oninvasive Blood Glucose Assay by Near-Infrared iffuse Reflectance Spectroscopy of the Human Inner Lip*

R. MARBACH, TH. KOSCHINSKY, F. A. GRIES, and H. M. HEISE†

Institut für Spektrochemie und angewandte Spektroskopie, Bunsen-Kirchhoff-Str. 11, W-4600 Dortmund 1, Germany (R.M., H.M.H.); and Diabetes-Forschungsinstitut, Auf'm Hennekamp 65, W-4000 Düsseldorf 1, Germany (Th.K., F.A.G)

Near-infrared (NIR) spectra of the human inner lip were obtained by using a special optimized accessory for diffuse reflectance measurements. The partial-least squares (PLS) multivariate calibration algorithm was applied for linear regression of the spectral data between 9000 and 5500 cm^{-1} (λ = 1.1-1.8 μm) against blood glucose concentrations determined by a standard clinical enzymatic method. Calibration experiments with a single person were carried out under varying conditions, as well as with a population of 133 different patients, with capillary and venous blood glucose concentration values provided. A genuine correlation between the blood glucose concentrations and the NIR-spectra can be proven. A time lag of about 10 min for the glucose concentration in the spectroscopically probed tissue volume vs. the capillary concentration can be estimated. Mean-square prediction errors obtained by cross-validation were in the range of 45 to 55 mg/dL. An analysis of different variance factors showed that the major contribution to the average prediction uncertainty was due to the reduced measurement reproducibility, i.e., variations in lip position and contact pressure. The results demonstrate the feasibility of using diffuse reflectance NIR-spectroscopy for the noninvasive measurement of blood glucose.

Index Headings: **Glucose; Near-infrared spectroscopy; Diffuse reflectance; Noninvasive monitoring; Multivariate calibration.**

INTRODUCTION

Blood glucose is an important parameter in medical diagnostics. More than 10% of the assays in an average clinical laboratory are concerned with glucose, making it the most frequently determined analyte in the hospital. A great number of measurements are carried out by patients suffering from diabetes mellitus, where periodic self-monitoring of their blood glucose levels is necessary in order to adjust calorie intake and/or administer insulin injections. In healthy subjects the glucose concentration is controlled in a closed-loop system provided by the insulin-producing β-cells of the pancreas, which also sense the current glucose level in the blood, so that the insulin secretion rate is increased within minutes after food intake, thus fixing the mean blood glucose concentration to a normal value of about 0.1 weight percent of blood or 100 mg/dL; whereas for diabetic patients with missing or insufficient insulin production, external control of the actual blood glucose level is required.

The development of a noninvasive, portable device for glucose self-monitoring is desirable for several reasons. For effective control, a type I diabetic patient often requires four and sometimes even more tests per day, with blood usually being taken from the fingertips. The removal of this daily constraint would considerably improve the patient's quality of life and, in this way, also strengthen his or her cooperation in achieving optimal metabolic control. The consequence of a noninvasive measuring device, allowing frequent measurements, is the possibility of a quasi-closed regulatory system with a nearly continuous watch on the patient's glucose level. The general health, medical care, and economic advantages lie in the long-term reduction in secondary physiological problems through the improved regulation of the blood glucose levels by more regular testing. Another concern is the reduction in the enormous costs of the disposable test strips which are required by the home monitoring devices currently on the market. Alternative methods are under investigation, and progress on invasive approaches for *in vivo* monitoring of glucose has been recently reviewed.[1]

Different approaches have been proposed to achieve the goal of an IR spectroscopic noninvasive blood glucose method which have also resulted in several patents.[2-5] However, only recent progress in spectrometry and chemometrics has led to acceptable results for the *in vitro* analysis of glucose in blood plasma.[6,7] Whereas mid-infrared spectroscopy using, in general, the attenuated total reflectance (ATR) technique can be applied successfully for the multicomponent analysis of several blood substrates including glucose, it lacks the penetration depth into tissue for a noninvasive measurement, necessary for monitoring blood glucose concentration transcutaneously. Alternatively, near-infrared spectroscopy can provide the penetration depth, at the cost of specificity due to fewer informative absorption signatures in the spectral ranges chosen for quantitative analysis. We have provided a comparison of the analytical performances for the use of Fourier transform spectroscopic data from either the long-wavelength NIR range or an optimized interval of the fingerprint MIR range, setting the bench marks for an IR-spectroscopic noninvasive assay for blood glucose.[6-8]

NIR experiments using radiation of shorter wavelengths around 1 μm can be carried out successfully in transmittance, even with a tissue layer thickness around 1 cm, since the absorptivity of the main tissue constituent (i.e., water) is reduced, although the optical scattering in the nonhomogeneous tissue will increase with shorter wavelength. The largest penetration depths are obtained

Received 5 April 1993.

* This paper contains part of the doctoral thesis by R. Marbach.

† Author to whom correspondence should be sent.

0003-7028/93/4707-0875$2.00/0

within the so-called therapeutic window between 0.6 μm and 1.3 μm.[9] This spectral range was considered by Robinson *et al.*,[10] who made use of three different experimental setups including an FT spectrometer and grating monochromator with array detector, either with or without a fiber-optics accessory. As only oral glucose tolerance tests were performed during the preliminary investigation, the analytical relevance might be rather limited, because the pickup of spurious drift effects can influence the calibration results when a randomization of the measurements cannot be fulfilled. We experienced this problem when carrying out apparently sensitive ATR experiments with lip tissue of several diabetic probands using mid-infrared spectroscopy.

