Journal of the Optical Society of Korea
Vol. 13, No. 2, June 2009, pp. 279-285

DOI: 10.3807/JOSK.2009.13.2.279

Glucose Prediction in the Interstitial Fluid Based on Infrared Absorption Spectroscopy Using Multi-component Analysis

Hyejeong Kim

Institute for Biomedical Electronics, Seoul National University of Technology, Seoul 139-743, Korea

Insup Noh

Department of Chemical Engineering, Seoul National University of Technology, Seoul 139-743, Korea

Gilwon Yoon*

Department of Electronics & Information Engineering, Seoul National University of Technology, Seoul 139-743, Korea

(Received April 23, 2008 : revised May 15, 2009 : accepted May 28, 2009)

Prediction of glucose concentration in the interstitial fluid (ISF) based on mid-infrared absorption spectroscopy was examined at the glucose fundamental absorption band of 1000 - 1500/cm (10 - 6.67 um) using multi-component analysis. Simulated ISF samples were prepared by including four major ISF components. Sodium lactate had absorption spectra that interfere with those of glucose. The rest NaCl, KCl and $CaCl_2$ did not have any signatures. A preliminary experiment based on Design of Experiment, an optimization method, proved that sodium lactate influenced the prediction accuracy of glucose. For the main experiment, 54 samples were prepared whose glucose and sodium lactate concentration varied independently. A partial least squares regression (PLSR) analysis was used to build calibration models. The prediction accuracy was dependent on spectrum preprocessing methods, and Mean Centering produced the best results. Depending on calibration sample sets whose sodium lactate had different concentration levels, the standard error prediction (SEP) of glucose ranged 17.19 ~ 21.02 mg/dl.

Keywords: Glucose, Infrared spectroscopy, Interstitial fluid, PLSR, Sodium lactate

OCIS codes : (170.6510) Spectroscopy, tissue diagnostics; (000.1430) Biology and medicine; (260.3060) Infrared

I. INTRODUCTION

Diabetes mellitus is one of the most common non-contagious diseases. Worldwide 221 million are expected to be diabetics in 2010 [1]. It is important for diabetic patients to monitor glucose level and to manage the disease systematically. In order to read glucose level, blood withdrawn from a finger or other part of the body using a lancet is placed on a strip. Though the measurement procedure is simple, pain and contamination problems are factors preventing a diligent disease-management.

Recently, studies on noninvasive or minimally invasive methods have been made, and one of the most widely used technologies is based on optics. Many different aspects of optics, including fluorescence and absorption spectroscopy, have been investigated in various fields [2-4]. For glucose measurement, targets are blood and interstitial fluid (ISF). Infrared light radiates into the body and the reflected or backscattered portion is measured. Blood has various components and, furthermore, light also interacts with other tissues such as skin. Therefore, noninvasive glucose measurement has to overcome a huge challenge of extremely low sensitivity and selectivity [2]. ISF can be extracted from skin using reverse iontophoresis or ultrasound [5, 6]. The extraction

*Corresponding author: gyoon@snut.ac.kr

of ISF does not cause pain. ISF contains inorganic ions such as Na^+, K^+ and Cl^-, but has a very low concentration of protein such as hemoglobin [7, 8]. ISF has been used as a medium for determining glucose concentration [9, 10]. Infrared absorption spectroscopy can be an excellent candidate for ISF glucose since there are no dominant scattering effects observed in this noninvasive measurement.

ISF contains several inorganic ions and sodium lactate in particular has absorption peaks in a wavelength region where glucose has also absorption peaks. It is suspected that sodium lactate influences glucose determination. However, interestingly enough, no systematic investigations have been reported on glucose determination under the influence of other interfering components. In this study, we introduced a method of determining glucose level in the ISF based on mid infrared spectroscopy. First, we analyzed sodium lactate's influence in predicting glucose concentrations using Design of Experiment (DOE). Then, we performed calibration and prediction modeling for samples where the concentrations of glucose and sodium lactate were varied independently.

