



(12)发明专利申请

(10)申请公布号 CN 109863401 A

(43)申请公布日 2019.06.07

(21)申请号 201780016684.6

(22)申请日 2017.03.17

(30)优先权数据

62/309,766 2016.03.17 US

(85)PCT国际申请进入国家阶段日

2018.09.11

(86)PCT国际申请的申请数据

PCT/US2017/022853 2017.03.17

(87)PCT国际申请的公布数据

W02017/161215 EN 2017.09.21

(71)申请人 长庚大学

地址 中国台湾桃园市龟山区文化一路259号

申请人 长庚医疗财团法人林口长庚纪念医院

(72)发明人 余兆松 陈怡婷 蒋维凡 萧永晋
史丽珠 张玉生 张凯评

(74)专利代理机构 中科专利商标代理有限责任公司 11021

代理人 崔亚松

(51)Int.Cl.

G01N 33/574(2006.01)

G01N 33/53(2006.01)

C07K 14/00(2006.01)

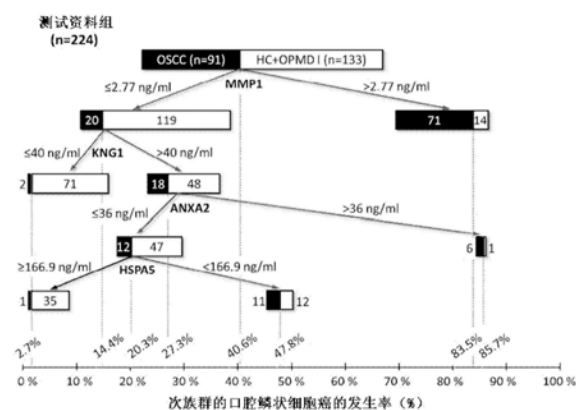
权利要求书3页 说明书33页
序列表7页 附图5页

(54)发明名称

一种用以诊断及预断癌症的方法

(57)摘要

本揭示内容是关于一种用以决定一个个体是否罹患癌症或有罹患癌症的风险的方法。该方法包含：由该个体取得一检体；决定至少二标的多肽的含量，该至少二标的多肽选自由ANXA2、HSPA5、KNG1及MMP1所组成的群组；以及基于该至少二标的多肽的含量来评估该个体是否罹患癌症或有罹患癌症的风险。本发明方法可用于诊断及预断口腔鳞状细胞癌的发生，据此及时对有需要的个体授予适当的治疗。



1. 一种用以决定一个体是否罹患口腔鳞状细胞癌或有罹患口腔鳞状细胞癌的风险的方法,包含:

(a) 由该个体取得一检体;

(b) 决定该检体中至少二标的多肽的含量,其中该至少二标的多肽是选自由ANXA2、HSPA5、KNG1及MMP1所组成的群组;

(c) 基于步骤(b)决定的该至少二标的多肽的含量来计算一风险指数;以及

(d) 基于步骤(c)的风险指数来决定该个体是否罹患口腔鳞状细胞癌或有罹患口腔鳞状细胞癌的风险。

2. 如权利要求1所述的方法,其中是利用逻辑式回归来计算该风险指数。

3. 如权利要求2所述的方法,其中是利用以下公式来计算该风险指数:

$$\text{风险指数} = \frac{e^{a+b1X1+b2X2+b3X3+b4X4}}{1 + e^{a+b1X1+b2X2+b3X3+b4X4}}$$

其中e为一数学常数,其为自然对数的基数;a为一常数值;X1、X2、X3及X4分别代表ANXA2、HSPA5、KNG1及MMP1的浓度;且b1、b2、b3及b4分别代表ANXA2、HSPA5、KNG1及MMP1的变异系数。

4. 如权利要求3所述的方法,其中

当该风险指数低于0.4时,则该个体未罹患口腔鳞状细胞癌或为罹患口腔鳞状细胞癌的低风险族群;以及

当该风险指数等于或高于0.4时,则该个体罹患口腔鳞状细胞癌或为罹患口腔鳞状细胞癌的高风险族群。

5. 如权利要求1所述的方法,其中该检体为唾液。

6. 如权利要求1所述的方法,其中是利用一试验来决定该至少二标的多肽的含量,且该试验是选自由酵素结合免疫吸附分析试验、利用条带进行的快速检测、西方墨点法、质谱分析法、蛋白微阵列、流式细胞仪分析、免疫荧光分析、免疫组织化学分析及多重探针监测试验所组成的群组。

7. 如权利要求6所述的方法,其中是利用液相层析串联质谱法以多反应监测模式来决定该至少二标的多肽的含量。

8. 一种用以诊断及治疗一罹患口腔鳞状细胞癌的个体的方法,包含:

(a) 由该个体取得一检体;

(b) 决定该检体中至少二标的多肽的含量,其中该至少二标的多肽是选自由ANXA2、HSPA5、KNG1及MMP1所组成的群组;

(c) 基于步骤(b)决定的该至少二标的多肽的含量来计算一风险指数;以及

(d) 当步骤(c)决定的风险指数等于或高于0.4时,则对该个体授予一有效量的抗癌治疗。

9. 如权利要求8所述的方法,其中该抗癌治疗是手术移除口腔鳞状细胞。

10. 如权利要求8所述的方法,其中是利用逻辑式回归来计算该风险指数。

11. 如权利要求10所述的方法,其中是利用以下公式来计算该风险指数:

$$\text{风险指数} = \frac{e^{a+b1X1+b2X2+b3X3+b4X4}}{1 + e^{a+b1X1+b2X2+b3X3+b4X4}}$$

其中e为一数学常数,其为自然对数的基数;a为一常数值;X1、X2、X3及X4分别代表ANXA2、HSPA5、KNG1及MMP1的浓度;且b1、b2、b3及b4分别代表ANXA2、HSPA5、KNG1及MMP1的变异系数。

12. 如权利要求8所述的方法,其中该检体为唾液。

13. 如权利要求8所述的方法,其中是利用一试验来决定该至少二标的多肽的含量,且该试验是选自由酵素结合免疫吸附分析试验、利用条带进行的快速检测、西方墨点法、质谱分析法、蛋白微阵列、流式细胞仪分析、免疫荧光分析、免疫组织化学分析及多重探针监测试验所组成的群组。

14. 如权利要求13所述的方法,其中是利用液相层析串联质谱法以多反应监测模式来决定该至少二标的多肽的含量。

15. 一种用以决定一生物检体是否包含癌化口腔鳞状细胞的方法,包含:

(a) 决定该生物检体中至少二标的多肽的含量,其中该至少二标的多肽是选自由ANXA2、HSPA5、KNG1及MMP1所组成的群组;

(b) 基于步骤(a)决定的该至少二标的多肽的含量来计算一风险指数;以及

(c) 基于步骤(b)的风险指数来评估该生物检体是否包含癌化口腔鳞状细胞。

16. 如权利要求15所述的方法,其中是利用逻辑式回归来计算该风险指数。

17. 如权利要求16所述的方法,其中是利用以下公式来计算该风险指数:

$$\text{风险指数} = \frac{e^{a+b1X1+b2X2+b3X3+b4X4}}{1 + e^{a+b1X1+b2X2+b3X3+b4X4}}$$

其中e为一数学常数,其为自然对数的基数;a为一常数值;X1、X2、X3及X4分别代表ANXA2、HSPA5、KNG1及MMP1的浓度;且b1、b2、b3及b4分别代表ANXA2、HSPA5、KNG1及MMP1的变异系数。

18. 权利要求15所述的方法,其中当该风险指数等于或高于0.4时,则该生物检体包含癌化口腔鳞状细胞。

19. 如权利要求15所述的方法,其中该检体为唾液。

20. 如权利要求15所述的方法,其中是利用一试验来决定该至少二标的多肽的含量,且该试验是选自由酵素结合免疫吸附分析试验、利用条带进行的快速检测、西方墨点法、质谱分析法、蛋白微阵列、流式细胞仪分析、免疫荧光分析、免疫组织化学分析及多重探针监测试验所组成的群组。

21. 如权利要求20所述的方法,其中是利用液相层析串联质谱法以多反应监测模式来决定该至少二标的多肽的含量。

22. 一种用以决定一个体是否罹患口腔鳞状细胞癌或有罹患口腔鳞状细胞癌的风险的药学套组,包含至少二试剂,用以决定该个体体内至少二标的多肽的含量,其中该至少二标的多肽是选自由ANXA2、HSPA5、KNG1及MMP1所组成的群组。

23. 如权利要求22所述的药学套组,其中该至少二试剂中的各试剂皆为一以同位速标记的多肽,且包含选自由序列编号:5、6、7及8所组成的群组的胺基酸序列。

24. 一种如权利要求22所述的药学套组的用途,其用以决定一个体是否罹患口腔鳞状

细胞癌或有罹患口腔鳞状细胞癌的风险。

25. 如权利要求24所述的用途,其中该至少二试剂中的各试剂皆为一以同位速标记的多肽,且包含选自由序列编号:5、6、7及8所组成的群组的胺基酸序列。

26. 如权利要求24所述的用途,其中该至少二试剂可用以决定该个体体内至少二标的多肽的含量,藉以计算一风险指数。

27. 如权利要求26所述的用途,其中是利用逻辑式回归来计算该风险指数。

28. 如权利要求27所述的用途,其中是利用以下公式来计算该风险指数:

$$\text{风险指数} = \frac{e^{a+b1X1+b2X2+b3X3+b4X4}}{1 + e^{a+b1X1+b2X2+b3X3+b4X4}}$$

其中e为一数学常数,其为自然对数的基数;a为一常数值;X1、X2、X3及X4分别代表ANXA2、HSPA5、KNG1及MMP1的浓度;且b1、b2、b3及b4分别代表ANXA2、HSPA5、KNG1及MMP1的变异系数。

29. 如权利要求25所述的用途,其中是利用一试验来决定该至少二标的多肽的含量,且该试验是选自由酵素结合免疫吸附分析试验、利用条带进行的快速检测、西方墨点法、质谱分析法、蛋白微阵列、流式细胞仪分析、免疫荧光分析、免疫组织化学分析及多重探针监测试验所组成的群组。

30. 如权利要求29所述的用途,其中是利用液相层析串联质谱法以多反应监测模式来决定该至少二标的多肽的含量。

31. 如权利要求26所述的用途,其中

当该风险指数低于0.4时,则该个体未罹患口腔鳞状细胞癌或为罹患口腔鳞状细胞癌的低风险族群;以及

当该风险指数等于或高于0.4时,则该个体罹患口腔鳞状细胞癌或为罹患口腔鳞状细胞癌的高风险族群。

一种用以诊断及预断癌症的方法

[0001] 相关申请的交叉引用

[0002] 本申请涉及并要求2016年3月7日提交的美国临时申请第62/309,766号的权益,该申请的内容通过引用整体并入本文。

技术领域

[0003] 本揭示内容是关于一用以侦测口腔鳞状细胞癌(oral squamous cell carcinoma,OSCC)的多标志分析组;更具体来说,本揭示内容是关于以四蛋白分析组、三蛋白分析组或二蛋白分析组藉由唾液检体来侦测口腔鳞状细胞癌及监测罹患口腔恶性潜能病变(oral potentially malignant disorder,OPMD)的病患的方法。

背景技术

[0004] 口腔癌是一种全球性的好发癌症,于世界各地造成日益严重的疾病问题。欧美及亚洲地区最常见的口腔癌发生位置为舌头及颊部区域。于2012年,估计约有300,400名新增的口腔癌病患,而约有145,400名病患因口腔癌而死亡。口腔癌发生率最高的地区为美拉尼西亚、中南亚及中东欧(每100,000人中约有9.1-22.9人)。超过三分之一的新增个案及半数的死亡人数皆发生于开发中国家。然而,若在分析西方欧洲国家的口腔癌发生率时一并将年龄列入考量因素,可发现西方罹患口腔癌的个体数亦逐渐攀升,显示在过去二十年间口腔癌呈现稳定增加的趋势。口腔鳞状细胞癌是一种最常见的口腔癌,估计占口腔癌总个数的90%以上。造成口腔鳞状细胞癌的主要致病因素包含吸烟、酗酒、非吸烟性的烟草使用及咀嚼槟榔。即使可藉由外科手术及管理等方法予以处理,多数国家口腔鳞状细胞癌的5年存活率仍约为50%。该统计数据显示超过60%的病患为第III或IV期的病患,且相较于其他癌症,口腔鳞状细胞癌较易造成第二原位性肿瘤的发生。癌症确诊时的分期是决定个体5年存活率的关键因素,第I期病患的5年存活率约为80%,相较之下,晚期病患的5年存活率则会显著下降。因此,相关领域亟需一种可早期诊断口腔鳞状细胞癌的方法。

[0005] 口腔鳞状细胞癌多发生于口腔的可见病变,且伴随着口腔上皮异常。世界卫生组织将该病变称为口腔恶性潜能病变。目前已确认超过20种口腔恶性潜能病变。相较于薄型同质性白斑症(thin homogeneous leukoplakia)及扁平苔癣(lichen planus),粘膜红斑(erythroplakia)、粘膜下纤维化(submucous fibrosis)、异质性白斑症(heterogeneous leukoplakia)及疣状增生(verrucous hyperplasia)等病变具有更高的恶性转形机率。口腔恶性潜能病变于各国的恶性转形率约为0.13%到17.5%。在台湾病患中,不同组织形态的总恶性转形率为4.32%,且恶性转形的平均发生时间为33.56个月。其中,上皮异常(24.4%)及疣状增生(20%)具有更高的转形率。口腔恶性潜能病变恶性转形为口腔鳞状细胞癌的过程相当缓慢,且几乎无法观察,造成病患无法及时注意,进而延误口腔鳞状细胞癌的诊断。此外,许多口腔恶性潜能病变的病灶包含潜在恶性细胞、尚未侵袭他处的恶性细胞、已侵袭他处的恶性细胞及正常细胞的混合性组织。该混合性组织反应了区域性的癌化现象,造成临床医疗人员在面对相同病灶时产生不同的判断,且使检体的诊断过程过于复

杂化。病理学家曾因此错误诊断病患的病情。基于该些因素,临床医疗人员往往不易诊断出早期的口腔鳞状细胞癌,因此医疗相关领域亟需可用以确认高风险的口腔恶性潜能病变及/或监测该病变恶性转形的方法。

[0006] 由于口腔鳞状细胞癌多由可观察的口腔恶性潜能病变转形而成,通常可藉由观察口腔粘膜及病理检验异常组织的检体,来侦测口腔鳞状细胞癌。最近一项于印度进行的观察性筛选口腔鳞状细胞癌或口腔恶性潜能病变的随机分配对照试验(191,873位参与者,15年观察期后确认553位罹患口腔鳞状细胞癌,且6749位罹患口腔恶性潜能病变)指出,口腔观察可减少接触烟草及酒精的病患的死亡率。过去二十年间,口腔鳞状细胞癌于台湾的发生率正逐渐提高;在1996到2009年之间,男性每年罹患口腔鳞状细胞癌的比例为24.64/100,000(将年龄列入考量因素后),该比例为世界最高的比例之一。2010后,中国台湾政府已推动台湾口腔癌筛检计画(Taiwan's Oral Cancer Screening Program),提供高风险族群(30岁以上有咀嚼槟榔或吸烟习惯的个体)每年免费的口腔检查。每年约有一百万参与者参与筛检活动,包含请临床医生或牙医进行口腔检查,转诊进行病理确认及后续的治疗(口腔癌筛检临床路径)。然而,2011到2012的筛检结果指出,相较于常规临床诊断,筛检仅能提升约3%的早期性(即第I期)口腔鳞状细胞癌的诊断率(表1)。该结果并不意外,原因在于第一线健康工作者通常不易判断何种口腔病变应转诊进行组织确认。此外,早期口腔鳞状细胞癌多与某些良性或发炎病状相似,且会与多种口腔恶性潜能病变共同存在;在此情况下,癌化病灶的分布或弥漫性粘膜下纤维化往往使临床医疗人员无法藉由组织切片来准确判断癌细胞。因此,相关领域亟需一种非侵入式的临床检测方法,用以有效检测包埋于口腔恶性潜能病变病灶的癌细胞。

[0007] 表1.台湾于2011-2012年间,经由检察筛选计画及非筛检的常规诊断所确认的口腔鳞状细胞癌病例

[0008]

由口腔粘膜筛检所确认的口腔鳞状细胞癌病例 ^a					
分期	0-1	2	3	4	全部
病例数	1553	942	521	1605	4621
(%)	(33.61)	(20.39)	(11.27)	(34.73)	(100)
由非筛检的常规诊断所确认的口腔鳞状细胞癌病例					
分期	0-1	2	3	4	全部
病例数	1130	695	449	1429	3703
(%)	(30.52)	(18.77)	(12.12)	(38.59)	(100)

[0009] ^a1,850,697位具有罹患口腔鳞状细胞癌风险的个体参与筛检。

[0010] 近年来,已有多种非侵入式的候选生物标志可用以检测口腔鳞状细胞癌。然而,多数的候选生物标志皆未经详细的分析评估,或利用适当的体液检体进行定量平行比对,据以确认何种候选生物标志可进一步于大规模检体族群进行临床验证。上述原因也部分解释了为何至今仍无分子生物标志获得政府健康机关的许可,可用以早期诊断及/或处理口腔鳞状细胞癌。有鉴于此,相关领域亟需一种新颖的生物标志,可用以预断及/或诊断口腔鳞状细胞癌,进而及时提供有需要的个体适当的治疗。

发明内容

[0011] 发明内容旨在提供本揭示内容的简化摘要,以使阅读者对本揭示内容具备基本的理解。此发明内容并非本揭示内容的完整概述,且其用意并非在指出本发明实施例的重要/关键元件或界定本发明的范围。

[0012] 本揭示内容的第一态样是关于一种用以决定一个体是否罹患口腔鳞状细胞癌或有罹患口腔鳞状细胞癌的风险的方法。该方法包含:

[0013] (a) 由该个体取得一检体;

[0014] (b) 决定该检体中至少二标的多肽的含量,其中该至少二标的多肽是选自由膜联蛋白A2(annexin A2,ANXA2)、热休克蛋白A5(heat shock protein A5,HSPA5)、激肽原-1(kininogen-1,KNG1)及基质金属蛋白酶-1(matrix metalloproteinase-1,MMP1)所组成的群组;

[0015] (c) 基于步骤(b)决定的该至少二标的多肽的含量来计算一风险指数(risk score);以及

[0016] (d) 基于步骤(c)的风险指数来决定该个体是否罹患口腔鳞状细胞癌或有罹患口腔鳞状细胞癌的风险。

[0017] 依据本揭示内容某些实施方式,是利用逻辑式回归(logistic regression)来计算风险指数。较佳地,是利用以下公式来计算风险指数:

$$[0018] \quad \text{风险指数} = \frac{e^{a+b_1X_1+b_2X_2+b_3X_3+b_4X_4}}{1 + e^{a+b_1X_1+b_2X_2+b_3X_3+b_4X_4}}$$

[0019] 其中e为一数学常数,其为自然对数的基数;a为一常数值;X1、X2、X3及X4分别代表ANXA2、HSPA5、KNG1及MMP1的浓度;且b1、b2、b3及b4分别代表ANXA2、HSPA5、KNG1及MMP1的变异系数。