The study of the absorption spectrum of glucose initiated our *in vitro* transmittance and *in vivo* diffuse reflectance experiments. As the absorptivities of glucose around 1.6 μm (overtone bands of OH- and C-H stretching modes) are still of an acceptable magnitude, we decided to check the potential of this spectral range for quantitative multivariate calibration. To ascertain whether the penetration depth into tissue would be sufficient, we carried out Monte Carlo simulations, simulating the optical propagation of NIR photons in human tissue. By these means, different measurement conditions, either diffuse reflectance or transmittance, could be tested under varying tissue layer thicknesses. Due to energy limitations in the long-wavelength NIR, when transmittance experiments were performed, only a diffuse reflectance experiment turned out to be capable of reaching the penetration depth and the spectroscopic signal-to-noise ratios necessary for a noninvasive glucose assay.[7] To meet our spectroscopic requirements for a successful *in vivo* experiment, we constructed an optimized diffuse reflectance accessory.[7,11]

For a noninvasive measurement, the inner lip promised the best opportunities for realization, because the stratum corneum is missing and the mucous tissue found here is rather permeable for small molecules such as glucose; moreover, the lip is a homogenous tissue rich in capillary blood vessels and well thermostatted. In view of the pitfalls posed by spurious correlations to this extremely ill-conditioned measurement problem, different single-person calibrations as well as an experiment including 133 mainly diabetic patients were performed to evaluate the near-infrared spectrometric noninvasive assay, the results of which will be reported.

EXPERIMENTAL

Diffuse reflectance near-infrared (NIR) spectra of the human lower inner lip were obtained by having probands place their lips against the plane surface of a hemispherical immersion lens installed in a novel optical accessory and thermostatted at 37°C. The focus diameter of the illuminated sample area was about 2 mm. More details about the accessory have been published elsewhere.[7,11] The experiments were performed with the use of a Fourier transform IR spectrometer (Model IFS-66 from Bruker) equipped for NIR measurements (tungsten lamp, CaF_2 beamsplitter). The liquid nitrogen-cooled InSb detector with an element-diameter size of 4 mm was purchased from Infrared Associates (Suffolk, U.K.). A total of 1200 interferograms providing a spectral resolution of 32 cm^{-1} were averaged, and the resulting measurement time was about 1 min.

For the reference measurements, single-beam spectra were recorded with Spectralon reflectance standards from Labsphere (North Sutton, NH, U.S.A) pressed against the immersion lens of the accessory following careful cleaning. This material provides nearly ideal diffuse, Lambertian characteristics suitable for our reflectance application in the NIR spectral range. In the interval of 0.4 to 1.6 μm, the rather constant reflectivity of the white standard (SRM-99L-100C) is better than 0.98, and for longer wavelengths it decreases slightly ($R_{\lambda=1.8\,\mu m} = 0.976$, $R_{\lambda=2.0\,\mu m} = 0.957$, and $R_{\lambda=2.2\,\mu m} = 0.944$).[12] For the lip reflectance measurements, gray Spectralon materials possessing reflectivity values between 5% and 15% were used, so that a greater radiant power could be considered without saturating the detector signal or the ancillary amplification and digitization electronics. However, slight variations in reflectance were noticed when the diffuse reflectance standard material was pressed against the accessory lens, due to slight deviations in surface contact and changes in surface roughness. Much better reproducibility was achieved by using the Spectralon diffuse white standard drilled out cylindrically, so that a cavity resulted, providing a blurred image at the sample focus of the rotational ellipsoid used for collection of the diffusely reflected radiation. The radiant power reflected onto the detector was thereby reduced to a constant fractional rate required for adjusting the interferogram maximum. Best reproducibility for the reference measurements was obtained by this procedure.

Capillary blood samples of 20 μL for the glucose determination in the laboratory were taken by puncture of the fingertip with capillary pipettes from Brand (Wertheim, Germany). For the experiments with different patients from the clinical department of the Diabetes Research Institute, venous blood was also collected with the use of Monovettes from Sarstedt (Nürnbrecht, Germany). The hemolyzing reagent solution for the blood samples contained digitonin and maleimide. The blood glucose concentration was determined by a standard clinical enzymatic method (test combination Glucoquant Glucose from Boehringer Mannheim, Germany) with the use of hexokinase/G6P-DH,[13] which was available on an ACP 5040 analyzer from Eppendorf (Hamburg, Germany). Plasma samples from the venous blood charges were measured by a Boehringer Mannheim/Hitachi 704 analyzer by means of standard enzymatic methodology, providing concentration data for total cholesterol and triglycerides, which are of interest for judging the variation in composition of the sample population studied.

