

*Perspectives***Noninvasive Glucose Sensing****Mark A. Arnold\* and Gary W. Small***Department of Chemistry and Optical Science and Technology Center, University of Iowa, Iowa City, Iowa 52242*

**The ability to measure glucose noninvasively in human subjects is a major objective for many research groups. Success will revolutionize the treatment of diabetes by providing a means to improve glycemic control, thereby delaying the onset of the medical complications associated with this disease. This article focuses on the current state of the art and attempts to identify the principal areas of research necessary to advance the field. Two fundamentally different approaches are identified for the development of noninvasive glucose sensing technology. The indirect approach attempts to measure glucose on the basis of its effect on a secondary process. The direct approach is based on the unique chemical structure of the glucose molecule. Advances for each approach are limited by issues of selectivity. Several critical parameters are discussed for the direct approach, including issues related to the optical path length, wavelength range, dimensionality of the multivariate calibration model, net analyte signal, spectral variance, and assessment of the chemical basis of measurement selectivity. A set of publication standards is recommended as a means to enhance progress toward a successful noninvasive monitor.**

Noninvasive glucose sensing refers to the measurement of *in vivo* glucose concentrations with no direct contact between the transduction element and the representative biological fluid. The most common strategy is to probe a region of tissue with a selected beam of electromagnetic radiation and then extract the concentration of glucose from the resulting spectrum. This approach is noninvasive in the sense that a sample does not need to be collected or handled, and there is no need to place the sensing device in immediate contact with a body fluid. The challenge is to make an accurate measurement without the aid of selective chemical reagents or physical separations.

It is important to realize at the outset that, to date, no one has proven the ability to measure glucose noninvasively in either human subjects or animal models. Nothing published in the peer-reviewed literature or described in the patent literature is proven to measure glucose selectively from noninvasive analytical information. Although many papers and patents claim this ability, none is able to provide the level of proof necessary to establish such a complex analytical measurement. In fact, nearly all the published accounts are most certainly not measuring glucose directly. While some studies might be based on the measurement of indirect effects caused by changes in glucose concentrations, many others are likely based purely on chance correlations induced by inadequate experimental design. In addition to poor experimental design, the common problems, which run across the different putative noninvasive technologies, include the misuse of multivariate calibration methods, a lack of attention to physical and chemical parameters, and a disregard for instrumental details.

An observation frequently noted at scientific meetings and review panels is that noninvasive glucose sensing has been under development for more than 15 years without a resulting product. Some evaluators conclude that this lack of success indicates that such measurements are impossible and further development is a waste of time and resources. In fact, the inability to produce an instrument during this period is more an indication of the difficulty of noninvasive measurements as opposed to their feasibility. Put simply, accurate noninvasive analytical measurements are not simple. Results from the past can teach us valuable lessons and guide us into the future, however. In our view, results from the past strongly indicate that a purely Edisonian approach will never be successful. Otherwise, pioneering researchers such as Robert Rosenthal, Reese Robinson, Michael Heise, Ralf Marbach, and others would have solved the basic problems years ago and a noninvasive glucose monitor would be on the market today. Clearly, success demands a detailed understanding of the underlying physics, chemistry, physiology, and instrumentation required for accurate noninvasive clinical measurements.

\* Corresponding author. Phone: 319-335-1368. Fax: 319-353-1115. E-mail: mark-arnold@uiowa.edu.

A critical evaluation of the published literature demonstrates that noninvasive glucose measurements have significant promise. Omar Khalil has written two comprehensive reviews of this literature,<sup>1,2</sup> where the first spans the period from 1989 to 1998 and the second from 1998 to 2003. These reviews highlight the major progress in the field and critically assess the feasibility of measuring glucose noninvasively. As a whole, published accounts in the peer-reviewed literature are encouraging that with further development, glucose-specific information can be extracted reliably from analytical information obtained noninvasively from human subjects. The objective of this paper is to identify the principal issues that must be considered in order to advance the field of noninvasive glucose sensing. This paper does not review the multitude of proposed approaches for noninvasive sensing, nor is the purpose of this article to evaluate the validity of the various reports in the peer-reviewed literature. Rather, the objective here is to categorize the two major approaches (direct and indirect) and to discuss issues related to the selectivity, accuracy, and feasibility of each approach. Critical parameters are identified with a focus on benchmarking these values and on establishing minimal or optimal values for successful noninvasive clinical measurements. A set of publication standards is proposed as a way to compare work within and between laboratories, thereby greatly facilitating progress.

## SIGNIFICANCE

The ability to measure glucose noninvasively in human subjects would revolutionize the daily treatment and management of diabetes. Such technology could be used to monitor blood glucose levels continuously, thereby providing information necessary to maintain euglycemia by insulin intervention. By all accounts, such glycemic control would significantly reduce the devastating and costly medical complications associated with diabetes.<sup>3–5</sup> Once successfully developed, noninvasive glucose monitoring will greatly impact the quality of life for millions of people with diabetes and will significantly reduce the health care cost burden on families and Society.

The challenge is to develop technology that can measure in vivo glucose concentrations with sufficient accuracy and precision for the reliable and effective clinical treatment of diabetes. The most demanding concentration range is 2–5 mM, where relatively minor inaccuracies can result in dangerously poor clinical decisions.<sup>6</sup> Current test strip glucose sensing technology provides relative accuracies of 6–10%,<sup>7</sup> and ideally, this level of performance must be matched by any successful noninvasive glucose sensing device.

Even a partially successful development of accurate noninvasive sensing technology will greatly impact the treatment of diabetes. A device that can noninvasively and continuously monitor

nocturnal glucose levels will certainly enhance the daily quality of life for people with type I diabetes. A nocturnal alarm would particularly help children with diabetes and their parents. Even if the final device must be calibrated daily with conventional test strip technology, a noninvasive monitor can provide a convenient and painless means to measure glucose frequently throughout the day. Of particular significance is the ability to measure rates and direction of glycemic change, which can be used to predict and avoid dangerous hypoglycemic events. Finally, noninvasive glucose sensing technology would greatly impact the research community by providing a means to follow in situ glucose concentrations in either human subjects or animal models.

## APPROACHES

Two fundamentally different approaches are being pursued for the noninvasive measurement of glucose. These approaches can be categorized as direct and indirect. Direct measurements are based on the unique chemical structure of the glucose molecule. Indirect measurements, on the other hand, involve measuring the effect of glucose on a measured physiological or physical parameter. The relative merits of each approach are discussed below.

*Indirect glucose measurements* are based on the effect of glucose on some secondary process. The most common example is the effect of glucose on the scattering properties of tissue. The original research reported by Kohl and co-workers clearly demonstrates a strong relationship between the light scattering properties of tissue and blood glucose concentrations.<sup>8,9</sup> More recently, optical coherent tomography (OCT) is under investigation as a means to measure the optical properties of tissue as a function of glucose concentration.<sup>10–13</sup> Other indirect approaches include impedance<sup>14</sup> and dielectric spectroscopy.<sup>15</sup>

Selectivity is the principal concern with all indirect approaches. While glucose might strongly impact the scattering and impedance properties of tissue, such properties are not exclusively modulated by glucose. Indeed, Kohl ultimately concluded that natural variations of other endogenous substances, such as albumin protein, can alter the light scattering properties of skin tissue just as much as glucose.<sup>9</sup> Years ago, noninvasive technology was developed to measure hypoglycemia based on skin sweat measurements.<sup>16</sup> Hypoglycemia generally induces skin sweat, which increased the salt content of the skin, thereby increasing the measured conductance. Sweat conductance was therefore proposed as a nocturnal alarm, where a threshold increase in skin conductivity set off an alarm to permit intervention and prevention of hypoglycemia. Of course, other physiological processes induce

(1) Khalil, O. S. *Clin. Chem.* **1999**, *45*, 165–177.

