

DEVELOPING WEARABLE BLOOD GLUCOSE METER USING RAMAN SPECTROSCOPY TECHNIQUE

by

Podchanan Rungthirakul

A Thesis Submitted in Partial Fulfillment of the Requirements for the
Degree of Master of Engineering in
Data Science and Artificial Intelligence

Examination Committee: Prof. Chaklam Silpasuwanchai (Chairperson)
Prof. Chantri Polprasert
Prof. Attaphongse Taparugssanagorn

Nationality: Thai

Previous Degree: Bachelor of Science in Industrial Chemistry
King Mongkut's University of Technology North Bangkok
Thailand

Scholarship Donor: AIT SCHOLARSHIPS

Asian Institute of Technology
School of Engineering and Technology
Thailand
Dec 2023

AUTHOR'S DECLARATION

I, Podchanan, declare that the research work carried out for this thesis was in accordance with the regulations of the Asian Institute of Technology. The work presented in it are my own and has been generated by me as the result of my own original research, and if external sources were used, such sources have been cited. It is original and has not been submitted to any other institution to obtain another degree or qualification. This is a true copy of the thesis including final revisions.

Date: Sep 20, 2023

Name: Podchanan

Signature: Podchanan R.

ABSTRACT

Continuous glucose monitoring (CGM) systems have been highlighted as an important component of optimal glycemic control in diabetic patients (Lee, Probst, Klonoff, & Sode, 2021). The commercial CGM technique currently includes the implantation of a sensor (Keenan, Mastrototaro, Voskanyan, & Steil, 2009). Its minimally intrusive nature prevents glucose monitoring from becoming ubiquitous. Raman spectroscopy, for example, has been investigated as a non-invasive method of measuring hyperglycemia in vivo. Wearable (constant, non-invasive, ubiquitous) self-monitoring blood glucose (SMBG) is therefore achievable. However, the development of wearable Raman-based SMBGs has received little attention. This development is fraught with difficulties. First, the best measurement locations to directly evaluate glucose scattering (wrist, forearm, nail fold, fingertip, and thenar) are unknown. Although previous research demonstrated excellent accuracy of glucose prediction from all five sites, this high accuracy may be due to the utilization of the whole spectrum, which may contain undesired signal aberrations that correlate with hyperglycemia (Kang et al., 2020). Second, we ran multiple trials from the previous study at measuring sites, but the resultant spectrum data was confusing. We went back to the foundation experiment, the calibration data from liquid samples, to overcome this difficulty and ensure that the experiment we devised was compatible with our Raman instrument. Third, appropriate features (engineering) are underutilized. Although previous work offered feature engineering strategies such as principal component analysis (Li et al., 2019) or protein/hemoglobin normalization (Kang et al., 2020; Shao et al., 2012), a lack of formal comparison makes understanding what works challenging. Along with that, a wearable SMBG prototype is being built and tested.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background of the Study

Diabetes is a metabolic disease that manifests as high blood glucose levels, leading to the degeneration of vital tissues and organ systems, such as the eyes, kidneys, and heart (H.Wu et al.,2022)(J.Wei et al.,2022). Monitoring blood glucose concentration levels in clinical therapy is one of the most effective approaches to postpone or avoid diabetic complications. The traditional finger-pick testing is invasive, and painful, and this method will increase the risk of infection (L.Tang et al.,2020).

From the foregoing, there is a great demand for non-invasive, sensitive, robust, and continuous blood glucose monitoring methods. Several studies have proposed Raman spectroscopy as a non-invasive method for detecting blood glucose concentration (N.Li et al.,2019)(A.Pors et al.,2023). According to the complexity of the raw spectrum acquired by Raman spectroscopy detection (e.g. fluorescence artifacts, spectral data from sample slide), chemometric techniques are necessary to extract the most relevant dataset from complicated Raman spectral data. However, this is typically mitigated by altering the laser intensity or measuring times.

Shao et al. (2012) demonstrated a high correlation ($R^2 = 0.91$) on Raman spectra between glucose solution content and ISF measured at mouse ear. Kang et al. (2020) developed a novel method for extracting glucose scattering by subtracting two Raman signals from two separate time intervals as a direct assessment of glucose in blood. Furthermore, they confirmed that the glucose peak (1125cm^{-1}) should be balanced with the protein and lipid peak (1450cm^{-1}) in order to correctly quantify the glucose concentration in blood. Using Raman spectra from a pig's ear, they established an $R^2 = 0.91$ correlation between actual and projected glycemic levels.

