Deep Learning Analysis for Blood Glucose Monitoring Using Near Infrared Spectroscopy

CF So[1]*, Kup-Sze Choi[2], Thomas KS Wong[3] and Joanne WY Chung[4]

[1]*Department of Mathematics and Information Technology, Hong Kong*

[2]*Centre for Smart Health, Hong Kong*

[3]*Guangzhou University of Chinese Medicine, China*

[4]*Department of Health and Physical Education, Hong Kong*

*Corresponding author: CF So, The Education University of Hong Kong, Hong Kong

ARTICLE INFO

ABSTRACT

Background and Objective: The race for the next generation of painless and reliable glucose monitoring for diabetes is on. Near infrared spectroscopy has become a promising technology among others for blood glucose monitoring. While advances have been made, the reliability and the calibration of non-invasive instruments could still be improved. The objective of this study was to set up a non-invasive blood glucose measurement device based on deep learning analysis to detect the spectral response from human tissue.

Methods: This study successfully adopted the four-stage framework of bio-signal processing to handle the near infrared spectroscopy that is used to measure blood glucose level. The major contributions the study makes include the selection of pre-processing methods (generalized least squares weighting pre-filter) and the algorithm for feature selection (genetic algorithm) and developing computational algorithm (partial least squares discriminant analysis using Monte Carlo) to improve the performance and accuracy rate of the calculation.

Results: An improved method based on Monte Carlo approach for the partial least squares is proposed. The overall classification rate of the model reached 75.2%. This algorithm outperforms conventional multivariate methods, whereby predicting the relationship between the response and the independent variables is more accurate, thus enhancing the reliability of the regression model.

Conclusion: The findings obtained in the study provide a useful reference for future development in non-invasive blood glucose measurement.

Keywords: Deep Learning Analysis; Near Infrared; Blood Glucose Monitoring; Partial Least Squares.

Introduction

Diabetes Mellitus (DM) is an intractable condition in which blood glucose levels cannot be regulated normally by the body alone [1] it has many complications, including heart disease and stroke, kidney failure, blindness or vision problems, diabetic neuropathy and diabetic foot. Treatment methods include dietary regulation to control blood glucose levels, oral medication, and insulin injection, and all of these treatments should rely on blood glucose measurement. Diabetic patients are encouraged to check their blood glucose levels several times per day [1]. Currently, the most common means of checking is by using a self-monitoring blood glucose meter [2]. In this way, diabetic patients can obtain a clear picture of their blood glucose levels for therapy optimization and for insulin dosage adjustment for those who need daily

injections. However many people dislike using sharp objects and seeing blood because there is a risk of infection. Over the long term, this practice may also result in damage to finger tissue. Given these realities, the advantages of a non-invasive technology are easily understood. The objective of this study is to construct a non-invasive blood glucose monitoring device based on Near-Infrared (NIR) sensing that is stable and easy to use to detect the spectral response from human tissue. NIR spectroscopy is selected because of its promising technology, among others, for non-invasive blood glucose monitoring [3]. This study adopts the four-stage conceptual framework of bio-signal processing [4] which shows that processing bio-signals should normally consist of

   a.   Signal acquisition or measurement

   b.   Signal transformation or signal pre-processing

   c.   Parameter selection or variable/feature selection

   d.   Signal classification or signal interpretation

In the first stage, signal acquisition, a non-invasive blood glucose monitoring device based on NIR spectroscopy is set up. NIR light is transmitted through an optical cable to the human skin surface and the reflected light is collected for computing blood glucose concentration in human blood. Second stage, also called pre-processing, the signals are transformed in a way that such problems can be simplified through a suitable choice of pre-processing method, and misleading results can be avoided. Since the pre-processing of NIR spectral data is important for the subsequent multivariate analysis, various normalization schemes and pre-processing techniques are evaluated and reported. The third stage delivers relevant variables, also called features, which can be used for decision making. The features are extracted by some complex searching algorithms to distinguish those signal features that contain discriminatory power, and normally they can reduce the size of data so as to compute the diagnostically most significant parameters. Once the signal parameters have been obtained, they are used for further decision making in the interpretation stage. The classification stage of bio-signal processing is the stage of identifying which set of categories a new measurement belongs to, on the basis of a set of training data containing relevant signal parameters whose category membership is known. This study presents an improved algorithm to deal with the signal classification stage that can provide more accurate classification results as well as enhance the stability of signal interpretation.