The partial least-squares (PLS) multivariate calibration algorithm was applied for a linear regression of the spectral data between 9000 and 5500 cm^{-1} (λ = 1.1–1.8 μm) against the probands' blood glucose concentration values. Mean-square prediction errors [$PRESS^{1/2} = (\Sigma(C_{ref,i}-C_{pred,i})^2/M)^{1/2}$ with M samples; PRESS = predicted error sum of squares] were estimated by cross-validation using the "leave one out" strategy. Programs for the input of spectral data, PLS calibration, and cross-validation were written in MATLAB (The Mathworks, South Natick, MA, U.S.A). Use was also made of the

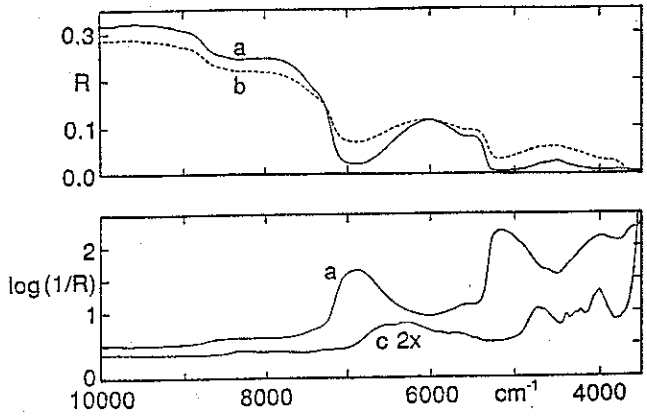

FIG. 1. Diffuse reflectance, resp. $\log(1/R)$ spectra for mucous lip tissue (trace *a*), for tongue tip tissue (trace *b*), and crystalline anhydrous glucose (trace *c*, enlarged for clarity).

signal processing toolbox. Further details to the calibration strategy including detection of outliers are described in Refs. 14 and 15.

RESULTS AND DISCUSSION

The diffuse reflectance accessory we used was particularly suited for measurement of bulky specimens. The high-throughput device was essential for the experiments planned. The refractive-index-matching immersion lens also promised reproducible sampling conditions, and the measurement time of 1 min was a compromise which could be tolerated by the patients. In the upper part of Fig. 1, two different reflectance spectra are shown; trace *a* gives a spectrum of the mucous tissue of the inner lip, whereas with trace *b* the spectrum of a tongue-tip is presented. Trace *b* is also an example of tissue not satisfactorily in contact with the immersion lens, so that Fresnel reflection contributes to the spectrum. The lower part of this figure gives the corresponding $\log(1/R)$ spectrum of the same lip tissue in comparison to a diffuse reflectance spectrum of crystalline glucose measured by the same device.

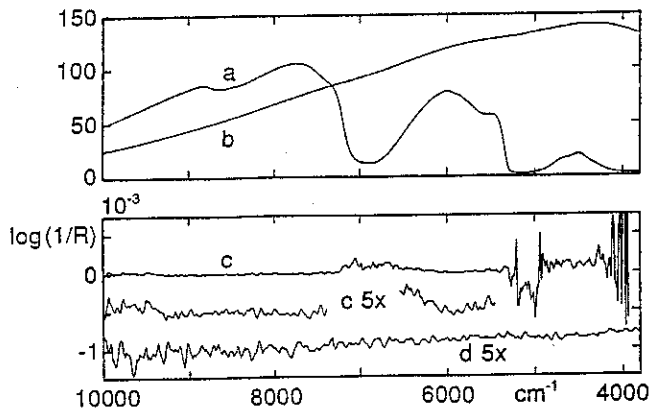

FIG. 2. Single-beam spectra of lip in arbitrary units (trace *a*) and a Spectralon gray standard with a reflectivity of about 0.13 (trace *b*) measured by the diffuse reflectance accessory (upper part); $\log(1/R)$ noise level estimated from two consecutive lip measurements (trace *c*) and single-beam measurements with the gray standard from above (trace *d*) (see lower part; the enlarged spectra are offset).

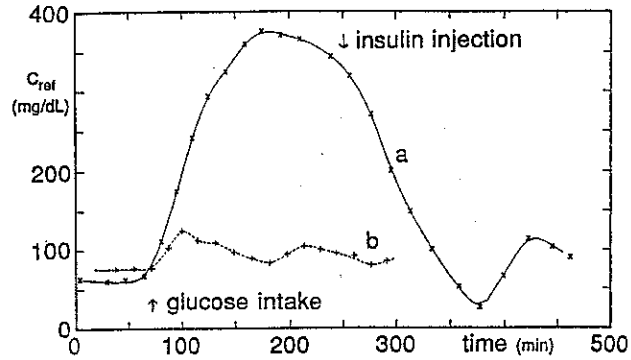

FIG. 3. Glucose time dependence for an oral glucose tolerance test (for details, see text) of a diabetic (solid curve) and a healthy test person (dashed curve); the same methodology was used for a calibration experiment with a single diabetic subject.

The energy limitations in the reflectance experiment for lip tissue are illustrated in the upper part of Fig. 2; trace *a* shows the single-beam spectrum of the FT spectrometer with lip tissue contacting the diffuse reflectance accessory, whereas for trace *b* a gray diffuse reflectance standard with a reflectivity of about 13% was used. In the lower part of Fig. 2, the corresponding $\log(1/R)$ noise levels for the use of two consecutive single-beam measurements either with lip tissue or the gray standard considered in trace *b* are displayed in traces *c* and *d*, respectively. The $\log(1/R)$ peak-to-peak noise level around 6000 cm^{-1} is smaller than $3 \cdot 10^{-5}$ absorbance units (AU). The glucose absorbance maximum at 6350 cm^{-1} measured for an aqueous solution (100 mg/dL) with a cell of 1 mm pathlength is about 10^{-4} AU, which underlines the ambitious goal of a noninvasive glucose assay.

