



(12)发明专利申请

(10)申请公布号 CN 110189829 A

(43)申请公布日 2019.08.30

(21)申请号 201910552296.0

(22)申请日 2019.06.21

(71)申请人 深圳供电局有限公司

地址 518000 广东省深圳市罗湖区深南东路4020号电力调度通信大楼

(72)发明人 许克豪 何群 夏英华 曹蓉
刘国恒 洪紫慧 黄苑 姜虹
梁鹏翼

(74)专利代理机构 深圳汇智容达专利商标事务所(普通合伙) 44238

代理人 潘中毅

(51)Int.Cl.

G16H 50/50(2018.01)

G16H 50/30(2018.01)

A61B 5/00(2006.01)

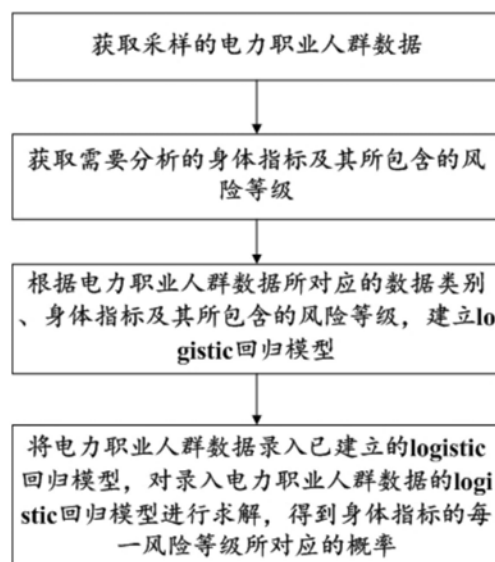
权利要求书2页 说明书12页 附图1页

(54)发明名称

一种电力职业人群重大慢性病风险评估方法及系统

(57)摘要

本发明提供一种电力职业人群重大慢性病风险评估方法,其包括下述步骤:获取采样的电力职业人群数据,电力职业人群数据所对应的数据类别包括:个人基本信息数据、个人生活习惯数据、心理健康状况数据、职业相关因素数据、体检数据、人资数据;获取需要分析的身体指标及其所包含的风险等级;根据电力职业人群数据所对应的数据类别、身体指标及其所包含的风险等级,建立logistic回归模型;将电力职业人群数据录入已建立的logistic回归模型,对录入电力职业人群数据的logistic回归模型进行求解,得到身体指标的每一风险等级所对应的概率。本发明可以对电力职业人群的健康状态进行评估和提示,对电力职业人群的健康管理具有重大意义。



1. 一种电力职业人群重大慢性病风险评估方法,其特征在于,包括下述步骤:

获取采样的电力职业人群数据,所述电力职业人群数据所对应的数据类别包括:个人基本信息数据、个人生活习惯数据、心理健康状况数据、职业相关因素数据、体检数据、人资数据;

获取需要分析的身体指标及其所包含的风险等级;

根据所述电力职业人群数据所对应的数据类别、所述身体指标及其所包含的风险等级,建立logistic回归模型;

将所述电力职业人群数据录入已建立的logistic回归模型,对录入所述电力职业人群数据的logistic回归模型进行求解,得到所述身体指标的每一风险等级所对应的概率。

2. 根据权利要求1所述的电力职业人群重大慢性病风险评估方法,其特征在于,所述身体指标包含有:身体质量指数、血压指数、血糖指数中的至少一种;所述身体质量指数包含的风险等级有:低体重、正常、肥胖前期、一级肥胖、二级肥胖;所述血压指数包含的风险等级有:正常、正常高压、高血压;所述血糖指数包含的风险等级有:正常、糖尿病前期、糖尿病。

3. 根据权利要求1所述的电力职业人群重大慢性病风险评估方法,其特征在于,logistic回归模型的数量为 $g-1$ 个, g 为所述身体指标对应的风险等级的数量,其中第 j 个logistic回归模型为:

$$\text{Logit}(\sum_{i=1}^j P_i) = \beta_{0j} + \sum_{n=1}^m \beta_n * X_n, m \text{ 为电力职业人群数据总量, } \beta_n \text{ 为logistic回归模型中}$$

每个自变量 X_n 的回归系数, β_{0j} 为logistic回归模型的截距值, β_i 和 β_{0j} 通过最大似然法对logistic回归模型进行拟合得到, X_n 为具体的电力职业人群数据。

4. 根据权利要求3所述的电力职业人群重大慢性病风险评估方法,其特征在于,对录入所述电力职业人群数据的logistic回归模型进行求解,具体为:

利用下述公式对logistic回归模型进行求解:

$$\text{当 } k=1 \text{ 时, } P_k = \frac{1}{1 + \exp(f(k))};$$

$$\text{当 } g > k > 1 \text{ 时, } P_k = \frac{1}{1 + \exp(f(k))} - \frac{1}{1 + \exp(f(k-1))};$$

$$\text{其中 } f(k) = -\text{Logit}(\sum_{i=1}^k p_i), P_k \text{ 为所述身体指标的第 } k \text{ 个风险等级对应的概率。}$$

5. 根据权利要求1所述的电力职业人群重大慢性病风险评估方法,其特征在于,还包括下述步骤:

利用logistic回归模型的回归系数计算用于衡量危险因素作用大小的优势比。

6. 一种电力职业人群重大慢性病风险评估系统,其特征在于,包括:

数据获取模块,用于获取采样的电力职业人群数据,所述电力职业人群数据所对应的数据类别包括:个人基本信息数据、个人生活习惯数据、心理健康状况数据、职业相关因素数据、体检数据、人资数据;

身体指标及风险等级获取模块,用于获取需要分析的身体指标及其所包含的风险等

级;

模型构建模块,用于根据所述电力职业人群数据所对应的数据类别、所述身体指标及其所包含的风险等级,建立logistic回归模型;

风险计算模块,用于将所述电力职业人群数据录入已建立的logistic回归模型,对录入所述电力职业人群数据的logistic回归模型进行求解,得到所述身体指标的每一风险等级所对应的概率。

7.根据权利要求6所述的电力职业人群重大慢性病风险评估系统,其特征在于,所述身体指标包含有:身体质量指数、血压指数、血糖指数中的至少一种;所述身体质量指数包含的风险等级有:低体重、正常、肥胖前期、一级肥胖、二级肥胖;所述血压指数包含的风险等级有:正常、正常高压、高血压;所述血糖指数包含的风险等级有:正常、糖尿病前期、糖尿病。

8.根据权利要求6所述的电力职业人群重大慢性病风险评估系统,其特征在于,logistic回归模型的数量为 $g-1$ 个, g 为所述身体指标对应的风险等级的数量,其中第 j 个logistic回归模型为:

$$\text{Logit}(\sum_{i=1}^j P_i) = \beta_{0j} + \sum_{n=1}^m \beta_n * X_n$$
, m 为电力职业人群数据总量, β_n 为logistic回归模型中每个自变量 X_n 的回归系数, β_{0j} 为logistic回归模型的截距值, β_i 和 β_{0j} 通过最大似然法对logistic回归模型进行拟合得到, X_n 为具体的电力职业人群数据。

9.根据权利要求8所述的电力职业人群重大慢性病风险评估系统,其特征在于,所述风险计算模块,进一步地用于:

利用下述公式对logistic回归模型进行求解:

$$\text{当 } k=1 \text{ 时, } P_k = \frac{1}{1 + \exp(f(k))};$$

$$\text{当 } g > k > 1 \text{ 时, } P_k = \frac{1}{1 + \exp(f(k))} - \frac{1}{1 + \exp(f(k-1))};$$

其中 $f(k) = -\text{Logit}(\sum_{i=1}^k p_i)$, P_k 为所述身体指标的第 k 个风险等级对应的概率。

10.根据权利要求6所述的电力职业人群重大慢性病风险评估系统,其特征在于,还包括:

优势比计算模块,用于利用logistic回归模型的回归系数计算用于衡量危险因素作用大小的优势比。

一种电力职业人群重大慢性病风险评估方法及系统

技术领域

[0001] 本发明涉及医疗技术领域,尤其涉及一种电力职业人群重大慢性病风险评估方法及系统。

背景技术

[0002] 医疗健康大数据是人类与医疗及生命健康相关的活动过程中产生的数据集合,对于健康数据的有效利用在疾病的预防、诊疗与康复中起到重要作用。尤其是在新时期“互联网+”在健康管理领域的不断发展,对于多维健康数据的重视不断加强。

[0003] 首先是客观的发展需求。重大慢性病,包括心血管疾病、糖尿病、肿瘤等给国家和企业带来沉重疾病负担,在生命的全周期开展慢性病的一级预防、二级预防和三级预防是减少患病率、致残率、提高健康寿命年的最有效措施。三级预防的关键有两点,一是找对影响疾病发生、发展的危险因素。电力人群具有职业特殊性,除了普通人群具有的危险因素外,还存在职业危险因素,但目前对这些因素认识不足。二是识别个体的健康状态、发现危险因素暴露并采取预防措施,这需要个体准确的数据和科学预防技术。

[0004] 第二个背景是企业健康数据的沉积和挖掘不足。大量的企业定期开展职工体检、工作环境监测等,数据量越积越多,但不具备大数据处理能力、统计处理能力和医疗卫生知识。一个能帮助企业快速识别人群健康状态并提示健康风险的工具对企业的健康管理有很大意义。

发明内容

[0005] 为解决上述技术问题,本发明提供一种电力职业人群重大慢性病风险评估方法及系统,可以对电力职业人群的健康状态进行评估和提示,对电力职业人群的健康管理具有重大意义。

[0006] 本发明提供了一种电力职业人群重大慢性病风险评估方法,其包括下述步骤:

[0007] 获取采样的电力职业人群数据,所述电力职业人群数据所对应的数据类别包括:个人基本信息数据、个人生活习惯数据、心理健康状况数据、职业相关因素数据、体检数据、人资数据;

[0008] 获取需要分析的身体指标及其所包含的风险等级;

[0009] 根据所述电力职业人群数据所对应的数据类别、所述身体指标及其所包含的风险等级,建立logistic回归模型;

[0010] 将所述电力职业人群数据录入已建立的logistic回归模型,对录入所述电力职业人群数据的logistic回归模型进行求解,得到所述身体指标的每一风险等级所对应的概率。

[0011] 优选地,所述身体指标包含有:身体质量指数、血压指数、血糖指数中的至少一种;所述身体质量指数包含的风险等级有:低体重、正常、肥胖前期、一级肥胖、二级肥胖;所述血压指数包含的风险等级有:正常、正常高压、高血压;所述血糖指数包含的风险等级有:正

常、糖尿病前期、糖尿病。

[0012] 优选地, logistic回归模型的数量为 $g-1$ 个, g 为所述身体指标对应的风险等级的数量, 其中第 j 个logistic回归模型为:

[0013] $\text{Logit}(\sum_{i=1}^j P_i) = \beta_{0j} + \sum_{n=1}^m \beta_n * X_n$, m 为电力职业人群数据总量, β_n 为logistic回归模型中每个自变量 X_n 的回归系数, β_{0j} 为logistic回归模型的截距值, β_i 和 β_{0j} 通过最大似然法对logistic回归模型进行拟合得到, X_n 为具体的电力职业人群数据。

[0014] 优选地, 对录入所述电力职业人群数据的logistic回归模型进行求解, 具体为:

[0015] 利用下述公式对logistic回归模型进行求解:

[0016] 当 $k=1$ 时, $P_k = \frac{1}{1 + \exp(f(k))}$;

[0017] 当 $g > k > 1$ 时, $P_k = \frac{1}{1 + \exp(f(k))} - \frac{1}{1 + \exp(f(k-1))}$, 其中 $f(k) = -\text{Logit}(\sum_{i=1}^k p_i)$, P_k 为

所述身体指标的第 k 个风险等级对应的概率。

[0018] 优选地, 还包括下述步骤:

[0019] 利用logistic回归模型的回归系数计算用于衡量危险因素作用大小的优势比。

[0020] 本发明还提供一种电力职业人群重大慢性病风险评估系统, 其包括:

[0021] 数据获取模块, 用于获取采样的电力职业人群数据, 所述电力职业人群数据所对应的数据类别包括: 个人基本信息数据、个人生活习惯数据、心理健康状况数据、职业相关因素数据、体检数据、人资数据;

[0022] 身体指标及风险等级获取模块, 用于获取需要分析的身体指标及其所包含的风险等级;

[0023] 模型构建模块, 用于根据所述电力职业人群数据所对应的数据类别、所述身体指标及其所包含的风险等级, 建立logistic回归模型;

[0024] 风险计算模块, 用于将所述电力职业人群数据录入已建立的logistic回归模型, 对录入所述电力职业人群数据的logistic回归模型进行求解, 得到所述身体指标的每一风险等级所对应的概率。

[0025] 优选地, 所述身体指标包含有: 身体质量指数、血压指数、血糖指数中的至少一种; 所述身体质量指数包含的风险等级有: 低体重、正常、肥胖前期、一级肥胖、二级肥胖; 所述血压指数包含的风险等级有: 正常、正常高压、高血压; 所述血糖指数包含的风险等级有: 正常、糖尿病前期、糖尿病。

[0026] 优选地, logistic回归模型的数量为 $g-1$ 个, g 为所述身体指标对应的风险等级的数量, 其中第 j 个logistic回归模型为:

[0027] $\text{Logit}(\sum_{i=1}^j P_i) = \beta_{0j} + \sum_{n=1}^m \beta_n * X_n$, m 为电力职业人群数据总量, β_n 为logistic回归模型中每个自变量 X_n 的回归系数, β_{0j} 为logistic回归模型的截距值, β_i 和 β_{0j} 通过最大似然法对logistic回归模型进行拟合得到, X_n 为具体的电力职业人群数据。

[0028] 优选地, 所述风险计算模块, 进一步地用于:

[0029] 利用下述公式对logistic回归模型进行求解：

[0030] 当 $k=1$ 时, $P_k = \frac{1}{1 + \exp(f(k))}$;

[0031] 当 $g > k > 1$ 时, $P_k = \frac{1}{1 + \exp(f(k))} - \frac{1}{1 + \exp(f(k-1))}$, 其中 $f(k) = -\text{Logit}(\sum_{i=1}^k p_i)$, P_k 为

所述身体指标的第 k 个风险等级对应的概率。

[0032] 优选地,还包括:

[0033] 优势比计算模块,用于利用logistic回归模型的回归系数计算用于衡量危险因素作用大小的优势比。

[0034] 实施本发明,具有如下有益效果:本发明将电力职业人群数据录入已建立的logistic回归模型,logistic回归模型由电力职业人群数据所对应的数据类别、身体指标及其所包含的风险等级构建而成的多分类有序logistic回归模型,对录入电力职业人群数据的logistic回归模型进行求解,得到所述身体指标的每一风险等级所对应的概率,可以对电力职业人群的健康状态进行评估和提示,对电力职业人群的健康管理具有重大意义。

附图说明

[0035] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0036] 图1是本发明提供的电力职业人群重大慢性病风险评估方法的流程图。

具体实施方式

[0037] 本发明提供一种电力职业人群重大慢性病风险评估方法,如图1所示,其包括下述步骤:

[0038] 获取采样的电力职业人群数据,电力职业人群数据所对应的数据类别包括:个人基本信息数据、个人生活习惯数据、心理健康状况数据、职业相关因素数据、体检数据、人资数据;

[0039] 获取需要分析的身体指标及其所包含的风险等级;

[0040] 根据电力职业人群数据所对应的数据类别、身体指标及其所包含的风险等级,建立有序分类的logistic回归模型;

[0041] 将电力职业人群数据录入已建立的logistic回归模型,对录入电力职业人群数据的logistic回归模型进行求解,得到身体指标的每一风险等级所对应的概率。

[0042] 具体地,上述的电力职业人群数据,可以通过调查问卷进行采样,或者从医院的数据库中,或者从企业的数据库中直接获取。

[0043] 进一步地,身体指标包含有:身体质量指数、血压指数、血糖指数中的至少一种;身体质量指数包含的风险等级有:低体重、正常、肥胖前期、一级肥胖、二级肥胖;血压指数包含的风险等级有:正常、正常高压、高血压;血糖指数包含的风险等级有:正常、糖尿病前期、