(2) Khalil, O. S. *Diabetes Technol. Ther.* **2004**, *6*, 660–697.

(3) Diabetes Control and Complications Trial Research Group. *N. Engl. J. Med.* **1993**, *329*, 977–986.

(4) *Diabetes in America*, 2nd ed.; NIH Publication No. 95-1468; National Diabetes Data Group, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases; 1995.

(5) American Diabetes Association. *Economic Costs of Diabetes in the U. S. in 2002*; *Diabetes Care* **2003**, *26*, 917–932.

(6) Clarke, W. L.; Cox, D.; Gonder-Frederick, L. A.; Carter, W.; Pohl, S. L. *Diabetes Care* **1987**, *10*, 622–628.

(7) Johnson, R. N.; Baker, J. R. *Clin. Chim. Acta* **2001**, *307*, 61–67.

(8) Kohl, M.; Cope, M.; Essenpreis, M.; Boecker, D. *Opt. Lett.* **1994**, *19*, 2170–2172.

(9) Kohl, M.; Essenpreis, M.; Cope, M. *Phys. Med. Biol.* **1995**, *40*, 1267–1287.

(10) Esenaliev, R. O.; Larin, K. V.; Larin, I. V.; Motamedi, M. *Opt. Lett.* **2001**, *26*, 992–994.

(11) Larin, K. V.; Larin, I.; Motamedi, M.; Gelikonov, Y.; Kuranov, R.; Esenaliev, R. O. *Proc. SPIE* **2001**, *4263*, 83–90.

(12) Larin, K. V.; Eleдрisi, M.S.; Motamedi, M.; Esenaliev, R. O. *Diabetes Care* **2002**, *25*, 2263–2267.

(13) Larin, K. V.; Motamedi, M.; Ashitkov, T. V.; Esenaliev, R. O. *Phys. Med. Biol.* **2003**, *48*, 1371–1390.

(14) Caduff, A.; Hirt, E.; Feldman, Y.; Ali, Z.; Heinemann, L. *Biosens. Bioelectron.* **2003**, *19*, 209–217.

(15) Hayashi, Y.; Livshits, L.; Caduff, A.; Feldman, Y. *J. Phys. D* **2003**, *36*, 369–374.

(16) Hansen, A. H.; Duck, C. D. *Diabetes Care* **1983**, *6*, 597–600.

skin sweat, thereby rendering this technology susceptible to false positives. Indeed, the frequency of false positives was too high and this technology was never accepted.

The principal challenge of all indirect approaches is to establish the selectivity of the measurement. For any putative indirect approach, all endogenous substances that modulate the monitored property must be identified and the extent of modulation determined. In vitro measurements will generally never be sufficient for such a characterization. While in vitro experiments can provide valuable information regarding the relative impact of glucose and possible interferences, the nature of the indirect effect likely will depend on a number of physiological or biochemical processes that cannot be adequately simulated in an in vitro experiment. An example is the recent attempt to judge the selectivity of OCT measurements for glucose by assessing the relative impact of glucose, urea, sodium chloride, and other clinical substances on the refractive index of aqueous solutions.<sup>17</sup> While such evaluations indicate that glucose more strongly impacts the refractive index of a relatively simple aqueous solution compared to the other tested substances, these measurements do not incorporate the effect of these substances on other physiological properties of the tissue. Other parameters that strongly impact light scattering include the size and shape of the tissue cells and the degree of hydration of the extracellular space. The relative impact of glucose and the test substances on these additional parameters must also be determined before selectivity can be assessed. The biological system must be evaluated as a whole in order to determine feasibility of measuring glucose by any indirect method.

Future research pertaining to these indirect approaches must focus on measurement selectivity. It should no longer be acceptable to show only the results of a multivariate calibration model generated by modulating glucose while monitoring the parameter of interest. A rigorous evaluation of measurement selectivity must be included. In fact, the physiological basis of measurement selectivity must be the central theme of these future investigations. Such efforts must vigorously attempt to identify interfering processes. Any selective system must be based on sound biochemical and physiological processes.

Furthermore, research to develop "better" instrumentation for measuring the proposed glucose-dependent parameter is simply unwarranted until the issues of selectivity and measurement accuracy are established. Developing a more robust or faster OCT system, for example, does not advance the field unless this approach is known to provide selective and accurate analytical measurements. In certain instances, technological or instrumental developments are required before adequate sensitivity and selectivity can be achieved by a given technique. Such arguments must be supported by a critical review of the literature coupled with a logical and detailed line of reasoning that strongly indicates improvements in the fundamental selectivity. Overall, it is imperative that research efforts focus on the principal limitations of our current understanding of the method in question, not the development of technology for the sake of the technology.

Direct approaches involve measuring an intrinsic property of the glucose molecule. Although some attention has been given

to polarimetry, particularly in the eye,<sup>18,19</sup> the vast majority of literature reports center on vibrational transitions within the glucose molecule. Selectivity and sensitivity are the central issues for all direct measurement schemes. From a spectroscopic standpoint, glucose is a relatively simple and mundane molecule with no exceptional spectroscopic features to exploit in distinguishing it from other endogenous molecules within the human body. The best differentiation stems from the unique vibrational spectrum of glucose compared to the corresponding spectra of the principal chemical components of human tissue. These components include water, proteins, amino acids, urea, fatty acids, triglycerides, lipid membranes, proteoglycans, and other classes of biomolecules. At millimolar concentrations, glucose is a minor component compared to these other substances.

The principal measurement techniques are absorption spectroscopy over either the near- or mid-infrared spectral regions and Raman spectroscopy.<sup>1,2</sup> Various data collection formats are described for collecting noninvasive spectra, including transmission, diffuse reflectance, transreflectance, and photoacoustic measurement geometries. Mid-infrared spectroscopy offers strong and distinctive absorption features for glucose. Unfortunately, these wavelengths of light are strongly absorbed by water and other tissue components, thereby limiting penetration into human tissue to tens of micrometers. These short penetration depths render mid-infrared spectroscopy impractical except for a limited number of thin and vascular measurement sites, such as the tympanic membrane.<sup>20</sup> Near-infrared spectroscopy covers the spectral region from 14 286 to 4000  $\text{cm}^{-1}$  or 0.7 to 2.5  $\mu\text{m}$ . Scattering and absorption processes influence the propagation of near-infrared light within human tissue. Penetration depths vary depending on the exact wavelength, but generally range from 1 mm at longer wavelengths to 10 mm or more at shorter wavelengths.<sup>21</sup> These longer penetration depths greatly expand the number of possible noninvasive measurement sites.<sup>22</sup> Molecular absorptions over the near-infrared spectrum correspond to combinations and overtones of fundamental vibrational transitions. The resulting absorption bands are weaker and broader compared to the fundamental bands. Although the near-infrared spectrum of glucose is unique relative to the principal chemical components of human tissue, spectral overlap is extensive, thereby demanding the use of multivariate calibration methods of analysis. Raman spectroscopy provides the ability to measure fundamental vibrational information at wavelengths that penetrate deep into human tissue.<sup>23</sup> Challenges for Raman spectroscopy include the presence of a strong background fluorescence signal and inherently low signal-to-noise ratios (SNR). Weak Raman cross sections limit the magnitude of the Raman scattering signal, and the instability of the excitation laser generates considerable noise and complex spectral variance.

