



# (12) 发明专利申请

(10) 申请公布号 CN 114807370 A

(43) 申请公布日 2022. 07. 29

(21) 申请号 202210476516.8

(22) 申请日 2022.04.29

(71) 申请人 中国医学科学院肿瘤医院  
地址 100021 北京市朝阳区潘家园南里17号

(72) 发明人 马飞 钱海利 莫红楠 钱天一 韩佳澍

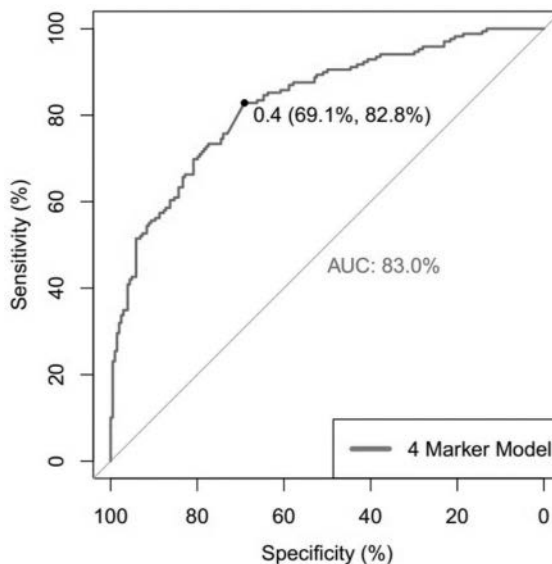
(74) 专利代理机构 北京预立生科知识产权代理有限公司 11736  
专利代理师 朱萍

(51) Int. Cl.  
C12Q 1/6886 (2018.01)  
G01N 33/574 (2006.01)  
G16H 50/30 (2018.01)  
G16B 20/00 (2019.01)

权利要求书2页 说明书12页 附图15页

(54) 发明名称  
一种新型的用于乳腺癌免疫治疗疗效精准预测的模型及其应用

(57) 摘要  
本发明公开了一种新型的用于乳腺癌免疫治疗疗效精准预测的模型及其应用,所述模型中包含IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin,根据该模型可对乳腺癌患者免疫治疗的预后和治疗效果进行准确预测,高评分提示所述乳腺癌患者采用免疫治疗策略可获益,本发明提供的模型对于预测乳腺癌免疫治疗患者的长期疗效具有指导意义,在临床上具有广泛的适用性。



1. 一组用于乳腺癌免疫治疗疗效预测的标志物组合,其特征在于,所述标志物组合包括IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin。

2. 检测样本中权利要求1所述的标志物组合表达水平的试剂在制备用于乳腺癌免疫治疗疗效预测的产品中的应用。

3. 根据权利要求2所述的应用,其特征在于,所述试剂包括通过组织质谱成像技术、数字成像技术、蛋白免疫技术、染料技术、核酸测序技术、核酸杂交技术、色谱技术、和/或二代测序技术检测样本中所述标志物组合表达水平的试剂;

优选地,所述试剂包括检测所述标志物组合mRNA表达水平的试剂;

更优选地,所述试剂为检测与所述标志物组合转录的mRNA互补的cDNA的水平的试剂;

最优选地,所述试剂为引物或探针;

优选地,所述试剂包括检测所述标志物组合蛋白表达水平的试剂;

更优选地,所述试剂为检测所述标志物组合编码的多肽或蛋白的水平的试剂;

最优选地,所述试剂为抗体、抗体片段或亲和性蛋白。

4. 一种用于乳腺癌免疫治疗疗效预测的风险预测模型,其特征在于,所述风险预测模型中包括权利要求1所述的标志物组合。

5. 根据权利要求4所述的风险预测模型,其特征在于,所述风险预测模型采用如下回归方程计算风险预测评分4RF-panel<sub>score</sub>:

风险预测评分4RF-panel<sub>score</sub> = IFN- $\gamma$  \* (-62.5851559) +  $\alpha$ SMA \* (-0.3814794) + collagen-I \* (28.1856122) + Vimentin \* (8.9254198) - 0.4248442;

其中,IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin为较DNA抗体通道的相对强度值;

优选地,所述风险预测评分4RF-panel<sub>score</sub> > 0.400时,受试者经免疫治疗后预后良好;

优选地,所述风险预测评分4RF-panel<sub>score</sub> < 0.400时,受试者经免疫治疗后预后不良。

6. 一种乳腺癌免疫治疗疗效预测的评估方法,其特征在于,所述方法包括如下步骤:

(1) 获取受试者的临床数据,所述临床数据包括IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin较DNA抗体通道的相对强度值;

(2) 向权利要求4或5所述的风险预测模型中输入步骤(1)中所述的临床数据,得到受试者的风险预测评分4RF-panel<sub>score</sub>;

(3) 根据步骤(2)得到的风险预测评分4RF-panel<sub>score</sub>对受试者的免疫治疗疗效进行预测评估:

优选地,所述风险预测评分4RF-panel<sub>score</sub> > 0.400时,受试者经免疫治疗后预后良好;

优选地,所述风险预测评分4RF-panel<sub>score</sub> < 0.400时,受试者经免疫治疗后预后不良。

7. 一种用于乳腺癌免疫治疗疗效预测的装置或系统,其特征在于,所述装置或系统包括:

(1) 数据获取模块,用于获取待测乳腺癌患者样本中的标志物组合较DNA抗体通道的相对强度值,所述标志物组合为IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin;

(2) 预测模块,用于将所述数据获取模块得到的标志物组合较DNA抗体通道的相对强度值数据作为输入数据提供给风险预测模型,所述风险预测模型基于乳腺癌患者样本中标志物组合较DNA抗体通道的相对强度值而对所述乳腺癌患者的免疫治疗疗效进行预后预测;

(3) 预测结果获取模块,用于获取所述预测模块中风险预测模型的输出结果,得到待测

乳腺癌患者的预后预测结果；

优选地,步骤(2)中所述的风险预测模型为权利要求4或5所述的风险预测模型；

优选地,所述风险预测评分 $4RF\text{-panel}_{score} > 0.400$ 时,待测乳腺癌患者经免疫治疗后预后良好；

优选地,所述风险预测评分 $4RF\text{-panel}_{score} < 0.400$ 时,待测乳腺癌患者经免疫治疗后预后不良。

8.一种计算机设备,其特征在于,所述计算机设备包括存储器和处理器,所述存储器存储有程序,所述处理器执行所述程序时实现如下方法:

获取待测乳腺癌患者样本中的标志物组合较DNA抗体通道的相对强度值数据,所述标志物组合为IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin；

将所述标志物组合较DNA抗体通道的相对强度值数据作为输入数据提供给风险预测模型；

输出待测乳腺癌患者的预后预测结果；

优选地,所述风险预测模型为权利要求4或5所述的风险预测模型；

优选地,所述风险预测评分 $4RF\text{-panel}_{score} > 0.400$ 时,待测乳腺癌患者经免疫治疗后预后良好；

优选地,所述风险预测评分 $4RF\text{-panel}_{score} < 0.400$ 时,待测乳腺癌患者经免疫治疗后预后不良。

9.一种计算机可读存储介质,其特征在于,所述计算机可读存储介质上存储有程序,所述程序被执行时实现如下方法:

获取待测乳腺癌患者样本中的标志物组合较DNA抗体通道的相对强度值数据,所述标志物组合为IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin；

将所述标志物组合较DNA抗体通道的相对强度值数据作为输入数据提供给风险预测模型；

输出待测乳腺癌患者的预后预测结果；

优选地,所述风险预测模型为权利要求4或5所述的风险预测模型；

优选地,所述风险预测评分 $4RF\text{-panel}_{score} > 0.400$ 时,待测乳腺癌患者经免疫治疗后预后良好；

优选地,所述风险预测评分 $4RF\text{-panel}_{score} < 0.400$ 时,待测乳腺癌患者经免疫治疗后预后不良。

10.如下任一方面的应用,其特征在于,所述应用包括:

(1) 权利要求1所述的标志物组合在制备用于乳腺癌免疫治疗疗效预测的产品中的应用；

(2) 权利要求1所述的标志物组合在构建用于乳腺癌免疫治疗疗效预测的风险预测模型中的应用；

(3) 权利要求4或5所述的风险预测模型在构建用于乳腺癌免疫治疗疗效预测的装置或系统中的应用；

(4) 权利要求4或5所述的风险预测模型在构建计算机设备或计算机可读存储介质中的应用。

## 一种新型的用于乳腺癌免疫治疗疗效精准预测的模型及其应用

### 技术领域

[0001] 本发明属于生物医学技术领域,涉及一种疾病免疫治疗疗效精准预测的模型,具体而言,涉及一种新型的用于乳腺癌免疫治疗疗效精准预测的模型及其应用。

### 背景技术

[0002] 乳腺癌(Breast cancer)是全球性疾病,发病率居全球女性恶性肿瘤之首,全球癌症调查相关数据显示,全球新发乳腺癌占有所有新发恶性肿瘤的24.2%,乳腺癌的病死率占恶性肿瘤病死率的15%,并呈现出持续增长的趋势(Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. CA: a cancer journal for clinicians, 2018, 68 (6) : 394-424.)。尽管在过去的30年中,随着新的抗癌药物的不断发现,以及诊断和治疗方法的不断发展,使得乳腺癌的死亡率降低了39%,但是由于乳腺癌基因具有高度变异性和特殊的生物学微环境,以及患者对传统治疗方法(包括手术、放疗、化疗、内分泌治疗等)的副作用难以耐受等原因,乳腺癌仍然是导致女性癌症死亡的第二大原因。因此,目前本领域迫切需要可提高乳腺癌患者生存率的有效治疗方法。

