

基于离散三维荧光光谱的糖尿病识别方法研究

张洋^{1,2}, 何腾超^{1,4}, 刘林³, 张健³, 张俊清³, 刘勇¹, 王贻坤^{1,4**}, 张元志^{1*}¹中国科学院合肥物质科学研究院安徽光学精密机械研究所, 安徽省生物医学光学仪器工程技术研究中心, 安徽省医用光学诊疗技术与装备工程实验室, 安徽 合肥 230031;²中国科学技术大学研究生院科学岛分院, 安徽 合肥 230026;³北京大学第一医院内分泌科, 北京 100034;⁴皖江新兴产业技术发展中心, 安徽 铜陵 244000

摘要 基于皮肤荧光光谱技术的糖基化终产物(AGE)检测方法被广泛用于糖尿病及其并发症的检测评估,然而生物组织体内特异性的吸收、散射特性以及多种荧光成分对于检测会产生干扰。本研究通过分析人体皮肤组织常见荧光成分的三维荧光光谱,确定皮肤组织荧光光谱测量的最佳激发波长组合,搭建集成组织漫反射光谱测量模块的离散三维荧光光谱测量系统,研究基于扩散理论的组织生理参数提取方法以及基于高斯多峰拟合的离散三维荧光光谱分离方法。在此基础上,开展临床社区队列研究,结果表明糖尿病组的黑色素、脱氧血红蛋白相对浓度均值低于正常对照组,而 500 nm 处约化散射系数高于正常对照组,差异有统计学意义,氧合血红蛋白相对浓度无明显差异。基于高斯多峰拟合,每位受试者共计得到 78 个荧光特征,剔除差异不明显的特征($p > 0.1$),最终得到 50 个荧光特征。汇总有差异的组织生理参数和荧光特征,采用 Logistic 回归分析建立糖尿病筛查模型,受试者工作特征(ROC)曲线分析结果显示,该方法用于糖尿病识别时,预测训练集 ROC 曲线下面积为 0.793,预测测试集的面积 0.799,而单波长皮肤荧光(Sf_{365})的面积为 0.731,结果表明相比单波长荧光,模型具有更好的诊断价值。

关键词 光谱学; 组织荧光光谱; 离散三维荧光光谱; 高斯多峰拟合; 生理参数

中图分类号 R318.51

文献标志码 A

doi: 10.3788/AOS202242.0117002

Diabetes Recognition Method Based on Discrete Three-Dimensional Fluorescence Spectrum

Zhang Yang^{1,2}, He Tengchao^{1,4}, Liu Lin³, Zhang Jian³, Zhang Junqing³,
Liu Yong¹, Wang Yikun^{1,4**}, Zhang Yuanzhi^{1*}¹Anhui Provincial Engineering Laboratory for Medical Optical Diagnosis & Treatment Technology and Instrument, Anhui Provincial Engineering Technology Research Center for Biomedical Optical Instrument, Anhui Institute of Optics and Fine Mechanics, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei, Anhui 230031, China;²Science Island Branch of Graduate School, University of Science and Technology of China, Hefei, Anhui 230026, China;³Endocrinology Department, Peking University First Hospital, Beijing 100034, China;⁴Wan Jiang New Industry Technology Development Center, Tongling, Anhui 244000, China

Abstract The advanced glycation end (AGE) product detection based on skin fluorescence spectroscopy is widely

收稿日期: 2021-05-21; 修回日期: 2021-06-24; 录用日期: 2021-07-15

基金项目: 安徽省重点研究与开发计划(202004a07020016)、安徽省自然科学基金(1908085QH365)、中国科学院合肥物质科学研究院院长基金(YZJJZX202009)、中国科学院科技服务网络计划(KFJ-STQZYD-184)

通信作者: *yzzhang@aiofm.ac.cn; **wyk@aiofm.ac.cn

used in the screening and evaluation of diabetes and its complications. However, this method is limited by the specific absorption and scattering characteristics in biological tissues and a variety of fluorescent components. In this study, the optimal excitation wavelength combination for measuring skin tissue fluorescence spectra was determined by analyzing the three-dimensional fluorescence spectra of common fluorescent components in human skin tissues. A discrete three-dimensional fluorescence spectrum measurement system was built that integrated the measurement module of tissue diffuse reflectance spectra. The methods of extracting tissue physiological parameters under the diffusion theory and separating discrete three-dimensional fluorescence spectra by multi-peak Gaussian fitting were studied. A clinical community cohort study was then carried out. The results showed that the mean relative concentrations of melanin and deoxyhemoglobin in the diabetic group were lower than those in the normal control group, while the reduced scattering coefficient at 500 nm was higher than that in the normal control group. These differences were statistically significant. No significant difference was found in the relative concentration of oxyhemoglobin. Under the multi-peak Gaussian fitting, 78 fluorescence features were obtained for each subject, and 50 fluorescence features remained after features with insignificant differences ($p > 0.1$) were eliminated. Differential tissue physiological parameters and fluorescence characteristics were pooled, and a diabetes screening model was built by the logistic regression analysis. According to the analysis results of the receiver operating characteristic (ROC) curve, when the proposed method was applied to diabetes screening, the area under the ROC curve of the prediction training set was 0.793, and that of the prediction test set was 0.799. In contrast, that of the single-wavelength skin fluorescence (Sf_{365}) was 0.731, which indicated that the proposed model had a better diagnostic value than that of the single-wavelength fluorescence.