First, experiments were performed with the use of oral glucose tolerance tests with supply of a sugar syrup (400 mL Dextro O.G.-T., Boehringer Mannheim, Mannheim, Germany), which is a mixture of mono- and oligosaccharides equivalent to 100 g anhydrous glucose after enzymatic cleavage. For the standardized test, three to six blood glucose determinations are usually carried out within the test duration of about three hours; the patient will start with an empty stomach. Figure 3 shows the response for a nondiabetic person (dashed curve), whereas the solid curve demonstrates the time dependence of the blood glucose concentration of a diabetic type I patient under these conditions, with the crosses representing the moments of blood sampling. The arrow near the glucose maximum indicates the time of an injection of regular insulin, leading to a diminished blood glucose concentration. For our experiments blood samples were taken from the test-person at time intervals of about 20 min. The concentration values obtained from these samples were interpolated with the use of a cubic spline function[16] providing the reference glucose values for three lip spectra recorded between two blood samplings made by puncture of the fingertip.

Multivariate calibration within a one-person experiment was carried out for a total of 133 spectral measurements taken during a two-day period with the blood glucose being monitored over 16 h. The blood glucose profiles were similar to Fig. 3; however, the insulin dosage was planned not to produce hypoglycemia, so that for the beginning and end of the test a normal glucose con-

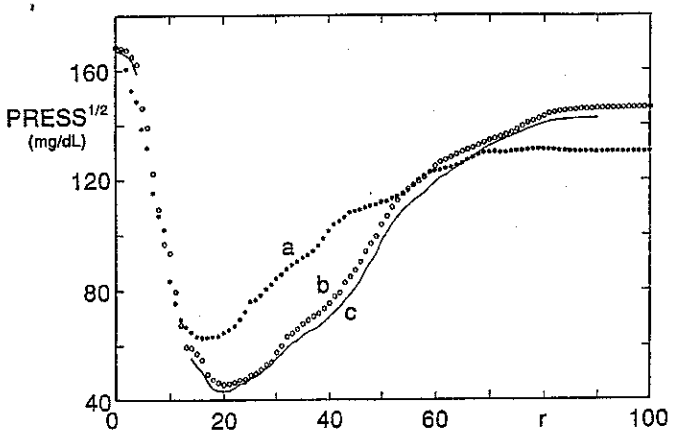

FIG. 4. PRESS$^{1/2}$ statistics of the population mean-square prediction error for capillary blood glucose vs. number of PLS factors chosen with calibration data collected during a one-person experiment using oral glucose tolerance testing; results are from calibrations with 132 spectra of spectral range from 8994 to 5477 cm^{-1} with a wavenumber spacing of 30.8 cm^{-1} using $\log(1/R)$ (trace *a*) and logarithmic single-beam spectra (trace *b*); results of trace *c* (solid curve) are given for low-pass time-filtered glucose concentration values (time constant T = 10 min).

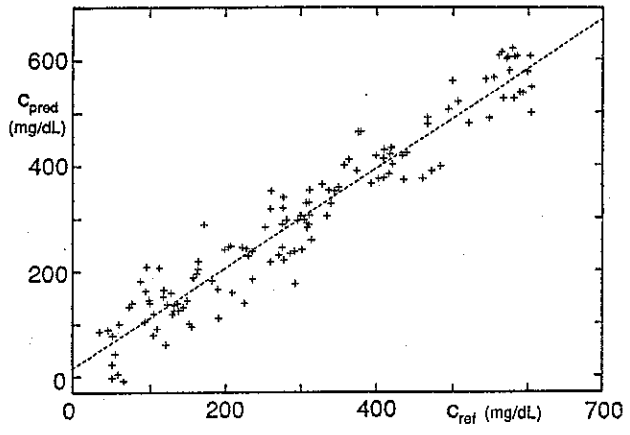

FIG. 5. Cross-validated predictions of glucose concentrations using the optimum PLS model of the noninvasive one-person experiment with oral glucose tolerance testing vs. interpolated glucose reference concentration values. Also given is the result of an *a posteriori* linear least-squares fit.

centration value of about 100 mg/dL resulted. In an effort to reduce the possible effect of spurious correlations, the sugar potion was split into two portions, so that intermediate plateaus, lasting for about one hour, appeared during the first day, with the result of a maximum glucose level of 420 mg/dL. For the second day, the total content of two syrup bottles was ingested, with the consequence that a maximum of 600 mg/dL was reached. The resulting reference concentration values are almost equally distributed between 30 mg/dL and 600 mg/dL, with an average value of $\bar{c}_{pop}$ = 301 mg/dL and a standard deviation of $\hat{\sigma}_{pop}$ = 167 mg/dL.

