

doi: 10.3788/gzxb20124106.0684

血红蛋白对胆红素光学检测的影响分析

蒋景英^a, 龚启亮^a, 徐可欣^b

(天津大学 精密仪器与光电子工程学院 a. 生物医学检测技术与仪器天津市重点实验室;

b. 精密测试技术及仪器国家重点实验室, 天津 300072)

摘要:针对实验得到的光谱数据,应用偏最小二乘回归方法进行多变量建模,结果发现:血红蛋白的加入使建立模型的主成分数从一个变为三个,而对模型贡献最大的波长从 459 nm 变为 432 nm 和 433 nm,说明偏最小二乘模型的原始光谱数据变量引进了血红蛋白的吸收噪音.结合偏最小二乘建模的原理与评价指标分析表明:较纯胆红素溶液而言,混合溶液模型预测均方根误差增大了 24.90%,相关系数减小了 0.29%,加入血红蛋白后的模型对胆红素浓度预测能力降低.进一步分析血红蛋白和胆红素及其混合溶液吸收光谱发现:在检测的紫外可见光波段两者吸收光谱发生重叠,两者混合溶液在 408 nm 处产生一个新的吸收峰,且相比同浓度血红蛋白吸收峰而言 408 nm 处吸收峰发生了减色效应,这是由于血红蛋白与胆红素混合会发生某些弱的键位结合引起的,新的吸收峰改变了建模所用的主成分数,降低了模型的预测能力.结果说明在胆红素经皮测量中必须采取措施减少血红蛋白的影响,才能提高胆红素经皮检测的准确度.

关键词:血红蛋白;胆红素;光谱检测;偏最小二乘建模

中图分类号:O433.4

文献标识码:A

文章编号:1004-4213(2012)06-0684-5

0 引言

新生儿肝脏发育不完全,胆红素的不正常代谢会导致一系列的疾病(如黄疸),所以血清总胆红素的测量是临床中新生儿最常见的检测项目之一^[1].传统有创的胆红素检测方法虽然准确度较高,在 ± 0.03 mg/dL左右,但需要取血,增加了新生儿疼痛感和感染的风险.经皮胆红素监测仪快速、简单、无创,在临床胆红素检测中发挥了重要作用,但由于其准确度比较低,绝大多数监测仪一般为 ± 1.5 mg/dL,所以只能作为疾病的初次排查^[2].

为了提高人体胆红素检测准确度,研究者们一直致力于研究减小各项可能的影响因素.人体血液中血红蛋白作为胆红素检测中最大的影响因素,因而被研究者们所重视.1970年Kapitulnik等人^[3]研究了血红蛋白对羊水中胆红素化学与光谱检测的干扰,结果表明血红蛋白会明显的降低光谱450 nm处的吸收值而对化学检测法影响不大.1980年Bruce等人^[4-5]分别研究了血红蛋白对两种经典的总胆红素重氮检测法的干扰机制,结果表明血红蛋白对J-G方法结果影响不大,而对M-E方法有一定的影响.1993年Onks等人^[6]研究了血红蛋白对经

皮胆红素监测仪TcB读数的影响,结果表明水溶液中血红蛋白浓度的增加会引起460 nm和550 nm处光密度的线性增加,并降低TcB的读数,但并不成线性变化.2004年杨培慧等人^[7]研究了胆红素与血红蛋白分子作用的发光光谱,结果表明两者在偏碱性环境下混合会发生结构变化和弱的键位结合,造成紫外光谱的变化.2006年Walker等人^[8]研究了血红蛋白对总胆红素在Roche/Hitachi717分析仪检测的干扰,Lo等人^[9]研究了血红蛋白对总胆红素、未结合胆红素和结合胆红素在融合分析仪中的影响,结果都表明血红蛋白对相应胆红素化学检测法的影响可以忽略不计.大量的研究结果表明,血红蛋白对胆红素传统的化学检测法影响不大,而对胆红素的光学检测具有明显的影响,但是却鲜有文献对其影响机制进行分析.

