

# Noninvasive Prediction of Glucose by Near-Infrared Diffuse Reflectance Spectroscopy

STEPHEN F. MALIN, TIMOTHY L. RUCHTI, THOMAS B. BLANK, SURESH N. THENNADIL, and  
STEPHEN L. MONFRE\*

**Background:** Self-monitoring of blood glucose by diabetics is crucial in the reduction of complications related to diabetes. Current monitoring techniques are invasive and painful, and discourage regular use. The aim of this study was to demonstrate the use of near-infrared (NIR) diffuse reflectance over the 1050–2450 nm wavelength range for noninvasive monitoring of blood glucose.

**Methods:** Two approaches were used to develop calibration models for predicting the concentration of blood glucose. In the first approach, seven diabetic subjects were studied over a 35-day period with random collection of NIR spectra. Corresponding blood samples were collected for analyte analysis during the collection of each NIR spectrum. The second approach involved three nondiabetic subjects and the use of oral glucose tolerance tests (OGTTs) over multiple days to cause fluctuations in blood glucose concentrations. Twenty NIR spectra were collected over the 3.5-h test, with 16 corresponding blood specimens taken for analyte analysis.

**Results:** Statistically valid calibration models were developed on three of the seven diabetic subjects. The mean standard error of prediction through cross-validation was 1.41 mmol/L (25 mg/dL). The results from the OGTT testing of three nondiabetic subjects yielded a mean standard error of calibration of 1.1 mmol/L (20 mg/dL). Validation of the calibration model with an independent test set produced a mean standard error of prediction equivalent to 1.03 mmol/L (19 mg/dL).

**Conclusions:** These data provide preliminary evidence and allow cautious optimism that NIR diffuse reflectance spectroscopy using the 1050–2450 nm wavelength range can be used to predict blood glucose concentrations noninvasively. Substantial research is still required to validate whether this technology is a viable tool for long-term home diagnostic use by diabetics.

© 1999 American Association for Clinical Chemistry

Diabetes is a leading cause of death and disability worldwide and afflicts an estimated 16 million Americans. Complications of diabetes include heart and kidney disease, blindness, nerve damage, and high blood pressure, with the estimated total cost to the United States economy alone exceeding \$90 billion per year (1). Long-term clinical studies have shown that the onset of complications can be substantially reduced through proper control of blood glucose (2).

A vital element of diabetes management is the self-monitoring of blood glucose by diabetics in the home environment. Unfortunately, current monitoring techniques discourage regular use because of the invasive and painful nature of sampling. Therefore, new methods for self-monitoring of blood glucose are required to improve the prospects for more rigorous control of blood glucose in diabetic patients.

Numerous approaches have been explored for measuring blood glucose, ranging from invasive methods such as microdialysis to noninvasive technologies that rely on spectroscopy. Each method has associated advantages and disadvantages, but only a few have received approval from certifying agencies. Unfortunately, noninvasive techniques for the self-monitoring of blood glucose have not been certified.

One method, near-infrared (NIR)<sup>1</sup> diffuse reflectance spectroscopy involves the illumination of a spot on the body with low-energy NIR light (750–2500 nm). The light is partially absorbed and scattered, according to its interaction with chemical components within the tissue, before being reflected back to a detector. The detected light is used to create a graph of  $-\log R/R_s$ , where  $R$  is the reflectance spectrum of the skin and  $R_s$  is the reflectance of the instrument calibrator. In infrared spectroscopy, this graph is analogous to an absorbance spectrum containing

Instrumentation Metrics, Incorporated, 2085 Technology Circle, Suite 302, Tempe, AR 85284.

\*Author for correspondence. Fax 602-755-9832; e-mail smonfre@ix.netcom.com. Received May 7, 1999; accepted June 17, 1999.

<sup>1</sup> Nonstandard abbreviations: NIR, near-infrared; SEP, standard error of prediction; OGTT, oral glucose tolerance test; PLS, partial least squares; YSI, Yellow Springs Instruments; SECV, standard error of prediction through cross-validation; and SEC, standard error of calibration.

quantitative information that is based on the known interaction of the incident light with components of the body tissue and will be referred to in this manner throughout the remainder of the text.

The absorbance spectrum contains a mixture of the spectral signatures of many tissue components, including water, fat, protein, and glucose. An example of a noninvasive, NIR sample spectrum is given in Fig. 1. The spectrum was collected using a custom-designed spectrometer, described below, with staring detectors. An 80% Spectralon reflectance calibrator was used to measure  $R_s$ , and the forearm of a human subject was scanned to obtain  $R$ .

The prediction of blood glucose concentration is accomplished by detecting the magnitude of light attenuation caused by the absorption signature of blood glucose as represented in the targeted tissue volume (capillary beds) of the skin. The process of calibration involves the development of a mathematical transformation or model, which is used to estimate the blood glucose concentration from the measured tissue absorbance spectrum.