At first, calibration modeling was performed with the use of $\log(1/R)$ signals in an optimized spectral interval between 9000 and 5475 cm^{-1} with a wavenumber spacing of 30.8 cm^{-1}, giving N = 115 spectral data points. We noticed that spectral data pretreatment such as baseline correction, smoothing, or taking spectral derivatives did not improve calibration results. The rank-dependent average prediction errors (PRESS$^{1/2}$) are given in Fig. 4. It should be noted that the rank of the calibration matrices used here is equivalent to the number of PLS factors chosen for modeling. As a gray standard was considered (see also Experimental section) for the reference measurement, much prediction variance could be attributed to this spectroscopic step when the $\log(1/R)$ signals for calibration were taken into account.

This point becomes even more evident when we look at the results for multivariate calibrations presented in the same figure with the logarithm of the single-beam lip spectra being considered; these can be used, because the single-beam characteristic background is rather featureless and stable, although we experienced some variations in signal amplitude during the course of this experiment. Another advantage of the single-beam spectra, as compared to the $\log(1/R)$ spectra, is that the noise level will be reduced by a slightly wavenumber-dependent factor of about $\sqrt{2}$ (see also Fig. 2). A minimum of the average prediction error occurs at rank r_{opt} = 20 with PRESS$^{1/2}$ = 45.6 mg/dL. In Fig. 5 a scatter plot of the predicted (cross-validated) concentrations vs. reference values is provided for the optimum model with respect to logarithmic single-beam data. The precision, especially for the lower concentration range, is rather poor; however, at high levels this limitation could be tolerated for a self-monitoring device for blood glucose. The *a posteriori* least-squares fit of the independent predictions vs. reference values gave c_{pred} = 17.0 + 0.945 c_{ref} with a coefficient of determination of R^2 = 0.92.

The tissue volume spectroscopically probed contains most of its glucose in the interstitial fluid of the intercellular space. Since the time dependence of the glucose concentration is known for this type of experiment (see Fig. 3), the time delay between the glucose concentrations in blood and probed mucous tissue, as expected for transport processes (e.g., by diffusion), can be simulated. A characterization of the glucose level in subcutaneous human tissue by microdialysis had been published recently.[17] For our delay simulation, the reference glucose signal, as approximated by the spline function mentioned above, was equidistantly sampled at intervals of one minute. A digital low-pass filtering was carried out by application of an impulse-invariant designed Butterworth filter of first order $y(n) = (1 - e^{-(1/T)})\ x(n) + e^{-(1/T)}\ y(n - 1)$ with a time constant of T. As shown in Fig. 4, the filtering of the time-dependent reference concentration improves the prediction results, as the PRESS$^{1/2}$ value decreases to 43.0 mg/dL for a time constant of T = 10 min. Extension to T = 20 min will cause a marginal deterioration in the result again (PRESS$^{1/2}$ = 43.7 mg/dL at a calibration model rank r_{opt} = 19). The plausibility of our results is in agreement with the findings of Jansson *et al.*,[17] who found similar glucose levels under steady-state conditions, whereas for rapid changes a delay of 8 min for the glucose concentration in the subcutaneous interstitial space compared to that in venous blood could be experimentally determined.

For support of the results obtained by using oral glucose tolerance tests, a randomized one-person experiment with the same type I diabetic patient was carried out. During two weeks, three different, but rather constant, glucose levels (low, medium and high) were daily

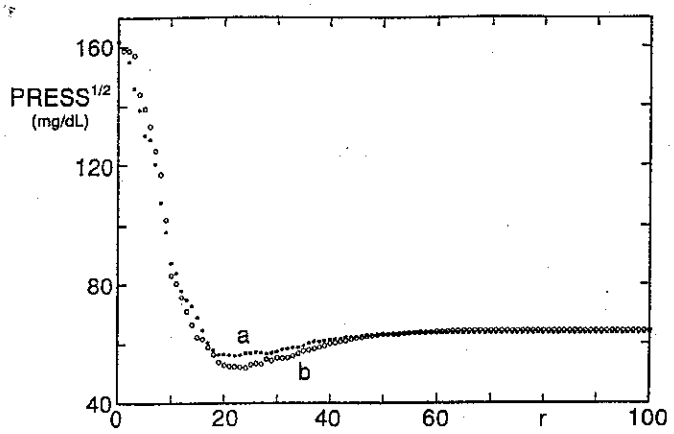

FIG. 6. PRESS$^{1/2}$ statistics of the population average prediction errors for capillary blood glucose using a one-person experiment with randomized sampling. Calibration was carried out with 216 spectra (same spectral range as for Fig. 4). Results are given for $\log(1/R)$ (trace *a*) and logarithmic single-beam data (trace *b*), respectively.

attempted, the sequence of which had been selected at random. For a low-level measurement the blood glucose concentration value was below 150 mg/dL; a high level was programmed to reach values greater than 400 mg/dL, whereas for a medium level any values in between were allowed. The following measurement cycles lasting up to 40 min were carried out in the morning, at noon, and in the evening before meal intake. Three capillary blood samples were taken with three lip spectra, each recorded in between two following samplings. Another lip spectrum was recorded before the first and after the last sampling step, so that a total of 8 spectra in each sequence were collected for calibration. For improving precision and reliability of the reference glucose concentration values, these were determined in duplicate in the clinical laboratory. The actual reference glucose value was calculated with the use of a spline fit approximation and interpolation at the average time of the spectral measurement. The variation in glucose concentration during one session was typically below 40 mg/dL, with a maximum of double that value. At the end of this experiment 219 diffuse reflectance spectra were available; the corresponding glucose concentration values were in the range between 34 and 590 mg/dL ($\bar{c}_{pop}$ = 268.6 mg/dL and $\hat{\sigma}_{pop}$ = 161.8 mg/dL).