目前市场上绝大部分经皮光学胆红素监测仪(如Bilitest和JM-103)都是使用双波长检测法,其利用血红蛋白在460 nm和550 nm吸收能力相同,通过差分来去除血红蛋白的影响,这种简单处理方法,导致其准确度很难达到临床监测的要求. Bilicheck使用多波长(137个波长)光谱检测法,利用胆红素浓度与检测光谱数据的线性相关性来建立

基金项目:国家自然科学基金(No. 60938002)资助

第一作者:蒋景英(1972-),女,副教授,博士,主要研究方向为生物医学光学检测. Email: jingying@tju.edu.cn

收稿日期:2011-11-07;修回日期:2012-02-24

相关模型并预测待检测人体胆红素浓度. 但鲜有文献对光谱检测中血红蛋白的具体影响机制进行分析研究. 本文研究了血红蛋白对胆红素光谱检测法尤其是在偏最小二乘回归 (Partial Least Squares Regression, PLS) 建模中对胆红素预测模型的影响, 并根据预测模型评价指标分析的结果及实验中光谱谱峰的变化情况, 对血红蛋白影响胆红素光学检测的机制进行了研究, 为进一步研制开发相关经皮胆红素光谱检测仪提供了理论依据.

1 偏最小二乘回归建模

偏最小二乘回归属于线性回归方法, 作为最常见的多变量建模方法之一, 主要用于样品浓度与光谱为线性关联的情况. 本文所分析的对象多为胆红素或其它成分的水溶液等, 而这些样品浓度的变化与光谱的变化基本成线性关系^[10-13].

1.1 主成分的提取

分析数据时, 光谱数据中往往包含了很多个测量波长, 较多的变量能携带更多的数据信息, 但同时也增加了分析问题的复杂性以及可能的多重相关性造成的信息重叠. 如何抓住系统的主要特征, 用较少的指标代替原本较多的变量, 又能综合反映原系统的信息, 主成分分析法 (Principle Component Analysis, PCA) 提供了有效的手段^[11].

主成分是对原变量的综合, 是原变量的线性组合, 对于标准化数据, 有^[12]

$$PC_h = \sum_{j=1}^p a_{hj} x_j \quad (1)$$

式中 h 为主成分 PC 个数, j 为原变量 x 个数.

使用 PCA 和 PLS 方法建立校正模型, 其中最困难的问题之一是如何确定能够建立模型所使用的主成分数目. 若建立模型时所使用的主成分数过少, 就不能反映样品被测组分产生的量测数据 (如光谱) 变化, 其模型预测准确度就会降低, 称为不充分拟合; 若使用过多主成分建立模型, 就会夹杂一些代表噪音, 使模型预测能力下降, 称为过度拟合^[13].

随着主成分数量增加, 数据可以推广到 m 维的主超平面, 从变量相关程度的角度看, 主超平面能够携带最多的原始信息, 对原始变量系统有最佳的解释能力, 在主超平面上, 最好地拟合原样本点集合. 因此, 主成分的数量和选择对模型的建立具有重要意义. 本文在 PLS 建模中, 采用完全交互验证法, 计算得到 PRESS 值确定主成分的数量并对模型进行验证.

1.2 PLS 模型的评价指标

模型的评价指标包括相关系数 (Correlation

coefficient) 和预测均方根误差 (Root Mean Square Error of Prediction, RMSEP), RMSEP 用来评价模型的预测能力. 相关系数越接近 1, RMSEP 值越小说明模型的拟合程度越好, 预测准确度越高^[13-14].

$$RMSEP = \sqrt{\frac{\sum (\hat{y}_i - y_i)^2}{n-1}} \quad (2)$$

式中 y_i 为真实浓度值, \hat{y}_i 为模型的预测浓度结果.

2 材料与实验方法

2.1 仪器与试剂

紫外光谱检测采用日本岛津公司生产的 UV-2550 紫外可见分光光度计; PLS 建模采用 CAMO 公司的 Unscrambler9.7 软件.