A further important factor is the location of the measurement. As prospective measuring sites, the forearm (Enejder et al., 2005; Scholtes-Timmerman, Bijlsma, Fokkert, Slingerland, and Veen, 2014), thenar (Lundsgaard-Nielsen and al., 2018), and nail fold (Li et al., 2019) have been chosen. While González Viveros et al. (2022) found that the forearm is the most effective location when compared to the wrist and index fin-

ger, it is unclear which site is the best owing to differences in equipment, parameters, and technique (e.g., how to preprocess) amongst articles. To provide a superior non-invasive methodology for continuous glucose monitoring in blood, we must overcome the interferences of complexed spectrum data and experimental design limitations (e.g. measurement method and Raman Spectroscopy's spec).

This study will expand on past work by (1) verifying the use of Raman scatterings for measuring blood glucose and (2) calibrating spectral data of blood, glucose solution, and skin (prospective measurement sites). All of this lays the groundwork for the development of the first wearable (continuous, non-invasive, ubiquitous) Raman-based self-monitoring blood glucose (SMBG) device, particularly for everyday users and intended for wider usage.

1.2 Statement of the Problem

The effective measurement locations are yet unknown. It is difficult to compare prior findings due to methodological variations. Furthermore, previous research had problematic technique, i.e., it may be stated that the high accuracy of glucose prediction is due to the utilization of the complete spectrum, which may contain unintentionally linked signal aberrations with hyperglycemia (Kang et al., 2020). There have been five successful measurements on the human body. However, the same approach was used to compare only the wrist, fingertip, and forearm (González Viveros et al., 2022). The nail fold is a suitable measurement site because the laser may easily enter the microvessel in the dermis due to its thin epidermal layer (Li et al., 2019). Thenar has been selected as a measurement location for a portable Raman-based SMBG instrument (Lundsgaard-Nielsen et al., 2018). Using the direct measurement of glucose technology, this study tries to determine the greatest correlation between glucose fingerprint and glycemic control (Kang et al., 2020).

Proper feature engineering is still being researched. Previous studies, in particular, relied heavily on two distinct methodologies. The first method employs statistical analysis on full spectrum data, together with extra preprocessing such as dimension reduction or feature significance (Enejder et al., 2005; González Viveros et al., 2022; Li et al., 2019; Scholtes-Timmerman et al., 2014). The second way is to normalize the spectrum using either the hemoglobin peak (1549 cm^{-1}) (Shao et al., 2012) or the protein and lipid peak (1450 cm^{-1}) (Kang et al., 2020) and then handpick a smaller group of spectra for

modeling (Kang et al., 2020; Shao et al., 2012). The purpose of this study is to compare these two methodological techniques on human beings.

There has been no effort on the creation of wearable Raman-based SMBG. There are various hurdles to developing wearables. In terms of measurement sites, the wrist is the best option since it can be easily integrated in a commonly used platform (smart band/watch). This option has the most pervasiveness but the lowest accuracy. A similar scenario exists in modeling. A more sophisticated model may match the data more correctly in exchange for efficiency. This effort aims to construct a wearable Raman-based SMBG for everyday usage in order to obtain CGM and to assess the amount of offerings required to reach this goal.

1.3 Objectives

This work has been divided into four studies.

1.3.1 Study 1: Measuring spectral data of solution

Objective: To validate spectral data from blood and glucose solutions at different concentrations.

Independent Variables:

- (a) Distilled water
- (b) Glucose solution
- (c) Blood
- (d) Blood with OGTT
- (e) Background noises (e.g. sample slide, air)

Dependent Variables: The most accurate direct measurement of glucose spectra (1125cm^{-1})

Outcome: The most accurate direct measurement of interested compounds spectra.

1.3.2 Study 2: Measuring spectral data of skin

Objective: To study the measuring sites

Independent Variables: Measuring sites are divided as below

- (a) Wrist
- (b) Forearm
- (c) Index fingertip
- (d) Index nail fold
- (e) Thenar

Dependent Variables: The most accurate direct measurement of glucose spectra (1125cm^{-1})

Outcome: Ranking of measuring sites

1.3.3 Study 3: Feature Engineering

Objective: Evaluate the impact of feature engineering and selection.

