Non-Invasive Blood Glucose Estimation using Handheld Near Infra-Red Device

Mahmud Iwan Solihin, Yaameen Shameem, Thein Htut, Chun Kit Ang
Muzaiyanah bt Hidayab

Abstract: Non-invasive blood glucose measurement would ease everyday life of diabetic patients and may cut the cost involved in their treatments. This project aims at developing a non-invasive blood glucose measurement using NIR (near infrared) spectroscopic device. NIR spectra data and blood glucose levels were collected from 45 participants, resulting 90 samples (75 samples for calibration and 15 samples for testing) in this project. These samples were then used to develop a predictive model using Interval Partial Least Square (IPLS) regression method. The results obtained from this project indicate that the handheld micro NIR has potential use for rapid non-invasive blood glucose monitoring. The coefficient of determination ($R^2$) obtained for calibration/training and testing dataset are respectively 0.9 and 0.91.

Index Terms: blood glucose, near infrared spectroscopy, non-invasive, handheld spectrometer.

I. INTRODUCTION

Diabetes is a non-curable disease which causes the body to lose its ability to maintain blood glucose at healthy levels. Hence, diabetic patients must manage their blood glucose levels by external means, which include regular blood glucose measurement. All the current methods that are used to measure blood glucose today are all invasive processes, which rely on extracting blood by piercing the body and causes pain to the patient [1].

Absorption spectra of chemical species (atoms, molecules, or ions) are generated when a beam of electromagnetic energy (i.e. light) is passed through a sample, and the chemical species absorbs a portion of the photons of electromagnetic energy passing through the sample. Lambert's beer law states that the absorbptive capacity of a dissolved substance is directly proportional to its concentration in a solution. The relationship can be expressed as shown in Eq. (1) [2].

$$A = \log_{10}\left(\frac{I}{I_0}\right) = \varepsilon lc$$

where: $A$ = absorbance
$\varepsilon$ = the molar extinction
$I$ = length of the path light must travel in the solution in centimeters
$c$ = concentration of a given solution

Spectroscopy has been widely applied nowadays into many fields as non-invasive/non-destructive quality assessment of substance [3][4][5]. Particularly, Near Infra-Red (NIR) spectroscopy can be used non-invasive quantitative measurement of the chemical properties of a material using this property stated in lamberts beer law. Several successful researches have been carried out to measure blood glucose non-invasively using NIR spectroscopy. Chemometric methods such as Partial Least Squares (PLS) regression and Principle Component Analysis (PCA) were used in these studies to find a linear relationship between blood glucose levels and NIR spectra [6]. However, almost all these studies used relatively large and seemingly un-practical spectrometers. This project aimed at developing a predictive model which can measure blood glucose levels by using a practical handheld micro NIR spectroscopic device, i.e. micro NIR spectrometer.

II. DEVELOPMENT OF THE PROJECT

Samples of NIR spectra and blood glucose levels were collected from 45 participants for this project. NIR spectra data that were collected using micro NIR spectroscopic instrument with wavelength range from 900-1700nm. Some participants have taken the data for more than once or twice at different time so that there are overall 90 data samples. Fig.1 shows the micro NIR instrument which is about hand palm size.

Figure 1. The handheld micro NIR device
The NIR spectra of the participants were recorded by scanning to their point fingers respectively. Blood glucose levels were measured using an ACCU_CHEK® Active Blood Glucose Monitor by Hoffmann-La Roche Ltd.

The retrieved data was recorded containing all the data from 45 participants (90 samples). The raw NIR spectra of their fingertips is shown in Fig. 2. To be noted that NIR absorbance signal is used in this study. PLS regression methods were then used to find a linear regression model which gave a linear relationship between the multivariate NIR data and the blood glucose levels. MATLAB® simulation software was used for generating PLS calibration models and testing these model’s accuracy.

Preprocessing methods such as Smoothing, Multiplicative Scatter Correction (MSC), Standard Normal Variate (SNV) Normalization and Derivation is normally used to remove physical additives from the spectra before calibration [7].

In this project, Extended (EMSC) preprocessing was applied into the raw spectra data. EMSC is preprocessing method that allows a separation of physical light-scattering effects from chemical (vibrational) light absorbance effects in spectra [8]. This model-based method is particularly useful in minimizing wavelength-dependent light scattering variation. After pre-treatment the corrected spectra become insensitive to light scattering variations and responds linearly to the analyte concentration. Through EMSC parameter estimation, an EMSC-corrected spectrum can be obtained, with only chemical absorbance part left after removal of baseline offset and wavelength-dependent variations [9]. The overall preprocessing stage used for generating calibration model is shown in Table I. The NIR spectra data after preprocessing is shown in Fig. 3.

Table I. The pre-processing for raw data

<table>
<thead>
<tr>
<th>Order of Preprocessing</th>
<th>Preprocessing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage-1</td>
<td>EMSC</td>
</tr>
<tr>
<td>Stage-2</td>
<td>Average Smoothing (windows=5)</td>
</tr>
<tr>
<td>Stage-3</td>
<td>Kennard–Stone dataset selection</td>
</tr>
</tbody>
</table>

Figure 2. All 90 raw NIR spectra data obtained by scanning fingertips

Variant PLS methods were used to increase the accuracy of the calibration methods that were generated. Such variations included Interval PLS (IPLS) and Back Interval PLS (BIPLS). These methods look at intervals within the spectrum to find the intervals with maximum co-variance [10]. IPLS is used in this study to make prediction of the blood glucose level from the obtained NIR spectra data of the participants.