[0020] 依据本揭示内容的实施方式,在步骤(d)中,当该风险指数低于0.4时,则该个体未罹患口腔鳞状细胞癌或为罹患口腔鳞状细胞癌的低风险族群;而当该风险指数等于或高于0.4时,则该个体罹患口腔鳞状细胞癌或为罹患口腔鳞状细胞癌的高风险族群。临床医疗人员可对风险指数等于或高于0.4的个体及时安排适当的病理检查及/或投予抗癌治疗(例如预防性治疗或治疗性治疗)。

[0021] 一般来说,该个体为一哺乳动物;较佳是一人类。依据本揭示内容的实施方式,该检体为唾液。

[0022] 本揭示内容的第二态样是关于一种用以决定一生物检体是否包含癌化口腔鳞状细胞的方法。本发明方法包含:

[0023] (a) 决定该生物检体中至少二标的多肽的含量,其中该至少二标的多肽是选自由ANXA2、HSPA5、KNG1及MMP1所组成的群组;

[0024] (b) 基于步骤(a)决定的该至少二标的多肽的含量来计算一风险指数;以及

[0025] (c) 基于步骤(b)的风险指数来评估该生物检体是否包含癌化口腔鳞状细胞。

[0026] 依据本揭示内容某些实施方式,是利用逻辑式回归来计算该风险指数。较佳地,是利用以下公式来计算该风险指数:

$$[0027] \quad \text{风险指数} = \frac{e^{a+b_1X_1+b_2X_2+b_3X_3+b_4X_4}}{1 + e^{a+b_1X_1+b_2X_2+b_3X_3+b_4X_4}}$$

[0028] 其中 e 为一数学常数,其为自然对数的基数; a 为一常数值; X_1 、 X_2 、 X_3 及 X_4 分别代表ANXA2、HSPA5、KNG1及MMP1的浓度;且 b_1 、 b_2 、 b_3 及 b_4 分别代表ANXA2、HSPA5、KNG1及MMP1的变异系数。

[0029] 依据本揭示内容一实施方式,当该风险指数等于或高于0.4时,则该生物检体包含癌化口腔鳞状细胞。

[0030] 依据本揭示内容某些实施方式,该检体为唾液。

[0031] 本揭示内容亦提供一种药组,以及其于诊断口腔鳞状细胞癌或评估口腔鳞状细胞癌发生风险的用途。本发明药组包含至少二试剂,用以决定该个体体内至少二标的多肽的含量,其中该至少二标的多肽是选自由ANXA2、HSPA5、KNG1及MMP1所组成的群组。依据本揭示内容一实施例,该至少二试剂中的各试剂皆为一以同位素标记的多肽,且包含选自序列编号:5、6、7及8所组成的组群的氨基酸序列。

[0032] 基于定量结果所产生的风险指数可作为口腔鳞状细胞癌指标。依据本揭示内容的实施方式,当该风险指数低于0.4时,则该个体未罹患口腔鳞状细胞癌或为罹患口腔鳞状细胞癌的低风险族群;而当该风险指数等于或高于0.4时,则该个体罹患口腔鳞状细胞癌或为罹患口腔鳞状细胞癌的高风险族群。

[0033] 例示性的可用以决定该至少二标的多肽的试验包含,但不限于,酵素结合免疫吸附分析试验(enzyme-linked immunosorbent assay,ELISA)、利用条带进行的快速检测(strip-based rapid test)、西方墨点法(western blotting)、质谱分析法(mass spectrometry)、蛋白微阵列(protein microarray)、流式细胞仪分析(flow cytometry)、免疫荧光分析(immunofluorescence)、免疫组织化学分析(immunohistochemistry)及多重探针监测试验(multiplex detection assay)。在本揭示内容一特定实施例中,是利用液相层析串联质谱法以多反应监测模式(liquid chromatography-tandem mass spectrometry with multiple reaction monitoring (MRM) mode,LC-MRM-MS)来决定该至少二标的多肽的含量。

[0034] 在参阅下文实施方式后,本发明所属技术领域中具有通常知识者当可轻易了解本发明的基本精神及其他发明目的,以及本发明所采用的技术手段与实施态样。

附图说明

[0035] 为了让本发明的上述与其他目的、特征、优点与实施例能更明显易懂,附图的说明如下:

[0036] 图1A的分群树状图是关于筛选的四种蛋白及其于各分截点的浓度截止值(cut-off value,单位为纳克/毫升);

[0037] 图1B的二维点状图是关于测试资料组(training set, $n=224$)的健康对照个体、口腔低恶性潜能病变病患(OPMD I)及口腔鳞状细胞癌病患(OSCC)的风险指数;

[0038] 图1C的二维点状图是关于试验组(test set, $n=106$)的健康对照个体、口腔低恶性潜能病变病患(OPMD I)及口腔鳞状细胞癌病患(OSCC)的风险指数;

[0039] 图1D阐述了测试资料组及试验组的曲线下面积(area under the curve,AUC)、灵敏度及专一性;

[0040] 图2的二维点状图阐述了相对于非口腔鳞状细胞癌族群(健康对照+口腔低恶性潜

能病变(OPMD),n=199),利用四蛋白分析组计算第I到IV期的口腔鳞状细胞癌病患(n分别为50、29、16及36)的风险指数;以及

[0041] 图3的二维点状图阐述了相对于非口腔鳞状细胞癌族群(健康对照+口腔低恶性潜能病变(OPMD I),n=199)及口腔鳞状细胞癌族群(OSCC,n=131),利用四蛋白分析组计算口腔高恶性潜能病变族群(OPMD II,n=130)的风险指数。

具体实施方式

[0042] 为了使本揭示内容的叙述更加详尽与完备,下文针对了本发明的实施态样与具体实施例提出了说明性的描述;但这并非实施或运用本发明具体实施例的唯一形式。实施方式中涵盖了多个具体实施例的特征以及用以建构与操作这些具体实施例的方法步骤与其顺序。然而,亦可利用其他具体实施例来达成相同或均等的功能与步骤顺序。

[0043] 虽然用以界定本发明较广范围的数值范围与参数皆是约略的数值,此处已尽可能精确地呈现具体实施例中的相关数值。然而,任何数值本质上不可避免地含有因个别测试方法所致的标准偏差。在此处,“约”通常是指实际数值在一特定数值或范围的正负10%、5%、1%或0.5%之内。或者是,“约”一词代表实际数值落在平均值的可接受标准误差之内,视本发明所属技术领域中具有通常知识者的考量而定。除了实验例之外,或除非另有明确的说明,当可理解此处所用的所有范围、数量、数值与百分比(例如用以描述材料用量、时间长短、温度、操作条件、数量比例及其他相似者)均经过“约”的修饰。因此,除非另有相反的说明,本说明书与附随申请专利范围所揭示的数值参数皆为约略的数值,且可视需求而更动。至少应将这些数值参数理解为所指出的有效位数与套用一般进位法所得到的数值。在此处,将数值范围表示成由一端点至另一端点或介于二端点之间;除非另有说明,此处所述的数值范围皆包含端点。

[0044] 除非本说明书另有定义,此处所用的科学与技术词汇的含义与本发明所属技术领域中具有通常知识者所理解与惯用的意义相同。此外,在不和上下文冲突的情形下,本说明书所用的单数名词涵盖该名词的复数型;而所用的复数名词时亦涵盖该名词的单数型。

[0045] “胺基酸序列相同度百分比”(Percentage(%) amino acid sequence identity)一词在本说明书是指在将一候选胜肽序列(candidate sequence)与一特定胜肽序列进行比对后,二序列间胺基酸残基的相同百分比;在比对时,是将二序列排比,且若有必要,加入适当的间隙(gap),藉以达到最大的相同度百分比,且不将保留性置换(conservative substitution)视为序列相同的一部份。可利用各式相关领域习知技艺人士所熟知的方法来进行比对,并决定二序列的相同百分比;举例来说,可以利用诸如BLAST、BLAST-2、ALIGN或Megalign(DNASTAR)等电脑软体来进行比对分析。相关领域习知技艺人士可使用适当的参数(包含各式演算法)来测量比对,以于胺基酸全长序列中产生最大的比对序列。在本说明书中,二胺基酸序列间的序列比对是利用国家生物技术讯息中心(Nation Center for Biotechnology Information,NCBI)线上提供的Blastp(蛋白-蛋白BLAST)电脑软体来进行分析。具体来说,一特定胺基酸序列A与一特定胺基酸序列B之间的胺基酸序列相同百分比(亦可表述为一特定胺基酸序列A与一特定胺基酸序列B具有某百分比的胺基酸序列相同度)是利用下列公式来计算:

[0046] $\frac{X}{Y} \times 100\%$

[0047] 其中X是经比对程式BLAST评比后,A与B序列间相互吻合的胺基酸残基数量,Y则是A或B序列中(取较短者)胺基酸残基的总数量。

[0048] “接受者操作特征曲线”(receiver operating characteristic curve,ROC curve)一词于本揭示内容中是指一利用真阳性率(true positive rate)与伪阳性率(false positive rate)绘出的曲线,用以决定一预断试验的截止值(cut-off value)。ROC曲线通常将(1-专一性)定义为x轴,将灵敏度定义为y轴。一高灵敏度的数值表示低伪阳性率。同理,一高专一性的数值亦表示低伪阳性率。“截止值”(cut-off value)一词于此是指一由ROC曲线所得的数值,用以表示该预断试验的敏感度及专一性达到平衡。一切点范围可以包含数个切点实施方式,其中各切点分别代表不同的灵敏度及其专一性间的平衡。

[0049] 在揭示内容中,“曲线下面积”(area under the curve,AUC)为相关领域习知技艺人士所知的名词,定义为ROC曲线下的面积。一AUC数值通常介于0.5-1.0之间,用以代表一预断试验的准确度;其中,AUC的数值愈高,表示该预断试验的预断效果愈佳。AUC数值通常会伴随95%信赖区间(信赖区间,CI),其是指一具有特定概率的统计范围,以使一设定参数可落于该范围内。

[0050] 在本揭示内容中,“评估”(assessing)一词是指决定一个体的健康状况的过程。该个体的健康状况可指诊断或预断该个体罹患癌症,或是具有罹患癌症的风险。

[0051] “风险”(risk)一词于此是指一项可能会产生不良结果的可能性,例如,口腔鳞状细胞癌的发生、进展或复发。依据一个体、其检体或其相关事件的分析结果,该名个体可被归类至“高度风险”(high risk)、或“低度风险”(low risk)族群。本揭示内容所述的风险指数是指,当一病患的风险指数 ≥ 0.4 时,则该名病患会归类为“高度风险”族群;亦即,他/她五年内罹患口腔鳞状细胞癌的可能性高于分析族群中其余的病患。当一病患的风险指数 < 0.4 时,则该名病患会归类为“低度风险”族群;亦即,他/她五年内罹患肝癌的可能性低于分析族群中其余的病患。

[0052] 在本揭示内容中,“预防性治疗”(prophylactic treatment或preventive treatment)是指防止或抑制一临床疾病或病状的发生,或是延缓一临床疾病或病状的早期临床病征的发生;举例来说,口腔鳞状细胞癌。依据本揭示内容的实施方式,“预防性治疗”一词是指对一有罹患口腔鳞状细胞癌倾向的个体投予防止性治疗。一般来说,罹患癌症倾向可能是由遗传因子、年龄、性别及损伤等因素所造成。

[0053] 在本揭示内容中,“治疗性治疗”(therapeutic treatment)是指对一已罹患一疾病(例如口腔鳞状细胞癌)的个体投予治疗,藉以产生治疗功效,例如减缓既有病征、减缓病征的潜在代谢原因、延缓或防止病状进一步的发生及/或减少病征可能或预期会产生的严重程度。

[0054] “个体”(subject)一词是指包含人类的动物,其可接受本发明方法的治疗。除非特定指出,否则“个体”(subject)一词同时意指不同年龄的男性及女性,例如孩童或成人。

[0055] 本揭示内容的第一态样是关于一种用以决定一个体是否罹患口腔鳞状细胞癌或有罹患口腔鳞状细胞癌的风险的方法。依据本揭示内容的实施方式,该方法包含以下步骤:

[0056] (a) 由该个体取得一检体;

[0057] (b) 决定该检体中至少二标的多肽的含量, 其中该至少二标的多肽是选自由 ANXA2、HSPA5、KNG1 及 MMP1 所组成的群组;

[0058] (c) 基于步骤 (b) 决定的该至少二标的多肽的含量来计算一风险指数; 以及

[0059] (d) 基于步骤 (c) 的风险指数来决定该个体是否罹患口腔鳞状细胞癌或有罹患口腔鳞状细胞癌的风险。

[0060] 在步骤 (a) 中, 先由个体取得一检体。该个体为一哺乳动物; 较佳为一人类。依据本揭示内容较佳的实施例, 该个体为一亚洲人。在本揭示内容一操作实施例中, 该个体为一华人。依据本揭示内容的实施方式, 该检体较佳为唾液。

[0061] 步骤 (b) 是用以决定检体中 ANXA2、HSPA5、KNG1 及 MMP1 中至少二种 (例如 ANXA2、HSPA5、KNG1 及 MMP1 任二种、三种或四种) 多肽的含量。依据本揭示内容某些实施方式, 是对 ANXA2、HSPA5、KNG1 及 MMP1 中任二种多肽进行定量分析 (可以是相对定量值或绝对定量值), 藉以产生可用以诊断或预断癌症的二标志分析组。该二标志分析组可以由 (1) ANXA2 及 HSPA5 多肽, (2) ANXA2 及 KNG1 多肽, (3) ANXA2 及 MMP1 多肽, (4) HSPA5 及 KNG1 多肽, (5) HSPA5 及 MMP1 多肽, 或是 (6) KNG1 及 MMP1 多肽所组成。依据本揭示内容某些实施方式, 是对 ANXA2、HSPA5、KNG1 及 MMP1 中任三种多肽进行定量分析 (可以是相对定量值或绝对定量值), 藉以产生可用以诊断或预断癌症的三标志分析组。该三标志分析组可以由 (1) ANXA2、HSPA5 及 KNG1 多肽, (2) ANXA2、HSPA5 及 MMP1 多肽, (3) ANXA2、KNG1 及 MMP1 多肽, 或是 (4) HSPA5、KNG1 及 MMP1 多肽所组成。依据本揭示内容其他实施方式, 是对 ANXA2、HSPA5、KNG1 及 MMP1 四种多肽进行定量分析 (可以是相对定量值或绝对定量值), 藉以产生四标志分析组。

[0062] 一般来说, 可利用习知技艺人士所熟知的方法来决定 ANXA2、HSPA5、KNG1 及/或 MMP1 的含量; 举例来说, 酵素结合免疫吸附分析试验、利用条带进行的快速检测、西方墨点法、质谱分析法、蛋白微阵列、流式细胞仪分析、免疫荧光分析、免疫组织化学分析或多重探针监测试验。依据本揭示内容一实施方式, 是利用液相层析串联质谱法以多反应监测模式 (一种蛋白质体学领域常用来专一且准确定量分析多肽的方法) 来决定 ANXA2、HSPA5、KNG1 及/或 MMP1 的含量。

[0063] 依据本揭示内容某些实施方式, 标的多肽 ANXA2 包含与序列编号: 1 具有至少 90% (即 90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 或 100%) 相似度的胺基酸序列; 标的多肽 HSPA5 包含与序列编号: 2 具有至少 90% 相似度的胺基酸序列; 标的多肽 KNG1 包含与序列编号: 3 具有至少 90% 相似度的胺基酸序列; 而标的多肽 MMP1 则包含与序列编号: 4 具有至少 90% 相似度的胺基酸序列。依据本揭示内容一操作实施例, 标的多肽 ANXA2 具有序列编号: 1 的胺基酸序列; 标的多肽 HSPA5 具有序列编号: 2 的胺基酸序列; 标的多肽 KNG1 具有序列编号: 3 的胺基酸序列; 而标的多肽 MMP1 具有序列编号: 4 的胺基酸序列。

[0064] 在步骤 (c) 中, 利用步骤 (b) 得到的二、三或四标志分析组来计算可作为风险指数的预期或然率。依据本揭示内容某些实施方式, 是利用逻辑式回归来分析二、三或四标志分析组, 以得到风险指数。依据本揭示内容较佳的实施方式, 是利用以下公式来计算风险指数:

$$[0065] \quad \text{风险指数} = \frac{e^{a+b1X1+b2X2+b3X3+b4X4}}{1 + e^{a+b1X1+b2X2+b3X3+b4X4}}$$

[0066] 其中 e 为一数学常数, 其为自然对数的基数; a 为一常数值; X1、X2、X3 及 X4 分别代表

ANXA2、HSPA5、KNG1及MMP1的浓度；且b1、b2、b3及b4分别代表ANXA2、HSPA5、KNG1及MMP1的变异系数。

[0067] 依据本揭示内容一操作实施例，公式的常数值及变异系数会随着标志分析组的不同而改变，可依据表12-13或表16-17的公式来计算特定标的多肽的风险指数。

[0068] 依据本揭示内容一实施方式，是利用包含二种标的多肽的二标志分析组来计算风险指数，其中该二种标的多肽是选自由ANXA2、HSPA5、KNG1及MMP1所组成的群组。依据本揭示内容另一实施方式，是利用包含三种标的多肽的三标志分析组来计算风险指数，其中该三种标的多肽是选自由ANXA2、HSPA5、KNG1及MMP1所组成的群组。依据本揭示内容较佳的实施方式，是利用包含ANXA2、HSPA5、KNG1及MMP1四种标的多肽的四标志分析组来计算风险指数。

[0069] 在步骤(d)中，是利用步骤(c)计算的风险指数来评估一个体是否罹患口腔鳞状细胞癌或有罹患口腔鳞状细胞癌的风险。依据本揭示内容某些实施方式，风险指数可用以区分非口腔鳞状细胞癌个体(例如健康个体或口腔恶性潜能病变病患)与口腔鳞状细胞癌病患。在某些实施方式中，当风险指数等于或高于0.4(≥ 0.4)时，代表该个体罹患口腔鳞状细胞癌(阳性预测值(positive predictive value, PPV)为75.5%-89.1%)；相反地，当风险指数低于0.4(< 0.4)时，则代表该个体未罹患口腔鳞状细胞癌(阴性预测值(negative predictive value, NPV)为81.9%-93.6%)；用以区别非口腔鳞状细胞癌个体及口腔鳞状细胞癌病患的准确度为80.9%-86.7%。依据一操作实施例，风险指数是与口腔鳞状细胞癌的分期数相关，其中相较于晚期口腔鳞状细胞癌，罹患早期口腔鳞状细胞癌的病患具有较低的风险指数。依据本揭示内容其他实施方式，风险指数可用以评估口腔鳞状细胞癌于口腔恶性潜能病变(例如口腔低或高恶性潜能病变)病患的发生风险。在有些实施方式中，当风险指数等于或高于0.4(≥ 0.4)时，该病患为罹患口腔鳞状细胞癌的高风险族群(转形率为37.8%)；或者是，当风险指数低于0.4(< 0.4)时，则该病患为罹患口腔鳞状细胞癌的低风险族群(转形率为7.8%)。

