



(12) 发明专利申请

(10) 申请公布号 CN 113994437 A

(43) 申请公布日 2022. 01. 28

(21) 申请号 202080041388.3

(22) 申请日 2020.06.04

(30) 优先权数据

19305730.4 2019.06.05 EP

(85) PCT国际申请进入国家阶段日

2021.12.03

(86) PCT国际申请的申请数据

PCT/EP2020/065562 2020.06.04

(87) PCT国际申请的公布数据

W02020/245322 EN 2020.12.10

(71) 申请人 法国公立援助医院

地址 法国巴黎

申请人 国家健康科学研究所 巴黎大学

发育研究院 索邦大学

(72) 发明人 J-E·萨朗 E·普里夫蒂

A·A·普利尼 J-D·楚克尔

C·丰克-布伦塔诺 A·莱纳特

I·当茹瓦 F·克塞特尔默阿纳

(74) 专利代理机构 北京戈程知识产权代理有限公司 11314

代理人 程伟 王锦阳

(51) Int.Cl.

G16H 50/20 (2006.01)

(54) 发明名称

用于检测尖端扭转型室性心动过速的风险的方法

(57) 摘要

本发明涉及用于检测和预测患者患有尖端扭转型室性心动过速事件的风险的方法和装置，具体地经由利用神经网络来检测和预测。

权利要求书3页 说明书18页 附图2页

1. 一种用于评估患者患有尖端扭转型室性心动过速事件的风险的方法, 包括:
 - a. 由处理装置接收表示受试患者的心电图波形的时间片段的信号数据;
 - b. 由处理装置经由配置的人工机器学习分类器分析所述信号数据以生成作为人工神经网络的输出的可能性;其中, 所述可能性与患者患有尖端扭转型室性心动过速事件的风险相关。
2. 一种用于检测给定患者存在尖端扭转型室性心动过速的发作性症状的风险的方法, 包括以下步骤:
 - a. 从患者获得心电图数据;
 - b. 将心电图数据应用于机器学习分类器, 所述机器学习分类器配置为对指示患者患有尖端扭转型室性心动过速事件的风险的心电图数据中的变化进行检测,
 - c. 从机器学习分类器获得输出, 其中输出提供了患者患有尖端扭转型室性心动过速事件的风险的可能性。
3. 根据权利要求2所述的方法, 其中, 将心电图数据发送到远程机器学习分类器, 并且将输出发送到患者和/或医生。
4. 根据权利要求1至3中任一项所述的方法, 其中, 患者具有在48小时内患有尖端扭转型室性心动过速事件的风险。
5. 根据权利要求1至4中任一项所述的方法, 其中, 患者具有在24小时内患有尖端扭转型室性心动过速事件的风险。
6. 一种用于产生能够评估患者患有尖端扭转型室性心动过速事件的风险和其基础机制的机器学习分类器的方法, 包括:
 - a. 将包括第一参数、第二参数以及第三参数的患者数据存储在电子数据库中, 所述第一参数是来自患者的心电图, 所述第二参数与患者患有尖端扭转型室性心动过速事件的风险相关, 所述第三参数与其原因相关;
 - b. 设置机器学习系统; 并且
 - c. 利用患者数据训练机器学习系统, 使得机器学习系统经过训练以在接触到来自患者的心电图时产生对患者患有尖端扭转型室性心动过速事件的风险和其原因的预测。
7. 一种用于产生能够评估患者患有尖端扭转型室性心动过速事件的风险和其原因的人工神经网络的方法, 包括:
 - a. 将包括第一参数、第二参数以及第三参数的患者数据存储在电子数据库中, 所述第一参数是来自患者的心电图, 所述第二参数与患者患有尖端扭转型室性心动过速事件的风险相关, 所述第三参数与其原因相关;
 - b. 设置互连以形成人工神经网络的节点的网络, 所述节点包括多个人工神经元, 每个人工神经元具有带相关联的权重的至少一个输入; 并且
 - c. 利用患者数据训练人工神经网络, 使得多个人工神经元的每个人工神经元的至少一个输入的相关联的权重响应于来自患者数据的多个不同数据集的相应第一参数、第二参数以及第三参数而调整, 使得人工神经网络经过训练以在接触到来自患者的心电图时产生对患者患有尖端扭转型室性心动过速事件的风险和其原因的预测。
8. 根据权利要求6或7所述的方法, 其中, 患者数据包括来自已经施用QT延长药物的患者的心电图和来自尚未施用QT延长药物的患者的心电图。

9. 一种用于确定物质在向患者施用之后诱发尖端扭转型室速效果的风险的体外方法，其包括以下步骤：

- a. 在施用所述组合物之后从患者获得心电图数据；
- b. 将心电图数据应用于机器学习分类器，所述机器学习分类器配置为对指示尖端扭转型室性心动过速事件的风险增加的心电图数据中的变化进行检测，
- c. 从机器学习分类器获得输出，其中所述输出提供患者患有尖端扭转型室性心动过速事件的风险，

其中，如果在施用物质之后获得患者患有尖端扭转型室性心动过速事件的风险，那么物质存在诱发尖端扭转型室速效应的风险。

10. 根据权利要求9所述的方法，其在含有大于或等于10的患者数目的患者群组上重复。

11. 一种用于确定患者中的先天性长QT综合症的本质的方法，包括以下步骤：

- a. 从患者获得心电图数据；
- b. 将心电图数据应用于机器学习分类器，所述机器学习分类器配置为对指示尖端扭转型室性心动过速事件的风险的心电图数据中的变化进行检测；
- c. 从机器学习分类器获得输出，其中所述输出提供长QT的本质的可能性且无论先天性长QT综合症是LQT2综合症还是LQT1或LQT3综合症。

12. 根据权利要求1至6以及8至11中任一项所述的方法，其中，所述机器学习分类器是神经网络，优选地是卷积神经网络。

13. 根据权利要求1至5以及8至12中任一项所述的方法，其中，通过以下步骤将患者指派到类别

- a. 利用来自相同患者的不同心电图信号重复方法；并且
- b. 将患者指派在输出的大多数指示最高概率的类别中。

14. 根据权利要求1至5以及8至12中任一项所述的方法，其中，通过以下步骤将患者指派到类别

- a. 使用根据权利要求3或4所述的方法获得的不同机器学习分类器，利用来自患者的心电图信号重复方法；并且
- b. 将患者指派在输出中的大多数指示最高概率的类别中。

15. 根据权利要求1至14中任一项所述的方法，其中，心电图信号来自一个单导联的心电图信号，优选地，单导联是DII导联、V4导联或V3导联。

16. 根据权利要求1至15中任一项所述的方法，其中，心电图信号来自多于一个导联的心电图信号，优选地至少3个导联，更优选地至少6个导联或至少8个导联。

17. 一种心脏监测装置，包括：

- a. 至少一个电极，其用于从患者获得心电图 (ECG) 信号；
- b. 处理单元，其包括在操作上联接到至少一个电极的至少一个处理器；以及
- c. 至少一个非易失性计算机可读介质，其包括程序指令，所述程序指令在由至少一个处理器执行时使得心脏监测装置：
 - i. 从至少一个电极获得心电图信号；
 - ii. 任选地将心电图信号标准化/归一化以获得标准化/归一化心电图信号；

iii. 任选地从对应于预定持续时间的标准化心电图信号提取数据以便获得提取到的标准化心电图信号;

iv. 将任选地标准化/归一化并提取的心电图信号提供到机器学习分类器;

由此从机器学习分类器获得输出,所述输出是患有尖端扭转型室性心动过速事件的可能性和其原因。

18. 一种用于评价患者患有尖端扭转型室性心动过速事件的风险的系统,其包括:

a. 体外除颤器,其配置为监测来自患者的心电图数据;

b. 一个或多个处理器,其配置为执行操作,操作包括:

i. 任选地执行用于给定时域的心电图数据的数学标准化;

ii. 通过机器学习分类器执行计算;并且

iii. 确定患者患有尖端扭转型室性心动过速事件的可能性和其原因。

用于检测尖端扭转型室性心动过速的风险的方法

技术领域

[0001] 本发明涉及检测患者患有尖端扭转型室性心动过速事件的风险以及这种风险之下的机制,具体地经由利用神经网络来检测。

背景技术

[0002] 尖端扭转型室性心动过速(Torsades de Pointes,TdP)是可以恶化为心室颤动和猝死的心律失常。其基本上在先天性长QT综合症中发生或因QT延长心肌和非心肌药物的罕见副作用而发生。

[0003] 当前检测策略基于QT测量(即,在QRS复合波的开端与T波的末端之间的间期,传统地,通常利用心电图(electrocardiogram,ECG)的导联II)。TdP风险是药物行业的主要问题,并且可以是市场撤回药物的原因。

[0004] 另一方法被描述在US 20190059764中,其公开了利用相对于心率校正的QT间期、T波的第一峰值的最大幅度以及T波的最大第一峰值与T波的末端之间的间期的算法。尽管这种算法是高效的,但是这种算法可能并非总是易于实施,这是由于其需要将摄入QT延长药物前后的ECG进行比较。

[0005] 寻找用于量化根据药物摄入产生TdP的风险的新颖方法很重要。实际上,仅查看QT不足以检测延长QT的药物的摄入,如同索他洛尔(Sotalol)的情况。索他洛尔是主要用于预防心房颤动、心绞痛以及高血压的钾电流抑制剂药物。索他洛尔可以导致心室复极的延长并且罕见地(1.8-4.8%)导致TdP。