II. EXPERIMENTAL RESULTS

Due to the highly absorbing nature of biological medium at the mid-infrared wavelength region, light absorbance can be given by the following equation.

$$\log(I_0/I) = \varepsilon \; c \; d \tag{1}$$

where I_0 is the incident light, I is measured light, ε is the extinction coefficient, c is the concentration of component, and d is pathlength of light travel.

For a medium with many components, Eq. (1) is more conveniently expressed in a matrix form. If E represents a combination of ε and d since we used a sample cell of fixed-pathlength ($d = 25$ μm), then

$$A_{m,n} = E_{m,p} \; C_{p,n} \tag{2}$$

$A_{m,n}$ is an m-by-n absorbance matrix where m is the number of wavelengths, n is the number of samples. $C_{p,n}$ is a concentration matrix and $E_{m,p}$ is a coefficient matrix. p represents the number of components in the sample.

The calibration process includes the spectrum measurement for a certain number of sample sets whose actual concentrations, so called reference values, should be obtained using other existing methods. A is known from absorption measurement and C is obtained from the existing method. A prediction process is to estimate values of C for an unknown sample using A and E obtained during the calibration process. In this study, a partial least squares regression (PLSR) analysis was used for calibration and prediction modeling [11, 12]. During matrix inversion, PLSR decomposes the matrices simultaneously, sharing the matrix information with each other. Therefore, PLSR overcomes the problem of multiple collinearity and provides more robust and reliable computation.

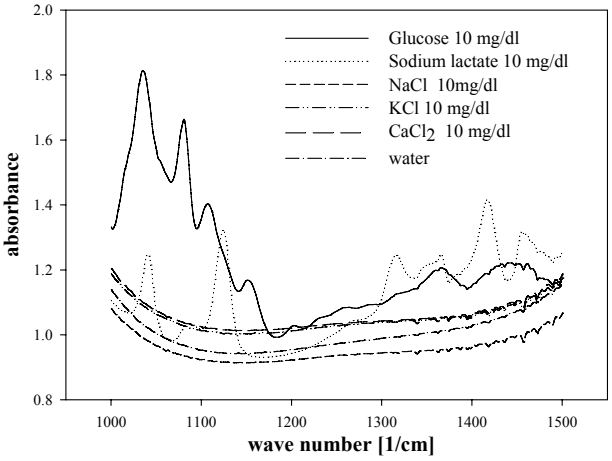

FIG. 1. Absorbance spectra of the ISF components measured by FT-IR spectrophotometer.

1. Influence of other substances in predicting glucose concentration

We prepared ISF samples based on Hartmann's solution. Hartmann's solution is isotonic with blood and is used for intravenous administration. Hartmann's solution contains sodium chloride (NaCl) 6.0 g, potassium chloride (KCl) 0.4 g, calcium chloride ($CaCl_2$) 0.27 g, sodium lactate ($NaC_3H_5O_3$) 3.22 g in 1 liter of injection water [13].

Fig. 1 shows absorption spectra of glucose and substances in Hartmann's solution. Glucose and sodium lactate have distinctive absorption peaks and some of the peaks are overlapping each other. This spectral overlapping implicates that the accuracy of glucose prediction can be interfered by sodium lactate. However, other substances (NaCl, KCl, $CaCl_2$, 3rd order distilled water (DW)) showed no apparent absorption peaks except relatively flat baseline changes.

We performed a preliminary experiment to examine the influence of sodium lactate in predicting glucose concentration. We used Design of Experiment (DOE) for this purpose. DOE is an analysis method that optimizes experimental conditions, often having a minimal number of experiments and obtaining maximally available information [14]. DOE also examines factors that influence experimental outcome.