[0070] 临床医疗人员可依据本发明方法所得的风险指数来及时诊断并对有需要的个体投予治疗，其中应对风险指数等于或高于0.4的个体投予一抗癌治疗(例如一预防性治疗或一治疗性治疗)或安排其进行后续的追踪检查。

[0071] 本揭示内容的第二态样是关于一种用以诊断及治疗一罹患口腔鳞状细胞癌的个体的方法。该方法包含利用上述方法的步骤(a)到(c)来决定一个体是否罹患口腔鳞状细胞癌，接着对风险指数等于或高于0.4的个体投予一有效量的抗癌治疗。一般来说，抗癌治疗可以是一预防性治疗(例如投予抗氧化剂)、一治疗性治疗(例如化疗、手术切除、放射线治疗及免疫治疗)或其组合。较佳地，抗癌治疗为手术切除口腔鳞状细胞癌。

[0072] 本揭示内容的第三态样是关于一种用以决定一生物检体是否包含癌化口腔鳞状细胞的方法。该方法包含：

[0073] (a) 决定该生物检体中至少二标的多肽的含量，其中该至少二标的多肽是选自由ANXA2、HSPA5、KNG1及MMP1所组成的群组；

[0074] (b) 基于步骤(a)决定的该至少二标的多肽的含量来计算一风险指数；以及

[0075] (c) 基于步骤(b)的风险指数来评估该生物检体是否包含癌化口腔鳞状细胞。

[0076] 用以评估生物检体的方法(即第三态样的方法)的步骤(a)到(b)分别与上述用以

评估由个体取得的检体的方法(即第一态样的方法)的步骤(b)到(c)相同,因此,为求精简,在此省略其相关叙述。

[0077] 在步骤(c)中,是利用步骤(b)计算的风险指数来分析生物检体。依据本揭示内容一实施方式,该生物检体为一口腔鳞状细胞癌检体;在该实施方式中,生物检体的风险指数等于或高于0.6。依据本揭示内容另一实施方式,该生物检体包含癌化口腔鳞状细胞;举例来说,一由口腔高恶性潜能病变病患分离的检体,其中检体中的癌化细胞无法以传统方法(例如活体组织切片)来侦测,或是未来(5年内)有转形成异常病灶的可能性;在该实施方式中,生物检体的风险指数等于或高于0.4,且低于0.6。

[0078] 依据本揭示内容某些实施方式,该检体为唾液。

[0079] 本揭示内容亦提供一种药学套组,可用以决定一个体是否罹患口腔鳞状细胞癌或有罹患口腔鳞状细胞癌的风险。本发明药学套组包含至少二试剂(例如二种、三种或四种试剂),可用以决定个体体内ANXA2、HSPA5、KNG1及MMP1中至少二多肽(例如ANXA2、HSPA5、KNG1及MMP1中任二种、三种或四种多肽)的含量。举例来说,本发明药学套组可包含二试剂,分别用以定量分析ANXA2、HSPA5、KNG1及MMP1中任二种多肽的含量。或者是,本发明药学套组可包含三试剂,分别用以定量分析ANXA2、HSPA5、KNG1及MMP1中任三种多肽的含量。具选择性地,本发明药学套组可包含四试剂,分别用以定量分析ANXA2、HSPA5、KNG1及MMP1的含量。

[0080] 依据使用需求的不同,各试剂可以是一多肽(例如抗体或以同位素标记的多肽)或一适体(aptamer)。依据本揭示内容一实施例,各试剂皆为一以同位素标记的多肽,其中用以定量分析ANXA2(序列编号:1)、HSPA5(序列编号:2)、KNG1(序列编号:3)及MMP1(序列编号:4)的试剂分别包含序列编号:5、6、7及8的胺基酸序列。

[0081] 可依据试剂类型的不同来决定分析ANXA2、HSPA5、KNG1及/或MMP1含量的方法。依据本揭示内容一实施方式,各试剂皆为一以同位素标记的多肽,且是以LC-MRM-MS来定量分析ANXA2、HSPA5、KNG1及/或MMP1含量。

[0082] ANXA2、HSPA5、KNG1及/或MMP1的定量结果可用以计算风险指数,藉以诊断口腔鳞状细胞癌或评估其发生风险。如上所述,可利用逻辑式回归来计算风险指数;较佳地,可利用以下公式来计算风险指数:

$$[0083] \quad \text{风险指数} = \frac{e^{a+b_1X_1+b_2X_2+b_3X_3+b_4X_4}}{1 + e^{a+b_1X_1+b_2X_2+b_3X_3+b_4X_4}}$$

[0084] 其中e为一数学常数,其为自然对数的基数;a为一常数值;X₁、X₂、X₃及X₄分别代表ANXA2、HSPA5、KNG1及MMP1的浓度;且b₁、b₂、b₃及b₄分别代表ANXA2、HSPA5、KNG1及MMP1的变异系数。

[0085] 依据本揭示内容一实施例,公式的常数值及变异系数会随着标志分析组的不同而改变,可依据表12-13或表16-17的公式来计算特定标的多肽的风险指数。

[0086] 依据本揭示内容的实施方式,当风险指数低于0.4时,该个体未罹患口腔鳞状细胞癌,或为罹患口腔鳞状细胞癌的低风险族群;而当风险指数等于或高于0.4时,则该个体罹患口腔鳞状细胞癌,或为罹患口腔鳞状细胞癌的高风险族群。

[0087] 下文提出多个实验例来说明本发明的某些态样,以利本发明所属技术领域中具有通常知识者实作本发明,且不应将这些实验例视为对本发明范围的限制。据信习知技艺者在阅读了此处提出的说明后,可在不需过度解读的情形下,完整利用并实践本发明。此处所

引用的所有公开文献,其全文皆视为本说明书的一部分。

[0088] 实施例

[0089] 材料及方法

[0090] 样本

[0091] 在收集预处理的唾液检体前,确保参与的个体皆有签署奇美医学中心人体试验委员会许可的知情同意书,同意本发明使用其唾液检体。由96位健康对照个体(具有正常黏膜)、103位具有口腔低恶性潜能病变(OPMD I)病患、130位具有口腔高恶性潜能病变(OPMD II)病患及131位罹患口腔鳞状细胞癌的病患收集唾液检体。由奇美医学中心(中国台湾,柳营)取得2008-2013年的检体(表2)。所有的个体皆有参与台湾口腔癌筛检计画。利用活体组织切片来诊断确认口腔鳞状细胞癌,并依据标准流程使病患接受常规检查。依据已知方法来分类口腔恶性潜能病变病例。将130位口腔高恶性潜能病变的个案分为:粘膜红白斑(erythroleukoplakia, $n=6, 4.6\%$)、具有高度口腔粘膜下纤维化的粘膜红斑(erythroplakia plus high-grade oral submucous fibrosis (OSF), $n=1, 0.8\%$)、异质性白斑症(heterogeneous leukoplakia, $n=5, 3.8\%$)、具有高度OSF的白斑症($n=7, 5.4\%$)、高度OSF($n=21, 16.2\%$)、斑点白斑症(speckle leukoplakia, $n=7, 5.4\%$)、疣状白斑症(verrucous leukoplakia, $n=1, 0.8\%$)、疣状增生(verrucous hyperplasia, $n=44, 33.8\%$)及具有OSF的疣状增生($n=38, 29.2\%$)。将103位口腔低恶性潜能病变的个案分为:白斑症($n=91, 88.3\%$)、苔藓样的变灶(lichenoid lesions, $n=4, 3.9\%$)及低度OSF($n=8, 7.8\%$)。

[0092] 表2. 参与实验的人口特征及其吸烟与咀嚼槟榔的情况

[0093]

	对照组	口腔低恶性潜能病变	口腔高恶性潜能病变	口腔鳞状细胞癌	p^*	全部
病例数量	96 (20.9%)	103 (22.4%)	130 (28.3%)	131 (28.5%)		460 (100.0%)
性别						
男性	96 (100.0%)	102 (99.0%)	129 (99.2%)	129 (98.5%)	0.6763^1	456 (99.1%)
女性	0 (0.0%)	1 (1.0%)	1 (0.8%)	2 (1.5%)		4 (0.9%)
年龄	48.75±11.84	49.49±10.71	51.36±10.51	52.51±9.65	0.0320^2	50.72±10.68
吸烟 (每天包数 × 年)	19.13±11.15	24.59±24.15	31.03±21.48	27.01±22.39	0.0030^2	25.96±21.09
槟榔 (每天颗数 × 年)	138.06±328.63	172.18±187.91	389.63±524.15	386.89±477.71	$<.0001^2$	287.66±430.67

[0094] ¹Fisher's正确检定

[0095] ²分析变异性 (ANOVA)

[0096] 组间的p值

[0097] 收集并处理唾液检体。简单来说,在口腔粘膜检查期间,收集未经刺激的全唾液。

确保供体于检体收集前至少1小时未进食、饮食、吸烟及使用口腔清洁产品。将各检体于4℃以3000×g的转速离心15分钟。对上清液投予蛋白酶抑制剂混合物(Sigma, St. Louis, MO, USA)后,将检体储存于-80℃。

[0098] 筛选标的蛋白的代理胜肽

[0099] 对每一标的蛋白筛选一代理胜肽。首先,由代表癌细胞株及初代细胞的分泌蛋白体(secretomes)以及口腔鳞状细胞癌的组织蛋白质体(tissue proteomes)的猎枪MS资料组(shotgun MS datasets)来挑选胜肽。接着挑选:(a)包含8到23个残基的特定胜肽,其中经由人类蛋白参考资料库(human protein reference database)确认该些残基不具有已知转译后修饰位置,且无连续或漏失的胰蛋白酶切割位点;(b)不具化学反应的胺基酸(例如半胱胺酸或甲硫胺酸)的胜肽;(c)不包含潜在会导致漏失切割的序列(例如RP或KP)的胜肽;以及(d)于MS2数据中具有高度辨识分数的胜肽。进一步利用MRMPilot软体(2.1版;AB Sciex, Forster City, CA, USA)来分析符合上述条件的胜肽,以预测其碎片离子是否可以MS进行侦测。对于无可用的实质证据且在猎枪MS资料组中无适合的胜肽的四种标的蛋白(DSG3、HGF、CRNN及TP53),是藉由生物资讯预测法来取得所有可能的胰蛋白酶胜肽,并利用上述条件来挑选其代理胜肽。

[0100] 胰蛋白酶分解及加入稳定同位素标记的标准化(stable isotope-labeled standard, SIS)胜肽

[0101] 利用LC-MRM-MS来分析经3重复处理的唾液检体3次。以BCA蛋白试验套组(Thermo Scientific Pierce, USA)来检测各唾液检体的蛋白浓度。将15微克的唾液蛋白溶解于15微升的25mM的碳酸氢铵(ammonium bicarbonate)中,再加入15微升的10%去氧胆酸钠(sodium deoxycholate, DOC)进行变性反应。以81.35微克的25mM碳酸氢铵稀释检体,并与12.4微升的50mM三(2-羧乙基)膦(Tris(2-carboxyethyl) phosphine, TCEP)于60℃反应30分钟进行还原反应后,再与13.75微升的100mM碘乙酰胺(iodoacetamide)于37℃作用30分钟进行烷基化反应。将改良的定序级胰蛋白酶(Promega, Madison, WI)以20:1蛋白/酵素的比例加至经还原及烷基化的检体中,并于37℃反应9小时以分解检体。中止胰蛋白酶分解反应,并将检体保存于-20℃以供后续研究分析使用。将49条SIS胜肽标准化混合物(参见后续内容)渗入各检体中,再以6微升的10%甲酸(formic acid)及1.5微升的10%三氟乙酸(trifluoroacetic acid, TFA)进行酸化反应以沉淀DOC。在渗入SIS胜肽标准化混合物时,是将等量的包含49[¹³C₆; ¹⁵N₂]离胺酸或[¹³C₆; ¹⁵N₄]精胺酸编码的SIS胜肽的SIS混合物加入分解物中,均匀混合。SIS胜肽是由UVic-Genome BC Proteomics Centre, BC Canada合成及纯化而得。ANXA2、HSPA5、KNG1及MMP1的SIS的纯度分别为93.4%、98.3%、98.2%及99.1%。表3阐述了特定标的多肽的SIS胜肽的序列及浓度,其中ANXA2、HSPA5、KNG1及MMP1的SIS胜肽分别为:“QDIAFAYQR”(序列编号:5)、“ITPSYVAFTPEGER”(序列编号:6)、“TVGSDTFYFSK”(序列编号:7)及“DIYSSFGFPR”(序列编号:8)。在SIS重胜肽中,将精胺酸(Arginine, R)及离胺酸(Lysine, K)的碳-12(¹²C)置换为碳-13(¹³C),将精胺酸及离胺酸的氮-14(¹⁴N)置换为氮-15(¹⁵N)。将酸化的检体于室温以16,000×g的速度离心2分钟,移除DOC后,将上清液冷冻于-20℃中以进行后续实验。依据使用操作方法并加入些微修改步骤,利用Waters Oasis HLB Elution Plate(Waters, MA)以固相萃取法来使检体去盐化并进行浓缩。简单来说,以乙腈(acetonitrile)润湿树脂后,利用平衡缓冲液(0.1%TFA及0.1%甲酸)进行平衡。加入

唾液蛋白分解物,经水洗涤后,以50微升的70%乙腈冲提2次。冷冻干燥冲提检体,再加入0.1%甲酸(体积比),使浓度调整为每微升0.25微升,以进行后续LC-MRM-MS分析。

[0102] LC-MRM-MS分析及取得数据

[0103] 利用nanoACQUITY UPLC系统(Waters,USA)来注入唾液胜肽。各检体的LC-MRM/MS分析(参见后续内容)耗费70分钟。将4微升检体(代表1微克的胜肽)注入解析管柱(nanoACQUITY UPLC C18,150微米×10毫米,1.7微米的颗粒大小;水),于97%缓冲液A(0.1%的溶于水的甲酸)(J.T.Baker,USA)及3%缓冲液B(0.1%的溶于乙腈的甲酸)(J.T.Baker)中,流速设定为每分钟1微升,持续10分钟。利用48分钟的线性浓度梯度(由3%到28%的缓冲液B)及5分钟的线性浓度梯度(由28%到38%的缓冲液B)及1分钟的线性浓度梯度(由38%到95%的缓冲液B)以每分钟400纳升的流速来分离检体。加入95%缓冲液B作用5分钟后,逐渐降至3%溶剂B(时间超过1分钟),再以3%缓冲液B重新平衡10分钟,以回复解析管柱。在各检体操作间注入空白溶剂(以每分钟400纳升的速度分析25分钟),以防止留存于UPLC管柱的检体相互干扰。

[0104] 利用具备由Analyst 1.5.1软体控制的纳米电喷雾游离源的AB/MDS Sciex 5500 QTRAP(all from AB Sciex,Singapore)来进行LC-MRM-MS分析。以下列参数进行侦测:离子喷雾电压为1900-2200伏特;气帘设定为20psi(UHP氮气);界面加热温度为150℃;MS操作压力为 3.5×10^{-5} Torr;Q1及Q3的单元分解(0.6-0.8Da,于半高的全宽度)。利用三组MRM离子对以下列条件针对各胜肽进行MRM侦测:碎片离子特定去团簇电位(fragment-ion-specific-tuned declustering potential,DP);进入电位(entrance potential,EP);碰撞能量(collision energy,CE);碰撞细胞输出电位(collision cell exit potential,CXP)及滞留时间。以1秒的标的循环时间及4分钟的MRM侦测窗来侦测所有的数据。在一次LC-MRM-MS运作中定量分析49条受测胜肽(对应于49种标的蛋白)的转移。Table 3总结了用以分析标的多肽ANXA2、HSPA5、KNG1及MMP1的MRM参数。

[0105] 表3.用以定量标的多肽的MRM参数

[0106]

蛋白	固定 SIS 胜肽浓度 (飞摩尔/每微克总蛋白)	胜肽	Q1/Q3 质量 (Da)	Q3 形式
ANXA2	5	QDIAFAYQR (序列编号: 5).1.轻	556.28 537.28	/ 2/y4
		QDIAFAYQR (序列编号: 5).2.轻	556.28 684.35	/ 2/y5
		QDIAFAYQR (序列编号: 5).3.轻	556.28 755.38	/ 2/y6
		QDIAFAYQR (序列编号: 5).1.重	561.28 547.29	/ 2/y4
		QDIAFAYQR (序列编号: 5).2.重	561.28 694.35	/ 2/y5
		QDIAFAYQR (序列编号: 5).3.重	561.28 765.39	/ 2/y6
HSPA5	5	ITPSYVAFTPEGER (序列编号: 6).1.轻	783.89 676.81	/ 2/y12(2+)
		ITPSYVAFTPEGER (序列编号: 6).2.轻	783.89 906.43	/ 2/y8
		ITPSYVAFTPEGER (序列编号: 6).3.轻	783.89 835.39	/ 2/y7
		ITPSYVAFTPEGER (序列编号: 6).1.重	788.9 681.83	/ 2/y12(2+)
		ITPSYVAFTPEGER (序列编号: 6).2.重	788.9 916.44	/ 2/y8
		ITPSYVAFTPEGER (序列编号: 6).3.重	788.9 / 845.4	2/y7
KNG1	5	TVGSDFYFSFK (序列编号: 7).1.轻	626.3 1051.47	/ 2/y9
		TVGSDFYFSFK (序列编号: 7).2.轻	626.3 792.39	/ 2/y6
		TVGSDFYFSFK (序列编号: 7).3.轻	626.3 907.42	/ 2/y7
		TVGSDFYFSFK (序列编号: 7).1.重	630.31 1059.49	/ 2/y9
		TVGSDFYFSFK (序列编号: 7).2.重	630.31 800.41	/ 2/y6
		TVGSDFYFSFK (序列编号: 7).3.重	630.31 915.43	/ 2/y7
MMP1	10	DIYSSFGFPR (序列编号: 8).1.轻	594.79 797.39	/ 2/y7
		DIYSSFGFPR (序列编号: 8).2.轻	594.79 960.46	/ 2/y8
		DIYSSFGFPR (序列编号: 8).3.轻	594.79 476.26	/ 2/y4
		DIYSSFGFPR (序列编号: 8).1.重	599.79 807.4	/ 2/y7
		DIYSSFGFPR (序列编号: 8).2.重	599.79	/ 2/y8

[0107]

	970.47	
DIYSSFGFPR (序列编号: 8).3.重	599.79	/
	486.27	2/y4

[0108] 将以底线标记的精氨酸(R)及离氨酸(K)的碳-12 (^{12}C)及氮-14 (^{14}N)分别置换为碳-13 (^{13}C)及氮-15 (^{15}N)。