Direct glucose measurements are further complicated by the heterogeneous distribution of glucose throughout a tissue matrix.

(17) Larin, K. V.; Akkin, T.; Esenaliev, R. O.; Motamedi, M.; Milner, T. E. *Appl. Opt.* **2004**, *43*, 3408–3414.

(18) Cameron, B. D.; Baba, J. S.; Coté, G. L. *Diabetes Technol. Ther.* **2001**, *3*, 201–207.

(19) Lane, J. T.; Toris, C. B.; Nakhle, S. N.; Chacko, D. M.; Wand, Y. L.; Yablonski, M. E. *Am. J. Ophthalmol.* **2001**, *132*, 321–327.

(20) Block, M. J. U.S. Patent 6,002,953, 1999.

(21) Heise, H. M. *Horm. Metab. Res.* **1996**, *28*, 527–534.

(22) Burmeister, J. J.; Arnold, M. A. *Clin. Chem.* **1999**, *45*, 1621–1627.

(23) Hanlon, E. B.; Manoharan, R.; Koo, T. W.; Shafer, K. E.; Motz, J. T.; Fitzmaurice, M.; Kramer, J. R.; Itzkan, I.; Dasari, R. R.; Feld, M. S. *Phys. Med. Biol.* **2000**, *45*, R1–R59.

As light propagates through the tissue, it interacts with many different compartments that can contain different concentrations of glucose. Major pools of glucose include extracellular and intracellular fluids as well as capillary, venous, and arterial blood. Most direct noninvasive measurements cannot distinguish between these different compartments, and the measured response corresponds to a weighted average across the entire tissue matrix. A notable exception is the proposal to measure glucose in arterial blood by focusing exclusively on vascular pulsations.<sup>24</sup> Research is needed to determine the best tissue site for noninvasive measurements and to characterize the impact of tissue heterogeneity on calibration factors and calibration stability.

### CRITICAL PARAMETERS

Many papers in the peer-reviewed literature and in the patent literature claim to measure glucose successfully from noninvasive spectra collected from human subjects.<sup>1,2</sup> Unfortunately, the validity of such claims cannot be judged because of a lack of supportive information. Fundamental information, such as optical path length, wavelength range, and spectral SNR, are typically not provided, thereby making it impossible to assess the validity of claims to be measuring glucose. In addition, essentially all of these claims are based on multivariate statistical methods of analysis, such as partial least-squares (PLS) regression or artificial neural networks, which are prone to false calibrations based on spurious correlations between endogenous glucose concentrations and spectral features within the data set.<sup>25</sup> In general, no attempts are made to explore the chemical basis of selectivity for these models. More disturbingly, many of these multivariate models are established with a disregard for the relationship between the experimental degrees of freedom and the number of independent variables used in constructing the model (i.e., the model dimensionality).

The critical parameters required to assess and thereby advance the field of noninvasive glucose sensing spectroscopy include optical path length, wavelength range, instrumental performance, model dimensionality, and the chemical basis of model selectivity. These issues are discussed below, primarily in the context of near-infrared spectroscopy, but the major points are relevant for all noninvasive spectroscopic approaches.

**Optical Path Length.** All quantitative spectroscopic measurements depend strongly on the number of analyte molecules in the optical path during the measurement. A noninvasive glucose measurement based on near-infrared spectroscopy is nothing more than an absorbance measurement, and as for all absorbance measurements, the analytical sensitivity is directly proportional to optical path length. Plainly stated, there must be a sufficient number of glucose molecules in the optical path to absorb enough light to distinguish the absorbance related to glucose from the background variance or noise.

According to the Beer–Lambert law ( $A = \epsilon bc$ ), the measured light absorbance ( $A$ ) depends on the molar absorptivity of the analyte ( $\epsilon$ ), optical path length ( $b$ ), and analyte concentration ( $c$ ). For a noninvasive measurement, the optical path length is the only major parameter that can be controlled by the operator. The molar absorptivity is an intrinsic property of the glucose molecule,

and the concentration range is dictated by the physiology. Given the strong dependency on optical path length, it is unfortunate that few publications or patents specify a value for this critical experimental parameter.

Optical path lengths are not trivial to measure for noninvasive tissue measurements. Both human and animal tissues highly scatter near-infrared wavelengths of light, and this scattering process greatly alters the effective path of photons through the tissue matrix. As near-infrared light propagates through living tissue, many of these photons experience multiple scattering sites that change the direction of the propagating photon, and in general, the path traveled by each photon increases with each scattering event.<sup>26</sup> So-called ballistic photons also are possible and correspond to a small fraction of the incident photons that are not scattered while propagating through the matrix. In general, multiple scattering events increase the average path a photon takes through the sample, thereby increasing the optical path length of the measurement.

Diffuse reflectance and transreflectance measurement geometries are the most commonly used for noninvasive measurements.<sup>1,2</sup> In these geometries, multiscattered photons exit the sample and are measured by the detection optics. In photoacoustic spectroscopy, the measurement is based on the detection of acoustic waves generated from heat produced during the absorption of modulated incident radiation.<sup>27</sup> For all three measurement geometries, the effective optical path length depends on the penetration depth of the incident light into the sample and is dependent on many chemical and physical parameters of the sample and the measurement environment. Monte Carlo simulations coupled with Mie theory are commonly used to model such systems and to estimate optical path lengths.<sup>28</sup> Information obtained from both diffuse reflectance and photoacoustic measurements is heavily weighted toward the tissue surface. Transreflectance measurements can be configured to probe more deeply below the tissue surface.<sup>29</sup>

Detection of light transmitted through the sample is an alternative measurement geometry. In this geometry, both multiscattered and ballistic photons are detected as the incident light propagates through the tissue. The optical path length corresponds to the weighted average of these two types of photons. Because multiscattering events increase the effective path length of the propagating photons, optical path lengths for transmission measurements are larger than the physical thickness of the sampled tissue.<sup>30</sup> Again, the exact nature of the propagating light path depends on many chemical and physical properties of the sample. Unlike diffuse reflectance and photoacoustic measurements, transmission measurements are not as heavily weighted toward the surface and give a value more representative of the bulk tissue.

As noted above, the propagation of light through tissue is complicated by the heterogeneous nature of the tissue matrix. In addition to the different pools of aqueous fluids (intracellular,

(24) Aldrich, T. K. U.S. Patent 6,615,064, 2003.

(25) Arnold, M. A.; Burmeister, J. J.; Small, G. W. *Anal. Chem.* **1998**, *70*, 1773–1781.

(26) Weiss, G. H.; Gandjbakhche, A. H.; Masoliver, J. *J. Mod. Opt.* **1995**, *42*, 1567–1574.

(27) MacKenzie, H. A.; Ashton, H. S.; Spiers, S.; Shen, Y.; Freeborn, S. S.; Hannigan, J.; Lindberg, J.; Rae, P. *Clin. Chem.* **1999**, *45*, 1587–1595.

(28) Heise, H. M.; Marbach, R.; Koschinsky, Th; Gries, F. A. *Artif. Organs* **1994**, *18*, 439–447.

(29) Segtnan, V. H.; Isaksson, T. *J. Near Infrared Spectrosc.* **2000**, *8*, 109–116.

