



# (12) 发明专利

(10) 授权公告号 CN 118406766 B

(45) 授权公告日 2024. 11. 01

(21) 申请号 202410873864.8

(22) 申请日 2024.07.02

(65) 同一申请的已公布的文献号  
申请公布号 CN 118406766 A

(43) 申请公布日 2024.07.30

(73) 专利权人 天津云检医学检验所有限公司  
地址 300458 天津市滨海新区高新六路39  
号渤龙湖总部基地二区32号楼  
专利权人 天津云检医疗器械有限公司

(72) 发明人 陈利民

(74) 专利代理机构 北京卫智易创专利代理事务  
所(普通合伙) 16015  
专利代理师 朱春野

(51) Int. Cl.  
C12Q 1/6886 (2018.01)  
C12N 15/11 (2006.01)

(56) 对比文件

Junpeng Cui等.Development and validation of a prognostic 9-gene signature for colorectal cancer.Front Oncol.2022,第12卷DEGs in TCGA training dataset部分,补充表2.

Sha Zhou等.TRIM25 regulates oxaliplatin resistance in colorectal cancer by promoting EZH2 stability.Cell Death & Disease.2021,第12卷结果部分.

Elizabeth Day等.IRS2 is a candidate driver oncogene on 13q34 in colorectal cancer.Int J Exp Pathol.2013,第94卷(第3期),摘要,结果部分.

审查员 李玲玲

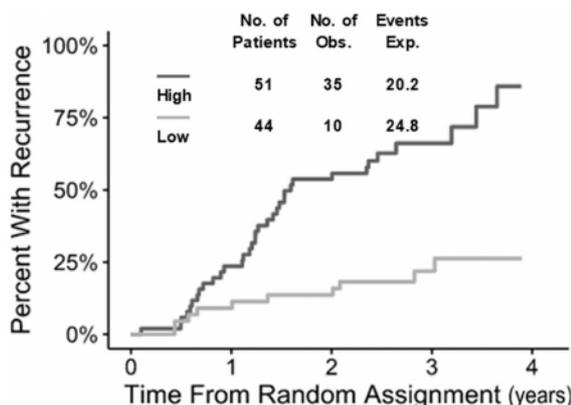
权利要求书1页 说明书14页  
序列表(电子公布) 附图6页

## (54) 发明名称

预测II期和III期结直肠癌预后临床结果的生物标志物、引物组及检测试剂盒

## (57) 摘要

本发明实施例公开了预测II期和III期结直肠癌预后临床结果的生物标志物、引物组及检测试剂盒;生物标志物为IRS2、TRIM25、MYBL2、KRT6B、FMNL2、NAT1、FAP、MYC、INHBA。并通过检测结直肠癌生物标志物水平可以得到结直肠癌RS评分,能够有效区分结直肠癌复发风险的高低;II期结直肠癌公式为(IRS2-TRIM25)+(IRS2-MYBL2)+(KRT6B-FMNL2)+(INHBA-MYC);III期结直肠癌公式为(IRS2-TRIM25)+(KRT6B-NAT1)+(FAP-MYC)+(INHBA-MYC),根据评分结果为患者提供更加个性化和有效的治疗方案。



1. 预测II期和III期结直肠癌预后临床结果的生物标志物,其特征在于,  
用于II期结直肠癌患者复发风险评估的生物标志物为基因对IRS2-TRIM25、IRS2-MYBL2、KRT6B-FMNL2和INHBA-MYC的组合;  
用于III期结直肠癌患者复发风险评估的生物标志物为基因对IRS2-TRIM25、KRT6B-NAT1、FAP-MYC和INHBA-MYC的组合。
2. 用于PCR检测权利要求1所述的生物标志物表达水平的引物组在制备用于预测II期和III期结直肠癌预后临床结果的检测试剂盒中的应用,其特征在于,所述引物组由IRS2引物组、TRIM25引物组、KRT6B引物组、MYBL2引物组、FMNL2引物组、NAT1引物组、FAP引物组、MYC引物组、INHBA引物组组成,上述引物组的序列分别如SEQ ID No.1~18所列。
3. 根据权利要求2所述的应用,其特征在于:  
用于II期结直肠癌患者复发风险评估的公式为:  
 $(IRS2-TRIM25) + (IRS2-MYBL2) + (KRT6B-FMNL2) + (INHBA-MYC)$  ;  
若评分高于临界值2.589,表明II期结直肠癌患者在初始治疗后处于预后不良的高风险;  
用于III期结直肠癌患者复发风险评估的公式为:  
 $(IRS2-TRIM25) + (KRT6B-NAT1) + (FAP-MYC) + (INHBA-MYC)$  ;  
若评分高于临界值-0.1356,表明III期结直肠癌患者的结直肠癌复发风险较高;  
其中,IRS2-TRIM25、IRS2-MYBL2、KRT6B-FMNL2、INHBA-MYC、KRT6B-NAT1、FAP-MYC表示基因的循环阈值差值  $\Delta Ct$ 。

## 预测II期和III期结直肠癌预后临床结果的生物标志物、引物组及检测试剂盒

### 技术领域

[0001] 本发明属于生物检测技术领域,具体涉及预测II期和III期结直肠癌预后临床结果的生物标志物、引物组及检测试剂盒。

### 背景技术

[0002] 在标准护理下预测癌症治疗结果能力可以帮助医生决定潜在的辅助治疗。例如,如果患者的预后治疗结果较差,可能会给予更积极的辅助治疗。目前,在实践中使用的大多数诊断试验或临床适应症都不是定量的,如免疫组化(ICH),它通常在不同的实验室产生不同的结果,这可能是由于试剂不标准和对结果的主观解释所导致的。

[0003] 能够根据复发风险和初始治疗后的临床结果对II期和III期结肠直肠癌(CRC)患者进行分层,临床医生就能够制定最适合特定患者的化疗方案。目前,辅助化疗是III期结肠癌患者的标准护理,但缺乏识别可能不需要奥沙利铂基础化疗的低风险III期结直肠癌患者的临床数据。对于II期结直肠癌患者,仅推荐高危患者进行辅助化疗。

[0004] 然而,目前定义高危患者特征的标准还不够充分。对于已知的II期结直肠癌,导致预后差的预后因素包括肿瘤浸润深度(T分期)、淋巴结数、错配修复(MMR)状态、肿瘤分级、淋巴和血管浸润(LVI)评分,可只有T分期、淋巴结数量和MMR状态有大量的临床证据支持其有限的使用。

[0005] 公司(Exact Sciences公司)开发了一种12个基因的复发评分检测,即:Oncotype DX结肠复发评分试验,该检测基于定量聚合酶链反应qPCR结果的评分系统,能够定量地显示复发风险的可能性。但是,使用单协变量Cox回归时,每四分位数范围的复发率评分IQR的风险比HR仅为1.38,而使用多协变量Cox回归时,风险比HR为1.43。Oncotype Dx对结直肠癌的复发评分比目前的临床预测指标T分期和MMR状态差,与检查的淋巴结数量接近。因此,12基因复发评分检测在结直肠癌中的临床应用很低,不常用于临床。

### 发明内容

[0006] 有鉴于此,一方面,一些实施例公开了预测II期和III期结直肠癌预后临床结果的生物标志物,生物标志物为IRS2、TRIM25、MYBL2、KRT6B、FMNL2、NAT1、FAP、MYC、INHBA。

[0007] 进一步,一些实施例公开了预测II期和III期结直肠癌预后临床结果的生物标志物,用于II期结直肠癌患者复发风险评估的生物标志物为IRS2-TRIM25、IRS2-MYBL2、KRT6B-FMNL2和INHBA-MYC的组合。

[0008] 一些实施例公开了预测II期和III期结直肠癌预后临床结果的生物标志物,其特征在于,用于III期结直肠癌患者复发风险评估的生物标志物为IRS2-TRIM25、KRT6B-NAT1、FAP-MYC和INHBA-MYC的组合。

[0009] 另一方面,一些实施例公开了用于PCR检测预测II期和III期结直肠癌预后临床结果的生物标志物的引物组,包括IRS2引物组、TRIM25引物组、KRT6B引物组、MYBL2引物组、