牛血红蛋白 (*Hb*, 国药集团化学试剂有限公司, 样品紫外吸收度 [406 nm, 0.125 g/L, 1 cm] ≥ 0.88), 配制成 2.5×10^{-4} mol/L 溶液贮存; 胆红素 (*BR*, 日本和光纯药标准校准品, 3 mL 标准液的浓度为 12.7 mg/dL), 每瓶制成 25 mL 浓度 1.524 mg/dL 溶液贮存. 所有试剂除注明外均为分析纯, 实验用水为蒸馏水.

2.2 实验方法

仪器进行预热和基线校准后, 分别对 *Hb* 溶液 10^{-6} mol/L、*BR* 溶液 0.2 mg/dL 及同浓度二者混合液进行紫外可见吸收光谱的测量; 配制 21 组具有固定浓度梯度的 *BR* 溶液不加入 *Hb*, 随后配制 21 组具有相同浓度梯度的 *BR* 溶液并加入一定浓度 *Hb*, 分别对它们进行紫外可见光谱扫描, 然后根据得到的光谱数据在 430~490 nm 范围分别建立 PLS 模型并进行胆红素浓度预测. 根据是否加入 *Hb* 溶液的光谱变化和对分别建立的 PLS 模型进行对比分析, 其中各样品体积为 5 mL, 加入的 *Hb* 浓度为 5×10^{-5} mol/L, 21 组 *BR* 溶液最小浓度为 0.030 48 mg/dL, 最大浓度为 0.640 08 mg/dL, 浓度梯度为 0.030 48 mg/dL.

3 实验结果与讨论

3.1 *Hb* 对 PLS 建模的影响

人体血液中正常血红蛋白含量为 110~200 g/L, 人体血液中正常胆红素含量为 0.2~1 mg/dL, 实验溶液中血红蛋白浓度 5×10^{-5} mol/L $\approx 3.224 8$ g/L 远小于人体正常血红蛋白含量, 而实验中胆红素浓度范围 0.030 48~0.640 08 mg/dL 处于人体正常胆红素含量范围内.

3.1.1 *Hb* 对 PLS 建模最大贡献波长及主成分数的影响

由图 1(a) 中可知, *BR* 溶液模型光谱变量的回

归系数在 459 nm 存在最大值,即在胆红素的吸收峰 460 nm 附近处光谱数据对模型建立的贡献最高;图 1(b)中 *Hb* 和 *BR* 混合液模型光谱变量的回归系数在 432 nm 和 433 nm 回归系数最高,可见 *Hb* 的加入影响了溶液的吸收峰,使建模的主成分从 460 nm 往 *Hb* 的吸收峰 406 nm 发生了偏移.表 1 中各波长变量的回归系数值定量的说明了这一结论.

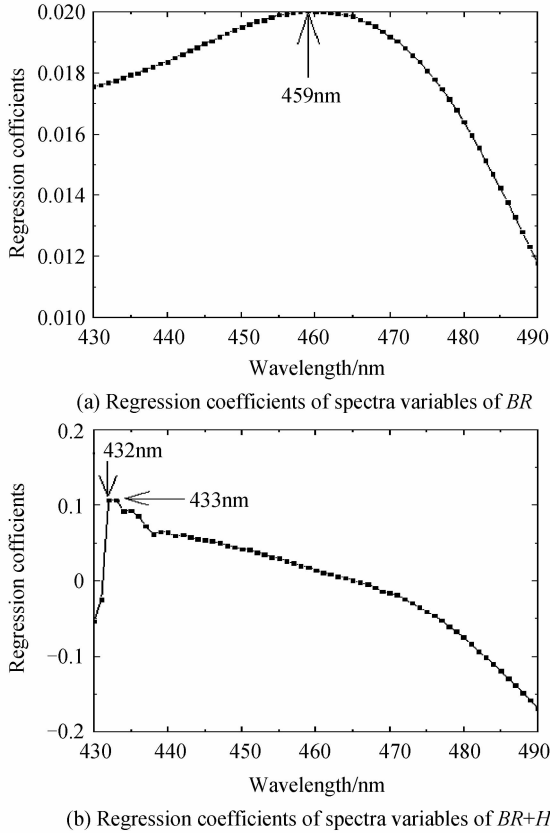


图 1 *BR* 及其与 *Hb* 混合液模型光谱变量的回归系数
Fig. 1 Regression Coefficients of Spectra variables for *BR* and *BR+Hb*

