰氨基酸的多肽的稳定性或循环半衰期。在一些实施方案中,与含有原本相同的未修饰氨基酸的多肽相比,此类修饰不显著改变含有修饰氨基酸的多肽的相关活性。如从上下文中可以清楚地看到,在一些实施方案中,术语“氨基酸”可用于指代游离氨基酸;在一些实施方案中,其可用于指代多肽的氨基酸残基,例如,多肽内的氨基酸残基。如本文所用,术语“类似物”通常是指与参考物质共享一个或多个特定结构特征、元素、组分或部分的物质。通常,“类似物”显示出与参考物质的显著结构相似性,例如共享核或共有结构,但也以某些离散方式不同。在一些实施方案中,类似物是可以从参考物质生成的物质,例如通过对参考物质的化学操纵。在一些实施方案中,类似物是可以通过执行与生成参考物质的合成过程基本上类似(例如,共享多个步骤)的合成过程生成的物质。在一些实施方案中,类似物是或可以通过执行不同于用于生成参考物质的合成过程生成的。

[0039] 如本文所用,术语“拮抗剂”通常可指其存在、水平、程度、类型或形式与靶标的水平或活性降低相关联的药剂或状况。拮抗剂可包括任何化学类别的药剂,包括例如小分子、多肽、核酸、碳水化合物、脂质、金属或显示出相关抑制活性的任何其他实体。在一些实施方案中,拮抗剂可以是“直接拮抗剂”,因为其与其靶标直接结合;在一些实施方案中,拮抗剂可以是“间接拮抗剂”,因为其通过除直接结合其靶标之外的机制施加其影响;例如,通过与靶标的调节剂相互作用,从而改变靶标的水平或活性)。在一些实施方案中,“拮抗剂”可称为“抑制剂”。

[0040] 如本文所用,术语“抗体”通常是指包括足以赋予与特定靶抗原的特异性结合的经典免疫球蛋白序列元件的多肽。在一些实施方案中,在自然界中产生的完整抗体是大约150kD的四聚体剂,由两个相同的重链多肽(每个约50kD)和两个相同的轻链多肽(每个约25kD)组成,它们相互缔合成通常称为“Y形”结构的物质。在一些实施方案中,每条重链由至少四个结构域(每个长约110个氨基酸)组成——一个氨基末端可变(VH)结构域(位于Y结构的顶端),接着是三个恒定结构域:CH1、CH2和羧基末端CH3(位于Y的茎的底部)。在一些实施方案中,称为“开关”的短区域连接重链可变区和恒定区。“铰链”将CH2和CH3结构域连接到抗体的其余部分。在一些实施方案中,该铰链区中的两个二硫键将两个重链多肽连接到完整抗体中的另一个。在一些实施方案中,每个轻链由两个结构域组成——氨基末端可变(VL)结构域,接着是羧基末端恒定(CL)结构域,由另一个“开关”彼此分开。在一些实施方案中,完整的抗体四聚体由两个重链-轻链二聚体组成,其中重链和轻链通过单个二硫键相互连接;另外两个二硫键将重链铰链区相互连接,从而使二聚体相互连接并形成四聚体。在一些实施方案中,天然产生的抗体也被糖基化,诸如在CH2结构域上。在一些实施方案中,天然抗体中的每个结构域都具有这样一种结构:其特征是由在压缩的反平行 β 桶(beta barrel)中彼此紧密堆积的两个 β 折叠(例如,3、4或5链折叠)形成的“免疫球蛋白折叠”。在一些实施方案中,每个可变结构域包含三个称为“互补决定区”的高变环(CDR1、CDR2和CDR3)和四个在一定程度上不变的“框架”区(FR1、FR2、FR3和FR4)。在一些实施方案中,当天然抗体折叠

时,FR区形成为结构域提供结构框架的 β 折叠,并且使得来自重链和轻链的CDR环区在三维空间中聚集在一起,使得它们形成位于Y结构顶端的单个高变抗原结合位点。在一些实施方案中,天然存在的抗体的Fc区结合到补体系统的元件,并且还结合到效应细胞上的受体,包括例如介导细胞毒性的效应细胞。在一些实施方案中,Fc区对Fc受体的亲和力或其他结合属性可以通过糖基化或其他修饰进行调节。在一些实施方案中,根据本公开产生或利用的抗体包括糖基化Fc结构域,包括具有修饰或工程化的此类糖基化的Fc结构域。出于本公开的目的,在某些实施方案中,包括天然抗体中发现的足够免疫球蛋白结构域序列的任何多肽或多肽复合物可以被称为或用作“抗体”,无论所述多肽是天然产生的(例如,由生物体对抗原反应产生的),还是通过重组工程化、化学合成或其他人工系统或方法产生的。在一些实施方案中,抗体是多克隆的;在一些实施方案中,抗体是单克隆的。在一些实施方案中,抗体具有小鼠、兔、灵长类动物或人类抗体所特有的恒定区序列。在一些实施方案中,抗体序列元件为人源化的、灵长类化的、嵌合的等。此外,如本文所用,术语“抗体”可以在适当的实施方案中(除非另有说明或由上下文清楚)指用于在替代呈现中利用抗体结构和功能特征的任何构建体或形式。例如,在一些实施方案中,根据本公开使用的抗体的形式选自但不限于完整IgA、IgG、IgE或IgM抗体;双特异性或多特异性抗体(例如,Zybodies[®]等);抗体片段,诸如Fab片段、Fab'片段、F(ab')₂片段、Fd'片段、Fd片段和分离的CDR或其集合;单链Fv;多肽-Fc融合体;单结构域抗体(例如,鲨鱼单结构域抗体,诸如IgNAR或其片段);骆驼样抗体;隐蔽抗体(例如,Probodyes[®]);小型模块化免疫药物(Small Modular ImmunoPharmaceuticals,“SMIPsTM”);单链或串联双体(TandAb[®]);VI-11-1s;Anticalins[®]; Nanobodies[®]微型抗体;BiTE[®];锚蛋白重复蛋白或DARPINS[®]; Avimers[®];DART;TCR样抗体;Adnectins[®]; Afflins[®]; Trans-bodies[®]; Affibodies[®]; TrimerX[®];微蛋白(MicroProteins);Fynomers[®], Centyrins[®];以及KALBITOR[®]。在一些实施方案中,抗体可能缺少如果其自然产生将具有的共价修饰(例如,附接聚糖)。在一些实施方案中,抗体可包含共价修饰(例如,附接聚糖、有效载荷[例如,可检测部分、治疗部分、催化部分等]或其他侧基[例如,聚乙二醇等])。

[0041] 如果一个事件或实体的存在、水平、程度、类型或形式与另一个相关,则两个事件或实体通常彼此“关联”,如所述术语在本文所使用的。例如,如果特定实体(例如,多肽、基因特征(genetic signature)、代谢物、微生物等)的存在、水平或形式与疾病、病症或病况的发生率或易感性(例如,在相关群体中)相关,则认为所述实体与特定疾病、病症或病况相关联。在一些实施方案中,如果两个或更多个实体直接或间接地相互作用,使得它们彼此物理接近或保持彼此物理接近,则它们彼此在物理上“关联”。在一些实施方案中,彼此物理关联的两个或更多个实体彼此共价连接;在一些实施方案中,彼此物理关联的两个或更多个实体不是彼此共价连接,而是非共价缔合,例如借助于氢键、范德华相互作用、疏水相互作用、磁性及其组合的机制。

[0042] 如本文所用,术语“生物样品”通常是指如本文所述从感兴趣的生物来源(例如,组织或生物体或细胞培养物)获得或衍生的样品。在一些实施方案中,感兴趣的来源包括生物

体,诸如动物或人类。在一些实施方案中,生物样品是或包括生物组织或流体。在一些实施方案中,生物样品可以是或包括骨髓;血液;血细胞;腹水;组织或细针活检样品;含有细胞的体液;游离漂浮核酸;痰;唾液;尿;脑脊液、腹腔液;胸膜液;粪便;淋巴;妇科流体;皮肤拭子;阴道拭子;口腔拭子;鼻拭子;冲洗液或灌洗液,诸如导管灌洗液或支气管肺泡灌洗液;抽吸物;刮屑;骨髓样本;组织活检样本;外科样本;粪便、其他体液、分泌物或排泄物;或其中的细胞等。在一些实施方案中,生物样品是或包括从个体获得的细胞。在一些实施方案中,获得的细胞是或包括来自获得样品的个体的细胞。在一些实施方案中,样品是通过任何适当的手段直接从感兴趣的来源获得的“原始样品”。例如,在一些实施方案中,原始生物样品通过选自以下的方法获得:活检(例如,细针穿刺或组织活检)、手术、收集体液(例如,血液、淋巴、粪便等)等。在一些实施方案中,如从上下文中可以清楚地看出,术语“样品”是指通过处理原始样品(例如,通过去除其一种或多种组分或通过向其中添加一种或多种药剂)而获得的制备物。例如,使用半渗透膜过滤。这样的“处理样品”可包括,例如,从样品中提取或通过使原始样品经受诸如mRNA的扩增或逆转录、某些组分的分离或纯化等技术而获得的核酸或蛋白质。

[0043] 如本文所用,术语“联合疗法”通常是指对象同时暴露于两种或更多种治疗方案(例如,两种或更多种治疗剂)的临床干预。在一些实施方案中,可同时施用两种或更多种治疗方案。在一些实施方案中,可顺序施用两种或更多种治疗方案(例如,在施用任何剂量的第二方案之前施用第一方案)。在一些实施方案中,以重叠给药方案施用两种或更多种治疗方案。在一些实施方案中,组合疗法的施用可涉及向接受另外一种或多种药剂或模式的对象施用一种或多种治疗剂或模式。在一些实施方案中,组合疗法不一定要单独药剂在单一组合物中(或甚至不一定同时)一起施用。在一些实施方案中,通过单独的施用途径(例如,一种药剂口服施用,并且另一种药剂静脉内施用)或在不同的时间点,单独(例如,以单独的组合物形式)向对象施用组合疗法的两种或更多种治疗剂或模式。在一些实施方案中,可通过相同的施用途径或同时在组合组合物中或甚至在组合化合物中(例如,作为单一化学复合物或共价实体的一部分)一起施用两种或更多种治疗剂。

[0044] 如本文所用,术语“相当的”通常是指两种或更多种药剂、实体、情况、条件集等可能彼此不相同,但足够相似以允许相互进行比较,使得可基于观察到的差异或相似性合理地得出结论。在一些实施方案中,相当的条件集、环境、个体或群体的特征在于多个基本上相同的特征和一个或少量的变化特征。可理解,在上下文中,在任何给定情况下,两种或更多种这样的药剂、实体、情况、条件集等被认为是相当的所需要的同一性程度。例如,当通过足够数量和类型的基本上相同的特征表征以保证以下合理结论时,环境、个体或群体的集合可以是彼此相当的:在不同的环境、个体或群体集合下或使用所述不同的环境、个体或群体集合获得的结果或观察到的现象中的差异是由那些发生变化的特征的变化引起的或指示所述特征的变化。

[0045] 如本文所用,短语“对应于”通常是指两个实体、事件或现象之间的关系,这两个实体、事件或现象共享足以可合理比较的特征,使得“对应”属性是明显的。例如,在一些实施方案中,该术语可参考化合物或组合物使用,以通过与适当的参考化合物或组合物进行比较来指定化合物或组合物中结构元件的位置或身份。例如,在一些实施方案中,聚合物中的单体残基(例如,多肽中的氨基酸残基或多核苷酸中的核酸残基)可被鉴定为“对应于”适当

的参考聚合物中的残基。例如,普通技术人员将理解,为了简单起见,通常基于参考相关多肽使用规范编号系统来指定多肽中的残基,使得例如,“对应于”位置190处的残基的氨基酸实际上不需要是特定氨基酸链中的第190个氨基酸,而是对应于参考多肽中190处的残基;多种途径可用于鉴定“对应”氨基酸。例如,存在各种序列比对策略,包括软件程序,例如BLAST、CS-BLAST、CUSASW++、DIAMOND、FASTA、GGSEARCH/GLSEARCH、GenoogLe、HMMER、HHpred/HHsearch、IDF、Infernal、KLAST、USEARCH、parasail、PSI-BLAST、PSI-Search、ScalaBLAST、SequIlab、SAM、SSEARCH、SWAPHI、SWAPHI-LS、SWIMM或SWIPE,其可以根据本公开用于例如鉴定多肽或核酸中的“对应”残基。

[0046] 如本文所用,术语“给药方案”通常是指单独施用给对象的例如按时间段分开的一组单位剂量(例如,多于一次)。在一些实施方案中,给定治疗剂具有推荐的给药方案,其可涉及一次或多次剂量。在一些实施方案中,给药方案包括多次剂量,其各自在时间上与其他剂量分开。在一些实施方案中,单次剂量彼此间隔相同长度的时间段;在一些实施方案中,给药方案包括多次剂量和分开单次剂量的至少两个不同的时间段。在一些实施方案中,给药方案内的所有剂量均具有相同的单位剂量量。在一些实施方案中,给药方案内的不同剂量具有不同的量。在一些实施方案中,给药方案包括第一剂量量的第一剂量,然后是与第一剂量量不同的第二剂量量的一次或多次另外剂量。在一些实施方案中,给药方案包括第一剂量量的第一剂量,然后是与第一剂量量相同的第二剂量量的一次或多次另外剂量。在一些实施方案中,当在相关群体中施用时,给药方案与期望的或有益的结果相关(例如,是治疗性给药方案)。

[0047] 如本文所用,术语“改善”、“增加”或“减少”,或其语法上相当的比较性术语,通常指示相对于相当的参考测量结果的值。例如,在一些实施方案中,使用感兴趣的药剂获得的评估值相对于使用相当的参考药剂获得的评估值可以是“改善的”。替代地或附加地,在一些实施方案中,相对于在不同条件下(例如,在诸如施用感兴趣的药剂的事件之前或之后)的相同对象或系统中,或在不同的相当对象(例如,在存在感兴趣的特定疾病、病症或病况的一个或多个指标的情况下,与感兴趣的对象或系统不同的相当的对象或系统,或在先前暴露于病况或药剂等)中获得的评估值,在感兴趣的对象或系统中获得的评估值可以是“改善的”。

[0048] 如本文所用,术语“药物组合物”通常是指与一种或多种药学上可接受的载体一起配制的活性剂。在一些实施方案中,活性剂以适合在治疗方案中向相关对象施用的单位剂量量存在(例如,已证明在施用时显示达到预定治疗效果的统计显著概率的量),或在不同的相当对象(例如,在存在感兴趣的特定疾病、病症或病况的一个或多个指标的情况下,或在暴露于病况或药剂之前等,与感兴趣的对象或系统不同的相当的对象或系统)中存在。在一些实施方案中,比较性术语是指统计学上相关的差异(例如,具有足以实现统计相关性的普遍性或幅度)。

[0049] 如本文所用,短语“药学上可接受的”通常是指在合理医学判断的范围内适合用于与人类和动物的组织接触而没有过度的毒性、刺激、过敏反应或其他问题或并发症,与合理的收益/风险比相称的那些化合物、材料、组合物或剂型。

[0050] 如本文所用,术语“参考”通常描述了与其进行比较的标准或对照。例如,在一些实施方案中,将感兴趣的药剂、动物、个体、群体、样品、序列或值与参考或对照药剂、动物、个

体、群体、样品、序列或值进行比较。在一些实施方案中,与感兴趣的测试或确定基本上同时地测试或确定参考或对照。在一些实施方案中,参考或对照是历史参考或对照,任选地体现在有形介质中。在一些实施方案中,在与所评估的那些相当的条件或情况下确定或表征参考或对照。可以确定存在足够的相似性来证明对特定的可能参考或对照的依赖或比较的情况。

[0051] 如本文所用,术语“治疗有效量”通常是指当作为治疗方案的一部分施用引起期望的生物应答的物质(例如,治疗剂、组合物或制剂)的量。在一些实施方案中,物质的治疗有效量是当向患有或易感疾病、病症或病况的对象施用足量以治疗、诊断、预防或延迟疾病、病症或病况的发作的量。如本领域普通技术人员将理解的,物质的有效量可根据诸如期望的生物终点、待递送的物质、靶细胞或组织等因素而变化。例如,用于治疗疾病、病症或病况的制剂中化合物的有效量是减轻、改善、缓解、抑制、预防、延迟疾病、病症或病况的一种或多种症状或特征的发作、降低其严重程度或降低其发生率的量。在一些实施方案中,以单剂量施用治疗有效量;在一些实施方案中,需要多个单位剂量来递送治疗有效量。

[0052] 如本文所用,术语“变体”通常是指表现出与参考实体的显著结构同一性,但与参考实体相比,在一个或多个化学部分的存在或水平上与参考实体在结构上不同的实体。在许多实施方案中,变体在功能上也不同于其参考实体。通常,特定实体是否被恰当地视为参考实体的“变体”取决于其与参考实体的结构同一性程度。任何生物或化学参考实体具有某些特征结构元件。根据定义,变体是共享一个或多个此类特征结构元件的独特化学实体。仅举几个例子,小分子可具有特征核结构元件(例如,大环核)或一个或多个特征侧基部分,使得小分子的变体是共享核结构元件和特征侧基部分但在其他侧基部分或核内存在的键类型上不同的变体(单对双、E对Z等),多肽可具有由在线性或三维空间中相对于彼此具有指定位置或有助于特定生物功能的多个氨基酸组成的特征序列元件,核酸可具有由在线性或三维空间中相对于彼此具有指定位置的多个核苷酸残基组成的特征序列元件。例如,由于氨基酸序列中的一个或多个差异或共价连接到多肽主链的化学部分(例如,碳水化合物、脂质等)中的一个或多个差异,变体多肽可不同于参考多肽。在一些实施方案中,变体多肽显示出与参考多肽具有至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%或99%的总体序列同一性。替代地或附加地,在一些实施方案中,变体多肽不与参考多肽共享至少一个特征序列元件。在一些实施方案中,参考多肽具有一种或多种生物活性。在一些实施方案中,变体多肽共享参考多肽的一种或多种生物活性。在一些实施方案中,变体多肽缺乏参考多肽的一种或多种生物活性。在一些实施方案中,与参考多肽相比,变体多肽表现出降低水平的一种或多种生物活性。在许多实施方案中,如果感兴趣的多肽具有与亲本相同但在特定位置有少量序列改变的氨基酸序列,则将感兴趣的多肽视为亲本或参考多肽的“变体”。例如,与亲本相比,变体中少于20%、15%、10%、9%、8%、7%、6%、5%、4%、3%、2%的残基被取代。在一些实施方案中,与亲本相比,变体具有10、9、8、7、6、5、4、3、2或1个取代残基。通常,变体具有非常少量(例如,少于5、4、3、2或1)的取代功能残基(例如,参与特定生物活性的残基)。此外,与亲本相比,变体可以具有具有不超过5、4、3、2或1个添加或缺失,并且通常没有添加或缺失。此外,任何添加或缺失可少于约25、约20、约19、约18、约17、约16、约15、约14、约13、约10、约9、约8、约7、约6,并且通常少于约5、约4、约3或约2个残基。在一些实施方案中,亲本或参考多肽是在自然界中发现的多肽。

[0053] A.所提供的分类器

[0054] 本公开提供了一种分类器以及这种分类器的开发,所述分类器可以鉴定(例如,预测)哪些患者将对特定疗法产生或不产生应答。在一些实施方案中,分类器被建立用于区分已接受抗TNF疗法(例如,特定抗TNF剂或方案)的有应答和无应答的先前对象。

[0055] 除其他事项外,本公开涵盖以下见解:特定基因集的表达水平,单独或彼此组合,任选地与特定临床特征或存在或不存在特定单核苷酸多态性相结合,可用于预测对抗TNF疗法的应答(例如,应答的一个或多个特征)。

[0056] 在一些实施方案中,本公开提供了一种分类器,其为或包括此类基因表达水平、临床特征或SNP,并证明已被建立用于区分对抗TNF疗法作出应答和不作出应答的对象。在一些实施方案中,所提供的分类器被建立用于通过对接受抗TNF疗法且其应答性已知(例如,先前确定)的历史(例如,先前)对象群体的回顾性分析在对抗TNF疗法有应答或无应答的对象(例如,抗TNF疗法原初对象)之间进行区分。在一些实施方案中,当应用于此类历史(例如,先前)群体时,以至少70%的准确度鉴定队列中至少50%的无应答者的分类器被视为“经验证”。在一些实施方案中,当应用于此类历史(例如,先前)群体时,以至少70%的准确度鉴定队列中至少60%的无应答者的分类器被视为“经验证”。在一些实施方案中,当应用于此类历史(例如,先前)群体时,以至少70%的准确度鉴定队列中至少70%的无应答者的分类器被视为“经验证”。在一些实施方案中,当应用于此类历史(例如,先前)群体时,以至少70%的准确度鉴定队列中至少80%的无应答者的分类器被视为“经验证”。在一些实施方案中,当应用于此类历史(例如,先前)群体时,以至少70%的准确度鉴定队列中至少90%的无应答者的分类器被视为“经验证”。在一些实施方案中,当应用于此类历史(例如,先前)群体时,以至少70%的准确度鉴定队列中至少99%的无应答者的分类器被视为“经验证”。

[0057] 在一些实施方案中,当应用于此类历史(例如,先前)群体时,以至少80%的准确度鉴定队列中至少50%的无应答者的分类器被视为“经验证”。在一些实施方案中,当应用于此类历史(例如,先前)群体时,以至少90%的准确度鉴定队列中至少50%的无应答者的分类器被视为“经验证”。在一些实施方案中,当应用于此类历史(例如,先前)群体时,以至少99%的准确度鉴定队列中至少50%的无应答者的分类器被视为“经验证”。

[0058] 在一些实施方案中,本公开提供了治疗患有疾病、病症或病况的对象的方法,其包括向已通过应用所提供的分类器确定可能对这种抗TNF疗法产生应答的一个或多个对象施用抗TNF疗法;替代地或另外地,在一些实施方案中,本公开提供了治疗患有疾病、病症或病况的对象的方法,其包括停止抗TNF疗法,或对通过应用提供的分类器确定不太可能对这种抗TNF疗法产生应答的一个或多个对象施用抗TNF疗法的替代方案。

[0059] 在一些实施方案中,所提供的分类器可以是或包括一种或多种基因的基因表达信息。替代地或另外地,在一些实施方案中,所提供的分类器可以是或包括一个或多个单核苷酸多态性(SNP)的存在或不存在或相关对象的一种或多种临床特征或特性。

[0060] 在一些实施方案中,通过评估以下各项开发分类器:表达水平与临床应答性或无应答性显著相关(例如,以线性或非线性方式)的一种或多种基因;一个或多个SNP的存在;以及至少一种临床特征。

[0061] 在一些实施方案中,如本文所述,分类器是通过对自己接受抗TNF疗法并已确定有应答(例如,应答者)或无应答(例如,无应答者)的患者(例如,先前对象)的生物样品的一

种或多种特征(例如,基因表达水平、存在或不存在一个或多个SNP等)进行回顾性分析来开发的;替代地或另外地,在一些实施方案中,分类器是通过对此类患者的一种或多种临床特征的回顾性分析来开发的,其可能涉及或可能不涉及任何生物样品的评估(并且可以例如通过参考医疗记录来完成)。在一些实施方案中,所有此类患者已接受相同的抗TNF疗法(任选地持续相同或不同的时间段);替代地或另外地,在一些实施方案中,所有此类患者已被诊断患有相同的疾病、病症或病况。在一些实施方案中,其生物样品在回顾性分析中被分析的患者已经接受了不同的抗TNF疗法(例如,使用不同的抗TNF剂或根据不同的方案);替代地或另外地,在一些实施方案中,其生物样品在回顾性分析中被分析的患者已被诊断为患有不同的疾病、病症或病况。

[0062] 许多统计学分类技术适于用作为执行上述分类的方法(例如,区分对抗TNF疗法有应答的对象和无应答的对象)。这些方法包括但不限于监督学习方法。

[0063] 在监督学习方法中,使用统计学分类方法分析或处理来自两个组或多个组的一组样品(例如,对抗TNF疗法有应答和无应答的样品)。本文所述的基因或特定SNP或变异体的不存在/存在、或基因或生物标志物的表达水平可用作区分两个组或更多个组的分类器的基础。然后可以分析或处理新样品,以便分类器可以将新样品与两个组或更多个组中的一个组相关联。

[0064] 常用的监督分类器包括但不限于神经网络(例如,人工神经网络、多层感知器)、支持向量机、k-最近邻、高斯混合模型、高斯、朴素贝叶斯、决策树和径向基函数(RBF)分类器。线性分类方法包括Fisher线性判别、逻辑回归、朴素贝叶斯分类器、感知器和支持向量机(SVM)。与根据本公开的方法一同使用的其他分类器包括二次分类器、k-最近邻、boosting、决策树、随机森林、神经网络、模式识别、贝叶斯网络和隐马尔可夫模型。常用于监督学习的其他分类器包括其改进或组合,也可以适用于本文所述的方法。

[0065] 使用监督方法的分类通常可以按照以下的方法进行。

[0066] 1) 收集训练集。这些可以包括,例如,本文所述的一种或多种基因或生物标志物的表达水平,所述基因或生物标志物来自对抗TNF疗法有应答或无应答的患者的样品。训练样品用于“训练”分类器。