[0003] 随着免疫学和生物学的不断发展,免疫疗法长期以来已成为继传统化学疗法、放射疗法和外科手术之后的另一种重要的乳腺癌治疗方法。在肿瘤微环境中,抗原刺激T淋巴细胞活化过程中,由于多种免疫检查点的存在,导致T细胞无法有效激活,免疫检查点在肿瘤的发生和发展过程中起着关键作用。免疫检查点抑制通路的确定促使免疫疗法不断发生重大的改进。目前多数肿瘤治疗主要是针对淋巴细胞激活基因-3 (LAG-3)、细胞毒T淋巴细胞相关抗原-4 (CTLA-4) 和程序性死亡受体-1 (PD-1/PD-L1) 等相关免疫检查点分子为靶点。虽然免疫检查点抑制剂 (Immune checkpoint inhibitor, ICI) (例如,抗PD-1/PD-L1抗体) 在乳腺癌的治疗中取得了令人瞩目的成就,但是仅限于少数肿瘤患者获益,相当比例的肿瘤患者对抗PD1/PD-L1抗体无反应。目前本领域面临的主要挑战是识别和开发可预测ICI疗效的生物标志物,以精准发现免疫治疗获益的肿瘤患者。

[0004] 通过免疫组化评估肿瘤PD-L1表达水平是目前最为常用的筛选评估手段,多项临床研究表明,PD-L1表达量升高的肿瘤患者采用ICI治疗后获益更多,因此一些研究将PD-L1表达视为主要评估目标,并且基于PD-L1表达水平在部分肿瘤中就ICI疗效预测具有良好的区分作用。但是对于PD-L1免疫组化检测,目前有研究发现,部分患者的PD-L1表达和ICI治疗响应或总生存期 (Overall survival, OS) 之间并无关联性,甚至许多PD-L1表达低至无法检测到的患者从ICI治疗中获得持久的临床益处。此外,将PD-L1表达作为免疫疗效预测因素还有以下缺陷:PD-L1的表达在肿瘤的不同发展阶段、不同区域会存在异质性,其检测结果可能受取样时间和取样位点的影响,且受前期或同期放化疗的影响,PD-L1检测和评价方法缺乏公认标准,以致不同研究中获得不一致结果。因此,目前基于PD-L1免疫组化的筛选手段有待进一步的改进。

## 发明内容

[0005] 鉴于此,本发明的目的在于提供一种新型的用于乳腺癌免疫治疗疗效精准预测的模型及其应用,以克服目前本领域存在的上述技术问题。

[0006] 本发明的上述目的通过以下技术方案得以实现:

[0007] 本发明的第一方面提供了一组用于乳腺癌免疫治疗疗效预测的标志物组合。

[0008] 进一步,所述标志物组合包括IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin。

[0009] 在本发明的具体实施例中,本发明通过收集实际临床样本进行分析验证发现,IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin组合能够准确用于乳腺癌患者乳腺癌免疫治疗疗效的预测中,且显著优于其他标志物组合,具有良好的诊断预测效能、敏感性和特异性,可用于临床上精准识别免疫治疗获益的乳腺癌患者,指导免疫治疗的临床应用。

[0010] 本发明的第二方面提供了检测样本中本发明第一方面所述的标志物组合表达水平的试剂在制备用于乳腺癌免疫治疗疗效预测的产品中的应用。

[0011] 进一步,所述试剂包括通过组织质谱成像技术、数字成像技术、蛋白免疫技术、染料技术、核酸测序技术、核酸杂交技术、色谱技术、和/或二代测序技术检测样本中所述标志物组合表达水平的试剂;

[0012] 优选地,所述试剂包括检测所述标志物组合mRNA表达水平的试剂;

[0013] 更优选地,所述试剂为检测与所述标志物组合转录的mRNA互补的cDNA的水平的试剂;

[0014] 最优选地,所述试剂为引物或探针;

[0015] 优选地,所述试剂包括检测所述标志物组合蛋白表达水平的试剂;

[0016] 更优选地,所述试剂为检测所述标志物组合编码的多肽或蛋白的水平的试剂;

[0017] 最优选地,所述试剂为抗体、抗体片段或亲和性蛋白。

[0018] 进一步,所述样本来源于受试者的组织样本、外周血样本、血清样本、血浆样本、尿液样本、唾液样本;

[0019] 优选地,所述样本来源于受试者的组织样本,在本发明的具体实施方案中,所述受试者优选为乳腺癌患者,所述组织样本优选为乳腺癌肿瘤组织样本。

[0020] 在本发明的具体实施方案中,所述免疫治疗优选为PD-L1单抗免疫治疗,所述乳腺癌优选为HER2阴性乳腺癌。

[0021] 此外,本发明还提供了一种用于乳腺癌免疫治疗疗效预测的产品。

[0022] 进一步,所述产品为体外诊断产品;

[0023] 优选地,所述体外诊断产品为体外诊断试剂盒。

[0024] 进一步,所述产品包含检测样本中本发明第一方面所述的标志物组合表达水平的试剂;

[0025] 优选地,所述试剂包括检测所述标志物组合mRNA表达水平的试剂、和/或检测所述标志物组合蛋白表达水平的试剂;

[0026] 更优选地,所述试剂包括检测与所述标志物组合转录的mRNA互补的cDNA的水平的试剂、检测所述标志物组合编码的多肽或蛋白的水平的试剂;

[0027] 最优选地,所述试剂为引物、探针、抗体、抗体片段和/或亲和性蛋白。

[0028] 进一步,所述产品还包含总RNA抽提试剂、逆转录试剂、和/或二代测序试剂。

[0029] 本发明中所述标志物组合表达水平的检测可以采用本领域已知的测定方法,包括但不限于检测所述标志物组合中的基因的RNA转录物的量或所述标志物组合中的基因编码的多肽的量的方法;

[0030] 优选地,所述基因的RNA转录物可通过本领域已知的方法转化为与其互补的cDNA,通过测定互补cDNA的量可以获得RNA转录物的量。基因的RNA转录物或与其互补的cDNA的量,可以针对样本中总RNA或总cDNA的量或者针对一组管家基因的RNA转录物或与其互补的cDNA的量来标准化;

[0031] 优选地,所述RNA转录物可以通过例如杂交、扩增或者测序的方法来检测和量化,包括但不限于将RNA转录物与探针或者引物杂交的方法,通过基于聚合酶链式反应(PCR)的各种定量PCR技术或测序技术检测RNA转录物或其对应cDNA产物的量的方法。所述定量PCR技术包括但不限于荧光定量PCR、实时PCR或半定量PCR技术。所述测序技术包括但不限于Sanger测序、二代测序、三代测序和单细胞测序等。优选地,所述测序技术为二代测序,更优选地为靶向RNA-seq的转录组测序技术。

[0032] 在本发明的具体实施方案中,所述免疫治疗优选为PD-L1单抗免疫治疗,所述乳腺癌优选为HER2阴性乳腺癌。

[0033] 本发明的第三方面提供了一种用于乳腺癌免疫治疗疗效预测的风险预测模型。

[0034] 进一步,所述风险预测模型中包括本发明第一方面所述的标志物组合。

[0035] 进一步,所述风险预测模型采用如下回归方程计算风险预测评分4RF-panel<sub>score</sub>:

[0036] 风险预测评分4RF-panel<sub>score</sub> = IFN- $\gamma$  \* (-62.5851559) + aSMA \* (-0.3814794) + collagen-I \* (28.1856122) + Vimentin \* (8.9254198) - 0.4248442;

[0037] 其中,IFN- $\gamma$ 、aSMA、collagen-I、Vimentin为较DNA抗体通道的相对强度值;

[0038] 优选地,所述风险预测评分4RF-panel<sub>score</sub> > 0.400时,受试者经免疫治疗后预后良好;

[0039] 优选地,所述风险预测评分4RF-panel<sub>score</sub> < 0.400时,受试者经免疫治疗后预后不良。

[0040] 在本发明的具体实施方案中,所述免疫治疗优选为PD-L1单抗免疫治疗,所述乳腺癌优选为HER2阴性乳腺癌。

[0041] 本发明通过对构建得到的风险预测模型在收集得到的训练集和验证集进行验证发现,其具有较高的诊断预测效能,在训练集中的AUC值高达0.843,在验证集中的AUC高达0.830,显著高于其它对比模型的AUC值(AUC=0.674-0.789),表明了本发明构建得到的四风险因素(4RF-panel)模型具有较好的预测乳腺癌患者免疫治疗疗效的能力。

[0042] 本发明的第四方面提供了一种乳腺癌免疫治疗疗效预测的评估方法。

[0043] 进一步,所述方法包括如下步骤:

[0044] (1) 获取受试者的临床数据,所述临床数据包括IFN- $\gamma$ 、aSMA、collagen-I、Vimentin较DNA抗体通道的相对强度值;

[0045] (2) 向本发明第三方面所述的风险预测模型中输入步骤(1)中所述的临床数据,得到受试者的风险预测评分4RF-panel<sub>score</sub>;

[0046] (3) 根据步骤(2)得到的风险预测评分4RF-panel<sub>score</sub>对受试者的免疫治疗疗效进行预测评估;