FMNL2引物组、NAT1引物组、FAP引物组、MYC引物组、INHBA引物组,上述引物组的序列分别如SEQ ID No.1~18所列。

[0010] 再一方面,一些实施例公开了用于预测II期和III期结直肠癌预后临床结果的检测试剂盒,包含用于PCR检测预测II期和III期结直肠癌预后临床结果的生物标志物的引物组。

[0011] 进一步,一些实施例公开的用于预测II期和III期结直肠癌预后临床结果的检测试剂盒,用于II期结直肠癌患者复发风险评估的生物标志物为IRS2-TRIM25、IRS2-MYBL2、KRT6B-FMNL2和INHBA-MYC的组合。

[0012] 一些实施例公开的用于预测II期和III期结直肠癌预后临床结果的检测试剂盒,用于III期结直肠癌患者复发风险评估的生物标志物为IRS2-TRIM25、KRT6B-NAT1、FAP-MYC和INHBA-MYC的组合。

[0013] 进一步,一些实施例公开的用于预测II期和III期结直肠癌预后临床结果的检测试剂盒,用于测量结直肠癌患者组织样本的生物标志物的循环阈值。

[0014] 一些实施例公开的用于预测II期和III期结直肠癌预后临床结果的检测试剂盒,生物标志物的循环阈值转换为结直肠癌RS评分。

[0015] 一些实施例公开的用于预测II期和III期结直肠癌预后临床结果的检测试剂盒,II期结直肠癌公式为 $(IRS2-TRIM25) + (IRS2-MYBL2) + (KRT6B-FMNL2) + (INHBA-MYC)$ ,若结直肠癌RS评分高于临界值2.589,表明II期结直肠癌患者在初始治疗后处于预后不良的高风险;III期结直肠癌公式为 $(IRS2-TRIM25) + (KRT6B-NAT1) + (FAP-MYC) + (INHBA-MYC)$ ,若结直肠癌RS评分高于临界值-0.1355,表明III期结直肠癌患者的结直肠癌复发风险较高。

[0016] 本发明实施例公开的预测II期和III期结直肠癌预后临床结果的生物标志物组、引物组及检测试剂盒,并通过检测结直肠癌生物标志物水平可以得到结直肠癌RS评分,能够有效区分结直肠癌复发风险的高低;可以更好地评估结直肠癌患者的治疗风险和预后,从而为患者提供更加个性化和有效的治疗方案。

## 附图说明

[0017] 图1、Kaplan-Meier估计单独手术患者的三年复发率(全队列患者);

[0018] 图2、Kaplan-Meier估计单独手术患者的三年复发率(II期患者);

[0019] 图3、Kaplan-Meier估计单独手术患者的三年复发率(III期患者);

[0020] 图4、mProbe复发评分(RS)组的无复发间隔(所有患者);

[0021] 图5、mProbe复发评分(RS)组的无复发间隔(II期患者);

[0022] 图6、mProbe复发评分(RS)组的无复发间隔(IIIA/B期患者);

[0023] 图7、Kaplan-Meier通过使用Oncotype Dx和评分系统估计单独手术患者风险组的三年复发率(全队列患者);

[0024] 图8、Kaplan-Meier通过使用Oncotype Dx和评分系统估计单独手术患者风险组的三年复发率(II期患者);

[0025] 图9、Kaplan-Meier通过使用Oncotype Dx和评分系统估计单独手术患者风险组的三年复发率(III期结肠癌患者);

[0026] 图10、根据Oncotype Dx复发评分(RS)组的无复发间隔(所有患者);

[0027] 图11、根据Oncotype Dx复发评分(RS)组的无复发间隔(II期患者)；

[0028] 图12、根据Oncotype Dx复发评分(RS)组的无复发间隔(IIIA/B期患者)。

### 具体实施方式

[0029] 在这里专用的词“实施例”，作为“示例性”所说明的任何实施例不必解释为优于或好于其它实施例。本发明实施例中性能指标测试，除非特别说明，采用本领域常规试验方法。应理解，本发明实施例中所述的术语仅仅是为描述特别的实施方式，并非用于限制本发明实施例公开的内容。

[0030] 本文所用的术语“基本”和“大约”用于描述小的波动。例如，它们可以是指小于或等于 $\pm 5\%$ ，如小于或等于 $\pm 2\%$ ，如小于或等于 $\pm 1\%$ ，如小于或等于 $\pm 0.5\%$ ，如小于或等于 $\pm 0.2\%$ ，如小于或等于 $\pm 0.1\%$ ，如小于或等于 $\pm 0.05\%$ 。在本文中以范围格式表示或呈现的数值数据，仅为方便和简要起见使用，因此应灵活解释为不仅包括作为该范围的界限明确列举的数值，还包括该范围内包含的所有独立的数值或子范围。例如，“1~5%”的数值范围应被解释为不仅包括1%至5%的明确列举的值，还包括在所示范围内的独立值和子范围。因此，在这一数值范围中包括独立值，如2%、3.5%和4%，和子范围，如1%~3%、2%~4%和3%~5%等。这一原理同样适用于仅列举一个数值的范围。此外，无论该范围的宽度或所述特征如何，这样的解释都适用。在本文中，包括权利要求书中，连接词，如“包含”、“包括”、“带有”、“具有”、“含有”、“涉及”、“容纳”等被理解为是开放性的，即是指“包括但不限于”。只有连接词“由……构成”和“由……组成”是封闭连接词。

[0031] 为了更好的说明本发明内容，在下文的具体实施例中给出了众多的具体细节。本领域技术人员应当理解，没有某些具体细节，本发明同样可以实施。在实施例中，对于本领域技术人员熟知的一些方法、手段、仪器、设备等未作详细描述，以便凸显本发明的主旨。

[0032] 在不冲突的前提下，本发明实施例公开的技术特征可以任意组合，得到的技术方案属于本发明实施例公开的内容。

[0033] 本发明公开了结直肠癌(CRC)生物标志物、CRC生物标志物组合，以及获取样本的CRC风险评估生物标志物水平表示的方法。这些组合和方法在评估II期和III期治疗结果的风险以及结直肠癌复发的可能性方面发挥作用。同时，还提供了用于实现上述方法的系统、设备和试剂盒。下文中充分描述的组合和方法的细节，将对本领域技术人员理解本文公开的技术方案变得显而易见。

[0034] 需要理解的是，本发明实施例的公开内容并不局限于所描述的特定方法或组合，因为这些可能会有所变化。同时，需要明白的是，此处使用的术语仅仅是为了描述特定的实施例，并不是为了限制。

[0035] 除非另有定义，本文中使用的所有技术和科学术语均应按照该发明所属技术领域的普通技术人员通常理解的含义来解释。尽管在实践或测试本发明时可以使用与本文描述的相似或等效的任何方法和材料，但现在将描述一些潜在的和首选的方法和材料。本文中提到的所有出版物均通过引用纳入本文，以披露和描述与引用出版物相关的方法和/或材料。如果存在矛盾，应当理解为本文披露的内容取代了纳入出版物的任何披露。

[0036] 本发明文件公开后，对那些在该技术领域具有专业技能的人是显而易见的，这里描述和说明的每一个具体实施例都具有可以轻易地从其他几个实施例的特征中分离出来

或者结合起来的独立组件和特征,而这并不会偏离本发明的范围或发明构思。任何提及的方法都可以按照所提及的事件顺序来执行,或者按照任何其他逻辑上可能的顺序来执行。

[0037] 本文中,“结直肠癌”或“CRC”指的是在结肠或直肠中生长的异常且失控的疾病。它也被称为结肠癌,而结肠是大肠或大肠。直肠是连接结肠到肛门的通道。所谓的“对结直肠癌进行风险评估”,或“将CRC风险确定作为临床体征之一”,通常意味着提供一个作为额外临床体征的CRC风险,例如,确定患者在初始治疗后CRC复发的风险;对疾病进展过程和/或疾病结果的风险评估,例如五年总体生存率或无进展生存率。