(30) Burmeister, J. J.; Chung, H.; Arnold, M. A. *Photochem. Photobiol.* **1998**, *67*, 50–55.

extracellular, and capillary blood), the tissue matrix also includes a significant amount of nonaqueous material. Assuming glucose is only located in these aqueous components, the effective optical path length of a noninvasive measurement must differentiate between the aqueous and nonaqueous fractions of the tissue.

A relatively simple and straightforward method is proposed for estimating the effective aqueous path length of light propagating through living tissue.<sup>30</sup> This method is based on an analysis of the absolute absorption spectrum of the sample. Absorbance of the sample is computed at all wavelengths by dividing the single-beam spectrum for the sample by a single-beam spectrum of air and taking the negative logarithm. The resulting spectrum in absorbance units is then regressed relative to a set of pure-component aqueous absorbance spectra of the major chemical constituents in the sample. The principal chemical components for living tissue include water, protein, fat, and terms to account for changes in tissue scattering, which can be accounted for by simple slope and y-intercept terms. Equation 1 shows the regres-

$$A_s = \beta_w A_w + \beta_p A_p + \beta_f A_f + \beta_{b1} A_{b1} + \beta_{b0} A_{b0} \quad (1)$$

sion expression where  $A_s$ ,  $A_w$ ,  $A_p$ ,  $A_f$ ,  $A_{b1}$ , and  $A_{b0}$  correspond to absorbance spectra of the sample, water, protein, fat, sloping line, and horizontal line, respectively, and the  $\beta$  terms are the corresponding regression coefficients. Each regression coefficient gives the relative thickness or amount of each component represented in the sample. If the pure-component water spectrum is taken from a 1-mm-thick sample of water and the water regression coefficient for the fitted spectrum is 0.5, then the effective aqueous optical path length of the near-infrared light propagating through this sample is estimated to be 0.5 mm.

The importance of determining the effective optical path length for noninvasive measurements cannot be understated. The limit of detection for an analyte depends on the background noise level and the analytical sensitivity. The fact that analytical sensitivity is directly proportional to path length makes it critical to know this value in order to support a claim of measuring glucose from noninvasive spectra. The ability to measure clinically relevant levels of glucose under conditions of the *in vivo* or noninvasive experiment must be demonstrated. In other words, the experimenter has an obligation to demonstrate that glucose can be measured with a detection limit of at least 2 mM under the conditions of path length and SNR used for the *in vivo* measurements. A corresponding *in vitro* experiment can be performed where the same path length and SNR are used to measure glucose in a set of well-designed standard solutions with random glucose concentrations relative to both time of measurement and composition of the chemical matrix. The exact nature and complexity of the chemical matrix used in such an *in vitro* experiment are important from the standpoint of rigor in demonstrating validity of the corresponding *in vivo* results. The more similar the *in vitro* chemical matrix is to the actual *in vivo* matrix, the greater confidence one has in the functionality of the *in vivo* model. Nevertheless, value can be obtained from simply performing this test experiment in a simple phosphate buffer matrix. While such a simple sample matrix suffers from the correlation of glucose and water concentrations through water displacement effects, it will certainly be true that if it is not possible to measure clinically

relevant levels of glucose in a simple phosphate buffer under the conditions of path length and SNR used in the *in vivo* experiment, the *in vivo* model is invalid. Such a simple test has rarely been used, but must become routine in order to legitimize *in vivo* studies.

Positive results from the above experiment are necessary but not sufficient to prove that glucose is being measured directly. Additional conditions must be met before such a claim can be made.

**Wavelength Range.** The near-infrared spectrum spans a wide range from 14 286 to 4000  $\text{cm}^{-1}$  or 0.7 to 2.5  $\mu\text{m}$ . Absorption features throughout this spectral range primarily correspond to overtones and combinations of molecular vibrations. The absorption properties of water play a critical role in the regions of the near-infrared spectrum available for noninvasive measurements. Strong water absorption bands centered at approximately 7500, 5200, and 3600  $\text{cm}^{-1}$  create three transmission windows through aqueous solutions and living tissue. These spectral windows are termed the short-wavelength region (14 286–7300  $\text{cm}^{-1}$ ), the first overtone region (6500–5500  $\text{cm}^{-1}$ ), and the combination region (5000–4000  $\text{cm}^{-1}$ ). Absorption features in the combination region correspond to first-order combination transitions associated with bending and stretching vibrations of C–H, N–H, and O–H functional groups. The first overtone region corresponds to the first-order overtone of C–H stretching vibrations, and the short-wavelength region includes numerous higher order combination and overtone transitions. For combination spectra, molar absorptivities are larger and bands are narrower compared to first overtone spectral features.<sup>31,32</sup> Near-infrared absorption features become significantly weaker and broader as the order increases, thereby greatly reducing the analytical utility of the short-wavelength region in terms of molecular vibrational information.

For many noninvasive reports, the issue of spectral range is not considered in detail. Although a nominal spectral range is specified in all cases, the actual useful spectral range is rarely identified. Sample thickness and wavelength optimization are two critical issues in this regard. The different spectral ranges permit different sample thicknesses depending on the scattering properties of the tissue matrix and the absorption properties of water. Thicker samples are possible for spectra collected at shorter wavelengths, while spectra at longer wavelengths are restricted to thinner samples. The necessary sample thickness is a complex relationship between the effective aqueous path length, desired limit of detection, and SNR of the measurement. As a rule, optimal sample thicknesses for the combination, first overtone, and short-wavelength spectral regions are on the order of 1, 5, and 10 mm, respectively, for transmission measurements. Estimated effective path lengths for diffuse reflectance measurements are significantly shorter.<sup>28</sup>

Effective analytical measurements demand that the sample thickness match the spectral region. For example, no practical spectroscopic information can be obtained from near-infrared spectra collected through a 5-mm-thick aqueous sample over the combination spectral range. Strong attenuation of the incident radiation will result in no light reaching the detector, thereby providing no useful analytical information. On the other hand,

(31) Amerov, A. K.; Chen, J.; Arnold, M. A. *Appl. Spectrosc.* **2004**, *58*, 1195–1204.

(32) Chen, J.; Arnold, M. A.; Small, G. W. *Anal. Chem.* **2004**, *76*, 5405–5413.

samples that are too thin will not have sufficient absorbance from glucose to provide a useful limit of detection. First overtone spectra collected from a 0.5-mm-thick aqueous sample will not provide sufficient analytical signal to measure clinical levels of glucose with most spectrometer systems. In many literature reports, this required match between spectral range and sample thickness is not considered. Near-infrared spectra are commonly collected over a broad spectral range with a single sample thickness. When the collected spectra encompass multiple spectral regions, it is not possible to match the sample thickness with each spectral region. Typically, a sample thickness is used that matches one of these spectral regions, which greatly reduces or eliminates the significance of the analytical information provided by the other spectral region. For example, near-infrared spectra collected from a 5-mm-thick aqueous sample over a spectral range from 6500 to 4000  $\text{cm}^{-1}$  will only contain useful analytical information from the first overtone spectral region (6500–5500  $\text{cm}^{-1}$ ). All information associated with the combination region will be lost in the poor SNR generated by the extreme loss of photons through the thick sample. This point must be recognized when analyzing *in vivo* data and when designing *in vitro* experiments.