[0067] 2) 确定所学习函数的输入“特征”表示。学习函数的准确性取决于输入对象的表示方式。例如,将输入对象转换为特征向量,其中包含描述该对象的多个特征。这些特征可能包括在来自患者或对象的样品中检测到的一组基因。

[0068] 3) 确定所学习函数的结构和相应的学习算法。选择一种学习算法,例如人工神经网络、决策树、贝叶斯分类器或支持向量机。学习算法用于构建分类器。

[0069] 4) 构建分类器(例如,分类模型)。在收集的训练集上运行学习算法。学习算法的参数可以通过优化在训练集的子集(称为验证集)上的性能或通过交叉验证来进行调整。在参数调整和学习之后,可以在独立于训练集的原初样品的测试集上测量算法的性能。所建立的模型可能涉及分配给各个特征的特征系数或重要性度量。

[0070] 在一些情况下,个体特征是个体基因或个体基因的水平。在一些情况下,基因的水平是归一化值、平均值、中间值、均值、调整后的平均值或其他调整后的水平或值。个体特征可以包括或由基因集或基因板组成,诸如本文提供的基因集。

[0071] 如上所述确定了分类器(例如,分类模型)后(“经训练”),即可以将其用于分类样

品,例如,根据本文所述的方法分析或处理的包含表达基因的患者样品。

[0072] 1. 基因表达

[0073] 在一些实施方案中,如本文所述的分类器的基因表达方面通过评估以下各项来确定:表达水平与临床应答性或无应答性显著相关(例如,以线性或非线性方式)的一种或多种基因;以及以下各项中的至少一个:一种或多种基因的表达序列中一个或多个单核苷酸多态性(SNP)的存在;或有应答和无应答的先前对象的至少一种临床特征。基因应答特征中可包括表达水平在应答者和无应答者群体之间表现出统计学显著性差异的基因。

[0074] 在一些实施方案中,本公开体现了这样一种见解,即在有应答与无应答对象之间鉴定或提供分类器的某些先前努力的问题根源是应答者和无应答者群体中的基因表达水平的比较强调或集中在(通常仅集中在)表现出群体间表达水平的最大差异(例如,大于2倍的变化)的基因上。本公开认识到,即使那些表达水平差异相对较小(例如,表达变化小于2倍)的基因也提供有用的信息,并且有价值地包括在本文所述实施方案中的分类器中。

[0075] 此外,在一些实施方案中,本公开体现了这样一种见解,即如本文所述对表达水平在应答者与无应答者群体之间显示出统计学显著性差异(任选地包括小的差异)的基因的相互作用模式的分析提供新的且有价值的信息,这从实质上提高分类器的质量和预测能力。

[0076] 在一些实施方案中,所提供的分类器是或包括可以用于确定对象是否将对特定疗法(例如,抗TNF疗法)作出应答(例如,表达水平与其相关)的基因或基因集。在一些实施方案中,分类器通过评估以下各项进行开发:表达水平与临床应答性或无应答性显著相关(例如,以线性或非线性方式)的一种或多种基因;以及以下各项中的至少一个:一个或多个单核苷酸多态性(SNP)的存在;以及有应答和无应答的先前对象的至少一种临床特征。

[0077] 在一些实施方案中,用于在分类器中使用或用于测量基因表达的一种或多种基因选自表1中的基因及其组合:

[0078] 表1

[0079]

ALPL
ATRAID
BCL6
CDK11A
CFLAR
COMMD5
GOLGA1
IL1B
IMPDH2
JAK3
KLHDC3

[0080]

LIMK2
NOD2
NOTCH1
SPINT2
SPON2
STOML2
TRIM25
ZFP36

[0081] 在一些实施方案中,用于分类器或用于测量基因表达的基因选自表1中的两个或更多个基因。在一些实施方案中,用于分类器或测量基因表达的基因选自表1中的两个或更多个、三个或更多个、四个或更多个、五个或更多个、六个或更多个、七个或更多个、八个或更多个、九个或更多个、十个或更多个、十一个或更多个、十二个或更多个、十三个或更多个、十四个或更多个、十五个或更多个、十六个或更多个、十七个或更多个、或全部十九个基因。

[0082] 在一些实施方案中,用于分类器或用于测量基因表达的基因选自表2中的一种或多种基因及其组合。

[0083] 表2

[0084]

ALPL
BCL6
CDK11A
CFLAR
IL1B
JAK3
LIMK2
NOD2
NOTCH1
TRIM25
ZFP36

[0085] 在一些实施方案中,用于分类器或用于测量基因表达的基因选自表2中的两个或更多个基因。在一些实施方案中,用于分类器或用于测量基因表达的基因选自表2中的两个或更多个、三个或更多个、四个或更多个、五个或更多个、六个或更多个、七个或更多个、八个或更多个、九个或更多个、十个或更多个、或全部十一个基因。

[0086] 在一些实施方案中,分类器中的基因表达模式可以使用mRNA或蛋白质表达数据集来鉴定或检测,例如,可以是或已经从经验证的生物数据(例如,来源于诸如基因表达综合数据库(Gene Expression Omnibus,“GEO”)的可公开获得的数据库的生物数据)制备。在一些实施方案中,可通过比较特定疗法(例如,抗TNF疗法)的已知有应答和已知无应答的先前对象的基因表达水平来衍生分类器。在一些实施方案中,某些基因(例如,特征基因)选自待用于开发分类器的该基因表达数据队列。

[0087] 在一些实施方案中,通过与出于所有目的通过引用并入本文的Santolini,“A personalized, multiomics approach identifies genes involved in cardiac hypertrophy and heart failure”, *Systems Biology and Applications*, (2018) 4:12; doi:10.1038/s41540-018-0046-3中报告的那些类似的方法鉴定特征基因或表达模式。在一些实施方案中,通过比较已知有应答和无应答的先前对象的基因表达水平并鉴定两组之间的显著变化来鉴定特征基因或表达模式,其中显著变化可以是表达上的巨大差异(例如,大于2倍的变化)、表达上的微小差异(例如,小于2倍的变化),或两者。在一些实施方案中,通过表达差异的显著性对基因进行排序。在一些实施方案中,通过基因表达与应答结果之间的Pearson相关性来测量显著性。在一些实施方案中,通过表达差异的显著性从排序中选择特征基因。在一些实施方案中,所选择的特征基因的数量小于所分析基因的总数。在一些实施方案中,选择200种或更少的特征基因。在一些实施方案中,选择100种或更少的基因。

[0088] 在一些实施方案中,特征基因是结合其在人类相互作用组(HI)(蛋白质-蛋白质相互作用的图谱)上的位置选择的或者特征基因的特征在于该位置。以这种方式使用HI包括认识到mRNA活性是动态的,并确定对理解某些疾病至关重要的蛋白质的实际上过度和不足的表达。在一些实施方案中,与对某些疗法(例如,抗TNF疗法)的应答相关联的基因可聚类在HI图谱上的离散模块中(例如,形成基因的聚类)。此类聚类的存在与基本的潜在疾病生物学的存在相关联。在一些实施方案中,分类器源自从HI图谱上的基因聚类中选择的特征基因。因此,在一些实施方案中,分类器源自与人类相互作用组图谱上对抗TNF疗法的应答相关联的基因聚类。

[0089] 在一些实施方案中,当映射到人类相互作用组图谱上时,与对某些疗法的应答相关联的基因表现出某些拓扑性质。例如,在一些实施方案中,多个基因与抗TNF疗法的应答相关联,并通过其在人类相互作用组图谱上的位置(例如拓扑性质,例如,它们的彼此邻近性)进行表征。

[0090] 在一些实施方案中,与对某些疗法(例如,抗TNF疗法)的应答相关联的基因可在HI图谱上彼此非常接近地存在。所述邻近基因不一定共享基本的潜在疾病生物学。也就是说,在一些实施方案中,邻近基因不共享显著的蛋白质相互作用。因此,在一些实施方案中,分类器源自人类相互作用组图谱上邻近的基因。在一些实施方案中,分类器源自人类相互作用组图谱上的某些其他拓扑特征。

[0091] 在一些实施方案中,当与HI图谱结合使用时,可通过扩散状态距离(Diffusion State Distance, DSD)(参见Cao等人, *PLoS One*, 8(10):e76339(Oct. 23, 2013),其出于所有目的通过引用并入本文)确定与对某些疗法(例如,抗TNF疗法)的应答相关联的基因。

[0092] 在一些实施方案中,通过(1)基于与已知应答者和已知无应答者相比的基因表达差异的显著性对基因进行排序;(2)从排序的基因中选择基因并将所选基因映射到人类相互作用组图谱上;以及(3)从映射到人类相互作用组图谱上的基因中选择特征基因来选择特征基因。因此,在一些实施方案中,特征基因的特征在于它们在应答者与无应答者对象或群体中的表达差异的相对排序。

[0093] 在一些实施方案中,将特征基因(例如,从Santolini方法中选择,或使用各种网络拓扑特性,包括但不限于基于聚类、接近性和扩散的方法选择)提供给概率神经网络或其他本文所述的分类器,从而提供(例如,“训练”)分类器。在一些实施方案中,概率神经网络实

施D.F.Specht在“Probabilistic Neural Networks,”*Neural Networks*,3(1):109-118 (1990) (其通过引用并入本文)中提出的算法。在一些实施方案中,概率神经网络以R统计语言编写,并且已知通过定量变量的向量描述的一组观察结果,将观察结果分类为给定数量的组(例如,应答者和无应答者)。该算法使用获自己知应答者和无应答者的特征基因的数据集进行训练,并提供新观察结果。在一些实施方案中,概率神经网络是源自 [pnn: Probabilistic neural networks v1.0.1](#) 的 [The Comprehensive R Archive Network](#) 的概率神经网络。在一些实施方案中,根据随机森林模型分析特征基因以提供分类器。

[0094] 2. 单核苷酸多态性

[0095] 本公开还包括这样的见解,即可以通过RNA序列数据鉴定单核苷酸多态性(SNP)。也就是说,通过将RNA序列数据与参考人类基因组进行比较,例如,通过将RNA序列数据映射到GRCh38人类基因组。在不希望受理论约束的情况下,认为与分类器中使用的RNA序列相关的SNP的存在可有助于鉴定对某些疗法(例如,抗TNF疗法)有应答或无应答的对象亚群。也就是说,可以使用网络医学和通路富集分析来分析区分性基因或含SNP的RNA的蛋白质产物。由包括在分类器中的区分性基因或含SNP的RNA编码的蛋白质可以在例如人类相互作用组的图谱上重叠,以通过鉴定某些区分性基因集合来帮助鉴定对象的某些亚群。

[0096] 在一些实施方案中,所提供的分类器和使用此类分类器的方法结合与单核苷酸多态性(SNP)相关的评估。在一些实施方案中,本公开提供了一种开发分类器以用于针对一个或多个治疗属性对对象分层的方法,其包括:分析针对至少一个治疗属性代表至少两个不同类别的对象中表达的RNA的序列数据;由序列数据评估一个或多个单核苷酸多态性(SNP)的存在;确定一个或多个SNP的存在与至少一个治疗属性相关;以及将一个或多个SNP包括在分类器中。

[0097] 在一些实施方案中,本公开提供了开发分类器以用于通过分析针对至少一个治疗属性代表至少两个不同类别的对象中表达的RNA的序列数据来针对一个或多个治疗属性对对象分层的方法的改进,其包括:由序列数据评估一个或多个单核苷酸多态性(SNP)的存在;以及确定一个或多个SNP的存在与至少一个治疗属性相关;以及将一个或多个SNP的存在包括在分类器中。

[0098] 在一些实施方案中,一种或多种SNP选自表3。

[0099] 表3

[0100]

chr1.161644258
chr1.2523811
chr11.107967350
chr17.38031857
chr7.1285800042
rs10774624
rs10985070
rs11889341
rs1571878
rs1633360
rs17668708
rs1877030
rs1893592
rs1980422
rs2228145
rs2233424
rs2236668
rs2301888
rs2476601
rs3 087243
rs3218251
rs331463
rs34536443
rs34695944
rs4239702
rs4272
rs45475795
rs508970
rs5987194
rs657075
rs67 15284
rs706778

[0101]	rs72634030
	rs73013527
	rs73 194058
	rs773125
	rs7752903
	rs8083786
	rs9653442

[0102] 在一些实施方案中, SNP选自表3中的两种或更多种SNP。在一些实施方案中, SNP选自表3中的两种或更多种、三种或更多种、四种或更多种、五种或更多种、六种或更多种、七种或更多种、八种或更多种、九种或更多种、十种或更多种、十一种或更多种、十二种或更多种、十三种或更多种、十四种或更多种、十五种或更多种、十六种或更多种、十七种或更多种、十八种或更多种、十九种或更多种、二十种或更多种、二十一种或更多种、二十二种或更多种、二十三种或更多种、二十四种或更多种、二十五种或更多种、二十六种或更多种、二十七种或更多种、二十八种或更多种、二十九种或更多种、三十种或更多种、三十一一种或更多种、三十二种或更多种、三十三种或更多种、三十四种或更多种、三十五种或更多种、三十六种或更多种、三十七种更多种、三十八种或更多或全部39种SNP。

[0103] 3. 临床特征

[0104] 在一些实施方案中, 分类器还可以结合附加信息, 例如以便进一步提高分类器在应答者与无应答者之间鉴定的预测能力。例如, 在一些实施方案中, 分类器通过评估以下各项进行开发或评估 (例如, 检测): 表达水平与临床应答性或无应答性显著相关 (例如, 以线性或非线性方式) 的一种或多种基因; 以及以下各项中的至少一个: 一种或多种基因的表达序列中一个或多个单核苷酸多态性 (SNP) 的存在; 或有应答和无应答的先前对象的至少一种临床特征。也就是说, 在一些实施方案中, 通过评估表达水平与临床应答性或无应答性显著相关 (例如, 以线性或非线性方式) 的一种或多种基因以及一种或多种基因的表达序列中一个或多个单核苷酸多态性 (SNP) 的存在来开发或评估 (例如, 检测) 分类器。在一些实施方案中, 通过评估表达水平与临床应答性或无应答性显著相关 (例如, 以线性或非线性方式) 的一种或多种基因以及有应答和无应答的先前对象的至少一种临床特征来开发或评估 (例如, 检测) 分类器。

[0105] 本公开还包括这样的见解, 即某些临床特征 (例如, BMI、性别、年龄等) 可以结合到本文提供的分类器中。在一些实施方案中, 所提供的分类器和使用此类分类器的方法结合了与临床特征相关的评估。在一些实施方案中, 本公开提供了一种开发分类器以用于针对一个或多个治疗属性对对象分层的方法, 其包括: 分析针对至少一个治疗属性代表至少两个不同类别的对象中表达的RNA的序列数据; 评估一种或多种临床特征的存在; 确定与所述临床特征相关的表达与至少一个治疗属性相关; 以及将一种或多种临床特征包括在分类器中。

[0106] 在一些实施方案中, 至少一种临床特征选自: 体重指数 (BMI)、性别、年龄、种族、先前疗法治疗、疾病持续时间、C反应蛋白 (CRP) 水平、抗环瓜氨酸肽的存在、类风湿因子的存在、患者整体评估、治疗应答率 (例如, ACR20、ACR50、ACR70) 以及其组合。

[0107] 在一些实施方案中, 临床特征选自表4。

[0108] 表4

[0109]

年龄
出生性别
疾病持续时间(以年为单位)
种族(包括白人、亚洲人、黑人、混血人种、美洲原住民、太平洋岛民和其他种族)
纤维肌痛史
慢性血管疾病史(包括急性冠脉综合征、冠状动脉疾病、充血性心力衰竭、高血压、心肌梗死、外周动脉疾病、中风、不稳定型心绞痛、心脏骤停、血运重建术和室性心律失常)
导致住院的严重感染史(包括粘液囊或关节感染、蜂窝组织炎、鼻窦

[0110]

炎、憩室炎、败血症、肺炎支气管炎、胃脑膜炎、尿路感染、上呼吸道感染和结核病)
癌症史(包括乳腺癌、肺癌、皮肤癌、淋巴瘤,但不包括非黑素瘤皮肤)
BMI
吸烟状态(包括从不吸烟、以前吸烟或目前吸烟)
强的松剂量
DMARD 剂量
基线时的 C 反应蛋白水平
基线时的 DAS28-CRP
基线时的肿胀的 28 关节计数
基线时的压痛的 28 关节计数
基线时的患者整体评估
基线时的医师整体评估
基线时的 CDAI
基线时的经过修改的健康评估问卷得分
基线时的患者疼痛评估
使用 DAS28-CRP 获得的基线时的 EULAR 应答(包括差、中或好)
抗 CCP 状态(阳性或阴性)
基线时的抗 CCP 滴度
类风湿因子状态(阳性或阴性)
基线时的类风湿因子滴度

[0111] 在一些实施方案中,临床特征选自表4中的两种或更多种临床特征。在一些实施方案中,临床特征选自表4中的两种或更多种、三种或更多种、四种或更多种、五种或更多种、六种或更多种、七种或更多种、八种或更多种、九种或更多种、十种或更多种、十一种或更多种、十二种或更多种、十三种或更多种、十四种或更多种、十五种或更多种、十六种或更多种、十七种或更多种、十八种或更多种、十九种或更多种、二十种或更多种、二十一种或更多种、二十二种或更多种、二十三种或更多种、二十四种或更多种、二十五种或更多或全部二十六种临床特征。

[0112] 4. 验证分类器

[0113] 替代地或另外地,在一些实施方案中,可以使用已知应答者和无应答者的队列,使用留一法交叉或k折交叉验证,在概率神经网络中训练分类器。在一些实施方案中,这样的过程在分析中留一个样品(例如,留一法),并且基于其余样品训练分类器。在一些实施方案

中,随后使用升级的分类器预测留下样品的应答概率。在一些实施方案中,例如,可以迭代地重复这样的过程,直到所有样品被留一次。在一些实施方案中,这样的过程将已知应答者和无应答者的队列随机分配为k个大小相等的组。在k个组中,保留一个组作为验证数据,以用于测试模型,并且其余组用作训练数据。这样的过程可以重复k次,其中k个组中的每一个只使用一次作为验证数据。在一些实施方案中,结果是训练集中的每个样品的概率得分。可以将这种概率得分与实际应答结果相关。递归操作曲线 (Recursive Operating Curves, ROC) 可以用于估计分类器的性能。在一些实施方案中,约0.6或更高的曲线下面积 (AUC) 反映了合适的经验验证分类器。在一些实施方案中,0.9的阴性预测值 (NPV) 反映了合适的经验验证分类器。在一些实施方案中,可以在完全独立 (例如,盲的) 队列中测试分类器,以例如确认适合性 (例如,使用留一法或k折交叉验证)。因此,在一些实施方案中,所提供的方法还包括验证分类器,例如,通过将应答概率分配给一组已知应答者和无应答者;并且针对盲组应答者和无应答者检查分类器。这些过程的输出是经过训练的分类器,其可用于确定对象是否将对特定疗法 (例如,抗TNF疗法) 产生应答。

[0114] 在一些实施方案中,分类器被建立用于区分已接受一种类型的疗法 (例如,抗TNF疗法) 的有应答和无应答的先前对象。该分类器可以预测对象是否将对给定疗法产生应答。在一些实施方案中,有应答和无应答的先前对象患有相同的疾病、病症或病况。

[0115] 在一些实施方案中,通过监测特定的临床特征来评估治疗的有效性。例如,在一些实施方案中,通过临床特征的统计分析在对象中验证治疗应答。在某些实施方案中,相关分类器的开发、验证或使用可以涉及或已经涉及对一个或多个临床参数 (例如,患者的表现或疾病状态) 的评估。本公开认识到,在这样的临床评估中可能发生变化,这些临床评估可以例如代表患者外部的输入 (例如,评估或解释患者特征或应答的应用的差异)。本公开通过提供一个或多个相关参数的患者自我评估来提供针对该已鉴定问题的解决方案。

[0116] 在一些实施方案中,分类器的验证包括临床特征的统计分析,以分析已被分类器如此分类并接受抗TNF疗法的患者的临床特征的变化。此类验证方法认识到,与本文描述的方法相比,临床变化的某些主观测量无法量化并且涉及自我评估。本公开涵盖这样的见解:患者的自我评估不一定是一致的,但可以提供关于随时间推移的治疗应答的有价值的信息。这种自我评估应答可用于确认患者是真正的应答者还是非应答者。例如,对患者队列的某些临床特征进行统计分析可以验证分类器的准确性。在一些实施方案中,临床特征的统计分析分析ACR50、ACR70、CDAI LDA、CDAI缓解、DAS28-CRP LDA和DAS28-CRP缓解及其组合中的一种或多种的变化。在一些实施方案中,统计分析是通过蒙特卡罗模拟来执行的。

[0117] 在一些实施方案中,使用先前已使用抗TNF疗法治疗但独立于用于制备分类器的对象队列的对象队列来验证分类器。在一些实施方案中,使用基因表达数据、SNP数据或临床特征升级分类器。在一些实施方案中,当在验证队列中以60%或更高的准确度预测90%或更多的无应答对象时,分类器被视为“经验证”。

[0118] 在一些实施方案中,分类器以至少60%的准确度预测对象的非应答性,从而在至少100名对象的群体中预测非应答性。在一些实施方案中,分类器在至少150名对象的群体中以至少60%的准确度预测对象的非应答性。在一些实施方案中,分类器在至少170名对象的群体中以至少60%的准确度预测对象的非应答性。在一些实施方案中,分类器在至少200名或更多对象的群体中以至少60%的准确度预测对象的非应答性。

[0119] 在一些实施方案中,分类器在至少100名对象的群体中以至少80%的准确度预测对象的非应答性。在一些实施方案中,分类器在至少150名对象的群体中以至少80%的准确度预测对象的非应答性。在一些实施方案中,分类器在至少170名对象的群体中以至少80%的准确度预测对象的非应答性。在一些实施方案中,分类器在至少200名或更多对象的群体中以至少80%的准确度预测对象的非应答性。在一些实施方案中,分类器在至少300名或更多对象的群体中以至少80%的准确度预测对象的非应答性。在一些实施方案中,分类器在至少350名或更多对象的群体中以至少80%的准确度预测对象的非应答性。

[0120] B. 检测基因特征或SNP

[0121] 可以进行使用经过训练的分类器来检测对象中的基因特征。换句话说,通过首先定义(来自分类器的)基因特征,可以使用多种方法确定一个对象或一组对象是否表达已建立的基因特征。例如,在一些实施方案中,从业者可以在施用疗法之前从对象获得血液或组织样品,并从所述血液或组织样品中提取和分析mRNA图谱(mRNA profile)。mRNA图谱的分析可以通过多种方法,包括但不限于基因阵列、RNA测序、纳米串(nanostring)测序、实时定量逆转录PCR(qRT-PCR)、珠粒阵列或酶联免疫吸附测定(ELISA)及其组合。因此,在一些实施方案中,本公开提供了确定将对象分类为应答者还是无应答者的方法,其包括通过微阵列、RNA测序、实时定量逆转录PCR(qRT-PCR)、珠粒阵列和ELISA及其组合中的至少一种测量基因表达。在一些实施方案中,本公开提供了确定将对象分类为应答者还是无应答者的方法,其包括通过RNA测序(例如,RNAseq)测量对象的基因表达。

[0122] 本公开还包括这样的见解,即可以通过RNA序列数据鉴定单核苷酸多态性(SNP)。也就是说,通过将RNA序列数据与参考人类基因组进行比较,例如,通过将RNA序列数据映射到GRCh38人类基因组。在不希望受理论约束的情况下,认为与分类器中使用的RNA序列相关的SNP的存在可有助于鉴定对某些疗法(例如,抗TNF疗法)有应答或无应答的对象亚群。也就是说,可以使用网络医学和通路富集分析来分析区分性基因和含SNP的RNA的蛋白质产物。由包括在分类器中的区分性基因和含SNP的RNA编码的蛋白质可以在例如人类相互作用组的图谱上重叠,以通过鉴定某些区分性基因集来帮助鉴定对象的某些亚群。

[0123] 在一些实施方案中,通过减去背景数据、校正批次效应并除以看家基因的平均表达来测量基因表达。参见Eisenberg&Levanon,“Human housekeeping genes,revisited,” Trends in Genetics,29(10):569-574(2013年10月),其出于所有目的通过引用并入本文。在微阵列数据分析的上下文中,背景差减是指从每个探针特征的荧光信号强度中减去芯片上与任何mRNA序列不互补的探针特征产生的平均荧光信号,例如由非特异性结合产生的信号。背景差减可以使用不同的软件包执行,诸如Affymetrix基因表达控制台(Gene Expression Console)。看家基因参与基本的细胞维持,并且因此,预期在所有细胞和条件下保持恒定的表达水平。感兴趣的基因(例如应答特征中的那些基因)的表达水平可以通过将表达水平除以一组选定看家基因的平均表达水平来归一化。该看家基因归一化程序针对实验变异性校准基因表达水平。此外,归一化方法,诸如针对不同批次微阵列的变异性进行校正的鲁棒多阵列平均(“RMA”)可以在Illumina或Affymetrix平台推荐的R软件包中获得。对归一化数据进行对数变换,并去除样品中检测率低的探针。此外,从分析中去除没有可用基因符号或Entrez ID的探针。