[0006] 在机器学习(Machine Learning,ML)的领域中,且更具体地,在深度学习(Deep Learning,DL)的领域中的最新进展已经证实了利用包括心电图的生物信号进行处理的能力。具体地,深度学习已经用于利用ECG信号的数种应用。具体地,应用卷积神经网络(Convolutional Neural Networks,CNN)架构以检测各种类型的心律失常(Acharya、Oh等人,2017年)、心肌梗塞(Acharya、Fujita、Lih、Hagiwara等人,2017年);(Acharya、Fujita、Oh等人,2017年);Liu等人,2018年)、冠状动脉疾病以及可除颤室性心律失常。这些研究中的大多数利用单导联方法,且在模型中不包括临床数据。

[0007] 另一最新研究基于单导联ECG端对端模型进行训练以检测高达12种不同的心律失常。数据库包括来自经认证的心脏造影技师标注的超过29000个受试者的记录。模型对来自328个患者的336条记录进行测试,其胜过一组6位委员会认证的心脏病科医师的平均分数(Rajpurkar等人,2017)。相同小组利用超过53,000个受试者的数据集重复分析。基于独立数据集测试的模型获得了97%的曲线下面积,而心脏病科医师得到了78%的曲线下面积(Hannun等人,2019年)。研究阵发性心房颤动的Pourbabaee、Roshtkhari以及Khorasani,2017年也利用了端对端模型。

[0008] 其它研究聚焦于具有多导联ECG输入的模型。为了检测泛发性前壁心肌梗塞,Liu等人2018年使用了利用5-折交叉验证引起>90%的准确性的4导联方法。Tison等人2018年替代地创建隐马氏模型(Hidden Markov Model)-CNN模型,隐马氏模型采取12导联作为输

入以便检测肺动脉高血压、肥厚型心肌病、心肌淀粉样变性以及二尖瓣脱垂。针对4种类别,曲线下面积ROC在77%到94%的范围内。应用了基于>97,000个患者的数据库的CNN模型以检测左心室功能障碍的Attia等人,2019年也利用了12导联。他们利用>52,000个患者的保留数据集且总体准确性为86%。此外,一部分被错误地分类为心室功能障碍的患者在未来的数年中出现低射血分数,从而表明模型能够在此病症变为可见之前检测到。

[0009] Attia等人2018年尝试通过分析接受多非利特(Dofetilide,抗心律失常药物)或安慰剂的42个患者的10-s记录,利用CNN评估血浆药物浓度。在药物消耗前后的特定时间点处记录这些记录。在此实验中,作者利用了来自Physionet存储库中的可用的两个不同预期随机对照试验的数据。在他们利用CNN的回归分析中,他们达到了0.85的相关性。然而,数据库相对较小(42个受试者)且在训练中没有利用交叉验证,从而导致可能的过度拟合。此外,这篇文件并没有公开这种系统实际上能够检测患者不久将患有尖端扭转型室性心动过速的发作性症状,这是由于报告数据仅聚焦于增加的QT,其为TdP的一个因素但不足以得出实际发生的真实风险。这篇文章进一步利用了来自12导联的ECG数据,而本文中所公开的方法表明,方法可以利用较少导联(最低为一个)生成的ECG进行操作。此外,作者并没有证实有可能区分先天性长QT的种类。最后,作者提及研究的局限性,且具体地,无法得出关于心律失常风险的发生的结论。

[0010] Yıldırım 等人2018年能够利用端对端模型来检测具有91%的准确性(测试集)的17种类型的心律失常。然而,此分析基于从45个个体记录的10秒的1000个单导联ECG。这意味着相同受试者在训练和测试中都具有数据,从而导致非独立数据集和过度拟合模型。其他作者应用了数据扩增以便增加数据集来检测5种心脏病症,且当利用10-折交叉验证程序测试时,准确性高于90%(Acharya、Oh等人,2017年)。这些作者还比较了降噪的数据库与原始数据库之间的性能,从而表明噪音去除并没有关联到更高分数。然而,交叉验证随机分离是在数据扩增之后完成的。信号在训练集与测试集之间并非完全独立的事实可能会导致过度拟合的表现。如本公开中所示出,这些作者并没有报告他们能够鉴别先天性长QT的本质或预测TdP在近期发生的风险,诸如在48小时内或随后的24小时内。

[0011] 寻找用于量化根据药物摄入产生TdP的风险的新颖方法很重要。实际上,仅查看QT不足以检测延长QT的药物,诸如索他洛尔的摄入,如示例中所示出。

[0012] 由于TdP事件的罕见且由于此事件基本上在摄入QT延长药物之后发生的事实,由本发明人研发的策略是研发能够检测索他洛尔(其为患有尖端扭转型室性心动过速风险的标记)的摄入的模型。利用因此研发的模型,发明人能够表明其能够鉴别先天性长QT综合症(易发生来自多态性室性心动过速的晕厥和猝死的心肌复极的遗传性失调症,即,TdP),且具体地,相比于LQT1或LQT3患者能够标识LQT2患者(因为临近的生物机制是不同的)。

[0013] 应注意,上文所提及的风险不是绝对风险,而是相对风险。在本发明的上下文中,检测到“处于风险中”的患者存在比作为整体的群体更高的风险。换句话说,患者存在由已经记录了TdP发作性症状的其他患者共享的特性。这意味着风险与患者将患有TdP的事实的检测的灵敏度有关。然而,由于测试并非100%特定的,因此一些患者可能存在该特性而不患有TdP发作性症状(误报);且由于测试并非100%敏感的,因此可能存在不具有该特性而将产生TdP事件的一些患者(漏报)。测试的质量由接受者操作特征曲线下面积(Area Under Received Operating Curve,AURC)表示。

[0014] 本文中所公开的方法和系统能够将患者分类为与患者在增加的风险类别(药物诱发风险、后天性风险、先天性长QT综合症及其类型)中的风险的类型(或存在风险的机制)相关的类别(风险增加/风险没有增加)和子类别。

发明内容

[0015] 本发明涉及一种用于产生能够预测或评估患者患有尖端扭转型室性心动过速事件的风险且优选地能够预测或评估其原因(或基础机制)的机器学习算法(或机器学习系统)的方法,其包括:

[0016] a. 将包括第一参数、第二参数以及第三参数的患者数据存储在电子数据库中,第一参数是来自患者的ECG,第二参数与患者患有尖端扭转型室性心动过速事件的风险相关,第三参数与其原因相关;

[0017] b. 设置机器学习系统;并且

[0018] c. 利用患者数据训练机器学习系统,使得机器学习系统经过训练以在接触到来自患者的ECG时产生对患者患有尖端扭转型室性心动过速事件的风险和其原因的预测。

[0019] 如示例中所证实,发明人已经表明利用来自患者的心电图使得能够将患者分类为患有尖端扭转型室性心动过速风险(如与索他洛尔的摄入相关的),且能够分类风险之下的机制(药物的摄入、先天性长QT综合症、先天性长QT综合症的类型)。利用如上文所公开的数据集有可能获得使检测这种风险(以及风险之下的机制)成为可能的系统的事实对于发明人来说是有出乎意料的。实际上,发明人表明ECG信号包括可以从机器学习系统提取的信息且允许针对TdP风险的患者的非常详细的分类。

[0020] 具体地,发明人已经表明利用这种系统使得有可能检测患者是否将在短延迟内(在48小时或24小时内)患有尖端扭转型室性心动过速的发作性症状。发明人还表明也可以利用从单导联获得的ECG来操作系统。

[0021] 因而,经过训练的系统的输出可以依赖于分类技术,其提供预测、概率或可能性,而不是确定性,即患者患有某一病症(这里冒着遭受TdP事件的风险,以及生物机制的本质)。举例来说,利用以下技术的一种或更多种的分类器可以用于确定这种风险的可能性:线性鉴别分析、神经网络、聚类技术、支持向量机、逻辑回归、朴素贝叶斯、随机森林、决策树等。

[0022] 然而,优选的是机器学习分类器是人工神经网络且优选地是卷积神经网络。

[0023] 因此,且更具体来说,本发明涉及一种用于产生能够预测或评估患者患有尖端扭转型室性心动过速事件的风险且优选地能够预测或评估其原因(或基础机制)的人工神经网络系统的方法,其包括:

[0024] a. 将包括第一参数、第二参数以及第三参数的患者数据存储在电子数据库中,第一参数是来自患者的ECG,第二参数与患者患有尖端扭转型室性心动过速事件的风险相关,第三参数与其原因相关;

[0025] b. 设置互连以形成人工神经网络的节点的网络,节点包括多个人工神经元,每个人工神经元具有带相关联的权重的至少一个输入;并且

[0026] c. 利用患者数据训练人工神经网络,使得多个人工神经元的每个人工神经元的至少一个输入的相关联的权重响应于来自患者数据的多个不同数据集的相应第一参数、第二

参数以及第三参数而调整,使得人工神经网络经过训练以在接触到来自患者的ECG时产生对患者患有尖端扭转型室性心动过速事件的风险和其原因的预测。

[0027] 优选地,神经网络利用这种输入数据并且训练机器学习系统,允许获得能够确定患者的ECG是否存在与处于患有TdP事件的风险中的患者的ECG相同的特性的模型。因此,这种模型可以用于实施如下文所公开的各种方法。