[0109] MRM,数据分析及产生校正曲线

[0110] 利用MultiQuant软体(2.1版;AB Sciex)以MQ4公式计算峰积分(peak integration)来处理所有MRM数据。在侦测数据时,使用预定MRM来减少循环时间及于各高峰产生更多点,藉以确保定量结果更为准确。各标的胜肽皆有一条标准曲线,其是以不同量的胰蛋白酶分解一标准唾液检体所得。如临床检体的操作,收集并混合由三个体(二位口腔鳞状细胞癌病患及一位对照个体)收集的唾液,再进行胰蛋白酶分解,以制备标准唾液检体。将固定量的SIS胜肽渗入标准唾液检体后,产生11个点(空白,以及A到J)稀释曲线,其中SIS胜肽浓度为固定,并以适当稀释的胰肽蛋白酶分解物来调整轻胜肽的浓度。于各临床唾液检体中加入固定量的49-SIS-胜肽混合物。依据内源性唾液胜肽的浓度及讯号强度来调整SIS混合物的组成,以确保定量的准确度。加入相同浓度的标准唾液检体(检体I)作为未知组(注入1微克的内源性胜肽),而检体A、B、C、D、E、F、G、H及J则分别对应于0.00001-、0.0001-、0.005-、0.01-、0.05-、0.1-、0.2-、0.5-及2倍的标准检体(检体I)浓度。将固定量的49 SIS混合物渗入检体A到I,其中是将0.5倍稀释的SIS混合物加至检体J中。已知各SIS胜肽的确切浓度,因此可藉由观察高峰区域的比例来决定未知检体中蛋白的浓度。

[0111] 重复3次独立操作重复(由分解到最终的LC-MRM-MS步骤)以得到各唾液检体及校正曲线的浓度点。利用标准 $1/x$ (x =浓度比)权值选项(可涵盖广泛的动态范围)来得到所有校正曲线的线性回归。以三组MRM离子对来测所各胜肽;其中一组是作为计量计,另二组则是用以确认滞留时间及显示任何讯号干扰。手动检查所有峰积分以确保正确的高峰侦测及准确的积分。在统计分析时,将未侦测到高峰的蛋白浓度设定为0。由三次独立实验所得的浓度来计算各标的蛋白浓度的平均值,并以飞摩尔/每微克(fmol/ μg)及纳克/每毫升(ng/ml)来表示唾液蛋白;其中是假设胰蛋白酶完全分解及100%的胜肽回收,而由各原型胜肽的摩尔含量来得该些数值。

[0112] 统计分析

[0113] 在适当的情况下,利用Fisher's正确检定(Fisher's exact test)、单向变异分析(one-way analysis of variance,ANOVA)及/或无母数Mann-Whitney 检定(nonparametric Mann-Whitney test)比对四组(健康对照组、口腔低恶性潜能病变、口腔高恶性潜能病变及口腔鳞状细胞癌)的分类和连续数据。当ANOVA结果为显著时,利用Student-Newman-Keuls事后多重比对(Student-Newman-Keuls post-hoc multiple comparison)来确认不同的平均值。在筛选口腔鳞状细胞癌的各生物标志或生物标志组合时,利用接受者操作特征曲线(receiver operating characteristic (ROC) curve)、曲线下面积(the area under the ROC curve,AUC)、正概度比(the positive likelihood ratio,LR+)及负概度比(the negative likelihood ratio,LR-)来评估检定效果。以ROC曲线得到的最高Youden指数(定义为 $\text{Se}+\text{Sp}-1$)来决定最佳截止值,其中该ROC曲线符合平滑无

母数法 (smooth nonparametric method) 以减少数据所产生的筛选偏好。

[0114] 利用三种常用以区分非疾病及疾病状态的统计法来得到生物标志分析组:k-最近邻辨别法 (k-nearest neighbors discrimination)、逻辑式回归及分类及回归树 (classification and regression tree, CART)。CART包含建构分析树及修正分析树二步骤,再利用二进制递归区分 (binary recursive partitioning) 来拟合数据。在建构分析树时,当程序到达节点分割 (splitting node),重复为各可能的预测值选择一适当的分割点,直到产生错误分类的最低成本。以Hold-Out法将检体以与人口特征相似的2:1的比例随机分组。以Bootstrap法重复刺激1000次,挑选测试资料组 (n=224) 中最佳的分析树大小及最重要的预测因子,再以独立的试验组 (n=106) 来确认最终的分析树形式。在建构分析树后,将筛选标志的连续数值 (数值变数) 依据其分割点的截止浓度二分为二元变数。当ANXA2、KNG1及MMP1的浓度高于截止值,或是HSPA5的浓度低于截止值时,将口腔鳞状细胞癌的阳性预测值设定为1,否则将阴性预测值设定为0。将以CART得到的用以筛选标志的二元变数作为逻辑式回归的共变量,并以下列公式计算罹患口腔鳞状细胞癌的预测概率:

$$[0115] \quad \text{风险指数} = \frac{e^{a+b_1X_1+b_2X_2+b_3X_3+b_4X_4}}{1+e^{a+b_1X_1+b_2X_2+b_3X_3+b_4X_4}},$$

[0116] 其中e为一数学常数,其为自然对数的基数;a为一常数值;X1、X2、X3及X4分别代表ANXA2、HSPA5、KNG1及MMP1的浓度;且b1、b2、b3及b4分别代表ANXA2、HSPA5、KNG1及MMP1的变异系数。利用SAS来进行辨别及逻辑式回归,以R (v3.0.3) 统计软体rpart来进行CART。

[0117] 实施例1产生候选生物标志分析组

[0118] 1.1筛选用以侦测口腔鳞状细胞癌的生物标志

[0119] 由测试资料组 (n=224) 来筛选与口腔鳞状细胞癌相关的生物标志,并由试验组 (n=106) 进行确认;其中该测试资料组及该试验组是由330位受测者所组成,包含口腔鳞状细胞癌病患 (n=131) 及非口腔鳞状细胞癌个体 (n=199),并以与人口特征相似的2:1的比例随机分组 (结果未显示)。利用逻辑式回归、最近邻辨别法及分类及回归树 (CART) 来分析所有的数据 (表4)。将由CART分析筛选的四种标的蛋白作为生物标志分析组,以侦测口腔鳞状细胞癌,其中该四种标的蛋白包含ANXA2、HSPA5、KNG1及MMP1 (第1A图)。MMP1为第一生物标志,其最能区别口腔鳞状细胞癌个体与非口腔鳞状细胞癌个体,在截止值为每毫升2.77纳克时,其可正确确认91位口腔鳞状细胞癌病患中的71位 (78%),仅对14产生伪阳性结果。接着将唾液中MMP1含量低于每毫升2.77纳克的个体 (n=139) 依据KNG1>每毫升40纳克及ANXA2>每毫升36纳克进行分类;该分析可正确确认剩余20位口腔鳞状细胞癌病患中的6位,而仅对1位产生伪阳性结果。最后,将唾液中KNG1>每毫升40纳克,而ANXA2<每毫升36纳克的个体 (n=59) 再依据HSPA5<每毫升166.9纳克进行分析;该分析可正确确认11位口腔鳞状细胞癌个体。因此,在测试资料组中,公式可正确辨识91个口腔鳞状细胞癌检体中的88个检体,以及133个非口腔鳞状细胞癌检体中的106个检体,灵敏度为96.7%,且专一性为79.7%。在试验组中,该四蛋白分析组侦测口腔鳞状细胞癌的灵敏度为87.5%,而专一性为78.8%。于测试资料组及试验组的正确度分别为86.6%及82.1% (表4)。

[0120] 表4. 比对三种用以建立标志分析组的统计方法

[0121]

	CART	逻辑式回归 [#]	最近邻辨别法
标志 s	ANXA2, HSPA5, KNG1, MMP1	ANXA2, HSPA5, KNG1, PRDX2	ANXA2, HSPA5, PRDX2, FLNA, KNG1, TIMP1
测试资料组 (n=224)			
准确度	86.61%	83.50%	78.10%
灵敏度	96.70%	75.80%	51.60%
专一性	79.70%	88.70%	96.20%
LR ⁺	4.76	6.70	13.60
LR ⁻	0.04	0.30	0.50
试验组(n=106)			
准确度	82.08%	85.80%	78.30%
灵敏度	87.50%	85.00%	57.50%
专一性	78.79%	86.40%	90.90%
LR ⁺	4.13	6.20	6.30
LR ⁻	0.16	0.20	0.50

[0122] [#]概率>0.4

[0123] 由于吸烟及咀嚼槟榔为二种影响口腔鳞状细胞癌发生的重要危险因子,接着利用蛋白标志来分析年龄/吸烟/咀嚼槟榔之间的关联性。然而,结果显示在460个体中,四种蛋白与危险习惯并无显著的关联性(表5)。因此,相较于以口腔检查进行的口腔癌筛选(专一性约为98%,不同国家具有不同的灵敏度(50%-99%)),本发明利用CART筛选的四蛋白分析组更适合用以侦测台湾口腔癌筛检计画中的口腔鳞状细胞癌病患。

[0124] 表5.唾液中四种生物标志的含量与吸烟或咀嚼槟榔于460个体中的关联性

[0125]

		年龄	吸烟	槟榔	19_MMP1	1_ANXA2	16_KNG1	13_HSPA5
年龄	皮尔森 r p	1	0.21845 <.0001	0.08635 0.0643	-0.00535 0.9089	0.08459 0.0699	0.11626 0.0126	0.13663 0.0033
吸烟	皮尔森 r p	0.21845 <.0001	1	0.27922 <.0001	0.03712 0.427	0.0706 0.1306	0.03178 0.4965	0.01341 0.7742
槟榔	皮尔森 r p	0.08635 0.0643	0.27922 <.0001	1	0.06318 0.1761	0.02778 0.5523	0.04162 0.3731	-0.0314 0.5017
19_MMP1	皮尔森 r p	-0.00535 0.9089	0.03712 0.427	0.06318 0.1761	1	0.20716 <.0001	0.51821 <.0001	0.27268 <.0001
1_ANXA2	皮尔森 r p	0.08459 0.0699	0.0706 0.1306	0.02778 0.5523	0.20716 <.0001	1	0.24789 <.0001	0.30029 <.0001
16_KNG1	皮尔森 r p	0.11626 0.0126	0.03178 0.4965	0.04162 0.3731	0.51821 <.0001	0.24789 <.0001	1	0.46617 <.0001
13_HSPA5	皮尔森 r p	0.13663 0.0033	0.01341 0.7742	-0.0314 0.5017	0.27268 <.0001	0.30029 <.0001	0.46617 <.0001	1

[0126] 1.2产生评分量表

[0127] 接着,依据四种蛋白标志的二元结果(即高于或低于内在截止值),利用逻辑式回归分析来计算可作为风险指数的预测概率。以Chi-square检定评估四蛋白分析组的变异显著性可得到MMP1、KNG1、ANXA2及HSPA5的p值分别<0.0001、<0.0001、0.0002及0.0007。在测试资料组中,该风险指数会由健康对照组的 0.16 ± 0.19 ,以及口腔低恶性潜能病变族群的 0.18 ± 0.29 ,显著增加为口腔鳞状细胞癌族群的 0.75 ± 0.24 ($p < 0.0001$) (第1B图),而于试验组亦可得到相似的结果,其中健康对照组为 0.21 ± 0.26 ,口腔低恶性潜能病变族群为 0.16 ± 0.22 ,而口腔鳞状细胞癌则为 0.74 ± 0.31 ($p < 0.0001$) (第1C图)。非口腔鳞状细胞癌检体及口腔鳞状细胞癌检体的ROC分析结果指出,测试资料组及试验组的AUC分别为0.926及0.91 (第1D图)。当将风险指数的截止值设定为0.4时,四标志评分量表于测试资料组产生的灵敏度为93.4%,专一性为80.5%。在试验组中,灵敏度仍然相当高(87.5%),而专一性则与测试资料组相同(80.5%)。

[0128] 实施例2多种标志的组合

[0129] 在建构CART树后(第1A图),将筛选标志的连续数值(数值变数)依据其分割点的截止浓度二分为二元变数。若ANXA2、KNG1及MMP1的浓度高于截止值,或是HSPA5的浓度低于截止值时,将口腔鳞状细胞癌的阳性预测值设定为1。相对地,若ANXA2、KNG1及MMP1的浓度低于截止值,或是HSPA5的浓度高于截止值时,则将口腔鳞状细胞癌的阴性预测值设定为0。将以CART得到的用以筛选标志的二元变数作为逻辑式回归的共变量,并以下列公式计算罹患口腔鳞状细胞癌的预测概率:

$$[0130] \quad \text{风险指数} = \frac{e^{a+b_1X_1+b_2X_2+b_3X_3+b_4X_4}}{1+e^{a+b_1X_1+b_2X_2+b_3X_3+b_4X_4}},$$

[0131] 其中e为一数学常数,其为自然对数的基数;a为一常数值;X1、X2、X3及X4分别代表ANXA2、HSPA5、KNG1及MMP1的浓度;且b1、b2、b3及b4分别代表ANXA2、HSPA5、KNG1及MMP1的变异系数。将由CART分析四种筛选标志得到的二元变数作为共变量,利用测试资料组检体进行逻辑式回归分析来得到个体罹患口腔鳞状细胞癌的预测概率,其中该测试资料组是由非口腔鳞状细胞癌组(健康对照+口腔低恶性潜能病变)及口腔鳞状细胞癌组所组成。除了四标志分析组之外,亦利用四种标志中任二种标志的组合来进行分析(表6)。此外,以四种筛选标志的连续数据(数值变数)进行逻辑式回归分析,藉以将四种蛋白组合为四或二标志分析组(表7)。该些结果指出,四种蛋白中每一种蛋白不论是基于二元变数或是基于数值变数所建构的四标志分析组,皆具有显著的意义(Sig.<0.05)。然而,在本分析中,利用HSPA5与ANXA2或MMP1结合所得到的二标志分析组不具有显著功效。

[0132] 表6. 依据由截止浓度决定的多种标志的二元变数利用逻辑式回归分析来得到标志分析组

[0133]

		B	S.E.	Wald	df	Sig.	Exp(B)
4- 标志 分析组	ANXA2 (>每毫升 36 纳克)	2.086	0.453	21.25	1	4.04E-06	8.06
	HSPA5 (<每毫升 166.9 纳克)	1.590	0.420	14.31	1	1.55E-04	4.90
	KNG1 (>每毫升 40 纳克)	3.105	0.561	30.68	1	3.04E-08	22.31
	MMP1 (>每毫升 2.772 纳克)	2.619	0.364	51.79	1	6.17E-13	13.72
	常数	-5.016	0.659	58.00	1	2.62E-14	0.01
2- 标志 分 析 组 #1	ANXA2 (>每毫升 36 纳克)	2.472	0.357	48.03	1	4.19E-12	11.84
	HSPA5 (<每毫升 166.9 纳克)	-0.296	0.263	1.26	1	2.61E-01	0.74
	常数	-0.800	0.202	15.70	1	7.44E-05	0.45
2- 标志 分 析 组 #2	ANXA2 (>每毫升 36 纳克)	2.163	0.380	32.44	1	1.23E-08	8.70
	KNG1 (>每毫升 40 纳克)	2.901	0.455	40.61	1	1.86E-10	18.19
	常数	-3.119	0.435	51.36	1	7.68E-13	0.04
2- 标志 分 析 组 #3	ANXA2 (>每毫升 36 纳克)	1.940	0.412	22.19	1	2.47E-06	6.96
	MMP1 (>每毫升 2.772 纳克)	2.950	0.321	84.22	1	4.42E-20	19.11
	常数	-1.985	0.214	86.17	1	1.65E-20	0.14
2- 标志 分 析 组 #4	HSPA5 (<每毫升 166.9 纳克)	0.657	0.312	4.44	1	3.50E-02	1.93
	KNG1 (>每毫升 40 纳克)	3.629	0.490	54.84	1	1.31E-13	37.66
	常数	-3.494	0.512	46.62	1	8.63E-12	0.03
2- 标志 分 析 组 #5	HSPA5 (<每毫升 166.9 纳克)	-0.101	0.311	0.10	1	7.47E-01	0.90
	MMP1 (>每毫升 2.772 纳克)	3.238	0.319	103.08	1	3.21E-24	25.48
	常数	-1.656	0.266	38.66	1	5.05E-10	0.19
2- 标志 分 析 组 #6	KNG1 (>每毫升 40 纳克)	2.328	0.476	23.88	1	1.02E-06	10.25
	MMP1 (>每毫升 2.772 纳克)	2.698	0.325	69.10	1	9.35E-17	14.85
	常数	-3.293	0.451	53.27	1	2.91E-13	0.04
3- 标志 分 析 组 #1	ANXA2 (>每毫升 36 纳克)	2.458	0.393	39.20	1	3.83E-10	11.69
	HSPA5 (<每毫升 166.9 纳克)	1.218	0.343	12.64	1	3.79E-04	3.38
	KNG1 (>每毫升 40 纳克)	3.639	0.512	50.55	1	1.16E-12	38.07
	常数	-4.296	0.559	59.07	1	1.52E-14	0.01
3- 标志 分 析 组 #2	ANXA2 (>每毫升 36 纳克)	2.050	0.431	22.59	1	2.01E-06	7.77
	HSPA5 (<每毫升 166.9 纳克)	0.322	0.340	0.90	1	3.43E-01	1.38
	MMP1 (>每毫升 40 纳克)	3.027	0.336	81.25	1	1.99E-19	20.64
	常数	-2.197	0.316	48.33	1	3.60E-12	0.11
3- 标志 分 析 组 #3	ANXA2 (>每毫升 36 纳克)	1.696	0.426	15.81	1	7.00E-05	5.45
	KNG1 (>每毫升 40 纳克)	2.153	0.486	19.61	1	9.48E-06	8.61
	MMP1 (>每毫升 2.772 纳克)	2.442	0.338	52.21	1	5.00E-13	11.50
	常数	-3.408	0.461	54.74	1	1.38E-13	0.03
3- 标志 分 析 组 #4	HSPA5 (<每毫升 166.9 纳克)	1.148	0.391	8.63	1	3.30E-03	3.15
	KNG1 (>每毫升 40 纳克)	3.056	0.546	31.31	1	2.19E-08	21.25
	MMP1 (>每毫升 2.772 纳克)	2.859	0.342	69.82	1	6.49E-17	17.44
	常数	-4.428	0.615	51.90	1	5.85E-13	0.01

[0134] B:变异系数;S.E.:邻近系数的标准差;Wald:Wald chi-square检定;df:Wald chi-square检定的自由度;Sig.:显著的p值;以及Exp(B):B系数的指数运算。

[0135] Table 7. 依据由截止浓度决定的多种标志的数值变数利用逻辑式回归分析来得
到标志分析组

[0136]