[0124] 在一些实施方案中,本公开提供了一种试剂盒,其包括被建立用于区分已接受抗

TNF疗法的有应答和无应答的先前对象的分类器。

[0125] C.使用分类器

[0126] 1.患者分层

[0127] 除其他外,本公开提供了用于预测对抗TNF疗法的应答性的技术。在一些实施方案中,所提供的技术表现出优于其他方法的跨队列的一致性或准确度。

[0128] 因此,本公开提供了用于患者分层、定义或区分应答者和无应答者群体的技术。例如,在一些实施方案中,本公开提供了用抗TNF疗法治疗对象的方法,在一些实施方案中,所述方法包括:向已通过分类器确定为有应答的对象施用抗TNF疗法,所述分类器被建立用于区分已接受抗TNF疗法的有应答和无应答的先前对象。

[0129] 在一些实施方案中,本公开提供了一种开发分类器以用于针对一个或多个治疗属性对对象分层的方法,其包括:分析针对至少一个治疗属性代表至少两个不同类别的对象中表达的RNA的序列数据;由序列数据评估一个或多个单核苷酸多态性(SNP)的存在;确定一个或多个SNP的存在与至少一个治疗属性相关;以及将一个或多个SNP包括在分类器中。

[0130] 本文描述的分类器可以通过分析对象的基因表达来使用。在一些实施方案中,通过微阵列、RNA测序、实时定量逆转录PCR(qRT-PCR)、珠阵列、ELISA和蛋白质表达及其组合中的至少一种来测量对象的基因。

[0131] 2.疗法监测

[0132] 此外,本公开提供了用于监测用于给定对象或对象队列的疗法的技术。由于对象的基因表达水平可随时间变化,因此可能希望在一个或多个时间点(例如,以指定的和/或周期性的间隔)评估对象。

[0133] 在一些实施方案中,通过监测特定的临床特征来评估治疗的有效性。例如,在一些实施方案中,通过临床特征的统计分析在对象中验证治疗应答。在某些实施方案中,相关分类器的开发、验证或使用可以涉及或已经涉及对一个或多个临床参数(例如,患者的表现或疾病状态)的评估。本公开认识到,在这样的临床评估中可能发生变化,这些临床评估可以例如代表患者外部的输入(例如,评估的应用或患者特征或应答的解释的差异)。本公开通过提供一个或多个相关参数的患者自我评估来提供针对该已鉴定问题的解决方案。

[0134] 在一些实施方案中,分类器的验证包括临床特征的统计分析,以分析已被分类器如此分类并接受抗TNF疗法的患者的临床特征的变化。此类验证方法认识到,与本文描述的方法相比,临床变化的某些主观测量无法量化并且涉及自我评估。本公开涵盖这样的见解:患者的自我评估不一定是一致的,但可以提供关于随时间推移的治疗应答的有价值的信息。这种自我评估应答可用于确认患者是真正的应答者还是非应答者。例如,对患者队列的某些临床特征进行统计分析可以验证分类器的准确性。在一些实施方案中,临床特征的统计分析分析ACR50、ACR70、CDAI LDA、CDAI缓解、DAS28-CRP LDA和DAS28-CRP缓解及其组合中的一种或多种的变化。在一些实施方案中,统计分析是通过蒙特卡罗模拟来执行的。

[0135] 在一些实施方案中,在一定时间下重复监测允许或实现对象的基因表达谱或特征中可能影响正在进行治疗方案的一个或多个变化的检测。在一些实施方案中,响应于对对象施用的特定疗法的持续、改变或暂停而检测变化。在一些实施方案中,疗法可以例如通过增加或减少已经用来治疗对象的一种或多种药剂或治疗的施用频率或量来改变。替代地或另外地,在一些实施方案中,疗法可通过添加具有一种或多种新药剂或治疗的疗法来改

变。在一些实施方案中,疗法可通过暂停或停止一种或多种特定药剂或治疗来改变。

[0136] 仅举一个实例,如果对象最初被归类为有应答(因为对象的基因表达通过分类器被确定为与疾病、病症或病况相关联),则可以施用给定的抗TNF疗法。以给定的间隔(例如,每六个月、每年等),可以再次对对象进行测试,以确保他们仍然满足对给定抗TNF疗法的“有应答”的标准。在给定对象的基因表达水平随时间变化,并且对象不再表达与疾病、病症或病况相关的基因,或者现在表达与无应答性相关的基因的情况下,可以改变对象的疗法以适应基因表达的改变。

[0137] 因此,在一些实施方案中,本公开提供了向先前通过分类器确定为对抗TNF疗法有应答的对象施用疗法的方法。

[0138] 在一些实施方案中,本公开提供了方法,所述方法进一步包括在施用之前通过分类器确定对象不是应答者;以及施用抗TNF疗法的替代疗法。

[0139] 在一些实施方案中,通过微阵列、RNA测序、实时定量逆转录PCR(qRT-PCR)、珠粒阵列、ELISA和蛋白质表达及其组合中的至少一种测量对象的基因。

[0140] 在一些实施方案中,对象患有选自以下各项的疾病、病症或病况:类风湿性关节炎、银屑病性关节炎、强直性脊柱炎、克罗恩病、溃疡性结肠炎、慢性银屑病、化脓性汗腺炎、多发性硬化症和幼年特发性关节炎及其组合。

[0141] 在一些实施方案中,抗TNF疗法是或包括施用英夫利昔单抗、阿达木单抗、依那西普、赛妥珠单抗、golimumab或其生物仿制药及其组合。在一些实施方案中,抗TNF疗法是或包括施用英夫利昔单抗或阿达木单抗。

[0142] 在一些实施方案中,有应答和无应答的先前对象患有相同的疾病、病症或病况。

[0143] 在一些实施方案中,施用抗TNF疗法的对象患有与先前有应答和无应答的先前对象相同的疾病、病症或病况。

[0144] 在一些实施方案中,所述疾病、病症或病况选自类风湿性关节炎、银屑病性关节炎、强直性脊柱炎、克罗恩病、溃疡性结肠炎、慢性银屑病、化脓性汗腺炎、多发性硬化症和幼年特发性关节炎及其组合。

[0145] 在一些实施方案中,所述疾病、病症或病况为类风湿性关节炎。

[0146] 在一些实施方案中,所述疾病、病症或病况为溃疡性结肠炎。

[0147] D. 治疗方法

[0148] 在一些实施方案中,针对其施用抗TNF疗法或停止抗TNF疗法(或施用替代疗法)的对象或群体是被确定为表现出一种或多种基因(并且在一些情况下的多种基因)的特定表达水平的对象或群体。在一些实施方案中,一种或多种基因被确定为具有低于特定阈值的表达水平;替代地或另外地,在一些实施方案中,一种或多种基因被确定为具有低于特定阈值的表达水平。在一些实施方案中,特定的基因集被确定为具有表达模式,其中相对于特定阈值评估每个基因(并且例如,被确定为高于、低于此类阈值或与其相当)。

[0149] 在一些实施方案中,本公开提供了一种治疗患有疾病、病症或病况的对象的方法,其包括向已被确定为表现出低于一种或多种基因的特定表达水平的对象施用抗TNF疗法的替代方案。

[0150] 在一些实施方案中,本公开提供了向已通过分类器确定为有应答的对象施用抗TNF疗法的方法,所述分类器被建立用于区分已接受抗TNF疗法的有应答和无应答的先前对

象(例如,其中已通过回顾性分析建立分类器,以在对其所接受的抗TNF疗法产生应答的患者与不产生应答的那些患者之间进行区分);其中所述分类器通过评估以下各项进行开发:表达水平与临床应答性 or 无应答性显著相关(例如,以线性或非线性方式)的一种或多种基因;以及以下各项中的至少一个:表达序列中一个或多个单核苷酸多态性(SNP)的存在;以及有应答和无应答的先前对象的至少一种临床特征。

[0151] TNF介导的病症目前通过抑制TNF,特别是通过施用抗TNF剂(例如,通过抗TNF疗法)进行治疗。在美国批准使用的抗TNF剂的实例包括单克隆抗体,诸如阿达木单抗(Humira®)、赛妥珠单抗(Cimiza®)、英夫利昔单抗(Remicade®)以及诱骗循环受体融合蛋白,诸如依那西普(Enbrel®)。这些药剂目前被批准用于根据在表5中列出的给药方案治疗适应症:

[0152] 表5

[0153]

适应症	阿达木单抗 ¹	赛妥珠单抗 ¹	英夫利昔单抗 ²	依那西普 ¹	戈利木单抗 ¹	戈利木单抗 ²
幼年特发性关节炎	10 kg (22 lbs) 至 <15 kg (33 lbs): 10 mg 每隔一周 15 kg (33 lbs) 至 <30 kg (66	N/A	N/A	每周 0.8 mg/kg, 每周最多 50 mg	N/A	N/A

[0154]

适应症	阿达木单抗 ¹	赛妥珠单抗 ¹	英夫利昔单抗 ²	依那西普 ¹	戈利木单抗 ¹	戈利木单抗 ²
	lbs): 20 mg 每隔一周 • ≥ 30 kg (66 lbs): 40 mg 每隔一周					
银屑病关节炎	40 mg 每隔一周	最初以及第 2 周和第 4 周时 400 mg, 随后每隔一周 200 mg; 维持剂量, 每 4 周 400 mg	第 0、2 和 6 周 5 mg/kg, 然后每 8 周一次	50 mg 每周一次, 联合或不联合甲氨蝶呤	50 mg 通过皮下注射施用, 每月一次	N/A
类风湿关节炎	40 mg 每隔一周	最初以及第 2 周和第 4 周时 400 mg, 随后每隔一周 200 mg; 维持剂量, 每 4 周 400 mg	与甲氨蝶呤联合使用, 第 0、2 和 6 周时 3 mg/kg, 然后每 8 周一次	50 mg 每周一次, 联合或不联合甲氨蝶呤	50 mg 每月一次	第 0 周和第 4 周, 经过 30 分钟静脉输注 2 mg/kg, 然后每 8 周一次
强直性脊柱炎	40 mg 每隔一周	最初以及第 2 周和第 4 周时 400 mg (分 2 次皮下注射, 每次 200 mg), 随后每隔一周 200 mg	第 0、2 和 6 周时 5 mg/kg, 然后每 6 周一次	50 mg, 每周一次	50 mg 通过皮下注射施用, 每月一次	N/A

[0155]

适应症	阿达木单抗 ¹	赛妥珠单抗 ¹	英夫利昔单抗 ²	依那西普 ¹	戈利木单抗 ¹	戈利木单抗 ²
		或每 4 周 400 mg				
成人 克罗恩病	- 初始剂量(第 1 天): 160 mg - 两周后(第 15 天)第二剂量: 80 mg - 两周后(第 29 天): 开始维持剂量, 每隔一周 40 mg	- 最初以及第 2 周和第 4 周时 400 mg - 以每周 400 mg 继续	第 0、2 和 6 周时 5 mg/kg, 然后每 8 周一次。	N/A	N/A	N/A
小儿 克罗恩病	17 kg (37 lbs) 至 < 40 kg (88 lbs): - 初始剂量(第 1 天): 80 mg - 两周后(第 15 天)第二剂量: 40 mg - 两周后(第 29 天): 开始维持剂	N/A	第 0、2 和 6 周时 5 mg/kg, 然后每 8 周一次。	N/A	N/A	N/A

[0156]

适应症	阿达木单抗 ¹	赛妥珠单抗 ¹	英夫利昔单抗 ²	依那西普 ¹	戈利木单抗 ¹	戈利木单抗 ²
	量，每隔一周 20 mg ≥ 40 kg (88 lbs): • 初始剂量(第 1 天): 160 mg • 两周后(第 15 天)第二剂量: 80 mg • 两周后(第 29 天): 开始维持剂量，每隔一周 40 mg					

[0157]

适应症	阿达木单抗 ¹	赛妥珠单抗 ¹	英夫利昔单抗 ²	依那西普 ¹	戈利木单抗 ¹	戈利木单抗 ²
溃疡性结肠炎	- 初始剂量(第 1 天): 160 mg - 两周后(第 15 天)第二剂量: 80 mg - 两周后(第 29 天): 开始维持剂量, 每隔一周 40 mg	N/A	第 0、2 和 6 周时 5 mg/kg, 然后每 8 周一次。	N/A	N/A	N/A
斑块型银屑病	80 mg 初始剂量; 从初始剂量后一周开始, 每隔一周 40 mg	N/A	N/A	50 mg 每周两次, 持续 3 个月, 随后 50 mg 每周一次	N/A	N/A
化脓性汗腺炎	- 初始剂量(第 1 天): 160 mg - 两周后(第 15 天)第二剂量: 80 mg - 第三剂量(第	N/A	N/A	N/A	N/A	N/A

	适应症	阿达木单抗 ¹	赛妥珠单抗 ¹	英夫利昔单抗 ²	依那西普 ¹	戈利木单抗 ¹	戈利木单抗 ²
[0158]		29 天) 及后续 剂量: 每周 40 mg					
	葡萄膜炎	80 mg 初始剂量; 从初始剂量后一周开始, 每隔一周 40 mg	N/A	N/A	N/A	N/A	N/A

[0159] 1. 通过皮下注射施用

[0160] 2. 通过静脉内输注施用

[0161] 本公开提供了与抗TNF疗法相关的技术, 包括表5中列出的那些治疗方案。在一些实施方案中, 抗TNF疗法是或包括施用英夫利昔单抗(**Remicade[®]**)、阿达木单抗(**Humira[®]**)、赛妥珠单抗(**Cimzia[®]**)、依那西普(**Enbrel[®]**)或其生物仿制药。在一些实施方案中, 抗TNF疗法是或包括施用英夫利昔单抗(**Remicade[®]**)或阿达木单抗(**Humira[®]**)及其组合。在一些实施方案中, 抗TNF疗法是或包括施用英夫利昔单抗(**Remicade[®]**)。在一些实施方案中, 抗TNF疗法是或包括施用阿达木单抗(**Humira[®]**)。

[0162] 在一些实施方案中, 抗TNF疗法是或包括施用生物类似抗TNF剂。在一些实施方案中, 抗TNF剂选自英夫利昔单抗生物仿制药例如CT-P13、BOW015、SB2、**Inflectra[®]**、**Renflexis[®]**和Ixifi[™], 阿达木单抗生物仿制药, 例如ABP 501 (AMGEVITA[™])、Adfrar和Hulio[™], 以及依那西普生物仿制药, 例如HD203、SB4 (**Benepali[®]**)、GP2015、**Erelzi[®]**和Intacept及其组合。

[0163] 在一些实施方案中, 治疗例如幼年特发性关节炎、银屑病关节炎、类风湿性关节炎、强直性脊柱炎、儿科克罗恩病、溃疡性结肠炎、斑块型银屑病、化脓性汗腺炎和葡萄膜炎包括表5中的抗TNF剂的给药方案。在一些实施方案中, 抗TNF剂包括例如表5中的阿达木单抗。在一些实施方案中, 阿达木单抗的给药方案包括例如最高达160mg或更多的初始剂量。在一些实施方案中, 阿达木单抗的给药方案包括例如最高达80mg或更多的第二剂量。在一些实施方案中, 阿达木单抗的给药方案包括例如每隔一周最高达40mg或更多的维持剂量。在一些实施方案中, 抗TNF剂包括例如表5中的赛妥珠单抗。在一些实施方案中, 赛妥珠单抗的给药方案包括例如最高达400mg或更多的第一初始剂量。在一些实施方案中, 赛妥珠单抗的给药方案包括例如在第2周最高达400mg或更多的第二初始剂量。在一些实施方案中, 赛妥珠单抗的给药方案包括例如在第4周最高达400mg的第三初始剂量。在一些实施方案中,

赛妥珠单抗的给药方案包括例如每隔一周最高达200mg或更多的维持剂量或每四周最高达400mg或更多的维持剂量。在一些实施方案中,抗TNF剂包括例如表5中的英夫利昔单抗。在一些实施方案中,英夫利昔单抗的给药方案包括例如最高达5mg/kg或更多的第一初始剂量。在一些实施方案中,英夫利昔单抗的给药方案包括例如在第2周最高达5mg/kg或更多的第二初始剂量。在一些实施方案中,英夫利昔单抗的给药方案包括例如第6周最高达5mg/kg或更多的第三初始剂量。在一些实施方案中,英夫利昔单抗的给药方案包括例如每6周或每8周最高达5mg/kg或更多的维持剂量。在一些实施方案中,抗TNF剂包括例如表5中的依那西普。在一些实施方案中,依那西普的给药方案包括例如最高达50mg或更多的初始剂量,每周两次,持续三个月。在一些实施方案中,依那西普的给药方案包括例如每周高达50mg或更多的维持剂量。在一些实施方案中,抗TNF剂包括例如表5中的戈利木单抗。在一些实施方案中,戈利木单抗的给药方案包括例如每月最高达50mg或更多的剂量。在一些实施方案中,戈利木单抗的给药方案包括例如最高达2mg/kg的第一初始剂量。在一些实施方案中,戈利木单抗的给药方案包括例如在第2周最高达2mg/kg或更多的第二初始剂量。在一些实施方案中,戈利木单抗的给药方案包括例如每8周最高达2mg/kg或更多的维持剂量。

[0164] 在一些实施方案中,本公开提供了治疗患有自身免疫性病症的对象的方法,该方法包括:向已通过分类器确定为有应答的对象施用抗TNF疗法,所述分类器被建立用于区分队列中的已接受抗TNF疗法的有应答的先前对象和无应答的先前对象;其中通过评估以下来开发分类器:表达水平与临床应答性 or 无应答性显著相关(例如,以线性或非线性方式)的一种或多种基因;以下至少一项:所述一种或多种基因的表达序列中一种或多种单核苷酸多态性(SNP)的存在;或有应答和无应答的先前对象的至少一种临床特征;并且其中分类器由与已接受抗TNF疗法的队列不同的独立队列验证。

[0165] 在一些实施方案中,对象先前已施用抗TNF疗法。在一些实施方案中,在所述施用之前至少1个月、至少2个月、至少3个月、至少4个月、至少5个月或至少6个月,对象已施用抗TNF疗法。

[0166] 在一些实施方案中,源自队列中已接受抗TNF疗法的对象的数据是一种类型(例如,微阵列、RNAseq等),并且用于在独立队列中验证分类器的数据来自不同类型(例如,微阵列、RNAseq等)。因此,一些实施方案,使用源自有应答和无应答的先前对象的微阵列分析来建立分类器。在一些实施方案中,使用源自独立队列的RNAseq数据来验证分类器。

[0167] E. 疾病、病症或病况

[0168] 一般而言,所提供的公开内容在考虑或实施抗TNF疗法的施用的任何情况下都是有用的。在一些实施方案中,所提供的技术可用于诊断或治疗患有与异常(例如,升高的)TNF表达或活性相关的疾病、病症或病况的对象。在一些实施方案中,所提供的技术可用于监测正在接受或已经接受抗TNF疗法的对象。在一些实施方案中,提供的技术鉴定对象是否会响应于给定的抗TNF疗法。在一些实施方案中,所提供的技术鉴定对象是否会对给定的抗TNF疗法产生抗性。

[0169] 因此,本公开提供了与治疗与TNF相关的各种病症(包括表5中列出的那些)相关的技术。在一些实施方案中,对象患有选自以下的疾病、病症或病况:类风湿性关节炎、银屑病关节炎、强直性脊柱炎、克罗恩病(成人或儿童)、溃疡性结肠炎、炎性肠病、慢性银屑病、斑块型银屑病、化脓性汗腺炎、哮喘、葡萄膜炎和幼年特发性关节炎及其组合。在一些实施方

案中,疾病、病症或病况是类风湿性关节炎。在一些实施方案中,疾病、病症或病况是银屑病关节炎。在一些实施方案中,疾病、病症或病况是强直性脊柱炎。在一些实施方案中,疾病、病症或病况是克罗恩病。在一些实施方案中,疾病、病症或病况是成人克罗恩病。在一些实施方案中,疾病、病症或病况是儿科克罗恩病。在一些实施方案中,疾病、病症或病况是炎症肠病。在一些实施方案中,疾病、病症或病况是溃疡性结肠炎。在一些实施方案中,疾病、病症或病况是慢性银屑病。在一些实施方案中,疾病、病症或病况是斑块型银屑病。在一些实施方案中,疾病、病症或病况是化脓性汗腺炎。在一些实施方案中,疾病、病症或病况是哮喘。在一些实施方案中,疾病、病症或病况是葡萄膜炎。在一些实施方案中,疾病、病症或病况是幼年特发性关节炎。

[0170] 在一些实施方案中,疾病、病症或病况是环状肉芽肿、脂性渐进性坏死、化脓性汗腺炎、坏疽性脓皮病、Sweet综合征、角膜下脓疱性皮肤病、系统性红斑狼疮、硬皮病、皮炎、白塞氏病、急性/慢性移植物抗宿主病、毛发红糠疹、干燥综合征、韦格纳肉芽肿病、风湿性多肌痛、皮炎和坏疽性脓皮病及其组合。

[0171] 此外,如所指出的,本公开提供了允许从业者可靠且一致地预测对象队列中的应答的技术。特别地,例如,在给定的对象队列中,一些抗TNF疗法的应答率低于35%。所提供的技术允许在对象队列中以超过65%的准确度预测应答率(例如,某些对象是否会对给定治疗产生应答)。在一些实施方案中,本文描述的方法和系统预测给定队列内65%或更多的无应答者(例如,不会对抗TNF疗法产生应答)对象。在一些实施方案中,本文描述的方法和系统预测给定队列内70%或更多的无应答者(例如,不会对抗TNF疗法产生应答)对象。在一些实施方案中,本文描述的方法和系统预测给定队列内80%或更多的无应答者(例如,不会对抗TNF疗法产生应答)对象。在一些实施方案中,本文描述的方法和系统预测给定队列内90%或更多的无应答者(例如,不会对抗TNF疗法产生应答)对象。在一些实施方案中,本文描述的方法和系统预测给定队列内100%的无应答者(例如,不会对抗TNF疗法产生应答)对象。

[0172] 计算机控制系统

[0173] 本公开提供了被编程以实现本公开的方法的计算机控制系统。图10示出了计算机系统1001,其被编程或以其他方式配置为生成或开发自身抗体谱或将自身抗体与特异性免疫应答谱进行比较。计算机系统1001可以调节本公开的各个方面,例如接收或生成序列读段、将序列与特定表位或自身抗体相关联、向用户输出关于自身抗体或谱的存在的结果,或疾病的预期进展。计算机系统1001可以是用户的电子设备或者相对于电子设备位于远程的计算机系统。电子设备可以是移动电子设备。

[0174] 计算机系统1001包括中央处理单元(CPU,本文也称为“处理器”和“计算机处理器”)1005,其可以是单核或多核处理器,或者用于并行处理的多个处理器。计算机系统1001还包括存储器或存储器位置1010(例如,随机存取存储器、只读存储器、闪存)、电子存储单元1015(例如,硬盘)、用于与一个或多个其他系统通信的通信接口1020(例如,网络适配器)以及外围设备1025,例如高速缓存、其他存储器、数据存储或电子显示适配器。存储器1010、存储单元1015、接口1020和外围设备1025通过诸如主板的通信总线(实线)与CPU 1005通信。存储单元1015可以是用于存储数据的数据存储单元(或数据存储库)。计算机系统1001可以借助于通信接口1020可操作地耦合到计算机网络(“网络”)1030。网络1030可以是英特

网、互联网或外联网、或者与英特网通信的内联网或外联网。在一些情况下,网络1030是电信或数据网络。网络1030可以包括一台或多台计算机服务器,其可以实现分布式计算,例如云计算。在一些情况下,在计算机系统1001的帮助下,网络1030可以实现对等网络,这可以使得耦合到计算机系统1001的设备能够充当客户端或服务器。

[0175] CPU 1005可以执行机器可读指令序列,其可以体现在程序或软件中。这些指令可以存储在诸如存储器1010的存储器位置中。这些指令可以被引导至CPU 1005,其可以随后对CPU 1005进行编程或以其他方式配置CPU 1005以实现本公开的方法。由CPU 1005执行的操作的示例可以包括获取、解码、执行和写回。

[0176] CPU 1005可以是电路(例如集成电路)的一部分。系统1001的一个或多个其他组件可以被包括在电路中。在一些情况下,电路是专用集成电路(ASIC)。

[0177] 存储单元1015可以存储文件,例如驱动程序、库和保存的程序。存储单元1015可以存储用户数据,例如用户偏好和用户程序。在一些情况下,计算机系统1001可以包括在计算机系统1001外部的一个或多个附加数据存储单元,例如位于通过内联网或英特网与计算机系统1001通信的远程服务器上。

[0178] 计算机系统1001可以通过网络1030与一个或多个远程计算机系统通信。例如,计算机系统1001可以与用户的远程计算机系统通信。远程计算机系统的示例包括个人计算机(例如便携式PC)、平板电脑或平板电脑(例如**Apple**[®] iPad、**Samsung**[®] Galaxy Tab)、电话、智能手机(例如**Apple**[®] iPhone、支持Android的设备、**Blackberry**[®]),或个人数字助理。用户可以通过网络1030访问计算机系统1001。

[0179] 本文描述的方法可以通过存储在计算机系统1001的电子存储位置上(例如存储器1010或电子存储单元1015上)的机器(例如,计算机处理器)可执行代码来实现。机器可执行或机器可读代码可以以软件的形式提供。在使用期间,代码可以由处理器1005执行。在一些情况下,代码可以从存储单元1015检索并且存储在存储器1010上以供处理器1005准备好访问。在一些情况下,可以排除电子存储单元1015,并且机器可执行指令被存储在存储器1010上。