Key words spectroscopy; tissue fluorescence spectrum; discrete three-dimensional fluorescence spectrum; multi-peak Gaussian fitting; physiological parameter

1 引 言

近年来随着人们生活水平的不断提高,生活方式、饮食习惯的不断改变,以及人口老龄化进程的加速,我国 2 型糖尿病 (T2DM) 的发病率显著升高。流行病学调查显示我国 18 岁及以上人群 T2DM 患病率已达 11.2%^[1]。针对高危人群进行 T2DM 筛查,控制 T2DM 的危险因素,有助于早期发现并预防 T2DM 的发生。糖基化终产物 (AGE) 是指在非酶促反应条件下,蛋白质、氨基酸、脂类或核酸等大分子物质的游离氨基与还原糖(葡萄糖、果糖等)的醛基经过缩合、重排、裂解、氧化修饰等一系列化学反应形成的一组稳定的终末产物^[2]。研究表明体内过量蓄积的 AGE 可以通过直接修饰胰岛素或与受体结合激活氧化应激/炎症信号通路的方式导致胰岛素抵抗^[3-4],AGE 还可以通过抑制 β 细胞胰岛素分泌、诱导 β 细胞凋亡等途径促进 β 细胞功能缺陷^[5],在胰岛素抵抗的基础上,导致 2 型糖尿病的发生。临床队列研究表明人体 AGE 蓄积量越高,糖尿病患病风险越大。因此,AGE 是糖尿病的早期致病因子,有助于早期识别糖尿病高风险人群^[6]。

传统的 AGE 检测方法,主要包括高效液相-质谱、酶联免疫吸附法 (ELISA)、放射免疫法等,这些方法均需要抽血或者皮肤组织活体采样,有创伤,而且过程繁琐,耗时很长,限制了 AGE 指标在临床上广泛应用^[7]。AGE 具有荧光特性,通常采用的激发

波长为 370 nm/发射波长 440 nm^[8]。Smit 等^[9]采用近紫外光 350~420 nm(中心波长 370 nm)照射人体前臂内侧皮肤,根据测量的皮肤荧光(420~600 nm)与激发波段反射光的比值来校正吸收、散射的影响,表征 AGE 的积累量,结果表明皮肤荧光法检测与 AGE 活检结果具有较好的相关性。Cleary 等^[10]采用中心波长为 375 nm 的激发光,测量皮肤荧光(435~655 nm)与激发光反射光以及漫反射光(435~655 nm)乘积的比值来表征 AGE 的蓄积量。

研究表明^[11-13]皮肤主要的荧光团有 280/340、295/345、335/390、370/460、440/530 nm,其中 280/340 nm、295/345 nm 的荧光主要来自酪氨酸和色氨酸基团。335/390 nm、370/460 nm 的荧光则主要来自 AGE、还原型辅酶 I (NADH)、胶原以及极少量的角蛋白。440/530 nm 的荧光主要来自弹性蛋白、黄素腺嘌呤二核苷酸 (FAD) 和卟啉。目前研究中大多采用 370 nm 照射皮肤组织,选择 420~655 nm 的宽波段荧光进行积分,代表皮肤 AGE 的荧光值,用于糖尿病筛查,损失了其他皮肤荧光光谱信息。因此,同时引入基于漫反射谱的组织生理参数提取以及离散三维荧光光谱纳入糖尿病筛查模型,预期可提高皮肤荧光的准确性。本文通过研究皮肤组织中的 AGE 以及其他荧光物质 NADH、FAD、胶原等的荧光特性,确定离散三维荧光光谱测量的激发波长组合,搭建多光谱检测系统,并开展

临床社区队列研究,收集正常人群和糖尿病人群的光谱数据,在训练集中采用 Logistic 回归分析方法,把具有统计学差异的组织生理参数和基于高斯多峰拟合的荧光光谱特征纳入糖尿病筛查模型,采用受试者工作特征(ROC)曲线评价基于离散三维荧光光谱的糖尿病识别方法的有效性。

2 原理与方法

2.1 离散三维荧光光谱测量系统

2.1.1 离散三维荧光光谱激发波长选择

采用去离子水溶解 AGE 标准品(上海玉博生物科技有限公司)、NADH (sigma, 美国)、FAD (sigma, 美国)以及胶原(sigma, 美国)。利用日立

F7000 光谱仪对 4 个样本进行三维荧光光谱扫描。去除三维荧光光谱中的瑞利和拉曼散射峰后,进行艾米特插值,得到如图 1 所示的三维荧光光谱。其中,AGE 三维荧光光谱特征显示,其激发波长覆盖 320~410 nm,发射波长覆盖 350~540 nm,最佳激发/发射波长 345/440 nm,范围与已知报道的一致。NADH 激发发射波长覆盖 240~390 nm,发射波长覆盖 400~600 nm,存在 260/450 nm 和 340/450 nm 两个激发/发射峰。FAD 激发波长覆盖 350~460 nm,发射波长覆盖 460~620 nm,存在 370/530 nm 和 450/530 nm 两个激发/发射峰。胶原激发波长覆盖 290~350 nm,发射波长覆盖 300~460 nm,最佳激发/发射波长 320/400 nm。