糖尿病。

[0044] 进一步地, logistic回归模型的数量为 $g-1$ 个, g 为身体指标对应的风险等级的数量, 其中第 j 个logistic回归模型为:

[0045] $\text{Logit}(\sum_{i=1}^j P_i) = \beta_{0j} + \sum_{n=1}^m \beta_n * X_n$, m 为电力职业人群数据总量, β_n 为logistic回归模型中每个自变量 X_n 的回归系数, β_{0j} 为logistic回归模型的截距值, β_i 和 β_{0j} 通过最大似然法对logistic回归模型进行拟合得到, X_n 为具体的电力职业人群数据。这里, 可以利用R软件相应的函数程序包拟合原始数据(电力职业人群数据), 输出结果。

[0046] 进一步地, 对录入电力职业人群数据的logistic回归模型进行求解, 具体为:

[0047] 利用下述公式对logistic回归模型进行求解:

[0048] 当 $k=1$ 时, $P_k = \frac{1}{1 + \exp(f(k))}$;

[0049] 当 $g > k > 1$ 时, $P_k = \frac{1}{1 + \exp(f(k))} - \frac{1}{1 + \exp(f(k-1))}$, 其中 $f(k) = -\text{Logit}(\sum_{i=1}^k p_i)$, P_k 为

身体指标的第 k 个风险等级对应的概率;

[0050] 当 $g=k$ 时, $P_k = 1 - \sum_{i=1}^{g-1} P_i$ 。

[0051] 电力职业人群重大慢性病风险评估方法还包括下述步骤:

[0052] 利用logistic回归模型的回归系数计算用于衡量危险因素作用大小的优势比。

[0053] 电力职业人群数据如下表所示:

[0054]

序号	数据类别	具体数据指标
1	个人基本信息	学历
2		婚姻状况
3		身体状况
4		直系亲属家族史
5		旁系亲属家族史
6	个人生活习惯	吃早餐频率
7		用餐习惯
8		叫外卖频率
9		吃速食类频率
10		食盐习惯
11		腌制品食用频率
12		油脂习惯
13		喝老火汤频率
14		保健品食用习惯
15		吃零食习惯

[0055]	16		含糖饮料饮用习惯
	17		运动习惯
	18		吸烟习惯
	19		饮酒习惯
	20		睡眠习惯
	21		口腔卫生
	22	心理健康状况	工作满意程度
	23		精神压力来源
	24		精神压力程度
	25		躯体化反应
	26		焦虑程度
	27		抑郁程度
	28	职业相关因素	服务对象
	29		劳动强度
	30		工作内容
	31		工作危险程度
	32		健康影响因素
	33		职业防护措施
	34	体检数据	数据示例
	35	人资数据	工种
	36		岗位序列
	37		职务类别

[0056] 身体质量指数对应的具体数据参数如下：

[0057]	数据类别	具体数据	参数代码
	体检数据	体质指数分类	BMI.catabc
	个人基本信息	性别	genderabc
		年龄	age.catabc
		工龄	workyear.cat

[0058]			abc
		学历	Nq1abc
		婚姻状况	Nq2.Nabc
		直系亲属是否患有肥胖	Nq4.7abc
		旁系亲属是否患有肥胖	Nq5.7abc
	人资数据	工种	GZModelCodeabc
		岗位序列	GWXLModelCodeabc
		职务级别	ZWJBModelCodeabc
	个人生活习惯	您通常的用餐习惯	Nq7abc
		您每日的食盐习惯	Nq10abc
		您日常使用油脂的习惯	Nq12abc
		您一周内有多少天吃零食,不包含水果	Nq15abc
		通常一周内您有多少天饮含糖饮料	Nq17abc
		通常一周内您有多少天锻炼	Nq18abc
		您是否吸烟	Nq23abc
		通常一周您有多少天饮酒	Nq29abc
	心理健康状况	工作带给您的压力程度	Nq41abc
		家庭生活带给您的压力程度	Nq43abc
		个人生活带给您的压力程度	Nq45abc
	职业相关因素	您工作的体力劳动强度	Nq51abc
		您工作的脑力劳动强度	Nq52abc

[0059] 一般的线性回归只适用于应变量为连续变量的情况,此时可以采用logistic回归分析。logistic回归是一种研究两水平或多水平的反应变量与影响因素间曲线关系的回归分析。logistic回归应变变量可以是二分类的,也可以是多分类的,二分类的情况具有一般性,但在实际中也可能会遇到多分类的情况。多分类logistic回归模型又可以分成两种不同的情况,第一种为反应变量是无序分类资料,如血型分为:A型、B型、AB型、O型;第二种为

反应变量是有序分类资料(等级资料),如疗效分为:痊愈、显效、微效、无效。由于应变变量BMI(身体质量指数)分为低体重、正常、肥胖前期、一级肥胖、二级肥胖等五个风险等级,BP(血压)分为正常、正常高压、高血压等三个风险等级,GLU(血糖)分为正常、糖尿病前期、糖尿病等三个风险等级。对该数据需要采用有序分类logistic回归模型进行分析。