[0180] 代码可以被预编译并被配置为与具有适于执行代码的处理器一起使用,或者可以在运行时期间被编译。可以以编程语言来提供代码,可以选择该编程语言来使得代码能够以预编译或编译后的方式执行。

[0181] 本文提供的系统和方法的方面,例如计算机系统1001,可以以编程来体现。该技术的各个方面可以被认为是通常采用承载在或体现在一种类型的机器可读介质中的机器(或处理器)可执行代码或相关数据的形式“产品”或“制品”。机器可执行代码可以存储在电子存储单元上,例如存储器(例如,只读存储器、随机存取存储器、闪存)或硬盘。“存储”类型介质可以包括计算机、处理器等的任何或所有有形存储器、或其相关联的模块,例如各种半导体存储器、磁带驱动器、磁盘驱动器等,其可以随时为软件编程提供非暂时性存储。软件的全部或部分有时可以通过英特网或各种其他电信网络进行通信。例如,这样的通信可以使得能够将软件从一个计算机或处理器加载到另一计算机或处理器中,例如从管理服务器或主计算机加载到应用服务器的计算机平台中。因此,可以承载软件元素的其他类型的介质包括光波、电波和电磁波,例如通过有线和光陆线网络以及通过各种空中链路在本地设

备之间的物理接口上使用。承载这种波的物理元件,例如有线或无线链路、光链路等,也可以被认为是承载软件的介质。如本文所使用的,除非限于非暂时性有形“存储”介质,否则诸如计算机或机器“可读介质”之类的术语是指参与向处理器提供指令以供执行的任何介质。

[0182] 因此,诸如计算机可执行代码之类的机器可读介质可以采用多种形式,包括但不限于有形存储介质、载波介质或物理传输介质。非易失性存储介质包括例如光盘或磁盘,诸如任何计算机等中的任何存储设备,诸如可用于实现附图中所示的数据库等。易失性存储介质包括动态存储器,例如这种计算机平台的主存储器。有形传输介质包括同轴电缆;铜线和光纤,包括构成计算机系统内总线的电线。载波传输介质可以采用电或电磁信号、或者声波或光波的形式,例如在射频(RF)和红外(IR)数据通信期间生成的那些波。因此,计算机可读介质的常见形式包括例如:软盘、软盘、硬盘、磁带、任何其他磁性介质、CD-ROM、DVD或DVD-ROM、任何其他光学介质、打孔卡纸带、具有孔图案的任何其他物理存储介质、RAM、ROM、PROM和EPROM、FLASH-EPROM、任何其他存储芯片或盒、传输数据或指令的载波、传输此类载波的电缆或链路,或计算机可以从中读取程序代码或数据的任何其他介质。许多这些形式的计算机可读介质可涉及将一个或多个指令的一个或多个序列传送到处理器以供执行。

[0183] 计算机系统1001可以包括电子显示器1035或与电子显示器1035通信,电子显示器1035包括用户界面(UI) 1040,用于提供例如选择自身抗体进行分析,与将自身抗体与特定生成的谱相关联的图表进行交互。UI的示例包括但不限于图形用户界面(GUI)和基于Web的用户界面。

[0184] 本公开的方法和系统可以通过一种或多种算法来实现。算法可以在由中央处理单元1005执行时通过软件来实现。该算法可以例如计算统计测量值以鉴定自身抗体并生成谱或预测治疗的功效和毒性。

[0185] 实施例

[0186] 实施例1-用于预测类风湿性关节炎中对肿瘤坏死因子- α 抑制剂的不充分应答的分子特征应答分类器。

[0187] 类风湿性关节炎(RA)是一种自身免疫性疾病,其特征导致关节破坏的慢性炎症。在对诸如甲氨蝶呤的合成疾病缓解抗风湿药物(csDMARD)应答不充分之后,临床指南建议许多具有相当疗效和安全性的靶向疗法之一,包括肿瘤坏死因子- α 抑制剂(TNFi)、IL-6抑制剂、Janus激酶(JAK)抑制剂和B或T细胞调节剂。丰富的治疗选择强调了对风湿病领域中精准医学的需求。由于临床指南并不相对于另一种而推荐一种治疗,因此治疗选择通常由行政指令驱动,并且TNFi疗法仍然是近90%RA患者的主要治疗方法。为每位患者匹配正确的靶向疗法,以达到低疾病活动度(LDA)或缓解的治疗到靶标(treat-to-target)目标是RA中一项未得到满足的关键医疗需求。

[0188] RA患者的子集对TNFi疗法有充分的应答:50-70%达到ACR20,30-40%达到ACR50,15-25%达到ACR70应答,并且10-25%达到缓解。许多研究试图在治疗开始前鉴定生物标志物并开发模型来预测对TNFi疗法的应答。未能在新患者群体和临床试验中验证和重现这些预测生物标志物的性能是一个典型的结果。患者群体、实验室方法、生成分子数据的程序之间的不同特征以及单队列回顾性血液研究固有的其他偏差不仅阻碍了风湿病学而且也阻碍了其他医学专业的精准医学进步。

[0189] 利用新的网络医学方法来发现生物标志物,开发了一种基于血液的分子特征测

试,该测试将下一代RNA测序数据与临床特征相结合以预测RA患者对TNFi疗法应答不充分的可能性。该分子特征测试在CERTAIN研究的一部分患者中进行的临床验证表明,具有无应答分子特征的患者在6个月时不太可能达到ACR50。

[0190] 将疾病相关蛋白质映射到人类相互作用组(即人类细胞中发生的成对蛋白质-蛋白质相互作用的网络图)上已经产生了对人类疾病生物学和对治疗的应答的新见解。通过从人类相互作用组分析中发现的与RA生物学相关的分子生物标志物的鉴定,本研究在两项前瞻性观察性临床研究(CERTAIN研究和NETWORK-004)中证明了分子特征应答分类器(MRSC)的TNFi应答预测性能不足。MRSC的验证是在来自未接受过靶向疗法的患者的总共391个血液样品和来自TNFi疗法暴露患者的113个血液样品中进行的。

[0191] 方法

[0192] 患者

[0193] 研究设计Corrona的概要在图6中描述。CERTAIN研究包括:345个RA患者PAXgene血液样品和临床测量,这是一项针对RA患者开始使用生物制剂的比较有效性研究。CERTAIN研究嵌套在Corona注册表中。在样品收集和参与研究之前获得了机构审查委员会或伦理委员会的批准,并且患者提供知情同意书。CERTAIN是一项调查生物制剂引发剂的比较有效性研究。对于这些分析,选择的样品来自在样品收集时未接受过靶向疗法并开始TNFi疗法的患者。92%(318/345)的患者被纳入先前的分类器训练和验证分析中。与CERTAIN研究的纳入标准一致,所有患者在开始生物治疗时的临床疾病活动指数(CDAI)均大于10。临床和分子数据用于生物标志物特征选择(100名患者)和队列内交叉验证(245名患者)。无论每个样品在先前的研究中如何使用,患者都被随机分配到这两个单独的分析中。

[0194] NETWORK-004:在登记之前,治疗风湿病学家将患者确定为TNFi疗法的候选者。符合条件的患者年龄>18岁,患有活动性RA(CDAI>10,肿胀关节计数>4),并且在基线前接受稳定剂量的甲氨蝶呤(>15mg/周)>10周。允许羟氯喹每天不超过400mg的剂量或来氟米特每天不超过20mg的剂量,只要在基线访视前剂量稳定至少4周即可。允许每天<10mg的泼尼松剂量,只要在基线前剂量稳定至少2周即可。第一次研究程序前<2周禁止使用关节内或肠外皮质类固醇。该研究得到了哥白尼集团独立审查委员会(批准号20191082)和当地审查委员会(如果需要)的批准。所有患者均提供了书面知情同意书。TNFi疗法的剂量和治疗由风湿病学家酌情决定。在3个月的随访中,如果临床护理认为适当,风湿病学家可以进行剂量调整。开始第二次TNFi疗法导致对象退出。COVID-19大流行造成的人员流失率比最初预计的要高。在入组的273名RA患者中,168名完成了为期24周的研究。146名患者拥有完整的临床和分子数据并被纳入分析。有关退出研究的患者的信息可参见补充表S1。在113名患者的3个月访视中收集的PAXgene血液样品被分析为TNFi暴露样品。

[0195] TNFi疗法的临床评价和应答

[0196] 特征选择应答定义:临床结果指标例如肿胀和压痛关节计数、患者和医师疾病评估具有固有的可变性。为了鉴定训练队列中已被高置信度分配有应答者和无应答者标签的患者子集,采用蒙特卡罗模拟方法来计算每位患者的置信度结果得分。模拟与实际报告结果之间至少有70%一致性的患者临床结果数据被认为具有高置信度。ACR和EULAR指标的高置信度临床结果用于特征选择。

[0197] CERTAIN研究检查了基线RNA测序数据和临床评估,以根据ACR、CDAI和DAS28-CRP

标准预测3个月和6个月随访时对TNFi疗法的应答。

[0198] NETWORK-004研究检查了在基线、3个月和6个月访问时收集的临床评估:28压痛和肿胀关节计数、患者疼痛的总体评估、患者疾病活动的总体评估、CDAI评分、健康评估问卷和C反应蛋白(CRP)。在基线时记录类风湿因子(RF)和抗环瓜氨酸蛋白(抗CCP)抗体血清状态。每次访问时均收集PAXgene RNA血管。根据ACR、CDAI和DAS28-CRP标准,使用3个月随访的RNA测序数据来预测6个月随访时对TNFi疗法的应答。

[0199] RNA制备和测序分析

[0200] 根据制造商的说明,使用稳定血液PAXgene管RNA分离试剂盒MagMax™ (Thermo Fisher Scientific)从PAXgene RNA管中的全血中提取RNA;使用KAPRNAHyperPrep试剂盒和RiboErase (HMIR) 珠蛋白处理100-1000ng RNA。使用Agilent D1000试剂对样品进行定量。对文库进行高均匀深度测序,目标为>700万个蛋白质编码读段。使用Illumina NextSeq DX 500和NovaSeq 6000仪器对CERTAIN队列进行测序。NETWORK-004样品使用Illumina NovaSeq 6000仪器进行测序,并使用根据临床实验室改进修正案(CLIA)进行的经验证的诊断测定。处理序列数据以确定整个基因组的基因表达。若要纳入分析,对于所有碱基,样品必须具有TapeStation RIN>4、RNA浓度>10ng/μL、测序文库产量>10nM、%完美碱基对指数>85、超过Phred评分30的%碱基>75、平均质量Phred评分>30、中位Phred评分>25、下四分位Phred评分>10。

[0201] 人类相互作用组分析和特征选择

[0202] 为了选择转录物生物标志物特征,从345名患者的队列中随机选择100个样品(CERTAIN研究)。使用随机森林算法通过96轮20%交叉验证计算机实验对蛋白质编码转录物进行排序。通过人类相互作用组分析进一步分析在70/96次迭代中排名前100的特征,以鉴定生物学相关的生物标志物。最终模型中使用了与人类相互作用组上的RA疾病模块重叠³⁵或与疾病模块具有大量连接的生物标志物。使用超几何检验评估连接的显著性。

[0203] 预测分类模型训练和验证

[0204] 评估未包括在生物标志物特征选择中的样品。使用机器学习,将本研究中确定的转录物与先前描述的生物标志物相整合以重新训练应答分类模型。为了评估模型性能,使用前馈人工神经网络进行10折交叉验证。模型构建是使用Python机器学习库sklearn中提供的MLPClassifier包完成的。

[0205] 统计分析

[0206] 使用Python 3.7.6和R版本3.6.1进行统计分析。连续数据用平均值、标准差、中值、最小值、最大值和可评估观测值的数量进行总结。用频率计数和百分比总结了分类变量。在适当的情况下,使用针对连续数据的t分布(分类变量的精确方法)确定置信区间(CI)。所有测试均在双面环境中进行。除非另有说明,假设检验均在两侧0.05显著性水平下进行。所有尝试都是为了限制丢失的数据。没有尝试估算缺失的数据。

[0207] 结果

[0208] 使用人相互作用组鉴定对TNFi疗法无应答的分子特征

[0209] 根据ACR50和EULAR应答定义(参见方法)在6个月时预测对TNFi疗法应答不充分的转录物是使用机器学习从随机选自Corona CERTAIN研究的100名未接受过RA靶向疗法的患者的基线血液样品数据确定的。为了确保转录物反映RA疾病生物学,将所选转录物编码的

蛋白质映射到成对蛋白质-蛋白质相互作用的人类相互作用组图上,以鉴定与RA疾病模块显著相关 (p值<0.05) 的转录物(图1)。TNFi疗法应答特征与由RA疾病相关蛋白组成的人类相互作用组的同一网络邻域重叠。这些特征包括与RA病理学相关的蛋白质,包括JAK3和白细胞介素-1 β (IL-1B)。对TNFi疗法无应答的分子特征包括23个特征:19个RNA转录物和4个临床特征(表6):

[0210] 表6

[0211]	ALPL
	ATRAID
	BCL6
	CDK11A
	CFLAR
	COMM5
	GOLGA1
	IL1B
	IMPDH2
	JAK3
	KLHDC3
[0212]	LIMK2
	NOD2
	NOTCH1
	SPINT2
	SPON2
	STOML2
	TRIM25
	ZFP36
	BMI
	性别
	患者总体
	评估
	抗 CCP

[0213] MSRC的队列内验证

[0214] 通过来自Corona CERTAIN研究的从未接受过靶向疗法的245名患者的独立队列的基线血液样品之间的队列内交叉验证来测试MSRC(表7)。这导致治疗开始后六个月时ACR50、ACR70、CDAI和DAS28应答的AUC值为0.63至0.67(图2A和表8)。在具有和不具有无应

答分子特征的患者之间观察到模型评分的对数似然比的显著差异 (p<0.001) (图2B)。此外，在缺乏无应答分子特征的患者中，根据CDAI和DAS28-CRP在6个月时实现LDA或缓解的患者比例更大(图2C)。

[0215] 表7

[0216]

特征	Corrona CERTAIN 研究特征选 择 (N =100)	Corrona CERTAIN 研究交叉验 证 (N = 245)	NETWOR K-004 研究 未接收过靶 向疗法 (N =146)	NETWOR K-004 研究 TNFi 暴露 (N =113)
年龄(岁)，平均值 (SD)	54 (12.4)	55 (12.3)	58 (14.1)	57 (14.9)
女性，n (%)	72 (72.0%)	179 (73.1%)	115 (78.8%)	87 (77.0%)
疾病持续时间 (年)，中位数(IQR)	1 (1,5)	2 (1, 6)	1 (0, 4.25)	1 (0,4)

[0217]

特征	Corrona CERTAIN 研究特征选 择 (N =100)	Corrona CERTAIN 研究交叉验 证 (N = 245)	NETWOR K-004 研究 未接收过靶 向疗法 (N =146)	NETWOR K-004 研究 TNFi 暴露 (N =113)
种族, n (%)				
白人	83 (83.0%)	213 (86.9%)	113 (77.4%)	92 (81.4%)
黑人	9 (9.0%)	13 (5.3%)	16 (11.0%)	12 (10.6%)
其他	8 (8.0%)	19 (7.8%)	13 (8.9%)	9 (8.0%)
CCP 阳性, n (%)	62 (62.0%)	154 (62.9%)	72 (49.3%)	54 (47.8%)
RF 阳性, n (%)	76 (76.0%)	172 (70.2%)	55 (37.7%)	48 (54.5%)
基线泼尼松, n (%)	30 (30.0%)	64 (26.1%)	33 (22.6%)	23 (20.4%)
泼尼松剂量, 中位数(IQR)	5 (5, 10)	5 (5, 10)	5 (5, 9.38)	5 (5, 5)
当前 csDMARD, n (%)				
甲氨蝶呤	56 (56.0%)	138 (56.3%)	113 (77.4%)	81 (71.7%)
>2 csDMARDs	7 (7.0%)	42 (17.1%)	11 (7.5%)	8 (7.1%)
无	15 (15.0%)	37 (15.1%)	32 (22.6%)	32 (28.3%)
TNFi 使用, n (%)				
阿达木单抗	36 (36.0%)	98 (40.0%)	48 (32.9%)	40 (35.4%)
依那西普	35 (35.0%)	76 (31.0%)	31 (21.2%)	25 (22.1%)
英夫利昔单抗	15 (15.0%)	48 (19.6%)	18 (12.3%)	16 (14.2%)
塞妥珠单抗	10 (10.0%)	17 (6.9%)	13 (8.9%)	7 (6.2%)
戈利木单抗	4 (4.0%)	6 (2.4%)	36 (24.7%)	25 (22.1%)

[0218] 表8

[0219]

交叉验证, 原初	AUC	优势比(95% CI; p-值)
ACR50, 6 个月	0.66	3.0 (1.6-5.5; 0.0002)
ACR70, 6 个月	0.66	3.4 (1.6-7.1; 0.0008)

[0220]

CDAI LDA, 6 个月	0.67	3.7 (2.2-6.4; <0.0001)
CDAI 缓解, 6 个月	0.67	3.4 (1.6-7.6; 0.0014)
DAS28-CRP LDA, 6 个月	0.64	2.5 (1.5-4.3; 0.0005)
DAS28-CRP 缓解, 6 个月	0.65	2.7 (1.6-4.7; 0.0003)

[0221] MSRC在未接受过靶向疗法的患者样品中的前瞻性观察性临床研究中的验证

[0222] 为了进一步验证MSRC预测对TNFi疗法应答不充分的可能性的能力,在多中心观察性临床研究中前瞻性地收集患者样品。经过临床和分子数据质量审查后,146名患者完成了为期24周的研究并被纳入分析。这些患者主要是女性(78.8%)和白人(80.1%),中位年龄为58岁(表7)。TNFi疗法选择由处方医生酌情决定,并且代表了该类别中的所有五种治疗选择(阿达木单抗32.9%、赛妥珠单抗8.9%、依那西普21.2%、英夫利昔单抗12.3%和戈利木单抗24.7%)。44.5%(65/146)的患者在基线时检测到无应答分子特征。

[0223] 根据6个月时对TNFi疗法的ACR50应答的主要终点,MSRC根据患者对TNFi疗法应答不充分的可能性对患者进行分层,AUC为0.64(图3A),优势比为4.1(95%CI:2.0-8.3;p值0.0001)(表9)。

[0224] 其他终点包括根据ACR70、DAS28-CRP缓解(<2.4)或LDA(<2.9)和CDAI缓解(<2.8)或LDA(<10)对3个月时的ACR50应答以及3和6个月时对治疗的应答进行评估。MSRC根据患者在两个时间点应答不充分的可能性和应答标准对患者进行分层,AUC值范围为0.59-0.74(图3A-图3B),显著优势比为3.0-9.1(p值<0.01)(表7)。描述具有无应答分子特征的患者是否未能实现ACR70或DAS28-CRP缓解的优势比在6个月时是显著的(p值<0.0001),但在3个月时则不显著(p值分别为0.07和0.34)。对于除3个月时DAS28-CRP缓解之外的所有应答标准,观察到具有和不具有无应答分子特征的患者之间模型评分的显著差异(p值<0.002)(图3B-3C)。此外,在不具有无应答分子特征的患者中,根据CDAI和DAS28-CRP定义在6个月时实现缓解和LDA的患者比例更大(图3E-3F)。

[0225] 表9

[0226]

	AUC	优势比(95%CI;p值)
前瞻性观察性,原初		
ACR50,3个月	0.67	3.5(1.6-7.7;0.002)
ACR50,6个月	0.64	4.1(2.0-8.3;0.0001)
ACR70,3个月	0.67	2.6(1.0-6.7;0.07)
ACR70,6个月	0.72	9.1(3.5-24.2;<0.0001)
CDAI LDA,3个月	0.65	3.0(1.4-6.1;0.005)
CDAI LDA,6个月	0.68	3.6(1.8-7.2;0.0002)
CDAI缓解,3个月	0.70	3.4(1.3-8.7;0.01)
CDAI缓解,6个月	0.74	8.4(3.0-23.9;<0.0001)
DAS28-CRP LDA,3个月	0.69	3.3(1.5-7.1;0.002)
DAS28-CRP LDA,6个月	0.67	3.4(1.7-6.8;0.0007)

DAS28-CRP缓解,3个月	0.59	1.5 (0.7-3.2;0.34)
DAS28-CRP缓解,6个月	0.74	5.2 (2.5-11.1; <0.0001)
前瞻性观察性,TNFi暴露		
ACR50,6个月	0.67	3.3 (1.5-7.4; <0.0001)
ACR70,6个月	0.76	7.3 (2.3-23.3;0.00004)
CDAI LDA,6个月	0.74	7.5 (3.2-17.6; <0.0001)
CDAI缓解,6个月	0.84	25.4 (3.2-200.6; <0.0001)
DAS28-CRP LDA,6个月	0.71	6.2 (2.6-14.9; <0.0001)
DAS28-CRP缓解,6个月	0.65	2.0 (0.9-4.6;0.1)

[0227] 在TNFi暴露的患者样品中的前瞻性观察性临床研究中MSRC的验证

[0228] 在完成24周研究的患者中,113名患者可获得3个月时的RNA血液样品。使用与未接受过靶向疗法的分析相同的MSRC,使用3个月的患者样品来预测对TNFi疗法的应答不充分。这些TNFi暴露的样品中的分子特征根据对治疗的应答不充分对患者进行分层,AUC值为0.65至0.84(图4A)。针对40.7%(46/113)的TNFi暴露患者检测到无应答分子特征,并且在有和没有无应答分子特征的患者之间观察到模型评分存在显著差异(p值<0.012)(图4B)。这对应于未能根据除DAS28-CRP缓解以外的所有标准对治疗产生应答的具有分子特征的患者中3.3-25.4的显著优势比(表9)。

[0229] 讨论

[0230] 许多靶向疗法选择可用于RA,但疗法选择是一个挑战,因为这些选择具有相似的治疗结果。非常需要精准医疗工具来确定哪些患者具有适合每种靶向疗法的疾病生物学。所描述的基于血液的MSRC分析RNA测序数据以及临床特征,以准确鉴定未接受过靶向疗法的患者和TNFi暴露的患者,这些患者不太可能对TNFi疗法产生充分应答。在开始首次靶向疗法的患者中,那些具有无应答分子特征的患者对TNFi疗法产生充分应答的可能性要低三到九倍(表7)。当患者接受TNFi疗法至少三个月后进行测试时,具有无应答分子特征的患者获得缓解的可能性降低了25倍。此外,根据包括ACR50、ACR70、DAS28-CRP和CDAI在内的多项临床验证指标,无应答分子特征可预测TNFi疗法应答不充分。MSRC可以在护理途径中的多个场合为提供者决策提供信息,例如在初始疗法选择之前或当靶向TNFi疗法未达到治疗目标并且正在考虑第二治疗或剂量递增时。通过验证多个应答目标定义,MSRC适合多个实践方案,使其易于在临床环境中理解、采取行动和操作。

[0231] 精准医学通过将治疗选择与患者独特的生物学相匹配而改善了肿瘤学和血液学方面的患者结果。然而,即使在这些领域,预测药物应答(特别是从血液)仍然是一个具有挑战性的技术问题。此外,机器学习和统计方法往往过度拟合研究人群的特征和属性。与肿瘤学不同,在RA患者护理中,对疾病组织的DNA分析和活检的依赖并不容易接受,因为临床研究之外的滑膜活检很少见,而且DNA序列变异在RA中提供的可操作信息有限。AMPLE、AVERT、GO-BEFORE和GO-FORWARD试验的研究使用了基线疾病评估如DAS28、RAPID3、CDAI或SDAI来预测响应于靶向疗法治的放射学进展或磁共振成像检测到的滑膜炎。报道了相当的AUC值(0.54-0.72),但优势比(1.01-1.65)低于本研究中观察到的值(3.0-25.4)。此外,本研究中交叉验证CERTAIN队列和前瞻性NETWORK-004队列之间的优势比一致,表明MSRC在研究和患者群体中具有可重复性和普适性。正如本研究结果所证明的那样,围绕精准医疗工具开

发的技术挑战强调了评估与疾病生物学相关的生物标志物和开发新方法(例如基于网络的方法)的重要性。

[0232] 本研究中使用的基于网络的方法揭示了疾病生物学中的新联系。对248名基于美国的风湿病学家的一项调查表明,风湿病学家可能会欢迎自身免疫性疾病领域的精准医学进步,并可能会发现这种预测性药物应答测试的价值。当风湿科医生收到表明无应答的样品MSRC结果时,TNFi疗法的选择下降了超过80%(从79.8%降至低至11.3%)。此外,大多数接受调查的风湿病学家报告说,测试结果可以增加他们对处方决策的信心,改善医疗决策并改变他们的治疗选择。对以预测对TNFi疗法的应答的精准医学工具为指导的治疗选择建模,以提高对靶向疗法的应答率并节省医疗成本。

[0233] MSRC中的RNA转录物评估疾病生物学的看似不同的方面,尽管这些方面在人类相互作用组上的同一网络邻域中是统一的,并且捕获RA的不同生物学和对TNFi疗法的应答。这些转录物编码的蛋白质影响生物过程,包括适应性和先天免疫细胞的细胞稳态、TNF- α 和其他分泌信号传导分子的产生、滑膜炎和骨破坏(图5)。TNF- α 生物学在MSRC中被可靠捕获,并且特征涉及TNF- α 的产生和释放(例如,COMMD5)以及上游或下游TNF- α 信号传导事件(例如,NOTCH1)。循环血细胞中表达的分子特征的鉴定表明,直接评估关节生理学或生物化学对于评估滑膜表型或对治疗的应答可能不是必需的。MSRC植根于RA疾病生物学,并且很容易推广到盲法研究中独立患者队列的分子表型。