[0028] 因此,本发明涉及一种计算机实施的心电图(electrocardiogram,ECG)分析方法,其包括:

[0029] a.从受试者获取的ECG信号中分离数据片段且提取这种数据片段;

[0030] b.任选地将数据片段归一化或标准化;

[0031] c.将数据片段应用于机器学习模型,机器学习模型配置为对指示尖端扭转型室性心动过速的风险的ECG数据中的变化进行检测;

[0032] d.从机器学习模型获得输出;

[0033] 其中输出提供患有尖端扭转型室性心动过速事件的风险的评估以及这种风险的基础机制的评估。

[0034] 此方法以及本文中所公开的其它方法在体外执行。

[0035] 本发明还涉及一种计算机实施的心电图(ECG)分析方法,其包括:

[0036] a.从受试者获得的ECG信号中分离数据片段且提取这种数据片段,其中数据片段为在特定时间窗口期间的ECG信号;

[0037] b.任选地将数据片段归一化或标准化;

[0038] c.将作为输入数据的数据片段应用于机器学习分类器,机器学习分类器配置为对指示尖端扭转型室性心动过速的风险的ECG数据中的变化进行检测;

[0039] d.从机器学习分类器获得输出;

[0040] 其中输出提供患有尖端扭转型室性心动过速事件的风险的评估以及这种风险是否与药物的摄入、先天性长QT综合症和先天性长QT综合症的类型相关联的评估。

[0041] 具体地,机器学习模型(或系统)是诸如前馈神经网络、卷积神经网络或递归神经网络的神经网络,此列表是非限制性的。优选地是卷积神经网络。

[0042] 步骤a对应于来自ECG信号的片段(部分)的提取(优选地在10秒窗口期间),以便获得可以用于如上文所公开的先前已经训练过的机器学习系统中的处理的数据片段。其可以在线或离线完成。

[0043] 在步骤b,有可能将数据片段归一化和/或标准化或者有可能完全不进行处理。举例来说,当已经基于多个导联获取了ECG信号时,可以利用针对患者的每个导联或按每个导联获取的信号的平均值执行标准化。可以利用存在于参考数据库中的所有患者的ECG信号或针对每个ECG执行归一化。

[0044] 在步骤c,将数据片段应用于如上文所公开的已经研发出的机器学习算法以便获得输出。还可以将其它数据与数据片段一起应用(诸如性别、患者的年龄、高钾血等)。当已经基于多个导联获得数据时,可以一次性键入来自所有导联的数据,分别作为ECG信号(复合ECG信号)或导联部分或所有导联。

[0045] 本发明还涉及一种用于评估患者患有尖端扭转型室性心动过速事件的风险和其基础机制的方法,其包括:

[0046] a.由处理装置接收表示受试患者的ECG波形的时间片段的信号数据;

[0047] b.由处理装置经由配置的人工机器学习算法(具体地,神经网络)分析所述信号数据以生成人工神经网络的输出;

[0048] 其中输出为患者患有尖端扭转型室性心动过速事件的风险和其原因(基础机制)的可能性。

[0049] 针对上文(和贯穿本公开)所公开的方法,输出提供患者属于训练期间定义类别的可能性。

[0050] 作为说明

[0051] -如果利用两种类别(已经摄取了索他洛尔或没有摄取索他洛尔)执行了神经网络的训练:那么输出将提供处于一种或另一类别的可能性;

[0052] -如果利用多于两种类别(例如5种类别:索他洛尔+、对照、长QT类型1、长QT类型2、长QT类型3)执行了神经网络的训练,那么输出将提供患者处于五种类别的每一种的可能性。明显的是,输出将有可能赞同一种类别,其指示患者可能属于此类别。

[0053] 患有尖端扭转型室性心动过速事件的风险与对于患者来说的患者处于更高风险下的类别的归属有关。换句话说,与被分类为对照的患者相比,风险为相对风险。当将患者的ECG被分类为和属于风险增加的患者(服用增加QT的药物的患者或患有长QT综合症的患者)的类别的患者的ECG一样时,患有TdP事件的风险被视为增加(与对照患者相比)。机器学习算法也能够鉴别分类,且因此能够提供关于增加TdP风险的基础机制的信息:服用索他洛尔患者或LQT-2患者的IKr通道的抑制、LQT-1患者的IKs通道中所涉及的KvLQT1基因(也称为KCNQ1)的突变、SCN5A中的突变、编码LQT-3患者的钠通道INa/INa后期的基因。

[0054] 因此,医生有可能适当地行动(停止药物治疗且检查ECG再次变为正常、针对长QT综合症患者请求基因确认以及开适当药物治疗的药方)。

[0055] 本发明还涉及一种心脏监测装置,其包括:

[0056] a.至少一个电极,其用于从患者获得心电图(ECG)信号;

[0057] b.处理单元,其包括在操作上联接到至少一个电极的至少一个处理器;以及

[0058] c.至少一个非易失性计算机可读介质,其包括程序指令,程序指令在由至少一个处理器执行时使得心脏监测装置:

[0059] i.从至少一个电极获得ECG信号;

[0060] ii.任选地将ECG信号标准化/归一化以获得标准化/归一化ECG信号;

[0061] iii.任选地从对应于时间的预定持续时间的标准化ECG信号提取数据以便获得提取到的标准化ECG信号;

[0062] iv.将任选地标准化/归一化并提取的ECG信号提供到机器学习分类器;

[0063] 由此从机器学习分类器获得输出,所述输出是患有TdP事件的可能性和其原因。

[0064] 在ii,ECG信号的标准化/归一化是任选的且可以如上文所公开地执行。步骤iii目的在于在机器学习分类器中仅利用在特定持续时间(通常10秒)内的信号。此步骤是任选的且并非总是必需的。

[0065] 如上文所指示,机器学习分类器优选地是神经网络,具体地是卷积神经网络。

[0066] 本发明还涉及一种用于评价患者患有尖端扭转型室性心动过速事件的风险的系统,其包括:

- [0067] a. 体外除颤器,其配置为监测来自患者的心电图 (ECG) 数据;
- [0068] b. 一个或多个处理器,其配置为执行操作,操作包括:
- [0069] i. 任选地执行用于给定时域的ECG数据的数学标准化;
- [0070] ii. 通过机器学习分类器执行计算;并且
- [0071] iii. 确定患者患有尖端扭转型室性心动过速事件的可能性和其原因。
- [0072] 应提醒,当利用多个导联时,ECG数据是由来自不同导联的所有这些多个ECG信号形成的向量。示例提供这样的关于8个导联的向量的说明。
- [0073] 在以上实施方案中,包括至少一个电极以及计算机/处理器或处理单元的除颤器或监测器可以位于不同位置处。在此实施方案中,将ECG信号发送到处理单元或处理器以供机器学习分类器对其处理。也有可能将输出发送到除颤器或监测器或另一装置(诸如智能电话)。在这种情况下,有可能在家执行患有尖端扭转型室性心动过速事件的风险或其演变(取决于通过机器学习分类器计算的可能性的演变)的监测。当具有最少电极的除颤器或装置和处理系统相距遥远时,优选的是加密这些系统之间的通信。
- [0074] 本发明还使得有可能评价药物、产品或感兴趣的组合物的尖端扭转型室速潜能。
- [0075] 在获得上市授权之前的临床试验期间,能够评价化合物的尖端扭转型室速潜能以便能够起草用于使用此药物的可能注意事项确实很重要。物质的尖端扭转型室速潜能是此物质在向患者施用之后诱发尖端扭转型室性心动过速的潜能。
- [0076] 在标题《非抗心律失常药物的QT/QTc间期延长和致心律失常潜能的E14临床评价》下,食品和药物管理局 (Food and Drug Administration, FDA) 和欧洲药物管理局 (European Medicines Agency, EMA) 已经公布由国际协调会议 (International Conference for Harmonization, ICH) 针对如何进行这些试验所产生的指导方针。这些指导方针在此代理的网站上是可访问的且附随着问题和答案。
- [0077] 指导方针的url地址为:
- [0078] <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm073153.pdf>
- [0079] 如上文所论述,QT间期是致心律失常效果的风险的不完全的生物标记。具体地,示例表明在利用QT来检测已经服用了索他洛尔的患者时获得的AUC为0.76,证明其可以获得改进。
- [0080] 因此,必需严格地表征关于延迟心肌复极的风险的新药物的效果。通过分析已经施用了感兴趣的化合物的患者的ECG,如上文所定义的机器学习分类器使得有可能检测是否存在这种风险。
- [0081] 因此,本发明涉及一种用于确定物质在向患者施用之后诱发尖端扭转型室速效果的风险的体外方法,其包括以下步骤:
- [0082] a. 在施用所述组合物之后从患者获得ECG数据;
- [0083] b. 将ECG数据应用于机器学习分类器(如上文所公开),机器学习分类器配置为对指示尖端扭转型室性心动过速事件的风险增加的ECG数据中的变化进行检测;
- [0084] c. 从机器学习分类器获得输出,其中输出提供患者患有尖端扭转型室性心动过速事件的风险的可能性;
- [0085] 其中如果在施用物质之后获得患者患有尖端扭转型室性心动过速事件的风险,那