Specific studies have demonstrated that NIR diffuse reflectance spectroscopy represents a feasible and promising approach to the noninvasive prediction of blood glucose concentrations (3–12). Robinson et al. (3) reported three different instrument configurations for measuring diffuse transmittance through the finger in the 600–1300 nm range. Meal tolerance tests were used to perturb the glucose concentrations of three subjects, and calibration

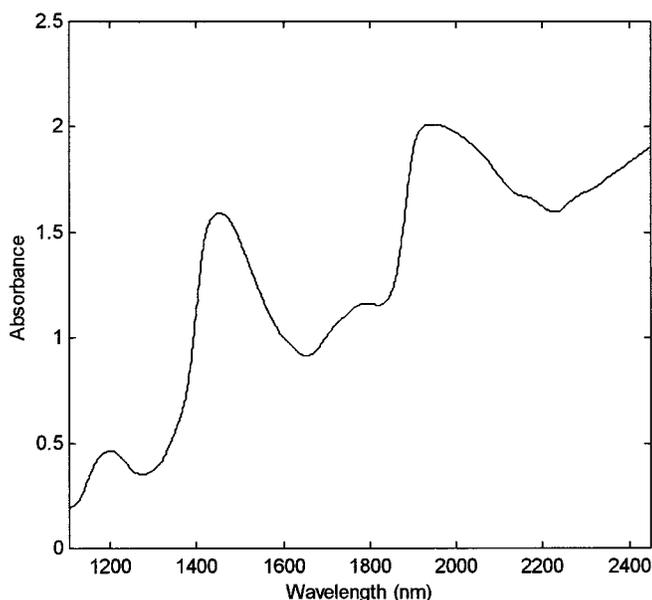


Fig. 1. Plot of the  $-\log_{10}R/R_s$  absorbance spectrum collected in diffuse reflectance mode on the forearm, using a custom-built scanning spectrometer with staring detectors.

The visually discernible absorbance features are attributable to water (1192, 1448, 1787, 1948, and 2600 nm), fat (1210, 1725, 1760, 2299, and 2342 nm), and protein (1700, 1740, 2056, 2174, and 2273 nm). The magnitude of absorbance attributable to glucose is four to five orders of magnitude less than the absorbance of water, with glucose absorbance bands located at 1613, 1689, 1732, 2105, 2273, and 2326 nm (27).

models were constructed specific to each subject on single days and tested through cross-validation. The absolute mean prediction errors ranged from 1.1 to 2.1 mmol/L (19.8 to 37.8 mg/dL).

Heise and co-workers (5,7) and Marbach et al. (6) presented results through a diffuse reflectance measurement of the oral mucosa in the 1111–1835 nm range with an optimized diffuse reflectance accessory (13). In vivo experiments were conducted on single diabetics using glucose tolerance tests and on a population of 133 different subjects. The best standard error of prediction (SEP) reported was 2.4 mmol/L (43 mg/dL) and was obtained from a 2-day single-person oral glucose tolerance test (OGTT) that was evaluated through cross-validation.

The studies reported by Jagemann et al. (8), Fischbacher et al. (9), Danzer et al. (10), and Muller et al. (12) recorded spectra in diffuse reflectance over the 800–1350 nm range on the middle finger of the right hand with a fiber-optic probe. Each experiment involved a diabetic subject and was conducted over a single day with perturbation of blood glucose concentrations through carbohydrate loading. Results, using both partial least-squares (PLS) regression and radial basis function neural networks were evaluated on single subjects over single days through cross-validation. Danzer et al. (10) reported a mean root mean square prediction error of 2.0 mmol/L (36 mg/dL) through cross-validation over 31 glucose profiles.

Burmeister et al. (11) collected absorbance spectra through a transmission measurement of the tongue in the 1429–2000 nm range. A study of five diabetic subjects was conducted over a 39-day period with five samples taken per day. Every fifth sample was used for an independent test set and the SEP for all subjects was  $>3.0$  mmol/L (54 mg/dL).

In all of these studies, limitations were cited that would affect the acceptance of such a method as a commercial product. These limitations included sensitivity, sampling problems, time lag, calibration bias, long-term reproducibility, and instrument noise. Additionally, accurate noninvasive estimation of blood glucose is limited at present by the available NIR technology and the dynamic nature of the sample: the skin and living tissue of the patient [for example, see Khalil (14)]. Chemical, structural, and physiological variations occur that produce dramatic and nonlinear changes in the optical properties of the tissue sample. The measurement is further complicated by the complex and varying background signals of other substances present in blood (tissue).

The research presented here is an advancement in the development a noninvasive diffuse reflectance NIR analysis for blood glucose and addresses several of the limitations of prior studies. The spectral range is extended to cover 1050–2450 nm, and the measurement is made with two different optical probes involving staring detectors and a fiber-optic accessory. The addition of the 2000–2450 nm range ensures accessibility to the glucose bands with

highest absorbance in the near-infrared (15, 16), and the change from staring optics to fibers dramatically reduces the problem of skin surface reflectance. The measurement is made on the forearm as opposed to the finger, tongue, or oral mucosa, and a custom-designed arm cradle was applied to minimize pressure and placement error, the details of which will be described elsewhere. For a diffuse reflectance measurement in the target wavelength range, this site was less susceptible to the effects of saliva, food, or movement during measurement than the tongue or lip. Furthermore, the forearm is expected to produce a more uniform spectral measurement between subjects than the fingers.

Two separate studies are presented aimed at the development of calibrations specific to individuals. The first involves the repeated measurement of subjects on a regular basis over time and the second uses OGTTs to rapidly change the subject's blood glucose concentration.