		B	S.E.	Wald	df	Sig.	Exp(B)
4-标志分析组	ANXA2	0.0366	0.009	15.05	1	1.05E-04	1.04
	HSPA5	-0.0016	0.001	5.48	1	1.92E-02	1.00
	KNG1	0.0012	0.001	4.01	1	4.52E-02	1.00
	MMP1	0.3357	0.065	27.02	1	2.02E-07	1.40
	常数	-2.2993	0.262	76.90	1	1.80E-18	0.10
2-标志分析组#1	ANXA2	0.0523	0.009	34.80	1	3.65E-09	1.05
	HSPA5	-0.0003	0.000	0.53	1	4.65E-01	1.00
	常数	-1.5116	0.197	59.01	1	1.57E-14	0.22
2-标志分析组#2	ANXA2	0.0330	0.008	18.80	1	1.45E-05	1.03
	KNG1	0.0028	0.001	17.75	1	2.52E-05	1.00
	常数	-1.8642	0.210	78.83	1	6.77E-19	0.16
2-标志分析组#3	ANXA2	0.0321	0.008	15.72	1	7.33E-05	1.03
	MMP1	0.3352	0.060	31.69	1	1.81E-08	1.40
	常数	-2.3902	0.256	86.84	1	1.17E-20	0.09
2-标志分析组#4	HSPA5	-0.0006	0.000	1.88	1	1.70E-01	1.00
	KNG1	0.0049	0.001	36.89	1	1.25E-09	1.00
	常数	-1.3170	0.181	52.83	1	3.63E-13	0.27
2-标志分析组#5	HSPA5	-0.0001	0.000	0.04	1	8.51E-01	1.00
	MMP1	0.4017	0.062	41.78	1	1.02E-10	1.49
	常数	-1.8367	0.221	69.29	1	8.50E-17	0.16
2-标志分析组#6	KNG1	0.0014	0.001	7.18	1	7.36E-03	1.00
	MMP1	0.3649	0.062	35.00	1	3.30E-09	1.44
	常数	-2.0499	0.218	88.02	1	6.49E-21	0.13
3-标志分析组#1	ANXA2	0.0405	0.009	21.13	1	4.30E-06	1.04
	HSPA5	-0.0014	0.001	7.22	1	7.21E-03	1.00
	KNG1	0.0036	0.001	19.25	1	1.14E-05	1.00
	常数	-1.7654	0.213	68.49	1	1.28E-16	0.17
3-标志分析组#2	ANXA2	0.0401	0.010	17.80	1	2.46E-05	1.04
	HSPA5	-0.0013	0.001	3.82	1	5.07E-02	1.00
	MMP1	0.3534	0.062	32.18	1	1.41E-08	1.42
	常数	-2.2545	0.260	75.23	1	4.19E-18	0.10
3-标志分析组#3	ANXA2	0.0282	0.008	11.79	1	5.94E-04	1.03
	KNG1	0.0008	0.001	2.54	1	1.11E-01	1.00
	MMP1	0.3183	0.061	27.34	1	1.71E-07	1.37
	常数	-2.4437	0.261	87.72	1	7.56E-21	0.09
3-标志分析组#4	HSPA5	-0.0006	0.001	1.20	1	2.74E-01	1.00
	KNG1	0.0016	0.001	7.42	1	6.46E-03	1.00
	MMP1	0.3770	0.064	34.64	1	3.97E-09	1.46
	常数	-1.9608	0.230	72.64	1	1.55E-17	0.14

[0137] B: 变异系数; S.E.: 邻近系数的标准差; Wald: Wald chi-square 检定; df: Wald

chi-square检定的自由度:Sig.:显著的p值;以及Exp(B):B系数的指数运算。

[0138] 藉由二元变数分析(表8及9)或数值变数分析(表10及11)得到的不同标志分析组的ROC,可用以比对非口腔鳞状细胞癌族群(健康对照+口腔低恶性潜能病变)及口腔鳞状细胞癌族群,比对口腔高恶性潜能病变族群及口腔鳞状细胞癌族群,以及比对非转形族群及口腔高恶性潜能病变病患中产生口腔鳞状细胞癌转形的族群。结果发现多数标志分析组可用以辨别口腔鳞状细胞癌与非口腔鳞状细胞癌,或是口腔鳞状细胞癌与口腔高恶性潜能病变(AUC 0.65-0.93),且除了二标志分析组#1、#3及#5的AUC值小于0.6(当用以辨识口腔高恶性潜能病变病患中非转形病患与口腔鳞状细胞癌转形病患时)之外,其余标志分析组皆可用以预测口腔高恶性潜能病变病患的恶性转形。

[0139] 表8.由二元变数得到的不同标志分析组的ROC于非口腔鳞状细胞癌及口腔鳞状细胞癌族群的分析结果

[0140]

检定结果变异	非口腔鳞状细胞癌(健康+口腔低恶性潜能病变, n=199)与口腔鳞状细胞癌(n=131)				
	AUC	标准差 ^a	渐近显著 ^b	渐近 95%信赖区间	
				下限	上限
4-标志分析组 (ANXA2, HSPA5, KNG1, MMP1)	0.922	0.015	1.90E-38	0.892	0.951
2-标志分析组#1 (ANXA2, HSPA5)	0.710	0.031	1.03E-10	0.650	0.771
2-标志分析组#2 (ANXA2, KNG1)	0.836	0.022	5.49E-25	0.793	0.879
2-标志分析组#3 (ANXA2, MMP1)	0.865	0.023	3.06E-29	0.821	0.909
2-标志分析组#4 (HSPA5, KNG1)	0.777	0.025	1.68E-17	0.728	0.826
2-标志分析组#5 (HSPA5, MMP1)	0.828	0.025	5.87E-24	0.779	0.878
2-标志分析组#6 (KNG1, MMP1)	0.884	0.019	3.97E-32	0.847	0.921
3-标志分析组#1 (ANXA2, HSPA5, KNG1)	0.863	0.020	5.45E-29	0.824	0.903
3-标志分析组#2 (ANXA2, HSPA5, MMP1)	0.878	0.021	3.00E-31	0.838	0.919
3-标志分析组#3 (ANXA2, KNG1, MMP1)	0.907	0.017	5.57E-36	0.875	0.940
3-标志分析组#4 (HSPA5, KNG1, MMP1)	0.900	0.017	9.62E-35	0.866	0.934

[0141] a. 在无母数假设情况下

[0142] b. 零假设: 实际区域=0.5

[0143] 表9. 由二元变数得到的不同标志分析组的ROC于特定族群的分析结果

[0144]

检定结果变异		口腔高恶性潜能病变(n=130)与口腔鳞状细胞癌(n=131)				
		AUC	标准差 ^a	渐近显著 ^b	渐近 95%信赖区间	
					下限	上限
二元变数	4-标志分析组 (ANXA2, HSPA5, KNG1, MMP1)	0.840	0.024	2.26E-21	0.792	0.888
	2-标志分析组#1 (ANXA2, HSPA5)	0.716	0.032	1.63E-09	0.653	0.779
	2-标志分析组#2 (ANXA2, KNG1)	0.765	0.029	1.29E-13	0.708	0.822
	2-标志分析组#3 (ANXA2, MMP1)	0.807	0.028	1.02E-17	0.753	0.861
	2-标志分析组#4 (HSPA5, KNG1)	0.650	0.035	2.84E-05	0.582	0.718
	2-标志分析组#5 (HSPA5, MMP1)	0.733	0.032	7.05E-11	0.670	0.797
	2-标志分析组#6 (KNG1, MMP1)	0.811	0.027	4.09E-18	0.758	0.864
检定结果变异		非- (n=70)与恶性- (n=18)转形				
		AUC	标准差 ^a	渐近显著 ^b	渐近 95%信赖区间	
					下限	上限
二元变数	4-标志分析组 (ANXA2, HSPA5, KNG1, MMP1)	0.716	0.057	4.82E-03	0.605	0.828
	2-标志分析组#1 (ANXA2, HSPA5)	0.571	0.077	3.52E-01	0.421	0.722
	2-标志分析组#2 (ANXA2, KNG1)	0.715	0.057	5.06E-03	0.604	0.826
	2-标志分析组#3 (ANXA2, MMP1)	0.531	0.078	6.87E-01	0.379	0.683
	2-标志分析组#4 (HSPA5, KNG1)	0.757	0.055	8.03E-04	0.649	0.866
	2-标志分析组#5 (HSPA5, MMP1)	0.494	0.078	9.42E-01	0.341	0.648
	2-标志分析组#6 (KNG1, MMP1)	0.693	0.060	1.18E-02	0.576	0.810

[0145] a. 在无母数假设情况下

[0146] b. 零假设: 实际区域=0.5

[0147] 表10. 由数值变数得到的不同标志分析组的ROC于非口腔鳞状细胞癌及口腔鳞状细胞癌族群的分析结果

[0148]

检定结果变异	非口腔鳞状细胞癌(健康+口腔低恶性潜能病变, n=199)与口腔鳞状细胞癌(n=131)				
	AUC	标准差 ^a	渐近显著 ^b	渐近 95%信赖区间	
				下限	上限
数值变数					
4-标志分析组 (ANXA2, HSPA5, KNG1, MMP1)	0.930	0.014	6.70E-40	0.902	0.958
2-标志分析组#1 (ANXA2, HSPA5)	0.822	0.023	4.49E-23	0.776	0.867
2-标志分析组#2 (ANXA2, KNG1)	0.865	0.020	2.96E-29	0.825	0.905
2-标志分析组#3 (ANXA2, MMP1)	0.917	0.016	1.10E-37	0.886	0.949
2-标志分析组#4 (HSPA5, KNG1)	0.891	0.018	3.18E-33	0.855	0.927
2-标志分析组#5 (HSPA5, MMP1)	0.857	0.025	5.67E-28	0.807	0.906
2-标志分析组#6 (KNG1, MMP1)	0.917	0.015	1.13E-37	0.887	0.948
3-标志分析组#1 (ANXA2, HSPA5, KNG1)	0.892	0.018	2.23E-33	0.856	0.928
3-标志分析组#2 (ANXA2, HSPA5, MMP1)	0.923	0.015	1.41E-38	0.892	0.953
3-标志分析组#3 (ANXA2, KNG1, MMP1)	0.922	0.015	2.01E-38	0.891	0.952
3-标志分析组#4 (HSPA5, KNG1, MMP1)	0.913	0.017	6.86E-37	0.879	0.946

[0149] a. 在无母数假设情况下

[0150] b. 零假设: 实际区域=0.5

[0151] 表11. 由数值变数得到的不同标志分析组的ROC于特定族群的分析结果

[0152]

检定结果变异		口腔高恶性潜能病变(n=130)与口腔鳞状细胞癌(n=131)				
		AUC	标准差 ^a	渐近显著 ^b	渐近 95%信赖区间	
					下限	上限
数值变数	4-标志分析组 (ANXA2, HSPA5, KNG1, MMP1)	0.870	0.022	5.02E-25	0.827	0.913
	2-标志分析组#1 (ANXA2, HSPA5)	0.780	0.028	4.98E-15	0.725	0.835
	2-标志分析组#2 (ANXA2, KNG1)	0.819	0.026	4.92E-19	0.769	0.869
	2-标志分析组#3 (ANXA2, MMP1)	0.855	0.023	3.14E-23	0.810	0.901
	2-标志分析组#4 (HSPA5, KNG1)	0.822	0.026	2.64E-19	0.771	0.872
	2-标志分析组#5 (HSPA5, MMP1)	0.802	0.029	3.19E-17	0.745	0.859
	2-标志分析组#6 (KNG1, MMP1)	0.862	0.023	4.65E-24	0.817	0.907
检定结果变异		非- (n=70)与恶性- (n=18)转形				
		AUC	标准差 ^a	渐近显著 ^b	渐近 95%信赖区间	
					下限	上限
数值变数	4-标志分析组 (ANXA2, HSPA5, KNG1, MMP1)	0.671	0.067	2.61E-02	0.539	0.802
	2-标志分析组#1 (ANXA2, HSPA5)	0.555	0.080	4.75E-01	0.398	0.711
	2-标志分析组#2 (ANXA2, KNG1)	0.734	0.055	2.28E-03	0.626	0.842
	2-标志分析组#3 (ANXA2, MMP1)	0.525	0.084	7.48E-01	0.361	0.688
	2-标志分析组#4 (HSPA5, KNG1)	0.810	0.048	5.48E-05	0.716	0.903
	2-标志分析组#5 (HSPA5, MMP1)	0.487	0.084	8.60E-01	0.321	0.652
	2-标志分析组#6 (KNG1, MMP1)	0.658	0.070	3.95E-02	0.521	0.795

[0153] a. 在无母数假设情况下

[0154] b. 零假设:实际区域=0.5

[0155] 由二元变数(表12)或数值变数(表13)分析得到不同标志分析组的风险指数,可用以比对非口腔鳞状细胞癌(健康对照+口腔低恶性潜能病变)族群及口腔鳞状细胞癌族群。当风险指数的截止值设定为0.4时,具有高阴性预测值(除了2-标志分析组#1、#2及#4之外,其余为81.9%-93.6%);设定为0.6时,具有高阳性预测值(除了3-标志分析组#1及2-标志分析组#4之外,其余为81.4%-93.4%)。该些结果指出,将风险指数的截止值设定为0.4或

0.6时,可高度准确(除了2-标志分析组#1、#2及#4之外,其余为78.8%-86.7%)地辨别非口腔鳞状细胞癌(健康对照+口腔低恶性潜能病变)个体及口腔鳞状细胞癌病患。

[0156] 表12. 依据截止浓度所决定的标志的二元结果,利用标志分析组由健康对照+口腔低恶性潜能病变个体区别口腔鳞状细胞癌的功效

[0157]

分析组	分析组中计算风险指数的标志 ($RS=e^{f(x)}/(1+e^{f(x)})$)	RS 的截止值	灵敏度	专一性	准确度	阳性预测值	阴性预测值
4-标志分析组							
ANXA2 (A), HSPA5 (H), KNG1 (K), MMP1 (M) $f(x)=2.086*X_A+1.590*X_H+3.105*X_K+2.619*X_M-5.016$	0.4	91.6%	80.4%	84.8%	75.5%	93.6%	
	0.6	74.0%	89.9%	83.6%	82.9%	84.0%	
3-标志分析组#1							
ANXA2 (A), HSPA5 (H), KNG1 (K) $f(x)=2.458*X_A+1.218*X_H+3.639*X_K-4.296$	0.4	71.8%	86.9%	80.9%	78.3%	82.4%	
	0.6	71.8%	86.9%	80.9%	78.3%	82.4%	
3-标志分析组#2							
ANXA2 (A), HSPA5 (H), MMP1 (M) $f(x)=2.050*X_A+0.322*X_H+3.027*X_M-2.197$	0.4	82.4%	84.9%	83.9%	78.3%	88.0%	
	0.6	75.6%	89.4%	83.9%	82.5%	84.8%	
3-标志分析组#3							
ANXA2 (A), KNG1 (K), MMP1 (M) $f(x)=1.696*X_A+2.153*X_K+2.442*X_M-3.408$	0.4	80.2%	87.9%	84.8%	81.4%	87.1%	
	0.6	80.2%	87.9%	84.8%	81.4%	87.1%	
3-标志分析组#4							
HSPA5 (H), KNG1 (K), MMP1 (M) $f(x)=1.1481*X_H+3.056*X_K+2.859*X_M-4.428$	0.4	85.5%	82.9%	83.9%	76.7%	89.7%	
	0.6	73.3%	91.0%	83.9%	84.2%	83.8%	
2-标志分析组#1							
ANXA2 (A), HSPA5 (H) $f(x)=2.472*X_A-0.296*X_H-0.800$	0.4	45.8%	94.0%	74.8%	83.3%	72.5%	
	0.6	45.8%	94.0%	74.8%	83.3%	72.5%	
2-标志分析组#2							
ANXA2 (A), KNG1 (K) $f(x)=2.163*X_A+2.901*X_K-3.119$	0.4	95.4%	54.3%	70.6%	57.9%	94.7%	
	0.6	45.8%	96.5%	76.4%	89.6%	73.0%	
2-标志分析组#3							
ANXA2 (A), MMP1 (M) $f(x)=1.940*X_A+2.950*X_M-1.985$	0.4	82.4%	84.9%	83.9%	78.3%	88.0%	
	0.6	75.6%	89.4%	83.9%	82.5%	84.8%	
2-标志分析组#4							
HSPA5 (H), KNG1 (K) $f(x)=-0.657*X_H+3.629*X_K-3.494$	0.4	95.4%	54.3%	70.6%	57.9%	94.7%	
	0.6	32.8%	89.9%	67.3%	68.3%	67.0%	
2-标志分析组#5							
HSPA5 (H), MMP1 (M) $f(x)=-0.101*X_H+3.238*X_M-1.656$	0.4	75.6%	89.4%	83.9%	82.5%	84.8%	
	0.6	75.6%	89.4%	83.9%	82.5%	84.8%	

[0158]

2-标志分析组#6

KNG1 (K), MMP1 (M)	0.4	73.3%	91.0%	83.9%	84.2%	83.8%
$f(x)=2.328*X_K+2.698*X_M-3.293$	0.6	73.3%	91.0%	83.9%	84.2%	83.8%

[0159] X_A 为由截止浓度决定的ANXA2的二元结果(>每毫升36纳克为正)[0160] X_H 为由截止浓度决定的HSPA5的二元结果(<每毫升166.9纳克为正)[0161] X_K 为由截止浓度决定的KNG1的二元结果(>每毫升40纳克为正)[0162] X_M 为由截止浓度决定的MMP1的二元结果(>每毫升2.772纳克为正)

[0163] 表13. 依据标志的数值浓度,利用标志分析组由健康对照+口腔低恶性潜能病变个体区别口腔鳞状细胞癌的功效

[0164]

分析组	分析组中计算风险指数的标志 ($RS=e^{f(x)}/(1+e^{f(x)})$)	RS 的 截止值	灵敏度	专一性	准确度	阳性预测 值	阴性预测 值
4-标志分析组							
ANXA2 (A), HSPA5 (H), KNG1 (K), MMP1 (M)		0.4	77.1%	93.0%	86.7%	87.8%	86.0%
$f(x)=0.0366*X_A-0.0016*X_H+$ $0.0012*X_K+0.3357*X_M-2.2993$		0.6	67.9%	96.0%	84.8%	91.8%	82.0%
3-标志分析组#1							
ANXA2 (A), HSPA5 (H), KNG1 (K)		0.4	70.2%	88.9%	81.5%	80.7%	81.9%
$f(x)=0.0405*X_A-0.0014*X_H+$ $0.0036*X_K-1.7654$		0.6	52.7%	96.0%	78.8%	89.6%	75.5%
3-标志分析组#2							
ANXA2 (A), HSPA5 (H), MMP1 (M)		0.4	75.6%	92.5%	85.8%	86.8%	85.2%
$f(x)=0.0401*X_A-0.0013*X_H+$ $0.3534*X_M-2.2545$		0.6	67.2%	96.5%	84.8%	92.6%	81.7%
3-标志分析组#3							
ANXA2 (A), KNG1 (K), MMP1 (M)		0.4	74.0%	93.5%	85.8%	88.2%	84.5%
$f(x)=0.0282*X_A+0.0008*X_K+$ $0.3183*X_M-2.4437$		0.6	67.9%	96.5%	85.2%	92.7%	82.1%
3-标志分析组#4							
HSPA5 (H), KNG1 (K), MMP1 (M)		0.4	74.0%	93.0%	85.5%	87.4%	84.5%
$f(x)=-0.0006*X_H+0.0016*X_K+$ $0.3770*X_M-1.9608$		0.6	66.4%	96.0%	84.2%	91.6%	81.3%
2-标志分析组#1							
ANXA2 (A), HSPA5 (H)		0.4	59.5%	86.9%	76.1%	75.0%	76.5%
$f(x)=0.0523*X_A-0.0003*X_H-1.5116$		0.6	41.2%	94.5%	73.3%	83.1%	70.9%
2-标志分析组#2							
ANXA2 (A), KNG1 (K)		0.4	67.2%	88.9%	80.3%	80.0%	80.5%
$f(x)=0.0330*X_A+0.0028*X_K-1.8642$		0.6	51.1%	95.0%	77.6%	87.0%	74.7%
2-标志分析组#3							
ANXA2 (A), MMP1 (M)		0.4	74.8%	94.0%	86.4%	89.1%	85.0%
$f(x)=0.0321*X_A+0.3352*X_M-2.3902$		0.6	67.2%	96.5%	84.8%	92.6%	81.7%
2-标志分析组#4							
HSPA5 (H), KNG1 (K)		0.4	61.8%	91.5%	79.7%	82.7%	78.4%
$f(x)=-0.0006*X_H+0.0049*X_K-1.3170$		0.6	44.3%	97.0%	76.1%	90.6%	72.6%
2-标志分析组#5							
HSPA5 (H), MMP1 (M)		0.4	71.0%	92.5%	83.9%	86.1%	82.9%
$f(x)=-0.0001*X_H+0.4017*X_M-1.8367$		0.6	64.9%	97.0%	84.2%	93.4%	80.8%