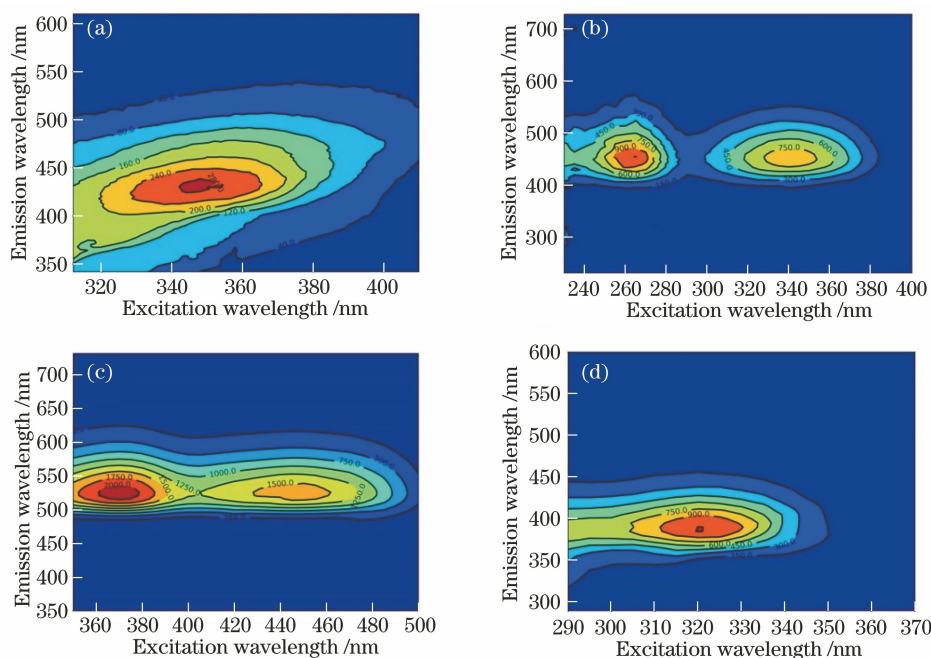


图 1 不同荧光成分的三维荧光光谱特征。(a) AGE;(b) NADH;(c) FAD;(d)胶原

Fig. 1 Three dimensional fluorescence spectra characteristics of different fluorescent components.

(a) AGE; (b) NADH; (c) FAD; (d) collagen

为实现以最优的结果代替连续三维荧光光谱测量^[14],本文遵循如下原则对最佳离散三维荧光光谱的激发波长组合进行选择:1)确保所选特征波段独立,避免共线性问题;2)最大程度反映原三维荧光光谱谱型,保留特征信息;3)最小化非特征区域和特征区域比例,突出组分差异。除此之外,考虑当前单色发光二极管技术水平,选择最佳组合包含以下波长:340、365、395、415、430、455 nm。

2.1.2 离散三维荧光光谱测量系统结构

为实现组织漫反射率以及荧光光谱同步测量,搭建了如图 2 所示测量系统,包括激发光源、宽带光源、传输光纤、光谱测量模块、控制模块等 5 个部分。

系统软件通过动态调节积分时间分次检测合适的皮肤荧光光谱和漫反射光谱。激发光源采用 6 个单色发光二极管(LED 1:中心波长 340 nm,半峰全宽 11 nm;LED 2:中心波长 365 nm,半峰全宽 9 nm;LED 3:中心波长 395 nm,半峰全宽 11 nm;LED 4:中心波长 415 nm,半峰全宽 14 nm;LED 5:中心波长 430 nm,半峰全宽 15 nm;LED 6:中心波长 455 nm,半峰全宽 18 nm);LED 1、LED 3、LED 5 和 LED 2,LED 4,LED 6 分别经过准直透镜、二向色镜和聚焦透镜后耦合进传输光纤。宽带光源采用大功率白光 LED,经球透镜、平凸透镜耦合进传输光纤。光谱探测模块采用紫外可见微型光纤光谱仪。

传输光纤为 1×4 的光纤束, 共计 7 根光纤, 2 根连接 LED 1、LED 3、LED5, 2 根连接 LED 2、LED 4、LED 6, 2 根连接白光 LED, 剩余 1 根连接光纤光谱

仪, 光纤束端面连接光纤光谱仪的光纤居中, 其余 6 根间隔排列。

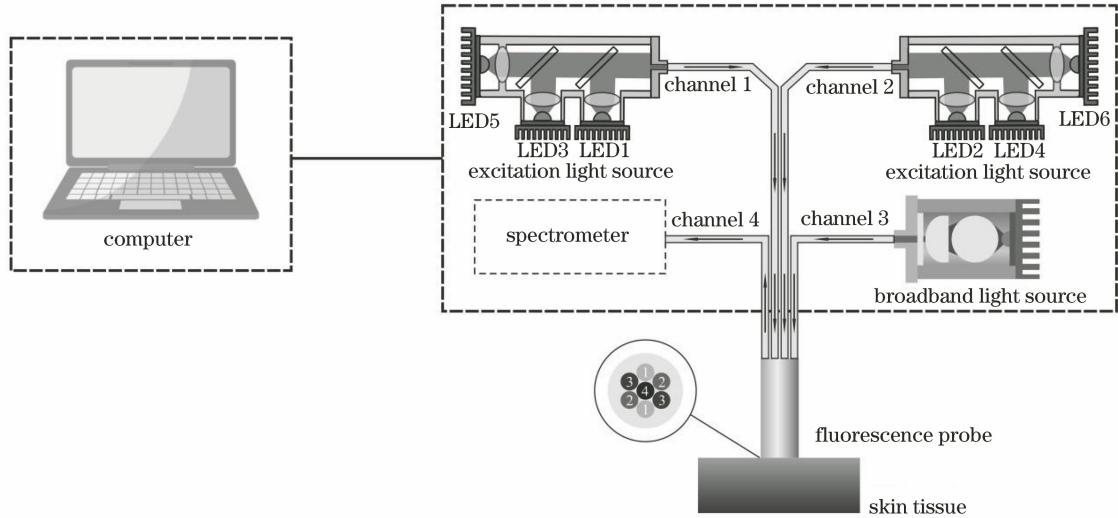


图 2 离散三维荧光光谱测量系统结构示意图

Fig. 2 Schematic of discrete three dimensional fluorescence spectrum measurement system

2.2 组织生理参数、荧光光谱特征提取方法

可见及近红外波段范围内, 组织吸收来源于黑色素和血红蛋白, 根据朗伯比尔定律, 组织吸收系数可表述为 (1) 式; 生物组织约化散射系数呈指数分布, 可由 (2) 式表示^[15-17]。

$$\mu_a(\lambda) = c_{\text{melanin}} \epsilon_{\text{melanin}}(\lambda) + c_{\text{HbO}_2} \epsilon_{\text{HbO}_2}(\lambda) + c_{\text{Hb}} \epsilon_{\text{Hb}}(\lambda), \quad (1)$$

$$u'_s(\lambda) = u'_s(500 \text{ nm}) \cdot (\lambda/500 \text{ nm})^{-b}, \quad (2)$$