The concentration level of glucose and sodium lactate was four factor levels (0, 1, 2, 3); 0, 1000, 5000 and 10000 mg/dl for glucose and 0, 100, 400 and 1000 mg/dl for sodium lactate. The maximum concentrations were higher than the allowable physiological level in order to enhance the effect of interferences. We tested using two different types of solution; 3rd DW (factor level 0) and 0.9% saline solution (factor level 1). We used soft-

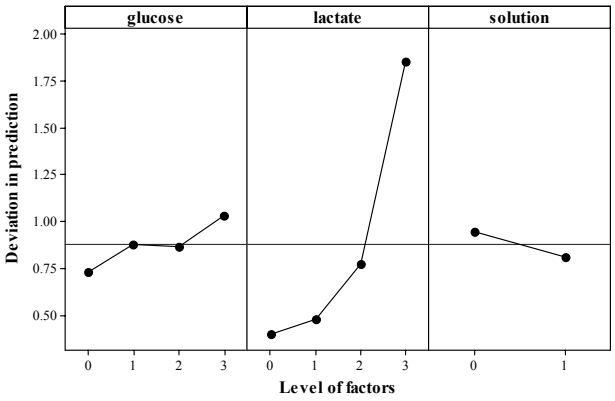

FIG. 2. Analysis of Design of Experiment using the Minitab^{TM} program. Deviations in predicted glucose concentrations were plotted with respect to glucose and lactate levels as well as solution type (water and saline).

ware, Minitab^{TM} (Minitab Inc., USA) where the number of experimental conditions was reduced to 16 based on Taguchi Method L_{16} ($4^2 \times 2^1$) array.

Fig. 2 shows the deviations in predicting glucose with respect to different factor levels. Deviation was defined as the magnitude of glucose prediction error and increased slightly when the glucose level increased. The medium type (DW or saline solution) did not make a noticeable difference. On the other hand, it was observed that sodium lactate influences glucose prediction. The higher the sodium lactate level was, the higher the glucose prediction error was as shown in Fig. 2.

2. Sample preparation and mid-infrared spectrum measurement

We maintained the same concentrations for NaCl, KCl and $CaCl_2$ as in Hartmann's solution since they did not have any specific absorption peaks in the frequency region of measurement. The concentrations of glucose and sodium lactate were varied in order to examine calibration and prediction modeling using PLSR regression analysis. A physiologically allowable concentration is up to 500 mg/dl for glucose and diabetic patients can have even 600 mg/dl sometimes. We set the glucose level between 0 – 1000 mg/dl in preparing the samples. The number of the glucose level was set to 18.

The maximum concentrations of sodium lactate were set to about twice the maximally allowable normal level. Hartmann's solution contains much higher concentration of lactic acid than the normal body fluid does. That is why it is also called Ringer's lactate solution. Normally, lactate is under 18 mg/dl in whole blood [15]. Lactate level can be even ten times higher than the normal level under stress or shock. But, in our analysis, we assumed that sodium lactate varies within the normal range. The concentration between 0 and 50 mg/dl was assigned for sample preparation and 9 concentration levels were set.

54 samples were prepared such that there was no concentration correlation between glucose and sodium lactate (correlation coefficient = -0.027, p-value = 0.846). The remaining three substances had the same concentrations as in Hartmann's solution. The concentrations of samples were summarized in Table 1.

Mid infrared spectra were measured using a FT/IR-4100 (Jasco co. Japan) spectrophotometer. A sample thickness of 25 μm was controlled by a Teflon spacer between the windows. The window had a thickness of

TABLE 1. Concentrations of the ISF substances in 54 samples. Concentrations of the other three components are the same (NaCl = 600 mg/dl, KCl = 40 mg/dl, $CaCl_2$ = 27 mg/dl). unit [mg /dl]

glucose	S. lactate	glucose	S. lactate	glucose	S. lactate	glucose	S. lactate
0	0	700	30	400	40	200	50
10	10	800	35	450	35	250	30
20	15	900	40	500	30	300	10
50	20	1000	50	600	0	350	25
100	25	0	50	700	20	400	35
150	30	10	40	800	15	450	40
200	35	20	35	900	10	500	35
250	40	50	30	1000	25	600	30
300	50	100	50	0	15	700	25
350	0	150	20	10	40	800	0
400	10	200	15	20	40	900	20
450	15	250	10	50	0	1000	15
500	20	300	20	100	0		
600	25	350	50	150	10		