么物质存在诱发尖端扭转型室速效果的风险。

[0086] 优选的是将在施用感兴趣的物质之后在各种时间点获取的多个ECG信号应用于机器学习分类器。

[0087] 也有可能将在施用感兴趣的物质之前获取的患者的ECG信号应用于机器学习分类器。

[0088] 通常,为了限制患者可变性,上文所公开的方法在含有大于或等于2的患者数目的患者群组上重复。因此,基于足够数目的患者进行这些研究使得获得的结果有可能呈现期望测试的分子(感兴趣的药物)的统计的真实性,从而消除了患者间的变化。

[0089] 可以对安慰剂(阴性对照)、阳性对照(诸如80毫克的索他洛尔)以及感兴趣的物质进行比较。

[0090] 优选地,80毫克的索他洛尔实际上用作阳性对照。此浓度下的此分子具有许多优点:

[0091] -最大浓度是熟知的(在经口施用之后大约三小时),其使得有可能容易地计划采取上文所测量的测量值的时刻。

[0092] -已知此分子用以抑制IKr通道。

[0093] -此浓度下的此分子没有诱发尖端扭转型室性心动过速的发作性症状,且因此对于患者来说是安全的。然而,已知的是,此分子在其临床使用剂量(使用高达320毫克)下可以诱发尖端扭转型室性心动过速。

[0094] -本文中所公开的机器学习分类器具体地适用于检测AUC为0.99(CI[0.98;1])的患者中的索他洛尔(80毫克)的摄入。

[0095] 当针对女性测试物质时,合理的是在月经周期的不同时间进行各种测量。这是因为已知的是,激素水平影响了女性中复极失调症(长QT)的风险。

[0096] 一般来说,如上文所指示,针对患者群组的每个患者进行以上步骤,且因此针对每个患者且针对每种物质获得风险。

[0097] 患者群组(对感兴趣的物质进行测试的患者的数目)优选地至少为10个患者,优选地至少20个或甚至50个或100个患者。本领域技术人员将能够确定患者群组中的受试者的最小数量,以便获得统计上显著的结果。

[0098] 同样地,在进行这些研究之前,优选地,本领域技术人员已经进行了对感兴趣的物质的药物动力学研究和剂量反应研究,以便确定用以执行ECG获取和机器学习分类的最好时间。

[0099] 本发明还涉及一种用于分类(确定)患者中的先天性长QT综合症的本质的方法,其包括步骤:

[0100] a. 从患者获得ECG数据;

[0101] b. 将ECG数据应用于机器学习分类器,所述机器学习分类器配置为对指示尖端扭转型室性心动过速事件的风险的ECG数据中的变化进行检测;

[0102] c. 从机器学习分类器获得输出,其中输出提供长QT的本质的可能性和其分类。

[0103] 具体地,可以对已经检测为患有先天性长QT综合症但尚未对其执行基因分析的患者执行此方法。

[0104] 因此,此方法使得有可能在获得基因确认(由于其可以花费几个月)之前极早地采

取治疗。

[0105] 应提醒,具有临床上某一LQTS(长QT综合症)诊断的大约75%的患者在三种主要LQTS易感基因(KCNQ1、KCNH2或SCN5A)的一种中具有突变,其中KCNQ1中的突变约占LQTS(类型1,LQT1)的35%的原因,KCNH2突变约占LQTS(LQT2)的30%以及SCN5A突变约占LQTS(LQT3)的10%。

[0106] 其它基因(AKAP9、CACNA1C、CALM1、CALM2、CAV3、KCNE1、KCNE2、KCNJ5、SCN4B以及SNTA1)为次要LQTS基因。

[0107] 针对已经执行了基因测试且已经在这些基因的一些中发现了突变的患者,机器学习分类器系统将能够通过分析这种患者的ECG来检测这种患者是否已经在实际上被分类为LQTS。因此,本文中所公开的系统使得标识被分类为未知意义的变型的给定突变更有可能是致病的还是良性的成为可能。

[0108] 因此,本发明涉及一种用于确定基因中感兴趣的突变是否易于与先天性LQTS相关联的方法,其包括:

[0109] a.将来自存在感兴趣的突变的患者的ECG数据应用于如本文中所公开的机器学习分类器;

[0110] b.从机器学习分类器获得输出分数;

[0111] c.根据输出分数确定患者是否患有延长QT;

[0112] 其中当输出分数将患者放置在患有延长QT的患者的类别中时,感兴趣的突变易于与先天性LQTS相关联。因此与先天性LQTS的关联性由临床表现的存在(QT延长特征的ECG特性的存在)确定。

[0113] 应提醒,机器学习分类器的输出是患者患有延长QT的可能性(当机器学习分类器已经利用来自已经接受QT延长药物的患者和来自尚未接受QT延长药物的患者的ECG进行训练时),且因此有可能将患者分类在这两种类别的一种;如果患者被分类在已经接受QT延长药物的患者的分类,那么其被视为患有延长QT。如果患者患有延长QT,如果突变因此易于关联到此病症、与此病症有关或导致此病症。如果患者不患有延长QT,那么突变有可能不导致LQTS或关联到LQTS。如上文所公开,如果利用先天性长QT和其类型1、2、3或其它类型)的具体的类别来执行神经网络的训练,那么结果将更加具体。然而,如示例中所表明,仅利用来自服用QT延长药物的患者的数据的训练也使得有可能标识先天性长QT。

[0114] 具体地,感兴趣的突变在如上文所引用的基因(主要的三个基因的任一个或在次要基因中)中。

[0115] 还应提醒,必须将此测试理解为概率性的而不是无条件地确定性的,且需要未来工作(诸如用于具有相同突变的其它患者的QT分析)来确认突变与LQTS的作用或关联性。然而,能够区分来自中性变型的危害性将帮助避开在良性等位基因的载体中的不必要干预,且聚焦于具有真正相关的突变的患者。

[0116] 由于本公开还表明本文中所公开的方法能够标识LQTS的本质,因此当通过机器学习分类器标识LQTS的本质时,也有可能恰当地处理患者。举例来说,诸如阿替洛尔(Atenolol)、纳多洛尔(Nadolol)、普萘洛尔(Propranolol)的 β -阻滞剂(I类或IIa类)可以用于LQTS(尤其用于LQT1和LQT2),而美西律(Mexiletine)、雷诺嗪(Ranolazine)或氟卡尼(flecainide)(非选择性电压栅控钠通道阻滞剂)可以用于LQT3。

[0117] 本发明还涉及一种用于检测(或评估)给定患者存在尖端扭转型室性心动过速的发作性症状的风险的方法,其包括步骤:

[0118] a.从患者获得ECG数据;

[0119] b.将ECG数据应用于机器学习分类器,机器学习分类器配置为对指示患者患有尖端扭转型室性心动过速事件的风险的ECG数据中的变化进行检测;

[0120] c.从机器学习分类器(具体地,神经网络系统)获得输出,其中输出提供患者患有尖端扭转型室性心动过速事件的风险的可能性。

[0121] 在具体实施方案中,通过电子/网络手段(诸如电缆或无线网络、GSM、3G或wifi)将ECG数据发送到远程机器学习分类器,且通过电子/网络手段(诸如通过电子邮件或通过电话装置上的短消息(SMS))将输出发送到患者装置和/或医生装置(电话、计算机)。具体地,将输出发送到医生,使得医生可以因此确定患者的TdP的风险且相应地行动。具体地,每日都执行这种方法。优选的是对测量ECG的机器与含机器学习分类器的远程服务器之间的通信进行加密,且对远程服务器与患者装置或医生装置之间的通信进行加密。

[0122] 也有可能通过利用如下文所公开的投票系统将患者分类或将患者指派在类别(长QT综合症的本质、TdP事件的风险增加)中。如示例中所表明,这种系统将增加方法的准确性。

[0123] 在优选的实施方案中,通过对在将患者的多个ECG信号应用于机器学习分类器(投票系统)时获得的风险分数求平均来计算患者的可能性(概率或评估)。机器学习分类器基于多个ECG样品的重复等效于医生查看不同ECG样品以得出结论。实际上,此方法使得有可能将患者分类在给定类别,该给定类别是鉴于由机器学习分类器计算的不同的可能性而最有可能的一种类别。处理来自相同患者的多个ECG有可能是通过模型处理的且当输出的大多数指示患者在那个类别中的概率更高时有可能影响类别中的患者。

[0124] 也有可能通过利用已经出于相同目的而独立地训练的不同机器分类器来将患者分类在给定类别中。利用此类型的投票系统等效于多个医生在做出决策之前查看相同数据。

[0125] 也有可能组合这两种投票方法。

[0126] 应提醒,机器学习系统使得有可能提供已经在训练过程期间利用的类别中的患者的概率。在操作阶段期间,能够将类别指派到患者(或将患者指派到类别)很重要。

[0127] 以上方法可以通过将患者指派到类别的方法完成,其中通过:

[0128] a.利用来自相同患者的不同ECG信号重复上文所公开的方法,并且

[0129] b.将患者指派在输出的大多数指示最高概率的类别中。

[0130] 也有可能利用将患者指派到类别的方法,其中通过:

[0131] a.使用根据上文所公开的方法训练和获得的不同机器学习分类器,利用来自患者的ECG信号重复方法;并且

[0132] b.将患者指派在输出的大多数指示最高概率的类别中。

[0133] 可以基于单个导联(优选地,V3或DII导联)获取ECG信号。然而,优选的是从多于一个导联,优选地从至少3个导联,更优选地从至少6个导联或从至少8个导联获取ECG信号。也可以单独地利用任何导联。DI、V4或V6导联尤为感兴趣。