Wavelength optimization is another process that must be considered when exploring glucose measurements from *in vivo* spectra. This process involves evaluating a set of calibration models generated with different spectral regions from within the collected spectrum. A grid search or numerical optimization strategy can be used to find the spectral region, or combination of spectral regions, that provides the best analytical results. In general, noisy regions with no analyte-specific information should be excluded from the optimized spectral range. Certainly, the selected spectral range must include known analyte absorption features and these features should be the same as those deemed most informative from complementary *in vitro* experiments.

**Model Dimensionality.** Multivariate calibration methods are necessary to extract analyte-specific information from complex *in vivo* spectra in which the glucose spectral signature overlaps with the signatures corresponding to other constituents of the sample matrix. Many related algorithms have been reported for accomplishing this modeling step. Each of these methods treats the *in vivo* spectrum as the summation of a series of underlying spectral shapes that are derived from the chemical constituents of the sample, the physical characteristics of the measurement site, and the instrumental signature associated with the measurement. The modeling step involves the determination or selection of these spectral shapes, the calculation of scores that encode how much each spectral shape contributes to a given spectrum, and the derivation of a numerical model that correlates the scores with analyte concentration. The calibration algorithms differ in the manner in which the spectral shapes are derived and the form of the model used to establish the correlation with analyte concentration.

Among the methods available, PLS regression has been used most widely for the analysis of near-infrared spectral data.<sup>33</sup> In the most common implementation of this method, the spectral variance is decomposed into a series of orthogonal factors (latent variables). These factors represent the underlying spectral shapes that comprise the data, weighted according to the covariance

between the spectral information and analyte concentration. Through this spectral decomposition step, the large dimensionality of each spectrum (e.g., hundreds of resolution elements) is reduced to a few (e.g., <20) latent variables. The calibration model is then formulated with the reference analyte concentrations serving as the dependent variable and the latent variable scores as the independent variables. The form of the model can range from a straightforward multiple linear regression model to a highly nonlinear model encoded with an artificial neural network. In each case, the actual calculation of the model requires the estimation of parameters (regression coefficients or weights) that relate the scores to the analyte concentration. The number of these estimated parameters (i.e., the model dimensionality) is limited by the number of independent samples, or degrees of freedom, available within the calibration data set. According to the ASTM Standard E 1655 on infrared multivariate calibration methods,<sup>34</sup> a minimum of six independent samples is needed for each estimated parameter in the model (section 17.5) and more samples per parameter are highly recommended. Models based on fewer samples per parameter are simply not acceptable and should not be published. An additional four independent samples per estimated parameter are recommended for use in validating the multivariate model (section 18.2.3). These validation samples should be separated in time from the calibration samples to allow the assessment of the effects of instrumental drift on model performance.

What constitutes an independent sample must be considered when establishing the degrees of freedom for a data set. Certainly, if multiple spectra are collected for the same sample, then all these spectra do not represent independent measurements, but must be considered replicates. So, the maximum permitted model dimensionality does not depend on the number of spectra in the calibration data set, but on the number of independent samples used to generate these spectra. In cases where the full spectral data set is split into calibration and validation or prediction spectra, any replicates for each sample must be kept together and not split between calibration and prediction.

Future publications should include a separate section dedicated to a discussion of degrees of freedom and model dimensionality for each of the presented calibration models.

**Net Analyte Signal (NAS).** Any analysis of selectivity for multivariate calibration models must recognize the NAS. The term NAS was originally defined by Lorber as the component of the pure-component analyte spectrum that is orthogonal to all sources of spectral variance within the sample matrix.<sup>35,36</sup> The NAS represents the component of the analyte spectrum that is unique compared to all other sources of spectral variance, including changes in the chemical matrix, alterations in the physical properties of the sample, and variations induced by instrumental or environmental artifacts.

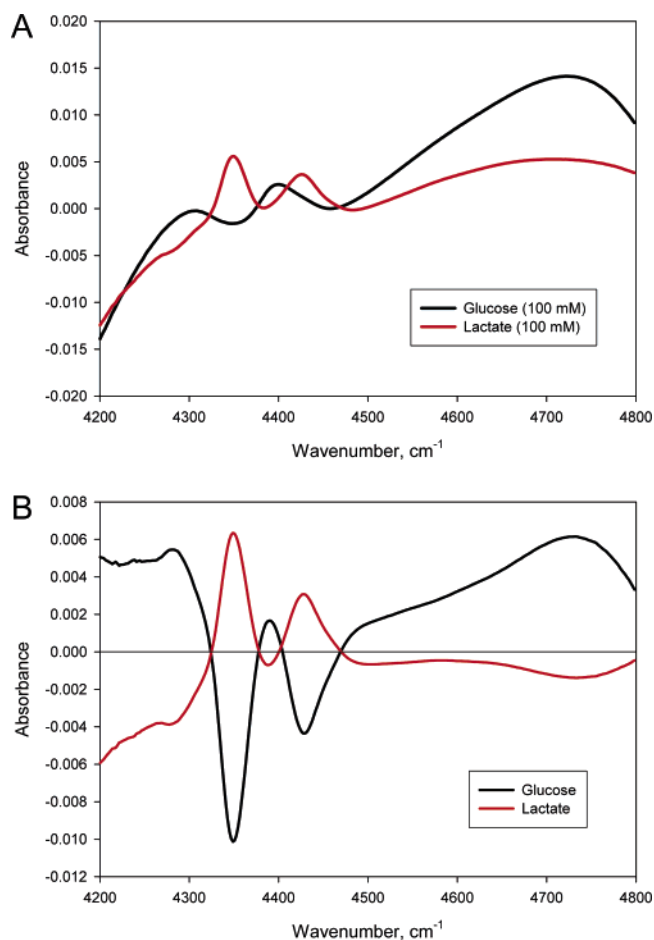
Vector analysis can be used to visualize the NAS. In this analysis, each spectrum is represented as a vector in a multidimensional space defined by the resolution elements of the spectrum. For example, an aqueous spectrum of glucose collected

(33) Haaland, D. M.; Thomas, E. V. *Anal. Chem.* **1988**, *60*, 1193–1202.

(34) American Society for Testing and Materials. Standard Practices for Infrared, Multivariate, Quantitative Analysis (E 1655-94). In *Annual Book of ASTM Standards*, Global Engineering Documents, Philadelphia, PA, 1995.

(35) Lorber, A. *Anal. Chem.* **1986**, *58*, 1167–1172.

(36) Lober, A.; Faber, K.; Kowalski, B. R. *Anal. Chem.* **1997**, *69*, 1620–1626.



**Figure 1.** Absorbance spectra (A) and NAS vectors (B) for glucose and lactate over the combination region of the near-infrared spectrum. *lp;*&-4q;1

from 4800 to 4200  $\text{cm}^{-1}$  with a  $2\text{-cm}^{-1}$  point spacing results in 300 resolution elements. The magnitude of absorbance at each resolution element is plotted in this 300-dimensional space. Such a plot results in a single point, and this glucose spectrum is then represented by a vector that extends from the origin to this point. The direction of this vector represents the shape of the glucose spectrum, while its length represents the magnitude of the absorbance, which is proportional to concentration. Spectral vectors can also be plotted in this manner for all nonanalyte sources of spectral variance within the data set. The NAS corresponds to that portion of the analyte spectral vector that is orthogonal to the subspace spanned by the vectors representing all nonanalyte sources of spectral variance. This component of the analyte spectral vector is a vector itself, and the NAS vector can be viewed as a spectrum by plotting the resulting absorbance values as a function of the corresponding spectral resolution elements.