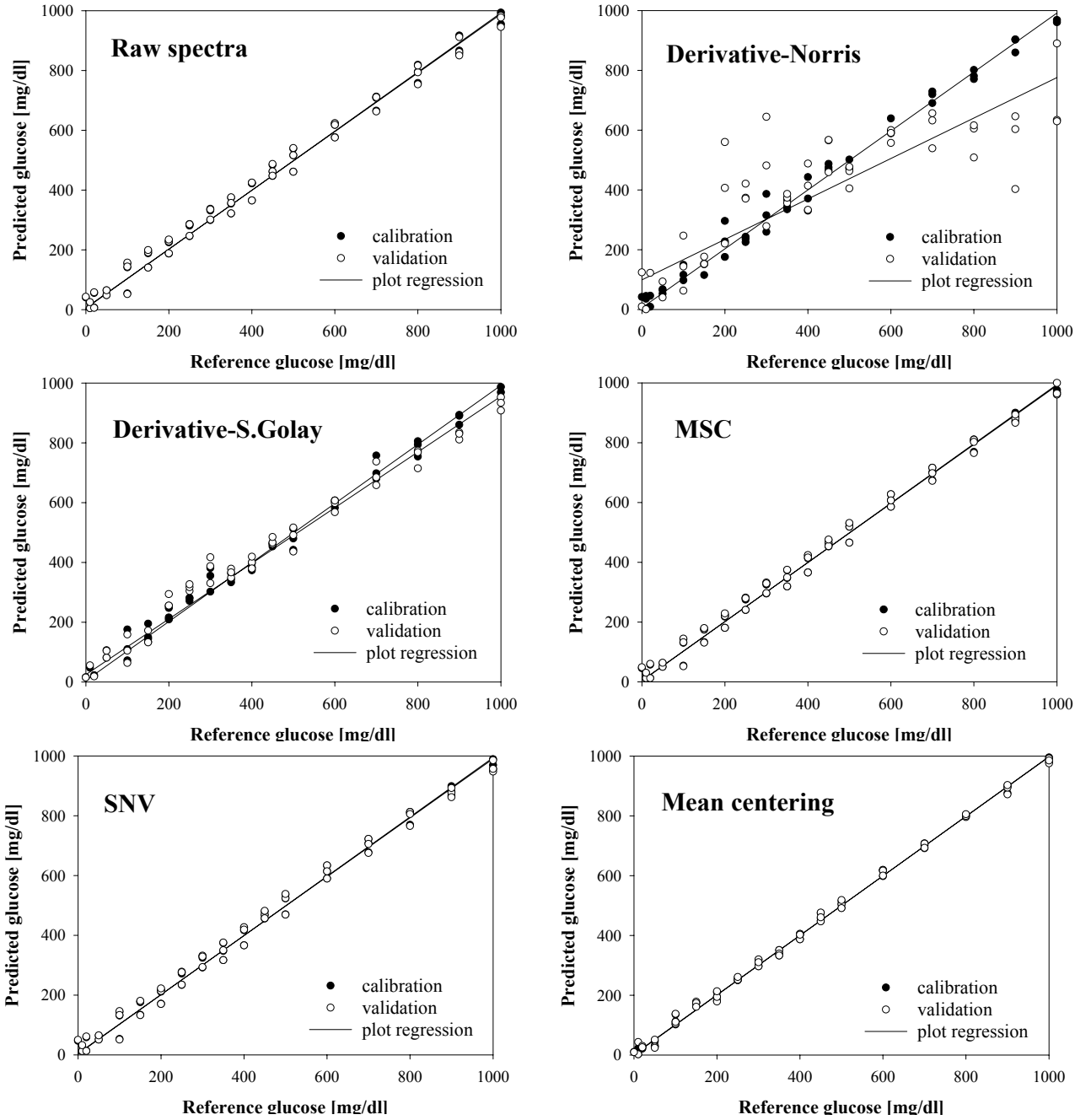

FIG. 3. Predicted glucose concentrations with respect to various spectrum preprocessing methods.