[0134] 发明人已经具体地确定出利用机器学习分类器,其已经配置(教导)为利用来自自己

经施用QT延长药物的患者和来自尚未施用QT延长药物的患者的ECG,使得有可能在实际上检测尖端扭转型室性心动过速发作性症状是否将在48小时或24小时内发生,或检测尖端扭转型室性心动过速发作性症状是否不会在48小时或24小时内发生。

[0135] 重要的是,应理解,患有尖端扭转型室性心动过速发作性症状的风险并不与延长(或增强型、或增加型)QT严格地相关。尽管这种延长QT增加风险(在服用增加型QT的患者群体相比于服用普通QT的患者群体中存在相对增加的风险的意义上),但这种增加本身并不是尖端扭转型室性心动过速将在实际上发生的事实的充分预测。

[0136] 群体风险(群体中增加的风险)本身没有提供群体中的给定患者的信息。

[0137] 然而,能够评定个体患者在近期是否处于患有尖端扭转型室性心动过速发作性症状的风险对医生来说很重要。还可以据称,确定个体患者在近期是否将不患有任何尖端扭转型室性心动过速发作性症状很重要。患者通常为施用药物的患者,该药物已知用以增加QT且因此增加尖端扭转型室性心动过速事件的风险。医生的职责是监测患者以便减少其个体风险。监测可以通过昂贵且占用医院中的床位的恒定医院监督进行。因此,医生将优选地能够排除患者将实际上在近期(在48小时或24小时内)患有尖端扭转型室性心动过速发作性症状的事实且仅密切地监测大部分处于风险中的患者。

[0138] 发明人已经表明,当能够利用高特异性确定尖端扭转型室性心动过速的风险的新ECG被提供时,具有如示例中所公开的数据库的机器学习系统的训练使得有可能获得超出训练的输出。

[0139] 应提醒:

[0140] (1) 灵敏度是在患有病症(真阳性的检测)的个体中诊断是阳性的概率:如果患者患有病症,那么测试是阳性的。

[0141] (2) 特异性是在不患有病症(存在真阴性的非检测)的个体中诊断是阴性的概率:如果患者不患有病症,那么测试是阴性的。

[0142] (3) 阳性预测值(Positive predictive value,PPV):是在诊断测试是阳性(即,患者不是假阳性)的情况下患有疾病的概率:如果测试为阳性,那么患者患有病症。

[0143] (4) 阴性预测值(Negative predictive value,NPV):是在诊断测试是阴性(患者不是假阴性)的情况下不患有病症的概率:如果测试是阴性,那么患者不患有病症。

[0144] 在目前情况下,病症(实际尖端扭转型室性心动过速事件)的患病率较低。然而,在高灵敏度的情况下,NPV将较高。因此有可能说,如果患有尖端扭转型室性心动过速事件的任何风险因而并没有通过系统检测到,那么个人在近期(在48小时或24小时内)不具有患有尖端扭转型室性心动过速事件的任何风险(或极小风险)。换句话说,发明人表明系统极为擅长检测由于高灵敏度而将不患有尖端扭转型室性心动过速的患者,即使一些患者可以由于不具有同样高的灵敏度的特异性而被检测为假阳性也是如此。因此有可能减小要监测的患者的数量,且聚焦于存在TdP事件的更高患病率(更高风险)的子组。

[0145] 此外,发明人表明测试是半量化的(即,分数越高,风险越高)。针对高于给定阈值的患者,个体风险将增加且个体可以增加监测,停止处于风险中的药物的施用(或减少剂量)和/或通过平衡镁血或钾血浓度来校正任何可能的水电解质问题,从而将患者维持在医院中或进入医院中以供监督。

[0146] 如下文所公开,机器学习分类器(优选地,神经网络)提供分数作为输出。发明人已

经表明分数越高,在48小时或24内尖端扭转型室性心动过速事件的风险越高,且其为高度特异性的。应注意,分数的值与QT的增加不完全相关:因此,分类器能够利用除仅QT长度以外的其它元素对患者进行分类。由模型提供的信息(作为分数)因此与通过测量QT长度传递的信息不同。

[0147] 有可能利用两个截断值定义三种类别。在一个截止值以下,患者的个体风险将极低,且可以因此减少监测。针对高于另一阈值的患者,将加强监测。针对两个阈值之间的患者,将照常维持监测。

[0148] 本发明还涉及一种用于确定受试者是否在48小时内处于患有尖端扭转型室性心动过速的发作性症状的风险中(或处于增加的风险中)的方法,其包括:

[0149] a. 从受试者获取的ECG信号中分离数据片段且提取这种数据片段;

[0150] b. 任选地将数据片段归一化或标准化;

[0151] c. 将数据片段应用于机器学习分类器,机器学习分类器配置为对指示QT延长药物的摄入的ECG数据中的变化进行检测;

[0152] d. 从机器学习分类器获得输出;

[0153] 其中输出提供在48小时内患有尖端扭转型室性心动过速事件的风险的评估。

[0154] 本发明还涉及一种用于确定受试者是否在24小时内处于患有尖端扭转型室性心动过速的发作性症状的风险下(或处于增加的风险下)的方法,其包括:

[0155] a. 从受试者获取的ECG信号中分离数据片段且提取此数据片段;

[0156] b. 任选地将数据片段归一化或标准化;

[0157] c. 将数据片段应用于机器学习分类器,所述机器学习分类器配置为对指示QT延长药物的摄入的ECG数据中的变化进行检测;

[0158] d. 从机器学习分类器获得输出;

[0159] 其中输出提供在24小时内患有尖端扭转型室性心动过速事件的风险的评估。如所指示,当输出更高时,风险更高。此外,方法可以包括步骤:将输出与所确定的阈值进行比较并且如果输出高于阈值,那么评定患有TdP的风险。

[0160] 应注意,发明人已经训练了模型,使得当分数更高时,尖端扭转型室性心动过速的风险更高。这与正如在训练时提供到机器学习分类器的数据的标注完全有关。利用不同标注可以引起另一模型,其风险将在分数更低时更高。要点是将其与阈值进行比较,比较使得能够将患者分类为“处于风险中”或“不处于风险中”,且阈值和分类将根据特异性、灵敏度以及系统的训练类型所做出的选项。

[0161] 综上所述,本公开表明:

[0162] -有可能利用来自多个或单个导联的ECG数据训练机器学习分类器;

[0163] -这种训练后的模型能够鉴别/分类已经接受QT延长药物的患者与尚未接受这种药物的患者;

[0164] -无论ECG输入数据是从一个导联还是从多个导联获得以及模型是利用相同导联或另一导联训练还是利用多个导联训练,都获得这种分类;

[0165] -模型也能够标识先天性长QT病症的存在或不存在,以及其本质;

[0166] -模型还能够预测TdP事件在24小时或48小时内发生的风险和/或预测TdP事件在24小时或48小时内不发生。

附图说明

[0167] 图1:A.ROC曲线,其指示在Sot+受试者与(从左到右)对照、LQT-1、LQT-3以及LQT-2受试者之间的分离。B.ROC曲线,其指示在LQT-2受试者与(从左到右)对照、LQT-1、LQT-3以及Sot+受试者之间的分离。

[0168] 图2:ROC曲线,其指示在TdP事件期间的ECG与对照受试者之间的分离。

具体实施方式

[0169] 实验设计

[0170] 训练数据库包括来自792个健康患者的在索他洛尔摄入(口服80毫克,Sot+)之前(n=4014;Sot-)和之后(n=5000)的9014个ECG信号。在不同时间点(在索他洛尔摄入之后的1、2、3以及5小时)取得ECG。

[0171] 分离数据集的子集以测试训练后的模型,训练后的模型由具有在索他洛尔摄入之前(n=999)和之后(n=1238)的ECG的198个健康患者构成。

[0172] 此外,由长先天性QT患者构成的第三数据集用于测试训练后的模型,训练后的模型具体地由LQT-1(n=266个患者且n=560个ECG)、LQT-2(n=188个患者且n=456个ECG)以及LQT-3(n=33个患者且n=67个ECGs)构成。

[0173] 群组描述

[0174] 利用患者的两个群组。

[0175] 第一群组含有来自990个健康受试者的在消耗80毫克的索他洛尔之前和之后在不同时间段的心电图(即,在索他洛尔摄入之后的2、3、4以及5小时)。每个患者具有不同数量的记录,这是由于这些记录在每个时间间期要么重复,要么不重复。每个受试者在基底和在T3处具有至少一条记录。在摄入之后的2、3以及5小时,在血清中测量索他洛尔浓度,但这种情形并不用于所有受试者。在索他洛尔摄入之前,所有信号关联到索他洛尔的零浓度。此第一群组用于训练卷积神经网络(Convolutional Neural Network,CNN)和ResNet模型,ResNet模型在Sot+/Sot-组中分类个体且预测血液中索他洛尔的浓度。

[0176] 还利用由长QT先天性患者构成的第二群组,具体地,第二群组由来自用以测试训练后的模型的第二研究的LQT-1(n=266个患者且n=560个ECG)、LQT-2(n=188个患者且n=456个ECG)以及LQT-3(n=33个患者且n=67个ECG)构成。