The calibration was optimized with the use of spectral data from the same interval that was chosen for the first experiment. Two lip spectra, which were conspicuously different by visual inspection, were deleted because of high leverage values, and a further one was deleted because of a significant Cook's distance value. The important calibration results are illustrated in Fig. 6, showing the rank-dependent PRESS$^{1/2}$ values for $\log(1/R)$ and logarithmic single-beam data. The improvement observed by using single-beam spectra over $\log(1/R)$ spectra is not as pronounced when the drilled-out Spectralon white standard is considered as reference for the latter spectral data, due to much improved spectroscopic measurement reproducibility in comparison to that of the initial study (see Experimental section). The optimum results for both models were: PRESS$^{1/2}$ = 55.9 mg/dL and PRESS$^{1/2}$ = 51.9 mg/dL. The calibration results with single-beam spectra still compare favorably with those from the $\log(1/R)$ spectra, although the duration of this experiment was two weeks, so that a deterioration in the prediction performance is to be expected, in comparison to results from the two-day experiment using oral glucose tolerance tests.

An estimate of the pure spectroscopic contribution to the overall prediction variance experienced for the one-person experiment with random blood glucose plateaus can be obtained by ratioing two consecutive single-beam reference spectra recorded before each lip spectrum. These baseline spectra are a reasonable approximation of the spectroscopic random noise, as well as for systematic instrumental drifting; 34 baseline spectra were deleted, 26 because of excessive time differences. The residual 184 baselines show a maximum deviation of $\pm 2 \cdot 2 \cdot 10^{-3}$ AU from the ideal zero line within the interval of 9000 to 5500 cm^{-1}. By scalar multiplication with the calibration regression vector, these spectral features provide an estimate of the pure spectroscopic prediction error under these experimental conditions. For the population of baselines studied, the spectroscopic prediction errors were normally distributed between ± 60 mg/dL with an average value being insignificant from zero and a standard deviation of $\hat{\sigma}_{spec}$ = 25.0 mg/dL. This result can be verified by the different optimum PRESS$^{1/2}$ values in Fig. 6, because the model using the $\log(1/R)$ values contains approximately double the variance from spectroscopic noise compared to that using single-beam data.

The standard deviation for the average reference concentration error can be estimated for this experiment to be around $\hat{\sigma}_{ref}$ = 15 mg/dL with the use of results from control measurements with reference sera.[6,7] For the uncorrelated error contributions we define

$$\text{PRESS}^{1/2} = \sqrt{\hat{\sigma}^2_{ref} + \hat{\sigma}^2_{spec} + \hat{\sigma}^2_{bio}}. \qquad (1)$$

The results from the variance decomposition according to Eq. 1 demonstrate that the residual "biological" error, defined with the last term, is on the average $\hat{\sigma}_{bio}$ = 47 mg/dL for the use of either single-beam or $\log(1/R)$ data. It is mainly the result of the unsatisfactory reproducibility of the lip measurement, i.e., lip positioning and contact pressure. This observation provides clues for further assay improvement.

Additionally, a large population of patients was studied, which also provided spectroscopic evidence for the biological variation of the lip tissue considered for the *in vivo* spectroscopy. For this experiment study the FT spectrometer was set up in a test room of the Diabetes Research Institute. A population of 133 randomly selected subjects was at our disposal, most of whom were diabetics; each subject was tested once, with the exception of six persons providing two- and one-person three-measurement sessions.

The session data are characterized by 73 male and 68 female test persons, also distinguishable between 45 smokers and 96 nonsmokers. Also recorded were sex, age, and date of last food intake. The data characterizing the population studied are summarized in Table I, which also provides data for total cholesterol and triglycerides, underlining the physiological variation in blood composition of the test persons studied.

Usually, three lip measurements of each person were

TABLE I. Concentration range C_{min}–C_{max}, population average values $\bar{c}_{pop}$ and standard deviations $\hat{\sigma}_{pop}$ for the population of 141 lip measurements (133 different patients).

	Age (years)	Venous glucose (mg/dL)	Capillary glucose (mg/dL)	Total cholesterol (mg/dL)	Triglycerides (mg/dL)
c_{min}–c_{max}	13–79	37–401	35–417	112–330	39–1360
$\bar{c}_{pop}$	44.9	148.7	154.2	210.8	174.5
$\hat{\sigma}_{pop}$	16.5	75.3	76.7	47.6	163.2

recorded with, as reference, a Spectralon gray standard material with a reflectivity of 5%. Disinfection of the accessory part in contact with the test person was carried out with the disinfectant "Spitacid" after each measurement. The total recording time was approximately 20 min, so that the variance contribution to the prediction error from the reference concentrations is partly due to the method uncertainties, but also influenced by the time delay between blood sampling and spectrum recording in view of the changing blood glucose.

The measurement conditions were not as favorable as during the one-person experiments due to frequent disruptions from patients and hospital staff experienced during this measurement effort. Spectrometer stability was reduced so that the reference measurements with the use of the Spectralon standard were vital for deriving the spectral calibration data. Purging quality of the spectrometer was slightly reduced, so that the spectral range between 7420 cm^{-1} and 7000 cm^{-1} was excluded from the calibration because of greater spectral variations arising from atmospheric water vapor absorptions.