[0060] 以5水平的BMI指数为例,应变变量取值为1、2、3、4、5,分别对应身体质量指数的五个风险等级,相应概率取值水平为 p_1 、 p_2 、 p_3 、 p_4 、 p_5 。即某职工经过模型预测,其BMI指数处于低体重、正常、肥胖前期、一级肥胖、二级肥胖的风险概率为 p_1 、 p_2 、 p_3 、 p_4 、 p_5 ($p_1+p_2+p_3+p_4+p_5=1$)。

[0061] 根据有序分类logistic回归原理,可建立4个有序回归logistic回归模型:

$$[0062] \quad \text{logit}(p_1) = \beta_{01} + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m$$

$$[0063] \quad \text{logit}(p_1+p_2) = \beta_{02} + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m$$

$$[0064] \quad \text{logit}(p_1+p_2+p_3) = \beta_{03} + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m$$

$$[0065] \quad \text{logit}(p_1+p_2+p_3+p_4) = \beta_{04} + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m$$

$$[0066] \quad P_1 = P_{(Y \leq 1)} = \frac{1}{1 + \exp[-\text{logit}(P_1)]}$$

$$[0067] \quad P_2 = \frac{1}{1 + \exp[-\text{logit}(P_1 + P_2)]} - \frac{1}{1 + \exp[-\text{logit}(P_1)]}$$

$$[0068] \quad P_j = P_{(Y \leq j)} - P_{(Y \leq j-1)} = \frac{1}{1 + \exp[-\text{logit}(P_1 + P_2 + \dots + P_j)]} - \frac{1}{1 + \exp[-\text{logit}(P_1 + P_2 + \dots + P_{j-1})]}$$

$$[0069] \quad P_5 = 1 - (P_1 + P_2 + P_3 + P_4)$$

[0070] 以身体质量预测指数为例,需要说明的是“==”表示判断该变量是否属于该组或者等于该值,若是则为1,否则为0,例如性别男对应的genderabc应该为1,性别女对应的genderabc应该为2,性别男的(genderabc==2)的值为0,性别男的(genderabc==1)的值为1。

$$[0071] \quad \begin{aligned} \text{logit}(p_1) = & -5.2527 + (-1.575729613) * (\text{genderabc} == 2) + (0.19705621) * \\ & (\text{age.catabc} == 2) + (0.189338625) * (\text{age.catabc} == 3) + (0.279641929) * (\text{age.catabc} == 4) + (0.098110654) * (\text{workyear.catabc} == 2) + (0.319055879) * (\text{workyear.catabc} == 3) + (-0.023581518) * (\text{Nq1abc} == 2) + (-0.530720192) * (\text{Nq1abc} == 3) + (-0.196035291) * \\ & (\text{Nq2.Nabc} == 2) + (0.182739173) * (\text{Nq2.Nabc} == 3) + (0.181107888) * (\text{GZModelcodeabc} == 2) + (-0.143577901) * (\text{GZModelcodeabc} == 3) + (0.007667259) * (\text{GZModelcodeabc} == 4) + (-0.664537602) * (\text{GWXLModelcodeabc} == 2) + (-0.872149689) * (\text{GWXLModelcodeabc} == 3) + (-0.890754864) * (\text{Nq4.7abc} == 2) + (-0.390571293) * (\text{Nq5.7abc} == 2) + (-0.10208856) * (\text{Nq7abc} == 2) + (-0.086848718) * (\text{Nq7abc} == 3) + (0.166935229) * (\text{Nq10abc} == 2) + (0.24967892) * (\text{Nq10abc} == 3) + (-0.676813202) * (\text{Nq10abc} == 4) + (0.071105476) * (\text{Nq12abc} == 2) + (0.323905897) * (\text{Nq12abc} == 3) + (1.415612305) * (\text{Nq12abc} == 4) + (0.115952426) * (\text{Nq15abc} == 2) + (0.392958029) * (\text{Nq15abc} == 3) + (0.218606797) * (\text{Nq15abc} == 4) + (-0.084661004) * (\text{Nq17abc} == 2) + (-0.238055324) * (\text{Nq17abc} == 3) + (-0.51163021) * (\text{Nq17abc} == 4) + (0.087371152) * (\text{Nq18abc} == 2) + (0.083748491) * (\text{Nq18abc} == 3) + (0.24237703) * (\text{Nq18abc} == 4) + (0.038279688) * \end{aligned}$$

$(Nq23abc == 2) + (0.399948596) * (Nq23abc == 3) + (0.016914244) * (Nq29abc == 2) +$
 $(0.347088562) * (Nq29abc == 3) + (-2.838408819) * (Nq29abc == 4) + (-0.018402592) * (Nq41abc == 2) + (-0.2772015) * (Nq41abc == 3) + (-0.027672228) * (Nq41abc == 4) + (-0.179128728) * (Nq43abc == 2) + (-0.414707117) * (Nq43abc == 3) + (-0.720949549) * (Nq43abc == 4) + (0.136823021) * (Nq45abc == 2) + (0.366859804) * (Nq45abc == 3) + (0.055216777) * (Nq45abc == 4) + (0.410833021) * (Nq51abc == 2) + (0.51874652) * (Nq51abc == 3) + (-0.175404155) * (Nq52abc == 2) + (-0.238831047) * (Nq52abc == 3)。$