[0038] 以下示例提出是为了向本领域技术人员提供关于如何制作和使用本发明技术方案的描述,并不打算限制发明范围,也不打算表明下面的实验是全部或唯一的实验。文中确保使用的数字(例如,数量、温度等)的准确性,但应该考虑到一些实验误差和偏差。

[0039] 在一些实施方式中,确定受试者的结直肠癌生物标志物表达水平的方法,包括:

[0040] a. 评估样本中的一组结直肠癌生物标志物,以确定样本中每种结直肠癌生物标志物的表达水平,其中,样本包括手术切除的肿瘤组织或结直肠癌患者的FFPE组织;结直肠癌生物标志物为IRS2、TRIM25、MYBL2、KRT6B、FMNL2、NAT1、FAP、MYC、INHBA;通常可以通过检测核酸序列,得到生物标志物的RNA或DNA的水平;通常生物标志物的表达水平还可以通过循环阈值表示,即Ct值;通常循环阈值可以表达为绝对浓度或中位数倍数(即MoM);

[0041] b. 根据组合中每个结直肠癌生物标志物的表达水平获得结直肠癌生物标志物水平。

[0042] 通常,样本组织/样本FFPE组织的生物标志物表达水平与肿瘤组织中结直肠癌复发风险相关,可以利用qPCR检测系统,例如ABI QuantStudio 4,准确测量生物标志物的表达水平,得出结直肠癌风险评分,根据风险评分可以区分高风险患者和低风险患者。

[0043] 在某些实施实例中,用于qPCR检测生物标志物表达水平的引物包括IRS2引物组、TRIM25引物组、KRT6B引物组、MYBL2引物组、FMNL2引物组、NAT1引物组、FAP引物组、MYC引物组、INHBA引物组;对于在原发肿瘤组织中出现的与结直肠癌复发相关的基因,利用定量PCR系统准确测量组织生物标志物基因表达,例如拷贝数,从而得出复发评分RS,将II期和III期CRC的高风险和低风险患者进行区分。PCR系统例如包括Applied Biosystem QuantStudio、7500 Real-Time PCR系统或Bio-Rad CFX系统。其中,IRS2引物组是指IRS2基因的特异性引物对,该特异性引物对可以在qPCR分析中,将IRS2基因的mRNA转换为cDNA,然后进行扩增,并进行qPCR检测、分析。

[0044] 基因名称与对应的引物组的序列分别如下:

[0045] 基因名称:IRS2;

[0046] SEQ ID No.1:CCTACGCCAGCATTGACTT;

[0047] SEQ ID No.2:TGACATCCTGGTGATAAAGCC;

[0048] 基因名称:TRIM25;

[0049] SEQ ID No.3:GGATGAGTTCGAGTTTCTGGAG;

[0050] SEQ ID No.4:GCCTTTTATCAGCTTGTGGTTC;

[0051] 基因名称:MYBL2;

[0052] SEQ ID No.5:GCCGAGATCGCCAAGATG;

[0053] SEQ ID No.6:CTTTTGATGGTAGAGTTCCAGTGATTC;

- [0054] 基因名称:KRT6B;
- [0055] SEQ ID No.7:GCCCTCACTTTTCTTCTCATCA;
- [0056] SEQ ID No.8:CATGTCTGAGTGCTGATAACTGT;
- [0057] 基因名称:FNML2;
- [0058] SEQ ID No.9:AGAATGAAGCCATGTCCAAGA;
- [0059] SEQ ID No.10:GTGTGAACTTGAGTATTTGCATCT;
- [0060] 基因名称:NAT1;
- [0061] SEQ ID No.11:AACTGAAGATCAACCTACTTTCAAC;
- [0062] SEQ ID No.12:GTTTCCAAGTCCAATTTGTTCTCA;
- [0063] 基因名称:FAP;
- [0064] SEQ ID No.13:CTGACCAGAACCACGGCT;
- [0065] SEQ ID No.14:GGAAGTGGGTCATGTGGG;
- [0066] 基因名称:MYC;
- [0067] SEQ ID No.15:TCCCTCCACTCGGAAGGACTA;
- [0068] SEQ ID No.16:CGGTTGTTGCTGATCTGTCTCA;
- [0069] 基因名称:INHBA;
- [0070] SEQ ID No.17:GTGCCCAGCCATATAGCA;
- [0071] SEQ ID No.18:CGGTAGTGGTTGATGACTGTTGA.

[0072] 一些实施例中,利用生物标志物区分II期和III期结直肠癌患者的复发风险。在II期结直肠癌风险评估中,生物标志物是IRS2-TRIM25、IRS2-MYBL2、KRT6B-FMNL2和INHBA-MYC基因表达的循环阈值差值 $\Delta Ct$ 。对于III期结直肠癌风险评估,生物标志物是IRS2-TRIM25、KRT6B-NAT1、FAP-MYC和INHBA-MYC的循环阈值差值 $\Delta Ct$ 。

[0073] 一些实施实例中,生物标志物组合的一组用于区分结直肠癌患者复发的风险。用于结直肠癌风险评估的生物标志物组合可以至少包括上生物标志物。通常,生物标志物组合可以包括所有生物标志物的任何组合,一般地,较小的生物标志物组合足以区分表现较差的高风险和低风险患者,但更经济;然而,更大的组合可能会提供更详细的信息,并且可以应用于在不同区域的人口群体中。

[0074] 一些实施实例公开了一种基于二分类评分系统计算复发评分的方法,通过对生物标志物区进行RS评分,以区分高风险患者和低风险患者。RS评分低表明患者不太可能复发,RS评分高表明患者有可能复发。

[0075] 在某些情况下,临床参数与本文描述的生物标志物联合用于结直肠癌复发风险评估。本发明公开了用于确定患者RS评分的方法以及其他复发临床预测因子的方法。

[0076] 一些实施实例中,通过检测结直肠癌生物标志物水平得到结直肠癌RS评分,其中RS评分可以通过几何均值、多元线性判别分析或分布式梯度增强决策树机器学习,例如XGBoost,由血液中生物标志物的测量值计算得到。在二元模型中,受试者工作特征ROC曲线可以由生物标志物组合或配对得出。RS评分可以根据本文描述的方法分为低风险CRC复发临床评分或高风险CRC复发临床评分。

[0077] 一些实施实例中,在结直肠癌患者受试者初期缓解后,对结直肠癌患者进行风险评估的方法,由受试者组织样本中结直肠癌生物标志物的表达水平表示;对于结直肠癌风险

评估,结直肠癌RS评分在两个范围内评估为高或低以确定初始治疗后复发的风险;若结直肠癌RS评分高于RS临界值2.589,表面为例,表明II期结直肠癌患者在初始治疗后处于结直肠癌预后不良的高风险;该临界值可根据人群调整;若结直肠癌RS评分高于临界值-0.1355,表明III期结直肠癌患者的结直肠癌复发风险较高;该临界值可根据人群调整。

[0078] 一些实施例公开了对结直肠癌患者的预后结果/复发/再发进行评估的方法,包括以下步骤:

[0079] 获取患者的肿瘤样本,包括甲醛固定石蜡包埋组织,即FFPE组织;

[0080] 测量肿瘤样本中每个生物标志物的水平,例如RNA/DNA拷贝数或浓度;

[0081] 将每个生物标志物的水平与生物标志物的相应基因对的值进行比较。通常,配对值范围可以代表来自同一受试者的一个或多个肿瘤样本的生物标志物水平,或者是来自一个或多个癌症患者的一个或多个肿瘤样本的生物标志物水平,与良好预后受试者的生物标志物的参考值相比,生物样本中生物标志物组合的差异水平表明患者是否有高风险复发。

[0082] 一些实施例中,结直肠癌患者的预后结果/复发/再发进行评估的方法还包括计算RS评分以区分高风险和低风险结直肠癌患者。

[0083] 通常,可以通过结直肠癌生物标志物水平得到结直肠癌RS评分,其中RS评分的获得方法包括:

[0084] a. 可以从两个不同基因之间的循环阈值C(t)的差异得到,例如,存在Ct差异值的生物标志物组为IRS2-TRIM25、IRS2-MYBL2、KRT6B-FMNL2、KRT6B-NAT1、FAP-MYC以及INHBA-MYC。

[0085] 可以通过几何平均值、多元线性判别分析或分布式梯度增强决策树机器学习,从测量的血液生物标志物的值中得到;或,

[0086] c. 可以是每对生物标志物的倍数差异,如 $\Delta C(t)$ 、浓度差值或Ct值的 $\Delta MoM$ ,归一化以拟合具有适当临界值的范围量值,以区分高风险和低风险患者。例如,可以通过II期结直肠癌公式(IRS2-TRIM25)+(IRS2-MYBL2)+(KRT6B-FMNL2)+(INHBA-MYC)和III期结直肠癌公式(IRS2-TRIM25)+(KRT6B-NAT1)+(FAP-MYC)+(INHBA-MYC)推导得出。

[0087] 通常,生物标志物可以使用特定的引物和报告系统进行测量,以确定生物标志物和生物标志物的基因表达,例如基因拷贝数。通常,确定生物标志物的拷贝数的方法包括但不限于定量PCR检测、逆转录定量PCR、基因微阵列、RNA或DNA测序、夹心法分析、磁性捕获、微球捕获、电泳印迹、表面增强拉曼光谱、流式细胞术或质谱法。

[0088] 一些实施例中,通过特定的DNA引物与生物标志物RNA的结合,然后通过逆转录和聚合酶链反应,使用可检测的报告分子来确定生物标志物的拷贝数;报告分子例如SYBR绿染料等。

[0089] 一些实施例公开了用于定量qPCR系统的试剂盒,该试剂盒用于测量肿瘤样本中的RNA/DNA生物标志物。该试剂盒可能包括一个容器,用于存放从患有II期和III期结直肠癌的人类患者收集和分离的生物样本。该试剂盒包含至少一个用于测量结直肠癌风险生物标志物的试剂,用于将试剂与生物样本或生物样本的一部分反应以测量生物样本中至少一个结直肠癌生物标志物。试剂可能包装在单独的容器中。

[0090] 一些实施例中,用于定量qPCR系统的试剂盒还包括一个或多个对照参考样本和试剂,用于执行qPCR以检测如本文所述的生物标志物拷贝数。

- [0091] 一些实施例中,用利用试剂盒进行检测生物标志物的定量qPCR系统,包括:
- [0092] 测量结直肠癌生物标志物的平台系统,如QuantStudio 4系统;
- [0093] 计算结直肠癌风险评分的计算表;
- [0094] 确定患者复发和不良预后的风险的指示标准。
- [0095] 一些实施例中,用于定量qPCR系统的试剂盒,生物标志物表达水平的拷贝数以循环阈值C(t)表示。
- [0096] 一些实施例中,检测肿瘤组织中生物标志物的评估结直肠癌患者复发风险的方法包括:
- [0097] 收集结直肠癌患者的原发肿瘤或FFPE肿瘤组织样本,测量肿瘤组织中的每个生物标记物的拷贝数;
- [0098] 比较结直肠癌患者原发肿瘤或FFPE肿瘤组织中的一组生物标记物的测量值与相应的配对基因的参考值,可以评估这些生物标记物在组织中的表达差异;
- [0099] 基于这些差异表达的数据,可以计算出患者的结直肠癌复发风险评分,即RS评分,进一步根据RS评分评估患者的复发风险。
- [0100] 通常,为了测量至少一个生物标志物,需要使用特异性引物结合到该生物标志物或其片段上,并通过PCR技术来确定生物标志物的拷贝数。这些引物是根据生物标志物的编码区域(外显子)序列设计的。
- [0101] 在一些实施例中,选择特定的引物对,这些引物对能够特异性地结合一系列特定的生物标志物,如IRS2、TRIM25、MYBL2、KRT6B、FMNL2、INHBA、MYC、NAT1和FAP等,以便进行精确的测量和分析。
- [0102] 一些实施例中,测量至少一对CRC生物标记物,通过寡核苷酸引物针对生物标记物进行检测,从而计算RS评分,其中引物特异性地结合到生物标志物或包含生物标记物的RNA/DNA序列的生物标志物片段;
- [0103] 引物从生物标志物目标的部分互补DNA序列中选择;
- [0104] 染料或染料/猝灭剂报告标签可以准确地解释目标生物标志物的循环阈值,如样本中的RNA或DNA数量;
- [0105] qPCR仪器,如qstudio4,基于生物标志物模板的扩增结果,得到各生物标志物的Ct值,从而确定各生物标志物的表达水平。
- [0106] 实施例1
- [0107] 1实验方法
- [0108] 1.1样本信息
- [0109] 本实施例1的患者队列、人口统计信息和临床标准
- [0110] 结直肠癌(CRC)患者验证队列、人口统计信息和临床标准的所有研究方案均已获得复旦大学肿瘤医院机构审查委员会(IRB)的批准。从具有结果信息的结直肠癌患者获得了甲醛固定石蜡包埋(FFPE)组织样本,这些样本被分类为高风险和低风险人群。此外,还招募了错配修复(MMR)功能正常的结直肠癌患者进行研究。在复旦大学肿瘤医院中选择并表征了FFPE原发癌组织以进行分析。对于每位患者,都取得了5个未染色的5微米切片。患者队列、人口统计和临床信息显示在表1中。
- [0111] 表1患者队列、人口统计和临床信息列表

特征	复发	未复发	测试方法	P 值
总数	45	50		
性别, N(%)			Chisq. (1 df) = 0.07	0.785
女	15 (33.3)	18 (36)		
男	30 (66.7)	32 (64)		
年龄组, N(%)			Chisq. (1 df) = 0.25	0.615
<60	14 (31.1)	18 (36)		
>=60	31 (68.9)	32 (64)		
肿瘤部位, N(%)			Chisq. (1 df) = 0.21	0.643
左侧	29 (72.5)	34 (68)		
右侧	11 (27.5)	16 (32)		
检查的结节数量, N(%)			Fisher's exact test	0.043
<12	8 (17.8)	2 (4)		
>=12	37 (82.2)	48 (96)		
正节点数, N(%)			Fisher's exact test	0.79
>=4	3 (8.6)	4 (8.9)		
0	23 (65.7)	26 (57.8)		
1-3	9 (25.7)	15 (33.3)		
状态, N(%)			Chisq. (1 df) = 0.29	0.588
II	20 (44.4)	25 (50)		
III	25 (55.6)	25 (50)		
肿瘤级别, N(%)			Fisher's exact test	0.748
高	3 (6.7)	1 (2)		
低	4 (8.9)	6 (12)		
中	7 (15.6)	8 (16)		
未知	31 (68.9)	35 (70)		

[0112]

[0113] 1.2候选标志物选择

[0114] 本实施例1采用微阵列数据集 (GEO) 的元分析,使用训练和验证机器学习算法分析

了与结直肠癌结果和风险相关的差异表达基因。总共使用了六个数据集,如下表2所列。

[0115] 表2数据集列表

	研究主题	样本大小	复发数量	基因数量	
[0116]	GSE17536	转移基因表达谱预测结肠癌患者的复发和死亡(Moffitt 样本)	177	36	19320
	GSE17537	转移基因表达谱预测结肠癌患者的复发和死亡(VMC 样本)	55	6	19320
	GSE17538	实验衍生的转移基因表达谱预测结肠癌患者的复发和死亡	232	42	19320
	GSE4526	基因表达特征与 III 期结直肠癌复发的预测	36	15	19320
	GSE18105	II 期和 III 期结肠直肠癌	111	18	19320
	GSE26906	APC 结肠 II 期	90	21	19320

[0117] 首先使用三个数据集GSE18105、GSE26906和GSE4526,将基因特征的数量从19320减少到4548;

[0118] 使用Cox回归和LASSO正则化方法,结合数据集GSE17536和GSE17537对4548个基因进行分层;使用bagging (B=100)改进LASSO回归;