[0177] 卷积神经网络架构

[0178] 根据输入研发两种类型的模型。仅利用ECG(M1)或具有关联到性别、年龄以及钾血症的数据的ECG(M2)。模型的架构相同。

[0179] 模型是基于卷积神经网络(CNN),卷积神经网络由11个卷积块构成,且每个块具有带相同数量的过滤器的两个Conv1D层。每个块的过滤器的数量分别为8、8、8、16、32、64、128、256、512、1024以及2048(总共11个块)。零填补用于两个Conv1D,以便具有带与输入长度相同的输出(选项“相同”)。利用亚当(Adam)优化器、二进制交叉熵损失以及早停法(耐心=50)。计算类别权重以平衡在训练之前的输出类别。其余模型参数具有默认值。在M2模型中,在完全连接神经网络之前,将临床数据串接至扁平层输出。

[0180] 研发基于相同CNN架构的不同深度学习模型。第一(M1)仅包括在索他洛尔摄入前后的ECG数据(下文分别表示为对照和Sot+),而另一(M2)还包括基本临床信息,即,年龄、性

别以及在研究基线下量化的钾血症。

[0181] 研发不同模型M1。尽管一些个体在测试阶段期间的评定中具有一些差异,但他们都存在相同的整个性能。

[0182] 准备ECG数据

[0183] 使用利用Cardioplug装置(Cardionics公司)获取的原始ECG。ECG关联到8个导联(I、II、V1、V2、V3、V4、V5、V6)。在500赫兹的样品大小的情况下,每条记录持续10秒。利用Biosig库(<http://biosig.sourceforge.net/index.html>)解析呈.SCP文件格式的数据。

[0184] 类似地,使用来自长QT患者的原始ECG数据,原始ECG数据利用从1992年到2018年的CalECG工具获取。使用python xmltodict库解析原始.xml文件。总体来看,在500赫兹下获取875个文件且在250赫兹下获取208份文件,总共1083条记录。利用三次插值将250赫兹信号上采样到500赫兹(来自scipy python库的interp1d功能)。

[0185] 除了在整个ECG级别执行的标准化(即,每个导联信号通过用于给定ECG的所有其它导联信号的平均值标准化)以外,不将过滤或任何其它转换应用于在这些研究中考虑到的数据。将数据存储在于训练模型的python字典中。

[0186] 利用的ECG数据对应于更大的ECG获取的连续10秒的片段。

[0187] 针对训练,键入数据作为3D张量(8个导联、每个导联的5000个时间点、Sot+或Sot-)。为了测试模型,键入数据作为2D张量(8个导联、每个导联的5000个时间点)。

[0188] 模型训练和评价

[0189] 利用Python中的Keras库创建所有探索的模型。在分类和回归实验两者中,将第一数据库分在训练(80%)集和保留(20%)集。在训练数据库中执行10×10折交叉验证以评估算法的通用化置信度。在每一折下,训练子集和测试子集具有不同的示例,且利用CV训练集的平均值和标准差来将两个集合(即,训练和测试)归一化以避免过度拟合。最终模型(M1、M2以及ResNet18)在整个训练集中进行训练且在保留和体外长qt先天性群组中进行测试。还提供所评估的CV经验和通用化性能。

[0190] 针对分类任务,训练数据集包括来自792个健康患者的在索他洛尔摄入(80毫克)之前(n=4014)和之后(n=5000)的9014个ECG信号。分离数据集的子集以测试训练后的模型,所述训练后的模型由具有在索他洛尔摄入之前(n=999)和之后(n=1238)的ECG的198个健康患者构成。分类作为二进制分类而执行:无药物(Sot-)相比于药物(Sot+)。将关联到包含物、基底以及T0的记录用作Sot-,而将在药物摄入之后(在T1、T2、T3、T4处)的记录用作Sot+。最后,长qt先天性数据集用于测试训练后的模型,所述训练后的模型具体地由LQT-1(n=266个患者且n=560个ECG)、LQT-2(n=188个患者且n=456个ECG)以及LQT-3(n=33个患者且n=67个ECG)构成。

[0191] 投票策略

[0192] 利用单信号分析和通过对来自给定患者和病症的多条记录的风险分数求平均来计算性能。由模型提供的输出是作为Sot+(已经摄取索他洛尔)的概率(可能性)。

[0193] 为了影响Sot+或Sot-类别的一种中的患者,通过模型处理来自相同患者的多个ECG,且如果输出的大多数指示作为Sot+(resp.Sot-)的概率高于0.5(50%),那么患者被影响为Sot+(resp.Sot-)。

[0194] 替代地,可以通过利用多个模型(利用类似架构或不同架构分别训练的)处理来自

患者的一个ECG来使用投票策略,且如果模型的大多数指示作为Sot+ (resp.Sot-) 的概率高于0.5 (50%),那么将患者影响到Sot+ (resp Sot-) 类别。还可以利用其它截断值。

[0195] 根据原始ECG信号准确地检测索他洛尔摄入

[0196] 模型的性能在单个ECG级别进行评价或由患者通过10乘以10-交叉验证折叠(CV)的每一个进行平均。观察到,CV中的平均测试准确性与保留准确性相当,训练过程(CV)中的总体经验准确性和经验也是如此,从而指示模型的性能的良好评估。

[0197] 模型M1的平均通用化准确性在针对每个ECG计算时为0.885 (sd=0.003),而在针对每个患者计算时(使用如上文所公开的投票策略)为0.945 (sd=0.004)。类似地,针对模型M2,平均通用化准确性在针对每个ECG计算时为0.89 (sd=0.003),而在针对每个患者计算时为0.947 (sd=0.003)。

[0198] 在M1与M2之间不存在统计差异,因此指示ECG独自含有关于索他洛尔的摄取的检测的所有信息。

[0199] 然而,在ECG处的通用化准确性与患者级别(投票, $p<1.03e-25$)之间存在非常显著的差异,从而指示多个信号获取的重要性。

[0200] 最后,这些模型的精确度非常高,针对M1和M2(投票)分别为0.963 (sd=0.004)和0.961 (sd=0.003),相同模型的查全率也是如此:0.934 (sd=0.005)和0.937 (sd=0.004)。

[0201] 总之,这些结果指示出仅利用原始心电图训练的CNN模型在预测健康受试者是否已经服用了索他洛尔(80毫克)时非常敏感且准确。

[0202] 索他洛尔摄入分类模型鉴别长QT先天性特征

[0203] 接下来,经过训练以鉴别索他洛尔摄入患者的最简单(M1)模型用于对来自LQT先天性群组的长QT先天性患者进行分类。模型提供了风险概率[0,1],所述风险概率基于阈值 ≥ 0.5 而在Sot-/Sot+类别中进行转换。

[0204] 首先,应注意到,大多数Sot+ECGs (95%)被分类为Sot+且大多数对照(87%)被分类为Sot-。投票结果显示如上文先前所论述的Sot+ (99%)和对照(93%)组的优良性能。基于这些观察结果,稍后仅论述将ECG信号整合于患者级别(投票)的结果。

[0205] 其次,应注意到,所有LQT患者大多数被分类为Sot+。LQT-2显示Sot+的最强比例(85%),其次为LQT-3 (67%)和LQT-1 (63%)。此外,执行模型M1的ROC分析以一起鉴别组。在Sot+相比于对照、LQT-1、LQT-2以及LQT-3组之间的比较的结果分别呈现0.99(置信区间,CI [0.98;1])、0.79 (CI [0.75;0.83])、0.59 (CI [0.53;0.64])以及0.75 (CI [0.66;0.84])的曲线下面积(Area Under Curve,AUC) (图1A)。在LQT-2相比于LQT-1、LQT-3、Sot+以及对照组之间的比较的结果分别呈现0.71 (CI [0.66;0.75])、0.68 (CI [0.58;0.77])、0.59 (CI [0.53;0.64])以及0.92 (CI [0.89;0.95])的AUC (图1B)。

[0206] 就比较而言,当仅评定QT时,Sot+与对照之间的AUC为0.7674 [0.7381;0.7968]。

[0207] Sot+相比于LQT-2的低AUC指示这种模型无法良好地鉴别这两个组,从而表明它们的相互类似性(IKr通道抑制)。然而,AUC高于0.5的事实表明在这两种病症之间存在一些差异且进一步的训练(其中一种类别针对Sot+受试者且另一种类别针对LQT-2受试者)将允许能够鉴别Sot+和LQT2患者。

[0208] 连同LQT-2相比于对照(0.92)的高AUC,这些结果指示在患者患有先天性长QT的情况下,利用索他洛尔数据训练的M1模型作为用以鉴别来自对照的LQT-2患者的替代是恰当

的且切合的。

[0209] 索他洛尔摄入分类模型在事件发生时标识患有尖端扭转型室性心动过速的患者

[0210] 两个患者在尖端扭转型室性心动过速的发作性症状之后不久住院。

[0211] 对这些患者执行ECG且将ECG信号应用于如上文所公开的神经网络分类器。

[0212] 对于患者1,处于尖端扭转型室性心动过速的风险下的可能性(模型的输出)为0.991629、0.989505以及0.072291(三个ECG)。

[0213] 对于患者2,可能性是

[0214] 访问1:0.991812、0.999446以及0.984214

[0215] 访问2:0.998443、0.999096以及0.999811。

[0216] 这些结果说明患有TdP发作性症状且在发作性症状不久之后被分类为Sot+的患者的示例。患者不患有先天性长QT且不服用诱发QT延长的任何药物。其QT延长是后天性。