Signal Acquisition

This study was based on clinical validation of blood glucose levels between laboratory results and NIR absorption spectroscopy analysis. In the *in-vitro* validation process, the human blood glucose level was employed as the gold standard, with the spectrometry method used for the testing. Subjects were recruited from the community by convenience. All subjects who were interested in

the study and able to provide a voluntary informed consent were recruited to the study. All known infectious disease subjects were excluded. The major equipment in the experiment was the NIR spectrometer from Control Development, Inc. (CDI, Foundation Drive, South Bend, IN), the probe, and the tungsten halogen light source. This study considers the wavelength range between 1,121-1,880nm for analysis.

All subjects were asked to fast overnight before participating in the study. Spectra were obtained using the assembled spectrometer at the left fourth finger. Simultaneous veni-punctures for blood glucose were performed. Since the aim was to collect high blood glucose levels as much as possible in this study, all subjects with DM history were required to have blood taken one more time, at least 30 minutes after they had had their breakfast. The non-DM subjects were asked to give their consent before being invited to have breakfast (with calories ranging from 233 to 298 Kcal), and then had their blood taken for glucose measurement. Unlike the DM subjects, this was optional for non-DM subjects. Five hundred and twelve subjects (225 male and 287 female) voluntarily participated in the study. The mean age was 52.33 (SD12.8). Among them, 219 (42.8%) suffered from DM. A total of 840 samples were collected.

Signal Transformation

Pre-processing of spectral data is an important procedure before chemometric analysis according to the four-stage framework of bio-signal processing because scaling differences arise from path-length effects, scattering effects, source variations, or other instrumental sensitivity effects in NIR spectroscopy that will influence the measurement of the recorded sample. Various pre-processing techniques such as normalization process and pre-filtering process are explored and evaluated.

Normalization

In spectroscopic measurement, scaling differences arise from path length effects, scattering effects, source variations, or other instrumental sensitivity effects. A normalization process attempts to correct for these kinds of effects by identifying some aspect of each sample that should essentially be constant from one sample to the next and correcting the scaling of all variables based on this characteristic. Several different types of norms exist. 1-norm is used to divide each variable by the sum of the absolute value of all variables for the given samples [5]. Another commonly used norm in 2-dimensional Euclidean space is the 2-norm. The Standard Normal Variate (SNV) normalization process is a weighted normalization [6]. It is different from the 1-norm or 2-norm mentioned above because not all samples contribute to the normalization equally. Multiplicative Scatter correction (MSC) is another method for normalization [6]. It is based on the idea of correcting the scatter level of all spectra of a group of samples to the level of an average spectrum [7].

Performance Comparison - Normalization

The performance analysis is assessed by the root mean square error (RMSE) and the correlation coefficient. RMSE helps describe the fit of the model to the training data. It is defined as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(\tilde{y}_i - y_i\right)}{n}} \qquad (1)$$

where $\tilde{y}_i$ are the values of the predicted result and n is the number of training samples.

The correlation coefficient (R) which is used for comparing the correlation between the predicted value and the actual value. It is a measure of how well the predicted values from a forecast model fit with the data. As the strength of the relationship between the predicted values and actual values increases so does the correlation coefficient. Therefore, the higher the correlation coefficient the better the predicted model. The correlation coefficient is given by the formula:

$$R = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}} \qquad (2)$$

where $\bar{x}$ and $\bar{y}$ are the samples means of variables x and y

The PLS regression model [8] was used to study the performance of various normalization processes as it is widely used in multivariate calibration method. The PLS regression model was constructed with the Latent Variables (LVs) set equal to 10 in this simulation. The cross-validation was performed to test the robustness of the normalization process. The RMSEcv of cross-validation are shown to determine which process provides the minimum RMSEcv. In addition, the Rcv for cross-validation show how well the prediction model can perform. (Table 1) shows the results for the four normalization processes. The results shows that 2-norm performs best, and that 1-norm and 2-norm provide very similar RMSE and R, while SNV and MSC deliver relatively higher RMSE than 1-norm and 2-norm and the correlation coefficients are relatively smaller than they are. Since 2-norm is a form of weighted normalization where larger values are weighted more heavily in the scaling and it is also the most commonly used normalization process in 2-dimensional Euclidean space, 2-norm was used in the following analysis to take advantage of the uneven weighting mechanism.

Table 1: Results for the four normalization processes.

Normalization Process	RMSEcv	Rcv
1-norm	2.9868	0.4435
2-norm	2.9865	0.4436
SNV	3.0291	0.419
MSC	3.0297	0.4187

Pre-Filter Analysis

Before training the quantitative model, it is possible to manipulate the spectra using various pre-filter methods to try to improve the performance of the quantitative model. Many of the interferences, often described as noise, are caused by known background signals and artefacts. Pre-filtering the signals to remove this noise or the effects of signal variance can be useful for obtaining a more accurate quantitative model [9]. A variety of pre-filter methods exist to remove the interferences [10]. This section evaluates various kinds of pre-filter methods for the NIR spectra to enhance performance of the PLS prediction model. The Savitzky-Golay (SG) derivation is commonly used for numerical derivation because it includes a smoothing step when it takes the derivative, in which it can improve the utility of data [11].

This method requires selection of the window size (filter width) and the order of the polynomial [12]. The Goal of a Generalized Least Squares (GLS) weighting pre-filter is to down-weight the differences between similar samples and thus make them appear more similar. A GLS weighting pre-filter can be used to remove variance from the spectral responses that are mostly orthogonal to the concentration information. The GLS weighting pre-filter method involves the calculation of a covariance matrix from the differences between similar samples. Orthogonal Signal Correction (OSC) was introduced to remove systematic variation from the spectral responses that is unrelated or orthogonal to concentration information [13]. Such variance is identified as some number of components of the spectral responses that have been made orthogonal to the concentration.

Performance Comparison - Pre-Filter Analysis

Various pre-filter methods mentioned previously were compared; each method had its own parameter settings and the results are listed in Table 2 The first SG derivative and second SG derivative methods were implemented. The GLS weighting pre-filter method with various α values were evaluated. Moreover, the OSC pre-filter with various PCs and the tolerance values is listed as well. The results with and without normalization are shown together for comparison. This experiment considered performing the pre-filter with and without normalization for their own parameter settings. Comparing the results with and without normalization showed that the normalization process can improve the RMSE and R for all pre-filter methods, particularly for the GLS weighting pre-filter method. Therefore the results further revealed the benefit of implementing the normalization process and the pre-filter process thereafter so as to provide better performance for analysis. According to the simulation results, GLS weighting pre-filter performed the best amongst all methods. It yielded the smallest RMSEcv and highest Rcv when α=0.005.

Table 2: Results for various pre-filter methods with and without 2-norm applied.

Methods		$RMSE_{cv}$	R_{cv}	$RMSE_{cv}$	R_{cv}
		With 2-Norm Applied		**Without 2-Norm Applied**	
1st SG Derivative	(order 2, window 15pt)	3.0523	0.4141	3.0266	0.428
	(order 6, window 7pt)	3.0263	0.4343	3.0259	0.4341
2nd SG Derivative	(order 2, window 15pt)	3.0597	0.4171	3.0627	0.4158
	(order 3, window 9pt)	3.1043	0.3887	3.1068	0.3834
	(order 5, window 9pt)	3.116	0.372	3.1212	0.3728
GLS weighting	(α=0.001)	3.0359	0.4668	7.8822	0.089
	(α=0.002)	2.9892	0.4784	7.0674	0.0952
	(α=0.005)	2.9539	0.4807	4.271	0.0185
	(α=0.01)	2.9487	0.4749	3.1751	0.416
	(α=0.02)	2.962	0.4617	3.03	0.4659
OSC pre-filter	(PC=1, tol=99.9%)	3.1551	0.4007	3.0239	0.4304
	(PC=8, tol=99.9%)	3.3747	0.0202	3.3713	0.1575
	(PC=10, tol=95%)	3.3405	0.1715	3.3298	0.2116