- [0047] 优选地,所述风险预测评分 $4RF\text{-panel}_{\text{score}} > 0.400$ 时,受试者经免疫治疗后预后良好;
- [0048] 优选地,所述风险预测评分 $4RF\text{-panel}_{\text{score}} < 0.400$ 时,受试者经免疫治疗后预后不良。
- [0049] 进一步,所述IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin较DNA抗体通道的相对强度值是通过乳腺癌患者的肿瘤组织进行28通道组织质谱成像检测得到的检测结果。
- [0050] 进一步,所述受试者优选为人类。
- [0051] 进一步,所述受试者优选为乳腺癌患者。
- [0052] 进一步,所述受试者的临床数据是通过28通道组织质谱成像检测受试者的样本得到的,所述样本优选为乳腺癌受试者的肿瘤组织样本。
- [0053] 在本发明的具体实施方案中,所述免疫治疗优选为PD-L1单抗免疫治疗,所述乳腺癌优选为HER2阴性乳腺癌。
- [0054] 本发明的第五方面提供了一种用于乳腺癌免疫治疗疗效预测的装置或系统。
- [0055] 进一步,所述装置或系统包括:
- [0056] (1) 数据获取模块,用于获取待测乳腺癌患者样本中的标志物组合较DNA抗体通道的相对强度值,所述标志物组合为IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin;
- [0057] (2) 预测模块,用于将所述数据获取模块得到的标志物组合较DNA抗体通道的相对强度值数据作为输入数据提供给风险预测模型,所述风险预测模型基于乳腺癌患者样本中标志物组合较DNA抗体通道的相对强度值而对所述乳腺癌患者的免疫治疗疗效进行预后预测;
- [0058] (3) 预测结果获取模块,用于获取所述预测模块中风险预测模型的输出结果,得到待测乳腺癌患者的预后预测结果;
- [0059] 优选地,步骤(2)中所述的风险预测模型为本发明第三方面所述的风险预测模型;
- [0060] 优选地,所述风险预测评分 $4RF\text{-panel}_{\text{score}} > 0.400$ 时,待测乳腺癌患者经免疫治疗后预后良好;
- [0061] 优选地,所述风险预测评分 $4RF\text{-panel}_{\text{score}} < 0.400$ 时,待测乳腺癌患者经免疫治疗后预后不良。
- [0062] 进一步,所述样本来源于受试者的组织样本、外周血样本、血清样本、血浆样本、尿液样本、唾液样本;
- [0063] 优选地,所述样本来源于受试者的乳腺癌患者的组织样本,在本发明的具体实施方案中,所述组织样本优选为乳腺癌肿瘤组织样本。
- [0064] 在本发明的具体实施方案中,所述免疫治疗优选为PD-L1单抗免疫治疗,所述乳腺癌优选为HER2阴性乳腺癌。
- [0065] 本发明的第六方面提供了一种计算机设备。
- [0066] 进一步,所述计算机设备包括存储器和处理器,所述存储器存储有程序,所述处理器执行所述程序时实现如下方法:
- [0067] 获取待测乳腺癌患者样本中的标志物组合较DNA抗体通道的相对强度值数据,所述标志物组合为IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin;
- [0068] 将所述标志物组合较DNA抗体通道的相对强度值数据作为输入数据提供给风险预

测模型；

[0069] 输出待测乳腺癌患者的预后预测结果；

[0070] 优选地，所述风险预测模型为本发明第三方面所述的风险预测模型；

[0071] 优选地，所述风险预测评分 $4RF\text{-panel}_{score} > 0.400$ 时，待测乳腺癌患者经免疫治疗后预后良好；

[0072] 优选地，所述风险预测评分 $4RF\text{-panel}_{score} < 0.400$ 时，待测乳腺癌患者经免疫治疗后预后不良。

[0073] 进一步，所述样本来源于受试者的组织样本、外周血样本、血清样本、血浆样本、尿液样本、唾液样本；

[0074] 优选地，所述样本来源于受试者的乳腺癌患者的组织样本，在本发明的具体实施方案中，所述组织样本优选为乳腺癌肿瘤组织样本。

[0075] 在本发明的具体实施方案中，所述免疫治疗优选为PD-L1单抗免疫治疗，所述乳腺癌优选为HER2阴性乳腺癌。

[0076] 本发明的第七方面提供了一种计算机可读存储介质。

[0077] 进一步，所述计算机可读存储介质上存储有程序，所述程序被执行时实现如下方法：

[0078] 获取待测乳腺癌患者样本中的标志物组合较DNA抗体通道的相对强度值数据，所述标志物组合为IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin；

[0079] 将所述标志物组合较DNA抗体通道的相对强度值数据作为输入数据提供给风险预测模型；

[0080] 输出待测乳腺癌患者的预后预测结果；

[0081] 优选地，所述风险预测模型为权利要求4或5所述的风险预测模型；

[0082] 优选地，所述风险预测评分 $4RF\text{-panel}_{score} > 0.400$ 时，待测乳腺癌患者经免疫治疗后预后良好；

[0083] 优选地，所述风险预测评分 $4RF\text{-panel}_{score} < 0.400$ 时，待测乳腺癌患者经免疫治疗后预后不良。

[0084] 进一步，所述样本来源于受试者的组织样本、外周血样本、血清样本、血浆样本、尿液样本、唾液样本；

[0085] 优选地，所述样本来源于受试者的乳腺癌患者的组织样本，在本发明的具体实施方案中，所述组织样本优选为乳腺癌肿瘤组织样本。

[0086] 在本发明的具体实施方案中，所述免疫治疗优选为PD-L1单抗免疫治疗，所述乳腺癌优选为HER2阴性乳腺癌。

[0087] 本发明的第八方面提供了如下任一方面的应用：

[0088] (1) 本发明第一方面所述的标志物组合在制备用于乳腺癌免疫治疗疗效预测的产品中的应用；

[0089] (2) 本发明第一方面所述的标志物组合在构建用于乳腺癌免疫治疗疗效预测的风险预测模型中的应用；

[0090] (3) 本发明第三方面所述的风险预测模型在构建用于乳腺癌免疫治疗疗效预测的装置或系统中的应用；



[0091] (4) 本发明第三方面所述的风险预测模型在构建计算机设备或计算机可读存储介质中的应用。

[0092] 相对于现有技术,本发明具有的优点和有益效果:

[0093] 本发明提供了一种新型的用于乳腺癌免疫治疗疗效精准预测的模型,所述模型可实现对乳腺癌患者免疫治疗疗效的精准预测,根据模型计算得到的结果可以非常明确地区分对乳腺癌免疫治疗有响应者和非响应者,能够有效帮助临床医生对乳腺癌患者在接受免疫治疗前PD-L1单抗免疫治疗的有效性进行精准的个体化评估,指导医生对乳腺癌患者进行精准的免疫治疗,从而为乳腺癌患者带来更好的生存获益,具有重要的临床应用价值。

#### 附图说明

[0094] 以下,结合附图来详细说明本发明的实施方案,其中:

[0095] 图1为IFN- $\gamma$ 在免疫治疗响应者和免疫治疗非响应者的肿瘤细胞之间差异表达的结果图;

[0096] 图2为HLA-DR在免疫治疗响应者和免疫治疗非响应者的肿瘤细胞之间差异表达的结果图;

[0097] 图3为PD1在免疫治疗响应者和免疫治疗非响应者的肿瘤细胞之间差异表达的结果图;

[0098] 图4为PDL1在免疫治疗响应者和免疫治疗非响应者的肿瘤细胞之间差异表达的结果图;

[0099] 图5为PDL2在免疫治疗响应者和免疫治疗非响应者的肿瘤细胞之间差异表达的结果图;

[0100] 图6为TGF- $\beta$ 在免疫治疗响应者和免疫治疗非响应者的肿瘤细胞之间差异表达的结果图;

[0101] 图7为CKAE1/AE3在免疫治疗响应者和免疫治疗非响应者的肿瘤细胞之间差异表达的结果图;

[0102] 图8为TNF- $\alpha$ 在免疫治疗响应者和免疫治疗非响应者的肿瘤细胞之间差异表达的结果图;

[0103] 图9为 $\alpha$ SMA在免疫治疗响应者和免疫治疗非响应者的肿瘤细胞之间差异表达的结果图;

[0104] 图10为collagen-I在免疫治疗响应者和免疫治疗非响应者的肿瘤细胞之间差异表达的结果图;

[0105] 图11为E-cadherin在免疫治疗响应者和免疫治疗非响应者的肿瘤细胞之间差异表达的结果图;

[0106] 图12为Vimentin在免疫治疗响应者和免疫治疗非响应者的肿瘤细胞之间差异表达的结果图;

[0107] 图13为Ki-67在免疫治疗响应者和免疫治疗非响应者的肿瘤细胞之间差异表达的结果图;

[0108] 图14为 $\beta$ -catenin在免疫治疗响应者和免疫治疗非响应者的肿瘤细胞之间差异表达的结果图;

[0109] 图15为FoxP3在免疫治疗响应者和免疫治疗非响应者的肿瘤细胞之间差异表达的结果图；

[0110] 图16为标志物CD68、CD45、TGF- $\beta$ 、Vimentin、CD8、CD3、CD11b、Ki-67、 $\beta$ -catenin、CD33、 $\alpha$ SMA、PDL2、HLA-DR、CD4、FoxP3、TNF- $\alpha$ 、collagen-I、CD15、E-cadherin、CD14、CD16、CD19、PDL1、CD11c、CKAE1/AE3、IFN- $\gamma$ 、PD1、CD56与乳腺癌免疫治疗基因相关性热图；

[0111] 图17为本发明构建的四风险因素(4RF-panel)模型(包含4种标志物:IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin)在训练集中对乳腺癌患者免疫治疗疗效进行预测的诊断效能结果图；

[0112] 图18为本发明构建的四风险因素(4RF-panel)模型(包含4种标志物:IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin)在验证集中对乳腺癌患者免疫治疗疗效进行预测的诊断效能结果图。

