Chemometrics and Intelligent Laboratory Systems 145 (2015) 84–92

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemolab

Determination of glycated hemoglobin using near-infrared spectroscopy combined with equidistant combination partial least squares

Yun Han [a], Jiemei Chen [b,*], Tao Pan [a,*], Guisong Liu [c]

[a] Department of Optoelectronic Engineering, Jinan University, Huangpu Road West 601, Tianhe District, Guangzhou 510632, China
[b] Department of Biological Engineering, Jinan University, Huangpu Road West 601, Tianhe District, Guangzhou 510632, China
[c] Department of Mathematics, Jinan University, Huangpu Road West 601, Tianhe District, Guangzhou 510632, China

ARTICLE INFO

Article history:
Received 11 April 2014
Received in revised form 20 April 2015
Accepted 21 April 2015
Available online 29 April 2015

Keywords:
Glycated hemoglobin
Near-infrared spectroscopic analysis
Equidistant combination PLS
Competitive adaptive reweighted sampling combined with PLS
Monte Carlo uninformative variable elimination by PLS

ABSTRACT

A novel near-infrared-spectroscopy-based quantification method for glycated hemoglobin (HbA1c), a major clinical diagnosis indicator of diabetes, was developed on the basis of simultaneous determination of hemoglobin (Hb) and absolute HbA1c content (Hb•HbA1c) in human hemolysate samples. Equidistant combination partial least squares (EC-PLS) method was proposed to perform wavelengths selection. Competitive adaptive reweighted sampling PLS (CARS-PLS) and Monte Carlo uninformative variable elimination PLS (MC-UVE-PLS) methods were also conducted for comparison. A randomness and stability dependent rigorous process of calibration, prediction, and validation was performed to produce objective and stable models. The search range covered the unsaturated region (780–1880 nm, 2090–2330 nm). For Hb and Hb•HbA1c, only 6 and 14 wavelengths were selected with EC-PLS, 23 and 30 wavelengths were selected with CARS-PLS, and 100 and 120 wavelengths were selected with MC-UVE-PLS, respectively.

The predicted values of relative percentage HbA1c were calculated from the predicted Hb and Hb•HbA1c values. The sensitivity and specificity for diabetes were 93.5% and 97.1% with EC-PLS, 91.3% and 94.1% with CARS-PLS, and 89.1% and 76.5% with MC-UVE-PLS, respectively. In three methods, EC-PLS not only employed the least wavelengths but also produced the best quantification accuracy for HbA1c. EC-PLS also achieved the highest classification accuracy for negative and positive samples for diabetes.

The results confirm the feasibility of HbA1c quantification based on the simultaneous analysis of Hb and Hb•HbA1c with NIR spectroscopy. This technique is rapid and simple when compared with conventional methods, and is a promising tool for screening diabetes in large populations.

1. Introduction

At present, diabetes poses an increasingly grave threat to human health. Therefore, the existing prevention measures must be strengthened. In conventional diagnosis and therapeutic monitoring of diabetes, several parameters such as postprandial glucose, fasting glucose, and oral glucose tolerance must be measured. However, the measured values of these parameters only correspond to the instantaneous blood glucose level. Glycated hemoglobin (HbA1c) is an effective indicator of long-term blood glucose level and has been widely used in the screening and diagnosis of diabetes, as well as for evaluating the effect of therapy on diabetes. Glycated hemoglobin contains HbA1 and other hemoglobin-glucose adducts. Futher, HbA1 is composed of HbA1a, HbA1b, and HbA1c. HbA1c, which is the major component of HbA1, is formed by a non-enzymatic irreversible process of combination of the aldehyde group of glucose with the amino-terminal valine of the β-chain of hemoglobin (Hb). This process involves a series of non-enzymatic reactions referred to as Maillard reactions [1]. The clinical value of HbA1c is expressed as a unit of relative percentage that is equal to the ratio of the absolute content of HbA1c to the amount of Hb. In clinical practice, HbA1c is referenced to a nondiabetic range of 4.0% to 6.0%. For phenotype-positive patients for diabetes, HbA1c > 6.0% [2].