[0234] 结论

[0235] MSRC的验证涉及对RNA测序数据的分析,该数据源于来自两项独立研究和患者群体的用TNFi疗法治疗的391名RA患者的血液样品,首次使用看似不同的RA生物学再现了分子生物标志物的预测能力。这些发现表明,直接评估关节生理学或生物化学对于预测应答可能并不重要。在未接受过靶向疗法的患者和TNFi暴露患者中,根据ACR50、ACR70、DAS28-CRP和CDAI的评估,具有无应答分子特征的患者在3或6个月时不太可能对TNFi疗法产生应答。当提供者使用MSRC测试结果对患者进行治疗分层时,具有对TNFi疗法的无应答分子特征的患者可以被指引到替代疗法,以避免费用和潜在的毒性,而没有可能的益处。那些缺乏这种特征的患者可以继续进行TNFi疗法,并且相对于未分层的人群可能会获得更高的缓解率。

[0236] 实施例2-用于预测类风湿性关节炎患者对肿瘤坏死因子抑制剂疗法无应答的RNA特征组的临床寿命

[0237] 作为一种潜在的使人衰弱的自身免疫性疾病,类风湿性关节炎(RA)是涉及关节恶化和慢性炎症的指示性临床表现。尽管这种疾病无法治愈,但RA患者确实有多种可用疗法以减轻症状并防止关节破坏。治疗指南表明,早期治疗干预对于延缓伴随组织结构损伤的关节功能永久性丧失非常重要。一旦患者被诊断患有RA,应用的第一疗程可以是合成疾病缓解抗风湿药物(csDMARD),其中甲氨蝶呤作为一线选择。根据治疗指南,csDMARD未能充分控制症状的RA患者可采用多种其他疗法,包括用于抑制白细胞介素6(IL-6)、Janus激酶(JAK)和肿瘤坏死因子 λ (TNF)的靶向药物。虽然靶向疗法被认为是csDMARD之外治疗的下一步,但在这些情况下,没有任何一种疗法比其他靶向疗法更值得推荐,并且疗法的选择可能取决于非临床选择因素。这一点通过>80%的具有通过csDMARD不能充分控制的症状的未接受过生物制剂的RA患者随后被引导至抗TNF疗法来证明。

[0238] 在对csDMARD应答不充分并随后开始抗TNF药物治疗的大患者亚群中,这些RA患者中的大约75-90%没有达到美国风湿病学会(ACR)指南中的低疾病活动度(LDA)或缓解的预期治疗目标。在刚刚开始使用TNF抑制剂(TNFi)的RA患者中,在50-70%的患者中见到症状改善20%(ACR20),在30-40%的患者中观察到50%ACR评分改善(ACR50),并且15-25%的患者达到70%的改善(ACR70)。据报道,在开始TNFi治疗的未接受过生物制剂治疗的RA患者中,只有不超过约10-25%能够实现RA症状的缓解,这表明TNFi疗法的广泛应用尚未满足对精准医疗以减少治疗循环的可能性的需求。鉴于RA随着时间的退行性性质,减少达到治疗目标的延迟可以为抗TNF治疗无应答的RA患者带来生活质量改善。

[0239] 在目前TNFi使用的情况下,先前的研究表明开始此类治疗的RA患者正在接受对其特定生物学而言次优的治疗方案。这导致医疗保健系统中浪费大量金钱用于支付无法从治疗中受益的患者的TNFi疗法费用,并且这也延迟了将这些RA患者引导至具有可能更适合他们的具体情况替代作用机制的疗法。鉴于目前的情况对大多数RA患者和医疗资源效率都没有帮助,引入能够预测患者对TNFi无应答性的经过验证的生物标志物组可能非常有益,并被风湿病学家良好接受。由Scipher Medicine开发的称为PrismRA的专有生物标志物组和预测算法可分析19个RNA转录物、抗环瓜氨酸肽(抗CCP)实验室测试以及3个临床指标(BMI、性别和患者整体评估),已被证明可以预测未接受过生物制剂治疗的RA患者对抗TNF疗法的应答。目前,患者开始TNFi治疗后PrismRA结果保持有效的时间长度尚不清楚。目前的研究旨在评估PrismRA预测在整个TNFi疗法的时间过程中的稳定性,旨在确定PrismRA评分在人群和个体水平上的长期临床意义。

[0240] 研究群体

[0241] 本文评估的患者群体的人口统计学概述于表10中。从330名类风湿性关节炎患者获得总共452个全血样品和伴随的临床测量。在RA患者开始抗TNF疗法后从其采集样品。该研究中的所有患者在开始TNFi疗程之前均未接受过RA生物制剂。在TNFi开始后3个月或6个月收集RA患者的血液样品,并对在两个时间点提供样品的患者进行横截面分析。在患者群体中,94名患者仅在TNFi开始后3个月提供样品,114名患者仅在6个月时间点提供样品,122名患者在3个月和6个月时间点均提供样品。三个患者组和两个样品收集时间点之间的叠加如图7所示。参与本研究的所有患者均提供了知情同意书,并在患者进行任何样品收集或研究参与之前获得了机构审查委员会的批准。所有患者的TNFi疗法和相关剂量的选择均由风湿病学家自行决定。

[0242] 临床评价和对抗TNF疗法的应答

[0243] 根据针对ACR、临床疾病活动指数(CDAI)和使用C反应蛋白的疾病活动评分28(DAS28-CRP)定义的标准,在基线、3个月和6个月就诊时评估对抗TNF疗法的临床应答。ACR50和ACR70的ACR测量被定义为当个体在28压痛关节计数、28肿胀关节计数以及用于评估RA患者的疾病状态的五个临床值中的至少三个表现出 $\geq 50\%$ 或 $\geq 70\%$ 的改善。每次访视时均采集PAXgene RNA血液管中的全血样品。在患者的基线采样点建立类风湿因子(RF)和抗环瓜氨酸蛋白(抗CCP)抗体血清状态测量。用于评估RA患者疾病状态的变量包括健康评估问卷残疾指数、患者总体评估、提供者总体评估、CRP和抗CCP水平以及患者报告的疼痛。临床评估数据还用于确定患者是否满足CDAI低疾病活动度(CDAI-LDA)、CDAI缓解(CDAI-R)、DAS28-CRP低疾病活动度(DAS28-CRP-LDA)和DAS28-CRP缓解(DAS28-CRP-R)的临床阈

值。

[0244] RNA分离、制备和测序分析

[0245] PAX-gene血液RNA管用于收集血液样品以用于总RNA分离。来自Thermo Fisher Scientific的稳定血液PAXgene管RNA分离试剂盒MagMaxTM根据制造商的方案用于RNA样品制备。使用具有RiboErase(HMR)珠蛋白的KAPA RNA HyperPrep试剂盒处理100-1000ng质量范围内的RNA。使用安捷伦生物分析仪(Agilent Bioanalyzer)自动电泳平台评估收集的RNA的质量,同时使用NanoDrop ND-8000分光光度计进行RNA定量。使用Illumina NovaSeq 6000平台和经临床实验室改进修正案(CLIA)验证的诊断测定对RNA样品进行测序。整个基因组的基因表达由经处理的序列数据确定。为了纳入样品分析,对于RNA分子中的所有碱基,RNA样品需要具有TapeStation RIN>4、RNA浓度 $\geq 10\text{ng}/\mu\text{L}$ 、测序文库产量 $\geq 10\text{nM}$ 、完美碱基对指数百分比>85、超过Phred评分30的碱基百分比>75、平均质量Phred评分>30、中位Phred评分>25并且下四分位Phred评分>10。

[0246] TNFi应答预测模型

[0247] 使用23个选定生物标志物的组,使用从RA患者收集的245个样品来训练TNFi疗法应答分类模型。使用Python机器学习库sklearn中提供的MLPClassifier进行模型构建。

[0248] 统计分析

[0249] 使用接受者操作特征(ROC)曲线下面积(AUC)来评估PrismRA生物标志物组的性能。用于优势比计算的MSRS模型截止值是根据之前的验证结果选择的。计算优势比。Python 3.7.6用于执行所有统计分析和数据处理程序。所有可分类为连续数据的值均以平均值、标准差、中值、最小值、最大值和适当的观察计数表示。对于分类变量,使用频率计数和百分比来汇总各个值。为了确定置信区间(CI),使用t分布获得连续数据CI,而使用精确方法确定分类变量的CI。除非另有说明,在所有情况下都采用0.05显著性水平进行双边检验。

[0250] PrismRA在整个抗TNF暴露时间过程中保持性能

[0251] 表10显示了本研究中评估的患者群体的人口统计数据。总共评估了从最近开始TNF疗法的330名RA患者收集的452个样品。在TNF开始后3个月或TNF开始后6个月收集样品。图7显示了在3个月和6个月时间点提供样品的患者的叠加。在该研究的330名患者中,94名患者仅在3个月的时间点提供了样品,122名患者在3个月和6个月的时间点均提供了样品,114名患者仅在6个月的时间点提供了样品。

[0252] 表10

[0253]	特征	n=330
	年龄, 平均值(SD)	55.5 (12.5)
	性别, 女性/男性	238/92
	CCP, 阳性/阴性/未知	211/104/15
	RF, 阳性/阴性/未知	233/83/14
	甲氨蝶呤, 是/否	235/95
	泼尼松, 是/否	75/255
	羟氯喹, 是/否	50/280
	磺胺吡啶, 是/否	10/320
	Arava, 是/否	22/308
	TNF, Humira [®] /Enbrel [®] /Remicade [®] /Simponi [®] /Cimzia [®]	124/107/65/9/25

[0254] 使用在两个时间点收集的患者数据,使用分子特征应答分类器(MSRC)来预测对TNF疗法的治疗应答。使用七种不同的临床可接受的应答定义(ACR20、ACR50、ACR70、CDAI-R、CDAI-LDA、DAS28-CRP-R和DAS28-CRP-LDA),在样品采集后+3个月和+6个月评估对TNF疗法的患者应答。更多细节参见材料和方法。图8显示了通过将MSRC评分与+3个月和+6个月的治疗结果进行比较而产生的ROC曲线。

[0255] 与使用6个月的数据(图8c/d)进行预测相比,当使用3个月的数据(图8a/b)时观察到相当的性能。在各种应答定义中,使用3个月数据时,AUC范围为0.66-0.73,并且使用6个月数据时,AUC范围为0.67-0.75。当将模型预测与数据收集时间后+3个月(图8a/c)和+6个月(图8b/d)的治疗应答结果进行比较时,观察到类似的表现,AUC范围分别为0.67-0.79和0.66-0.76。表11总结了每个样品和终点定义中观察到的优势比。对于所有评估的模型,观察到应答者和无应答者之间的分数分布存在统计学显著差异($p<0.001$)。

[0256] 表11

	时间点	终点	AUC	OR	[CI]
[0257]	3 个月	ACR50+3	0.73	4.83	[2.42-9.62]
	3 个月	das28crp_r+3	0.71	4.65	[2.36-9.15]
	3 个月	das28crp_lda+6	0.67	3.53	[1.87-6.64]
	3 个月	das28crp_lda+3	0.72	5.73	[2.99-10.99]
	3 个月	cdai_r+6	0.68	3.12	[1.35-7.18]
	3 个月	cdai_r+3	0.72	5.03	[1.86-13.54]
	3 个月	das28crp_r+6	0.66	4.42	[2.1-9.29]
	3 个月	cdai_lda+3	0.72	5.49	[2.98-10.11]
	3 个月	cdai_lda+6	0.67	3.04	[1.67-5.53]
	3 个月	ACR70+6	0.68	3.65	[1.52-8.76]
	3 个月	ACR70+3	0.73	4.63	[1.85-11.61]
	3 个月	ACR50+6	0.72	3.63	[1.88-6.99]
	6 个月	cdai_lda+3	0.7	4.32	[2.21-8.43]
	6 个月	cdai_lda+6	0.71	5.98	[2.91-12.26]
	6 个月	ACR70+3	0.75	7.71	[2.64-22.48]
[0258]	6 个月	cdai_r+3	0.79	7.19	[2.46-21.0]
	6 个月	cdai_r+6	0.76	7.14	[2.45-20.87]
	6 个月	das28crp_lda+3	0.71	4.69	[2.16-10.18]
	6 个月	ACR50+6	0.68	3.34	[1.69-6.59]
	6 个月	das28crp_lda+6	0.67	4.8	[2.07-11.13]
	6 个月	das28crp_r+3	0.67	5.22	[2.14-12.72]
	6 个月	ACR50+3	0.75	6.36	[2.92-13.87]
	6 个月	ACR70+6	0.74	5.18	[2.08-12.88]
	6 个月	das28crp_r+6	0.71	8.5	[2.82-25.58]

[0259] 通过比较提供3个月样品和6个月样品两者的122名患者之间的模型性能来进一步评估MSRC预测的稳定性(图9)。为了确保结果一致,在TNF疗法开始后+9个月(从3个月样品收集时间点起+6个月和从6个月样品收集时间点起+3个月)评估应答。在各种应答定义中,当使用TNF开始后3个月收集的数据时,AUC范围为0.66-0.74,并且当使用TNF开始后6个月收集的数据时,AUC范围为0.65-0.73。在这两种情况下,观察到有应答者和无应答者之间的MSRC评分分布存在统计学显著差异($p<0.001$)。

[0260] PrismRA预测在整个抗TNF暴露时间过程中表现出稳定性

[0261] 为了在个体基础上评估PrismRA应答预测的纵向稳定性,我们首先研究了可获得3个月和6个月数据的122名患者中应答结果标签的稳定性。表12详细列出了考虑TNF开始后3个月和6个月收集的数据时结果之间的一致性程度。平均而言,在不同终点定义中,+3个月的结果在73.9%的时间内保持一致,而+6个月的结果在80.9%的时间内保持一致。在不一

致的患者中,从无应答者到有应答者的比例与从无应答者到有应答者的比例相似。对于+3个月的结果,有平均14%从无应答者到有应答者变化,有12%从有应答者到无应答者变化。对于+6个月的结果,有8.5%从无应答者到有应答者变化,有10.5%从有应答者到无应答者变化。

[0262] 表12

终点	NR-NR	R-R	一致	NR-R	R-NR	不一致
ACR50+3	52	33	72.0%	18	15	28.0%
	(44.1%)	(28.0%)		(15.3%)	(12.7%)	
ACR50+6	55 (49.1%)	37 (33.0%)	82.1%	8 (7.1%)	12 (10.7%)	17.9%
ACR70+3	79 (66.9%)	17 (14.4%)	81.4%	11 (9.3%)	11 (9.3%)	18.6%
ACR70+6	77 (68.1%)	20 (17.7%)	85.8%	8 (7.1%)	8 (7.1%)	14.2%
cdai_lda+3	27 (22.7%)	55 (46.2%)	68.9%	19 (16.0%)	18 (15.1%)	31.1%
cdai_lda+6	31 (27.2%)	59 (51.8%)	78.9%	11 (9.6%)	13 (11.4%)	21.1%
cdai_r+3	85 (71.4%)	17 (14.3%)	85.7%	11 (9.2%)	6 (5.0%)	14.3%
cdai_r+6	80 (70.2%)	19 (16.7%)	86.8%	6 (5.3%)	9 (7.9%)	13.2%
das28crp_r+3	54 (47.4%)	29 (25.4%)	72.8%	15 (13.2%)	16 (14.0%)	27.2%
das28crp_r+6	53 (50.0%)	31 (29.2%)	79.2%	11 (10.4%)	11 (10.4%)	20.8%
das28crp_lda+3	33 (28.9%)	43 (37.7%)	66.7%	23 (20.2%)	15 (3.2%)	33.3%
das28crp_lda+6	34 (32.1%)	46 (43.4%)	75.5%	12 (11.3%)	14 (13.2%)	24.5%

[0265] 通过将使用TNF疗法开始后三个月收集的数据做出的预测与使用TNF疗法开始后六个月收集的数据做出的预测进行比较,来评估在整个TNF疗法时间过程中MSRC预测的稳定性。在可获得3个月数据和6个月数据二者的122名患者中,97名患者(81.5%)在两个时间点之间具有一致的类别预测,而22名患者(18.5%)在两个时间点之间具有不同的预测。在发生变化的18.5%中,有9名从无应答者变为有应答者,12名从有应答者变为无应答者。

[0266] 讨论

[0267] 类风湿性关节炎的生物疗法可以针对一系列不同的靶标(TNF、IL-6和JAK),并且当患者对疗法有应答时具有大致相同的益处,然而风湿病学家最常见的生物治疗选择是TNFi。在没有额外的临床指导的情况下,TNFi的主要使用肯定会继续下去,这意味着90%对csDMARD没有充分应答的患者将接受有70%可能性无法达到RA治疗到靶标阈值的药物。然而,通过实施可以确定患者是否会对TNFi疗法产生应答的临床组,可以减轻这些负面结果。PrismRA预测生物标志物组已在之前的研究中经过临床验证,可成功将RA患者分类为TNFi应答者或无应答者。

[0268] 本研究的一个目标是确定PrismRA MSRC在TNF疗法的整个时间过程中的功效。在群体水平上,模型性能良好并且反映了之前的验证结果。MSRC能够区分无应答者和应答者,当使用3个月数据时AUC范围为0.61-0.69,并且当使用6个月数据时AUC范围为0.64-0.73。这与Mellors等人的观察结果非常吻合,其中生物标志物应答组以6.57的优势比成功鉴定出TNFi无应答者。此外,Cohen等人之前对MSRC的验证成功对患者成为TNFi无应答者的可能性进行了分层,根据6个月时的ACR50应答终点,优势比为4.1。

[0269] 无论在整个抗TNF疗法时间线中何时收集样品,+3个月和+6个月结果的预测性能保持一致。PrismRA结果的稳定性表明,生物标志物组可以在TNFi治疗过程中的任何时间使用,同时仍然提供关于患者在接下来的3至6个月内对疗法的应答的有效预测。这项研究并不是仅在治疗开始时提供有效的治疗指导并随着时间的推移而变化,而是显示了该生物标志物组在TNFi疗法时间过程中的长期有效性。

[0270] 该研究的结果确实具有一定程度的变异性,这可以从+3个月结果的大约74%和+6个月结果的大约81%的一致性率推断出。错误预测的发生或患者在连续测量中在有应答者和无应答者之间切换都凸显了了解自然结果变异性以正确表征模型性能的重要性。

[0271] 在不一致的结果中,从应答者变为无应答者的患者比例与从无应答者变为应答者的患者比例大致相同。平均而言,无应答者转变为应答者的比例在+3个月时间点为14%,并且在+6个月时间点为8.5%,而应答者转变为无应答者的比例被发现在+3个月时间点为12%,并且在+6个月时间点为10.5%。由于在测试时为无应答者的接受抗TNF疗法的患者有较小可能性可以转变状态成为应答者,因此一些临床医生可能倾向于让无应答患者继续接受TNFi,以防万一他们最终有应答。然而,RA疾病进展的时间敏感性和使人衰弱的性质可以保证临床决策的成功概率高于等待无效疗法获得最终缓解的机会。美国风湿病学会在2021年指南中调整了bDMARD类抗TNF疗法的建议,其中未达到目标改善的患者应改用不同药物类别的bDMARD,而不是同一类别中的不同bDMARD。据报道,在TNF抑制剂的替代品中,IL-6、IL-6受体和JAK的抑制剂等选项的效果与TNF抑制剂大致相同(29-36)。此外,据报道,即使在使用抗TNF疗法且无应答后,其他类别的bDMARD对RA患者仍然有效,因此在初始TNFi治疗后可能从有应答者转变为无应答者的患者仍然可以使用生物制剂作为有效的治疗选择(37-39)。更新的ACR建议和TNFi疗法有效替代品的可用性证实了将患者分层为TNFi疗法有应答者或无应答者的动机,这样就不会浪费时间尝试人群应答率较差的疗法或继续使用患者不在有应答的药物。在一项关于PrismRA组的感知的临床实用性的研究中,Pappas等人发现,能够对患者TNFi应答进行分类的MSRC可能会受到风湿病学家的好评。在接受调查的248名临床医生中,92%认为测试结果可以提高他们在决定RA患者治疗时的信心,大约80%的接受调查的风湿病学家同意这种类型的生物标志物组可以改善医疗决策。

[0272] 当将使用3个月数据做出的预测与使用6个月数据做出的预测进行比较时,MSRC表现出高水平的一致性。在这两个时间点进行的测量之间进行比较的结果在所有情况下保持一致性>81%。这些数据表明,即使在TNFi时间过程中的不同点给出MSRC,该组仍将产生一致的结果。结果证实了在治疗选择时使用MSRC来鉴定TNFi疗法无应答者的有效性和实用性(24,26,40-43),这项纵向研究表明,医生在初始测试后可能会或可能不会重新测试他们的患者。相应地,当因疾病进展或当前疗法的副作用导致需要改变处方时,建议患者接受PrismRA测试。计划进行测量PrismRA测试间隔的进一步研究,以更好地定义患者测试建议。

[0273] 知道PrismRA测试具有持续的准确性,可以避免指示重复测试或不确定结果的进一步浪费和低效,并且可以实现基于其分子谱将患者置于最有效治疗的成本节省。Bergman等人描述了在临床决策中通过PrismRA预测TNF抑制剂无应答所节省的成本的近似值。与在RA患者治疗12个月中使用PrismRA分层相比对标准护理生物药剂治疗成本进行建模时,可以获得无效治疗花费的成本降低22%,并且RA治疗总体成本降低5%。在符合医疗保险资格的人群中,这些节省相当于每名患者每年在无效治疗上的支出减少了6668美元。考虑到PrismRA分层直接减少无效治疗并且通过减缓疾病进展带来间接价值,因此RA可能带来的多重益处增强了在临床领域推进精准医疗的价值。

[0274] 前面已经描述了其中描述的主题的某些非限制性实施方案。因此,应当理解,本说明书中描述的实施方案仅是对其中报告的主题的说明。对所示实施方案的细节的引用并不旨在限制权利要求的范围,权利要求本身叙述了被认为是必要的那些特征。

[0275] 可以预期的是,所要求保护的主题的系统和方法涵盖使用来自其中所描述的实施方案的信息而开发的变化和改编。相关领域的普通技术人员可以对其中描述的系统和方法进行适应、修改或两者。

[0276] 在整个说明书中,当系统被描述为具有、包括或包含特定组件时,或者当方法被描述为具有、包括或包含特定步骤时,可以预期的是,另外存在由本发明涵盖的基本上由所列举的组件组成或由所列举的组件组成系统,并且存在由本发明涵盖的基本上由所列举的处理步骤组成或由所列举的处理步骤组成的方法。

[0277] 应当理解,步骤的顺序或执行某些动作的顺序并不重要,只要其中描述的主题的任何实施方案保持可操作即可。此外,两个或更多个步骤或动作可以同时进行。

[0278] 虽然本文已经示出和描述了本发明的优选实施方案,但是对于本领域技术人员来说显而易见的是,仅通过示例的方式提供了这样的实施方案。现在,本领域技术人员将构思到许多变化、改变和替代而不背离本发明。应当理解,在实践本发明时可以采用本文描述的本发明实施方案的各种替代方案。所附权利要求旨在限定本发明的范围,并且由此涵盖这些权利要求及其等同物的范围内的方法和结构。