From the total of 390 lip spectra recorded, 9 spectra were identified as outliers. Three extreme cases had been noticed already by visual inspection and six due to significant Cook's distance values (three of these from a young female person, possibly due to lipstick which should have been removed prior to the measurement). Calibration modeling was done in terms of only $\log(1/R)$ values. The limitation of the spectral data to the lower interval limit of 5925 cm^{-1} is based on the greater spectral variations found in the range below, which are illustrated in Fig. 7. The upper part of this figure shows the difference spectrum of the average of all 182 spectra from female individuals to that of the 199 recorded from male persons. For comparison, the lower part shows two diffuse reflectance spectra of minced beef; one sample contained a greater percentage of fat (see trace *c*), which is evident from the more intense doublet absorption band below 5900 cm^{-1}. This absorption doublet was not detectable for many persons and could be found more often for female individuals, which is underlined by the difference spectrum in Fig. 7 (trace *a*). The effect was not reproducible for different measurements of a single person, indicating local differences in the fat content of the tissue spectroscopically probed. A discrimination analysis for these spectral features based on factor analysis was not successful since the effects are rather small.

The important calibration results are summarized in Fig. 8, which provides the rank dependence of the mean-square prediction errors obtained with either capillary or venous glucose concentration values; the optimum PRESS$^{1/2}$ values are 57.9 mg/dL and 55.4 mg/dL, respectively. The better prediction performance of the calibration model for venous glucose is another argument for the inherent delay of the time-dependent glucose concentration due to transport processes into the probed tissue volume discussed above, as a similar mechanism is responsible for the time lag between glucose in venous and capillary blood. An extension of the spectral calibration data using the range up to 11,000 cm^{-1} resulted in no significant improvement. Obviously, the information gain compared to the adverse effect of additional noise is negligible.

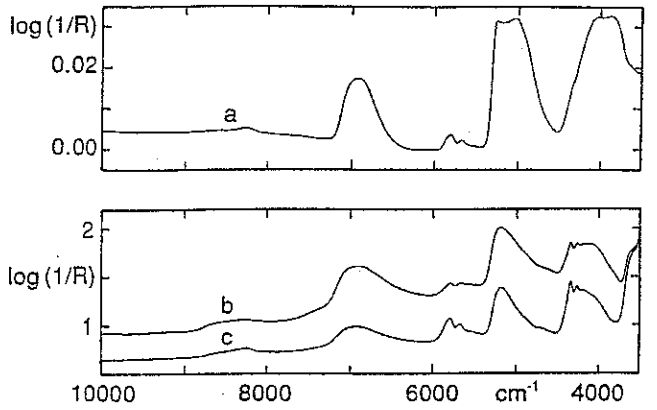

FIG. 7. Log(1/*R*) difference spectrum between the average spectrum from 182 female and from 199 male lip tissue spectra (trace *a*); log(1/*R*) spectra of lean minced beef (trace *b*) and of meat with greater fat content (trace *c*, see lower part).

We are aware of the rather small improvements when the rank-dependence of the PRESS$^{1/2}$ values are studied. It would have been advantageous to deal with a population with greater variation in the blood glucose reference concentrations. For the *in vitro* glucose assay described in Refs. 6–8, the standard deviation for the sample population was about 90 mg/dL obtained with diabetic patients, whereas the standard deviation experienced with patients hospitalized in a general hospital is only about 45 mg/dL.[15] This consideration sets limits for a random

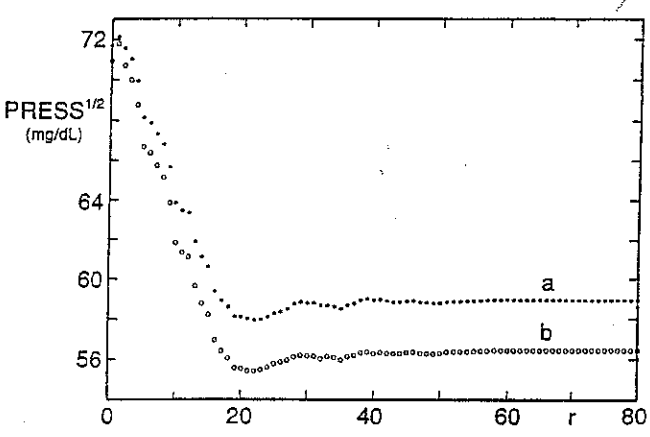

FIG. 8. PRESS$^{1/2}$ statistics of the population mean-square prediction error vs. number of PLS factors chosen; results are for capillary reference glucose concentration values (trace *a*) and for venous values (trace *b*). Calibration was carried out using 381 $\log(1/R)$ spectra from 133 different patients. Spectral data were considered in the interval 8994 to 7421 cm^{-1} and 7004 to 5924 cm^{-1} with a wavenumber spacing of 30.8 cm^{-1}.

multiple-person calibration experiment. However, the use of the large number of standards allows for a statistically sound assertion of the correlation of glucose with the NIR spectra measured, which can be underlined by the small increase in the PRESS$^{1/2}$ values for greater model complexity.