Variable Selection

The high dimensionality of spectral data increases the difficulty of using quantitative regression models. Reducing this high dimensionality helps reduce variable numbers, potentially improves the accuracy by removing irrelevant spectral information, and reduces computation time and cost. In this section, two heuristic optimization algorithms, sequential Floating Selection (SFS) and Genetic Algorithm (GA), are applied to variable selection for the spectral data. Sequential Floating Selection (SFS) is a well-known suboptimal search algorithm that is very efficient and effective even for problems of high dimensionality involving non-monotonic Feature Selection [14]. Genetic Algorithm (GA) is a search heuristic that mimics the process of natural evolution [15]. This heuristic is routinely used to generate useful solutions to optimization and search problems.

Performance Comparison - Variable Selection

Table 3: Results of various methods for features (wavelength) selection.

Methods	Variable Used	RMSEcv	Rcv
None	760	3.0359	0.4668
SFS	96	2.9068	0.4979
SFS	709	3.0189	0.4722
GA	273	2.8553	0.5189
GA	178	2.89	0.5126
GA	120	2.7748	0.5483

The findings from previous section were applied to perform these features selection methods. Considering the cross validation results of SFS and GA feature selections in Table 3, the study found that the GA method that selected 120 features (wavelength) provided the smallest RMSEcv and the highest Rcv amongst other results. In addition, the runtime of SFS and GA were also noted. The analysis results must be interpreted with caution, especially for the GA method; it usually considered the entire set of selected features (120 wavelengths) to be used as a whole in the experiment because a feature may only be helpful to prediction when used in conjunction with other features included in an individual feature subset.

Signal Classification

In the classification approach, the sample properties that relate to spectral variations belong to several different groups or classes. As this study focuses on the use of supervised methods, Linear Discriminant analysis, artificial neural network, support vector machine, partial least squares discriminant analysis (PLS-DA), and Monte Carlo PLS-DA (MC-PLS-DA) will therefore be discussed. Linear Discriminant Analysis (LDA) is used to classify analyte concentration based on the spectral data [16]. The idea is to project a dataset onto a lower dimensional space with good class separability to avoid overfitting. An Artificial Neural Network (ANN) is a statistical modelling inspired in the natural neurons and is commonly used in modelling complex relationships among independent variables and dependent variables. ANN is a self-adaptive and data driven modelling technique although it resembles regression analysis but has much more flexibility because it is not restricted by any statistical assumptions or pre-specified algorithms [17]. Support Vector Machine (SVM) is a supervised learning system that uses a hypothesis space of linear functions in a high dimensional feature space, trained with a learning algorithm from optimization theory that has originally been used for classification analysis [18]. With a suitable kernel, SVM can separate in the feature space the data that were non-separable in the original input space. PLS regression is a well-established tool in chemometric analysis [8]. This technique is commonly used in spectral quantitative analysis. Scientific research often involves using variables that are easily (or cheaply) measured to explain or predict the behavior of response variables that are often much

more difficult (or expensive) to acquire. When the factors are many in number and are highly collinear such as in spectroscopy, PLS is a robust method used to construct predictive models. The advantage of the PLS is that the spectral and concentration information are included in the calculation of the factors and the scores. The idea of the MC-PLS-DA [19] is to create a large number of PLS-DA prediction models using different training subsets that are selected randomly from the whole training data set by the Monte Carlo method [20].