表 1 光谱变量的回归系数

Table 1 Regression Coefficients of spectra variables

| Wavelength/nm | PC_01 (a) | PC_03 (b) |
|---------------|-----------|-----------|
| 430 | 0.017 56 | -0.053 44 |
| 431 | 0.017 6 | -0.024 65 |
| 432 | 0.017 68 | 0.107 |
| 433 | 0.017 75 | 0.107 |
| 434 | 0.017 85 | 0.092 47 |
| 435 | 0.017 93 | 0.092 76 |
| 436 | 0.017 99 | 0.085 73 |
| 437 | 0.018 09 | 0.072 41 |
| 438 | 0.018 19 | 0.061 89 |
| ... | ... | ... |
| 457 | 0.019 97 | 0.023 12 |
| 458 | 0.020 01 | 0.019 61 |
| 459 | 0.020 03 | 0.017 73 |
| 460 | 0.020 01 | 0.014 15 |

通过对比 *BR* 溶液模型和混合液模型,发现建模主成分个数从一个主成分 *PC_01* 变为三个主成分 *PC_03*,数据贡献最大波长由 459 nm 变为了 432 nm,433 nm,使建立模型的光谱数据变量带入了血红蛋白的吸收噪音,因此血红蛋白的加入降低了模型对胆红素浓度的预测能力.

3.1.2 *Hb* 对 PLS 模型评价指标的影响

表 2 中 *BR* 溶液模型的 $RMSEP=0.016 184$, *BR* 和 *Hb* 混合液模型的 $RMSEP=0.021 549$ mg/dL,它们的值都比实验中胆红素浓度梯度 0.030 48 mg/dL 小,而相关系数都大于 0.99,所以模型模拟的结果有意义.其中加入了 *Hb* 的混合液模型比不加入 *Hb* 的 *BR* 溶液模型的 $RMSEP$ 值增大了 24.90%;而加入了 *Hb* 的混合液模型比不加入 *Hb* 的 *BR* 溶液模型的相关系数减小了 0.29%.根据 PLS 评价指标性质,相关系数越大越接近 1, $RMSEP$ 值越小说明模型拟合程度越好预测准确度越高,因此血红蛋白的加入降低了模型对胆红素浓度的预测能力 ($RMSEP$ 增大 24.90%) 和模型的拟合能力 (相关系数减小 0.29%).

表 2 PLS 模型的评价指标

Table 2 Evaluation indices of PLS model

| PLS model | RMSEP(mg/dL) | Correlation coefficient |
|--------------|--------------|-------------------------|
| <i>BR</i> | 0.016 184 | 0.996 149 |
| <i>BR+Hb</i> | 0.021 549 | 0.993 239 |

3.2 *Hb* 和 *BR* 的吸收光谱

有机物化合物分子中存在化学键和未参与成键的电子,正是这些电子的跃迁形成了有机化合物的电子光谱,即紫外和可见光谱.