2-标志分析组#6						
[0165]	KNG1 (K), MMP1 (M)	0.4	74.0%	93.0%	85.5%	87.4%
	$f(x)=0.0014 \cdot X_K+0.3649 \cdot X_M-2.0499$	0.6	65.6%	96.0%	83.9%	91.5%
					84.5%	80.9%

[0166] X_A 为ANXA2的数值浓度(纳克/每毫升)

[0167] X_H 为HSPA5的数值浓度(纳克/每毫升)

[0168] X_K 为KNG1的数值浓度(纳克/每毫升)

[0169] X_M 为MMP1的数值浓度(纳克/每毫升)

[0170] 实施例3第I-IV期的口腔鳞状细胞癌病患的风险指数

[0171] 本研究包含50位第I期、29位第II期、16位第III期及36位第IV期的口腔鳞状细胞癌病患。利用四标志建立的评分量表来计算该些病患的风险指数。如第2图所示,风险指数会由早期到晚期逐渐增加(第I期: 0.63 ± 0.29 ;第II期: 0.78 ± 0.23 ;第III期: 0.83 ± 0.23 ;以及第IV期: 0.85 ± 0.20)。更重要的是,相较于非口腔鳞状细胞癌族群(健康对照+口腔低恶性潜能病变;平均分数为 0.17 ± 0.24),第I期口腔鳞状细胞癌病患的风险指数会显著上升($p < 0.0001$)。此外,84% (42/50)、97% (28/29)、94% (15/16)及97% (35/36)的第I、II、III及IV期的口腔鳞状细胞癌病患具有 >0.4 的风险指数(表14),该结果指出四蛋白分析组的评分量表具有侦测第I期口腔鳞状细胞癌病患的显著族群($>80\%$)的良好潜力。

[0172] 表14. 在第I到IV期的口腔鳞状细胞癌病患中,风险指数 >0.4 的的个体的百分比

[0173]

口腔鳞状细胞癌的 期数	病患数	风险指数 > 0.4	
		阴性病患数(%)	阳性病患数(%)
I	50	8 (16%)	42 (84%)
II	29	1 (3%)	28 (97%)
III	16	1 (6%)	15 (94%)
IV	36	1 (3%)	35 (97%)
全部	131	11 (8.4%)	120 (91.6%)

[0174] 实施例4口腔高恶性潜能病变病患的风险指数及其后续追踪结果

[0175] 基于第口腔高恶性潜能病变的病灶可能包含潜在性恶性细胞、恶性细胞及正常细胞的混合物,一般临床医疗人员并不易由口腔高恶性潜能病变中辨识出口腔鳞状细胞癌。然而,口腔高恶性潜能病变族群的平均风险指数(0.32 ± 0.33)高于非口腔鳞状细胞癌族群的平均风险指数(健康对照+口腔低恶性潜能病变; 0.17 ± 0.24),且低于口腔鳞状细胞癌族群的平均风险指数(0.75 ± 0.26) (第3图)。值得注意的是,42% (55/130)的口腔高恶性潜能病变病的风险指数 >0.4 (表15)。该结果与口腔高恶性潜能病变病灶可能包含恶性细胞的推论一致。

[0176] 表15. 在非口腔鳞状细胞癌(健康对照+口腔低恶性潜能病变)及口腔高恶性潜能病变族群中,风险指数 >0.4 的个体的百分比

[0177]

组别	病患数	风险指数 > 0.4	
		阴性病患数(%)	阳性病患数(%)
健康对照+口腔低恶性潜能病变	199	160 (80%)	39 (20%)
口腔高恶性潜能病变	130	75 (58%)	55 (42%)

[0178] 除了需要侦测口腔鳞状细胞癌之外,另一个重要的需求为预测及监控口腔恶性潜能病变(特别是口腔高恶性潜能病变族群)的恶性转形。在本研究233位口腔恶性潜能病变病患中,持续对153位(65位口腔低恶性潜能病变及88位口腔高恶性潜能病变)病患追踪13.5到76.6个月,以了解其疾病的恶化状况。18位口腔高恶性潜能病变族群于1.2到65.5个月之间恶性转形为口腔鳞状细胞癌;这些病患包含1位粘膜红白斑、1位具有粘膜下纤维化的粘膜红斑、1位粘膜下纤维化、1位斑点白斑症、4位疣状增生,以及10位具有粘膜下纤维化的疣状增生。相较之下,在口腔低恶性潜能病变族群的追踪观察中,并无观察到恶性转形。口腔低恶性潜能病变病患的恶性转形为11.8% (18/153),该数值落在先前研究报告范围内。利用二元变数(表16)或数值变数(表17)分析得到不同标志分析组于口腔高恶性潜能病变族群的风险指数。相较于风险指数 <0.4 的病患,风险指数 ≥ 0.4 的病患具有较高的口腔鳞状细胞癌转形率(除了2-标志分析组#1之外)。举例来说,依据Table 16的4-标志分析组,37位病患的风险指数 ≥ 0.4 ,其中37.8% (37位中的14位)在追踪时转形为口腔鳞状细胞癌。该转形率远高于51位风险指数 <0.4 的口腔高恶性潜能病变病患(7.8%; 4/51) (表16)。在18位口腔鳞状细胞癌转形病患中,77.8% (14/18)的风险指数 >0.4 。

[0179] 表16.88位口腔高恶性潜能病变个体的口腔鳞状细胞癌转形率及其二元标志分析组的风险指数

[0180]

分析组	分析组中计算 风险指数的标志 ($RS=e^{f(x)}/(1+e^{f(x)})$)	转形为口腔鳞状细胞癌的病患百分比	
		RS<0.4	RS≥0.4
4-标志分析组			
ANXA2 (A), HSPA5 (H), KNG1 (K), MMP1 (M) $f(x)=2.086*X_A+1.590*X_H+3.105*X_K+2.619*X_M-5.016$		7.8% (4/51)	37.8% (14/37)
3-标志分析组#1			
ANXA2 (A), HSPA5 (H), KNG1 (K) $f(x)=2.458*X_A+1.218*X_H+3.639*X_K-4.296$		12.5% (7/56)	34.4% (11/32)
3-标志分析组#2			
ANXA2 (A), HSPA5 (H), MMP1 (M) $f(x)=2.050*X_A+0.322*X_H+3.027*X_M-2.197$		19.0% (12/63)	24.0% (6/25)
3-标志分析组#3			
ANXA2 (A), KNG1 (K), MMP1 (M) $f(x)=1.696*X_A+2.153*X_K+2.442*X_M-3.408$		17.9% (12/67)	28.6% (6/21)
3-标志分析组#4			
HSPA5 (H), KNG1 (K), MMP1 (M) $f(x)=1.1481*X_H+3.056*X_K+2.859*X_M-4.428$		9.3% (5/54)	38.2% (13/34)
2-标志分析组#1			
ANXA2 (A), HSPA5 (H) $f(x)=2.472*X_A-0.296*X_H-0.800$		20.5% (16/78)	20.0% (2/10)
2-标志分析组#2			
ANXA2 (A), KNG1 (K) $f(x)=2.163*X_A+2.901*X_K-3.119$		2.6% (1/39)	34.7% (17/49)
2-标志分析组#3			
ANXA2 (A), MMP1 (M) $f(x)=1.940*X_A+2.950*X_M-1.985$		19.0% (12/63)	24.0% (6/25)
2-标志分析组#4			
HSPA5 (H), KNG1 (K) $f(x)=-0.657*X_H+3.629*X_K-3.494$		15.9% (10/63)	32.0% (8/25)
2-标志分析组#5			
HSPA5 (H), MMP1 (M) $f(x)=-0.101*X_H+3.238*X_M-1.656$		19.1% (13/68)	25.0% (5/20)
2-标志分析组#6			
KNG1 (K), MMP1 (M) $f(x)=2.328*X_K+2.698*X_M-3.293$		18.1% (13/72)	31.3% (5/16)

[0181] X_A 为以截止浓度决定的ANXA2的二元结果(>每毫升36纳克为正)[0182] X_H 为以截止浓度决定的HSPA5的二元结果(<每毫升166.9纳克为正)[0183] X_K 为以截止浓度决定的KNG1的二元结果(>每毫升40纳克为正)[0184] X_M 为以截止浓度决定的MP1的二元结果(>每毫升2.772纳克为正)

[0185] 表17.88位口腔高恶性潜能病变个体的口腔鳞状细胞癌转形率及其数值标志分析组的风险指数

[0186]

分析组	分析组中计算风险指数的标志 ($RS=e^{f(x)} / (1+e^{f(x)})$)	转形为腔鳞状细胞癌的病患百分比	
		RS<0.4	RS≥0.4
4-标志分析组			
ANXA2 (A), HSPA5 (H), KNG1 (K), MMP1 (M) $f(x)=0.0366*X_A-0.0016*X_H+$ $0.0012*X_K+0.3357*X_M-2.2993$		19.5% (15/77)	27.3% (3/11)
3-标志分析组#1			
ANXA2 (A), HSPA5 (H), KNG1 (K) $f(x)=0.0405*X_A-0.0014*X_H+$ $0.0036*X_K-1.7654$		19.1% (13/68)	25.0% (5/20)
3-标志分析组#2			
ANXA2 (A), HSPA5 (H), MMP1 (M) $f(x)=0.0401*X_A-0.0013*X_H+$ $0.3534*X_M-2.2545$		20.0% (15/75)	23.1% (3/13)
3-标志分析组#3			
ANXA2 (A), KNG1 (K), MMP1 (M) $f(x)=0.0282*X_A+0.0008*X_K+$ $0.3183*X_M-2.4437$		19.5% (15/77)	27.3% (3/11)
3-标志分析组#4			
HSPA5 (H), KNG1 (K), MMP1 (M) $f(x)=-0.0006*X_H+0.0016*X_K+$ $0.3770*X_M-1.9608$		17.6% (13/74)	35.7% (5/14)
2-标志分析组#1			
ANXA2 (A), HSPA5 (H) $f(x)=0.0523*X_A-0.0003*X_H-1.5116$		18.3% (13/71)	29.4% (5/17)
2-标志分析组#2			
ANXA2 (A), KNG1 (K) $f(x)=0.0330*X_A+0.0028*X_K-1.8642$		19.4% (14/72)	25.0% (4/16)
2-标志分析组#3			
ANXA2 (A), MMP1 (M) $f(x)=0.0321*X_A+0.3352*X_M-2.3902$		19.5% (15/77)	27.3% (3/11)
2-标志分析组#4			
HSPA5 (H), KNG1 (K) $f(x)=-0.0006*X_H+0.0049*X_K-1.3170$		16.4% (12/73)	40.0% (6/15)
2-标志分析组#5			
HSPA5 (H), MMP1 (M) $f(x)=-0.0001*X_H+0.4017*X_M-1.8367$		17.6% (13/74)	35.7% (5/14)
2-标志分析组#6			

[0187]

KNG1 (K), MMP1 (M)

18.4% (14/76)

33.3% (4/12)

 $f(x)=0.0014 \cdot X_K + 0.3649 \cdot X_M - 2.0499$ [0188] X_A 为ANXA2的数值浓度(纳克/每毫升)[0189] X_H 为HSPA5的数值浓度(纳克/每毫升)[0190] X_K 为KNG1的数值浓度(纳克/每毫升)[0191] X_M 为MMP1的数值浓度(纳克/每毫升)

[0192] 不同形式的口腔恶性潜能病变病灶具有复杂的病理特性,常会使口腔鳞状细胞癌的早期侦复杂化而不易诊断。约有1400篇文献探讨过候选蛋白,相较于健康对照个体,该些候选蛋白于口腔鳞状细胞癌病患的体液或组织中具有较高的表现量。然而,目前尚未有分子标志可实际应用于临床阶段来侦测早期疾病及/或提供转形高风险病灶(即口腔高恶性潜能病变的病患)早期警戒。本研究提供了由四种蛋白所组成的分析组,该些蛋白可直接由唾液侦测,且可有效辨识台湾口腔癌筛检计画中口腔鳞状细胞癌病患(包含第I期)及非口腔鳞状细胞癌病患。该四蛋白分析组可用以评估临床上可疑或口腔高恶性潜能病变病灶的恶性转形风险,进而防止诊断延误。在持续追踪88位口腔高恶性潜能病变个体的过程中,18位于5年内罹患癌症,其中14位在首次诊断罹患口腔高恶性潜能病变时,由唾液检体检测出具有高风险指数(>0.4)。

[0193] 本揭示内容提供可作为临床试验的应用基础,其利用本发明四标志分析组来:(i) 侦测高风险族群(例如参与台湾口腔癌筛检计画的个体)的口腔鳞状细胞癌的发生;(ii) 评估在临床可疑病灶中恶性细胞出现的风险;(iii) 筛选进一步追踪观察的口腔高恶性潜能病变病患;以及(iv) 监控治疗反应及复发。该四蛋白分析组可作为诊断辅助,以减少因病患或第一线医疗人员所造成的诊断延误。当将风险指数的截止值设定为0.4及0.6时,可分别于辨别口腔鳞状细胞癌与非口腔鳞状细胞癌时,产生高灵敏度(91.6%)及高专一性(90%)的辨识结果(表18),该设定值可用以侦测高风险族群的口腔鳞状细胞癌。基于该结果,(i) 高度风险指数(≥ 0.6)的个体需要再次进行活体组织切片或全面检测隐匿性肿瘤;(ii) 中度风险指数(≥ 0.4 且 < 0.6)的个体需要每年追踪检查2次;(iii) 低度风险指数(< 0.4)的个体可依据近期追踪计画处理(每2年检查1次);(iv) 低风险指数(< 0.4)且具有正常粘膜的个体可作为回归管理的重要指标,例如延长后续检查的间隔时间。

[0194] 表18. 将风险指数的截止值设定为0.4及0.6来分辨非口腔鳞状细胞癌(健康对照+口腔低恶性潜能病变)与口腔鳞状细胞癌族群的评估结果(灵敏度及专一性)

[0195]		健康对照+口腔恶性潜能病变与口腔鳞状细胞癌	
	AUC	0.922	
	截止值分数	≥ 0.4	≥ 0.6
	灵敏度	91.60%	74.05%
	专一性	80.40%	89.95%

[0196] 本揭示内容的数项发现十分值得注意。在过去20年间,超过一千篇的研究在探讨包含口腔鳞状细胞癌等头颈部癌的生物标志。然而,仅有少数的生物标志可应用至临床阶段。此结果反应了该些研究缺乏于充分的个体及对照检体中,比对候选生物标志,藉以确认可提供足够预期值以适当引导医疗照顾的生物标志。本揭示内容提一种解决方法,藉由:(i)使用密集的文獻内容来筛选候选蛋白,该些候选蛋白皆经本研究或其他研究以不同形态的临床检体进行测式;以及(ii)比对高风险族群的罹癌(口腔鳞状细胞癌)检体及对照组(健康对照及口腔低恶性潜能病变)检体,其具有相似的风险因子(吸烟或咀嚼槟榔)。四种筛选蛋白于唾液(其为口腔癌的特异性位置)的侦测灵敏度为87.5-93.4%。此外,标志分析组可成功侦测88%(79位中的70位)罹患早期(第I或II期)口腔鳞状细胞癌的病患,以及92%(131位中的120位)罹患口腔鳞状细胞癌的病患。该结果指出由疾病位置非侵入式地收集唾液以检测蛋白生物标志来早期侦测口腔鳞状细胞癌的可能性。

[0197] 许多交互作用的因子会延误口腔鳞状细胞癌的诊断,影响病患的预后结果及存活率。举例来说,口腔鳞状细胞癌会发生于口腔的所有组织中,而不同形式的口腔恶性潜能病变皆会使诊断变得更为复杂化。近期将口腔检查病灶位置作为诊断的第一线方法,但该策略的成功性却往往依据个人经验的不同而有所差异。在某些情况下,罹患严重粘膜下纤维化的病患无法完全张开嘴巴来进行全面性的检查或活体组织切片。此外,某些个体可能具有多种形式的病灶。再者,临床上采用的病理检验通常局限于单一检体(活体组织切片),而错失包覆于混合型口腔恶性潜能病变的病灶中的癌细胞。本发明四蛋白分析组于唾液检体(包含由整口收集的混合物)的应用分析将可克服该些问题。

[0198] 本发明蛋白分析组的专一性约为80%。未来,可进一步与其他类型的唾液癌细胞标志(例如对肿瘤具有专一性的mRNA或DNA突变)合并使用,以提升其检测功效。至于限制方面,本发明四组检测族群皆仅包含少量的受试者,且所有的受测者皆由单一医院的二个临床单位收集而来。未来将需要由多个医院收集更大量的检体来进行临床试验,以评估四蛋白分析组于早期侦测口腔鳞状细胞癌的功效。此外,本研究是采用回溯性样本。未来将需要利用特定用途样本进行前瞻性研究以评估新生物标志分析组的临床适用性。

[0199] 对一成功的生物标志评估研究而言,诸如多重测定、高灵敏度及专一性,以及广泛的侦测范围是相当重要的评估指标。在本揭示内容中,是利用LC-MRM-MS(一种用以定性及定量检测的技术)来分析唾液中蛋白(胜肽)的含量。本研究可侦测28种候选蛋白标志,且侦测浓度范围为每毫升1纳克到每毫升2000纳克。侦测限制与抗体(例如使用于ELISA的抗体)

的专一性一样好,唯LC-MRM-MS可避免因偏离标的抗体作用所引发的偏差结果。

[0200] 总结上述,早期侦测口腔鳞状细胞癌对于成功及具成本效益的疾病控制,以及病患的管理相当重要。目前相关领域仍缺乏可用以临床诊断或监控口腔鳞状细胞癌的分子标志。本揭示内容阐述了具临床应用性的唾液蛋白生物标志分析组的研发及确认结果,可用以早期侦测口腔鳞状细胞癌及监控口腔高恶性潜能病变病患。本研究近期将进行临床试验,可由另外二家医院收集更大量的口腔鳞状细胞癌病患及口腔高恶性潜能病变病患。