The motivation for modeling of instrument responses during an OGTT lies in the expectation that tissue chemistry and hydration do not vary substantially during the test. A relatively constant tissue chemistry gives rise to a relatively constant background, which improves the accessibility of the analyte signal of glucose by reducing the complexity of the calibration model. Alternatively, models based on calibration spectra with modest background variation cannot be expected to generalize to spectra taken on tissue containing a chemical matrix not represented in the calibration set. Another undesirable aspect of modeling using OGTTs lies in the potential existence of ancillary correlations that can give rise to "false" predictions (17). Despite the pitfalls of OGTTs, we decided that the simplicity of the interferences would provide a unique opportunity to study the *in vivo* measurement.

Although the OGTT studies reported here are similar to prior studies, this work is the first reported effort that uses OGTT data conducted on different days to construct independent calibration and test sets. Two days of tolerance test data are used to build a calibration model that is applied to predict blood glucose concentrations on a third day. In addition, instrument drift and temperature, subject (arm) temperature, and interfering analytes are analyzed to ensure that the glucose predictions are not a product of spurious correlations to these factors.

## Materials and Methods

### INSTRUMENTATION

All NIR spectra were collected using a custom-built scanning NIR spectrometer. The instrument collected intensity spectra in diffuse reflectance from the forearm in the wavelength range 1050–2450 nm. The spectral sampling interval was 1 nm, and the signal-to-noise ratio at the peak intensity was  $\sim 90$  dB [ $10 \log_{10}$  (peak intensity/root-mean-square noise)]. The detectors used in the study were a combination of indium-gallium-arsenide and extended indium-gallium-arsenide detectors. The single-subject experiment used an optical configuration involv-

ing staring detectors, whereas the OGTT experiments used a simple fiber-optic interface to the skin similar to those reported previously in the literature [for example, see Jagemann et al. (8)]. Reference spectra were recorded before each sample measurement by scanning an 80% Spectralon reflectance material from Labsphere.

A cradle was developed to position the arm over the sample interface in a reproducible location with a reproducible degree of pressure. The arm was placed onto the instrument 30 s before each measurement and removed subsequent to the measurement period. The subjects remained seated during the experiment. After the 30-s period, the diffuse reflectance spectra were collected over the next 60 s, with eight scans equaling one spectrum.

The internal temperature of the instrument was measured during each scan with thermistors [Yellow Springs Instruments (YSI)] and custom electronics. Five different areas of the instrument were measured with the lamp assembly acting as the main heat source.

The skin temperature was measured using a YSI 4000 precision thermometer with a skin temperature probe. The location of the temperature measurement was  $\sim 5$  mm away from the illumination point. The skin temperature was taken before and after each 60-s measurement.

### SINGLE-SUBJECT EXPERIMENTS

Seven diabetic subjects were scanned regularly over a 35-day period with the custom-built scanning NIR spectrometer with staring detectors. Subjects were scanned at most one time per day. Venous glucose was measured by a hexokinase method (Covance Central Laboratories Services, Indianapolis, IN). A calibration model for estimating blood glucose concentrations from the  $-\log_{10}(R/R_s)$  absorbance spectra was generated for each subject using the measurements and standard multivariate calibration methods. Model performance was evaluated through cross-validation using the "leave-one-out" strategy and calculating the standard error values. The cross-validation procedure is used to iteratively predict the glucose concentration of each sample by using all other samples to construct the calibration model. After each sample has been predicted, the standard error of cross-validation (SECV) is computed as the root mean square error of the cross-validation glucose predictions (18).

### GLUCOSE TOLERANCE TESTING

A series of experiments was conducted to examine the modeling and prediction of glucose concentrations using OGTTs. Because of the risk of underlying secondary correlation, major blood constituents and instrument-related variables were tracked to determine their contribution to glucose calibration models.

Three nondiabetic participants were subjected to five OGTT experiments at 2–3 day intervals over the period of 2 weeks. The test protocol included a 6-h fasting period, followed by a 3.5-h OGTT after the ingestion of a beverage containing 100 g of D-glucose (Fisher Healthcare). The

NIR scanning spectrometer with the fiber-optic interface was used to collect and average two replicate spectra consisting of 16 scans per 2-min measurement. Scans of an 80% reflectance calibrator were collected before each noninvasive scan and used to convert intensity spectra to absorbance spectra as described previously.

Sixteen venous blood specimens were collected between the noninvasive scans, and the measured blood analytes were interpolated in time to estimate their respective concentrations at the instance of each spectral measurement. We measured blood glucose (in duplicate, YSI 2300 glucose analyzer), hematocrit and hemoglobin (Sysmex KX-21 hematology analyzer), and other blood analytes, including triglycerides, cholesterol, total bilirubin, total protein, and albumin (Olympus AU400).

Algorithms were developed using the Matlab, Ver. 5.3, software package with the Chemometrics Toolbox, Ver. 5.1, from Eigenvector Technologies. Preprocessing was performed using smoothed first derivatives and an in-house wavelength standardization algorithm. Model parameters such as wavelength ranges and numbers of scores were selected on the basis of best predictive performance on the data used in glucose prediction.

The data for each participant were separated into calibration and test sets by days (each test set consisted of the data associated with an entire day). Calibration models were developed through PLS regression (19). Model performance is reported as the standard error of calibration (SEC) and SEP corresponding to the root mean square error of the glucose predictions over the calibration and test sets, respectively.