[0119] 然后使用GSE17536、GSE17537和GSE17538对分层后的基因进行重新训练,鉴定出二十九个与结直肠癌复发风险和结果密切相关的基因,对二十二个基因进行了验证,以确定它们在甲醛固定石蜡包埋组织(CRC FFPE)中通过q-PCR进行分析扩增的能力,还包括与复发风险相关的六个肠干细胞基因;然后,将十二个Oncotype Dx基因也纳入发现小组作为对照,总共有四十个基因被纳入qPCR验证小组。四十个基因为IRS2、TRIM25、MET、GSTM3、KRT6B、PKIB、VLDLR、ADAMTS5、ANAPC5、DPT、FAM111A、NADKD1、NAT1、WNT5A、FOXN3、CCL20、KLK6、DCBLD2、PKIG、DCTN5、SAMHD1、C5orf4、BGN、FAP、INHBA、Mki67、MYC、MYBL2、GADD45B、ATP5E、GPX1、PGK1、VDAC2、UBB、BCL2、FNML2、IGFBP4、SLC03A1、CXCR4和VEGFA。

[0120] 样本RNA提取及qPCR检测

[0121] 提取甲醛固定石蜡包埋组织中的mRNA,并进行实时荧光定量PCR检测,对具有错配修复功能的结直肠癌FFPE组织切片用脱蜡溶液处理。使用Deparaffinization Solution (Qiagen,19093)以及RNeasy\_FFPE Kit (Qiagen,73504)对甲醛固定石蜡包埋组织进行处理,以保持对后续qPCR分析有用的mRNA。使用40个基因的特异性引物将mRNA转换为cDNA。然后,将cDNA进行15个循环的预扩增,并将产物稀释至1:5用于qPCR检测、分析。

[0122] 一些实施例中,qPCR体系如下表3。

[0123] 表3 qPCR体系组成列表

	组成	体积, $\mu\text{L}$
	2×iTaQ Universal SYBR Green Supermix	10
	Primer F(10 $\mu\text{M}$ )	0.5
[0124]	Primer R(10 $\mu\text{M}$ )	0.5
	Template	2
	Nuclease-free H <sub>2</sub> O	7
	Volume	20

[0125] qPCR扩增反应程序如下表4。

[0126] 表4 qPCR扩增反应程序

步骤	温度, $^{\circ}\text{C}$	时间	循环数
Pre- Denaturation	95	3min	1
Denaturation	95	30s	40
Annealing/Extension (plate read)	60	40s	
Melt curve analysis	Use instrument default setting		

[0128] 1.4统计分析

[0129] 所有统计分析均使用R软件(版本4.1.2)进行,包括Bioconductor包(版本3.5)和R-Studio包(版本2022.02.0)。所有测试均为双侧对数秩检验;P值<0.05且单变量风险比>2被认为具有统计学意义。

[0130] 四十个基因与参照基因配对,并使用每个上调基因与下调基因之间的Ct值差异进行标准化。对所有基因对的总和组合进行详尽搜索,以通过测试假设来确定最佳基因对面板:对于II期和III期患者队列,无复发患者比例在低表达差异组中将显著高于高表达差异组。ROC曲线上的最佳截断点通过使用Youden指数,比较II期和III期患者队列中复发患者和非复发患者的面板RS得分确定。根据最佳截断面板得分定义的高风险组和低风险组,通过Kaplan-Meier估计法评估仅手术患者的三年复发情况。

[0131] 结果

[0132] 2.1患者人口统计学特征和特点:

[0133] 本实施例1关注95名具有错配修复功能的II期和III期结直肠癌患者。其中,有45名复发和50名未复发的结直肠癌患者。对于II期结直肠癌,有20名复发和25名未复发的患者。对于III期结直肠癌,有25名复发和25名未复发的患者。患者的人口统计信息和临床特征在表1中描述,没有显著差异。

[0134] 与CRC风险评估相关的差异表达基因

[0135] 从微阵列数据集分析和甲醛固定石蜡包埋验证小组中鉴定与CRC风险评估相关的差异表达基因,总共从微阵列数据集元分析中确定了29个基因,它们是ADAMTS5、ANAPC5、

C5orf4、CCL20、CRNDE、DCBLD2、DNER、DPT、FAM111A、GSTM3、IRS2、KLK6、KRT6B、MET、NADKD1、NAT1、NMU、NOB1、PKIB、PKIG、PTPLA、TPBG、TRIM25、VLDLR、WNT5A、ZFYVE27、DCTN5、FOXN3和SAMHD1。

[0136] 测试这29种基因在结直肠癌甲醛固定石蜡包埋组织中能否被检测到的能力,发现其中22个基因具有可检测的水平,并被纳入小组。另外与结直肠癌复发相关的6个肠干细胞基因,以及作为对照纳入小组的Oncotype Dx基因,最终选定的基因小组包括IRS2、TRIM25、MET、GSTM3、KRT6B、PKIB、VLDLR、ADAMTS5、ANAPC5、DPT、FAM111A、NADKD1、NAT1、WNT5A、FOXN3、CCL20、KLK6、DCBLD2、PKIG、DCTN5、SAMHD1、C5orf4、BGN、FAP、INHBA、Mki67、MYC、MYBL2、GADD45B、ATP5E、GPX1、PGK1、VDAC2、UBB、BCL2、FNML2、IGFBP4、SLC03A1、CXCR4和VEGFA。对这些基因在每个FFPE样本中的表达进行了单独评估,所有表达数据都以FFPE组织中表达的C(t)值表示。

[0137] 进一步,选定能够预测结直肠癌预后临床结果的生物标志物,例如IRS2、TRIM25、MYBL2、KRT6B、FMNL2、INHBA、MYC、NAT1和FAP等,以便进行精确的测量和分析。

[0138] 采用2个内参基因与基因之间的关系进行单变量分析,并计算配对基因的 $\Delta C_t$ 值的p值、AUC(曲线下面积)和风险比(HR),筛选出能够区分复发与非复发情况的基因对。选择标准如下表5。

[0139] 表5 单变量分析选择标准

	p 值	AUC	HR
[0140]	<0.05	>0.6	1.6

[0141] 最终确定15个基因对符合以上标准,见下表6。

[0142] 表6单变量分析选择结果列表

基因对	p 值	AUC	HR
IRS2-TRIM25	0.001	0.692	1.7
IRS2-MYBL2	0.001	0.692	1.7
IRS2-FMNL2	0.004	0.672	1.7
KRT6B-TRIM25	0.029	0.63	1.9
KRT6B-ANAPC5	0.011	0.652	1.9
KRT6B-NAT1	0.023	0.636	1.7
KRT6B-BCL2	0.026	0.633	1.8
KRT6B-FMNL2	0.041	0.622	1.7
KRT6B-IGFBP4	0.023	0.636	1.7
FAP-TRIM25	0.002	0.684	1.7
FAP-MYC	0.002	0.681	2.1
FAP-SLC03A1	0.003	0.679	1.7
INHBA-TRIM25	0.002	0.683	1.8
INHBA-MYC	0.03	0.629	1.7
INHBA-SLC03A1	0.001	0.693	1.7

[0144] 基于以上15个基因对,计算所有可能的组合,在II期和III期患者队列中最终挑选出风险比最高的生物标志物组合,见表7。

[0145] 表7 II期和III期队列中风险比最高的生物标志物组合

	p 值	AUC	HR	p 值	Feature
II 期	0.011	0.722	2.5	0.0095	IRS2-TRIM25, IRS2-MYBL2 KRT6B-FMNL2, INHBA-MYC
III 期	0	0.805	4.2	0.00007	IRS2-TRIM25, KRT6B-NAT1 FAP-MYC, INHBA-MYC

[0147] 2.3构建和验证II期和III期结直肠癌风险评估/复发组合

[0148] 根据基因表达差异对标志物进行配对,作为预测结直肠癌风险的生物标志物,并进行详尽搜索以对风险预测值最大化。

[0149] 对于II期结直肠癌,基于Youden指数的2.589作为参考值,基因表达差异组合:

[0150] (IRS2-TRIM25) + (IRS2-MYBL2) + (KRT6B-FMNL2) + (INHBA-MYC)

[0151] 上述CRC风险评估公式可以对II期结直肠癌复发风险做出最有效预测,如表8所列,对II期结直肠癌的预测风险比(HR)为2.5。

[0152] 对于III期结直肠癌,基于Youden指数的-0.1356为参考值,基因表达差异组合:

[0153] (IRS2-TRIM25)+(KRT6B-NAT1)+(FAP-MYC)+(INHBA-MYC)

[0154] 上述CRC风险评估公式可以对结直肠癌复发风险做出最有效预测,如表8所列,对III期结直肠癌复发的预测风险比为4.2。

[0155] 若采用现有技术中的Oncotype Dx组合,从本发明实施例1的qPCR数据中得出的II期和III期结直肠癌两个阶段的风险比(HR)为1.6。

[0156] 表8 mProbe和Oncotype Dx的复发组合性能

			mProbe 面板			Oncotype Dx 复发值		
	复发	未复发	P	AUC	HR	P	AUC	HR
[0157] 总数	45	50	<0.001	0.766	3.4	0.071	0.608	1.6
II 期	20	25	0.011	0.722	2.5	0.117	0.638	1.6
III 期	25	25	<0.001	0.805	4.2	0.361	0.424	1.6

[0158] 表8中,mProbe II期基因组合为 IRS2-TRIM25、IRS2-MYBL2、KRT6B-FMNL2、INHBA-MYC;mProbe III期基因组合为 IRS2-TRIM25、KRT6B-NAT1、FAP-MYC、INHBA-MYC;Oncotype Dx基因组合为FAP、GADD45B、MKi67、MYBL2、MYC、BGN、INHBA;内参基因为GPX1、PGK1、UBB、ATP5E、VDAC2。

[0159] 图1~图3为Kaplan-Meier估计单独手术患者的三年复发率,根据mProbe RS评分分类风险组观察结肠癌患者的复发情况,其中,图1为全队列患者,图2为II期患者,图3为III期患者;

[0160] 图4~图6为mProbe复发评分(RS)组的无复发间隔,其中,图4为所有患者,图5为II期患者,图6为IIIA/B期患者;

[0161] 图7~图9为Kaplan-Meier通过使用Oncotype Dx和评分系统估计单独手术患者风险组的三年复发率,其中,图7为全队列患者,图8为II期患者,图9为III期结肠癌患者。

[0162] 图10~图12为根据Oncotype Dx复发评分(RS)组的无复发间隔,其中,图10为所有患者,图11为 II期患者,图12为 IIIA/B期患者。

[0163] 根据II期和III期结直肠癌基因组合和RS评分系统对结直肠癌患者进行分类,对应的结直肠癌患者所处阶段的参考值设定为2.589和-0.1356。如图1和图2所示,基于配对基因的基因表达差异的qPCR分析,随后基于RS评分系统,能够准确预测并区分仅手术患者的三年复发高风险和低风险人群。

[0164] 以Oncotype Dx组合和公式作为对照进行平行分析,如图3和图4所示。结果显示与Oncotype Dx的结果相似,风险比(HR)约为1.38,本实施例1中观察到的风险比为1.6。数据表明,本实施例1的结直肠癌风险评估和复发的小组和RS模型的性能远远优于现有技术中的Oncotype Dx组合和模型。

[0165] 本发明实施例公开的预测II期和III期结直肠癌预后临床结果的生物标志物组、

引物组及检测试剂盒,并通过检测结直肠癌生物标志物水平可以得到结直肠癌RS评分,能够有效区分结直肠癌复发风险的高低;可以更好地评估结直肠癌患者的治疗风险和预后,从而为患者提供更加个性化和有效的治疗方案。

[0166] 本发明实施例公开的技术方案和实施例中公开的技术细节,仅是示例性说明本发明的发明构思,并不构成对本发明实施例技术方案的限定,凡是对本发明实施例公开的技术细节所做的常规改变、替换或组合等,都与本发明具有相同的发明构思,都在本发明权利要求要求的保护范围之内。