[0072] $\text{logit}(p1+p2) = -1.8182 + (-1.575729613) * (\text{genderabc} == 2) + (0.19705621) * (\text{age.catabc} == 2) + (0.189338625) * (\text{age.catabc} == 3) + (0.279641929) * (\text{age.catabc} == 4) + (0.098110654) * (\text{workyear.catabc} == 2) + (0.319055879) * (\text{workyear.catabc} == 3) + (-0.023581518) * (Nq1abc == 2) + (-0.530720192) * (Nq1abc == 3) + (-0.196035291) * (Nq2.Nabc == 2) + (0.182739173) * (Nq2.Nabc == 3) + (0.181107888) * (GZModelcodeabc == 2) + (-0.143577901) * (GZModelcodeabc == 3) + (0.007667259) * (GZModelcodeabc == 4) + (-0.664537602) * (GWXLModelcodeabc == 2) + (-0.872149689) * (GWXLModelcodeabc == 3) + (-0.890754864) * (Nq4.7abc == 2) + (-0.390571293) * (Nq5.7abc == 2) + (-0.10208856) * (Nq7abc == 2) + (-0.086848718) * (Nq7abc == 3) + (0.166935229) * (Nq10abc == 2) + (0.24967892) * (Nq10abc == 3) + (-0.676813202) * (Nq10abc == 4) + (0.071105476) * (Nq12abc == 2) + (0.323905897) * (Nq12abc == 3) + (1.415612305) * (Nq12abc == 4) + (0.115952426) * (Nq15abc == 2) + (0.392958029) * (Nq15abc == 3) + (0.218606797) * (Nq15abc == 4) + (-0.084661004) * (Nq17abc == 2) + (-0.238055324) * (Nq17abc == 3) + (-0.51163021) * (Nq17abc == 4) + (0.087371152) * (Nq18abc == 2) + (0.083748491) * (Nq18abc == 3) + (0.24237703) * (Nq18abc == 4) + (0.038279688) * (Nq23abc == 2) + (0.399948596) * (Nq23abc == 3) + (0.016914244) * (Nq29abc == 2) + (0.347088562) * (Nq29abc == 3) + (-2.838408819) * (Nq29abc == 4) + (-0.018402592) * (Nq41abc == 2) + (-0.2772015) * (Nq41abc == 3) + (-0.027672228) * (Nq41abc == 4) + (-0.179128728) * (Nq43abc == 2) + (-0.414707117) * (Nq43abc == 3) + (-0.720949549) * (Nq43abc == 4) + (0.136823021) * (Nq45abc == 2) + (0.366859804) * (Nq45abc == 3) + (0.055216777) * (Nq45abc == 4) + (0.410833021) * (Nq51abc == 2) + (0.51874652) * (Nq51abc == 3) + (-0.175404155) * (Nq52abc == 2) + (-0.238831047) * (Nq52abc == 3)。$

[0073] $\text{logit}(p1+p2+p3) = -0.7690 + (-1.575729613) * (\text{genderabc} == 2) + (0.19705621) * (\text{age.catabc} == 2) + (0.189338625) * (\text{age.catabc} == 3) + (0.279641929) * (\text{age.catabc} == 4) + (0.098110654) * (\text{workyear.catabc} == 2) + (0.319055879) * (\text{workyear.catabc} == 3) + (-0.023581518) * (Nq1abc == 2) + (-0.530720192) * (Nq1abc == 3) + (-0.196035291) * (Nq2.Nabc == 2) + (0.182739173) * (Nq2.Nabc == 3) + (0.181107888) * (GZModelcodeabc == 2) + (-0.143577901) * (GZModelcodeabc == 3) + (0.007667259) * (GZModelcodeabc == 4) + (-0.664537602) * (GWXLModelcodeabc == 2) + (-0.872149689) * (GWXLModelcodeabc == 3) + (-0.890754864) * (Nq4.7abc == 2) + (-0.390571293) * (Nq5.7abc == 2) + (-0.10208856) * (Nq7abc == 2) + (-0.086848718) * (Nq7abc == 3) + (0.166935229) * (Nq10abc == 2) + (0.24967892) * (Nq10abc == 3) + (-$

$0.676813202) * (Nq10abc == 4) + (0.071105476) * (Nq12abc == 2) + (0.323905897) * (Nq12abc == 3) + (1.415612305) * (Nq12abc == 4) + (0.115952426) * (Nq15abc == 2) + (0.392958029) * (Nq15abc == 3) + (0.218606797) * (Nq15abc == 4) + (-0.084661004) * (Nq17abc == 2) + (-0.238055324) * (Nq17abc == 3) + (-0.51163021) * (Nq17abc == 4) + (0.087371152) * (Nq18abc == 2) + (0.083748491) * (Nq18abc == 3) + (0.24237703) * (Nq18abc == 4) + (0.038279688) * (Nq23abc == 2) + (0.399948596) * (Nq23abc == 3) + (0.016914244) * (Nq29abc == 2) + (0.347088562) * (Nq29abc == 3) + (-2.838408819) * (Nq29abc == 4) + (-0.018402592) * (Nq41abc == 2) + (-0.2772015) * (Nq41abc == 3) + (-0.027672228) * (Nq41abc == 4) + (-0.179128728) * (Nq43abc == 2) + (-0.414707117) * (Nq43abc == 3) + (-0.720949549) * (Nq43abc == 4) + (0.136823021) * (Nq45abc == 2) + (0.366859804) * (Nq45abc == 3) + (0.055216777) * (Nq45abc == 4) + (0.410833021) * (Nq51abc == 2) + (0.51874652) * (Nq51abc == 3) + (-0.175404155) * (Nq52abc == 2) + (-0.238831047) * (Nq52abc == 3)。$