III. RESULTS AND DISCUSSIONS

A. Calibration Process

Calibration of the model was carried by finding 10 PLS components and cross validating the results by the Venetian Blind method to find the RMSECV (Root Mean Squared Error of Cross Validation) of the calibration model. There are 75 data used for calibration, while 15 data is used to test the prediction model.

An acceptable calibration model was generated using calibration data set with 2 PLS components as shown in Fig. 4. Maximum regression and minimum RMSECV (about 0.37) were obtained effectively using IPLS on a global model. Adding more PLS components does not improve the resulting RMSECV as shown in Figure 4. The calibration result is shown in Fig. 5 with $R^2=0.9$ indicating a good fit of the data.

B. Prediction Testing

The calibration model that was obtained with the lowest RMSECV was used to predict blood glucose level that were not used during the calibration phase, i.e. testing dataset. After prediction the RMSEP (RMSE prediction) by Eq. (2) and $R^2$ coefficient of the predicted variables was analyzed to assess the accuracy of the prediction model, the results are shown in Fig. 6.

The prediction results obtained an acceptable $R^2$ coefficient. The RMSEP was quite small and consistent with that in calibration phase, meaning the deviation between the predicted and measured results were relatively small.
\[ \text{RMSE} = \sqrt{\frac{\sum (Y_{\text{pred}} - Y_{\text{test}})^2}{N}} \]  

where: \( Y_{\text{pred}} \) = Predicted Y variable  
\( Y_{\text{test}} \) = Y test variable  
\( N \) = Number of variables

Figure 4. RMSECV versus PLS components for global model

![Figure 4. RMSECV versus PLS components for global model](image)

Another way of measuring the accuracy of a glucometer is given by plotting the measured values against the predicted values on the Clarke Error Grid (CEG). The CEG was created by medical professionals as a clinical practice guide to grade the error of glucose measurement devices. The grid is divided into 5 areas which correspond to different clinical judgement errors that might occur due to using the glucose measurement device [6]. Fig. 7 shows the CEG plot for the optimum result as discussed earlier in this study.

The areas A and B on the Clarke Error Grid correspond to acceptable glucose measurement accuracy in terms of medical decision that will be made based on the results [11]. Since the prediction model’s results fall within these areas (black dotted marks), the model can be considered accurate. Furthermore, the accuracy of the prediction model can be increased by collecting more NIR spectral data in a controlled environment, research has shown that temperature and light conditions greatly affect the quality of NIR spectral data [12]. The variance in blood glucose data can also be increased to get a better prediction model [13].

![Figure 5. Calibration results using IPLS with R^2=0.9](image)

![Figure 6. Prediction results using test dataset (R^2=0.91)](image)

![Figure 7. Calibration results error grid analysis](image)

The overall prediction results for both calibration and testing data sets is shown in Table II.

<table>
<thead>
<tr>
<th>Data</th>
<th>RMSE</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibration dataset</td>
<td>0.29</td>
<td>0.90</td>
</tr>
<tr>
<td>Testing dataset</td>
<td>0.21</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Table II. Prediction results using IPLS for both data sets
IV. CONCLUSIONS

Diabetes is a disease that causes blood glucose levels in the blood to rise, which is caused due to the body’s inability to produce the glucose control hormone insulin. There is no cure for diabetes, and once a person develops the disease, the person has to rely on external means to manage the body’s blood glucose levels for the rest of his/her life. Blood glucose management is a crucial task for diabetic patients, which currently involves blood glucose measurement by invasive means.

Several researches have been done on the ability of using NIR spectroscopy to predict blood glucose levels. This project aimed at developing a noninvasive predictive model using a handheld micro NIR device. A predictive model was worked upon using Interval Partial Least Squares (IPLS) regression methods, and the obtained calibration models were tested. These tests provided evidence that there is a linear relationship between NIR spectral data retrieved from the fingertip and the blood glucose concentration. This proves that the handheld spectrometer can be used as a potential tool to measure blood glucose non-invasively.

The results indicate that the calibration and prediction model the handheld micro NIR has potential for rapid non-invasive blood glucose monitoring with coefficient of determination ($R^2$) obtained for calibration and testing dataset are respectively 0.9 and 0.91. Further improvement can be done on: increasing accuracy especially on calibration data and testing the device on-site in the real situation. Further improvement will be done on this project such as adding more samples involving lower and higher variations of blood glucose level and hardware software integration for on-site testing of the calibrated instrument.

ACKNOWLEDGMENT

The authors would like to express the gratitude to Centre of Excellence for Research, Value Innovation and Entrepreneurship (CERVIE) – UCSI University for supporting this project and would also like to thank all the participants who volunteered to give blood samples for this project.

REFERENCES


AUTHORS PROFILE

Mahmud Iwan Solihin is an assistant professor in mechanical/mechatronics engineering department specialized in applications of artificial intelligence/machine learning, optimizations, control systems, robotics and automation. In the past, he was a research assistant for an eScienceFund project under MOSTI (Ministry of Science Technology & Innovation). His current research interest is mainly on applications of machine learning and optimization algorithms in various filed including agriculture, medical field, sensor networks, etc., toward sustainable engineering development in the future.

Ang Chun Kit is currently the dean of Faculty of Engineering and Built Environment, UCSI University, Kuala Lumpur. His research interest includes robotics and artificial intelligence applications in various field including medical and agricultural sectors.

Thein Htet was an international student from Myanmar who studied Mechatronic Engineering in UCSI University, Malaysia. He is currently working on projects for Internet of Things (IoT) industry as software engineer.

Yaameen Shameem was an international student who studied Mechatronic Engineering in UCSI University, Malaysia. He is currently working at electrical power company Sielco (Maldives) as an assistant manager.