3 mm and a diameter of 32 mm and was made of BaF_2. The windows were attached to a sample holder. Once the whole set (PIKE, USA) was assembled, sample solution was injected into one of two holes using a syringe. Absorption spectra between 700-7800 cm^{-1} (14.3 − 1.28 um) were measured. However, spectral analysis was done at 1000 - 1500 cm^{-1} (10 − 6.67 um) that included glucose fundamental absorption peaks.

3. Spectral analysis and concentration prediction

First, different methods of spectral data preprocessing were applied to measured spectra. We used a commercial software, the Unscrambler v9.7 (CAMO co., Norway). Five different data preprocessing methods were applied; Derivatives-Norris (DN), Derivatives-S Golay (DG), Multiplicative Scatter Correction (MSC), Standard Normal Variate (SNV) and Mean Centering (MC). The results of calibration and full-cross-validation in predicting glucose concentration were shown in Fig. 3. Measured spectra with no preprocessing treatment were designated as 'raw' spectra.

Mean Centering was better than any other preprocessing

TABLE 2. Calibration and validation results with respect to the different spectral preprocessing methods.

preprocessing method	calibration		validation	
	R^2	RMSEC [mg/dl]	R^2	RMSEP [mg/dl]
raw spectra	0.9955	32.56	0.9948	34.98
Derivative-Norris	0.9953	33.37	0.8896	161.68
Derivative-S.Golay	0.9952	33.56	0.9898	49.23
multiplicative scatter correction (MSC)	0.9968	27.36	0.9963	29.45
standard normal variate (SNV)	0.9964	29.19	0.9958	31.48
mean centering (MC)	0.9979	14.14	0.9974	16.06

TABLE 3. Statistics on calibration samples and the results of calibration and validation.

calibration samples					calibration		validation	
group	sample number	correlation coefficient*	p- value	sodium lactate range**	R^2	RMSEC [mg/dl]	R^2	RMSEP [mg/dl]
Cal #A	27	0.102	0.611	0 - 50	0.9975	14.98	0.9962	19.01
Cal #B	27	0.006	0.977	25 - 50	0.9958	18.30	0.9939	22.81
Cal #C	27	0.267	0.179	0 - 25	0.9984	13.22	0.9974	17.58

* Correlation coefficient between glucose and sodium lactate concentrations
** unit [mg/dl]

TABLE 4. Prediction of glucose concentration from three calibration models.

calibration model	predicted sample	prediction	
		R^2	RMSEP [mg/dl]
Cal #A	all 54 samples	0.9967	17.84
Cal #B		0.9953	21.02
Cal #C		0.9969	17.19

methods producing $R^2 = 0.9979$ and RMSEC (root mean squares standard error of calibration) = 14.14 mg/dl for calibration, $R^2 = 0.9974$ and RMSEP (root mean squares standard error of prediction) = 16.06 mg/dl for validation. MSC was the next best. Derivative-Norris and Derivative-S Golay produced high values of RMSEP. We also examined the cases where a spectrum was treated with more than one preprocessing method. However, the results were no better. In this work, we decided to choose Mean Centering as a data preprocessing method. Then, we performed a PLSR analysis of calibration and prediction.