[0074] $\text{logit}(p_1 + p_2 + p_3 + p_4) = 1.6484 + (-1.575729613) * (\text{genderabc} == 2) + (0.19705621) * (\text{age.catabc} == 2) + (0.189338625) * (\text{age.catabc} == 3) + (0.279641929) * (\text{age.catabc} == 4) + (0.098110654) * (\text{workyear.catabc} == 2) + (0.319055879) * (\text{workyear.catabc} == 3) + (-0.023581518) * (Nq1abc == 2) + (-0.530720192) * (Nq1abc == 3) + (-0.196035291) * (Nq2.Nabc == 2) + (0.182739173) * (Nq2.Nabc == 3) + (0.181107888) * (GZModelcodeabc == 2) + (-0.143577901) * (GZModelcodeabc == 3) + (0.007667259) * (GZModelcodeabc == 4) + (-0.664537602) * (GWXLModelcodeabc == 2) + (-0.872149689) * (GWXLModelcodeabc == 3) + (-0.890754864) * (Nq4.7abc == 2) + (-0.390571293) * (Nq5.7abc == 2) + (-0.10208856) * (Nq7abc == 2) + (-0.086848718) * (Nq7abc == 3) + (0.166935229) * (Nq10abc == 2) + (0.24967892) * (Nq10abc == 3) + (-0.676813202) * (Nq10abc == 4) + (0.071105476) * (Nq12abc == 2) + (0.323905897) * (Nq12abc == 3) + (1.415612305) * (Nq12abc == 4) + (0.115952426) * (Nq15abc == 2) + (0.392958029) * (Nq15abc == 3) + (0.218606797) * (Nq15abc == 4) + (-0.084661004) * (Nq17abc == 2) + (-0.238055324) * (Nq17abc == 3) + (-0.51163021) * (Nq17abc == 4) + (0.087371152) * (Nq18abc == 2) + (0.083748491) * (Nq18abc == 3) + (0.24237703) * (Nq18abc == 4) + (0.038279688) * (Nq23abc == 2) + (0.399948596) * (Nq23abc == 3) + (0.016914244) * (Nq29abc == 2) + (0.347088562) * (Nq29abc == 3) + (-2.838408819) * (Nq29abc == 4) + (-0.018402592) * (Nq41abc == 2) + (-0.2772015) * (Nq41abc == 3) + (-0.027672228) * (Nq41abc == 4) + (-0.179128728) * (Nq43abc == 2) + (-0.414707117) * (Nq43abc == 3) + (-0.720949549) * (Nq43abc == 4) + (0.136823021) * (Nq45abc == 2) + (0.366859804) * (Nq45abc == 3) + (0.055216777) * (Nq45abc == 4) + (0.410833021) * (Nq51abc == 2) + (0.51874652) * (Nq51abc == 3) + (-0.175404155) * (Nq52abc == 2) + (-0.238831047) * (Nq52abc == 3)。$

[0075] 当得到每个变量的回归系数 β_j 值后,根据logistic回归的原理,回归系数 β_j 表示自变量 X_j 每变化一个单位时,logit(p)的改变量。 β_j 与衡量危险因素作用大小的优势比(odds ratio,OR)有一个对应关系。即 $OR_j = \exp(\beta_j)$ 。

[0076] 以BMI指数中的自变量“工作的体力劳动强度”为例：将自变量中的“高强度”等级作为参照变量，“一般强度”和“低强度”对其的OR值分别为1.50和1.68，即表示“一般强度”和“低强度”的体力劳动强度其使BMI指数升高一个等级的可能性是“高强度”体力劳动强度的1.50倍和1.68倍，是其危险因素（即OR>1时是危险因素）。相反的，变量“直系亲属是否患有肥胖”中，选“有肥胖”为参照，“无肥胖”的OR值为0.41，即“直系亲属无肥胖”的职工BMI指数升高一个等级的可能性是“直系亲属有肥胖”的职工的0.41倍，更不容易患肥胖（即OR<1时是保护因素）。

[0077] 本发明还提供一种电力职业人群重大慢性病风险评估系统，其包括：数据获取模块、身体指标及风险等级获取模块、模型构建模块、风险计算模块。

[0078] 数据获取模块用于获取采样的电力职业人群数据，电力职业人群数据所对应的数据类别包括：个人基本信息数据、个人生活习惯数据、心理健康状况数据、职业相关因素数据、体检数据、人资数据。

[0079] 身体指标及风险等级获取模块用于获取需要分析的身体指标及其所包含的风险等级。

[0080] 模型构建模块用于根据电力职业人群数据所对应的数据类别、身体指标及其所包含的风险等级，建立logistic回归模型。

[0081] 风险计算模块用于将电力职业人群数据录入已建立的logistic回归模型，对录入电力职业人群数据的logistic回归模型进行求解，得到身体指标的每一风险等级所对应的概率。

[0082] 进一步地，身体指标包含有：身体质量指数、血压指数、血糖指数中的至少一种；身体质量指数包含的风险等级有：低体重、正常、肥胖前期、一级肥胖、二级肥胖；血压指数包含的风险等级有：正常、正常高压、高血压；血糖指数包含的风险等级有：正常、糖尿病前期、糖尿病。

[0083] 进一步地，logistic回归模型的数量为g-1个，g为身体指标对应的风险等级的数量，其中第j个logistic回归模型为：

[0084]
$$\text{Logit}\left(\sum_{i=1}^j P_i\right) = \beta_{0j} + \sum_{n=1}^m \beta_n * X_n$$
，m为电力职业人群数据总量， β_n 为logistic回归模型中每个自变量 X_n 的回归系数， β_{0j} 为logistic回归模型的截距值， β_i 和 β_{0j} 通过最大似然法对logistic回归模型进行拟合得到， X_n 为具体的电力职业人群数据。