Independent Variables:

1. Feature Engineering technique
 - (a) No engineering
 - (b) PCA (baseline)
 - (c) Normalization with protein and lipid peak (1450cm^{-1})
 - (d) Normalization with hemoglobin peak (1549cm^{-1})

2. Feature selection and modeling

- (a) Full spectrum + PLS (baseline)
 - (b) single 1125cm^{-1} peak + LR
 - (c) 911, 1060, 1125cm^{-1} peak + MLR

Dependent Variables: Glycemic spectra

Outcome: The influence of feature engineering, selection, and modeling on predictive accuracy.

1.3.4 Study 4: Designing and developing wearable Raman-based SMBG

Objective: Design and develop a prototype of a wearable Raman-based SMBG.

Outcome: A prototype

1.3.5 Study 5: Device Evaluation

Objective: To evaluate the prototype by repeat section 1.3.3 experiment with the prototype.

Dependent Variable: Glycemic

Outcome: Prototype achieves glycemic prediction correlation $R^2 > 0.8$ with actual glycemic.

1.4 Organization of the Study

The document is structured as follows. Chapter 2 is a Literature Review, while Chapter 3 is a Methodology.

CHAPTER 2

LITERATURE REVIEW

The Raman spectra of amorphous glucose (glucose fingerprint) when in combination with water (glucose solution) and blood (blood glucose) are examined. Following that, we will look at measurement sites, preprocessing strategies, and data modeling.

2.1 Glucose Fingerprint

When glucose has been included in the analyte, the "glucose fingerprint" shows as a distinctive Raman scattering spectrum. Peaks in Raman scattering in glucose solution (glucose and water) are 796, 1060, 1125, and 1366 cm^{-1} respectively (Figure 2.1). These peaks rise as glucose concentration rises, demonstrating a quantitative characteristic of Raman spectroscopy (Shao et al., 2012). When measuring Raman scattering of ISF, the glucose fingerprint is also visible in a more complex analyte such as blood. (González Viveros et al., 2022; Kang et al., 2020; Scholtes-Timmerman et al., 2014) as show in Figure 2.4. Kang et al. (2020) extract the glucose fingerprint, which contains a peak at 1125 cm^{-1} , from two noisy Raman signals. As a result, Raman spectroscopy can identify blood glucose.

2.2 Measuring sites

According to the thickness of human skin changes from place to place, Raman scattering measurements done at different locations may yield different findings (González Viveros

Figure 2.1

Raman spectra of glucose solution at different concentration (?).

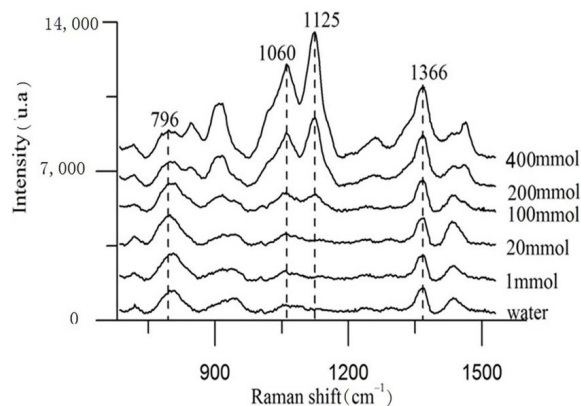


Table 2.1

Assignments of Raman peaks that are identified in the spectra of the microvessels and blood (????)

Peak Position (cm ⁻¹)		Assignments	Components
Microvessels	Blood		
650	643	P:C-S str	Ascorbic acid
758	752	ν_{15}	Trp
837	827	γ_{10}	Fructose
858	855	$\nu(C - C)$	Tyr, lac
885	-	-	-
902	898	p:C-C skeletal	Tyr
945	940	$\nu(C - C)$	Crtic acid
978	971	p: Skeletal vibr	Fibrin
1004	1004	ν -ring	Phe
1027	1026	$\delta(= C_b H_2)$ asym	Lac
1130	1129	ν_5 ,	Lac
1163	1157	ν_{44}	Heme
1217	1212	$\nu_5 + \nu_{18}$	Heme
1320	1321	p: CH ₂ twist	Try
1332	1341	ν_{41}	Trp
1424	1423	ν_{28}	Acetates
1448	1450	$\delta(= CH_2/CH_3)$	Trp
1551	1546	ν_{11}	Heme
1608	1603	$\nu(C = C)_{\text{venyl}}$	Heme
1660	1653	Amide I	Heme

Figure 2.2

A trace of glucose fingerprint in Raman scattering of in vivo blood (?).