The stability of the corresponding coefficients is calculated by using the regression coefficients of these models. The MC-PLS-DA prediction model is then obtained by averaging the PLS-DA models so that the outliers can be removed. In the proposed method, the raw data are randomly divided into two parts – the training set X_t and the prediction set X_p. The prediction set X_p is only used to evaluate the robustness of the prediction model and not for training. Different subsets are randomly selected from the training set X_t to construct a large number of PLS-DA models. The size of each subset is 70% of that of the whole training set X_t. These PLS-DA models are validated by the prediction set X_p. The top 5% of the models are selected to create the final MC-PLS-DA model by averaging the prediction of these models. The top performing model is defined as the one with the highest classification rate. The averaged PLS-DA prediction model is obtained by

$$\overline{B} = \frac{1}{n}\sum_{i=1}^{n} B_i \quad (3)$$

where n is the number of top 5% of the PLS-DA models, and B_(i) is the prediction model of the MC-PLS-DA model.

Performance Evaluation - Signal Classification

In this study, the blood glucose level (BGL) was divided into two classes, with BGL falling between 4mmol/L and 7mmol/L considered as normal BGL, and BGL greater than 7mmol/L defined as high BGL; that is, suffering from DM. The performance of the MC-PLS-DA model can be evaluated by the sensitivity, specificity, and overall accuracy of the model. Sensitivity is a measure of the ability of the classifier to identify normal BGL. Specificity is a measure of the ability to identify high BGL. Accuracy is a measure of the ability to identify the correct BGL. 840 Samples were randomly divided into the training data set and the prediction data set. The former contained 756 samples (90% of the total number of samples) while the latter contained 84 samples. 530 samples were then randomly taken from the training data set to create a PLS-DA model, where the correlation between the predicted and reference values was maximized.

All the samples in the prediction data set were used to fit the training model and evaluate the classification rate. By the same token, another 530 samples were drawn the same subset of training data to create a new PLS-DA model. The procedure was the carried out repeatedly until 1,000 PLS-DA models were built. The top 5% (50) of the models, in terms of classification rate, were selected to

create the MC-PLS-DA model with the same training and prediction data sets. The entire algorithm was then executed repeatedly for a total of 20 runs. At each run, the 840 samples were randomly divided to produce different training and prediction data sets, so that a total of 20 MC PLS DA models were obtained. The sensitivity, specificity, and the accuracy of these 20 MC-PLS-DA models were averaged to evaluate the performance of the algorithm. Based on the findings obtained from the previous experiments, a MC-PLS-DA model was created to predict the BGL for NIR spectral data collected. The number of LVs and PLS-DA models was 6 and 5,000 respectively with reference to [19-25].

The predictor matrix was obtained by averaging the training results. The performance of the MC-PLS-DA model was compared to that of conventional PLS-DA, LDA, NN and SVM. The same training and prediction sets were used for the comparison (Table 4) shows the sensitivity, specificity, and the overall accuracy of the five methods. The results indicate that MC-PLS-DA outperforms the other methods in all three aspects. The overall classification rate of the MC-PLS-DA model reached 75.2%, which was the highest accuracy among all methods. The findings in these experiments suggest that the proposed MC-PLS-DA is a feasible method for classifying NIR spectral data and can be used to monitor the BGL of DM patients through non-invasive measurement. With the enhancement achieved by the Monte Carlo method, the MC-PLS-DA method is more stable and accurate than the conventional PLS-DA method.

Table 4: Sensitivity, specificity and the accuracy of the five methods.

Method	Sensitivity	Specificity	Accuracy
LDA	64.60%	65.30%	64.80%
NN	72.00%	66.50%	69.30%
SVM (RBF)	72.90%	68.60%	70.70%
PLS-DA	72.50%	68.00%	70.20%
MC-PLS-DA	78.00%	72.50%	75.20%

Discussion

This study successfully adopted the four-stage framework of bio-signal processing to handle the NIR spectroscopy that is used to measure blood glucose level. A vigorous systematic approach from signal acquisition to signal interpretation has been established. These include, in the first stage, the assembly of the specific equipment for NIR spectroscopy and suggested the left fourth finger as the measurement site to acquire NIR spectral data. In addition, the study proposed using a 2-norm normalization process, together with the GLS weighting pre-filtering process, during the signal transformation stage so as to minimize the disturbance noise caused by the previous stage. After that, the study suggested using the genetic algorithm for wavelength selection in the variable selection stage. Finally, in the signal interpretation stage, improved methods based on the Monte Carlo approach for the PLS discriminant analysis were introduced, which were called MC-PLS-