[0085] 风险计算模块进一步地用于：

[0085] 风险计算模块进一步地用于：

[0086] 利用下述公式对logistic回归模型进行求解：

[0087] 当k=1时，
$$P_k = \frac{1}{1 + \exp(f(k))} ;$$

[0088] 当g>k>1时，
$$P_k = \frac{1}{1 + \exp(f(k))} - \frac{1}{1 + \exp(f(k-1))}$$
，其中 $f(k) = -\text{Logit}\left(\sum_{i=1}^k p_i\right)$ ， P_k 为

身体指标的第k个风险等级对应的概率。

[0089] 电力职业人群重大慢性病风险评估系统还包括：优势比计算模块。

[0090] 优势比计算模块用于利用logistic回归模型的回归系数计算用于衡量危险因素

作用大小的优势比。

[0091] 本发明将电力职业人群数据录入已建立的logistic回归模型,logistic回归模型由电力职业人群数据所对应的数据类别、身体指标及其所包含的风险等级构建而成的多分类有序logistic回归模型,对录入电力职业人群数据的logistic回归模型进行求解,得到所述身体指标的每一风险等级所对应的概率,可以对电力职业人群的健康状态进行评估和提示,对电力职业人群的健康管理具有重大意义。

[0092] 以上内容是结合具体的优选实施方式对本发明所作的进一步详细说明,不能认定本发明的具体实施只局限于这些说明。对于本发明所属技术领域的普通技术人员来说,在不脱离本发明构思的前提下,还可以做出若干简单推演或替换,都应当视为属于本发明的保护范围。

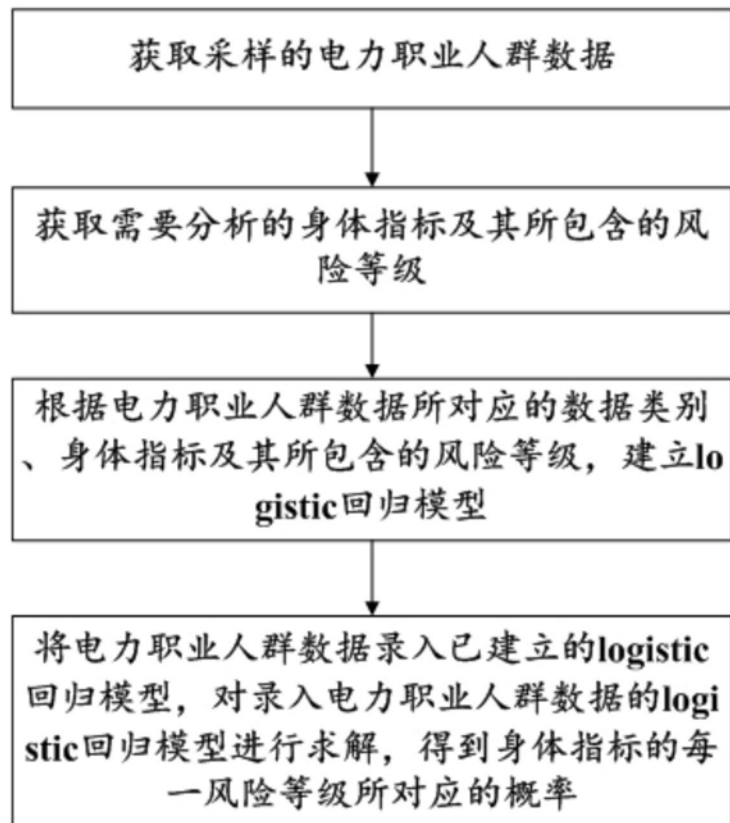


图1

专利名称(译)	一种电力职业人群重大慢性病风险评估方法及系统		
公开(公告)号	CN110189829A	公开(公告)日	2019-08-30
申请号	CN201910552296.0	申请日	2019-06-21
[标]申请(专利权)人(译)	深圳供电局有限公司		
申请(专利权)人(译)	深圳供电局有限公司		
当前申请(专利权)人(译)	深圳供电局有限公司		
[标]发明人	许克豪 何群 夏英华 曹蓉 刘国恒 黄苑 姜虹 梁鹏翼		
发明人	许克豪 何群 夏英华 曹蓉 刘国恒 洪紫慧 黄苑 姜虹 梁鹏翼		
IPC分类号	G16H50/50 G16H50/30 A61B5/00		
CPC分类号	A61B5/7275 A61B2503/20 G16H50/30 G16H50/50		
外部链接	Espacenet SIPO		

摘要(译)

本发明提供一种电力职业人群重大慢性病风险评估方法，其包括下述步骤：获取采样的电力职业人群数据，电力职业人群数据所对应的数据类别包括：个人基本信息数据、个人生活习惯数据、心理健康状况数据、职业相关因素数据、体检数据、人资数据；获取需要分析的身体指标及其所包含的风险等级；根据电力职业人群数据所对应的数据类别、身体指标及其所包含的风险等级，建立logistic回归模型；将电力职业人群数据录入已建立的logistic回归模型，对录入电力职业人群数据的logistic回归模型进行求解，得到身体指标的每一风险等级所对应的概率。本发明可以对电力职业人群的健康状态进行评估和提示，对电力职业人群的健康管理具有重大意义。