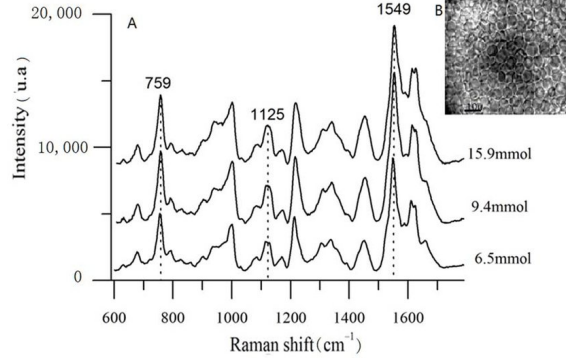
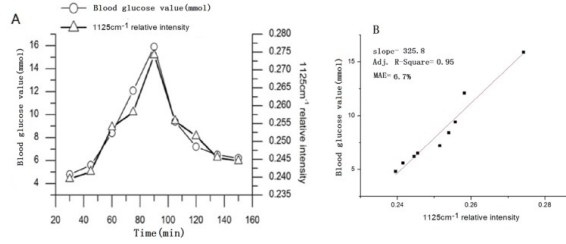


Figure 2.3

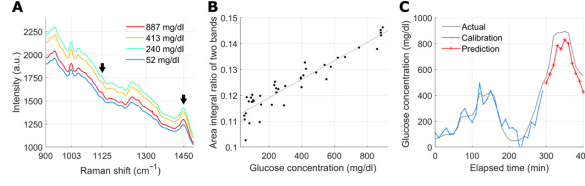
(A) Blood glucose value with 1125 cm^{-1} relative intensity. (B) Concentration-dependent Raman relative intensities of glucose (1125 cm^{-1}) (?)



et al., 2022; Li et al., 2019). If the laser is focused on the blood vessels, the glucose fingerprint signal will be greater (Shao et al., 2012). The stratum corneum (SC), epidermis, and dermis are the three layers of human skin (Li et al., 2019). When detecting glucose in ISF rather than blood, the following issues may arise: (1) ISF glucose lags after blood glucose (Cengiz & Tamborlane, 2009; Steil et al., 2005), and (2) the concentration of glucose in ISF is much lower than blood (O’Kane, 2012). From the foregoing, the signal from glucose Raman scattering is weak (Li et al., 2019). Considering a better technique, it would be to measure the Raman signal in locations where the laser can better penetrate the dermis, because the dermis contains microvessels (Cutolo, Grassi, & Matucci Cerinic, 2003; Ingegnoli, Smith, Sulli, & Cutolo, 2018). ? finds the greater $R^2 = 0.98$ prediction correlation when nailfold is used as a measuring site. The result is better than the forearm ($R^2 = 0.83$) (??) and fingertip ($R^2 = 0.91$) (?). ? measures Raman glucose at forearm, wrist, and index fingertip and achieves the root-mean-square error (RMSE) of 56.31 ± 4.28 , 58.22 ± 1.03 , and 56.65 ± 8.99 ml/dL respectively.

Figure 2.4

(B) Showing the linear relationship between normalized 1125 cm^{-1} with blood glucose ($R^2 = 0.94$) (?)



2.3 Preprocessing techniques and Data modeling

Preprocessing is largely divided into two categories. Using either the extracted features or the whole Raman spectrum as an input. The complete spectrum analysis with partial least squares (PLS) is a popular method. (????). The results vary from $R^2 = 0.62$ (?), $R^2 = 0.83$ (??), and RMSE of around 65 ± 0.4 ml/dL in ?. The preprocessing can be done by hand-pick (??) or automatically (??) pick the features. Hand-pick features were performed by normalizing the spectra with either protein and lipid peaks (1450 cm^{-1}) (?) or hemoglobin peaks (1549 cm^{-1}) (?). It is demonstrated that the normalized 1125 cm^{-1} has a linear correlation with in vivo blood glucose at $R^2 = 0.95$ (?). In ?, normalized 911 , 1060 , 1125 cm^{-1} with 1450 cm^{-1} were utilized as inputs of the more complex multiple linear regression (MLR) which achieved $R^2 = 0.91$.

CHAPTER 3

METHODOLOGY

3.1 Study 1: Confirming the parameters

The goal of this study is to determine how measurement locations and schemes impact Raman spectra. The findings of this study will be utilized to plan the experiment in section. ??.

3.1.1 Equipment

The Raman instrument, which includes a 785 nm laser and a 10x objective lens, will be utilized to evaluate the Raman spectra. The Accu-Chek® Guide Meter is a standard blood glucose meter for SMBG (?) and will be used to check glycemic control.