### 具体实施方式

[0113] 下面结合具体实施例,进一步阐述本发明,仅用于解释本发明,而不能理解为对本发明的限制。本领域的普通技术人员可以理解为:在不脱离本发明的原理和宗旨的情况下可以对这些实施例进行多种变化、修改、替换和变型,本发明的范围由权利要求及其等同物限定。下述实施例中所使用的实验方法如无特殊说明,均为常规方法;下述实施例中所用的试剂、生物材料等,如无特殊说明,均可从商业途径得到。

[0114] 为了更好地解释本发明,此处对本发明中涉及到的专业术语进行如下解释。

[0115] 在本发明的上下文中,所使用的术语“PD-L1单抗免疫治疗”,是指乳腺癌治疗的一种治疗方法,所述PD-L1单抗药物包括阿特珠单抗(Atezolizumab)、阿维单抗(Avelumab)、度伐单抗(Durvalunab)等等。PD-L1单抗免疫治疗的机理如下:在机体对肿瘤细胞的免疫过程中,浸润在肿瘤组织中的T淋巴细胞过表达负性调节分子PD-1,或肿瘤细胞过表达其配体PD-L1,作为抑制性第二信号分子,二者相互作用在一定程度上抑制了机体的细胞免疫反应,从而使肿瘤得以逃避免疫系统的监视和清除(肿瘤免疫逃避的部分机制)。PD-1抗体免疫疗法的作用机制就是针对PD-1或PD-L1设计特定的蛋白质抗体,这些抗体可以与PD-1或PD-L1结合,从而阻止T细胞表面PD-1和肿瘤细胞PD-L1的相互作用,进而部分恢复T细胞功能,从而增强T细胞杀死肿瘤细胞的作用。对HER2阴性乳腺癌患者进行PD-L1单抗免疫治疗可使患者获得生存获益,但是仍有部分患者对PD-L1单抗免疫治疗的反应不佳,因此,本发明提供了一种新型的用于乳腺癌免疫治疗疗效精准预测的模型,以实现乳腺癌患者免疫治疗疗效的精准预测。

[0116] 在本发明的上下文中,所使用的术语“包括”、“包含”、“具有”、“含有”或“涉及”,是指包含性的(Inclusive)或开放式的,且不排除其它未列举的元素或方法步骤。术语“由...组成”被认为是术语“包含”的优选实施方案。如果在文中某一组被定义为包含至少一定数目的实施方案,这也应被理解为揭示了一个优选地仅由这些实施方案组成的组。

[0117] 在本发明的上下文中,所使用的术语“样本”,是指获得自或衍生自受试者(例如感兴趣的个体)的组合物,其包含有待根据例如物理,生化,化学和/或生理特点来表征和/或鉴定的细胞和/或其它分子实体。例如,样本是指来源于感兴趣的受试者的任何样本,预计或已知其包含待表征的细胞和/或分子实体。样本包括但不限于,组织样本(例如肿瘤组织

样本),原代或培养的细胞或细胞系、细胞上清、细胞裂解物、血小板、血清、血浆、玻璃体液、淋巴液、滑液、滤泡液、精液、羊水、乳液、全血、血液衍生的细胞、尿液、脑脊髓液、唾液、痰液、泪液、汗液、粘液、肿瘤裂解物、组织培养液、组织提取物、匀浆化的组织、肿瘤组织、细胞提取物、及其组合。作为优选的实施方式,所述样本选自受试者的组织样本。

[0118] 在本发明的上下文中,所使用的术语“AUC(受试者工作特征曲线下面积)”,是诊断规程的性能或精确性的一项指标。诊断方法的精确性由它的受试者工作特征(ROC)描述得最好。ROC图是源自观察的整个数据范围上连续改变决策阈的所有灵敏度/特异性对的线图。实验室测试的临床性能取决于它的诊断精确性,或将受试者正确分类入临床有关亚组的能力。诊断精确性测量测试正确辨别所调查的受试者的两种不同状况的能力。

[0119] 在本发明的上下文中,所使用的术语“表达水平”,是指生物样本中多核苷酸或氨基酸产物或蛋白质的量。“表达”通常是指基因编码的信息转化为在细胞中存在和运转的结构的过程。因此,本文所用的标志物的“表达”指转录为多核苷酸、翻译为蛋白质、或甚至蛋白质的翻译后修饰。转录的多核苷酸、翻译的蛋白质或翻译后修饰的蛋白质的片段也视为表达的,无论它们是源自通过选择性剪接产生的转录物或降解的转录物,还是源自蛋白质的翻译后加工(例如,通过蛋白水解)。“表达的基因”包括转录为多核苷酸(如mRNA)然后翻译为蛋白质的那些,以及转录为RNA但不翻译为蛋白质的那些(例如,转运RNA和核糖体RNA)。

[0120] 在本发明的上下文中,所使用的术语“计算机设备”,包括但不限于任何一种可与用户通过键盘、触摸板或声控设备等方式进行人机交互的个人电脑、服务器等终端。本文中的计算设备还可以包括移动终端,其包括但不限于任何一种可与用户通过键盘、触摸板或声控设备等方式进行人机交互的电子设备,例如,平板电脑、智能手机、个人数字助理(Personal Digital Assistant,PDA)、智能式穿戴式设备等终端。计算设备所处的网络包括(但不限于)互联网、广域网、城域网、局域网、虚拟专用网络(Virtual Private Network,VPN)等。

[0121] 在本发明的上下文中,所使用的术语“存储器”,包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器(ROM)、可编程ROM(PROM)、电可编程ROM(EPROM)、电可擦除可编程ROM(EEPROM)或闪存。其上存储有操作系统的代码。例如,存储器上还存储有代码或指令,通过运行这些代码或指令,可以实现本文中公开的实施例提供的结直肠癌预后预测和诊断的风险评估模型。易失性存储器可包括随机存取存储器(RAM)或者外部高速缓冲存储器。

[0122] 在本发明的上下文中,所使用的术语“处理器”,可以包括一个或者多个微处理器、数字处理器。所述处理器可调用存储器中存储的程序代码以执行相关的功能。所述处理器又称中央处理器(CPU,Central Processing Unit),可以是一块超大规模的集成电路,是运算核心(Core)和控制核心(Control Unit)。

[0123] 实施例1与乳腺癌患者对免疫治疗(PD-L1单抗免疫治疗)反应性相关的标志物的筛选

[0124] 1、研究人群

[0125] 本发明回顾性收集了中国医学科学院肿瘤医院2020年1月至2021年10月期间根据AJCC第8版标准诊断为HER2阴性乳腺癌的患者26例,采集患者的肿瘤组织样本,均同患者及

患者家属说明了实验情况并同其签订了知情同意书,并通过组织伦理委员会的同意。

[0126] 纳入标准:

[0127] 1) 在签署知情同意书当天年满18岁的男性或女性。

[0128] 2) 经当地病理学医生证实,受试者患有局部乳腺浸润性导管癌,分期为T1c-T2且临床淋巴结分期cN1-cN3,或T3-T4/cN0-cN3,或M1。

[0129] 3) 按照最近的美国临床肿瘤学会/美国病理学家学会指南,患有HER2-的乳腺导管癌患者。

[0130] 4) 女性受试者未怀孕,未进行哺乳,并且末次给药后至少12个月避孕。

[0131] 5) 具有适当的器官功能。

[0132] 排除标准:

[0133] 1) 具有需要类固醇治疗的肺炎(非感染性)病史或当前患有肺炎。

[0134] 2) 具有需要系统性治疗的活动性感染。

[0135] 3) 在研究治疗首次给药前72小时内尿妊娠试验阳性。

[0136] 4) 既往曾接受过抗PD-1,抗PD-L1或抗PD-L2药物治疗或作用于另一个刺激性或协同抑制性T细胞受体的靶向药物。

[0137] 5) 被诊断为免疫缺陷或在试验治疗首次给药前7天内正在接受长期全身类固醇治疗(每日剂量超过10mg泼尼松等效物)或任何形式的免疫抑制治疗。

[0138] 6) 在过去的2年内患有需要全身性治疗的活动性自体免疫疾病(如,使用改善疾病药物、皮质激素类药物或免疫抑制性药物)。

[0139] 7) 已知具有活动性结核病史。

[0140] 8) 在筛选时进行超声心动,证实左心室射血分数(LVEF) < 50%或低于机构的正常限值。

[0141] 9) 已知具有人类免疫缺陷病毒(HIV)感染病史。

[0142] 10) 已知有乙型肝炎病史或已知有活动性丙型肝炎病毒感染。

[0143] 2、对乳腺癌患者的肿瘤组织进行28通道组织质谱成像

[0144] 所述乳腺癌患者均经PD-L1单抗免疫治疗。首先,本实施例对免疫治疗(PD-L1单抗免疫治疗)患者肿瘤组织进行28通道组织质谱成像并绘制单细胞分辨率图谱。根据本发明的发明人研究组(PMID:34653365)前期已发表的单细胞测序研究结果筛选多重标记成像中常用的肿瘤微环境标记物。主要包括细胞身份鉴定蛋白、细胞功能标志物、免疫调节蛋白等。最终形成了免疫微环境抗体标志物组合包括28个标志物的IMC panel。

[0145] 对纳入研究的患者肿瘤组织进行组织质谱成像。研究首先通过病理形态分析区分了病理微解剖区域,每例患者肿瘤组织分别对肿瘤组织3个结构区域各进行2处ROI (Region of interest) 重复扫描。成像数据通过MCD文件存储,可利用Cytomapper程序包绘制通道合成伪彩色照片。

[0146] 对乳腺癌组织质谱成像数据进行细胞识别和聚类。使用CellProfiler的内建机器学习算法进行图像单细胞分割。后使用无监督学习算法Phenograph基于单细胞形态学及共有蛋白标记物进行聚类,得到共25个细胞亚群,包括上皮来源肿瘤细胞、免疫细胞及间质细胞。经标记物可视化后合并相同亚群,最终得到23类具有生物意义的细胞亚群注释。