Examples of NAS spectra are presented in Figure 1 for glucose and lactate. Aqueous absorbance spectra for glucose and lactate are superimposed in Figure 1A for comparison. These spectra were collected from aqueous buffer solutions composed of 100 mM glucose and 100 mM lactate, respectively, and reference spectra used in the absorbance calculation were generated from blank phosphate buffer. The glucose spectrum is characterized by three absorption features centered at 4300, 4400, and 4720  $\text{cm}^{-1}$

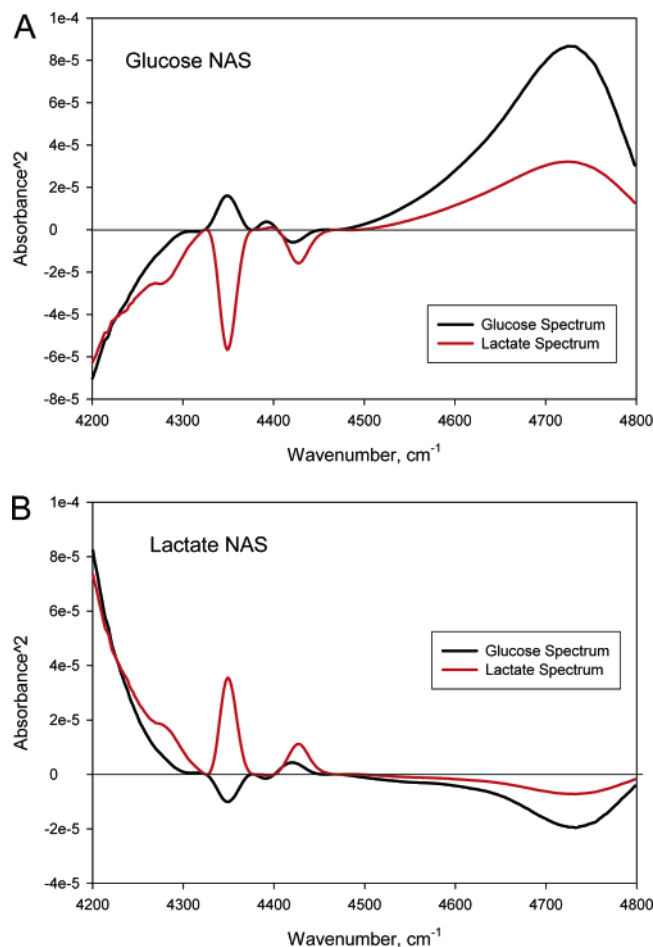
and the lactate spectrum shows bands centered at 4350, 4425, and a small, broad band with a maximum around 4710  $\text{cm}^{-1}$ . For both spectra, negative absorbance values are observed at lower wavenumbers due to water displacement effects and the relative magnitude of absorptivities between water and solute at these frequencies.<sup>31</sup> A comparison of the spectral traces reveals significant differences between position, width, and magnitude of the absorption features; however, considerable overlap is evident, particularly at the higher wavenumbers.

The corresponding NAS spectral vectors are plotted in Figure 1B, where these NAS vectors were determined by simple vector projections as described above. A notable characteristic of NAS spectral vectors is the presence of both positive and negative features. Simplistically, the positive features correspond to regions of the analyte spectrum that are distinctive relative to the interfering component(s). Negative features correspond to distinguishing spectral features of the interfering component(s) that must be “subtracted” in order to account for the interfering species. For the glucose NAS spectrum, glucose is the analyte and lactate is the interfering substance; hence, the NAS spectrum for glucose shows positive features at 4280, 4390, and 4732  $\text{cm}^{-1}$ , closely matching the positions of the dominant glucose absorption bands, and negative features centered at 4348 and 4427  $\text{cm}^{-1}$ , corresponding to distinctive lactate absorption bands.

In fact, the structure of the NAS spectral vector is more complicated than simply positive features for the analyte and negative features for the interference(s). To eliminate all contributions from an interfering species, the NAS vector must have the property where the interfering species contributes equally in magnitude to the positive and negative components of the integrated signal. Only in this way will the presence of the interfering component be exactly canceled regardless of its concentration, thereby permitting a selective measurement.

The selectivity of the NAS can be demonstrated by analyzing the product of the NAS spectrum and the pure-component spectra of the analyte and interfering compound. Figure 2 shows the resulting spectra obtained by performing a point-by-point multiplication of the NAS spectral vector and the pure-component spectral vectors for glucose and lactate, respectively. Figure 2A shows product spectra for the glucose NAS, and Figure 2B presents product spectra for the lactate NAS. Again, both positive and negative regions are evident for both situations where the pure-component spectrum corresponds to the analyte or the interference. Selectivity can be demonstrated by comparing the sum of these positive ( $\Sigma_{\text{pos}}$ ) and negative values ( $\Sigma_{\text{neg}}$ ). The values of  $\Sigma_{\text{pos}}$ ,  $\Sigma_{\text{neg}}$ ,  $\Sigma_{\text{net}} = \Sigma_{\text{pos}} + \Sigma_{\text{neg}}$ , and  $\Sigma_{\text{pos}}/\Sigma_{\text{neg}}$  are summarized in Table 1 for the product spectra presented in Figure 2. For the interfering species,  $\Sigma_{\text{pos}} = -\Sigma_{\text{neg}}$ , thereby effectively canceling the impact of the interference. This fact is evident for lactate relative to the NAS for glucose and for glucose relative to the NAS for lactate. In both cases,  $\Sigma_{\text{net}}$  is essentially zero and  $\Sigma_{\text{pos}}/\Sigma_{\text{neg}}$  is unity. For the analyte, however,  $\Sigma_{\text{net}} > 0$ . This result is necessary because  $\Sigma_{\text{net}}$  is directly related to analyte concentration through an appropriate proportionality constant.

The magnitude of the NAS signal is a critical parameter in determining the success of quantitative measurements. Taken over the entire spectrum, the magnitude of the NAS vector is directly proportional to the analytical sensitivity. As defined by Lorber,<sup>35</sup>



**Figure 2.** Product spectra obtained by point-by-point multiplication of the NAS spectrum for glucose (A) or lactate (B) with pure-component spectra for glucose and lactate.

**Table 1. Analysis of Positive and Negative Components of the Product Spectra from NAS and Pure-Component Spectra for Glucose and Lactate**

pure-comp spectrum	$\Sigma_{\text{pos}}^a$ (AU <sup>2</sup> )	$\Sigma_{\text{neg}}^b$ (AU <sup>2</sup> )	$\Sigma_{\text{net}}^c$ (AU <sup>2</sup> )	$ \Sigma_{\text{pos}}/\Sigma_{\text{neg}} ^d$
Glucose NAS Vector				
glucose	$7.30 \times 10^{-3}$	$-1.56 \times 10^{-3}$	$5.74 \times 10^{-3}$	4.69
lactate	$2.72 \times 10^{-3}$	$-2.72 \times 10^{-3}$	$4.23 \times 10^{-10}$	1.00
Lactate NAS Vector				
glucose	$1.53 \times 10^{-3}$	$-1.53 \times 10^{-3}$	$2.57 \times 10^{-9}$	1.00
lactate	$2.29 \times 10^{-3}$	$-5.28 \times 10^{-4}$	$1.76 \times 10^{-3}$	4.33

<sup>a</sup> Sum ( $\Sigma_{\text{pos}}$ ) computed over positive regions of the product spectra in Figure 2. <sup>b</sup> Sum ( $\Sigma_{\text{neg}}$ ) computed over negative regions of the product spectra in Figure 2. <sup>c</sup>  $\Sigma_{\text{net}} = \Sigma_{\text{pos}} + \Sigma_{\text{neg}}$ . <sup>d</sup> Absolute value of the ratio of  $\Sigma_{\text{pos}}$  to  $\Sigma_{\text{neg}}$ .

the multivariate sensitivity is  $\|\text{NAS}\|/c^0$ , where  $\|\text{NAS}\|$  is the vector magnitude of the NAS and  $c^0$  is the analyte concentration in the spectrum used to compute the NAS. This definition of multivariate sensitivity is analogous to the product of  $\epsilon b$  in the Beer–Lambert law. Under conditions of constant SNR and optical path length, the magnitude of the NAS vector determines the limit of detection.