[0201] 虽然上文实施方式中揭露了本发明的具体实施例,然其并非用以限定本发明,本发明所属技术领域中具有通常知识者,在不悖离本发明的原理与精神的情形下,当可对其进行各种更动与修饰,因此本发明的保护范围当以附随申请专利范围所界定者为准。

序列表

<110> 长庚大学
长庚医疗财团法人林口长庚纪念医院

<120> 一种用以诊断及预断癌症的方法

<130> PC197031

<150> US62309766

<151> 2016-03-17

<160> 2

<170> SIPOSequenceListing 1.0

<210> 1

<211> 339

<212> PRT

<213> ANXA2多肽(Artificial Sequence)

<400> 1

[0001]

```

Met Ser Thr Val His Glu Ile Leu Cys Lys Leu Ser Leu Glu Gly Asp
1          5          10          15
His Ser Thr Pro Pro Ser Ala Tyr Gly Ser Val Lys Ala Tyr Thr Asn
20
Phe Asp Ala Glu Arg Asp Ala Leu Asn Ile Glu Thr Ala Ile Lys Thr
35          40          45
Lys Gly Val Asp Glu Val Thr Ile Val Asn Ile Leu Thr Asn Arg Ser
50          55          60
Asn Ala Gln Arg Gln Asp Ile Ala Phe Ala Tyr Gln Arg Arg Thr Lys
65          70          75          80
Lys Glu Leu Ala Ser Ala Leu Lys Ser Ala Leu Ser Gly His Leu Glu
85          90          95
Thr Val Ile Leu Gly Leu Leu Lys Thr Pro Ala Gln Tyr Asp Ala Ser
100          105          110
Glu Leu Lys Ala Ser Met Lys Gly Leu Gly Thr Asp Glu Asp Ser Leu
115          120          125
Ile Glu Ile Ile Cys Ser Arg Thr Asn Gln Glu Leu Gln Glu Ile Asn
130          135          140
Arg Val Tyr Lys Glu Met Tyr Lys Thr Asp Leu Glu Lys Asp Ile Ile
145          150          155          160
Ser Asp Thr Ser Gly Asp Phe Arg Lys Leu Met Val Ala Leu Ala Lys
165          170          175
Gly Arg Arg Ala Glu Asp Gly Ser Val Ile Asp Tyr Glu Leu Ile Asp
180          185          190
Gln Asp Ala Arg Asp Leu Tyr Asp Ala Gly Val Lys Arg Lys Gly Thr
195          200          205
Asp Val Pro Lys Trp Ile Ser Ile Met Thr Glu Arg Ser Val Pro His
210          215          220
Leu Gln Lys Val Phe Asp Arg Tyr Lys Ser Tyr Ser Pro Tyr Asp Met
225          230          235          240
Leu Glu Ser Ile Arg Lys Glu Val Lys Gly Asp Leu Glu Asn Ala Phe
245          250          255
Leu Asn Leu Val Gln Cys Ile Gln Asn Lys Pro Leu Tyr Phe Ala Asp
260          265          270

```

Arg Leu Tyr Asp Ser Met Lys Gly Lys Gly Thr Arg Asp Lys Val Leu
 275 280 285
 Ile Arg Ile Met Val Ser Arg Ser Glu Val Asp Met Leu Lys Ile Arg
 290 295 300
 Ser Glu Phe Lys Arg Lys Tyr Gly Lys Ser Leu Tyr Tyr Tyr Ile Gln
 305 310 315 320
 Gln Asp Thr Lys Gly Asp Tyr Gln Lys Ala Leu Leu Tyr Leu Cys Gly
 325 330 335
 Gly Asp Asp

<210> 2
 <211> 654
 <212> PRT
 <213> HSPA5多肽(Artificial Sequence)

<400> 2
 Met Lys Leu Ser Leu Val Ala Ala Met Leu Leu Leu Leu Ser Ala Ala
 1 5 10 15
 Arg Ala Glu Glu Glu Asp Lys Lys Glu Asp Val Gly Thr Val Val Gly
 20 25 30
 Ile Asp Leu Gly Thr Thr Tyr Ser Cys Val Gly Val Phe Lys Asn Gly
 35 40 45
 Arg Val Glu Ile Ile Ala Asn Asp Gln Gly Asn Arg Ile Thr Pro Ser
 50 55 60
 Tyr Val Ala Phe Thr Pro Glu Gly Glu Arg Leu Ile Gly Asp Ala Ala
 65 70 75 80
 Lys Asn Gln Leu Thr Ser Asn Pro Glu Asn Thr Val Phe Asp Ala Lys
 85 90 95
 Arg Leu Ile Gly Arg Thr Trp Asn Asp Pro Ser Val Gln Gln Asp Ile
 100 105 110
 Lys Phe Leu Pro Phe Lys Val Val Glu Lys Lys Thr Lys Pro Tyr Ile
 115 120 125
 Gln Val Asp Ile Gly Gly Gly Gln Thr Lys Thr Phe Ala Pro Glu Glu
 130 135 140
 Ile Ser Ala Met Val Leu Thr Lys Met Lys Glu Thr Ala Glu Ala Tyr
 145 150 155 160
 Leu Gly Lys Lys Val Thr His Ala Val Val Thr Val Pro Ala Tyr Phe
 165 170 175
 Asn Asp Ala Gln Arg Gln Ala Thr Lys Asp Ala Gly Thr Ile Ala Gly
 180 185 190
 Leu Asn Val Met Arg Ile Ile Asn Glu Pro Thr Ala Ala Ala Ile Ala
 195 200 205
 Tyr Gly Leu Asp Lys Arg Glu Gly Glu Lys Asn Ile Leu Val Phe Asp
 210 215 220
 Leu Gly Gly Gly Thr Phe Asp Val Ser Leu Leu Thr Ile Asp Asn Gly
 225 230 235 240
 Val Phe Glu Val Val Ala Thr Asn Gly Asp Thr His Leu Gly Gly Glu
 245 250 255
 Asp Phe Asp Gln Arg Val Met Glu His Phe Ile Lys Leu Tyr Lys Lys
 260 265 270
 Lys Thr Gly Lys Asp Val Arg Lys Asp Asn Arg Ala Val Gln Lys Leu
 275 280 285
 Arg Arg Glu Val Glu Lys Ala Lys Arg Ala Leu Ser Ser Gln His Gln
 290 295 300
 Ala Arg Ile Glu Ile Glu Ser Phe Tyr Glu Gly Glu Asp Phe Ser Glu

[0002]

305	310	315	320
Thr Leu Thr Arg	Ala Lys Phe Glu Glu	Leu Asn Met Asp	Leu Phe Arg
	325	330	335
Ser Thr Met Lys	Pro Val Gln Lys	Val Leu Glu Asp	Ser Asp Leu Lys
	340	345	350
Lys Ser Asp Ile	Asp Glu Ile Val	Leu Val Gly Gly	Ser Thr Arg Ile
	355	360	365
Pro Lys Ile Gln	Gln Leu Val Lys	Glu Phe Phe Asn	Gly Lys Glu Pro
	370	375	380
Ser Arg Gly Ile	Asn Pro Asp Glu	Ala Val Ala Tyr	Gly Ala Ala Val
385	390	395	400
Gln Ala Gly Val	Leu Ser Gly Asp	Gln Asp Thr Gly	Asp Leu Val Leu
	405	410	415
Leu Asp Val Cys	Pro Leu Thr Leu	Gly Ile Glu Thr	Val Gly Gly Val
	420	425	430
Met Thr Lys Leu	Ile Pro Arg Asn	Thr Val Val Pro	Thr Lys Lys Ser
	435	440	445
Gln Ile Phe Ser	Thr Ala Ser Asp	Asn Gln Pro Thr	Val Thr Ile Lys
	450	455	460
Val Tyr Glu Gly	Glu Arg Pro Leu	Thr Lys Asp Asn	His Leu Leu Gly
465	470	475	480
Thr Phe Asp Leu	Thr Gly Ile Pro	Pro Ala Pro Arg	Gly Val Pro Gln
	485	490	495
Ile Glu Val Thr	Phe Glu Ile Asp	Val Asn Gly Ile	Leu Arg Val Thr
	500	505	510
Ala Glu Asp Lys	Gly Thr Gly Asn	Lys Asn Lys Ile	Thr Ile Thr Asn
	515	520	525
Asp Gln Asn Arg	Leu Thr Pro Glu	Glu Ile Glu Arg	Met Val Asn Asp
	530	535	540
Ala Glu Lys Phe	Ala Glu Asp Lys	Lys Leu Lys Glu	Arg Ile Asp
545	550	555	560
Thr Arg Asn Glu	Leu Glu Ser Tyr	Ala Tyr Ser Leu	Lys Asn Gln Ile
	565	570	575
Gly Asp Lys Glu	Lys Leu Gly Gly	Lys Leu Ser Ser	Glu Asp Lys Glu
	580	585	590
Thr Met Glu Lys	Ala Val Glu Glu	Lys Ile Glu Trp	Leu Glu Ser His
	595	600	605
Gln Asp Ala Asp	Ile Glu Asp Phe	Lys Ala Lys Lys	Lys Glu Leu Glu
	610	615	620
Glu Ile Val Gln	Pro Ile Ile Ser	Lys Leu Tyr Gly	Ser Ala Gly Pro
625	630	635	640
Pro Pro Thr Gly	Glu Glu Asp Thr	Ala Glu Lys Asp	Glu Leu
	645	650	

<210> 3

<211> 644

<212> PRT

<213> KNG1多肽(Artificial Sequence)

<400> 3

Met Lys Leu Ile	Thr Ile Leu Phe	Leu Cys Ser Arg	Leu Leu Leu Ser
1	5	10	15
Leu Thr Gln Glu	Ser Gln Ser Glu	Glu Ile Asp Cys	Asn Asp Lys Asp
	20	25	30
Leu Phe Lys Ala	Val Asp Ala Ala	Leu Lys Lys Tyr	Asn Ser Gln Asn
	35	40	45

[0004]	Gln	Ser	Asn	Asn	Gln	Phe	Val	Leu	Tyr	Arg	Ile	Thr	Glu	Ala	Thr	Lys
	50						55					60				
	Thr	Val	Gly	Ser	Asp	Thr	Phe	Tyr	Ser	Phe	Lys	Tyr	Glu	Ile	Lys	Glu
	65					70					75					80
	Gly	Asp	Cys	Pro	Val	Gln	Ser	Gly	Lys	Thr	Trp	Gln	Asp	Cys	Glu	Tyr
					85					90					95	
	Lys	Asp	Ala	Ala	Lys	Ala	Ala	Thr	Gly	Glu	Cys	Thr	Ala	Thr	Val	Gly
					100				105					110		
	Lys	Arg	Ser	Ser	Thr	Lys	Phe	Ser	Val	Ala	Thr	Gln	Thr	Cys	Gln	Ile
			115					120					125			
	Thr	Pro	Ala	Glu	Gly	Pro	Val	Val	Thr	Ala	Gln	Tyr	Asp	Cys	Leu	Gly
		130					135					140				
	Cys	Val	His	Pro	Ile	Ser	Thr	Gln	Ser	Pro	Asp	Leu	Glu	Pro	Ile	Leu
	145					150					155					160
	Arg	His	Gly	Ile	Gln	Tyr	Phe	Asn	Asn	Asn	Thr	Gln	His	Ser	Ser	Leu
					165					170					175	
	Phe	Met	Leu	Asn	Glu	Val	Lys	Arg	Ala	Gln	Arg	Gln	Val	Val	Ala	Gly
				180					185					190		
	Leu	Asn	Phe	Arg	Ile	Thr	Tyr	Ser	Ile	Val	Gln	Thr	Asn	Cys	Ser	Lys
			195					200					205			
	Glu	Asn	Phe	Leu	Phe	Leu	Thr	Pro	Asp	Cys	Lys	Ser	Leu	Trp	Asn	Gly
		210					215					220				
	Asp	Thr	Gly	Glu	Cys	Thr	Asp	Asn	Ala	Tyr	Ile	Asp	Ile	Gln	Leu	Arg
	225					230					235					240
	Ile	Ala	Ser	Phe	Ser	Gln	Asn	Cys	Asp	Ile	Tyr	Pro	Gly	Lys	Asp	Phe
					245					250					255	
	Val	Gln	Pro	Pro	Thr	Lys	Ile	Cys	Val	Gly	Cys	Pro	Arg	Asp	Ile	Pro
				260					265					270		
	Thr	Asn	Ser	Pro	Glu	Leu	Glu	Glu	Thr	Leu	Thr	His	Thr	Ile	Thr	Lys
			275					280					285			
	Leu	Asn	Ala	Glu	Asn	Asn	Ala	Thr	Phe	Tyr	Phe	Lys	Ile	Asp	Asn	Val
		290				295						300				
	Lys	Lys	Ala	Arg	Val	Gln	Val	Val	Ala	Gly	Lys	Lys	Tyr	Phe	Ile	Asp
	305					310					315					320
	Phe	Val	Ala	Arg	Glu	Thr	Thr	Cys	Ser	Lys	Glu	Ser	Asn	Glu	Glu	Leu
					325					330					335	
	Thr	Glu	Ser	Cys	Glu	Thr	Lys	Lys	Leu	Gly	Gln	Ser	Leu	Asp	Cys	Asn
				340					345					350		
	Ala	Glu	Val	Tyr	Val	Val	Pro	Trp	Glu	Lys	Lys	Ile	Tyr	Pro	Thr	Val
				355				360					365			
	Asn	Cys	Gln	Pro	Leu	Gly	Met	Ile	Ser	Leu	Met	Lys	Arg	Pro	Pro	Gly
		370					375					380				
	Phe	Ser	Pro	Phe	Arg	Ser	Ser	Arg	Ile	Gly	Glu	Ile	Lys	Glu	Glu	Thr
	385					390					395					400
	Thr	Val	Ser	Pro	Pro	His	Thr	Ser	Met	Ala	Pro	Ala	Gln	Asp	Glu	Glu
					405					410					415	
	Arg	Asp	Ser	Gly	Lys	Glu	Gln	Gly	His	Thr	Arg	Arg	His	Asp	Trp	Gly
				420					425					430		
	His	Glu	Lys	Gln	Arg	Lys	His	Asn	Leu	Gly	His	Gly	His	Lys	His	Glu
			435					440					445			
	Arg	Asp	Gln	Gly	His	Gly	His	Gln	Arg	Gly	His	Gly	Leu	Gly	His	Gly
		450				455						460				
	His	Glu	Gln	Gln	His	Gly	Leu	Gly	His	Gly	His	Lys	Phe	Lys	Leu	Asp
	465					470					475					480
	Asp	Asp	Leu	Glu	His	Gln	Gly	Gly	His	Val	Leu	Asp	His	Gly	His	Lys
					485					490					495	

His Lys His Gly His Gly His Gly Lys His Lys Asn Lys Gly Lys Lys
 500 505 510
 Asn Gly Lys His Asn Gly Trp Lys Thr Glu His Leu Ala Ser Ser Ser
 515 520 525
 Glu Asp Ser Thr Thr Pro Ser Ala Gln Thr Gln Glu Lys Thr Glu Gly
 530 535 540
 Pro Thr Pro Ile Pro Ser Leu Ala Lys Pro Gly Val Thr Val Thr Phe
 545 550 555 560
 Ser Asp Phe Gln Asp Ser Asp Leu Ile Ala Thr Met Met Pro Pro Ile
 565 570 575
 Ser Pro Ala Pro Ile Gln Ser Asp Asp Asp Trp Ile Pro Asp Ile Gln
 580 585 590
 Ile Asp Pro Asn Gly Leu Ser Phe Asn Pro Ile Ser Asp Phe Pro Asp
 595 600 605
 Thr Thr Ser Pro Lys Cys Pro Gly Arg Pro Trp Lys Ser Val Ser Glu
 610 615 620
 Ile Asn Pro Thr Thr Gln Met Lys Glu Ser Tyr Tyr Phe Asp Leu Thr
 625 630 635 640
 Asp Gly Leu Ser

<210> 4
 <211> 469
 <212> PRT
 <213> MMP1多肽(Artificial Sequence)

[0005] <400> 4
 Met His Ser Phe Pro Pro Leu Leu Leu Leu Leu Phe Trp Gly Val Val
 1 5 10 15
 Ser His Ser Phe Pro Ala Thr Leu Glu Thr Gln Glu Gln Asp Val Asp
 20 25 30
 Leu Val Gln Lys Tyr Leu Glu Lys Tyr Tyr Asn Leu Lys Asn Asp Gly
 35 40 45
 Arg Gln Val Glu Lys Arg Arg Asn Ser Gly Pro Val Val Glu Lys Leu
 50 55 60
 Lys Gln Met Gln Glu Phe Gly Leu Lys Val Thr Gly Lys Pro Asp
 65 70 75 80
 Ala Glu Thr Leu Lys Val Met Lys Gln Pro Arg Cys Gly Val Pro Asp
 85 90 95
 Val Ala Gln Phe Val Leu Thr Glu Gly Asn Pro Arg Trp Glu Gln Thr
 100 105 110
 His Leu Thr Tyr Arg Ile Glu Asn Tyr Thr Pro Asp Leu Pro Arg Ala
 115 120 125
 Asp Val Asp His Ala Ile Glu Lys Ala Phe Gln Leu Trp Ser Asn Val
 130 135 140
 Thr Pro Leu Thr Phe Thr Lys Val Ser Glu Gly Gln Ala Asp Ile Met
 145 150 155 160
 Ile Ser Phe Val Arg Gly Asp His Arg Asp Asn Ser Pro Phe Asp Gly
 165 170 175
 Pro Gly Gly Asn Leu Ala His Ala Phe Gln Pro Gly Pro Gly Ile Gly
 180 185 190
 Gly Asp Ala His Phe Asp Glu Asp Glu Arg Trp Thr Asn Asn Phe Arg
 195 200 205
 Glu Tyr Asn Leu His Arg Val Ala Ala His Glu Leu Gly His Ser Leu
 210 215 220
 Gly Leu Ser His Ser Thr Asp Ile Gly Ala Leu Met Tyr Pro Ser Tyr