[0147] 3、实验结果

[0148] 结果显示,筛选得到的细胞身份鉴定蛋白标志物包括:CKAE1/AE3、 $\beta$ -catenin、E-cadherin、CD45、CD3、CD8、CD4、FoxP3、CD11b、CD11c、CD14、CD15、CD16、CD19、CD33、CD56、CD68、Vimentin、 $\alpha$ SMA、collagen-I;所述细胞功能蛋白标志物包括:dsDNA、Ki-67、HLA-DR、IFN- $\gamma$ ;所述免疫调节蛋白标志物包括:PD1、PDL1、PDL2、TGF- $\beta$ 、TNF- $\alpha$ ;即本发明经筛选鉴定得到的与乳腺癌患者对PD-L1单抗免疫治疗反应性相关的标志物包括:CKAE1/AE3、 $\beta$ -catenin、E-cadherin、CD45、CD3、CD8、CD4、FoxP3、CD11b、CD11c、CD14、CD15、CD16、CD19、CD33、CD56、CD68、Vimentin、 $\alpha$ SMA、collagen-I、dsDNA、Ki-67、HLA-DR、IFN- $\gamma$ 、PD1、PDL1、PDL2、TGF- $\beta$ 、TNF- $\alpha$ ;其中和乳腺癌肿瘤细胞相关的标志物包括15种:IFN- $\gamma$ 、HLA-DR、PD1、PDL1、PDL2、TGF- $\beta$ 、CKAE1/AE3、TNF- $\alpha$ 、 $\alpha$ SMA、collagen-I、E-cadherin、Vimentin、Ki-67、 $\beta$ -catenin、FoxP3(见图1-图15);基因相关性热图的结果见图16,结果显示上述标志物在乳腺癌肿瘤细胞中的表达强度与免疫治疗反应显著有关。

[0149] 结果显示,不同免疫治疗反应的乳腺癌患者组织质谱成像特征:在接受PD-L1单抗治疗后病灶出现明显缩小的转移性TNBC患者中,其基线肿瘤组织微环境中富含CD8+T细胞及单核巨噬细胞,B细胞亦具有差异趋势;免疫治疗应答不佳、疾病出现进展的患者,其基线肿瘤组织微环境中则可见到更多髓系细胞。Treg细胞亚群在组间则没有显著差异。在单细胞表达谱分析中,我们发现有效组(免疫治疗响应组)患者肿瘤细胞表达更多TGF- $\beta$ 及Ki-67,而无效组(免疫治疗非响应组)患者肿瘤细胞中更常见高表达CKAE1/AE3、 $\alpha$ SMA等标志物,推测高分化程度可能与更低免疫原性有关。通过无偏分析技术鉴定了免疫细胞的一个子群,该子集在多个样品中共同表达TGF- $\beta$ ,并在免疫治疗效果较好的患者中富集。

[0150] 实施例2用于乳腺癌患者对免疫治疗(PD-L1单抗免疫治疗)疗效预测的模型的建立与验证

[0151] 1、四风险因素(4RF-panel)模型的建立

[0152] 将实施例1中收集的乳腺癌患者随机平均分配到训练队列( $n=13$ )和验证队列( $n=13$ )中,4RF-panel模型是在训练队列中使用四个与免疫治疗疗效相关的临床风险因素的Logistic回归分析方法构建得到的,所述评估受试者免疫治疗疗效的预测评分模型如下所示:

[0153] 预测评分 $4RF-panel_{score} = IFN-\gamma * (-62.5851559) + \alpha SMA * (-0.3814794) + collagen-I * (28.1856122) + Vimentin * (8.9254198) - 0.4248442$

[0154] 其中,\*代表乘号,IFN- $\gamma$ 表达量的单位为较DNA抗体通道相对强度值, $\alpha$ SMA表达量的单位为较DNA抗体通道相对强度值,collagen-I表达量的单位为较DNA抗体通道相对强度值,Vimentin表达量的单位为较DNA抗体通道相对强度值。

[0155] 在训练队列中,采用0%假阳性率和100%真阳性率的点的预测评分作为最佳截断值(Cut-off),并将其直接应用于训练队列、验证队列中,以将患者分为低评分组和高评分组。

[0156] 2、采用四风险因素(4RF-panel)模型进行免疫治疗疗效的预测及其预测诊断效能的评估

[0157] 为了确定上述构建得到的四风险因素(4RF-panel)模型在临床上预测乳腺癌患者对PD-L1单抗免疫治疗疗效的有效性,本实施例采用4RF-panel模型对训练队列中的每位乳腺癌患者进行了评分,得到预测评分 $4RF-panel_{score}$ 。为了进一步验证4RF-panel模型的预

测能力,本实施例对其诊断效能(ROC曲线)进行了分析。

### [0158] 3、统计分析

[0159] 数据在正态分布时以均数±标准差(SD)表示,在非正态分布时以中位数和四分位数间距(IQR)表示,在分类时以百分比表示,采用Wilcoxon符号秩检验(非正态分布变量)、Student's t检验(正态分布变量)和卡方检验(分类或二元变量)对两组中的基线特征的差异进行比较分析,所有基线特征均按复发次数分层,使用受试者工作特征曲线(AUROC)分析下的面积确定各个变量和已确定的危险因素组的疗效的预测性能,并采用DeLong检验进行比较,P<0.05为差异具有统计学意义,所有统计分析均采用R语言软件(v4.0.3)和Bioconductor软件(v3.13)进行。

### [0160] 4、实验结果

[0161] 本发明构建的四风险因素(4RF-panel)模型在训练集中预测免疫治疗疗效的AUC值高达0.843(见图17),在验证集中预测免疫治疗疗效的AUC值高达0.830(见图18),其中,在训练集最佳截断值为0.5,在验证集最佳截断值为0.4,评分大于截断值时表明受试者能够从免疫治疗中获益。

[0162] 对比例1本发明建立的用于乳腺癌患者对免疫治疗(PD-L1单抗免疫治疗)疗效预测的模型与其他模型的性能比较

### [0163] 1、实验方法

[0164] 本发明建立的用于乳腺癌患者对免疫治疗(PD-L1单抗免疫治疗)疗效预测的模型(目标模型),即四风险因素(4RF-panel)模型包含4种标志物:IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Vimentin。

[0165] 此外,在本实施例中还构建了另外几组乳腺癌患者对免疫治疗(PD-L1单抗免疫治疗)疗效预测的模型(即对比模型),以用于和目标模型的诊断效能进行比较。对比模型1中包含4种标志物:IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、E-cadherin,对比模型2中包含4种标志物:IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、 $\beta$ -catenin,对比模型3中包含4种标志物:IFN- $\gamma$ 、 $\alpha$ SMA、E-cadherin、Ki-67,对比模型4中包含4种标志物:IFN- $\gamma$ 、 $\alpha$ SMA、E-cadherin、 $\beta$ -catenin,对比模型5中包含4种标志物:IFN- $\gamma$ 、collagen-I、E-cadherin、Vimentin,对比模型6中包含4种标志物:IFN- $\gamma$ 、collagen-I、E-cadherin、Ki-67,对比模型7中包含5种标志物:IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、E-cadherin、Ki-67,对比模型8中包含5种标志物:IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、E-cadherin、 $\beta$ -catenin,对比模型9中包含5种标志物:IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、Ki-67、 $\beta$ -catenin,对比模型10中包含5种标志物:IFN- $\gamma$ 、 $\alpha$ SMA、E-cadherin、Ki-67、 $\beta$ -catenin,对比模型11中包含6种标志物:IFN- $\gamma$ 、TNF- $\alpha$ 、 $\alpha$ SMA、collagen-I、E-cadherin、Ki-67,对比模型12中包含6种标志物:IFN- $\gamma$ 、TNF- $\alpha$ 、 $\alpha$ SMA、E-cadherin、Ki-67、 $\beta$ -catenin,对比模型13中包含6种标志物:IFN- $\gamma$ 、TNF- $\alpha$ 、collagen-I、E-cadherin、Ki-67、 $\beta$ -catenin,对比模型14中包含6种标志物:IFN- $\gamma$ 、 $\alpha$ SMA、collagen-I、E-cadherin、Ki-67、 $\beta$ -catenin,本实施例分别采用上述模型对所有的乳腺癌患者进行了ROC曲线的分析。

### [0166] 2、实验结果

[0167] 结果表明,包含4-6个其他和乳腺癌免疫治疗疗效相关的标志物的对比模型的AUC值均低于0.800(见表1),而本发明构建的四风险因素(4RF-panel)模型预测免疫治疗疗效的AUC值高达0.843(见表1和图17和图18),其中,在训练集中的AUC=0.843,在验证集中的

AUC=0.830,显著高于包含4-6个标志物的对比模型的AUC值(AUC=0.674-0.789),表明了本发明建立的四风险因素(4RF-pane1)模型具有较好的预测乳腺癌患者免疫治疗疗效的能力,具有良好的诊断预测效能。

[0168] 表1不同预测模型预测效能的结果对比

治疗疗效预测模型	诊断预测效能 (AUC 值)
<b>目标模型 (4 个标志物)</b>	<b>0.843</b>
对比模型 1 (4 个标志物)	0.711
对比模型 2 (4 个标志物)	0.722
对比模型 3 (4 个标志物)	0.682
对比模型 4 (4 个标志物)	0.674
对比模型 5 (4 个标志物)	0.789
对比模型 6 (4 个标志物)	0.729
[0169] 对比模型 7 (5 个标志物)	0.740
对比模型 8 (5 个标志物)	0.722
对比模型 9 (5 个标志物)	0.747
对比模型 10 (5 个标志物)	0.709
对比模型 11 (6 个标志物)	0.761
对比模型 12 (6 个标志物)	0.744
对比模型 13 (6 个标志物)	0.770
对比模型 14 (6 个标志物)	0.747