## Results and Discussion

### SINGLE-SUBJECT STUDIES

Spectra collected on each subject were converted to absorbance units and preprocessed via multiplicative scatter correction (20) and Savitsky-Golay 55-point smoothing (21). Calibration models were developed through PLS using the wavelength ranges 1100–1380, 1450–1850, and 2050–2375 nm. However, only three of the seven subjects produced data sets that yielded statistically significant blood glucose calibrations. The first (subject 1) was a type 1 diabetic with a body mass index of 28 kg/m<sup>2</sup>. A total of 14 NIR spectra were accumulated for this subject with associated blood glucose concentrations ranging from 5.1 to 24.8 mmol/L (91 to 446 mg/dL). The test set results from cross-validation are depicted in Fig. 2. The plot is that of the predicted vs measured blood glucose concentrations. A “perfect” calibration would have all of the predictions falling on the theoretical 45° line (solid line). The diverging dashed lines indicate the 10% error-of-reading boundaries required for proper blood glucose monitoring.

The cross-validation results obtained using the first subject have a mean prediction error of 9.1% with a SECV of 1.6 mmol/L (28 mg/dL). Although evident from Fig. 2, the statistical significance of the cross-validation results

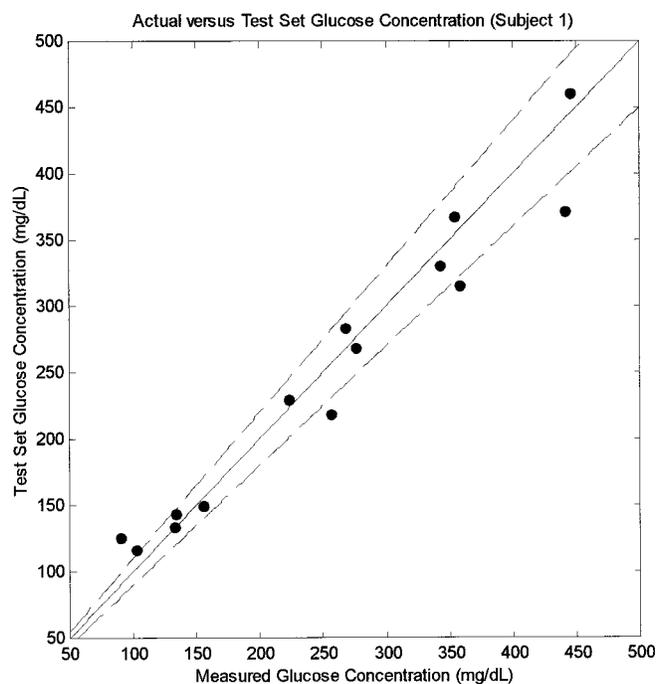


Fig. 2. Cross-validation results obtained from subject 1.

SEP = 1.6 mmol/L (28 mg/dL); mean error = 9.1%;  $t = 15.3$ . To convert results (mg/dL) to mmol/L, divide by 18.

was confirmed through a  $t$ -test on the slope of the predicted vs reference glucose concentrations. A correlation analysis between the measured blood glucose concentration and four experimental parameters, including chronological time (sample order), internal instrument temperature, reference scan intensity, and time of day, produced a maximum  $r^2$  of 0.13 for the time of day and  $r^2 < 0.05$  for the other variables. This indicates that it is unlikely that this calibration is related to instrument variation.

Although the results are encouraging, concern arises from the fact that four measurements (29%) fell outside the 10% error boundaries. Investigation of these specific samples provided no conclusions concerning the chemical or sampling variation that might have led to these errors.

The second subject was a type 2 diabetic with a body mass index of 34 kg/m<sup>2</sup>. The subject's blood glucose concentrations ranged from 4.6 to 16.3 mmol/L (82 to 294 mg/dL) over the 17 NIR scans that were recorded. A graphical representation of the test set cross-validation results is shown in Fig. 3.

Although the SECV for subject 2 is similar to that for subject 1 (1.7 and 1.6 mmol/L, respectively), it is obvious that the measurement does not provide an opportunity for proper blood glucose monitoring. In this experiment, the mean error-of-reading was 17.6%. Review of the graph demonstrates that sampling issues exist with this particular subject because it appears that the calibration model is not accurate for values  $< 8.3$  mmol/L (150 mg/dL). Although several potential reasons exist for this result (e.g., sampling, tissue hydration, and tissue temperature),

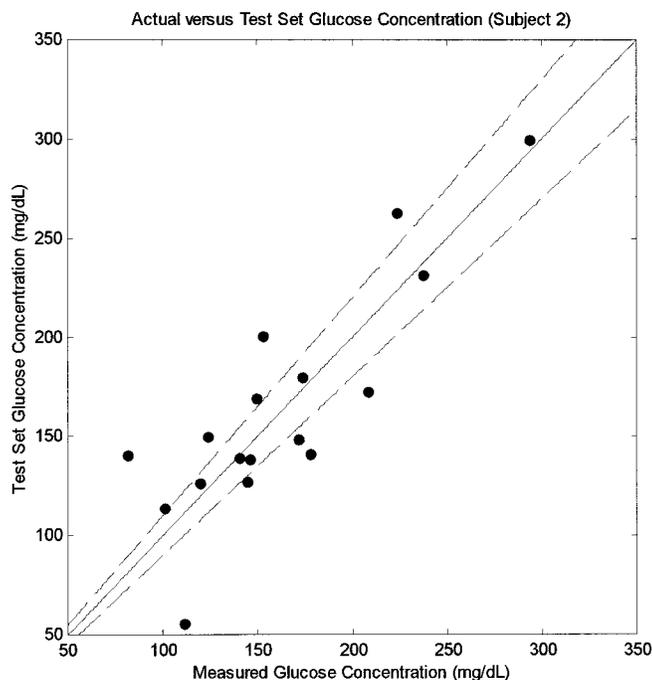


Fig. 3. Cross-validation results obtained from subject 2.