进一步对 *Hb* 溶液、*BR* 溶液及二者同浓度混合液进行紫外可见吸收光谱扫描.图 2 中得到检测波段 *Hb* 溶液的吸收峰为 406 nm, *BR* 溶液的吸收峰为 460 nm. *Hb* 与 *BR* 混合液的紫外吸收光谱有一定的变化,混合液在 408 nm 处有明显吸收峰,且相

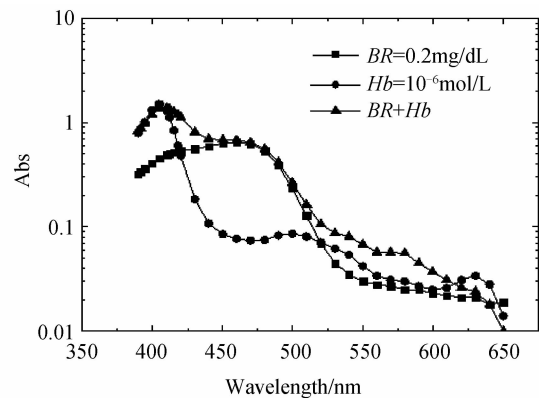


图 2 *Hb*、*BR* 及二者混合液的紫外可见吸收光谱
Fig. 2 UV absorption spectra of *Hb* and *BR*

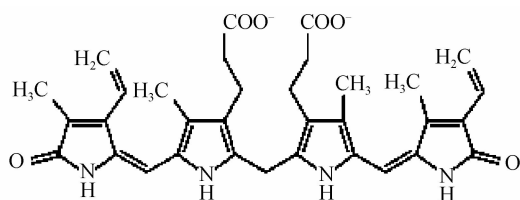


图3 BR的化学结构

Fig.3 Chemical structure of BR

对于 406 nm 处同浓度血红蛋白吸收峰而言吸收降低,产生减色效应.结合图 3 中 BR 的化学结构可知,BR 含有二吡咯发色团,其能量移动会导致光吸收的劈裂,且 BR 的两个-COOH 会与 Hb 中氨基酸残基上的-NH₂ 结合,会形成复杂的卷曲结构,所以两者混合溶液在 408 nm 处形成吸收峰.此吸收峰影响了建模中的光谱数据,引进了新的主成分数,造成了模型预测能力的下降.

4 结论

本文对比分析了不加入 Hb 和加入 Hb 光谱数据的建模结果.得到以下结论:Hb 加入会造成 BR 和 Hb 混合液建模主成分数的增加并对建模贡献较大波长具有影响,使相应 PLS 模型的 RMSEP 值增大,并使模型相关系数降低,降低了模型的预测能力和模型的拟合能力.结合 BR 的化学结构,发现 Hb 与 BR 混合有弱的键位结合作用,影响了二者混合液的紫外可见吸收光谱谱峰,所以造成了建模中主成分数的变化.说明了血红蛋白影响了胆红素光谱检测的准确度,只有研究去除血红蛋白的影响,才能提高胆红素经皮检测的准确度.