In a typical measurement, two elements greatly impact the NAS magnitude and, therefore, the sensitivity and limit of detection of

the measurement. The first is the complexity of the sample matrix. More spectral overlap is expected for samples with greater chemical complexity, and this greater spectral overlap reduces the magnitude of the NAS vector. Second, the NAS magnitude is adversely affected by spectral noise. Noise is omnivariant, which results in a loss of NAS across all spectral elements when absorbance from the analyte is comparable in magnitude to the noise.

Various methods are available to determine the NAS for a given analyte in a particular sample matrix. The direct calculation method demonstrated above can be difficult to implement for samples composed of complex chemical and physical matrixes. In addition, a functional NAS must incorporate spectral variances introduced by both instrumental and environmental factors. As noted below, many of these types of spectral variance are complicated, which makes them difficult to incorporate accurately into the NAS calculation. When the calibration model is based on latent variable scores obtained from methods such as PLS or principal component analysis (PCA), Lorber,<sup>36</sup> Faber,<sup>37</sup> and Ferré<sup>38,39</sup> have all proposed algorithms for estimating the NAS. Use of these algorithms to estimate the NAS requires good experimental design, however. Under less than ideal conditions, however, both PLS and PCA can incorporate spuriously correlated spectral variance into the model, thereby corrupting the estimate of the NAS. A ramification of this occurrence is that the computed calibration model is based, in part, on nonselective information. Overall, while extremely valuable for investigating the selectivity of multivariate calibration models, detailed interpretation of NAS vectors is very difficult. The presence of complex shapes and subtle features within the NAS vectors render quantitative interpretation and rigorous comparisons virtually impossible.

**Spectral Variance.** Random and pseudorandom variations within the spectra establish a background variance against which glucose must be distinguished. The three major sources of variance for noninvasive spectra have instrumental, chemical, and physical bases. Instrumental variance is related to the SNR of the instrument and to the stability or drift of the instrumentation over time. Chemical variance corresponds to changes in the chemical composition of the sample. Chemical variance is generally minimal for in vitro experiments where the chemical composition of the sample is controlled but can be a major source of variance during in vivo experiments. An example of in vivo chemical variance is the movement of endogenous molecules, such as water, fat, or protein, out of the optical path in response to a pressure gradient established by the spectrometer-to-human interface. Physical variance involves changes in matrix parameters that influence the collected spectra. Temperature and scattering are two prime examples of physical parameters that greatly influence spectra and that can be difficult or nearly impossible to control during in vivo measurements. Temperature is even difficult to control adequately during in vitro experiments because near-infrared spectra are known to be sensitive to temperature variations less than  $\pm 0.1$  °C.<sup>40</sup>

(37) Faber, N. K. *Anal. Chem.* **1998**, *70*, 5108–5110.

(38) Ferré, J.; Brown, S. D.; Rius, F. X. *J. Chemom.* **2001**, *15*, 537–553.

(39) Ferré, J.; Faber, N. K. *Chemom. Intell. Lab. Syst.* **2003**, *69*, 123–136.

(40) Hu, S.-K. B.; Arnold, M. A.; Wienczek, J. M. *Anal. Chem.* **2000**, *72*, 696–702.

Again, few published accounts provide information about the degree of spectral variance within a data set, yet the degree of spectral variance is critical in assessing the validity of the measurement. In essence, spectral variance must be sufficiently low to allow reliable detection of light absorbed by glucose molecules in the optical path. In practice, the relationship between spectral variance and optical path length defines the limit of detection. Clearly, standard methods are needed to quantify and benchmark spectral variance for the purpose of documenting progress and establishing limits.

All other parameters being equal, lower detection limits and superior calibration performance are expected for data sets with lower spectral variance. An interesting dichotomy exists, however, where better calibration models are possible from data sets with greater spectral variance. Spurious correlations with glucose concentration are the basis for this counterintuitive possibility.<sup>25</sup> Powerful multivariate calibration methods, such as PLS, are designed to accentuate small correlations between the spectral data and analyte concentrations. Such correlations are determined strictly on the basis of statistical analysis without regard for the chemical nature of the spectral variance. Greater spectral variance increases the chances of spurious correlations, thereby resulting in seemingly superior calibration models. Of course, basing calibration models on spurious spectral variance must be avoided, which makes it imperative to characterize the degree of spectral variance within each data set.

As discussed above for the optical path length, complementary *in vitro* experiments are needed to establish the minimal spectral variance required for a given limit of detection under conditions of *in vivo* measurements. Composition of the chemical matrix will be critical for these *in vitro* measurements and must match the *in vivo* matrix as closely as possible. Alternatively, detailed studies can be used to establish the practical relationship between limit of detection and spectral variance as a function of path length and the NAS.

Spectral variance can be partially characterized by analyzing 100% lines. A 100% line is generated by collecting two spectra from the exact same sample and dividing one by the other. Ideally, these spectra should be identical because they originate from the same sample, and therefore, the ratio between spectral intensity should be unity at all wavelengths. The result is a line with zero slope at 100% transmittance. In reality, such a plot will not be flat but will include noise that is associated with the instrumental SNR. Curvature or sloping 100% lines indicate differences in physical conditions between spectra and offsets from 100% are indications of drift in the instrumentation. Short- and long-term variance can be characterized by analyzing 100% lines generated from spectra collected over different periods of the experiment.

For *in vitro* experiments, short-term variance is best determined from 100% lines computed from back-to-back triplicate spectra for each sample. These intrasample spectra can be used to determine the instrumental SNR, where the instrumental SNR is inversely proportional to the root-mean-square (rms) noise computed from these 100% lines. This rms noise varies across the near-infrared spectrum according to the absorption spectrum of water and the thickness of the sample, thereby requiring a report of values over multiple short segments of the full spectrum. Long-term variance can be characterized from 100% lines gener-

ated with repeated spectra periodically collected from the same sample throughout the experiment. One possibility is to analyze 100% lines generated from buffer spectra collected at specific times during the experimental data collection period. Compared to 100% lines generated from back-to-back spectra, 100% lines computed from spectra collected over time will likely demonstrate more curvature, nonzero slopes, and offsets, thereby providing information related to variations in sample temperature, scattering, and instrumental drift.

Analysis of 100% lines is considerably more complicated for *in vivo* experiments, where conditions are nearly impossible to reproduce. Nevertheless, analysis of ratios of sequential *in vivo* spectra gives an excellent indication of the degree of spectral variance as a function of time. Such an analysis can provide information related to the source of the principal variance. For example, sloping baselines indicate variations in temperature or scattering in the *in vivo* environment. Furthermore, positive and negative absorption features within 100% lines generated from sequential spectra indicate increases and decreases, respectively, in the amount of a particular substance (e.g., protein or fat) within the optical path of the measurement.<sup>41</sup> Finally, chemical variations within noninvasive spectra can be characterized by the above-described regression technique where the *in vivo* spectra are fitted to a set of pure aqueous spectra of the principal chemical components of the sample matrix<sup>30</sup> (See eq 1 and corresponding discussion).