[0170] 上述实施例的说明只是用于理解本发明的方法及其核心思想。应当指出,对于本领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以对本发明进行若干改进和修饰,这些改进和修饰也将落入本发明权利要求的保护范围内。

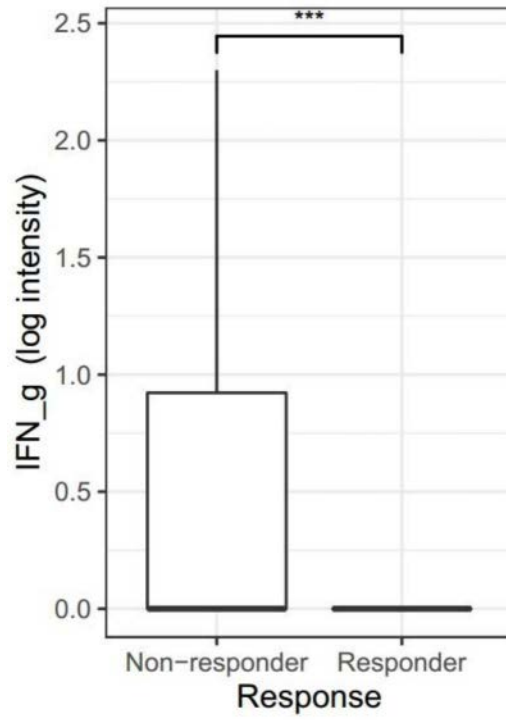


图1

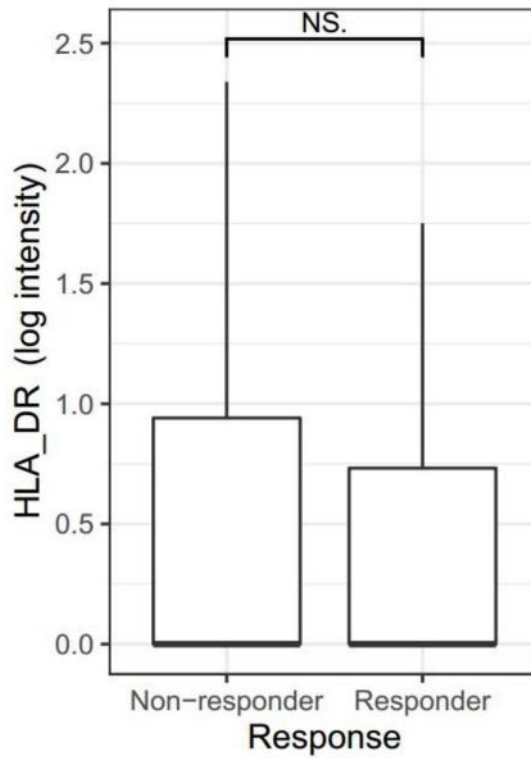


图2



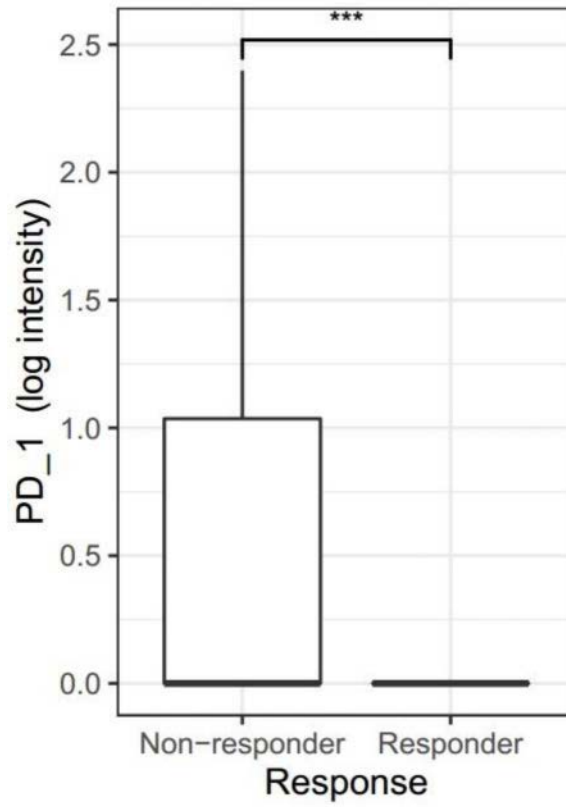


图3

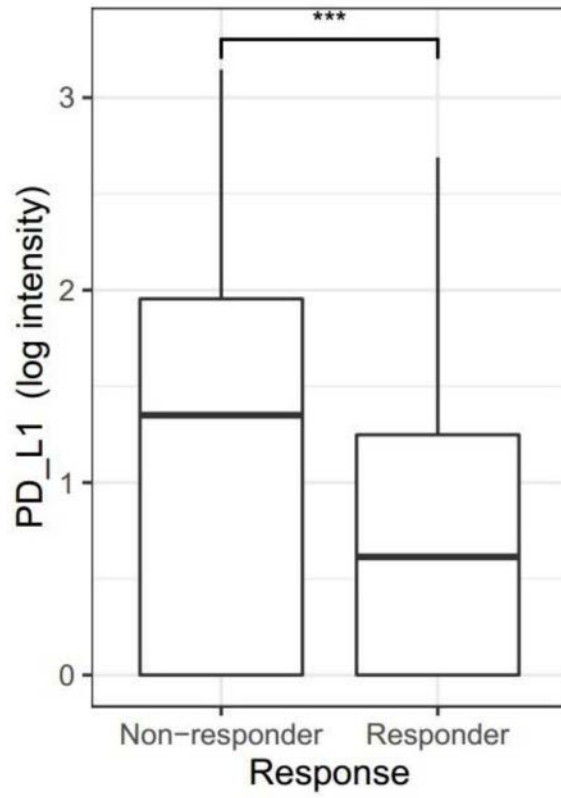


图4