SEP = 1.7 mmol/L (31 mg/dL); mean error = 17.6%;  $t = 6.3$ . To convert results (mg/dL) to mmol/L, divide by 18. *Solid line*, theoretical 45° line. *Dashed lines*, 10% error-of-reading boundaries.

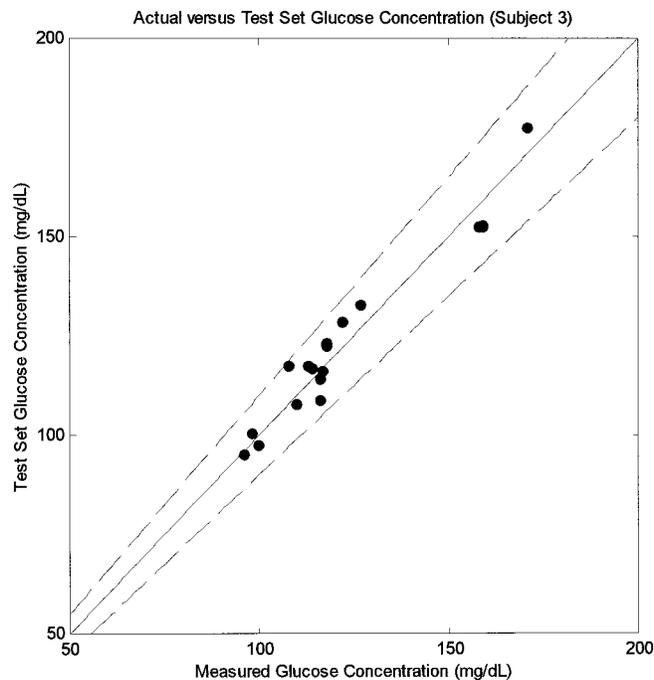


Fig. 4. Cross-validation results obtained from subject 3.

SEP = 0.95 mmol/L (17 mg/dL); mean error = 3.6%;  $t = 17.1$ . To convert results (mg/dL) to mmol/L, divide by 18. *Solid line*, theoretical 45° line. *Dashed line*, 10% error-of-reading lines.

there is no known physical parameter or variable that provides conclusive evidence as to the reason for the poor accuracy of the calibration model. These results demonstrate the variability accompanying any attempts to make in vivo measurements with NIR diffuse reflectance, as is evident with the next subject.

The third subject investigated was a type 2 diabetic with a body mass index of 34 kg/m<sup>2</sup> and a blood glucose of 5.4–9.5 mmol/L (97–171 mg/dL) in 18 samples. The cross-validation results are displayed in Fig. 4. Unlike the previous two subjects, the SECV for subject 3 was almost 50% lower [0.95 mmol/L (17.1 mg/dL)], and the mean error-of-prediction was only 3.6%. Such a result would easily make this an acceptable measurement for blood glucose. In this study, all readings fell within the 10% error-of-reading grid of the plot. Correlation analysis with the four experimental parameters described above did not reveal a significant systematic relationship that would explain the prediction results ( $r^2 < 0.16$ ).

The data from the remaining four diabetics produced no statistically relevant calibration models. Investigation of the differences between the successful ( $n = 3$ ) and unsuccessful ( $n = 4$ ) experiments revealed that issues existed within several areas of instrument design and patient interfacing. The key areas that were identified are discussed below and include detector electronics, sample illumination, and tissue condition.

We found that the detector electronics were major contributors to system noise. We also determined that the resolution of the 16-bit A/D converter was not sufficient

to provide signal resolution at low concentrations. Redesigns have reduced the noise of detector electronics, and modifications are underway to increase the resolution of the A/D converter. These modifications are expected to improve measurement precision at low concentrations.

Investigation into the chosen method of sample illumination demonstrated that surface reflectance contributed substantially to a decrease in dynamic range of the instrument. The redesigned (fiber-optic) has an improved dynamic range and potential for increasing the amount of light collected, thus increasing the signal-to-noise ratio of the measurement.

The experimental results discussed above led to further study of the effect of tissue condition on the NIR spectral measurement. Chemical, structural, and physiological variations occur that produce dramatic changes in the optical properties of the tissue sample (22–24). Such variations, for example, may be related to hydration or changes in the volume fraction of blood in the tissue. Very tight, dry skin appears to scatter light more efficiently than soft, supple skin (25). These variations greatly affect the tissue volume sampled and the depth of light penetration, thus affecting the accuracy of the blood glucose measurement. Studies are underway to investigate the possible modeling of spectral variance resulting from changes in skin condition within and between subjects.