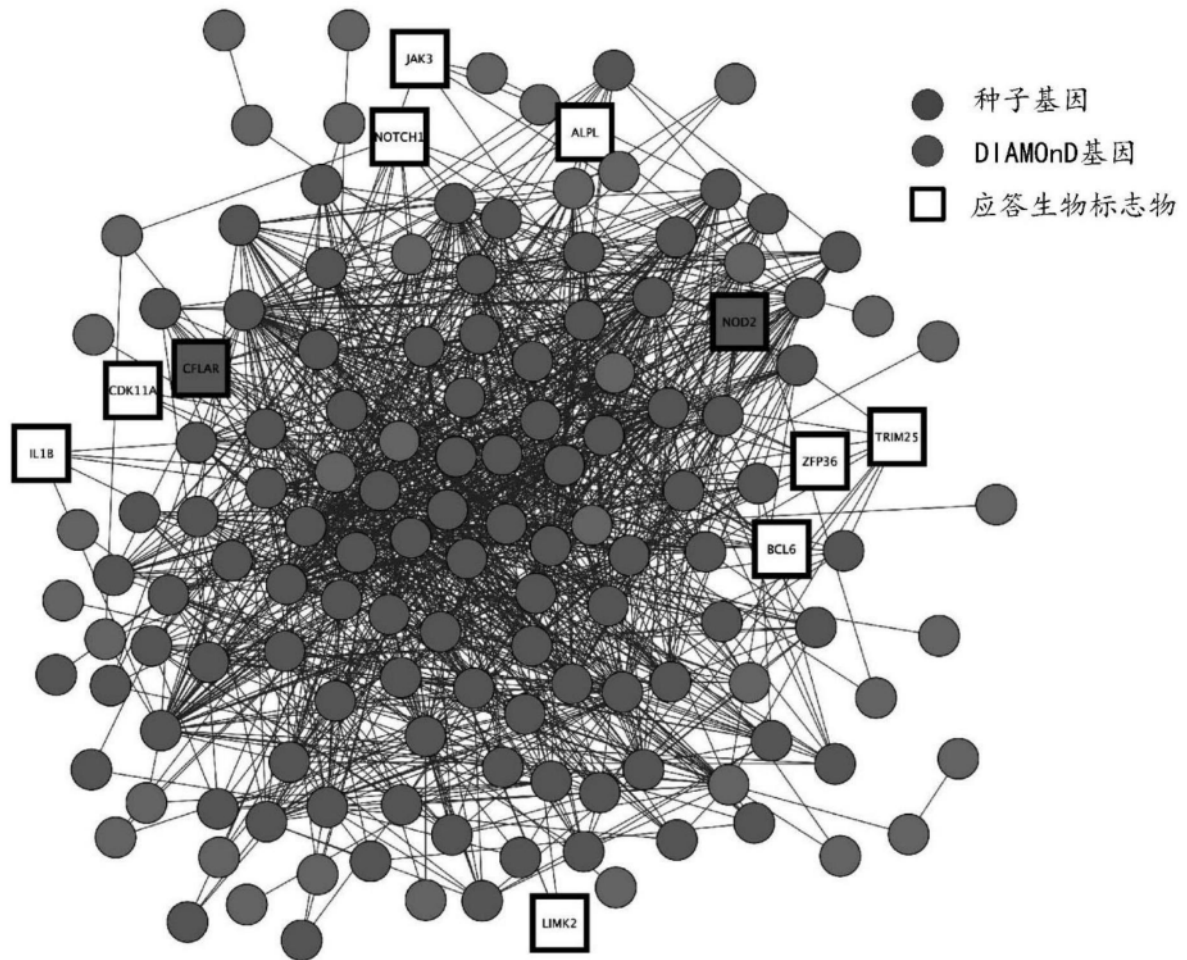


图1

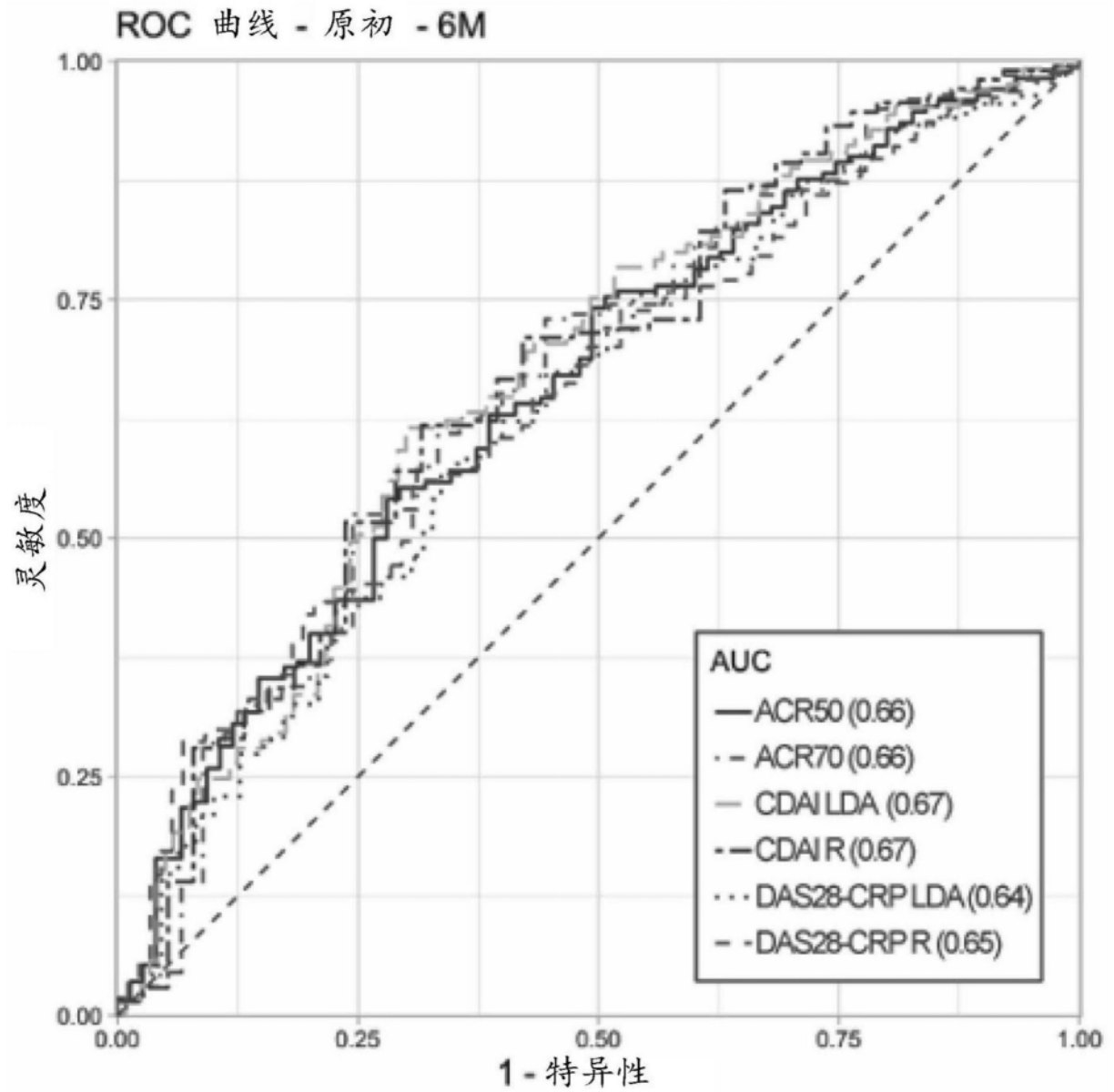


图2A

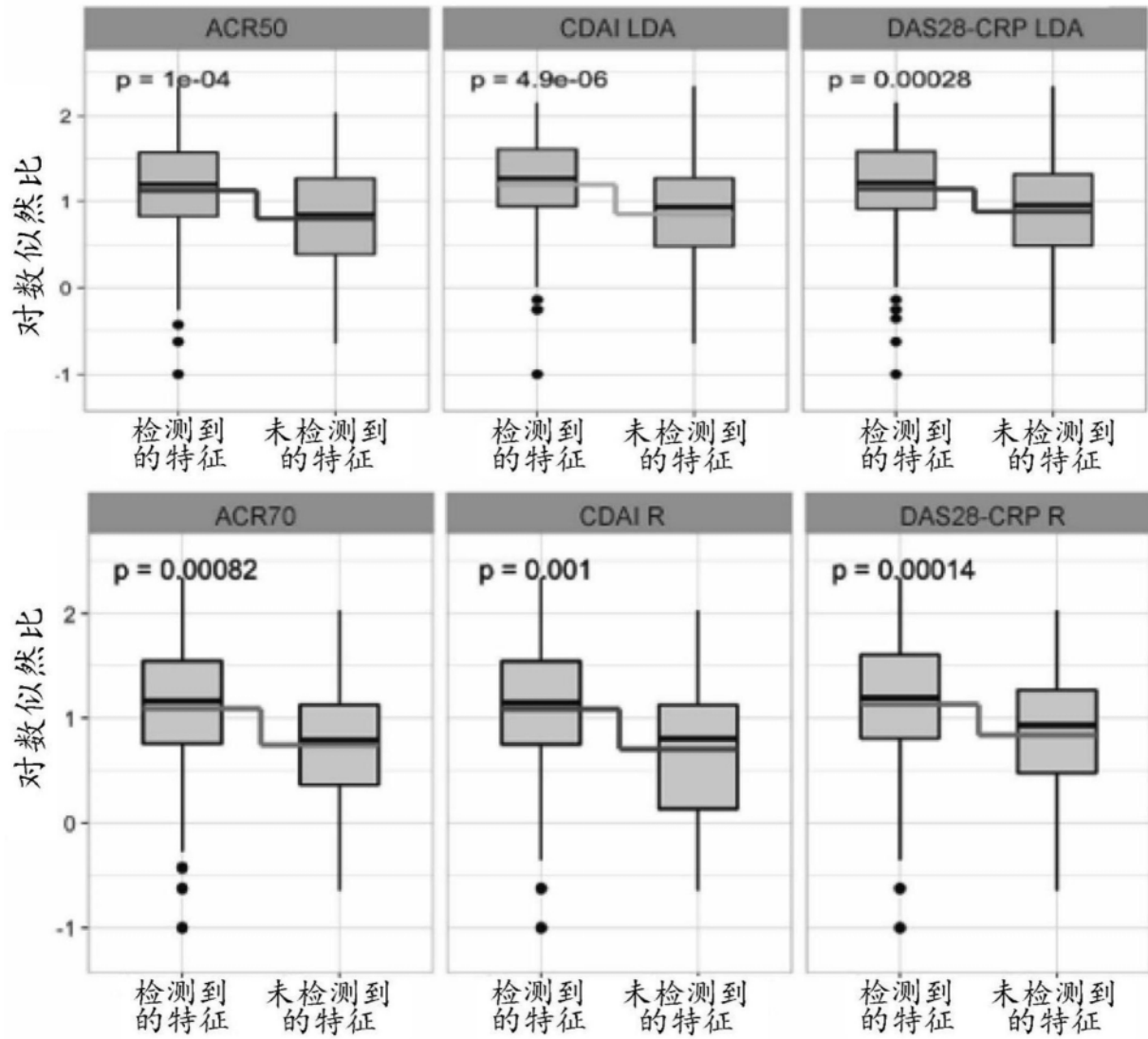


图2B

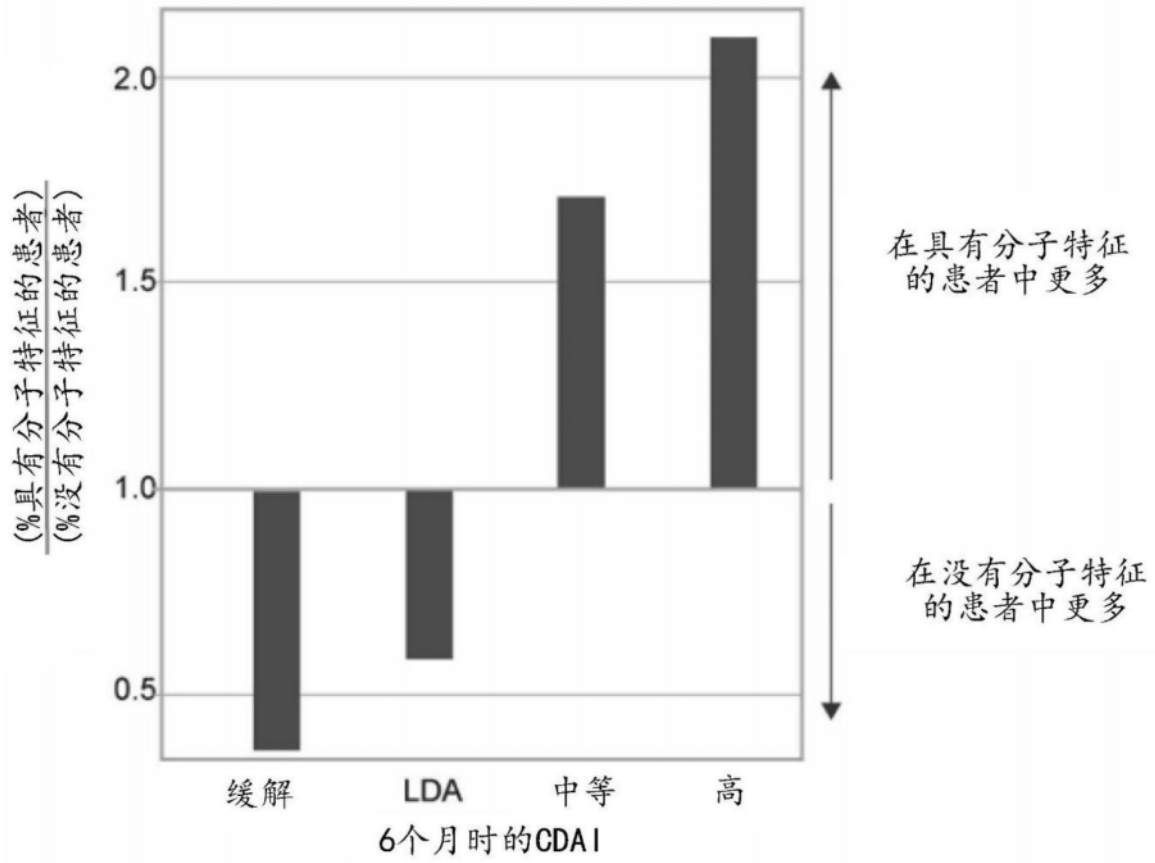


图2C

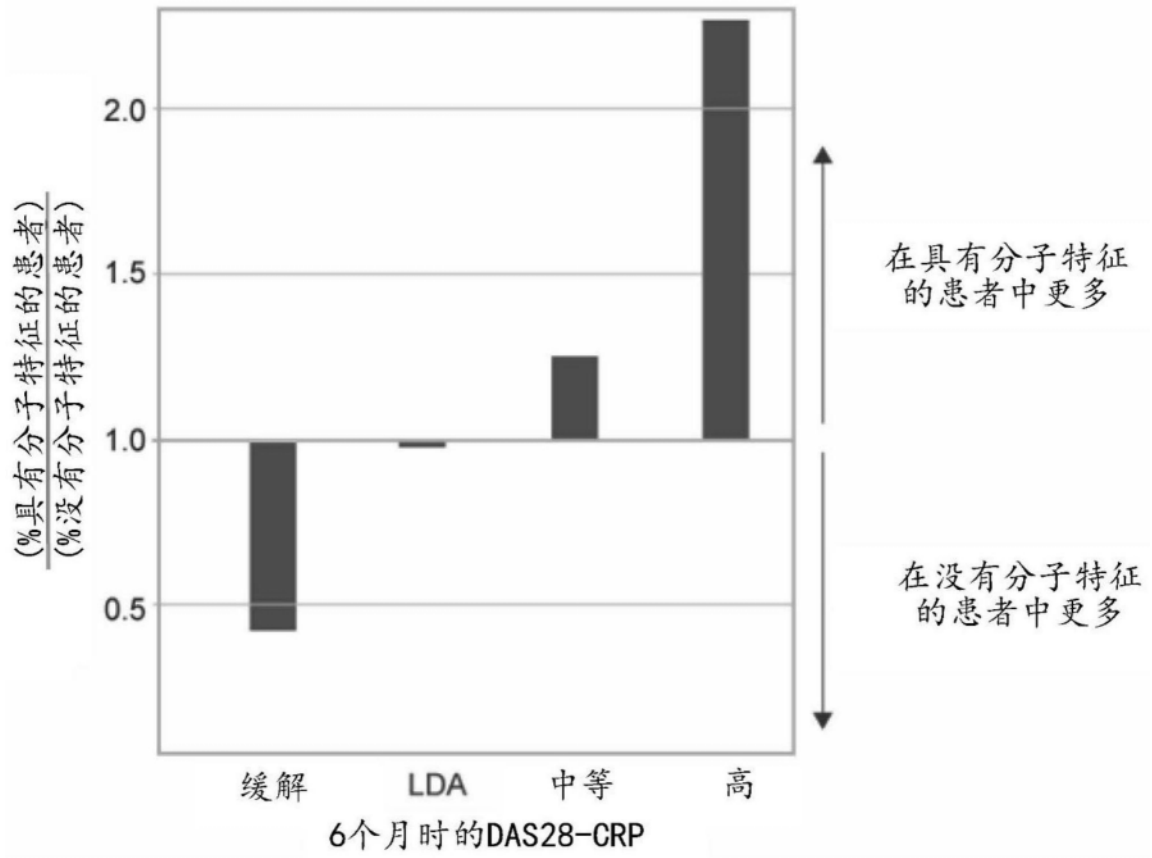


图2D

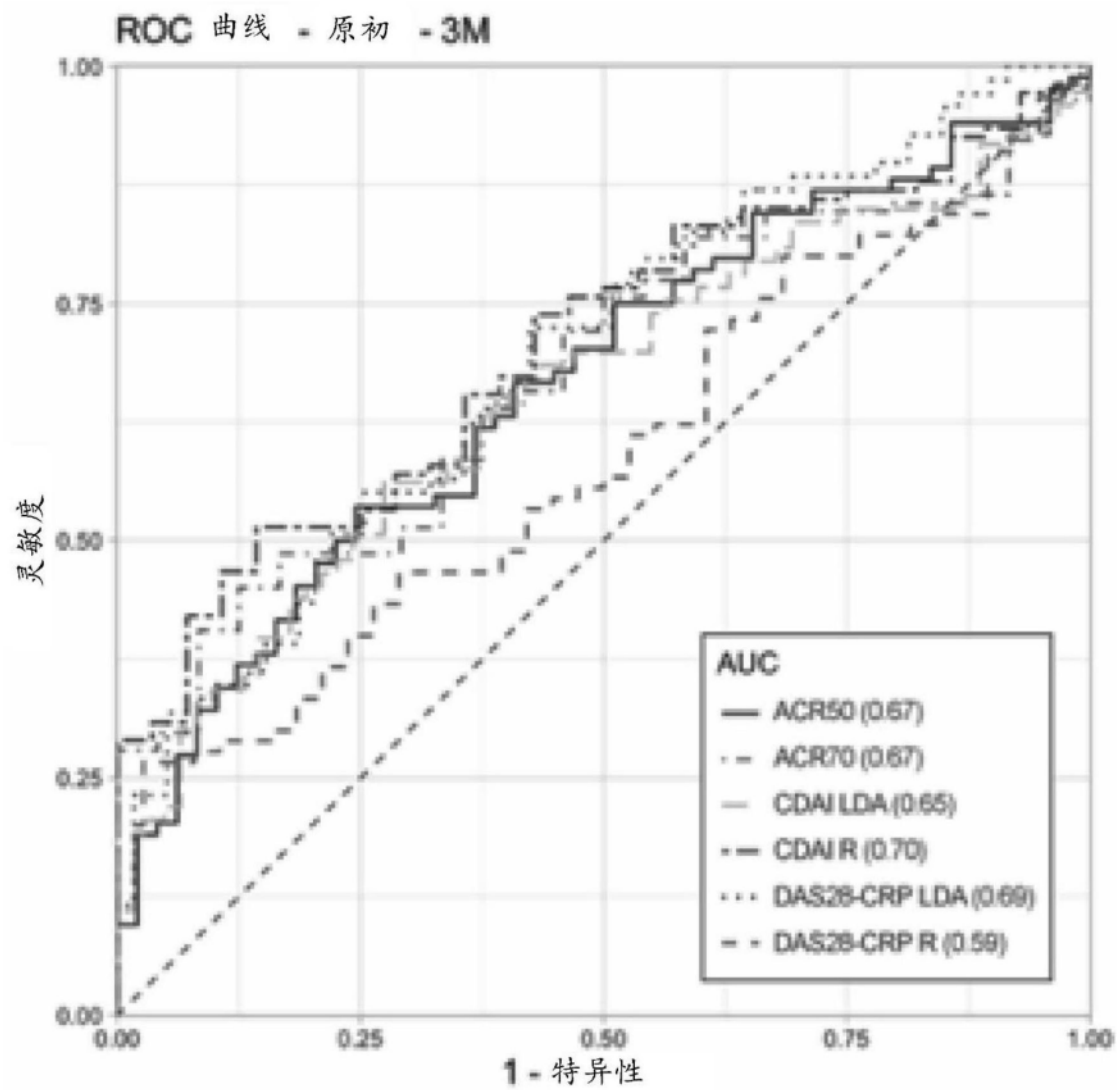


图3A

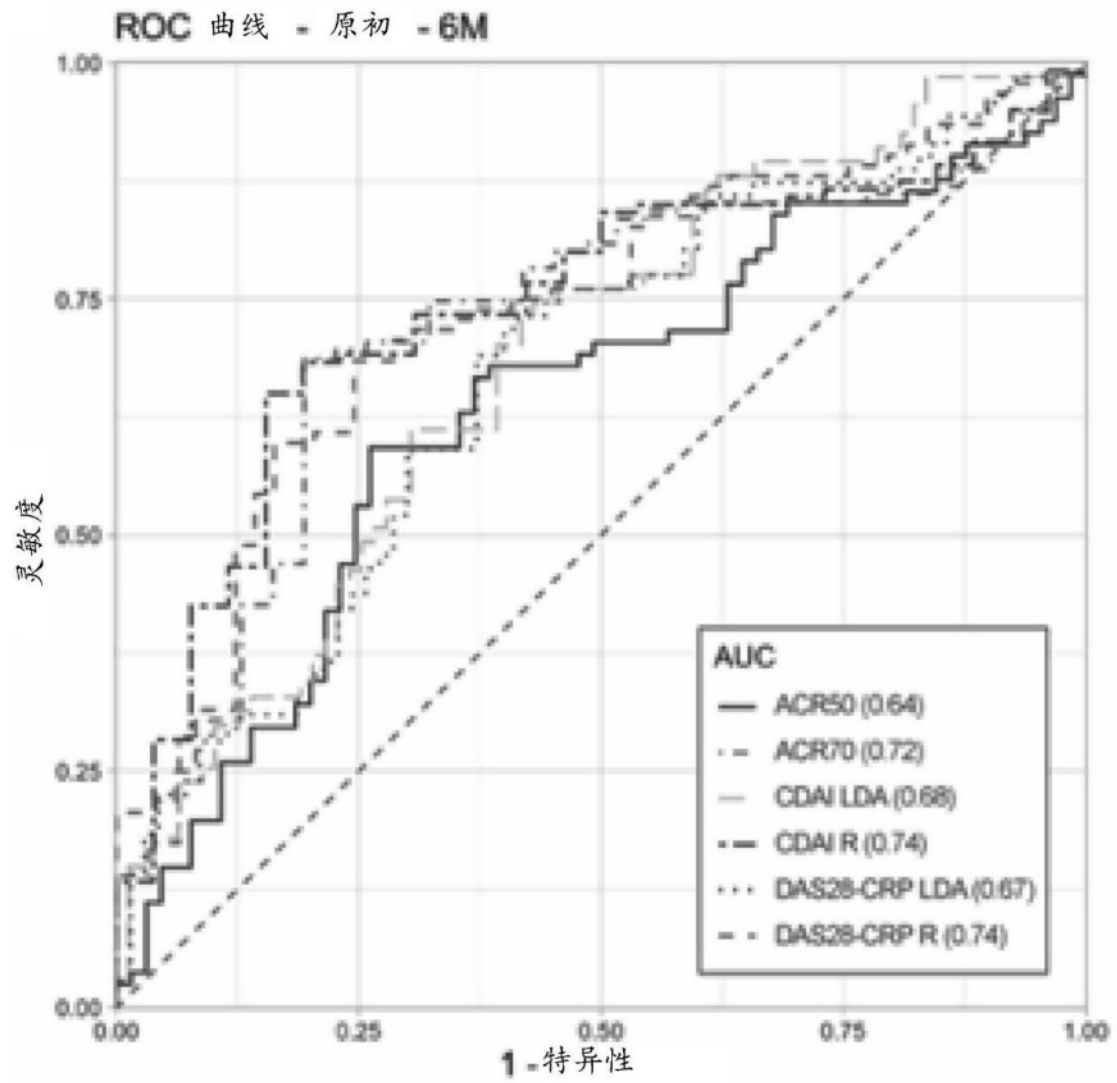


图3B

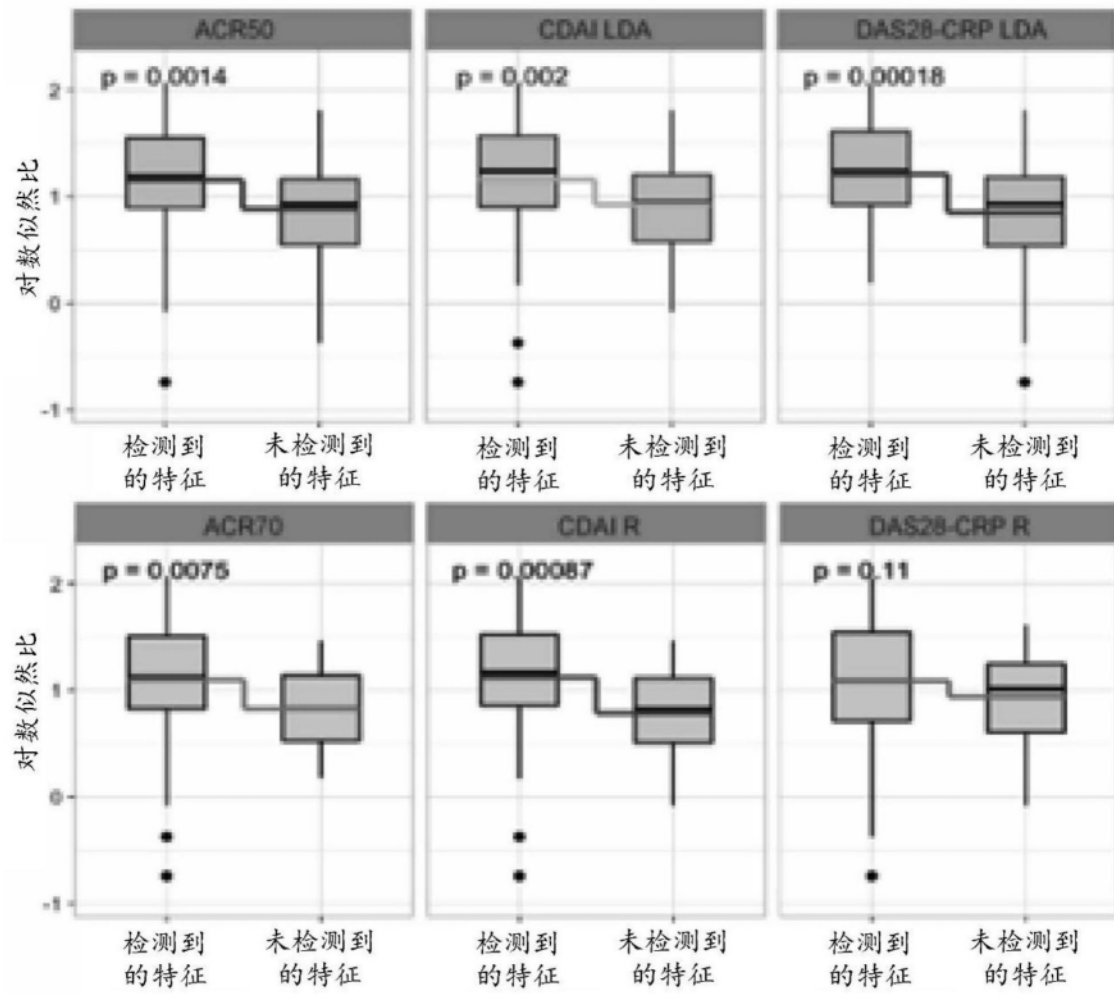


图3C

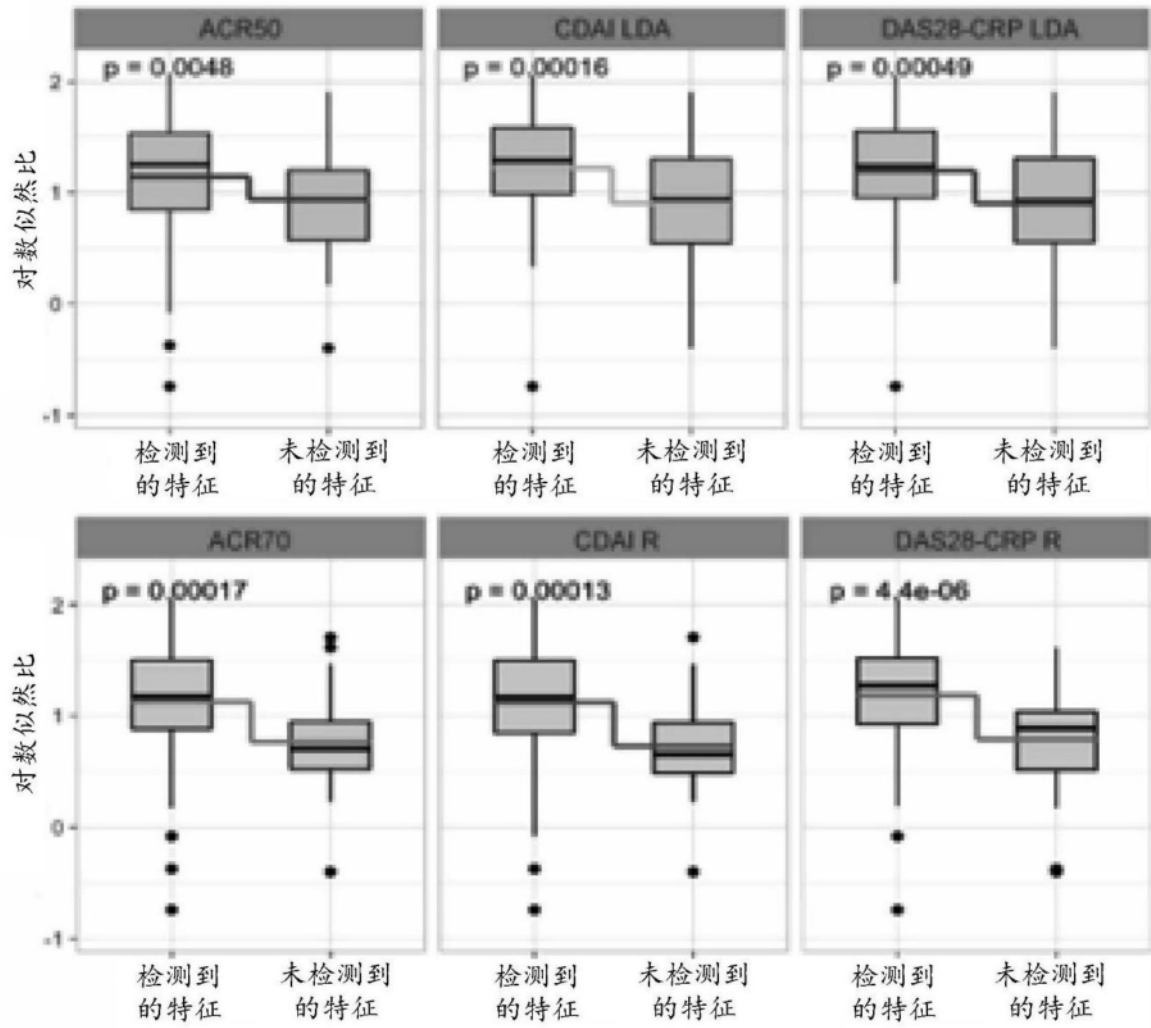


图3D

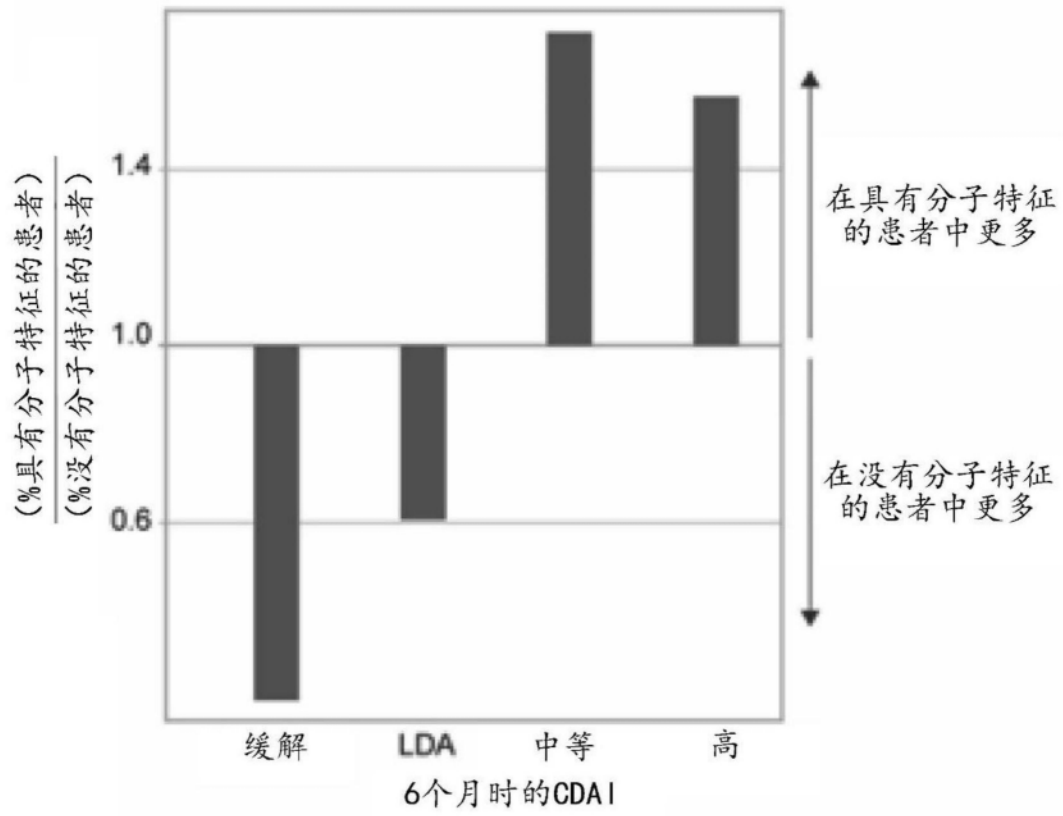


图3E

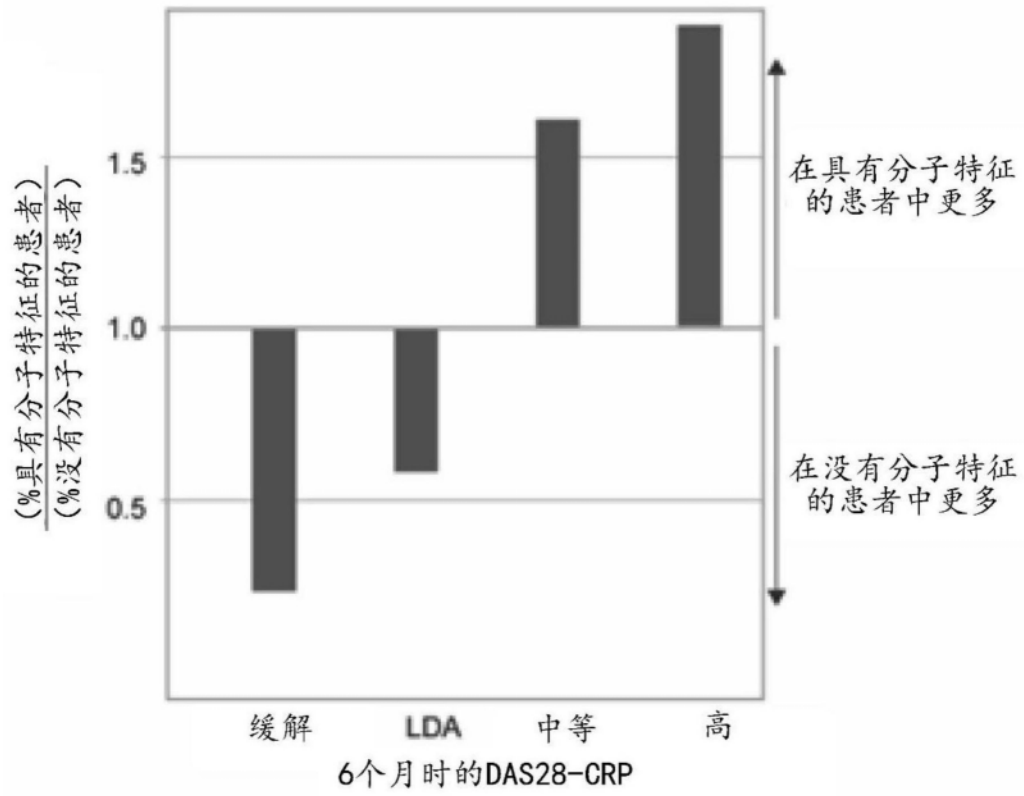