DA for classification analysis to predict the result (Figure 1) shows the proposed methods together with the four-stage framework for easy reference. The capability to determine blood glucose non-invasively in humans is a complex and difficult analytical problem. Many researchers and industry groups have attempted to measure blood glucose by various non-invasive methods (Table 5) lists the manuscripts that the experimental setup used NIR spectroscopy with multivariate analysis in testing human subjects. It is relatively simple to measure data and study the correlation with glucose solution tests or blood glucose levels under controlled conditions in research laboratories with only a few subjects/samples. The challenge here is to measure these variables in normal and practical environments with large datasets of different subjects.

Table 5: List of NIR spectroscopy experiments in human tests.

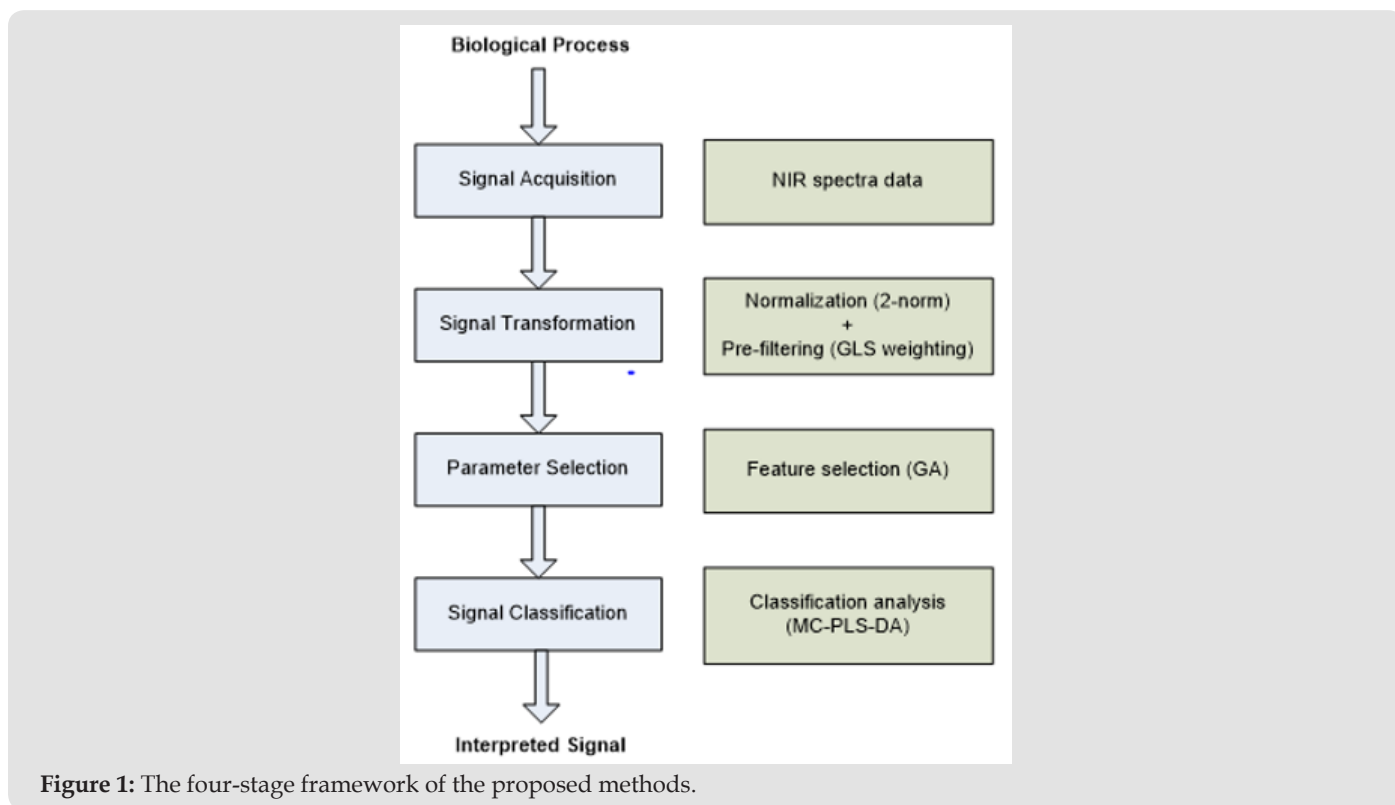| Method	Details	Results	Reference
Absorption of skin (wavelength range from 1,300 – 1,900nm)	Six subjects, OGTT, one calibration model per subject, forearm measurement	Overall correlation coefficient was 0.934, standard error of prediction was 1.32mmol/L, EGA: Zone A: 71.3%, Zone B: 21.3% and Zone D: 7.4%	[21]
Absorption of skin (wavelength range from 1,200 – 1,900nm)	One subject, OGTT, skin measurement	Correlation coefficient was 0.928, standard error of prediction was 1.79mmol/L, EGA: Zone A: 87.5%, Zone B: 8.3% and Zone D: 4.2%	[22]
Absorption of skin (wavelength range from 1,212 – 1,889nm)	One subject, OGTT, finger measurement	Correlation coefficient was 0.89, RMSEV was 1.12mmol/L, EGA: Zone A: 86%, Zone B: 7% and Zone D: 7%	[23]
Absorption of skin (wavelength range from 900 – 1,700nm)	23 subjects, OGTT, fingertip measurement	EGA: Zone A: 93.7%, Zone B: 6.3% and Zone C-E: 0%	[24]
Absorption of skin (wavelength range from 1,550 – 1,800nm)	One subject, 4 days finger measurement	Correlation coefficient was 0.8785, RMSEp was 0.8682mmol/L, EGA: 100% fall within Zone A & B	[25]