In order to examine the influence of sodium lactate level in predicting glucose concentration, we made three different calibration groups out of a total of 54 samples. As shown in Table 2, calibration model A (Cal #A) consisted of 27 samples whose sodium lactate concentrations were distributed in the entire range between 0 and 50 mg/dl. A correlation coefficient between glucose and sodium lactate levels was 0.102 and p-value was 0.611. This indicated no correlation between glucose and sodium lactate concentrations among the samples of Cal #A. The Cal #B sample group had higher sodium lactate concentrations having a distribution of 25 - 50 mg/dl. The Cal #C sample group had lower sodium lactate concentrations (0 -25 mg/dl). The samples of Cal #B and Cal #C group were also arranged such that there was no concentration correlation between glucose and sodium lactate with a p-value of 0.977 and 0.179 respectively.

RMSECs of three different calibration models were computed using PLSR analysis. The error of Cal #C (samples with low sodium lactate level) was the smallest (13.22 mg/dl). Cal #B with high sodium lactate level produced the highest RMSEC (18.30 mg/dl). Cal #B had higher RMSEC even compared with Cal #A that had a distribution of sodium lactate in the entire range (0 – 50 mg/dl). We observed the same trend with validation analysis where the Cal #B group had 22.81 mg/dl as RMSEP value. Validation was done based on one-leave-out cross validation where each sample was predicted from the model constructed using the rest samples. This was repeated for every sample in the group. Finally, using three calibration models, glucose was predicted

for the entire set of 54 samples (Table 4). The results showed the same trend where the calibration model with lower sodium lactate level produced the smallest RMSEP.

III. DISCUSSION AND CONCLUSION

The interstitial body fluid contains several major components other than glucose. Most often, glucose prediction has been investigated based on one component system, i.e., only glucose level was varied while the concentrations of the remaining components were fixed. Absorption spectra and DOE Taguchi analysis revealed that sodium lactate has absorption peaks in the glucose fundamental absorption band suggesting a potential interference. In biomedical spectroscopy, the interference among the components is one of the important problems. It will be very time-consuming to prepare samples where the concentration of each component varies independently. With a limited number of samples based on the DOE Taguchi method, we selected sodium lactate as an interfering component in predicting a target concentration. Then, we proceeded to the main experiment. Even though the DOE methods have been used in the quality control industry, we believe that it is an effective screening-tool for designing a main experiment. Then we prepared samples based on a two-component system where the concentrations of glucose and sodium lactate were varied independently. The rest of the components remained the same (NaCl = 600 mg/dl, KCl = 40 mg/dl, $CaCl_2$ = 27 mg/dl).

Often, a few- to several-wavelength measurement can be use to predict a target concentration. However, this so-called discrete wavelength method is valid only when there is no optical interference by other substances. For example, we calculated RMSEC and RMSEP at the two wavelengths (1040/cm and 1075/cm). For glucose, RMSEC was 87.69 mg/dl and RMSEP was 91.24 mg/dl. For comparison, our PLSR analysis produced 14.14 mg/dl and 16.06 mg/dl for RMSEC and RMSEP respectively. For sodium lactate, R^2 was smaller than 0.1 and we could not use two-wavelength any more due to very poor correlation coefficients. That is why statistical analysis like PLSR based on many-wavelength has been used for a weakly absorbing and multi-component system.

One salient feature was the importance of preprocessing of measured optical spectra. Depending on different methods, RMSEC varied between 14.14 ~ 33.56 mg/dl and RMSEP ranged from 16.06 mg/dl to 161.68 mg/dl. MSC and SNV are the methods of targeting on error reduction induced by medium scattering and they proved to be less effective. Mean Centering was a choice of data preprocessing with having the smallest RMSEC and RMSEP. It is speculated due to the facts that optical spectrum is dominated by absorption at the mid infrared, and that most variations in absorption profile are shown in the form of baseline shift in absorption spectrum.