式中, μ_a 为吸收系数, c_{melanin} 、 c_{HbO_2} 、 c_{Hb} 分别为黑色素、氧合血红蛋白、脱氧血红蛋白的相对浓度, $\epsilon_{\text{melanin}}$ 、 ϵ_{HbO_2} 、 ϵ_{Hb} 分别为黑色素、氧合血红蛋白、脱氧血红蛋白的摩尔消光系数, u'_s 为约化散射系数, b 为散射指数。

已知组织吸收系数和约化散射系数时, 根据半无限媒质、外推边界条件下扩散方程的近似解, 可计算从距光源径向距离 r_t 处出射的漫反射光强

$$R_{\text{sim}}(r_t, \lambda) = \frac{a'}{4\pi} \left[z_0 \left(u_{\text{eff}} + \frac{1}{r_1} \right) \frac{\exp(-\mu_{\text{eff}} r_1)}{r_1^2} + (z_0 + 2z_b) \left(u_{\text{eff}} + \frac{1}{r_2} \right) \frac{\exp(-\mu_{\text{eff}} r_2)}{r_2^2} \right], \quad (3)$$

式中, $z_0 = 1/\mu'_s$, $z_b = 2/[3(\mu_a + \mu'_s)]$, $r_1 = \sqrt{z_0^2 + r_t^2}$, $r_2 = \sqrt{(z_0 + 2z_b)^2 + r_t^2}$ 。 $a' = \mu'_s/(\mu_a + \mu'_s)$, 表示组织反照率; $\mu_{\text{eff}} = \sqrt{3\mu_a(\mu_a + \mu'_s)}$, 表示有效衰减系数。对于特定 r_t , 在测量得到组织漫反射

率 $R_{\text{meas}}(\lambda)$ 后, 结合 (1)~(3) 式, 通过最小二乘法最小化 $\sum_{i=0}^n [R_{\text{meas}}(\lambda_i) - R_{\text{sim}}(\lambda_i)]^2$, 即可得到组织生理参数 c_{melanin} 、 c_{HbO_2} 、 c_{Hb} 、 u'_s 、 b 的最优解^[18]。

对离散三维荧光光谱中每条荧光光谱进行高斯多峰拟合^[19-20]。光谱曲线可描述为

$$I_{F_{\text{sim}}} = \sum_{i=1}^n \sum_{j=1}^{m_i} \frac{A_{i,j}}{\sqrt{2\pi\sigma_{i,j}}} \exp \left[-\frac{(x - \omega_{i,j})^2}{2\sigma_{i,j}^2} \right], \quad (4)$$

式中, n 为离散三维荧光光谱激发波长数, m_i 为每组荧光光谱拟合的高斯曲线数, $A_{i,j}$ 为高斯曲线幅值, 表示光谱曲线下覆盖面积, $\omega_{i,j}$ 为高斯曲线中心位置, 表示光谱曲线中心波长, $\sigma_{i,j}$ 为高斯曲线方差, 表示光谱曲线半峰全宽。逐激发波长进行高斯拟合, 即可得到 $A_{i,j}$ 、 $\omega_{i,j}$ 、 $\sigma_{i,j}$ 。

2.3 研究对象与方法

选取来自“八五期间”以首钢集团附属多个社区为主体开展的慢病随访监测项目人群(首钢地区队列人群, 该人群为社区自然人群), 于 2018 年 9 月~12 月参与随访的人群, 排除标准: 1) 交流障碍, 有严重失语、视听障碍患者等妨碍评估进行的躯体功能缺陷; 2) 有严重精神疾病, 难以配合调查者; 3) 行动不便者; 4) 孕妇或其他研究者判断不适合参与本次调查者。检测部位为受试者前臂内侧皮肤, 需避开瘢痕、苔藓样硬化和畸形皮肤, 避免涂抹护肤品及防晒霜。采集研究对象的人口学特征(年龄、性别等)、

疾病史、体格检查(体质指数 BMI 等)、血液生化[空腹血糖 (FPG, F_{PG})、总胆固醇 (TC)、甘油三酯 (TG)、尿素氮(BUN)、肌酐(CREA)等]及多光谱数据检测结果。

采用 SPSS 19.0 软件进行统计学分析。计量资料以($\bar{x} \pm s$)表示,其中 \bar{x} 表示平均值, s 表示标准差。其中计数资料采用卡方检验,组间均值比较采用 t 检验, $p < 0.05$ 表明差异具有统计学意义, p 值越小,差异越显著。根据中国 2 型糖尿病防治指南(2020 年版),以空腹血糖和口服葡萄糖耐量试验(OGTT)2h 血糖(2h PG, $F_{2h PG}$)为诊断标准,糖尿病的诊断标准为 $F_{PG} \geq 7.0$ mmol/L 或 $F_{2h PG} \geq 11.1$ mmol/L;血糖正常人群的诊断标准为 $F_{PG} < 6.1$ mmol/L 且 $F_{2h PG} < 7.8$ mmol/L。以 ROC 曲线来评价建立的模型用于诊断识别糖尿病的价值。ROC 曲线是以真阳性率(Sensitivity,灵敏度,即预测为正且实际为正的样本占有所有正例样本的比例)为纵坐标,假阳性率(1-特异性,1-Specificity,预测为正但实际为负的样本占有所有负例样本的比例)为横坐标绘制的曲线,通过计算 ROC 曲线下的面积进行两种方法的比较,曲线下面积越大,诊断价值越高。