Figure 1: The four-stage framework of the proposed methods.

Method

a) Details Results Reference

The proposed multivariate analysis by making use of classification analysis for non-invasive blood glucose monitoring is the new initiative for the purpose of DM screening. The prediction output in this method is not represented in terms of exact value of the BGL, simply to classify it into normal or high BGL. The MC-PLS-DA integrates the Monte Carlo method into the conventional PLS-DA to enhance performance. It exhibits better sensitivity, specificity, and overall accuracy rate when compared to others. The classification output results for the relationship between the response and the independent variables is more accurate, hence enhancing the reliability of the classification model, especially for such a large training set in this study. These advantages make MC-PLS-DA a possible way for non-invasive blood glucose measurement using NIR spectroscopy.

Conclusion

In this century, particularly in wealthy developed countries, diabetes is a complex group of syndromes that have a disturbance in the body's use of glucose. It can be controlled by an appropriate regimen that includes patient education, weight management, diet control, sensible exercise, oral medication, and insulin therapy. However, these diabetes management and medications rely heavily on blood glucose measurement. With ever-improving advances in diagnostic technology, the race for the next generation of bloodless, painless, accurate and consistent blood glucose measurement instruments has begun. Nevertheless, many hurdles remain before these products reach commercial markets. Considerable progress has been made in the development of non-invasive blood glucose monitoring devices. In principle, the approach can be used for screening purposes for DM prevention. However, for diabetes that needs frequent testing, using invasive blood glucose measurement via finger pricking remains a practical way to provide suitable information for diabetes management at this current stage.

References

1. (2011) International Diabetes Federation. Guideline for the management of postmeal glucose in Diabetes. International Diabetes Federation, Brussels, Belgium.

2. (2016) American Diabetes Association. Standards of medical care in diabetes. Diabetes Care 39(1): 1-112.

3. So CF, Choi Kup Sze, Wong TKS, Chung JWY (2012) Recent advances in noninvasive glucose monitoring. Medical Devices: Evidence and Research 5: 45-52.