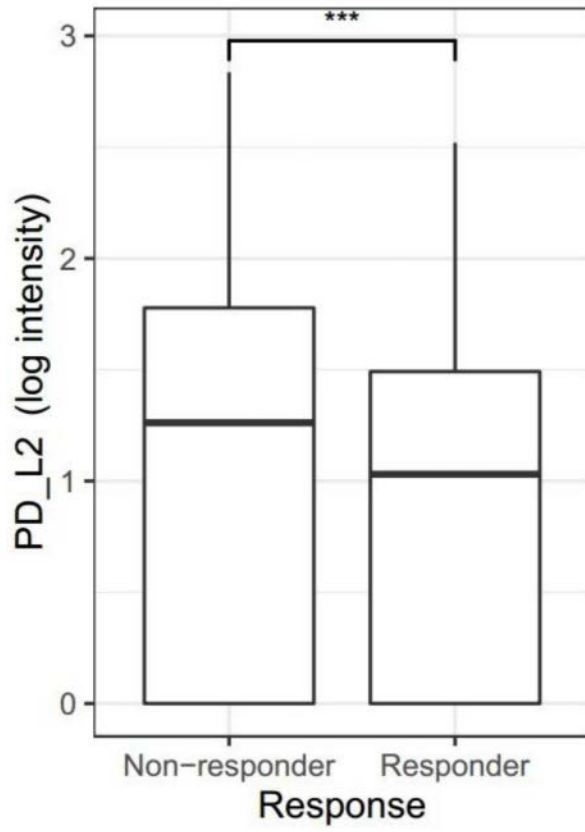


图5

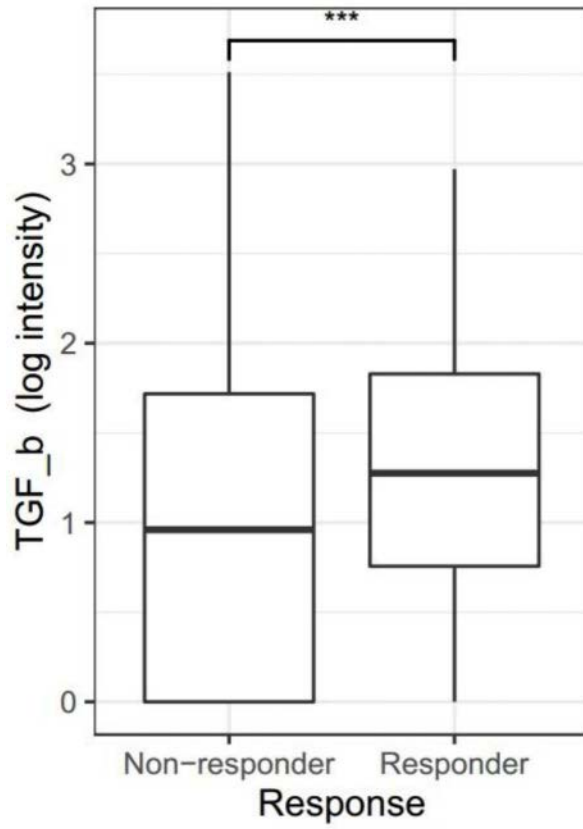


图6

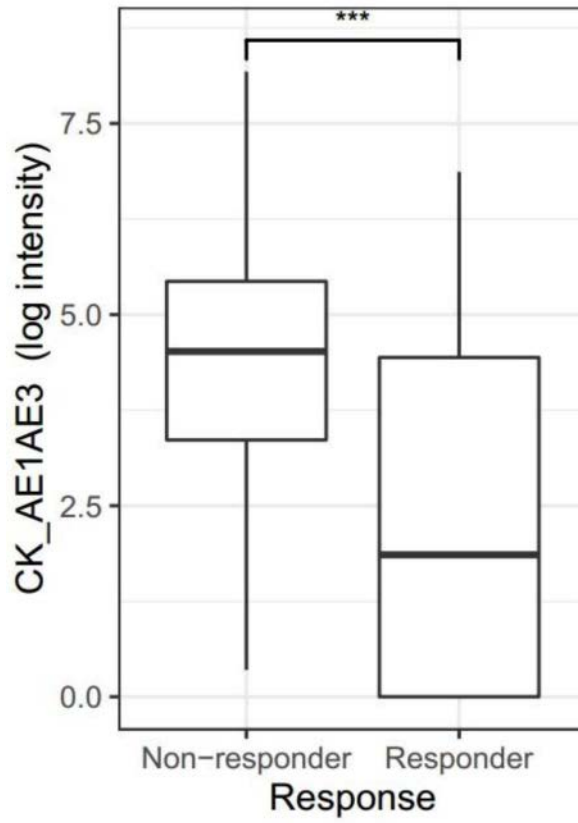


图7

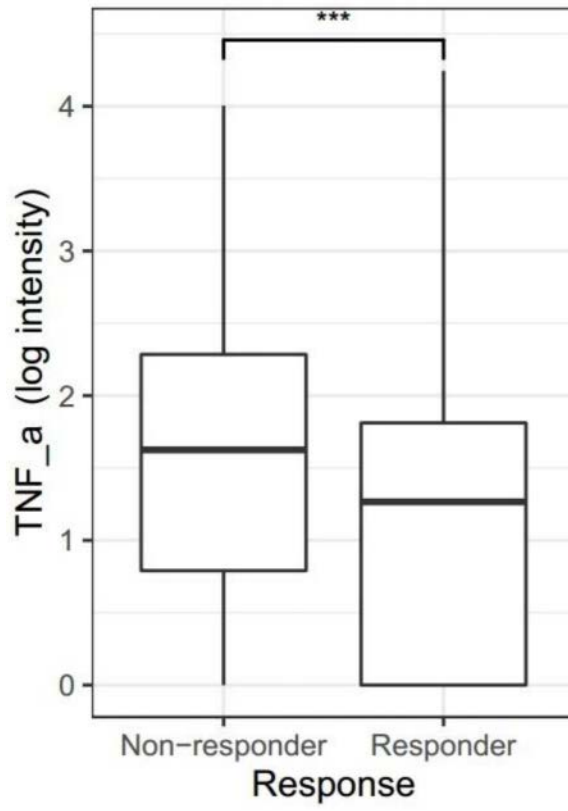


图8

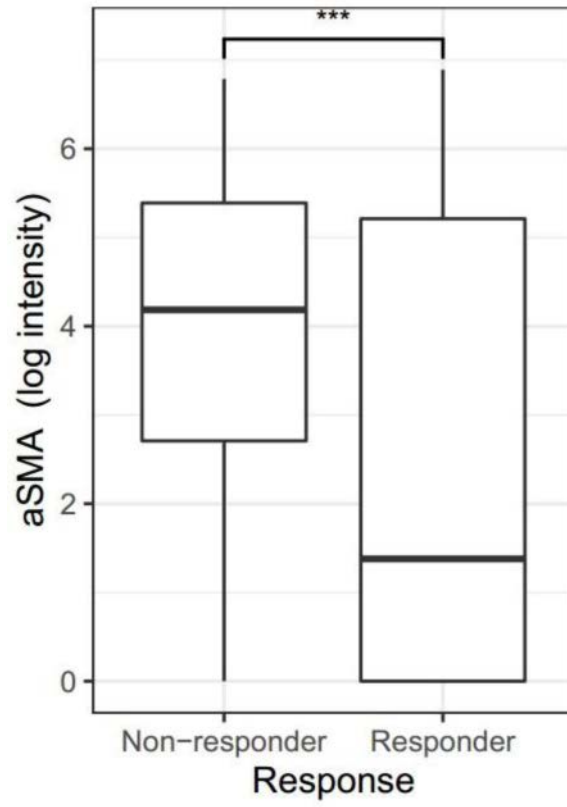


图9

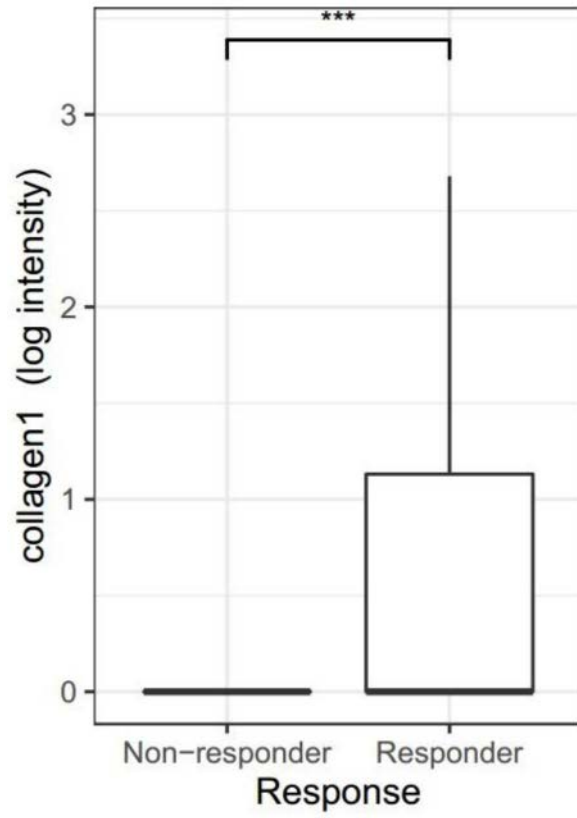


图10



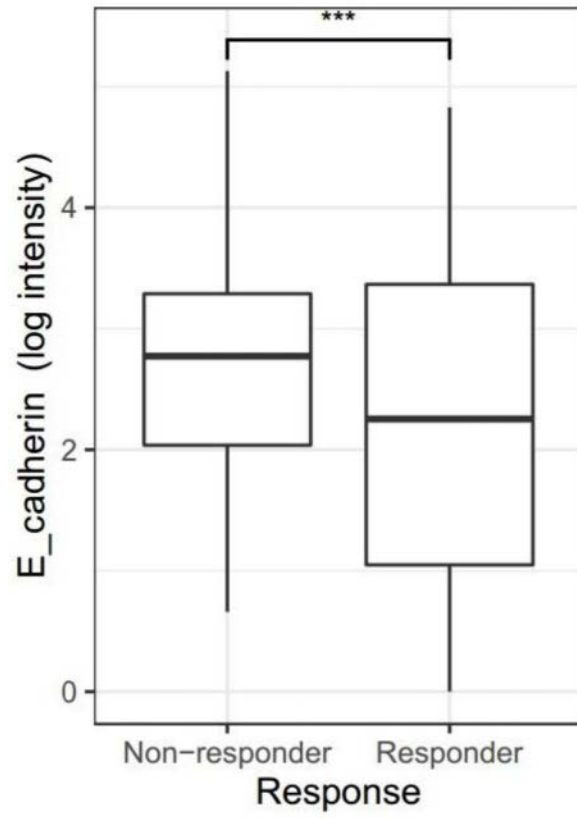


图11