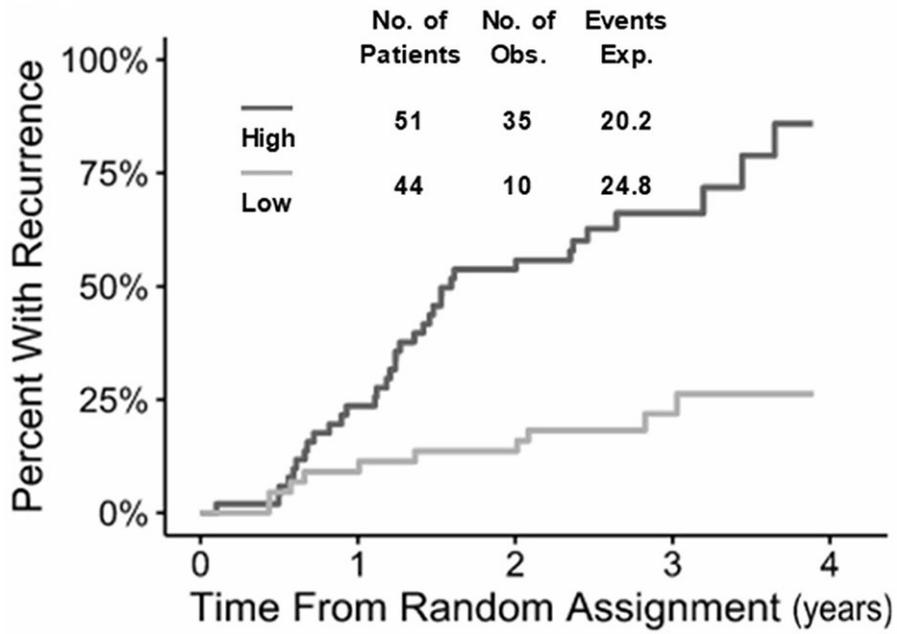


图 1

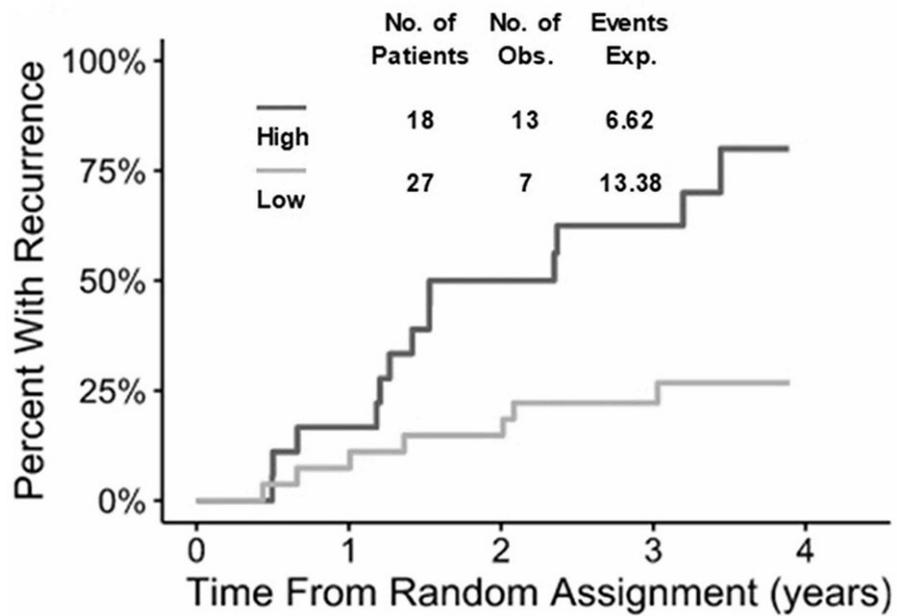


图 2

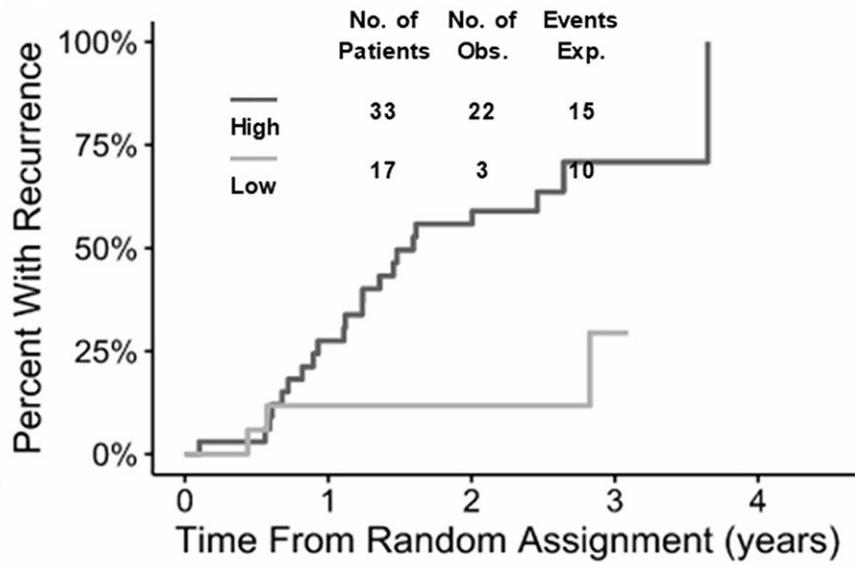


图 3

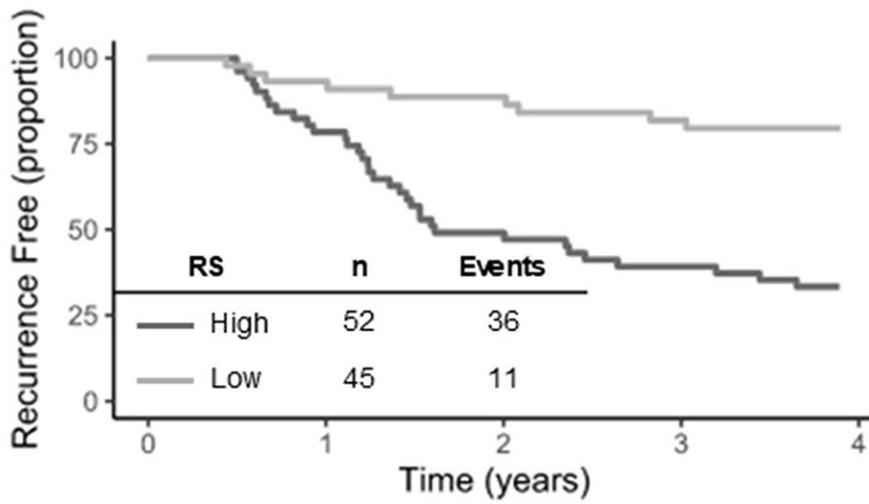


图 4

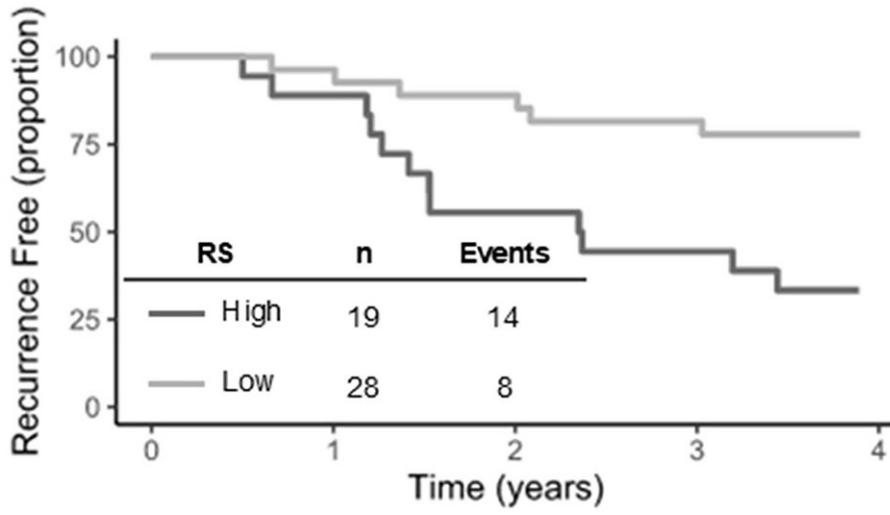


图 5

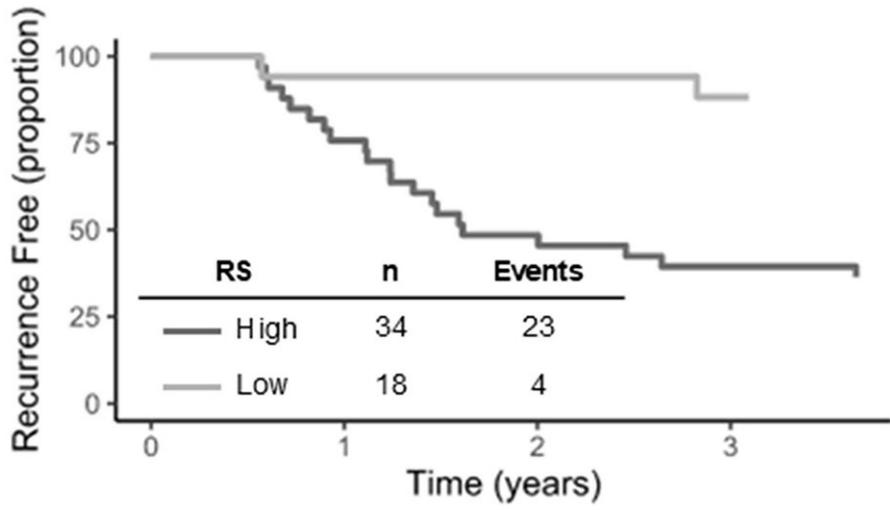


图 6