参考文献

- [1] KAZMIERCZAK S C, ROBERTSON A F, CATROU P G, *et al.* Direct spectrophotometric method for measurement of bilirubin in newborns: comparison with HPLC and an automated diazo method[J]. *Clinical Chemistry*, 2002, **48**(7): 1096-1097.
- [2] HO E, LEE S, CHOW C, *et al.* Bilirubin transcutaneous bilirubinometer: a screening tool for neonatal jaundice in the chinese population[J]. *Hong Kong Medical Journal*, 2006, **12**(2): 99-102.
- [3] KAPITULNIK J, KAUFMANN N A, BLONDHEIM S H. Chemical versus spectrophotometric determination of bilirubin in amniotic fluid, and the influence of hemoglobin and metheme pigments[J]. *Clinical Chemistry*, 1970, **16**(9): 756-759.
- [4] SHULL B C, LEES H, LI P K. Mechanism of interference by hemoglobin in the determination of total bilirubin. i. method of malloy-evelyn[J]. *Clinical Chemistry*, 1980, **26**(1): 22-25.
- [5] SHULL B C, LEES H, LI P K. Mechanism of interference by hemoglobin in the determination of total bilirubin. ii. method of jendrassik-grof[J]. *Clinical Chemistry*, 1980, **26**(1): 26-29.
- [6] ONKS D, SILVERMAN L, ROBERTSON A. Effect of melanin, oxyhemoglobin and bilirubin on transcutaneous bilirubinometry[J]. *Acta Paediatr*, 1993, **82**(1): 19-21.
- [7] YANG Pei-hui, ZHENG Zhi-wen, CAI Ji-ye. Spectral analyses for the molecular interaction between bilirubin and hemoglobin[J]. *Chinese Journal of Luminescence*, 2004, **25**(3): 247-251.
杨培慧, 郑志雯, 蔡继业. 胆红素与血红蛋白分子作用的发光光谱分析[J]. *发光学报*, 2004, **25**(3): 247-251.
- [8] WALKER M J, BORSSES D M. Performance characteristics of a newly reformulated total bilirubin procedure displaying minimal effects from hemoglobin interference [J]. *Clinical Chemistry*, 2006, **52**(6): A199.
- [9] LO S, DOUMAS B. Effect of hemoglobin on the measurement of (TBIL), unconjugated (Bu) and conjugated (Bc) bilirubins in the fusion analyzer[J]. *Clinical Chemistry*, 2006, **52**(6): A36.
- [10] WU Wei, YANG Xiao-min, CHEN Mo, *et al.* Face hallucination based on partial least squares [J]. *Acta Photonica Sinica*, 2009, **38**(11): 3025-3033.
吴炜, 杨晓敏, 陈默, 等. 基于偏最小二乘算法的人脸图像超分辨率技术[J]. *光子学报*, 2009, **38**(11): 3025-3033.
- [11] WANG Shuang, HUANG Min, ZHU Qi-bing. Optimal wavelength selection of hyperspectral scattering images based on UVE-PLS projection analysis[J]. *Acta Photonica Sinica*, 2011, **40**(3): 428-432.
王爽, 黄敏, 朱启兵. 基于无信息变量和偏最小二乘投影分析的高光谱散射图像最优波段选择[J]. *光子学报*, 2011, **40**(3): 428-432.
- [12] CHEN Zhong, ZHU Aio-shi, CHEN De-zhao. Combined prediction model based on partial least squares regression and its application to near infrared spectroscopy quantitative analysis[J]. *Chinese Journal of Analytical Chemistry*, 2007, **35**(7): 978-982.
成忠, 诸爱士, 陈德钊. 组合偏最小二乘回归方法在近红外光谱定量分析中的应用[J]. *分析化学*, 2007, **35**(7): 978-982.
- [13] 陆婉珍. 现代近红外光谱分析技术[M]. 2 版. 北京: 中国石化出版社, 2007: 44-180.
- [14] 徐可欣, 高峰, 赵会娟. 生物医学光子学[M]. 北京: 科学出版社, 2007: 135-137.

Influence of Hemoglobin on Non-invasive Optical Bilirubin Sensing

JIANG Jing-ying^a, GONG Qi-liang^a, XU Ke-xin^b

(a. *Tianjin Key Laboratory of Biomedical Detecting Techniques and Instruments*; b. *State Key Laboratory of Precision Measuring Technology and Instruments, School of Precision Instrument and Opto-electronics Engineering, Tianjin University, Tianjin 300072, China*)

Abstract: The spectral data were analyzed by using the Partial Least Squares Regression. The result shows that the addition of hemoglobin changes the number of principle component from one into three, while the largest contribution wavelength changes from 459 nm to 432 nm and 433 nm. This indicates that absorption noise of hemoglobin was introduced in the spectral data used for PLS model. Furthermore, compared with pure bilirubin solution, the Root Mean Squared Error of Prediction(RMSEP) of the model for mixed solution increases 24.90%, and the correlation coefficient of the model for mixed solution reduces 0.29%. It means that the addition of hemoglobin decreases the prediction ability of the model. Meantime, a new absorption peak at 408 nm appears, which seems there is a hypochromic effect, due to some weak key combinations between hemoglobin and bilirubin. And the new absorption peak changes the number of principal components and reduces the predictive ability of the model. Overall, the influence of hemoglobin should be reduced in order to improve the transcutaneous detection accuracy of bilirubin.

Key words: Hemoglobin; Bilirubin; Spectral measurement; Partial least squares regression modeling