图3F

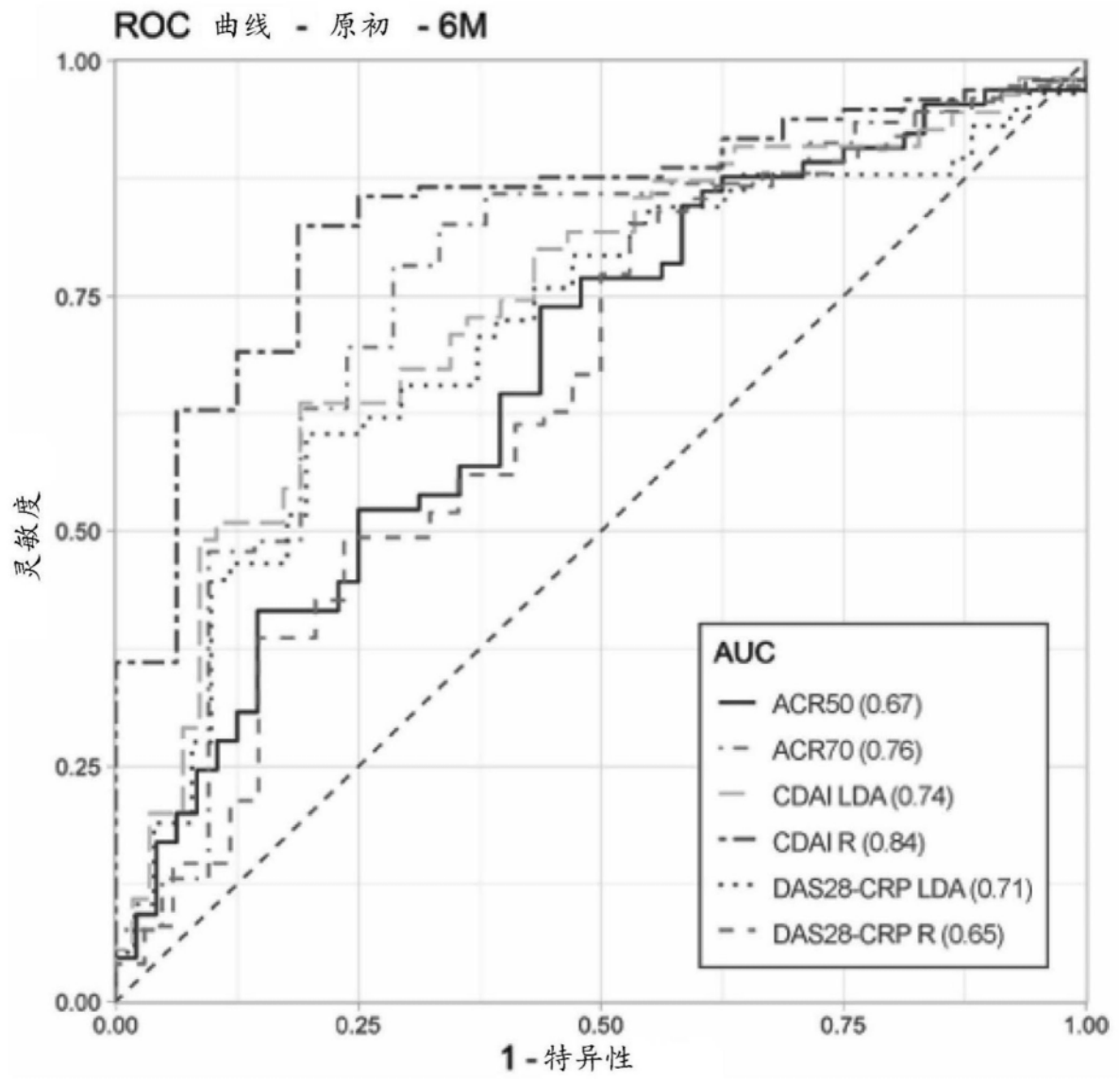


图4A

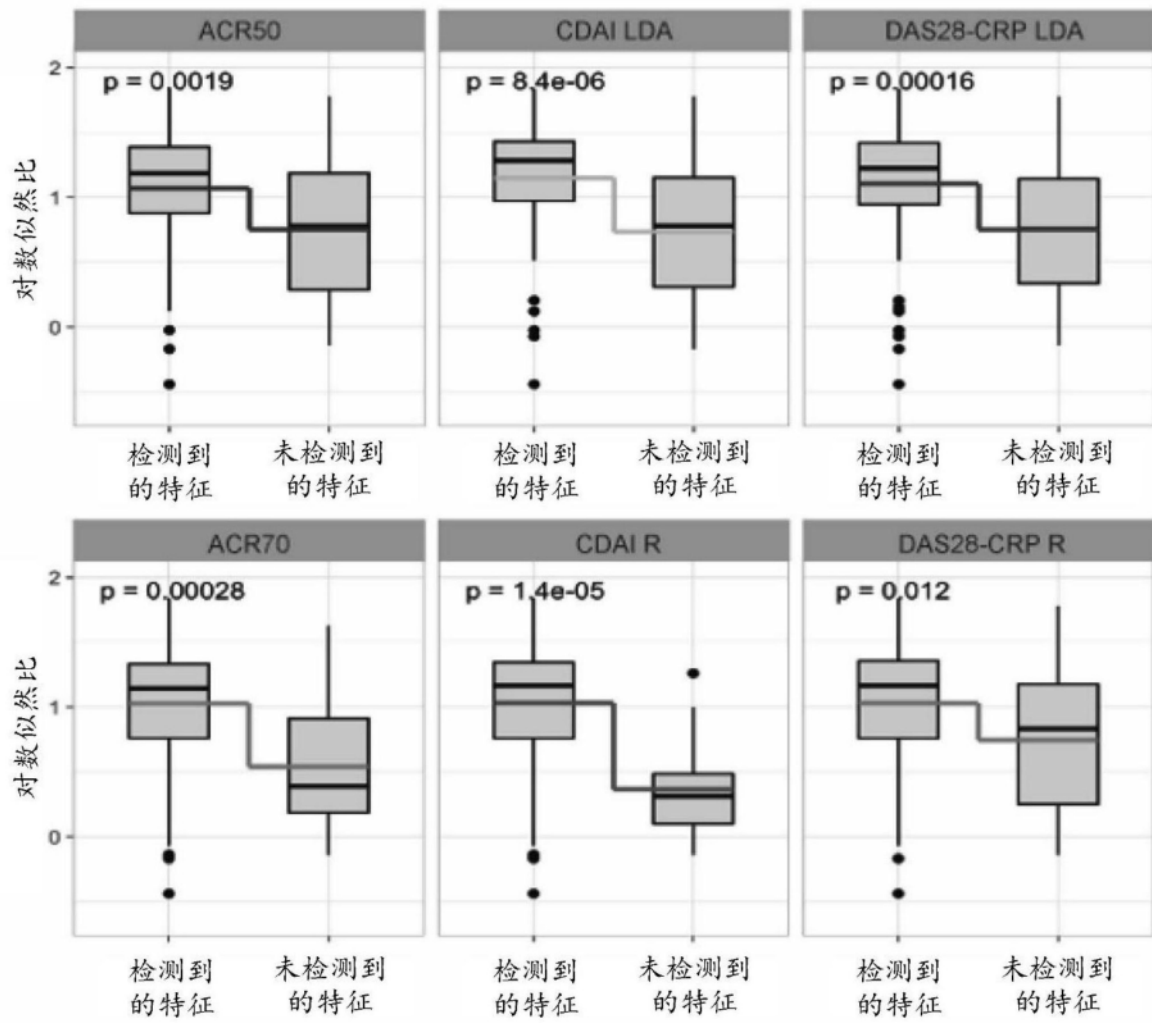


图4B

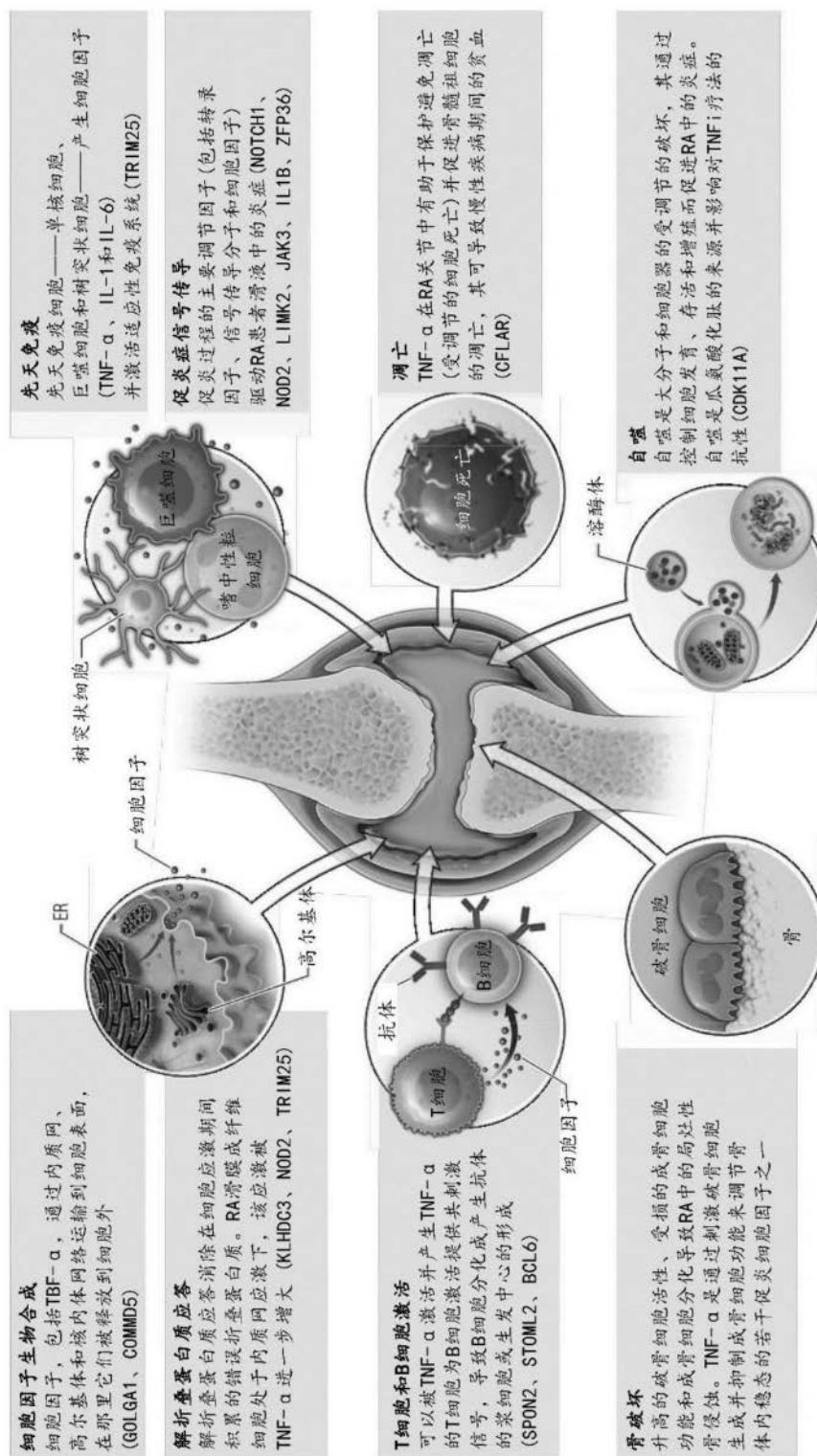
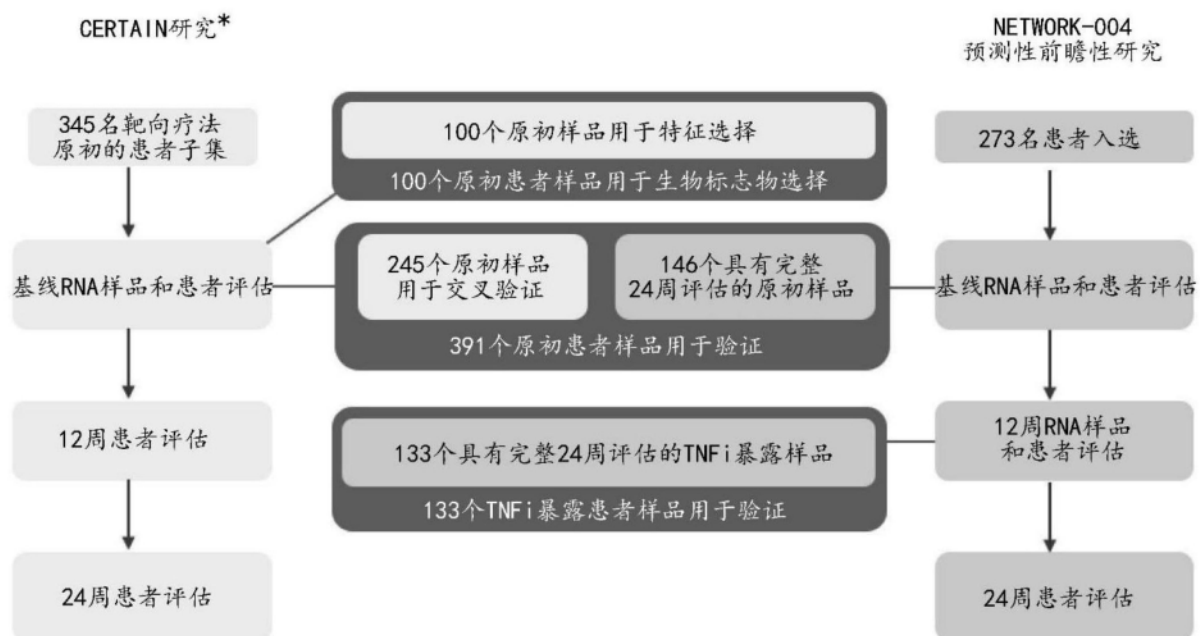


图5



*指出的评估和随访仅突出了这些分析中包括的Corrona CERTAIN研究的子集。对于CERTAIN研究的全部细节，参见 Pappas, D.A., Kremer, J.M., Reed, G. et al. Design characteristics of the CORRONA CERTAIN study: a comparative effectiveness study of biologic agents for rheumatoid arthritis patients. BMC Musculoskelet Disord 15, 113 (2014).