[0217] 因此,上文所描述的模型能够仅在心肌事件之后将患者分类为“处于尖端扭转型室性心动过速的风险中”,这使得有可能做出回顾性诊断。

[0218] 论述

[0219] 以上研究表明ECG信号承载重要信息且受到抑制IKr通道的索他洛尔摄入影响。

[0220] 实际上,能够提供血液中索他洛尔(或另一药物)的先前摄入的评估可以引起对TdP的风险的更好理解。

[0221] 已经将QT延长表明为关联到TdP,且因此许多研究将其视为TdP替代。然而,最新研究证实,由于高假阳性和假阴性速率(Hondeghem,2018年;同样参见单独利用QT鉴别Sot+和对照的AUC),因此仅QT延长并不是可信赖的预测物。

[0222] 如果患者确实服用或确实没有服用索他洛尔,不管药物消耗之后的时间(在4小时内),那么本文中所研发的模型能够准确地分类,其总体通用化准确性>94%。已经表明,多次获得与投票方法一起改进了分类。本身,此证实心电图内含有丰富信息的存在,比仅QT更丰富。

[0223] 接下来,已经表明这种替代模型还可以鉴别基于突变LQT-1、LQT-2以及LQT-3的类型的在不同程度下的长qt先天性患者。由于这些突变罕见且难以精简,因此这非常重要。

[0224] 这种分类方法在顺应对照的情况下且在药物监测的背景下调整剂量非常有用。此外,考虑到降低医疗保健相关成本的重要性,此系统将比标准血液分析更便宜、更实际且更快。其将允许具有持续的药物评估,而不是离散的结果。

[0225] 其它研发

[0226] 研发了各种模型:

[0227] 当仅利用一个导联时,已经表明可以获得与上文详细描述の利用来自8个导联的ECG的模型相同类型的结果(准确性)。

[0228] 研发另一模型,所述模型利用来自对照、已经摄取了索他洛尔的患者、LQT-1、LQT-2以及LQT-3患者的ECG作为训练集。此模型将提供患者在五种类别的每一种中的概率作为输出。利用投票系统(优选地通过提供患者的多个ECG或替代地通过利用多个CNN模型),有可能为医生提供信息并且允许适当护理。

[0229] 还可以利用多种TdP潜在地诱发药物以能够鉴别患者已经服用的药物来研发并丰富模型。

[0230] 尖端扭转型室性心动过速事件的检测

[0231] 发明人将模型应用于来自41个患者的群组的ECG ($n=1080$)，患者具有随访记录 (平均值=26、sd=16, 可分析的1065个) 和心脏病科医师因而标注的至少1个TdP事件。相比于距离事件超过24小时的相同个体的ECG, 在TdP事件的24小时内的风险分数更高。

[0232] 风险分数中的此差异显著不同 (AUROC: 0.61 ± 0.029 , 图2, $p < 0.0001$)。应注意, 使用来自处于TdP的风险下的已经接受药物 (QT延长药物, 主要为IKr阻滞剂) 的患者的数据来绘制此接受者操作特征曲线。处于0.61的AUROC表明有可能在这些患者内鉴别在24小时内患有TdP事件的患者, 尽管这些患者由于药物治疗都患有增加的QT。

[0233] 在所研发的模型中, 输出为从0到100的分数。分数越高, 风险越高。

[0234] 表1提供用于各种阈值的一些特异性和灵敏度值。

[0235]	阈值	特异性	灵敏度
	0.5000	0.6309	100.0
	3.500	3.365	99.12
[0236]	5.500	4.837	98.25
	10.50	8.202	95.61

[0237] 表1用于根据在ECG时不患有TdP的患者的ECG的阈值, 在24h内检测尖端扭转型室性心动过速发作性症状的灵敏度和特异性。

[0238] 根据医生是否希望具有高特异性和高灵敏度, 由医生做出对阈值的选择。如上文所指示, 有可能利用两个截断值或阈值以使得针对低于第一截断值的患者具有良好NPV且增加高于第二截断值的患者的监测。

[0239] 机器的进一步的训练将在不降低灵敏度的情况下提高特异性。

[0240] 此数据表明利用来自已经施用QT延长药物的患者的数据适当地训练机器学习分类器 (这里为卷积神经网络) 使得有可能检测将患有TdP事件的人, 或替代地可以预测其在24h内不会患有TdP事件的人。这与Attia等人的结论形成对比, Attia等人说他们无法关于心律失常风险 (TdP) 的发生得出结论。

[0241] 具有一个导联的机器学习分类器

[0242] 利用仅来自一个导联的ECG数据训练如上文所公开的卷积神经网络 (CNN)。

[0243] 利用以下导联DI、DII、V1、V2、V3、V4、V5以及V6的每一个单独地进行这种训练。因此其引起获得9个模型 (一个模型针对每个单独导联且一个模型具有通过对来自八个导联的信号求平均的复合ECG)。

[0244] 每个训练后的系统然后通过输入用于训练其的导联的ECG数据进行测试, ECG数据来自如上文所公开的测试集。患者的分类与利用来自所有导联 (复合模型) 的具有范围介于0.88到0.98的AUROC的ECG数据训练的CNN同样准确。

[0245] 还应注意, 当将仅从一个导联获得的数据应用于基于来自另一导联 (或所有导联) 的ECG数据训练的模型时, 模型也是在运转。具体地, 当将从DII导联获得的数据应用于利用来自DII导联的ECG数据训练的模型时, AUROC (表示用以鉴别已经施用QT延长药物的患者与

尚未施用这种种类的药物的患者的能力)为0.98。当将这种DII ECG数据应用于复合训练模型(利用来自8个导联的数据训练)时,AUROC为0.94。当将这种DII ECG数据应用于V4 ECG训练模型(利用来自V4导联的数据训练)时,AUROC为0.90。当将V4 ECG数据应用于利用来自V5导联的数据训练的模型时,AUROC为0.90。

[0246] 此数据表明利用来自单个导联的ECG或利用从多个导联获得的复合ECG训练的模型可以与无论是来自单独导联还是来自复合导联的任何ECG数据一起使用,其中AUROC平均值(将来自复合ECG或作为输入或在训练后的模型中的来自单独导联dI、dII、V1、V2、V3、V4、V5或V6的ECG的数据的所有组合考虑在内)是0.877,其标准差为0.067。

[0247] 这些结果表明机器学习分类器可以在只利用仅来自单个导联的ECG数据恰当地训练时提供良好的分析输出。此为患者可以自己放置单个电极来测量ECG的应用开辟了道路,ECG可以被发送到集中式服务器且可以通过机器分类器进行分析,以便检测患者患有TdP的风险且在检测到这种风险的情况下允许医生行动。

[0248] 先天性长QT (LQT) 的鉴别

[0249] 还对利用来自独特导联的ECG数据训练的模型进行测试以确定它们鉴别LQT1、LQT2以及LQT3患者的能力。

[0250] 已经表明,有可能利用仅来自一个导联的ECG数据鉴别这种患者,其中V6和DI导联在针对这些导联的范围介于0.63到0.70的AUROC的情况下是最好的。此指示出利用延长QT药物训练的系统可以在先天性LQT的不同类型之间进行鉴别。

[0251] 总之,以上数据表明有可能利用ECG数据训练机器分类器,ECG数据为复合ECG数据或来自单个导联的ECG数据,且训练后的机器分类器可以:

[0252] -标识在随后的24h内患有TdP事件的处于高风险下的患者;

[0253] -与已经用于训练的ECG数据不同的ECG数据一起利用;

[0254] -用于标识先天性LQT的本质(类型)。

[0255] 参考文献

[0256] Acharya,U.R、Fujita,H.、Liu,O.S.、Hagiwara,Y.、Tan,J.H.以及Adam,M. (2017年),利用卷积神经网络使用心动过速ECG片段的不同间期自动检测心律失常(Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network),信息科学,405,81-90。

[0257] Acharya,U.R.、Fujita,H.、Oh,S.L.、Hagiwara,Y.、Tan,J.H.以及Adam,M. (2017年),应用深度卷积神经网络以利用ECG信号自动检测心肌梗塞(Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals),信息科学,415,190-198。

[0258] Acharya,U.R.、Oh,S.L.、Hagiwara,Y.、Tan,J.H.、Adam,M.、Gertych,A.以及San Tan,R (2017年),用以对心跳进行分类的深度卷积神经网络模型(A deep convolutional neural network model to classify heartbeats),生物学和药物中的计算机,89,389-396。

[0259] Attia等人,2019年,J心血管电生理现象(J Cardiovasc Electrophysiol),2019年2月28日。

[0260] Attia,Z.I.、Sugrue,A.、Asirvatham,S.J.、Ackerman,M.J.、Kapa,S.、Friedman,

P.A.以及Noseworthy,P.A. (2018年),利用表面心电图的深度学习(神经网络)分析来无创性评定多非利特血浆浓度:概念研究证明(Noninvasive assessment of dofetilide plasma concentration using a deep learning(neural network)analysis of the surface electrocardiogram:A proof of concept study),PloS one,13(8),e0201059。

[0261] Hannun等人,2019年,自然医学,2019年1月;25(1):65-69。

[0262] Hondeghem,L.M. (2018年),药物诱发的QT延长和尖端扭转型室性心动加速:完全排它性的关系或友好分离时间?(Drug-Induced QT Prolongation and Torsades de Pointes:An All-Exclusive Relationship or Time for an Amicable Separation?),药物安全,41(1),11-17。

[0263] Liu等人,2018年,IEEE J生物医学健康信息(IEEE J Biomed Health Inform.)。2018年9月;22(5):1434-1444。