3.1.2 Studying Measuring schemes

We have previously performed the measurements on the nailfold and index fingertip, however the findings are still confusing due to variances in the specifications of our Raman instruments. To avoid issues caused by ambiguous spectrum data, we must first check the spectral from a liquid sample (e.g., glucose solution, blood) before evaluating the Raman spectroscopy at four different locations: the wrist, forearm, index fingertip, and index nailfold. The several measurement systems will be put to the test to discover which one gives the greatest scattering signal while avoiding fluorescence interference. The best measurement techniques will be utilized for the remainder of the Raman scattering evaluation.

3.1.3 Data collection

The solution's spectral data will be used to measure the concentration of glucose in distilled water and blood. The Oral Glucose Concentration will regulate the glucose concentration in blood.

- (1) Distilled water
- (2) Glucose solution (diluted in distilled water) 70, 80, 90, 100 mg/dL for imitating normal fasting blood glucose condition.
- (3) Glucose solution (diluted in distilled water) 110, 120, 130 mg/dL for imitating fasting blood glucose fluctuations. These concentrations indicate a proclivity for diabetes.
- (4) Blood from participants in normal fasting blood glucose condition. (participants will

be restricted to fasting for eight hours.)

(5) Blood from individuals in the OGTT-controlled fasting blood glucose fluctuations condition

The healthy volunteers' glucose levels will be manipulated using the Oral Glucose Tolerance Test (OGTT). Participants must fast for at least eight hours before to the test. After arriving at the experiment location, participants will be asked to acclimate for 30 minutes. The instructor will repeat the experiment technique during the acclimatization period. Data will be collected at the desired measurement point using both a Raman Instrument and conventional SMBG equipment. When the acclimatization period is over, the first sample will be taken. The OGTT begins with the subjects consuming a 250 ml of water containing 75 g of glucose in five minutes. Then, over the following two hours,

(1) take Raman spectra from the interesting measuring site every five minutes

(2) collect blood samples from the interested measuring site every 20 minutes.

There will be 25 Raman samples and eight blood samples in total.

Each participant has to repeat the experiment until all four measuring sites are measured. The experiment has to be done on another day.

3.1.4 Metric

The glucose spectra will be extracted by subtracting two Raman signals. The remaining signal is the change in glucose concentration, as shown in the derivation below.

The Raman spectra (RS) contains glucose fingerprint (G) and tissue spectra (T).

$$RS = G + T \quad (3.1)$$

Then, the subtraction of two Raman signal can be represented as follows;

$$\Delta RS = RS_1 - RS_2 \quad (3.2)$$

where RS_i is the Raman spectra measured at time i . Then, substituting Equation ?? with Equation ?? will derive the follows

$$\Delta RS = \Delta G + \Delta T \quad (3.3)$$

Given the same measurement site, the tissue spectra will be the same. Thus, ΔT is 0. Then, it is obvious that

$$\Delta RS = \Delta G \quad (3.4)$$

As a result, the optimum measurement site is the one with the highest correlation between ΔG and real blood glucose fluctuations.

3.2 Study 2: Raman scattering analysis of blood glucose

The objective of this study is to create a model of the correlation between Raman spectra and blood glucose levels. The same equipment from section ?? will be used. The measuring site and scheme are chosen based on the result of Section ??.

3.3 Data Collection

The identical data collecting and experiment approach as in section ?? will be employed. In this study, we raise the number of participants to 15. Participants must be between the ages of 20 and 35, 36 to 50, and 51 to 65. A total of $15 \times 25 = 375$ Raman samples and $15 \times 8 = 120$ blood samples are collected.

3.4 Preprocessing and Data Modeling

The acquired Raman spectra will be normalized by their intensity of 1450 or 1549 cm^{-1} . Thus, there are three preprocessing options: (1) without normalization, (2) normalize with 1450 cm^{-1} , and (3) normalized with 1549 cm^{-1} . The Linear Regression (LR) model will be used to assess the linearity of 1125 cm^{-1} with glucose concentration. The MLR model with $911, 1060, 1125 \text{ cm}^{-1}$ as inputs we will use for multivariate analysis. As a baseline model, full spectrum with PLS will also be employed.

In total, there will be $3 \times 3 = 9$ combinations to compare.

3.5 Metric

The Pearson correlation will be used to evaluate the model's performance. Furthermore, resource use during prediction will be monitored.