图6

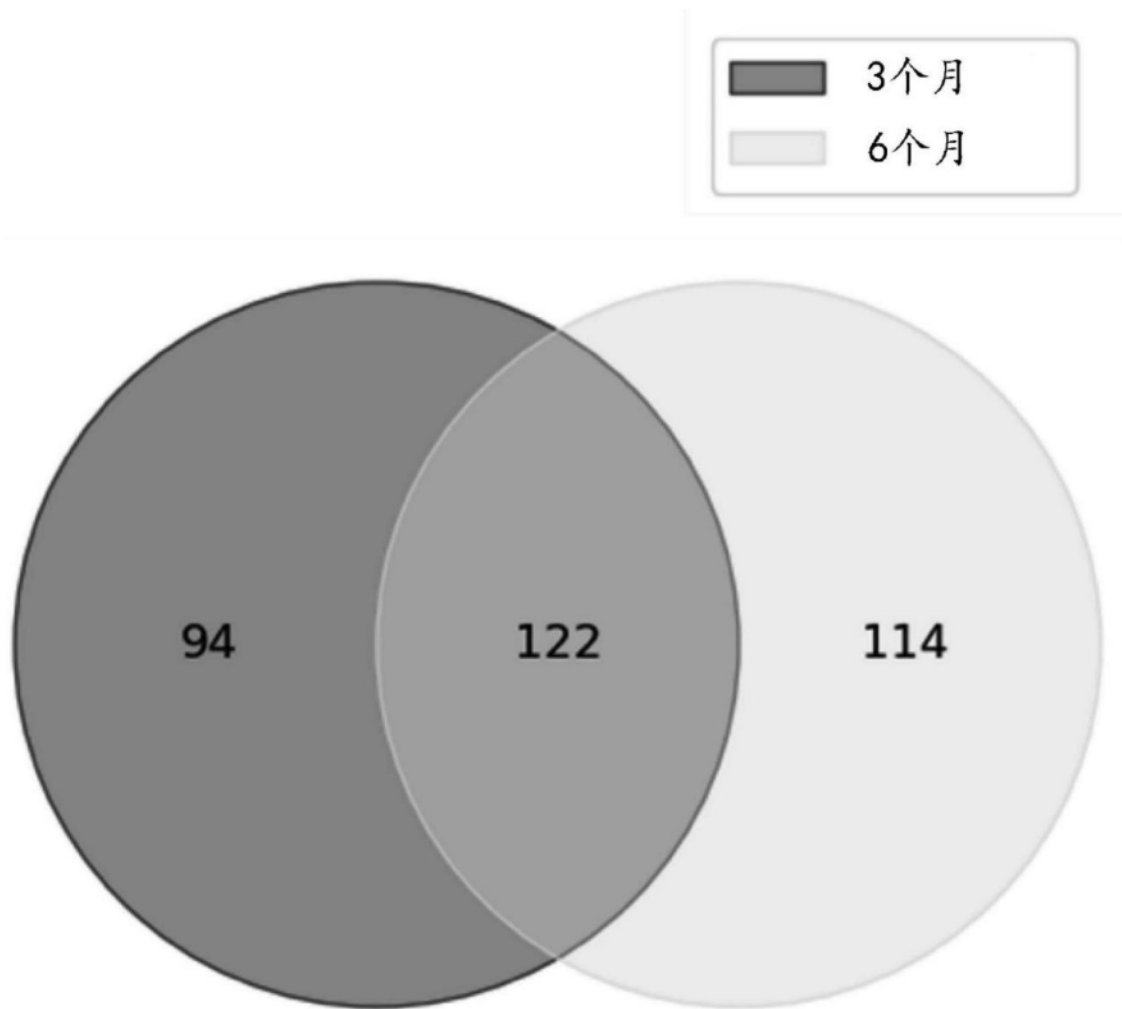


图7

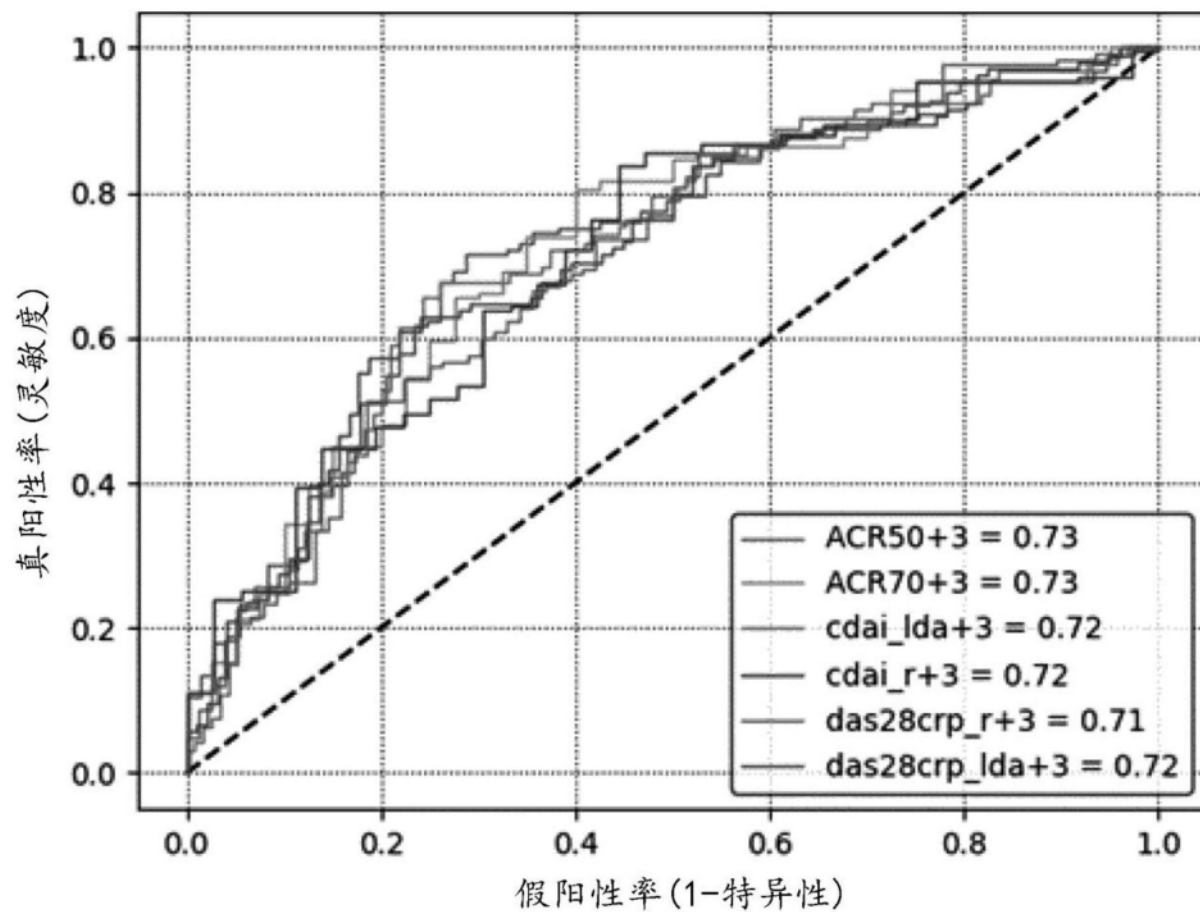


图8A

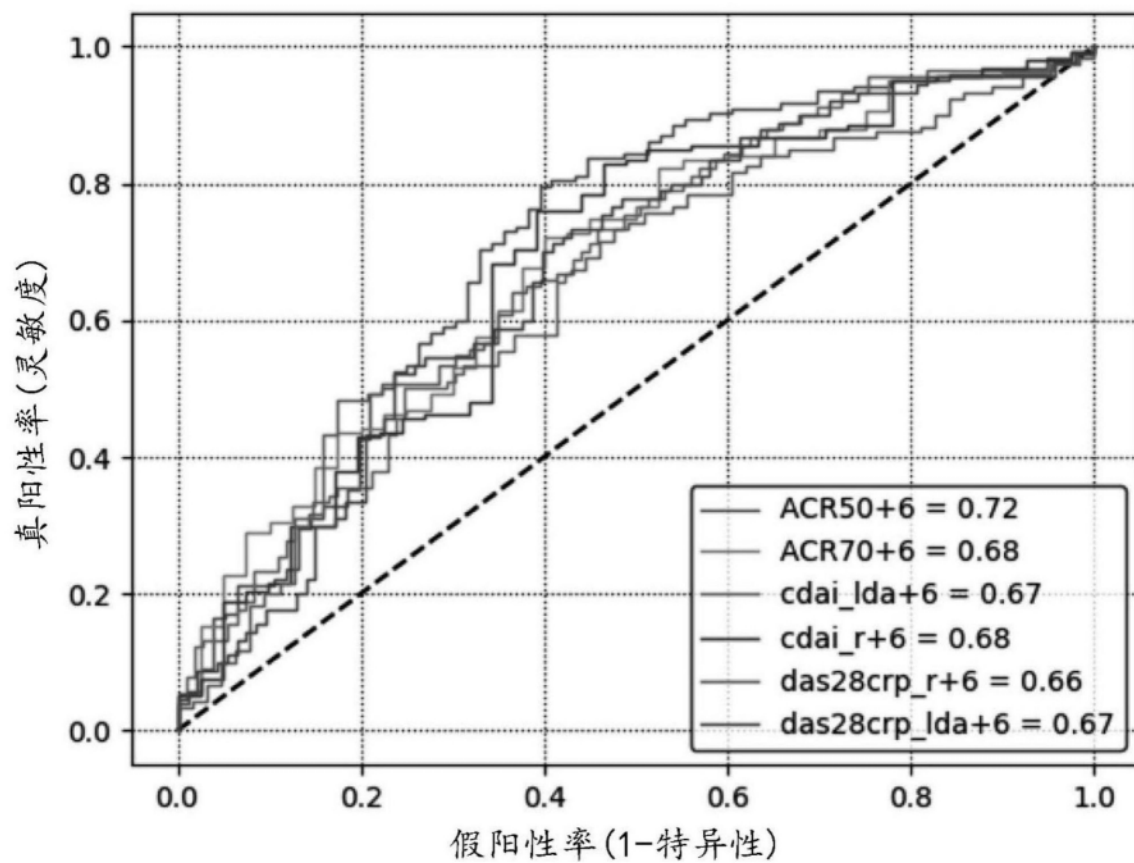


图8B

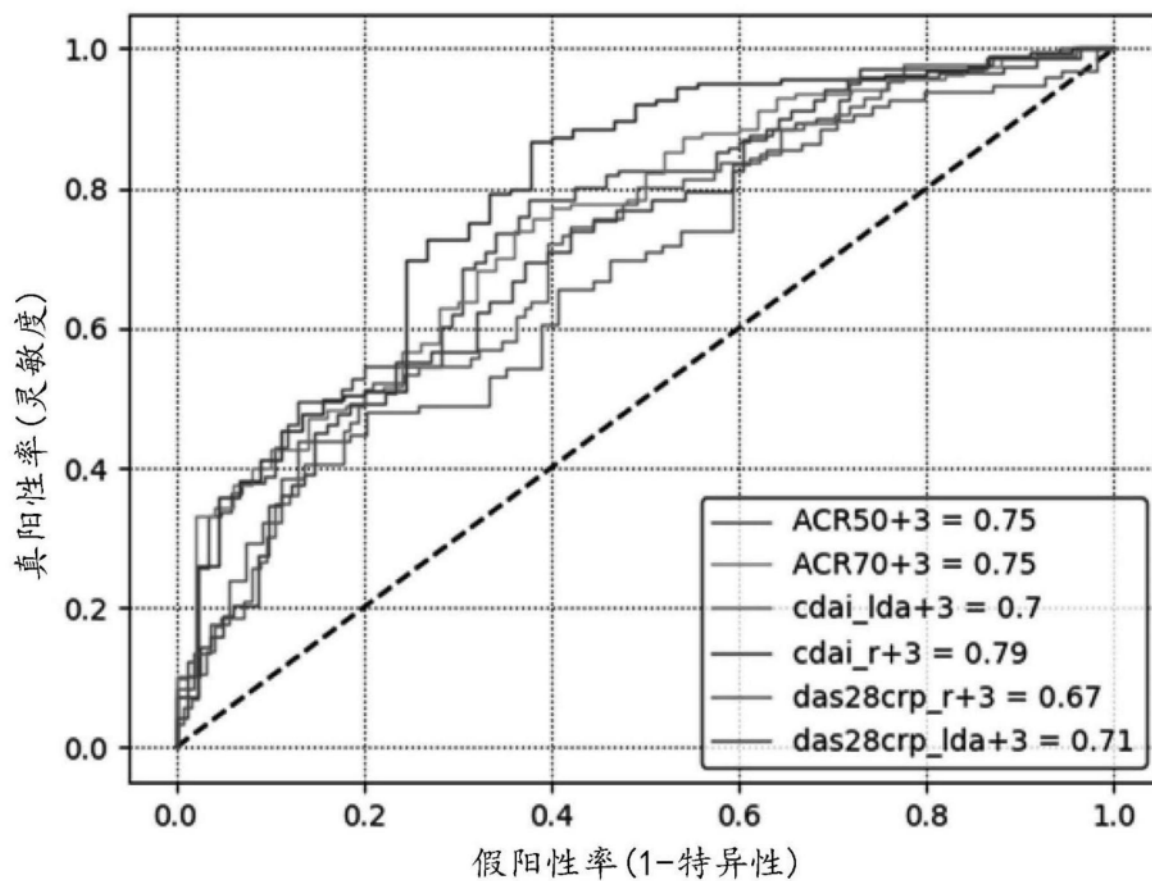


图8C

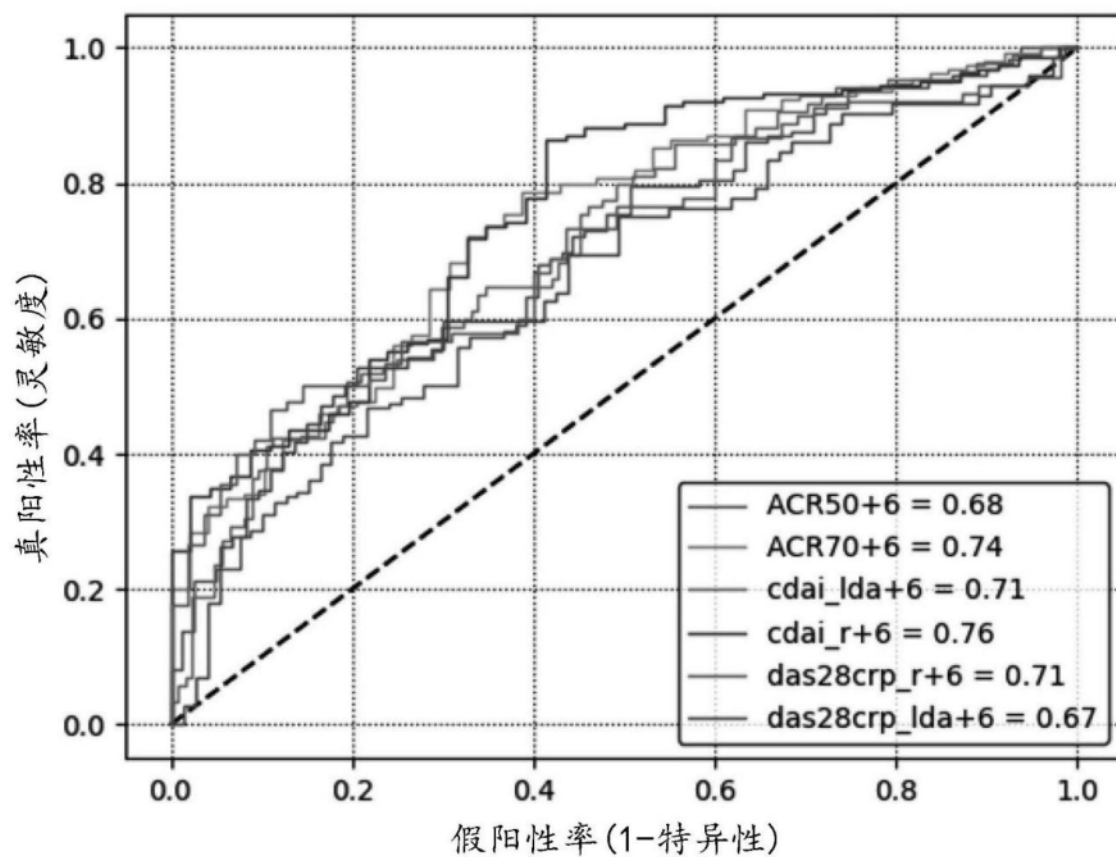


图8D

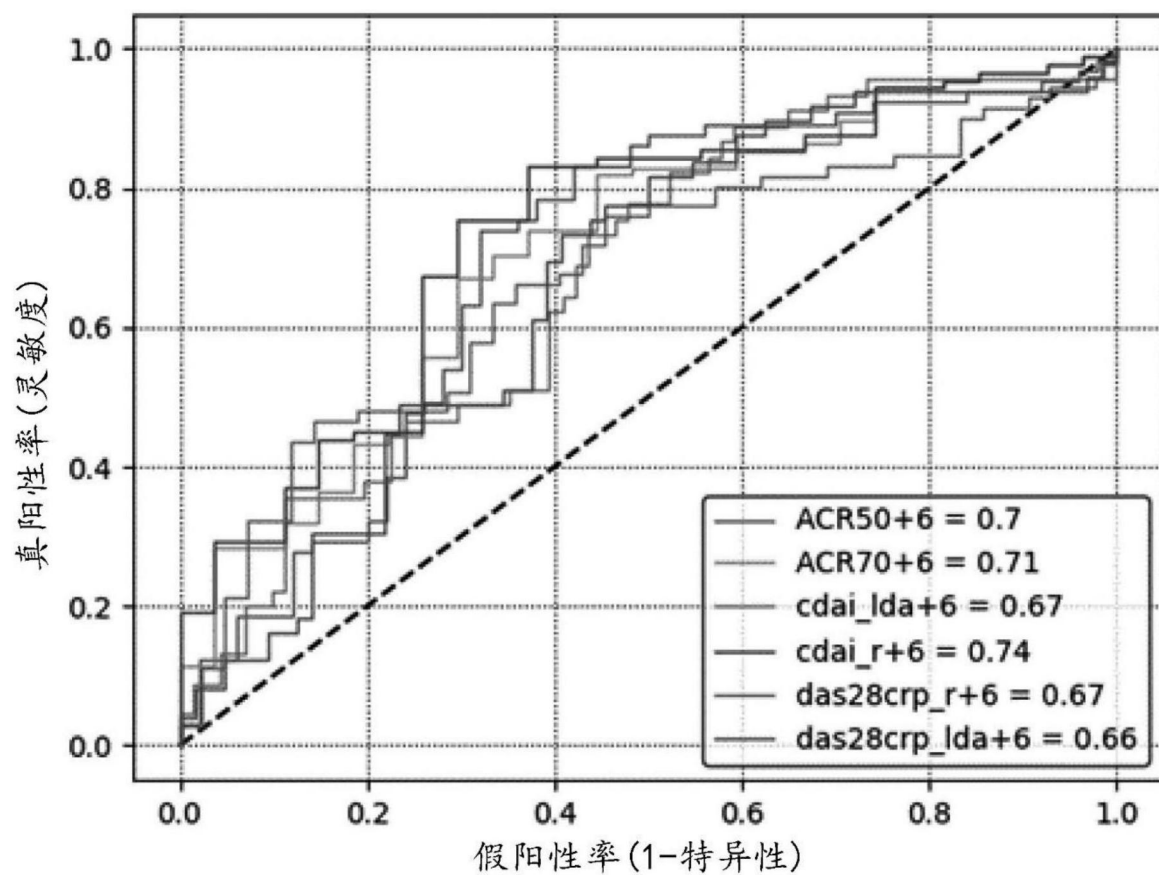


图9A

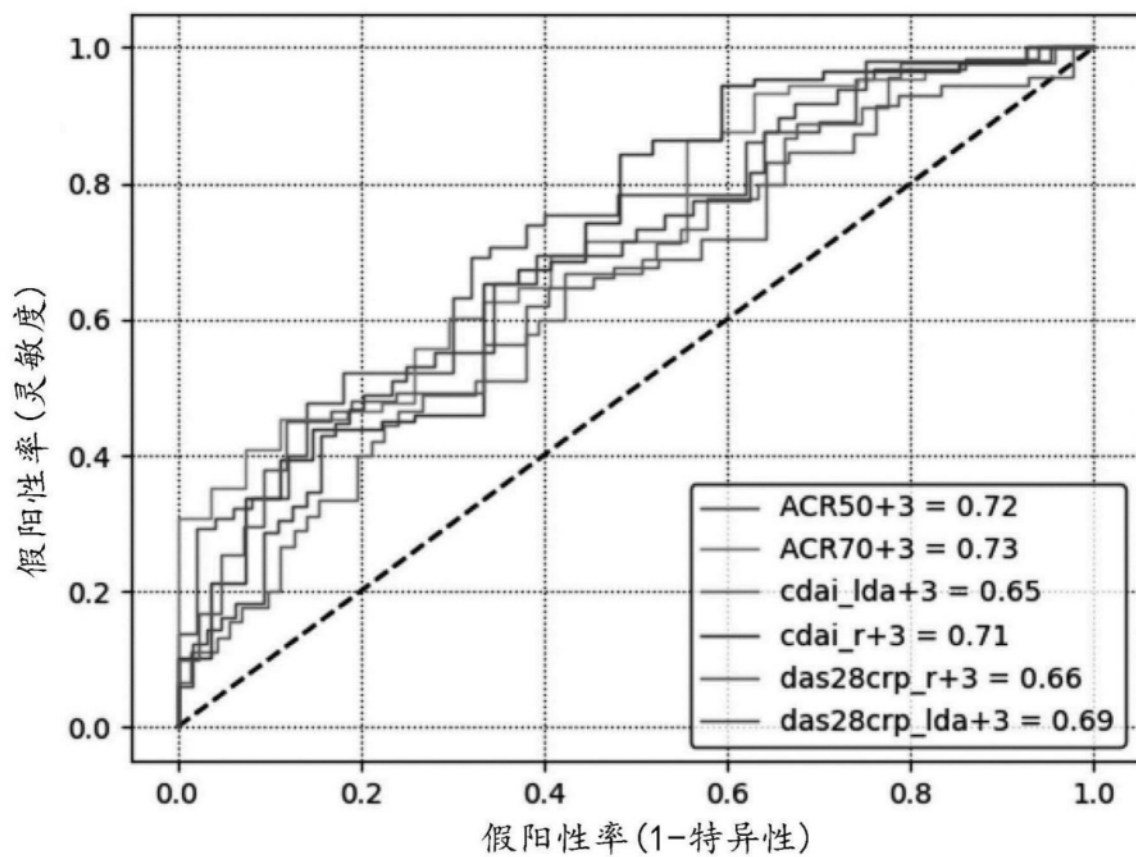


图9B

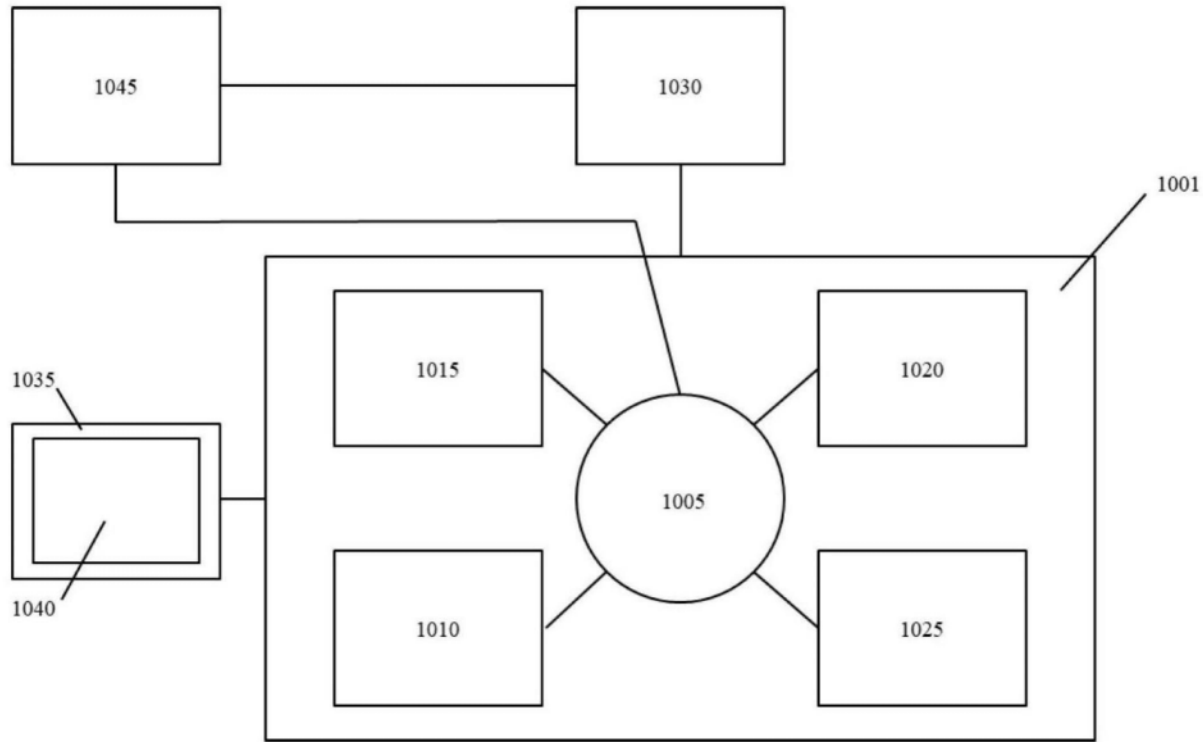


图10