Currently, numerous methods of HbA1c assessment are in use [3]. These methods are classified into two main groups on the basis of assay principles. The first group includes methods based on the charge differences between glycated and nonglycated components. The second group contains methods that separate components depending on the structural differences between glycated and nonglycated components. Typically, results obtained with these techniques vary slightly depending on the particular hemoglobinopathy and assay methodology used. Moreover, these methods are complicated and require chemical reagents and expertise for operation. Therefore, new simple methods are required for assessing long-term glycemic levels.

Near-infrared (NIR) spectroscopy primarily reflects absorption of overtones and combination of vibrations of X-H functional groups (such as C-H, O-H, and N-H). This rapid, simple technique is commonly used in many fields. Both glucose and hemoglobin molecules contain

* Corresponding authors.
E-mail addresses: tchjm@jnu.edu.cn (J. Chen), tpan@jnu.edu.cn (T. Pan).

http://dx.doi.org/10.1016/j.chemolab.2015.04.015
0169-7439/

different X-H functional groups that have significant absorption in the NIR region. Chemical-free and rapid analysis of glucose and hemoglobin using NIR spectroscopy has been the focus of previous studies [4–10]. The glycated hemoglobin process (Maillard reactions) involves some functional groups that contain hydrogen. Therefore, NIR spectroscopy can be used to obtain information about glycated Hb.

Previously, refractive index measurements method was used to investigate the effect of presence of glucose and glycation of proteins on the optical properties of water solutions of hemoglobin and albumin with different glucose concentrations in [1]. In addition, the NIR absorbance spectra of water solutions of hemoglobin and albumin were studied with different glucose concentrations. However, using NIR spectroscopy to directly measure blood glycated hemoglobin is more convenient for clinical application.

As a relative percentage of total Hb, HbA1c and spectral absorbance do not fulfill Beer's law because the absolute content (Hb•HbA1c) of samples with the same relative percentage (HbA1c) may vary when the amount of total Hb is different. Indirect measurement of HbA1c is considered. Because Hb levels can be measured using NIR spectroscopy [7–10], we assume that Hb•HbA1c can also be measured with the same method. By simultaneously measuring the two indicators, the predictive value of HbA1c can be obtained. In this study, experiments were conducted to confirm the feasibility of simultaneous quantitative analysis of Hb•HbA1c and Hb for blood samples using NIR spectroscopy.

Because human blood is a complex system with multiple components, spectroscopic analysis of some components in human blood must mitigate noise disturbance of other components. For the rapid and chemical-free measurement of a complex system using NIR spectroscopy, appropriate spectral wavelength selection method is an important, albeit difficult, aspect. Therefore, it is important to improve the effectiveness of spectral prediction, reduce method complexity, and design a specialized spectrometer with a high signal-to-noise ratio (SNR). Further, appropriate chemometric methods must be employed to optimize wavelength.

Partial least squares (PLS) regression is a popular multivariate calibration method that has been widely applied in multicomponent spectral analysis, especially in vibrational spectroscopy, such as NIR. The moving window PLS (MW-PLS) method has been proven to be an effective method of spectral analysis with high prediction capability [5, 10–15]. In addition, stacked-PLS method [16] is also a well-performed and PLS-based waveband selection method. However, the continuous-mode models based on PLS typically lead to high model complexity.

Non-adjacent wavelengths are typically selected with discrete mode, which is designed to minimize colinearity problems. Based on the achievements of multiple linear regression (MLR) and MW-PLS, the equidistant combination multiple linear regression (EC-MLR), has been proposed in our previous study [17,18], which is chosen as the equidistant discrete wavelength selection method. EC-MLR improves the conditioning of MLR by minimizing colinearity effects when the data gap selection is appropriate. EC-MLR method is in the quasi-continuous mode with a low degree of freedom and low computational complexity, which inherits the merits of both the continuous and discrete modes, which can easily undergo spectral preprocessing to improve prediction capability.