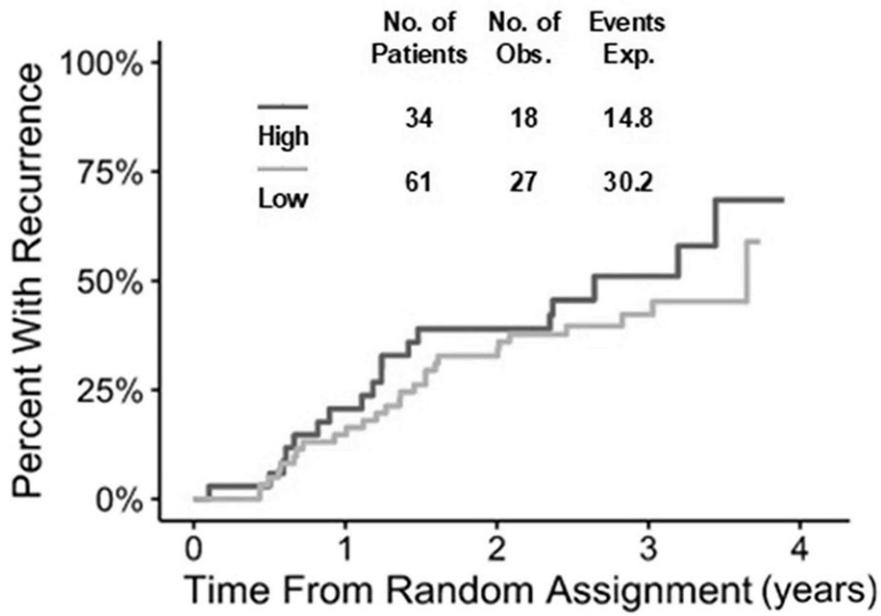


图 7

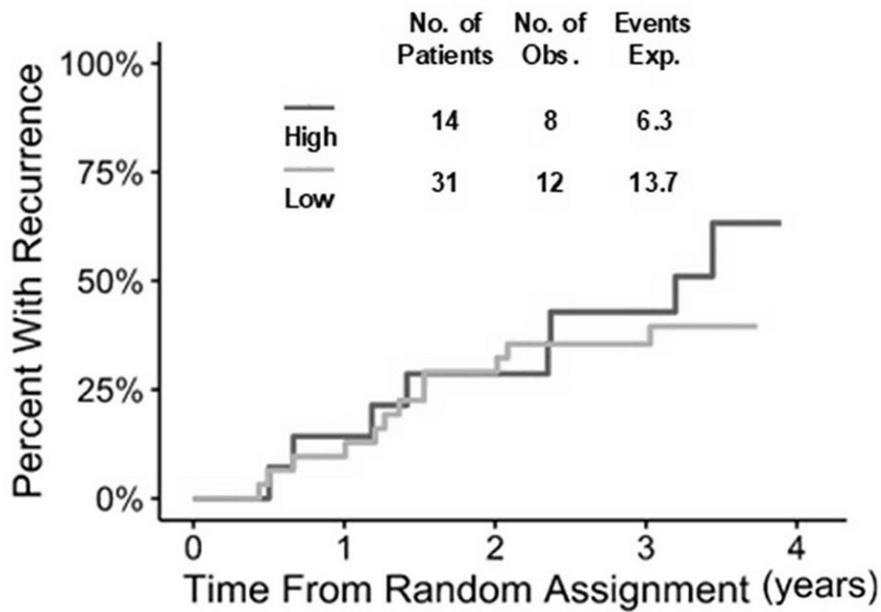


图 8

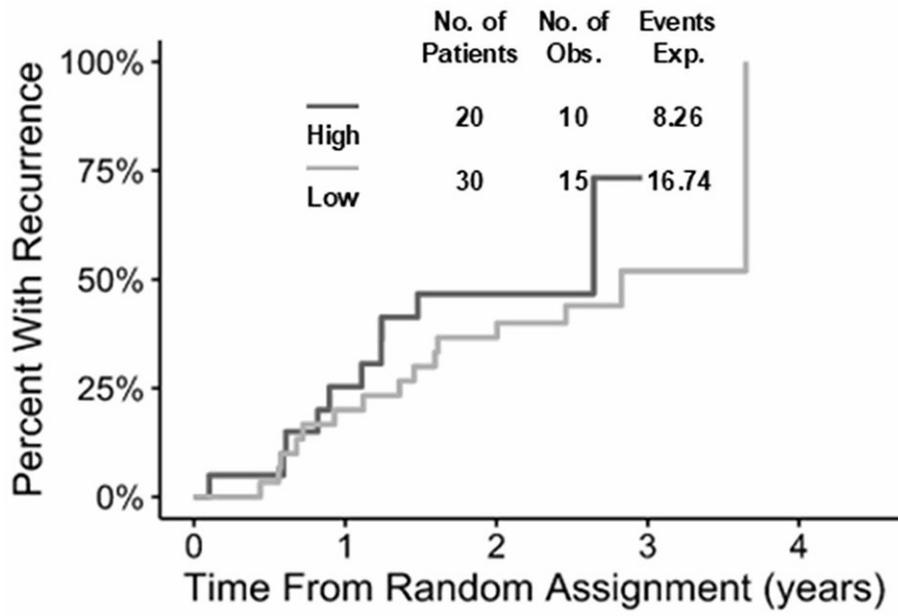


图 9

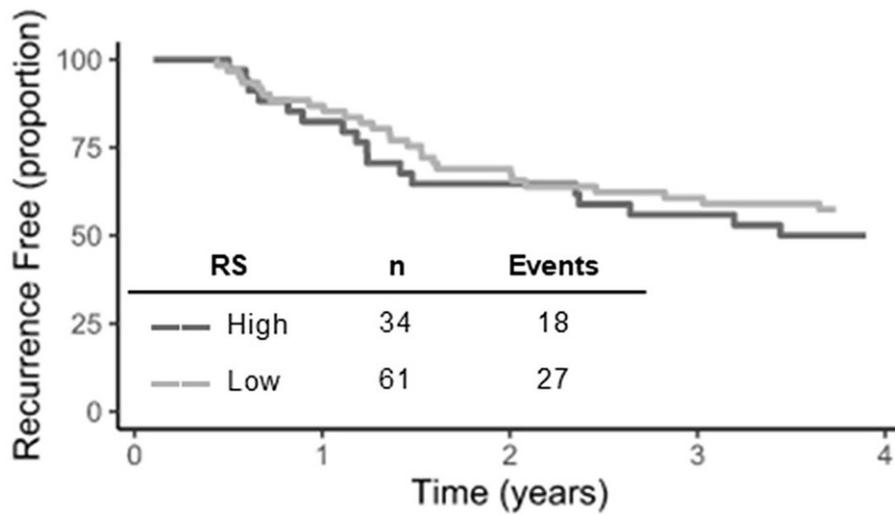


图 10

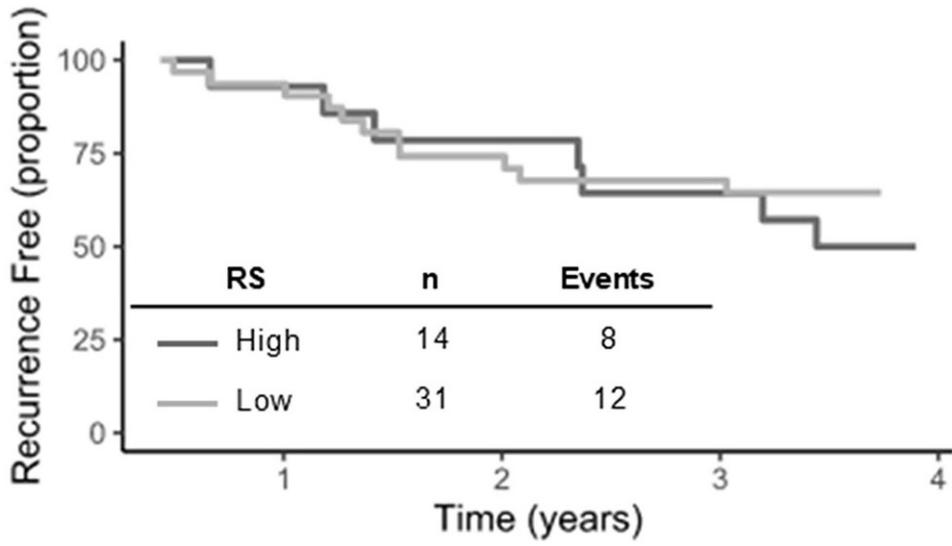


图 11

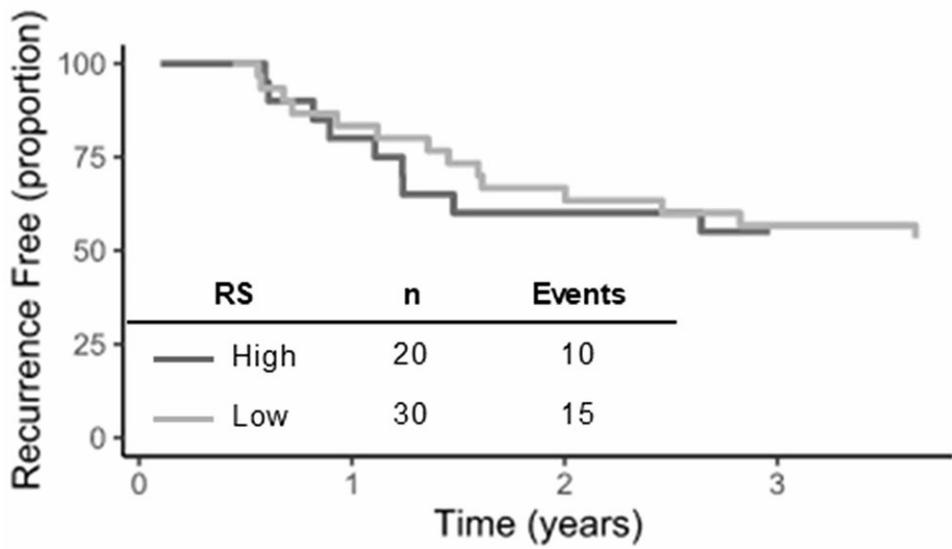


图 12