3 试验结果及分析

本试验研究共纳入受试者 1897 例,男性 703 例,女性 1194 例,平均年龄(61.68 ± 7.3)岁,平均 BMI 为(26.1 ± 16.7) kg/m^2 。其中正常对照组受试者 1216 例和糖尿病患者 681 例。一般资料见表 1,结果显示糖尿病人群的年龄、FPG、OGTT 2h PG、TG、BUN、CREA 均高于正常对照组,男性糖尿病的发病率也显著高于女性,差异有统计学意义。

采用上述离散三维荧光光谱测量系统同步测量

表 1 人群一般特征

Table 1 General characteristics of the population

Group	Nomal $n=1216$	T2DM $n=681$	p
Gender (male/female)	386/830	317/364	<0.001
Age /year	59.9 ± 6.5	64.9 ± 7.5	<0.001
BMI /(kg/m^2)	25.6 ± 16.5	27.0 ± 17.0	0.072
FPG /(mmol/L)	5.2 ± 0.4	8.4 ± 2.6	<0.001
2h PG /(mmol/L)	6.1 ± 1.0	15.3 ± 4.0	<0.001
TC /(mmol/L)	5.4 ± 1.0	5.2 ± 1.0	<0.001
TG /(mmol/L)	1.5 ± 1.3	1.8 ± 1.3	<0.001
BUN /(mmol/L)	5.4 ± 1.3	5.8 ± 1.4	<0.001
CREA /($\mu\text{mol}/\text{L}$)	62.7 ± 13.0	65.5 ± 16.4	<0.001

不同激发波长照射时受试者前臂内侧组织荧光光谱、漫反射光谱,以及标准漫反射白板(可见波段漫反射率约为 99%)反射谱。计算组织漫反射光谱和白板漫反射光谱比率,得到组织漫反射率。根据(1)~(3)式,最小化实测组织漫反射率与漫射近似模型输出的漫反射率的均方误差,得到受试者组织生理参数进行提取。图 3 为 3 位糖尿病患者与 3 位正常对照人群实测漫反射率。结果表明,在可见波段范围内不同受试者组织漫反射率曲线形状与强度均不同;在合适的生理参数组合下,漫射近似模型输出的漫反射率可以与实测漫反射率吻合。

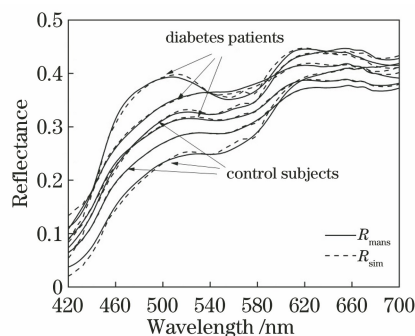


图 3 实测组织漫反射率(R_{meas})与漫射近似模型生成的漫反射率(R_{sim})

Fig. 3 Measured tissue diffuse reflectance (R_{meas}) and diffuse reflectance generated by diffuse approximation model (R_{sim})

提取受试者前臂内侧组织黑色素(melanin)、氧合血红蛋白(HbO_2)、脱氧血红蛋白(Hb)相对浓度以及 500 nm 处组织约化散射系数(μ_s)和散射指数(b)后,对上述指标进行统计分析,具体结果如图 4 所示。糖尿病组的组织黑色素、脱氧血红蛋白浓度均值低于正常对照组, t 检验结果显示两组人群黑色素浓度存在明显差异($p < 0.05$),脱氧血红蛋白浓度存在显著差异($p < 0.005$)。两组人群氧合血红蛋白则无明显差异($p > 0.05$)。临床研究表明糖尿病患者容易发生微循环功能障碍,而微循环是血液与组织间实现物质交换的主要场所和功能单位^[21]。糖尿病微循环障碍伴随着持续的血流动力学异常,最终会导致微循环缺血、缺氧,本研究虽然未观测到氧合血红蛋白出现显著降低,但是脱氧血红蛋白是降低的,有统计学差异,两者总和也是下降的,与微循环障碍缺血现象一致。此外,有研究表明 2 型糖尿病患者的血清 α -黑色素细胞刺激素水平相比正常人群明显降低,该激素能促进存在于皮肤的黑色素细胞产生黑色素^[22],本研究结果显示 2 型糖

尿病人群的黑色素水平降低,可能是由于病理状态下该激素水平的下降导致的结果。此外,糖尿病组前臂内侧组织 500 nm 处的约化散射系数高于正常对照组, t 检验结果显示二者存在明显差异 ($p <$

0.05)。糖尿病组的组织散射指数相比正常对照组有升高的趋势 ($p = 0.07$)。因此,把组织黑色素、脱氧血红蛋白相对浓度以及 500 nm 处组织约化散射系数和散射指数 (b) 纳入糖尿病筛查模型。