It should be noted that the calibration results are achieved by using a broad spectral range including the water overtone absorption band around 7000 cm^{-1}. For the *in vitro* analysis of blood plasma samples, the water regions with greater absorption had to be excluded to provide a satisfactory calibration result on the whole.[7,8] Obviously, there is a significant difference between the temperature dependence of "free" water and "bound" water in the tissue. For the latter state, the water molecules are mostly hydrogen bonded to proteins, so that temperature sensitivity is reduced. In the literature, this dependence, to our knowledge, has not yet quantitatively been described. A note can be found pointing out that the absorption peak for "free" water at 2.94 μm is shifted to 3.05 μm for tissue water.[18]

CONCLUSION

A genuine correlation between capillary blood glucose concentration and the NIR spectrum has been proven, and a time lag of about 10 min of the glucose concentration in the spectroscopically probed tissue volume has been established. The mean-square prediction error (PRESS$^{1/2}$) was estimated between 45 and 55 mg/dL with the use of the "leave-one-out" strategy of cross-validation. An analysis of error variance showed this limitation to be due to the reduced measurement reproducibility (i.e., variations in lip position and contact pressure), whereas the signal-to-noise ratio would have allowed for a PRESS$^{1/2}$ value of approximately 25 mg/dL for $\log(1/R)$ signals. Measurement irreproducibility occurring for single patients is of the same order of magnitude as those between different individuals.

The results show that NIR spectroscopy is feasible for the noninvasive measurement of blood glucose and may be used for building a self-monitoring device usable by diabetic patients at home. However, for such a device, sampling reproducibility is of paramount importance and has to be further improved if the prediction performance is to achieve clinical acceptance. A different approach using a fiber-optic accessory may be worth realizing. Furthermore, the physiological variance of glucose in tissue has to be properly investigated before a final statement about the realization is possible.

ACKNOWLEDGMENTS

The authors are indebted to Prof. Dr. med. H. Reinauer and Dr. med. C. Niederau from the Abteilung für Klinische Biochemie des Diabetes-Forschungsinstituts for providing the analytical reference data. Financial support by the Deutsche Forschungsgemeinschaft, the Ministerium für Wissenschaft und Forschung des Landes Nordrhein-Westfalen, and the Bundesminister für Forschung und Technologie is gratefully acknowledged.

1. G. Reach and G. S. Wilson, Anal. Chem. **64,** 381A (1992).
2. C. Dähne, Spectrophotometric Apparatus for the Non-invasive Determination of Glucose in Body Tissues, European Patent EP O 160 768 B1 (date of patent May 3, 1989).
3. R. D. Rosenthal, L. N. Paynter, and L. H. Mackie, Non-invasive Measurement of Blood Glucose, U.S. Patent No. 5,086,229 (date of patent Feb. 4, 1992).
4. R. H. Russell, and J. W. Brasch, Sr., Non-invasive Determination of Glucose Concentration in Body of Patients, U.S. Patent No. 5,070,874 (date of patent Dec. 10, 1991).
5. M. R. Robinson, K. J. Ward, R. P. Eaton, and D. M. Haaland, Method of and Apparatus for Determining the Similarity of a Biological Analyte from a Model Constructed from Known Biological Fluids, U.S. Patent No. 4,975,581 (date of patent Dec. 4, 1990).
6. H. M. Heise, R. Marbach, Th. Koschinsky, and F. A. Gries, Appl. Spectrosc., paper submitted.
7. R. Marbach, Ph.D. Thesis, University of Dortmund (1993).
8. R. Marbach and H. M. Heise, in *Proceedings of the 8th International Conference on Fourier Transform Spectroscopy,* H. M. Heise, E. H. Korte, and H. W. Siesler, Eds., SPIE Vol. 1575 (SPIE, Bellingham, Washington, 1992), pp. 507–508.
9. B. C. Wilson and S. L. Jacques, IEEE J. Quantum Electron. **26,** 2186 (1990).
10. M. R. Robinson, R. P. Eaton, D. M. Haaland, G. W. Koepp, E. V. Thomas, B. R. Stallard, and P. L. Robinson, Clin. Chem. **38,** 1618 (1992).
11. R. Marbach and H. M. Heise, in *Proceedings of the 8th International Conference on Fourier Transform Spectroscopy,* H. M. Heise, E. H. Korte, and H. W. Siesler, Eds. SPIE Vol. 1575 (SPIE, Bellingham, Washington, 1992), pp. 288–289.
12. Catalogue data from Labsphere, North Sutton, New Hampshire.
13. F. H. Schmidt, Klin. Wschr. **39,** 1244 (1961).
14. R. Marbach and H. M. Heise, Chemometrics Int. Lab. Syst. **9,** 45 (1990).
15. H. M. Heise, R. Marbach, G. Janatsch, and J. D. Kruse-Jarres, Anal. Chem. **61,** 2009 (1989).
16. C. H. Reinsch, Numer. Math. **10,** 177 (1967).
17. P.-A. Jansson, J. Fowelin, U. Smith, and P. Lönnroth, Am. J. Physiol. Metab. **18,** E218 (1988).
18. G. Yoon, A. J. Welch, M. Motamedi, and M. C. J. van Germert, IEEE J. Quantum Electron. **23,** 1721 (1987).