For the main experiment, another 54 samples within the physiological range were prepared. To examine the influence of sodium lactate level, different calibration models were built based on the sodium lactate level. Indeed, the sample set consisting of samples with high sodium lactate levels (25 - 50 mg/dl) produced the highest errors. This feature was consistent for all cases of calibration, validation and prediction. Prediction accuracy is more critical than that of calibration or validation because the glucose prediction of a sample will be what matters in practice. Prediction results summarized in Table 4 are the most important. Calibration models made from Cal #C (0 - 25 mg/dl sodium lactate) produced a similar SEP compared to Cal #A (0 - 50 mg/dl sodium lactate). They were 17.19 and 17.84 mg/dl. On the other hand, Cal #B (25 - 50 mg/dl) produced 21.02 mg/dl. This value was 22% and 18% higher than Cal #C and Cal #A, respectively. If a well-designed calibration sample group is used, our observation was that the prediction error of glucose could be slightly under 18 mg/dl.

ACKNOWLEDGMENT

This study was supported by a grant of the Korea Healthcare Technology R&D Project, Ministry for Health, Welfare and Family Affairs, Republic of Korea (A060484).

REFERENCES

1. N. H. Jo, "Prevalence of diabetes and management status in Korean population," The Korea Journal of Internal Medicine **68**, 10-17 (2005).
2. R. McNichols and G. L. Cote, "Optical glucose sensing in biological fluids : an overview," J. Biomedical Optics **5**, 5-16 (2000).
3. S. H. Kim, D. S. Choi, and D. Kim, "Single detection of fluorescence resonance energy transfer using confocal microscopy," J. Opt. Soc. Korea **12**, 107-111 (2008).
4. C. C. Byeon, M.-K. Oh, H. Kang, D.-K. Ko, and J. Lee, "Coherent absorption spectroscopy with supercontinuun for semiconductor quantum well structure," J. Opt. Soc. Korea **11**, 138-141 (2007).
5. A. Sieg, R. H. Guy, and M. B. Delgado-Charro, "Simultaneous extraction of urea and glucose by reverse iontophoresis in vivo," Pharmaceutical Research **10**, 1805-1810 (2004).
6. S. Mitragotri, M. Coleman, J. Kost, and R. Langer, "Transdermal extraction of analytes using low-frequency ultrasound," Pharmaceutical Research **4**, 466-470 (2000).
7. H. K. Seong, *Physiology* (Medical culture co., Seoul, Korea, 1991), pp. 268-275.
8. N. Fogh-Andersen, B. M. Altura, B. T. Altura, and O. Siggaard-Andersen, "Composition of interstitial fluid," Clinical Chemistry **41**, 1522-1525 (1995).

9. J. P. Bantle and W. Thomas, "Glucose measurement in patients with diabetes mellitus with dermal interstitial fluid," J. Lab. Clin. Med. **130**, 436-441 (1997).
10. A, Sieg, R. H. Guy, and M. B. Delgado-Charro, "Noninvasive and minimally invasive methods for transdermal glucose monitoring," Diabetes Technology & Therapeutics **7**, 174-197 (2005).
11. H. Chung, M. A. Arnold, M. Rhiel, and D. W. Murhammer, "Simultaneous measurements of glucose, glutamine, ammonia, lactate, and glutamate in aqueous solutions by near-infrared spectroscopy," Applied Spectroscopy **50**, 270-276 (1996).
12. K. J. Jeon, I. D. Hwang, S. Hahn, and G. Yoon, "Comparison between transmittance and reflectance measurements in glucose determination using near infrared spectroscopy," Journal of Biomedical Optics **11**, 014022-1~ 014022-7 (2006).
13. http://drugs-about.com/drugs/compound-sodium-lactate /compound-sodium-lactate.pdf
14. S. H. Park, *Modern Design of Experiments* (Minyongsa, Korea, 1995), pp. 452-453.
15. D. Lafrance, L. C. Lands, and D. H. Burns, "Measurement of lactate in whole human blood with near-infrared transmission spectroscopy," Talanta **60**, 635-641 (2003).