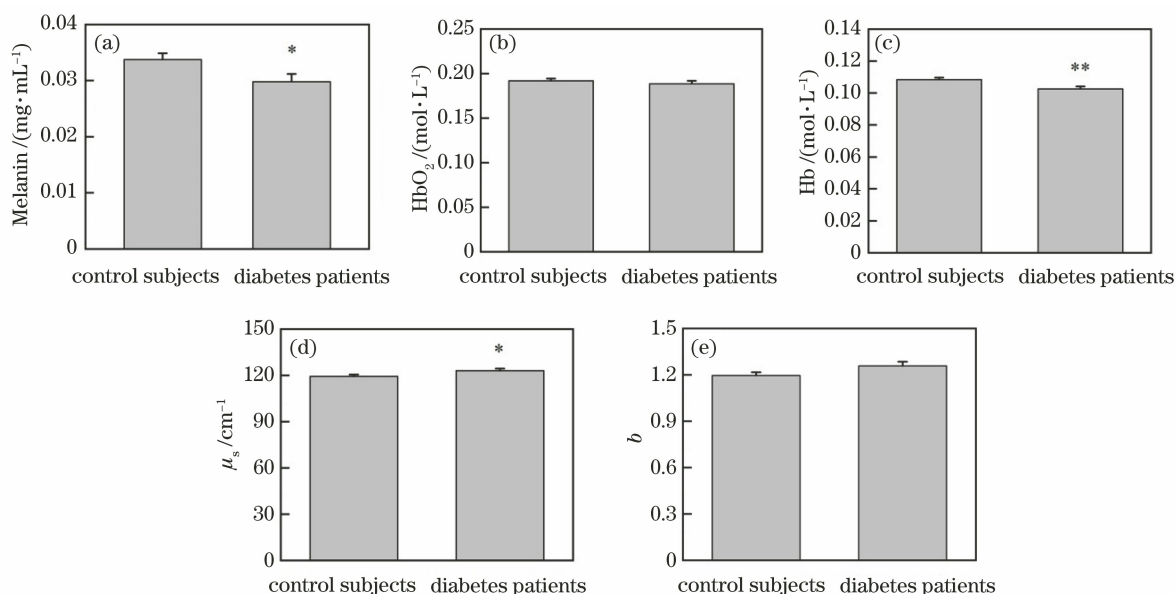


图 4 糖尿病组与正常对照组的受试者前臂内侧组织生理参数均值比较。(a)黑色素(melanin)相对浓度;(b)氧合血红蛋白(HbO₂)相对浓度;(c)脱氧血红蛋白(Hb)相对浓度;(d)500 nm 处组织约化散射系数(μ_s);(e)组织约化散射指数 b

Fig. 4 Comparison of physiological parameters of volar side forearm between diabetic group and normal control group.

(a) Relative concentration of melanin; (b) relative concentration of oxyhemoglobin (HbO₂); (c) relative concentration of deoxyhemoglobin (Hb); (d) tissue reduced scattering coefficient at 500 nm (μ_s); (e) tissue reduced scattering exponent b

对受试者前臂内侧组织离散三维荧光光谱中每条荧光发射谱进行高斯拟合。基于离散三维荧光光谱测量系统结构,(4)式中 n 表示离散三维荧光光谱激发波长数,取值为 6; m_i 为每组荧光光谱拟合的高斯曲线数。测试发现,为保证荧光光谱高斯拟合效果,激发波长取 340、365、395、415、430、455 nm 时,对应 m 取值为 5、5、4、4、4、4。即把离散三维荧光光谱分解为 26 条高斯曲线,每条高斯光谱曲线具有强度、中心波长位置、方差 3 个参数,共计得到 78 个荧光特征。图 5 为受试者前臂内侧组织离散三维荧光光谱高斯拟合结果,图 5(a)~(f)分别对应激发波长 365、340、395、415、430、455 nm。结果显示,实测荧光光谱与拟合得到的高斯曲线叠加状态基本重合,每条荧光光谱的曲线下面积等于被分解的数条高斯光谱曲线下面积之和,离散三维荧光光谱得到较好分解。

为提高识别效率,降低没有差异的光谱信息,同时尽可能地保留更多光谱特征,本研究对糖尿病组和正常对照组受检者的荧光特征进行均值 t 检验,

选择剔除其中均值差异显著不明显的特征 ($p > 0.1$),最终得到 50 个差异荧光特征,本次研究剔除的光谱信息需要通过更大的样本人群进行进一步的评估筛选。将得到的 50 个差异荧光特征与组织黑色素、脱氧血红蛋白相对浓度以及 500 nm 处组织约化散射系数和散射指数一同纳入糖尿病筛查模型。把 1897 例受试者随机划分为 2 组人群,其中 70% 作为训练集,30% 作为测试集。在训练集中进行 Logistic 回归建模,建模时采用 L2 正则化降低过拟合风险,采用五折交叉验证优化惩罚参数。

构建 Logistic 回归模型后,对测试集中的受试者数据进行预测,得到受试者划分为糖尿病人群的概率,将其与受试者实际患病情况进行 ROC 曲线分析。此外,根据 Cleary 课题组^[10]的方法,采用皮肤原始荧光与激发光反射光以及漫反射光乘积的比值,计算中心波长 365 nm 激发波段的皮肤荧光值 Sf_{365} ,并在总人群中进行 ROC 曲线分析。ROC 曲线结果如图 6 所示,模型预测训练集受试者 ROC 曲线覆盖面积为 0.793,预测测试集受试者 ROC 曲