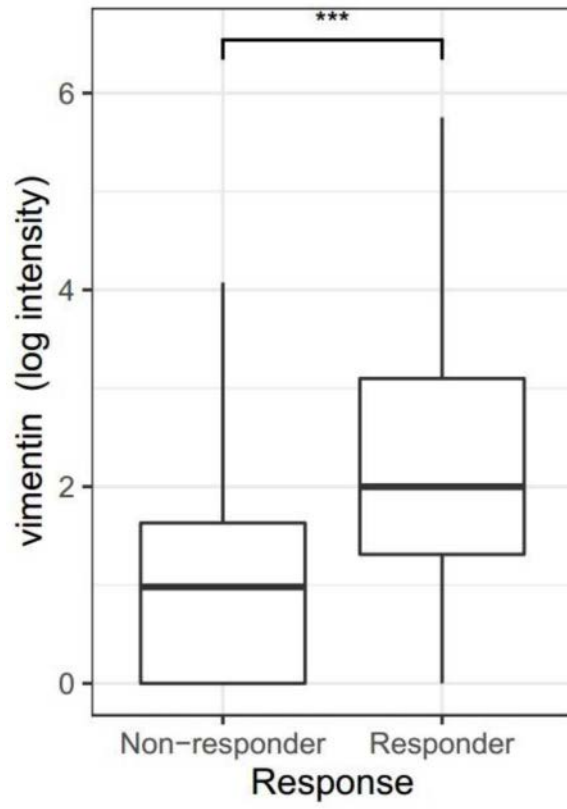


图12

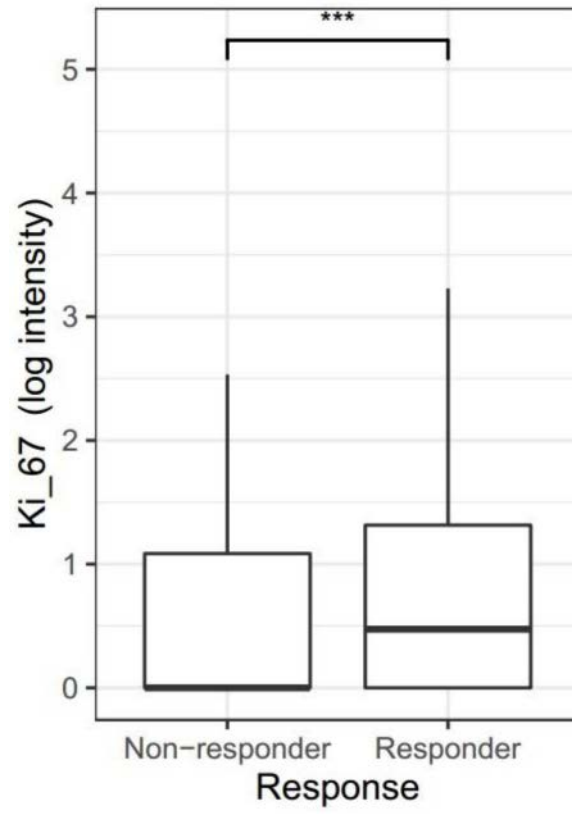


图13

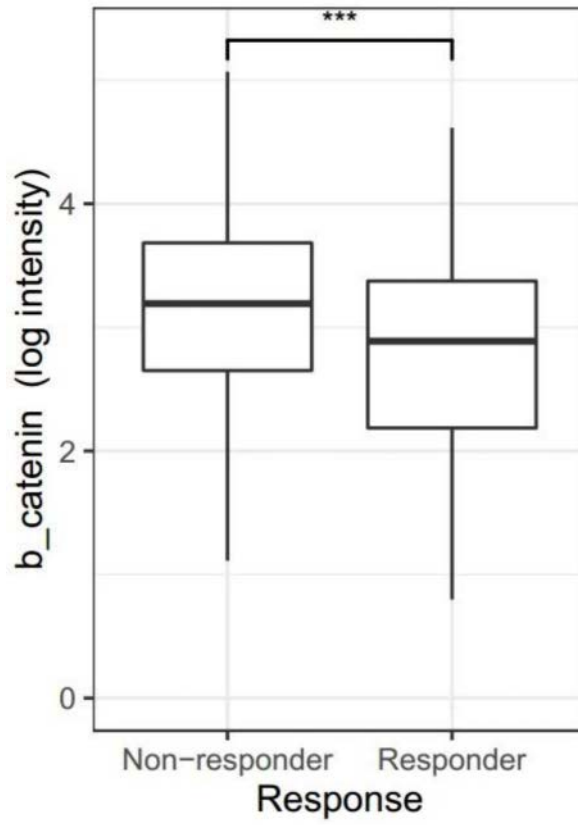


图14

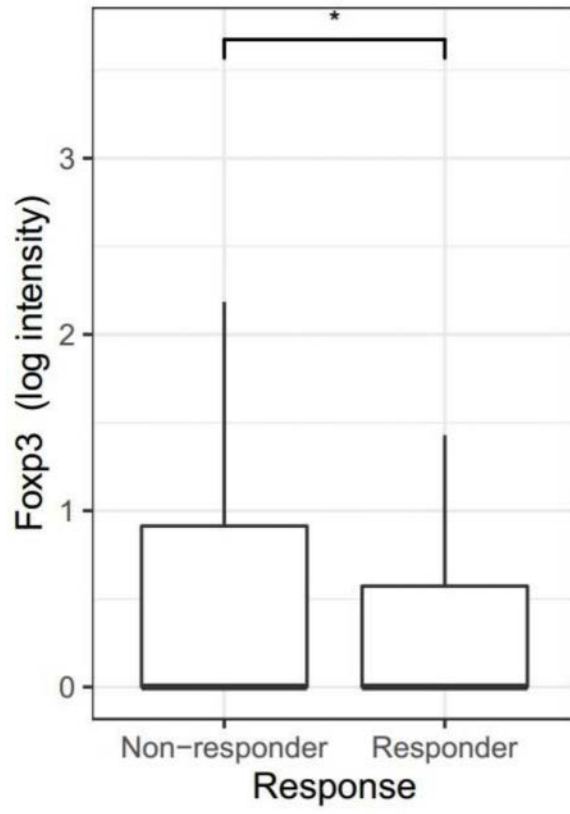


图15

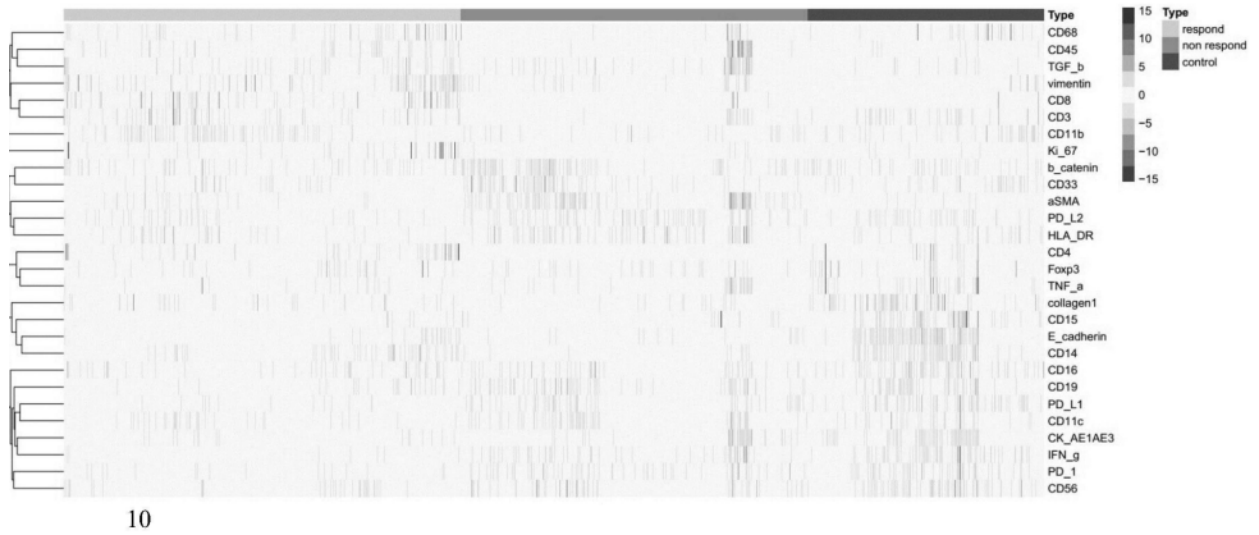


图16

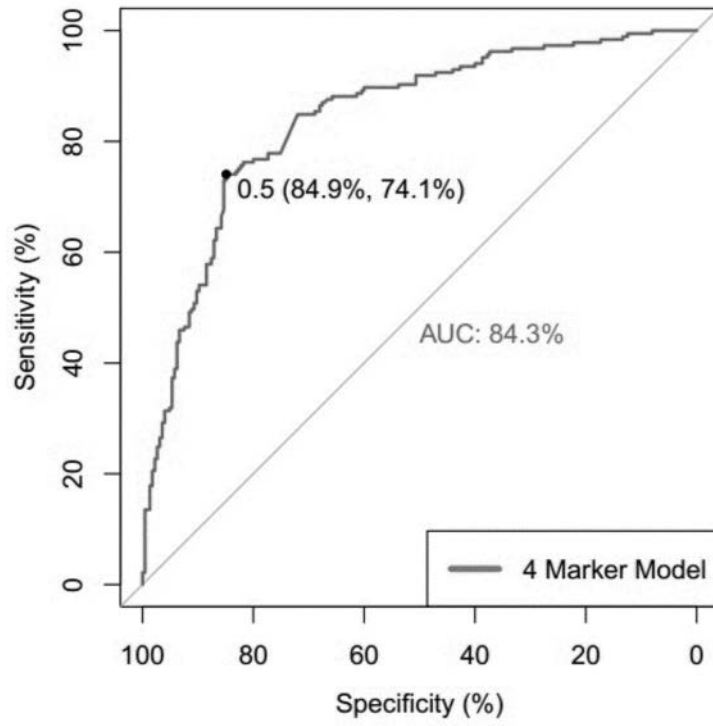


图17

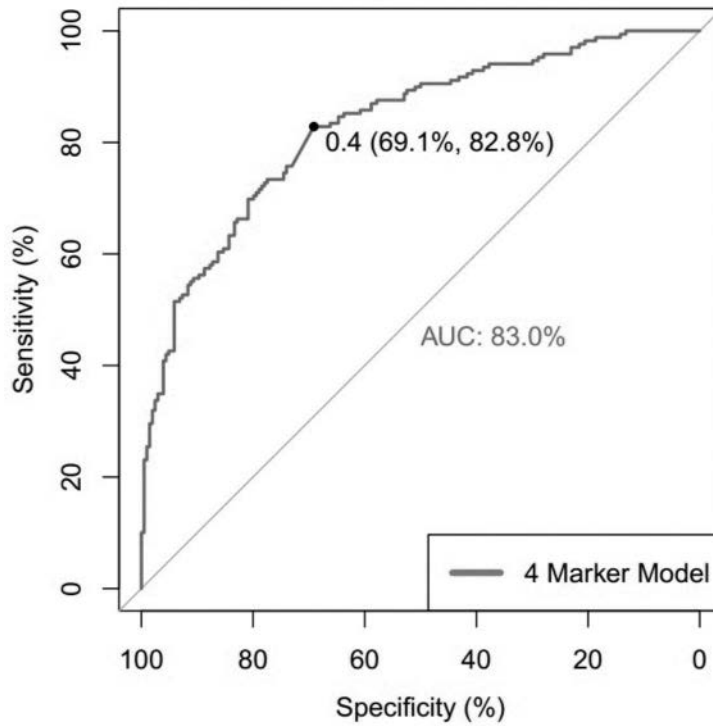


图18