[0264] Pourbabaee,Roshtkhari以及Khorasani,2017年,基于系统、人以及控制论的IEEE会刊:系统(IEEE Transactions on Systems,Man,and Cybernetics:Systems),PP(99):1-10,2017年6月。

[0265] Rajpurkar,P.、Hannun,A.Y.、Haghpanahi,M.、Bourn,C.以及Ng,A.Y. (2017年),利用卷积神经网络的心脏病科医师级心律失常检测(Cardiologist-level arrhythmia detection with convolutional neural networks),arXiv preprint arXiv:1707.01836。

[0266] Tison等人,2018年,Int J医学信息(Int J Med Inform.),2018年12月;120:1-7。

[0267] Yıldırım 等人,2018年,生物学和医学中的计算机(Comput Biol Med.),2018年11月1日;102:411-420。

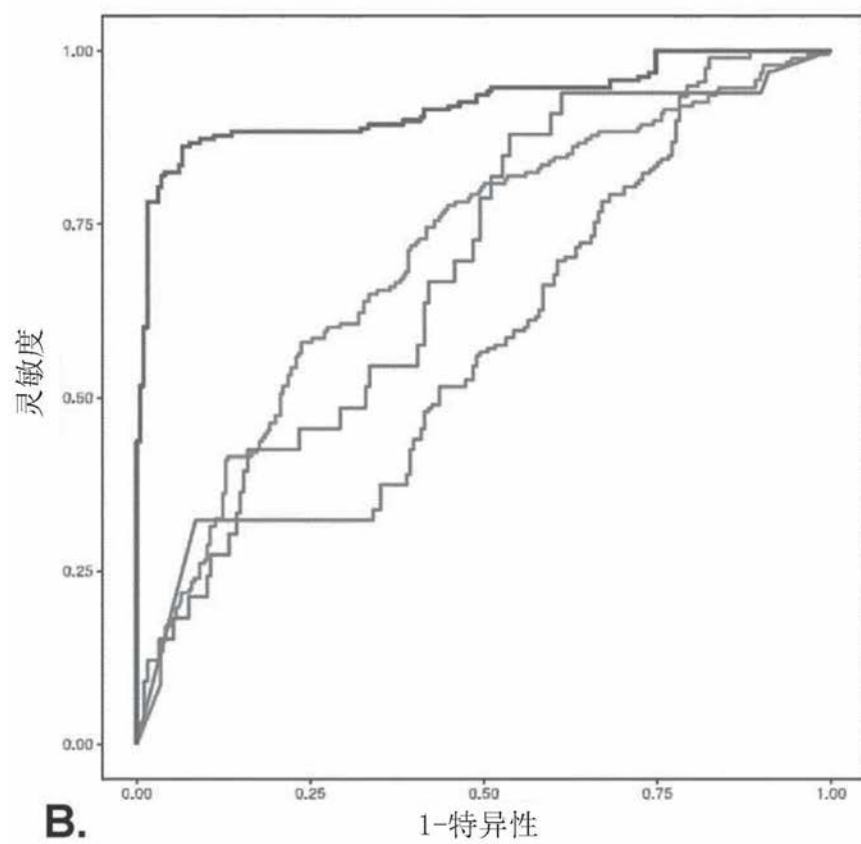
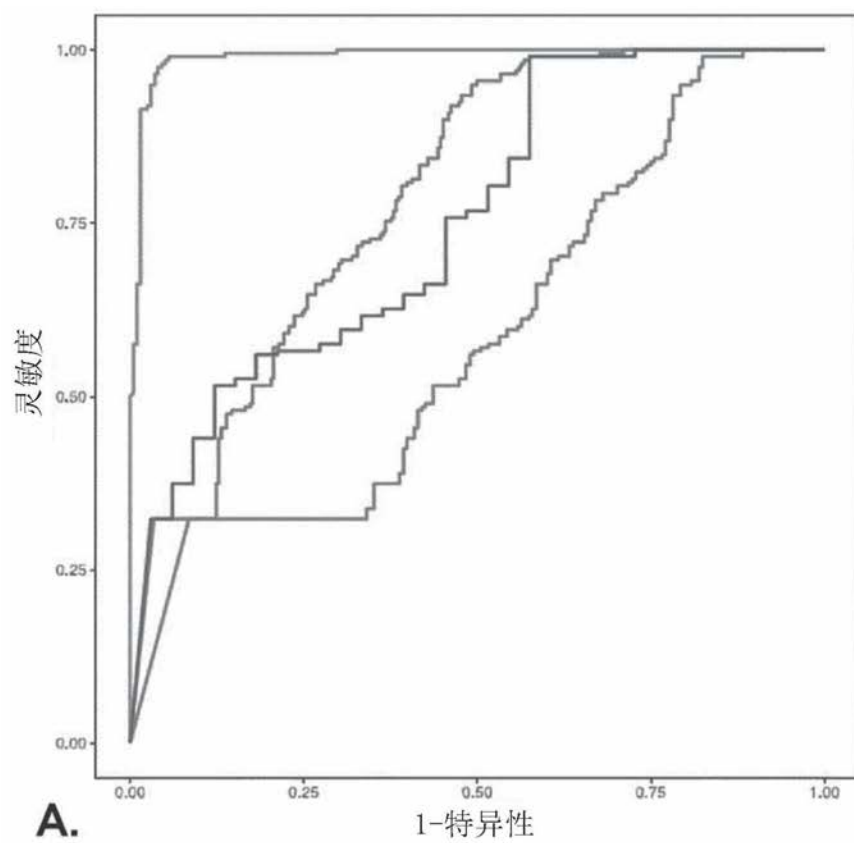


图1

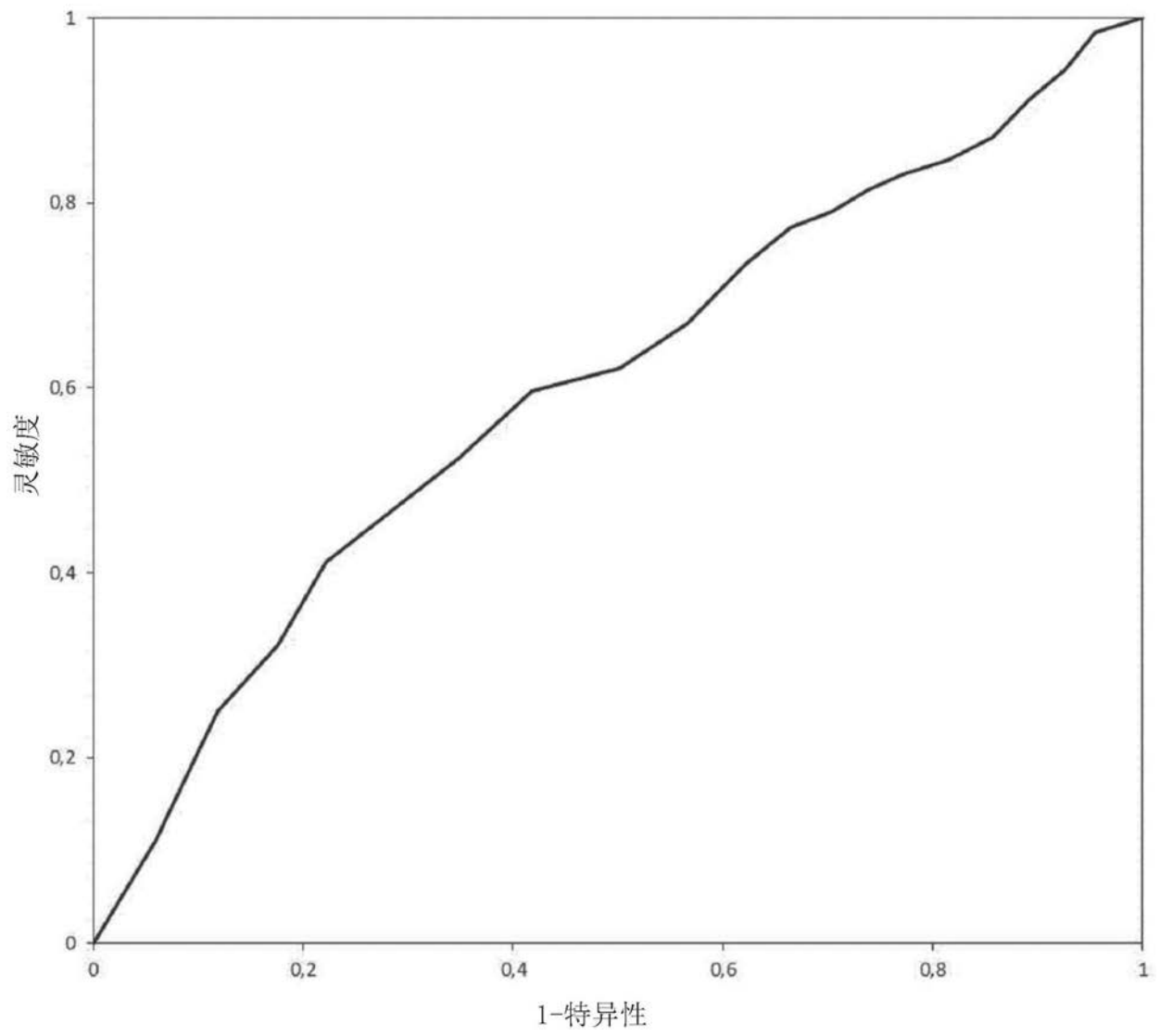


图2