#### GLUCOSE TOLERANCE TESTING

For each subject, the data from three of the five OGTT experiments were selected for analysis based on the

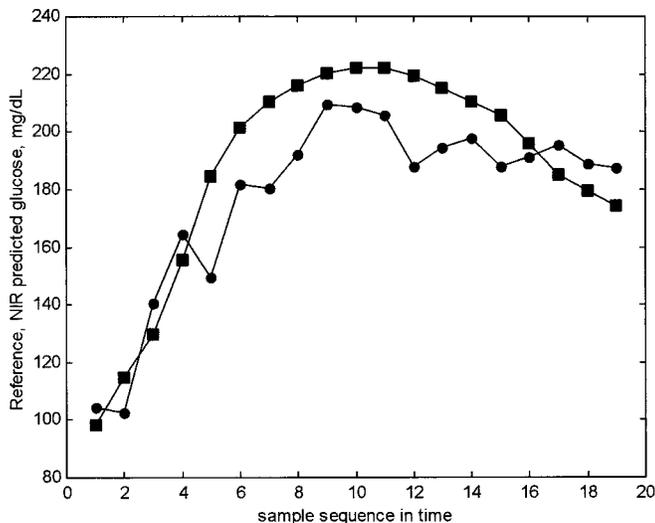


Fig. 5. Reference and prediction time profiles for subject 1 tolerance testing.

■, reference YSI glucose; ●, NIR-predicted glucose. To convert results to mmol/L, divide by 18.

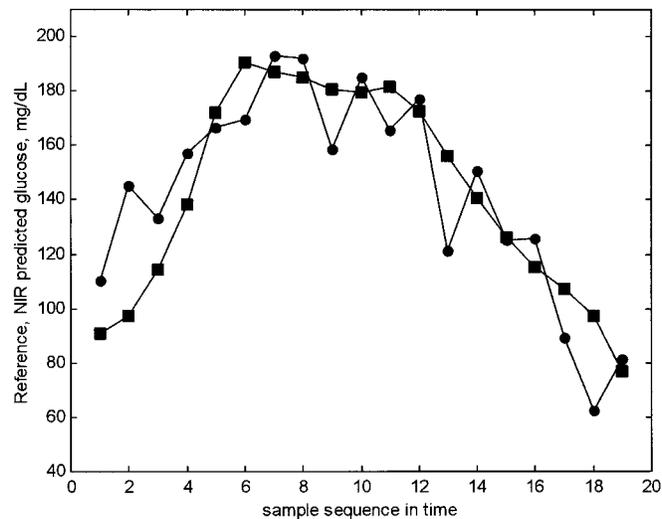


Fig. 6. Reference and prediction time profiles for subject 2 tolerance testing.

■, reference YSI glucose; ●, NIR-predicted glucose. To convert results to mmol/L, divide by 18.

criteria that instrument temperature variation had to be  $<1.5^{\circ}\text{C}$  over the experiment and the reference blood glucose concentrations had to have an insignificant correlation with skin temperature. The selection criteria were motivated by the potential for ancillary correlations between blood glucose concentrations and instrument drift or skin temperature (17). Before processing, data were rejected solely on the basis of the observed temperature variation of the skin and the internal instrument temperature.

The data associated with each subject were separated into independent calibration and test sets by selecting the data from two of the days for calibration and the remaining day for validation. However, the selection of calibration and test sets was complicated by the strong tendency of scans to cluster according to the day of measurement. Clustering of spectra is believed to be related to variations in skin hydration, skin surface smoothness, and day-to-day variations in skin surface temperature. The selection of the calibration set data was directed toward bounding the test set data in the multivariate calibration space. Spectral similarity was examined using hierarchical and K-means clustering (26). Failure to consider this aspect invariably led to dramatic increases in prediction error.

Optimization of the wavelength regions was performed separately for each subject. Although the resulting ranges varied only slightly between subjects, as shown below, the necessity of this step implies an inability to produce a universal calibration applicable to all subjects with the present measurement system. Therefore, future efforts must be directed toward instrument standardization and calibration transfer to ensure the potential for widespread application.

Subject 1 was calibrated using spectra collected on days 1 and 3, and predictions were made from the spectra

taken on day 2. The calibration model was developed using PLS (with eight scores) and spectral regions that included the second overtone at 1100–1300 nm, the first overtone at 1660–1770 nm, and the combination band region at 2050–2350 nm. Predictions produced a SEC of 1.1 mmol/L (19 mg/dL) and a SEP of 1.0 mmol/L (18 mg/dL). The tolerance test sample number for day 2 vs the YSI reference values and the predicted glucose are plotted in Fig. 5.

To determine whether instrument drift caused a secondary correlation to blood glucose concentrations, a calibration model was developed using only the reflectance reference spectra. Each reference spectrum was assigned the glucose value of the subject at that point in time when the reference spectrum was collected. The same spectral ranges defined above were used to calibrate the reference spectra and led to a SEP of 7.6 mmol/L (138 mg/dL). This result combined with the SD of 2.2 mmol/L (39 mg/dL) in the YSI glucose reference values implies that instrument-related variation, such as baseline drift, did not contribute to the statistical significance of the glucose predictions. Here, the unusually high SEP obtained using reference scans was traced to differences in the correlation of the instrument drift in the calibration and test sets. The reference spectra used for calibration were positively correlated to glucose, but the test set reference spectra were negatively correlated with glucose, leading to very poor prediction by the calibration model based on reference spectra.

Subject 2 was also calibrated using spectra collected on days 1 and 3 with the predictions being generated from spectra taken on day 2. The spectral regions used included the second overtone at 1130–1270 nm, the first overtone at 1660–1830 nm, and the combination band region at 2050–2360 nm. A calibration model was developed using PLS

with six scores and produced a SEC of 1.2 mmol/L (22 mg/dL) and a SEP of 1.1 mmol/L (20 mg/dL). The tolerance test sample number for day 2 vs the YSI glucose reference values and the glucose predicted by NIR are plotted in Fig. 6. The most accurate calibrations developed using the reference scans and the same wavelength ranges yielded a SEP of 2.4 mmol/L (44 mg/dL). This result combined with the SD in glucose reference values of 2.4 mmol/L (44 mg/dL) can be used to infer that instrument drift was not a significant factor in developing the glucose calibration.