Figure 3.1

An example of wearable (?).



3.6 Study 3: Designing and developing wearable blood glucose device

Objective: Design and develop a prototype of a wearable SMBG.

Outcome: A prototype.

3.7 Study 4: Device Evaluation

Objective: To evaluate the prototype, we redo Section ?? experiment with our prototype.

Independent Variables: Raman scattering of blood

Dependent Variables: Glycemic

Outcome: Prototype achieves glycemic prediction correlation $R^2 > 0.8$ with actual glycemic.

CHAPTER 4

EXPERIMENT

Based on ?? and ??, To remove undesirable spectral data from acquired spectra, we first measure the background noises (e.g., sample slide, air) as the starting spectral value from the Raman instrument as shown in Figure 4.1.

Second, we collected blood sample from index finger in fasting condition of 5 to 6 hours for 6 times on sample slide. From the above mentioned, in Figure 4.2 we can see the sample slide peak obviously because of difficulty of uncontrollably blood shedding on it. To eliminate mistakes from this peak in future experiments, we must carefully manage how blood is placed on the sample slide and keep it centered on one location.

We normalized the sample spectra with the lipid peak (1450 cm^{-1}) and the hemoglobin peak (1549 cm^{-1}), as illustrated in Figures 4.3 and 4.4. Although the preprocessing experiment did not go as expected, and the blood sample did not show a glucose peak (1125 cm^{-1}) obviously, it will be fully developed in future study.

Figure 4.1

Raman spectra background noises.

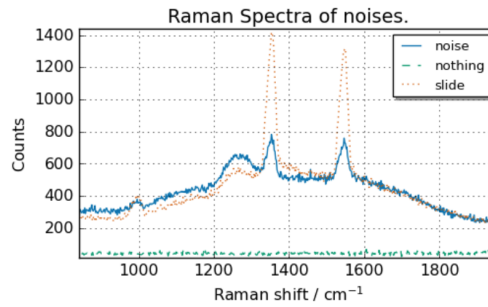


Figure 4.2

Raman spectra of blood sample from index finger.

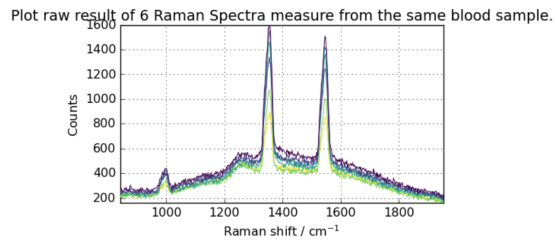


Figure 4.3

Raman spectra of blood sample with lipid normalized.

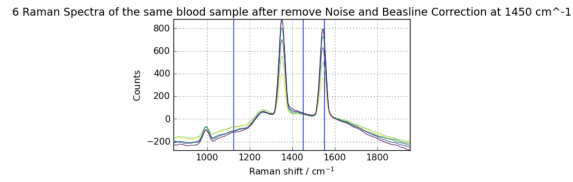
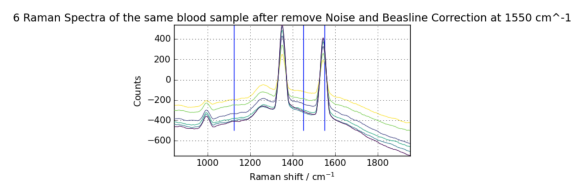


Figure 4.4

Raman spectra of blood sample with hemoglobin normalized.



CHAPTER 5

CONCLUSION

In this research study, we describe a preprocessing experiment based on previous research studies on direct glucose measurement. In this work, we tested the hypothesis that the whole spectrum should be normalized for protein/hemoglobin (Kang et al., 2020; Shao et al., 2012). We only tried to obtain spectral data from the desired measurement locations (e.g., nailfold and index finger) in earlier studies (Li et al., 2019) (González Viveros et al., 2022), but the findings are still ambiguous. Based on the above, we decided to return to fundamental studies by measuring spectrum data of the glucose peak (1125 cm^{-1}) on liquid samples and re-design tests to confirm that our Raman equipment provides precision and accuracy before measuring glycemic spectra via skin (interested measuring sites).

However, there remain uncharted areas with our experiment because it has not been properly tested. The experiment only used blood from the index finger after a fast of 5 to 6 hours. The acquired spectra were normalized for protein and hemoglobin, and background noise was removed. In any case, we want to repeat the entire experiment as a calibration before measuring spectral data via skin with participants in the future.