To date, no accepted protocols are available to characterize spectral variance within a data set. Before such protocols are established, all current publications and reports must provide, at a minimum, the instrumental SNR under the conditions of the experiment. For *in vitro* experiments, this information can be obtained from rms noise values calculated from back-to-back triplicate spectra for every sample. The consistency of these rms noise values can be determined by plotting the rms noise as a function of sample number. The mean and standard deviation of the composite rms noise values, and the corresponding SNR, should be reported for each experimental setup. For *in vivo* experiments, the same type of information can be generated by collecting a series of spectra from an appropriate tissue phantom. A simple, yet effective, phantom can be constructed by combining layers of water and a material that uniformly attenuates the incident light. Teflon blocks are excellent neutral density filters for near-infrared wavelengths and can be used as the phantom component. Thicknesses of the water and Teflon layers must be selected carefully. The water layer thickness should match the effective aqueous optical path length for the *in vivo* spectra, and the Teflon layer thickness must attenuate the incident light sufficiently so that the radiant power at the detector matches that observed for the *in vivo* spectra.

**Basis of Measurement Selectivity.** As noninvasive sensing technology is developed for *in vivo* clinical measurements, or for any other analytical application, it is critical to understand the chemical basis of measurement selectivity. Few papers address this aspect of such measurements. More commonly, calibration models generated from multivariate statistics are accepted without further investigation into the chemical or spectroscopic basis of

(41) Burmeister, J. J.; Arnold, M. A.; Small, G. W. *Diabetes Technol. Ther.* **2000**, *2*, 5–16.

the predicted concentrations. Part of the problem is a void in the availability of procedural tools that can be used to examine the selectivity of such calibration models.

Recently, we introduced the pure component selectivity analysis (PCSA) method as a means to probe the selectivity of multivariate calibration models.<sup>42</sup> In this method, the calibration model is generated in the normal way by using a set of calibration spectra to establish the multivariate relationship between spectral variance and analyte concentration. These calibration spectra originate from samples that encompass all the chemical variance expected for subsequent prediction or unknown samples. For the PCSA method, a set of pure-component aqueous spectra is collected for the analyte and all other major components of the chemical matrix. These pure-component spectra are sequentially subjected to the multivariate calibration model, and the output is inspected. Ideally, the output from all nonanalyte solutes should be zero, which would indicate that none of the spectral features associated with this nonanalyte component is used by the model to predict the analyte concentration. Likewise, output from the analyte pure-component spectra should produce concentrations that match analyte concentrations in these pure-component solutions. Correct concentrations from the analyte pure-component spectra indicate that all the required calibration information resides within the analyte spectrum. In this case, the analyte spectrum provides the chemical basis of selectivity for the measurement.

Very different PCSA outputs are expected when the analyte spectrum is not the sole basis of the calibration selectivity. For example, concentration correlations between the analyte and a major cosolute within the sample matrix can produce models where the analyte concentration predictions are based partially on the analyte spectrum and partially on the cosolute spectrum. Alternatively, in the absence of sufficiently strong absorption signals from the analyte, apparently functioning calibration models can be generated on the basis of even slight concentration correlations. Modeling experiments indicate that the PCSA method can detect these situations. Still, more testing and development of the PCSA method are necessary to more fully characterize its relative merits and to better understand its fundamental limitations.

The concept of the PCSA method is general, and this method should be applicable to any type of multivariate calibration scheme, including artificial neural networks. In addition, this method is not restricted to near-infrared spectroscopy, but should be equally informative for all types of data used in conjunction with multivariate calibration processes. Finally, additional characterization tools must be developed to better establish the chemical basis of selectivity for these complex models.

## CONCLUSIONS

The complexity of the in vivo environment is tremendous, and this complexity prevents direct measurements without a set of validating experiments. For indirect measurement schemes, the required validating experiments must address measurement selectivity by determining and examining the physiological basis of selectivity. All interfering species and competing biochemical pathways must be identified and their impact established. For

direct glucose detection, all in vivo findings must be consistent with complementary in vitro results, in terms of path length, SNR, wavelength range, model dimensionality, and the NAS.

Unfortunately, clinical researchers have published accounts of in vivo noninvasive glucose experiments without rigorous testing or model validation. Other clinicians generally accept these unsubstantiated findings as proof of successful noninvasive glucose measurements. It is extremely important that all researchers (clinicians and nonclinicians alike) critically evaluate these previous claims and reject them without considerably more information and validation. Propagating unsubstantiated claims in the literature can wrongfully legitimize this work, thereby adversely affecting general progress in the field. Everyone must realize it is relatively simple to measure a noninvasive signal that correlates with in vivo glucose concentrations; it is considerably more difficult, however, to construct accurate glucose calibration models that are based on a selective element of the glucose molecule and that can prospectively provide accurate glucose concentrations in human subjects.

Clearly, a set of publication standards must be developed in order to enhance progress toward a noninvasive glucose monitor. As a starting point, we offer the following list of issues that must be addressed in all future publications, regardless of journal.

(1) *Spectral Range.* The spectral range from which valid analytical information is derived must be identified and reported. The useful range can be determined by using optimization procedures or grid search methods to test regions for analytical utility. This range must be consistent with known glucose absorption bands. In addition, the rms noise of 100% lines should be characterized and reported over the critical segments of this spectral range.

(2) *Degrees of Freedom.* The dimensionality of any multivariate calibration model must be justified by an analysis of the degrees of freedom provided by the calibration data set. Specific information must be provided regarding the number of independent measurements. As a minimum, ASTM Standard E 1655 for multivariate calibrations should be followed.

(3) *Path Length.* The effective aqueous path length must be estimated and its relevance discussed. Relevance can be judged by comparing analytical performance relative to complementary in vitro experiments under the same conditions of path length and SNR.

(4) *Spectral Variance.* The degree of spectral variance must be characterized for both in vivo and complementary in vitro spectral data. Minimally, instrumental signal-to-noise ratios must be reported as determined from analysis of 100% lines from replicate measurements.

(5) *Chemical Basis of Selectivity.* An assessment of the chemical basis of selectivity is necessary to justify any proposed in vivo calibration model. Such assessments can utilize estimation and interpretation of the NAS or may incorporate methods such as the PCSA technique described above. Corroborating findings from well-designed complementary in vitro experiments will boost confidence in such assessments.

Benchmark experimental values are critical for progress toward the development of noninvasive sensing technology for glucose or any other clinical analyte. The reporting of values for the above-listed parameters in all subsequent publications will provide a

(42) Arnold, M. A.; Small, G. W.; Xiang, D.; Qiu, J.; Murhammer, D. W. *Anal. Chem.* **2004**, *76*, 2583–2590.

foundation against which future progress can be measured. Improvements can be judged quantitatively and the corresponding impact determined with confidence. Certainly, improvements in benchmarking procedures are needed and must be developed. Particularly, improvements are required for the assessment of chemical selectivity for complex multivariate calibration models. In addition, better methods are needed to characterize spectral

variance in general and to identify and quantify major sources of variance within sets of spectral data.

Received for review March 11, 2005. Accepted June 21, 2005.

AC050429E