4. Van Bemmel J, Musen M, Helder J (1997) Handbook of Medical Informatics. Germany: Springer-Verlag.

5. Suli E, Mayers D (2003) An introduction to numerical analysis. Cambridge. Cambridge University Press. New York.

6. Fearn T, Riccioli C, Garrido Varo A, Guerreo Ginel J (2009) On the geometry of SNV and MSC. Chemometrics and Intelligent Laboratory Systems 96(1): 22-26.

7. Naes T, Isaksson T, Fearn T, Davies T (2002) A user-friendly guide to multivariate calibration and classification. Chichester UK: NIR Publications.

8. Hoskuldsson A (1988) PLS regression methods. Journal of Chemometrics.

9. Skibsted E, Boelens H, Westerhuis J, Witte D, Smilde A (2004) New indicator for optimal reprocessing and wavelength selection of near-infrared spectra. Applied Spectroscopy 58(3): 264-271.

10. Rinnan A, Berg F, Engelsen S (2009) Review of the most common pre-processing techniques for near-infrared spectra. Trends in Analytical Chemistry 28(2): 1201-1222.

11. Orfanidis S (1996) Introduction to signal processing. Englewood Cliffs NJ: Prentice Hall.

12. Savitzky A, Golay M (1964) Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry 36(8): 1627-1639.

13. Wold S, Antti H, Lindgren F, Ohman J (1998) Orthogonal signal correction of near-infrared spectra. Chemometrics and Intelligent Laboratory 44(1): 175-185.

14. So CF, Choi Kup Sze, Chung JWY, Wong TKS (2013) Modified sequential floating selection for blood glucose monitoring using near infrared spectral data. Journal of Applied Spectroscopy 80(1): 291-294.

15. Sivanandam S, Deepa S (2008) Introduction to genetic algorithms. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.

16. Huberty C, Olejnik S (2006) Applied MANOVA and discriminant analysis (2nd edn.). Hoboken NJ: John Wiley & Sons.

17. Rumelhart D, Mcclelland J (1986) Parallel distributed processing: explorations in the microstructure of cognition. Volume 1. Foundations. Cambridge MA: MIT Press.

18. Kecman V (2001) Learning and soft computing: support vector machines, neural networks, and fuzzy logic models. Mass, Cambridge: MIT Press.

19. So CF, Choi KupSze, Chung JWY, Wong TKS (2013) An extension to the discriminant analysis of near-infrared spectra. Medical Engineering & Physics 35(2): 172-177.

20. Dunn W, Shultis J (2012) Exploring Monte Carlo methods. Amsterdam. Boston: Elsevier.

21. Maruo K, Tsurugi M, Tamura M, Ozaki Y (2003) *In vivo* noninvasive measurement of blood glucose by near-infrared diffuse-reflectance spectroscopy. Applied Spectroscopy 57(10): 1236-1244.

22. Maruo K, Tsurugi M, Chin J, Ota T, Arimoto H, et al. (2003) Noninvasive blood glucose assay using a newly developed near-infrared system. Journal of Selected Topics in Quantum Electronics pp. 322-330.

23. Kasemsumran S, Du Y, Maruo K, Ozaki Y (2006) Improvement of partial least squares models for *in vitro* and *in vivo* glucose quantifications by using near-infrared spectroscopy and searching combination moving window partial least squares. Chemometrics and Intelligent Laboratory Systems pp. 97-103.

24. Yamakoshi Y, Ogawa M, Tamakoshi T, Satoh M, Nogawa M, et al. (2007) A new non-invasive method for measuring blood glucose using instantaneous differential near infrared spectrophotometry. Proceedings of the 29th Annual International Conference of the IEEE EMBS pp. 2964-2967.

25. Chuah Z, Paramesran R, Thambiratnam K, Poh S (2010) A two-level partial least squares system for non-invasive blood glucose. Chemometrics and Intelligent Laboratory Systems 104: 347-351.

Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles

https://biomedres.us/