The calibrations developed on subject 3 utilized spectra collected on days 2 and 3, and predictions were performed using the spectra acquired on day 4. The spectral regions used included the second overtone at 1100–1270 nm, the first overtone at 1670–1830 nm, and the combination band region at 2160–2340 nm. A PLS calibration model with a SEC of 1.0 mmol/L (18 mg/dL) and a SEP of 0.97 mmol/L (17 mg/dL) was developed using 12 scores. The tolerance test sample number for day 2 vs the YSI reference values and the glucose predictions are plotted in Fig. 7. The best predictive calibrations developed using the reflectance calibrator spectra and the same wavelength ranges yielded a SEP of 2.3 mmol/L (41 mg/dL), and the SD in glucose reference values was 1.9 mmol/L (34 mg/dL).

An examination of the correlation of some major blood analytes and other sources of false correlation with glucose during each of the three tolerance studies is given in Table 1. The modest correlation of glucose with those variables can be used to eliminate these particular variables as sources of false correlation. Instrument drift can also be eliminated as a source of false correlation given the poor predictive performance of calibrations devel-

**Table 1. Results from a correlation analysis between blood glucose concentration and other variables.**

Analyte	$r^2$		
	Subject 1	Subject 2	Subject 3
Elapsed time	0.0387	0.0258	0.0395
Skin temperature	0.0088	0.0446	0.0222
Hemoglobin	0.0363	0.0253	0.0122
Hematocrit	0.0076	0.0848	0.0247
Total protein	0.0745	0.0614	0.0328
Albumin	0.007	0.0466	0.0328
Total bilirubin	0.0019	0.118	0.2219
Triglycerides	0.164	0.1231	0.0526
Cholesterol	0.1213	0.0548	0.0712

oped using reference scans collected with each noninvasive scan.

### Conclusions

The results reported here lead to cautious optimism that noninvasive glucose measurement using NIR spectroscopy in the 1050–2450 nm range is possible. The limited success (three of seven subjects) obtained from the single-patient studies demonstrated the need for better understanding of tissue optics and how long-wavelength NIR light propagates through tissue. The single-patient studies also motivated the need for larger data sets on individual subjects and led to replacement of the lens-based optical interface with a fiber-optic system and the addition of an anatomically designed arm support. The use of the new sampling interface and the arm support substantially decreased sampling variation, although variability in tissue sampling remains an issue.

The glucose tolerance test data have been used to develop models that led to accurate blood glucose prediction on spectra collected on alternate days (three of three subjects) using independent calibration and test set data. Sampling errors inherent to the in vivo measurement and the relatively small glucose signal prevent the extraction of direct evidence that glucose was measured in this study. In addition, the complexity of the NIR spectrum makes it difficult to identify or extract unique glucose absorbances from the NIR in vivo spectrum, which would ultimately be required to claim a direct measurement of blood glucose. However, correlation between glucose measured during a tolerance test and the associated spectral response has been established and used to predict in vivo blood glucose concentrations during a tolerance test. The common sources of false correlation, such as instrument drift and major blood analytes, have been eliminated as contributors to the prediction results. The potential for correlation to other unmeasured sources of variation (e.g., physiological factors) has not fully been addressed.

The OGTT experiments also permitted a unique opportunity to gain a better understanding of experimental constraints, the impact of physiological variations on the

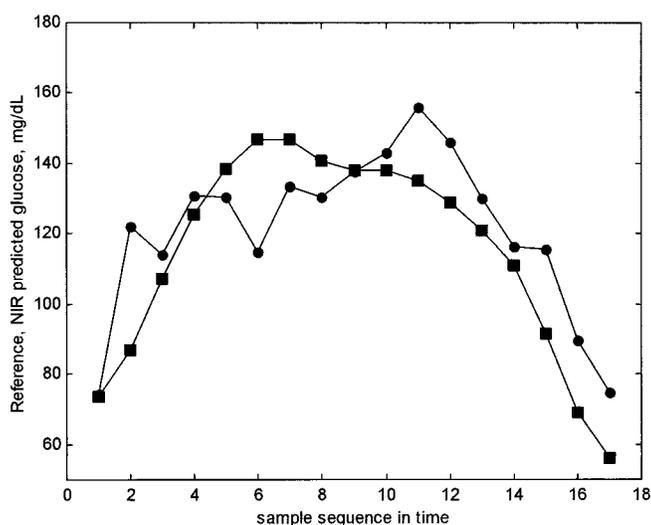


Fig. 7. Reference and prediction time profiles for subject 3 tolerance testing.

■, reference YSI glucose; ●, NIR-predicted glucose. To convert results to mmol/L, divide by 18.

noninvasive measurement, and the response of the skin to the measurement process. It has become evident that the need to control the ambient conditions of the experiment room and the instrument temperature complicates the already difficult task of making an effective *in vivo* measurement. Furthermore, the difficulty of reproducing noninvasive measurements in a way that ensures consistency in the weighted sampling of the tissue volume elements that contribute to the spectrum cannot be overstated. Many factors, including variations in skin surface roughness, hydration of the skin surface and underlying tissue, the effect of tissue displacement (contact pressure) by the probe interface, and variations in skin temperature, can contribute to significant changes in the sampling of the tissue volume elements. Variables that are internal to the tissue sample may not be controllable, but their impact on the measurement must be compensated.