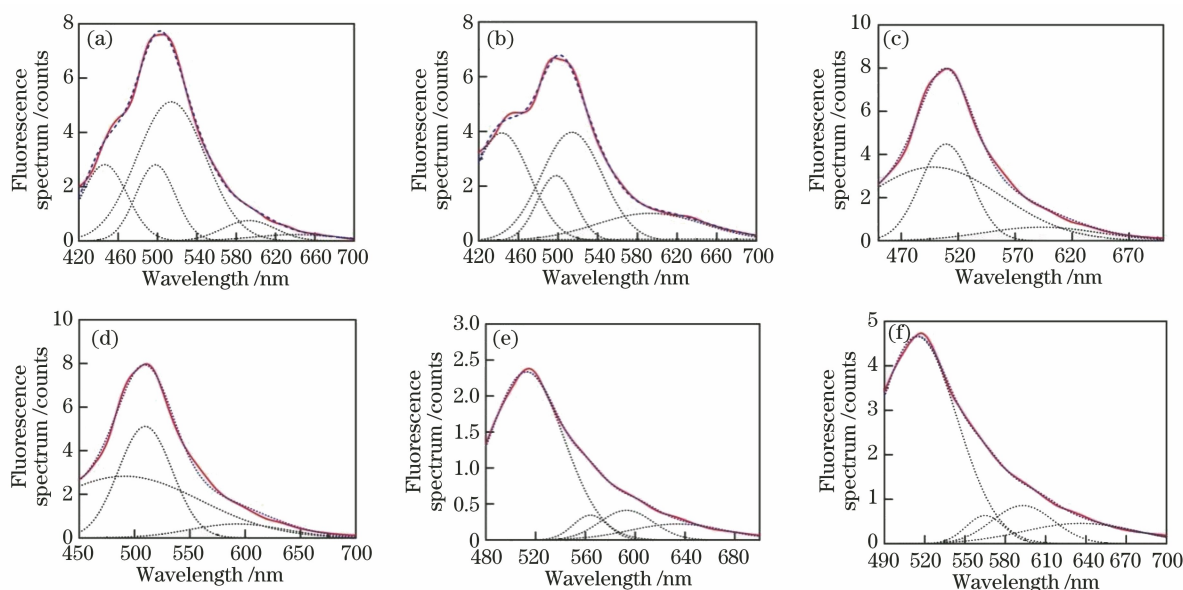


图 5 组织离散三维荧光光谱的高斯拟合结果。(a) 365 nm, (b) 340 nm, (c) 395 nm, (d) 415 nm, (e) 430 nm, (f) 455 nm

Fig. 5 Gaussian fitting results of tissue discrete three-dimensional fluorescence spectrum.

(a) 365 nm, (b) 340 nm, (c) 395 nm, (d) 415 nm, (e) 430 nm, (f) 455 nm

线覆盖面积为 0.799, Sf_{365} 的曲线下面积为 0.731, 结果表明相比单波长荧光, 模型具有更好的诊断价值。

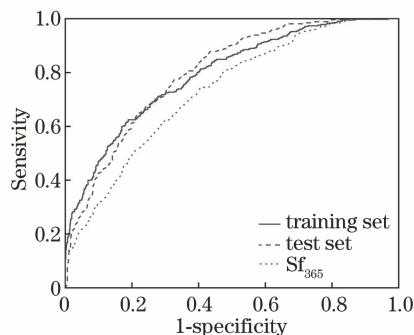


图 6 测试集和训练集模型以及皮肤荧光(Sf_{365})的糖尿病识别 ROC 曲线

Fig. 6 ROC curve of test set, training set model and skin fluorescence (Sf_{365}) for diabetes recognition

4 结 论

本研究通过分析生物组织中常见荧光成分 (AGE、NADH、FAD 以及胶原) 的三维荧光光谱特征, 确定 340、365、395、415、430、455 nm 作为最优激发波长组合, 尽可能多地保留了组织荧光光谱信息, 同时也降低了测量系统的复杂度。此外, 为降低生物组织特异性吸收、散射特性的干扰, 提高测量系统的准确性和适用范围, 设计、搭建并集成有组织漫反射光谱测量功能的离散三维荧光光谱测量系统, 研究了组织生理参数以及组织荧光光谱特征提取方

法。利用搭建的离散三维荧光光谱测量系统, 开展临床社区队列研究, 共纳入受试者 1897 例, 其中正常对照组受试者 1216 例和糖尿病组患者 681 例。临床研究结果显示糖尿病组的黑色素、脱氧血红蛋白相对浓度低于正常对照组, 差异有统计学意义。两组人群氧合血红蛋白则无明显差异。此外, 糖尿病组前臂内侧组织 500 nm 处的约化散射系数高于正常对照组, 组织散射指数相比正常对照组有升高的趋势 ($p=0.07$)。结果提示糖尿病人群与正常人群的生理参数存在差异, 可能导致了个体组织吸收、散射的差异化。因此, 需要把有差异的组织生理参数纳入糖尿病筛查模型。对受试者前臂内侧组织离散三维荧光光谱中每条荧光发射谱进行高斯拟合, 每个人共计得到包含高斯光谱曲线强度、中心波长位置、方差等信息的 78 个荧光特征, 剔除两组间差异不明显的特征 ($p>0.1$), 最终得到 50 个荧光特征。汇总生理参数后采用 Logistic 回归建模, 结果显示模型预测训练集受试者 ROC 曲线覆盖面积为 0.793, 测试集受试者 ROC 曲线覆盖面积为 0.799, Sf_{365} 的曲线下面积为 0.731, 临床研究结果表明基于离散三维荧光光谱的糖尿病识别方法具有更好的诊断价值。

参 考 文 献

- [1] Chinese Diabetes Society. Guideline for the prevention and treatment of type 2 diabetes mellitus