	225				230					235				240		
	Thr	Phe	Ser	Gly	Asp	Val	Gln	Leu	Ala	Gln	Asp	Asp	Ile	Asp	Gly	Ile
					245					250					255	
	Gln	Ala	Ile	Tyr	Gly	Arg	Ser	Gln	Asn	Pro	Val	Gln	Pro	Ile	Gly	Pro
				260					265					270		
	Gln	Thr	Pro	Lys	Ala	Cys	Asp	Ser	Lys	Leu	Thr	Phe	Asp	Ala	Ile	Thr
			275					280					285			
	Thr	Ile	Arg	Gly	Glu	Val	Met	Phe	Phe	Lys	Asp	Arg	Phe	Tyr	Met	Arg
		290					295					300				
	Thr	Asn	Pro	Phe	Tyr	Pro	Glu	Val	Glu	Leu	Asn	Phe	Ile	Ser	Val	Phe
	305					310					315				320	
	Trp	Pro	Gln	Leu	Pro	Asn	Gly	Leu	Glu	Ala	Ala	Tyr	Glu	Phe	Ala	Asp
				325						330					335	
	Arg	Asp	Glu	Val	Arg	Phe	Phe	Lys	Gly	Asn	Lys	Tyr	Trp	Ala	Val	Gln
				340					345					350		
	Gly	Gln	Asn	Val	Leu	His	Gly	Tyr	Pro	Lys	Asp	Ile	Tyr	Ser	Ser	Phe
			355					360					365			
	Gly	Phe	Pro	Arg	Thr	Val	Lys	His	Ile	Asp	Ala	Ala	Leu	Ser	Glu	Glu
		370					375					380				
	Asn	Thr	Gly	Lys	Thr	Tyr	Phe	Phe	Val	Ala	Asn	Lys	Tyr	Trp	Arg	Tyr
	385					390					395				400	
	Asp	Glu	Tyr	Lys	Arg	Ser	Met	Asp	Pro	Gly	Tyr	Pro	Lys	Met	Ile	Ala
				405						410					415	
	His	Asp	Phe	Pro	Gly	Ile	Gly	His	Lys	Val	Asp	Ala	Val	Phe	Met	Lys
				420					425					430		
	Asp	Gly	Phe	Phe	Tyr	Phe	Phe	His	Gly	Thr	Arg	Gln	Tyr	Lys	Phe	Asp
			435					440					445			
[0006]	Pro	Lys	Thr	Lys	Arg	Ile	Leu	Thr	Leu	Gln	Lys	Ala	Asn	Ser	Trp	Phe
		450					455					460				
	Asn	Cys	Arg	Lys	Asn											
	465															
	<210>	5														
	<211>	9														
	<212>	PRT														
	<213>	ANXA2的SIS胜肽(Artificial Sequence)														
	<400>	5														
	Gln	Asp	Ile	Ala	Phe	Ala	Tyr	Gln	Arg							
	1					5										
	<210>	6														
	<211>	14														
	<212>	PRT														
	<213>	HSPA5的SIS胜肽(Artificial Sequence)														
	<400>	6														
	Ile	Thr	Pro	Ser	Tyr	Val	Ala	Phe	Thr	Pro	Glu	Gly	Glu	Arg		
	1					5					10					
	<210>	7														
	<211>	11														
	<212>	PRT														
	<213>	KNG1的SIS胜肽(Artificial Sequence)														
	<400>	7														

	Thr	Val	Gly	Ser	Asp	Thr	Phe	Tyr	Ser	Phe	Lys
	1				5					10	

[0007] <210> 8
<211> 10
<212> PRT
<213> MMP1的SIS胜肽(Artificial Sequence)

<400> 8
Asp Ile Tyr Ser Ser Phe Gly Phe Pro Arg
1 5 10

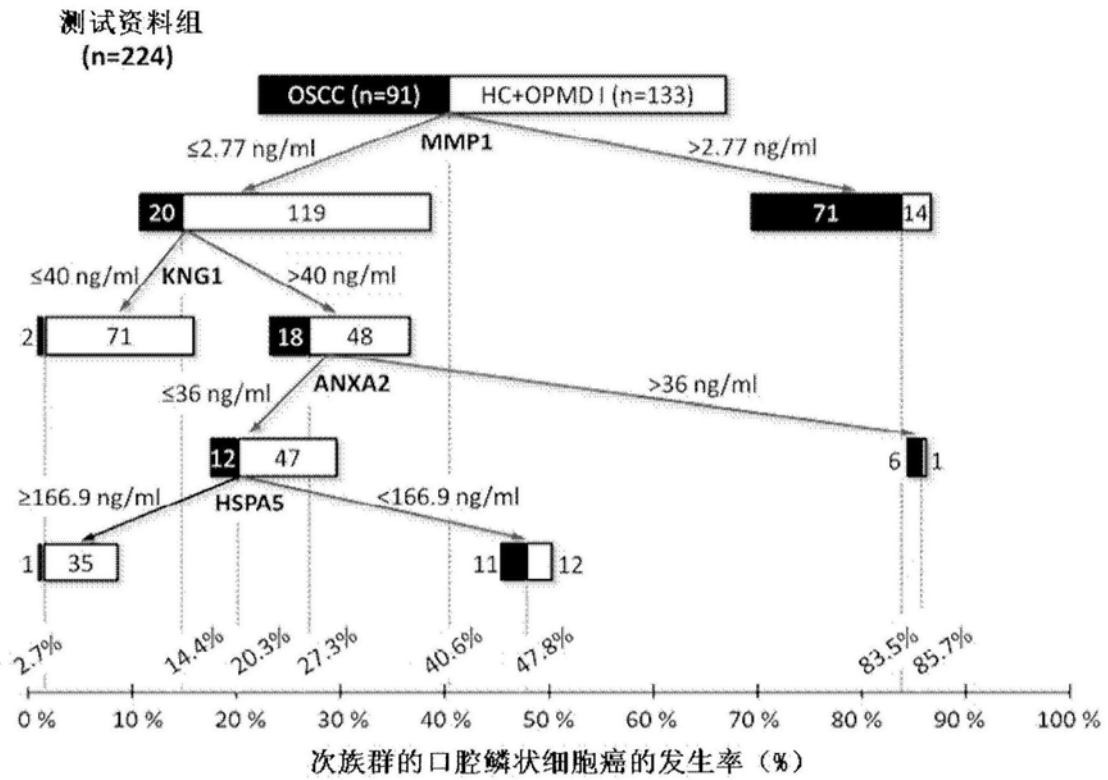


图1A

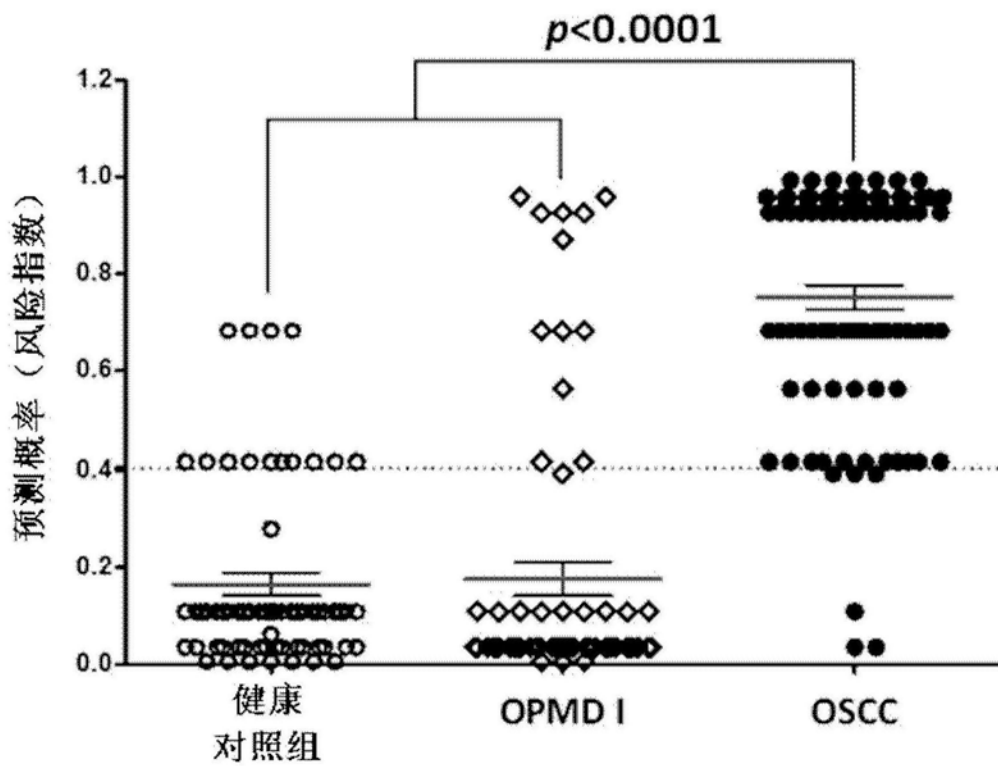


图1B

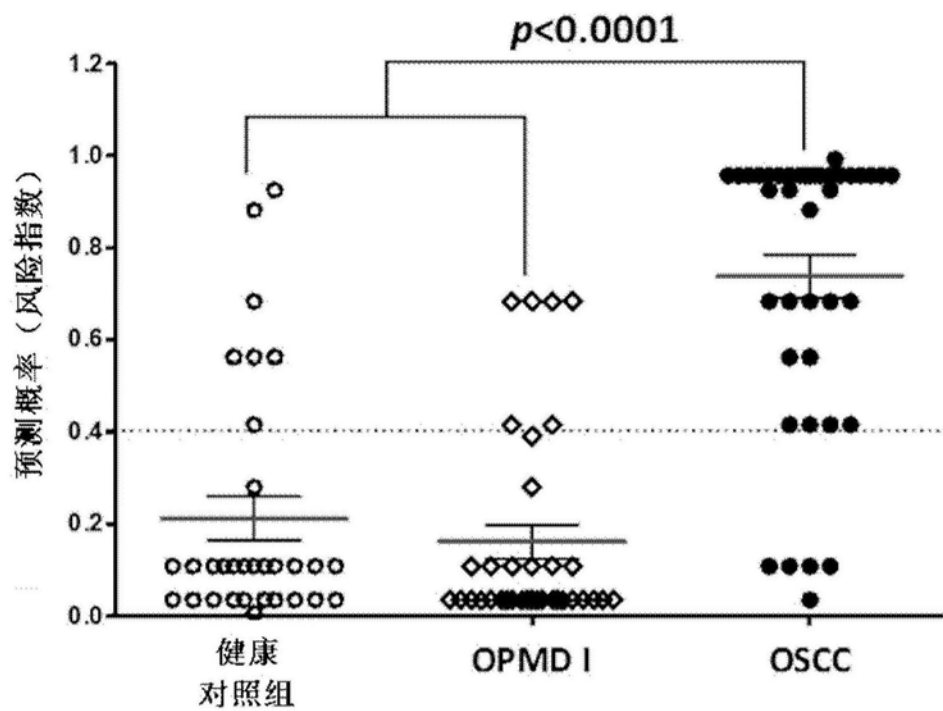


图1C

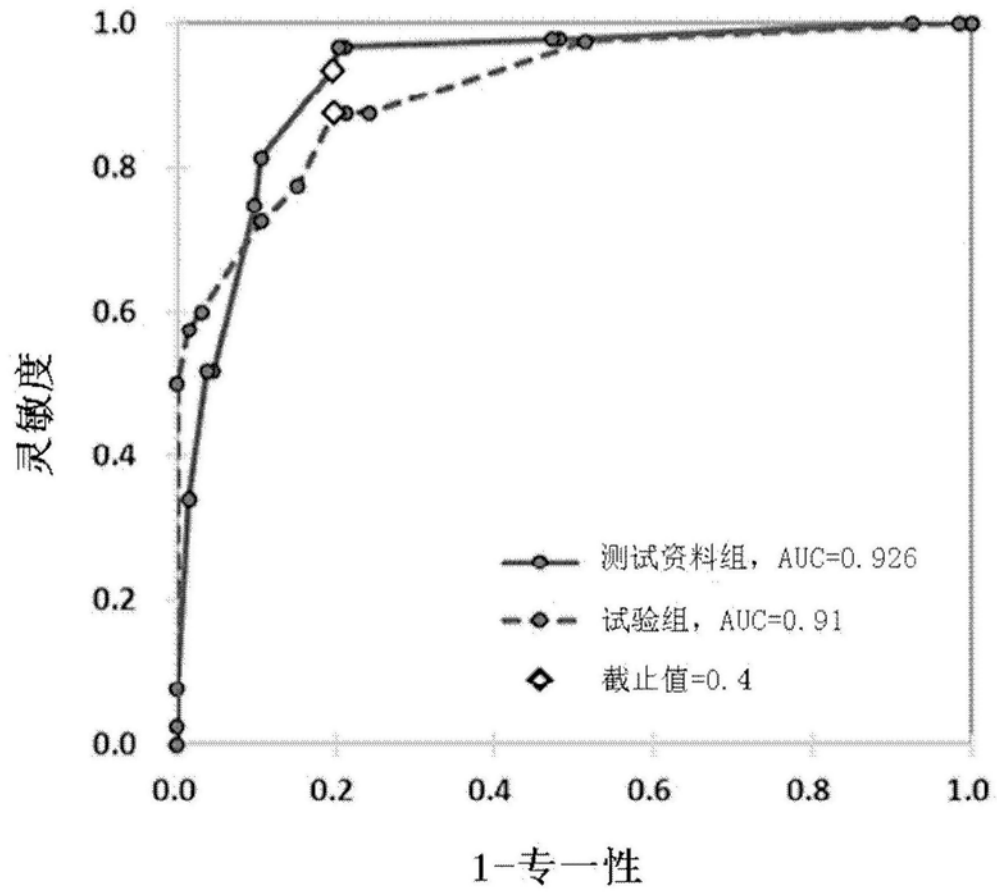


图1D

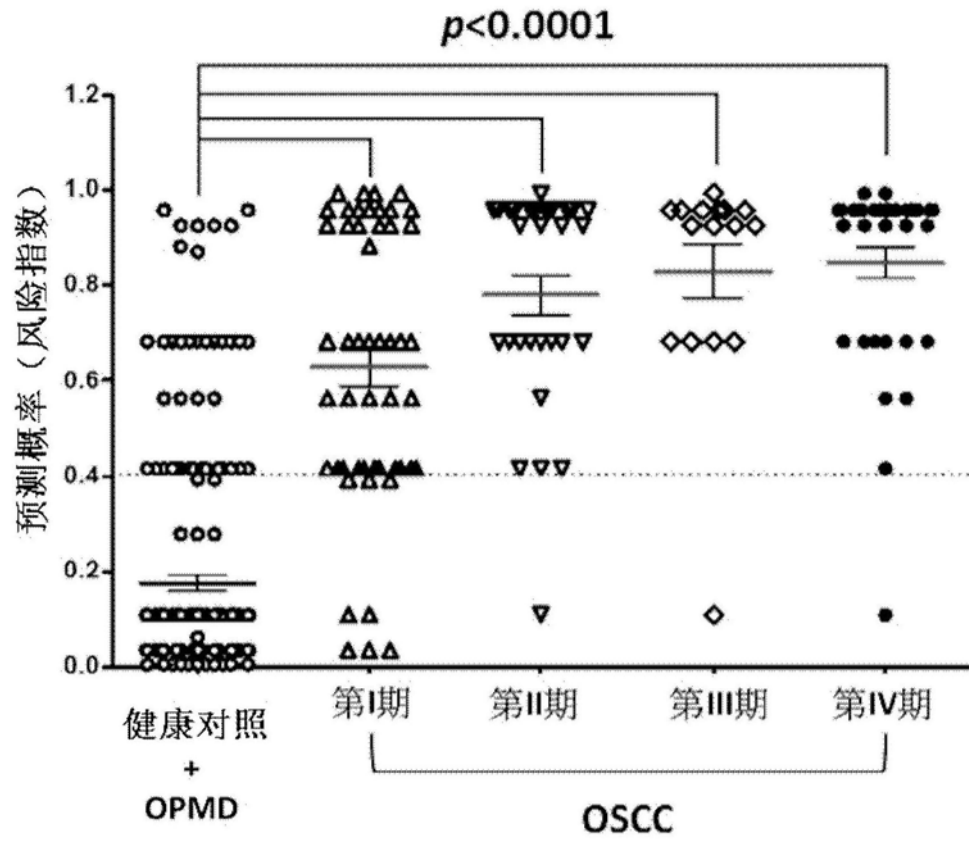


图2

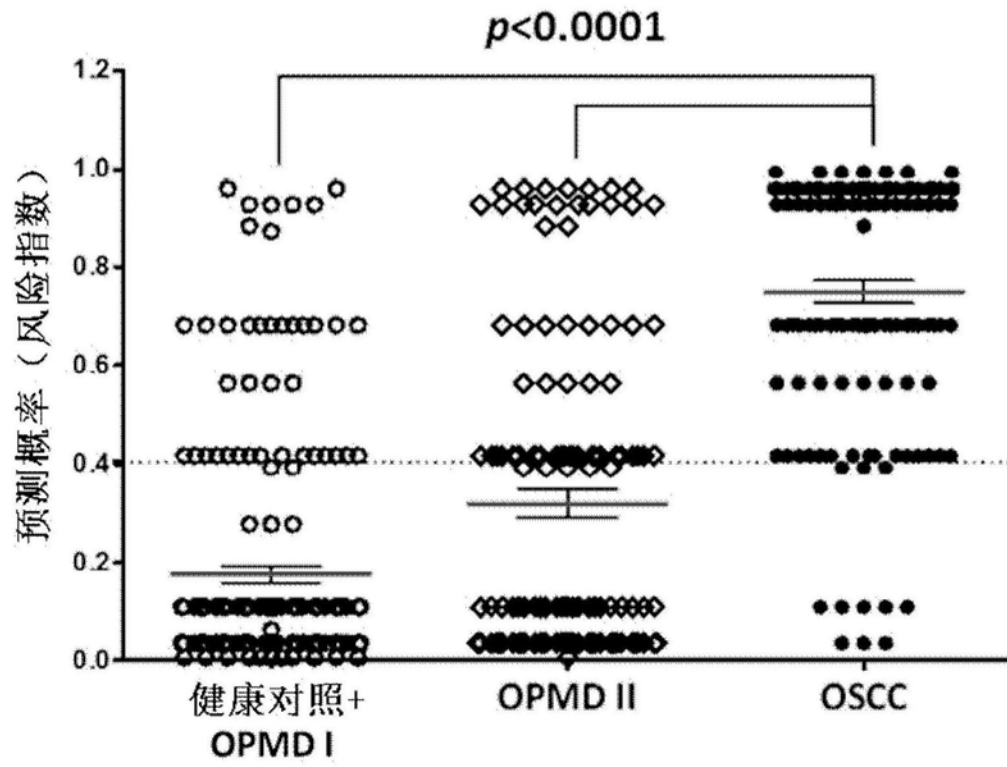


图3

专利名称(译)	一种用以诊断及预断癌症的方法		
公开(公告)号	CN109863401A	公开(公告)日	2019-06-07
申请号	CN201780016684.6	申请日	2017-03-17
[标]申请(专利权)人(译)	长庚大学 长庚医疗财团法人林口长庚纪念医院		
申请(专利权)人(译)	长庚大学 长庚医疗财团法人林口长庚纪念医院		
当前申请(专利权)人(译)	长庚大学 长庚医疗财团法人林口长庚纪念医院		
[标]发明人	余兆松 陈怡婷 蒋维凡 萧永晋 史丽珠 张玉生 张凯评		
发明人	余兆松 陈怡婷 蒋维凡 萧永晋 史丽珠 张玉生 张凯评		
IPC分类号	G01N33/574 G01N33/53 C07K14/00		
优先权	62/309766 2016-03-17 US		
外部链接	Espacenet SIPO		

摘要(译)

本揭示内容是关于一种用以决定一个体是否罹患癌症或有罹患癌症的风险的方法。该方法包含：由该个体取得一检体；决定至少二标的多肽的含量，该至少二标的多肽选自由ANXA2、HSPA5、KNG1及MMP1所组成的群组；以及基于该至少二标的多肽的含量来评估该个体是否罹患癌症或有罹患癌症的风险。本发明方法可用于诊断及预断口腔鳞状细胞癌的发生，据此及时对有需要的个体授予适当的治疗。