The use of OGTTs to study the development of robust calibration models will undoubtedly involve a more detailed study of human physiological responses to glucose tolerance testing. Further study of tolerance tests is expected to lead to refined strategies for *in vivo* measurement. Alternative test strategies must also be studied to address the limitations of calibrations based on glucose tolerance testing. Studies are currently underway to examine additional means of reducing sampling variation and to address the sensitivity of the measurement to ambient conditions.

### References

1. Diabetes statistics. National Institutes of Health Publication No. 98-3926. Bethesda, MD: National Institutes of Health, 1997.
2. The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med* 1993;329:977–86.
3. Robinson MR, Eaton RP, Haaland DM, Keep GW, Thomas EV, Stalled BR, Robinson PL. Non-invasive glucose monitoring in diabetic patients: a preliminary evaluation. *Clin Chem* 1992;38:1618–22.
4. Ward KJ, Haaland DM, Robinson MR, Eaton RP. Postprandial blood-glucose determination by quantitative mid-infrared spectroscopy. *Appl Spectrosc* 1992;46:959–65.
5. Heise HM, Marbach R, Koschinsky TH, Gries FA. Non-invasive blood glucose sensors based on near-infrared spectroscopy. *Artif Org* 1994;18:439–47.
6. Marbach R, Koschinsky TH, Gries FA, Heise HM. Non-invasive glucose assay by near-infrared diffuse reflectance spectroscopy of the human inner lip. *Appl Spectrosc* 1993;47:875–81.
7. Heise HM, Marbach R. Effect of data pretreatment on the non-invasive blood glucose measurement by diffuse reflectance near-IR spectroscopy. *SPIE Proc* 1994;2089:114–5.
8. Jagemann KU, Fischbacher C, Danzer K, Muller UA, Mertes B. Application of near-infrared spectroscopy for non-invasive determination of blood/tissue glucose using neural network. *Z Phys Chem* 1995;191S:179–90.
9. Fischbacher C, Jagemann KU, Danzer K, Muller UA, Papenkrodt L, Schuler J. Enhancing calibration models for non-invasive near-infrared spectroscopic blood glucose determinations. *Fresenius J Anal Chem* 1997;359:78–82.
10. Danzer K, Fischbacher CH, Jagemann KU, KJ Reichelt. Near-infrared diffuse reflection spectroscopy for non-invasive blood-glucose monitoring. *LEOS Newslett* 1998;12:9–11.
11. Burmeister J, Arnold MA, Small G. Human non-invasive measurement of glucose using near infrared spectroscopy [Abstract]. *Pittcon* 1998, March 1–5, New Orleans, LA, 1998.
12. Muller UA, Mertes B, Fischbacher CF, Jagemann KU, Danzer K. Non-invasive blood glucose monitoring by means of new infrared spectroscopic methods for improving the reliability of the calibration models. *Int J Artif Organs* 1997;20:285–90.
13. Marbach R, Heise HM. Optical diffuse reflectance accessory for measurements of skin tissue by near-infrared spectroscopy. *Appl Optics* 1995;34:610–21.
14. Khalil OS. Spectroscopic and clinical aspects of noninvasive glucose measurements. *Clin Chem* 1999;45:165–77.
15. Arnold MA, Small GW. Determination of physiological level of glucose in an aqueous matrix with digitally filtered Fourier transform near-infrared spectra. *Anal Chem* 1990;62:1457–64.
16. Pan S, Chung H, Arnold MA. Near-infrared spectroscopic measurement of physiological glucose levels in variable matrices of protein and triglycerides. *Anal Chem* 1996;68:1124–35.
17. Arnold MA, Burmeister J, Small GW. Phantom glucose calibration models from simulated non-invasive human near-infrared spectra. *Anal Chem* 1998;70:1773–81.
18. Martens H, Naes T. *Multivariate calibration*. New York: John Wiley & Sons, 1989:419 pp.
19. Geladi P, Kowalski BR. Partial least-squares regression: a tutorial. *Anal Chim Acta* 1986;185:1–17.
20. Geladi P, McDougall D, Martens H. Linearization and scatter-correction for near-infrared reflectance spectra of meat. *Appl Spectrosc* 1985;39:491–500.
21. Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem* 1964;36:1627–39.
22. Anderson RR, Parrish JA. The optics of human skin. *J Investig Dermatol* 1981;77:13–9.
23. Cheong WF, Prael SA, Welch AJ. A review of the optical properties of biological tissues. *IEEE J Quantum Electron* 1990;26:2166–85.
24. Van Gemert MJC, Jacques SL, Sterenborg HJCM, Star WM. Skin optics. *IEEE Trans Biomed Eng* 1989;36:1146–54.
25. Martin K.A. Direct measurement of moisture in skin by NIR spectroscopy. *J Soc Cosmet Chem* 1993;44:249–61.
26. Massart DL, Van deginste BGM, Deming SN, Michotte Y, Kaufman L. *Chemometrics: a textbook*. Amsterdam, The Netherlands: Elsevier Science, 1988:488 pp.
27. Hazen KH. Glucose determination in biological matrices using near-infrared spectroscopy [PhD Thesis]. Iowa City, IA: University of Iowa, 1995:315 pp.