- in China (2020 edition) [J]. Chinese Journal of Diabetes, 2021, 13(4): 315-409.
- 中华医学会糖尿病学分会. 中国 2 型糖尿病防治指南(2020 年版)[J]. 中国糖尿病杂志, 2021, 13(4): 315-409.
- [2] Vlassara H, Uribarri J. Advanced glycation end products (AGE) and diabetes: cause, effect, or both? [J]. Current Diabetes Reports, 2013, 14(1): 1-10.
- [3] Cai W, Ramdas M, Zhu L, et al. Oral advanced glycation endproducts (AGEs) promote insulin resistance and diabetes by depleting the antioxidant defenses AGE receptor-1 and sirtuin 1 [J]. Proceedings of the National Academy of Sciences of the United States of America, 2012, 109 (39): 15888-15893.
- [4] Nowotny K, Jung T, Höhn A, et al. Advanced glycation end products and oxidative stress in type 2 diabetes mellitus [J]. Biomolecules, 2015, 5 (1): 194-222.
- [5] Coughlan M T, Yap F Y, Tong D C, et al. Advanced glycation end products are direct modulators of β -cell function [J]. Diabetes, 2011, 60 (10): 2523-2532.
- [6] Waateringe R P, Fokkens B T, Slagter S N, et al. Skin autofluorescence predicts incident type 2 diabetes, cardiovascular disease and mortality in the general population [J]. Diabetologia, 2019, 62(2): 269-280.
- [7] Ye C S, Wang Y K, Ye S D. Property and detection of advanced glycation end products [J]. International Journal of Pathology and Clinical Medicine, 2013, 33 (2): 175-178.
- 叶成松, 王贻坤, 叶山东. 晚期糖基化终末产物的特性及其检测 [J]. 国际病理科学与临床杂志, 2013, 33(2): 175-178.
- [8] Zhang Y, Ni J S, Zhang Y Z, et al. Tissue intrinsic fluorescence spectrum recovery algorithm and its application in diabetes screening [J]. Chinese Journal of Lasers, 2018, 45(7): 0707001.
- 张洋, 倪敬书, 张元志, 等. 组织固有荧光光谱复原算法及其在糖尿病筛查中的应用研究 [J]. 中国激光, 2018, 45(7): 0707001.
- [9] Meerwaldt R, Graaff R, Oomen P H N, et al. Simple non-invasive assessment of advanced glycation endproduct accumulation [J]. Diabetologia, 2004, 47 (7): 1324-1330.
- [10] Cleary P A, Braffett B H, Orchard T, et al. Clinical and technical factors associated with skin intrinsic fluorescence in subjects with type 1 diabetes from the diabetes control and complications trial/epidemiology of diabetes interventions and complications study [J]. Diabetes Technology & Therapeutics, 2013, 15(6): 466-474.
- [11] Franco W, Gutierrez-Herrera E, Kollias N, et al. Review of applications of fluorescence excitation spectroscopy to dermatology [J]. The British Journal of Dermatology, 2016, 174(3): 499-504.
- [12] Gillies R, Zonios G, Anderson R R, et al. Fluorescence excitation spectroscopy provides information about human skin *in vivo* [J]. Journal of Investigative Dermatology, 2000, 115(4): 704-707.
- [13] Beisswenger P J, Howell S, MacKenzie T, et al. Two fluorescent wavelengths, 440_{ex}/520_{em} nm and 370_{ex}/440_{em} nm, reflect advanced glycation and oxidation end products in human skin without diabetes [J]. Diabetes Technology & Therapeutics, 2012, 14(3): 285-292.
- [14] Zuo Z L, Zhao N J, Meng D S, et al. Identification of petroleum organic matter in soil based on three-dimensional fluorescence spectroscopy [J]. Laser & Optoelectronics Progress, 2019, 56(22): 222601.
- 左兆陆, 赵南京, 孟德硕, 等. 基于三维荧光光谱的土壤中石油类有机物分类识别 [J]. 激光与光电子学进展, 2019, 56(22): 222601.
- [15] Zhang Y Z, Hou H Y, Zhang Y, et al. Tissue intrinsic fluorescence recovering by an empirical approach based on the PSO algorithm and its application in type 2 diabetes screening [J]. Biomedical Optics Express, 2018, 9(4): 1795-1808.
- [16] Hennessy R, Lim S L, Markey M K, et al. Monte Carlo lookup table-based inverse model for extracting optical properties from tissue-simulating phantoms using diffuse reflectance spectroscopy [J]. Journal of Biomedical Optics, 2013, 18(3): 037003.
- [17] Feng W, Shi R, Zhang C, et al. Lookup-table-based inverse model for mapping oxygen concentration of cutaneous microvessels using hyperspectral imaging [J]. Optics Express, 2017, 25(4): 3481-3495.
- [18] Liu Y, Zhang Y Z, Hou H Y, et al. Tissue intrinsic fluorescence spectrum recovering based on diffusion theory [J]. Spectroscopy and Spectral Analysis, 2016, 36(12): 3836-3841.
- 刘勇, 张元志, 侯华毅, 等. 基于扩散理论的生物组织固有荧光光谱复原方法研究 [J]. 光谱学与光谱分析, 2016, 36(12): 3836-3841.
- [19] Hu Y, Fu Y, Li X Y, et al. Fluorescence spectra and fluorescence saturation intensity analysis of hepatic cell, hepatoma carcinoma cell and hepatic fibrosis cell [J]. Spectroscopy and Spectral Analysis, 2020, 40 (2): 478-482.
- 胡悦, 付芸, 李欣阳, 等. 肝、肝癌及肝纤维细胞的

- 荧光光谱及其荧光饱和强度分析[J]. 光谱学与光谱分析, 2020, 40(2): 478-482.
- [20] Li T T, Yang F, Li S G, et al. Extraction method of line-structured light stripe center based on Gauss-Lorenz decomposition peak fitting [J]. Laser & Optoelectronics Progress, 2019, 56(7): 071201.
李涛涛, 杨峰, 李石庚, 等. 基于高斯-洛伦兹分峰拟合的线结构光条中心的提取方法[J]. 激光与光电子学进展, 2019, 56(7): 071201.
- [21] Chinese Association of Diabetes and Microcirculation. Chinese experts consensus for drug therapy of microcirculatory dysfunction in diabetes mellitus: 2021 updated[J]. Chinese Journal of the Frontiers of Medical Science (Electronic Version), 2021, 13(4): 49-57.
中国微循环学会糖尿病与微循环专业委员会. 糖尿病微循环障碍临床用药专家共识(2021年版)[J]. 中国医学前沿杂志(电子版), 2021, 13(4): 49-57.
- [22] Li B X, Zhang Y L, Zhang X, et al. Changes of serum leptin and α -MSH levels in type 2 diabetes patients complicating with different grades of hypertension[J]. Chinese Circulation Journal, 2018, 33(9): 851-855.
李宝新, 张云良, 张鑫, 等. 2 型糖尿病合并不同分级高血压患者血清瘦素和 α -黑色素细胞刺激素水平的变化[J]. 中国循环杂志, 2018, 33(9): 851-855.