In addition, as we all know, PLS method has better prediction effect than MLR method in the case of the same wavelength combinations. Because PLS can be integrated screening variables, and it has been very popular and easy to use. In this study, EC-MLR method was improved to the EC-PLS method and was applied to the analysis of glycated hemoglobin. Indeed, the EC-PLS method also is a discrete wavelength selection method. There are many well-performed and PLS-based methods for discrete wavelength selection, such as competitive adaptive reweighted sampling combined with PLS (CARS-PLS) [19], Monte Carlo uninformative variable elimination by PLS (MC-UVE-PLS) [20], and so on.

In this study, the proposed EC-PLS method was employed to select appropriate NIR wavebands and wavelength combinations for Hb and Hb•HbA1c, which correspond to the information of glycated Hb. EC-PLS was performed for appropriate quasi-continuous wavelength selection. While discrete wavelength selection through the CARS-PLS and MC-UVE-PLS methods were also conducted for comparison.

Table 1
Statistical analysis of measured Hb and HbA1c values of 240 samples.

Sample types	Number of samples	Hb (g L^{-1})				HbA1c (%)			
		Min	Max	Mean	SD	Min	Max	Mean	SD
All samples	240	77	165	130.8	14.2	4.6	10.8	6.32	0.98
Negative	104	78	165	128.0	14.2	4.6	6.0	5.57	0.35
Positive	136	77	156	133.0	13.9	6.1	10.8	6.90	0.92

Note: SD is the abbreviations of standard deviation.

2. Materials and methods

2.1. Experimental materials, instruments, and measurement methods

A total of 240 human peripheral blood samples were collected and placed in 0.2% ethylenediaminetetraacetic acid-containing tubes. The Hb values of these samples were measured with a BC-3000Plus automatic blood cell analyzer (Shenzhen Mairui Ke Technology Co., Ltd., China). The HbA1c values of the samples were measured with an ADAMS™ A1c HA-8160 automatic glycated Hb analyzer (ARKRAY, Inc., Japan) using high-pressure liquid chromatography analysis method. The obtained values were used in the calibration, prediction, and validation sets as the reference values for spectroscopic analysis. The statistical analysis of the measured Hb and HbA1c values of the 240 samples is given in Table 1. On the basis of the cut-off value of HbA1c (6.0%), 104 negative and 136 positive samples were obtained.

Scattering and noise disturbance may occur when light passes through the samples because the peripheral blood samples are highly viscous. Furthermore, Hb is contained within an erythrocyte, and the noise from the cell membrane must be overcome to determine Hb levels. Therefore, the accuracy of spectral analysis of peripheral blood samples may decrease. Previously in [9], we compared the effects of peripheral blood and hemolysate samples on the prediction capability of the method for Hb. The peripheral blood samples were configured to be 2 ×, 3 ×, 4 ×, 5 ×, and 6 × dilute hemolytic solution samples. Six sample groups (including the peripheral blood samples group) were obtained, and the Hb calibration and prediction models were then established for each group based on the second derivatives of the spectra with 11 points SG smoothing. The results showed that the group with the peripheral blood samples had significantly lower prediction accuracy whereas the prediction was in good agreement for each hemolysate

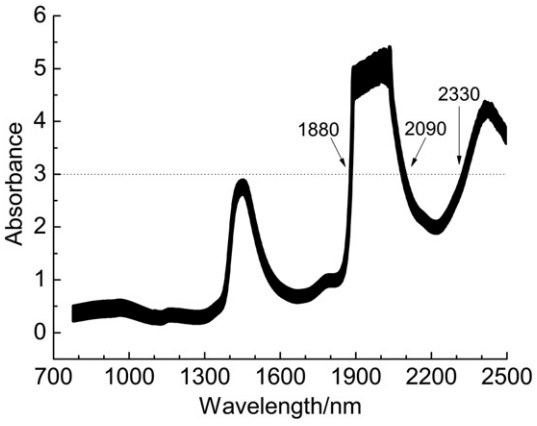

Fig. 1. NIR spectra of 240 human hemolysate samples in the entire scanning region (780–2498 nm).