# Near-Infrared Spectrometry of Microorganisms in Liquid Pharmaceuticals

**Leonard J. Galante and Michael A. Brinkley**

*Glaxo, Inc., Five Moore Drive, Research Triangle Park, North Carolina 27709*

**James K. Drennen and Robert A. Lodder***

*College of Pharmacy, University of Kentucky Medical Center, Lexington, Kentucky 40536-0082*

**Biotechnology and pharmaceutical research have created a number of new and potentially life-saving drugs. Many of these drugs are formulated as injectable products. Some drug products do not survive autoclaving or other means of terminal sterilization. An aseptic filling process is typically used to sterilize such products, but it is less reliable than autoclaving, making detection of unsterile units even more essential. Invasive microbiological methods and turbidimetry are currently employed as inspection techniques. These processes are time-consuming, destroy product, and may not detect low levels of contamination. Near-IR light scattering is proposed as a new method of determining low levels of contamination noninvasively and nondestructively. The method is used successfully in the current study to detect contamination by a species of yeast, mold, and bacteria in intact plastic infusion bags at levels as low as three colony-forming units per milliliter for yeast. By use of the near-IR method, each injectable unit can be evaluated with its integrity maintained, allowing the product to be dispensed or evaluated by another analytical method.**

## INTRODUCTION

Biotechnology has recently developed a number of new products that are now available for treating diseases. Many of the biotechnology products of pharmaceutical interest are fragile molecules (e.g., proteins) that cannot withstand exposure to the digestive tract or be formulated as a tablet or other oral dosage form. These products must be formulated as parenteral drugs (injectables) in a sterile vial or intravenous (IV) bag. Many parenteral drugs, including those derived from biotechnology and those used in conventional drug formulations, are also sensitive to the high temperatures required for heat sterilization (autoclaving).

For these types of products, maintaining the stability of the drug (preventing decomposition) and ensuring the sterility of the product (preventing microbial growth) can be challenging problems. Preservative systems and sterilization procedures must be well monitored (*1*) and tested by validated microbiological methods (*2*, *3*). Intravenous bags or vials can be filled aseptically as an alternative to heat sterilization. This process involves filling the product container with a solution of drug and excipients after it has passed through a series of filters (usually 0.22 $\mu$m) to remove any microorganisms.

The number of parenteral products produced aseptically has increased over the last few years. This trend is alarming the Food and Drug Administration (FDA) because aseptic-filling procedures are not as rugged or safe as terminal sterilization. Autoclaving provides a sterility assurance level of $10^{-6}$ or better (probability of an unsterile unit), while aseptic filling generally achieves an assurance level of only $10^{-3}$ (*4*, *5*). Accordingly, the FDA is trying to discourage the use of aseptic filling by requiring manufacturers of aseptically filled products to submit methods and data justifying why terminal sterilization cannot be used. These manufacturers must also describe the microbiological monitoring and control procedures used to assure sterility (*4*, *6*).

The challenge to the analyst is to determine which products are contaminated and prevent their use. Perhaps the simplest method of assessing sterility involves the incubation of an injectable product for a period of time until the microorganisms in the container grow sufficiently that turbidity develops. The turbidity is then detected by visual examination. Unfortunately, it can take a significant amount of time for turbidity to develop. Furthermore, the materials of some product containers hinder visual inspection. For example, some IV bags are composed of polymers that become translucent when the bags are stored too closely together because water becomes trapped in the wall of the bag.

Ordinarily, sterility assurance and microbial identifications are accomplished by selecting a number of units within a specific lot (batch sampling) and testing them by conventional microbiological methods. These methods are invasive and destroy the units of the product being examined. These methods also tend to be imprecise, laborious, and time-consuming.

An analytical method that enables the detection of low levels of microorganisms in parenteral products without laborious random testing or the need for long incubation times would represent a significant advance in the analysis of parenteral products. Such a method could be used to detect contamination by bacteria, yeast, or molds in drug vials and IV bags. An analytical method based on near-infrared (near-IR) light scattering is proposed in this work as a method for detecting small quantities of microorganisms in sealed IV bags. This light-scattering method is noninvasive and nondestructive, preventing possible contamination of bags by the analytical method itself.

Near-IR methods provide information on molecular structure, are well suited to the analysis of aqueous samples, and have been used in the detection of contaminated products (*7*, *8*). The energy of near-IR photons is low (approximately 1 eV), making destruction of even protein drugs by the probe beam unlikely. An analysis of the distribution quantiles of near-IR spectral data provides a powerful means of interpreting light-scattering results. The principal advantage of the near-IR light scattering method is that every single unit of the product can be examined for sterility without invading the sample and destroying the product. Furthermore, the method appears to be able to differentiate between different classes of microorganisms as well as to isolate their location inside the container and determine their concentration in solution.

* Author to whom correspondence should be addressed.

0003-2700/90/0362-2514$02.50/0 

## THEORY

Spectral data are collected at wavelengths $\mathbf{N}_{(m)} = \{1, 2, ..., w\}$ on each sample bag. Treatment of collected spectral data $\mathbf{I}$ begins with a smoothing process designed to reduce spectral noise

$$\mathbf{I}_{(1)} = \mathrm{W}(\mathrm{W}(\mathbf{I})) \tag{1}$$

where W represents a linear smoothing operation in which $i_{ij} = (i_{ij-2} + i_{ij-1} + i_{ij} + i_{ij+1} + i_{ij+2})/5$. Calculation of the first derivative of the smoothed spectra removes baseline variations from the spectra of the bag

$$\mathbf{I}_{(\mathrm{d1})i} = \left|\frac{\mathrm{d}\mathbf{I}_{(1)i}}{\mathrm{d}\mathbf{N}_{(m)}}\right| \tag{2}$$

A region of interest (i.e., a wavelength region where scattering is expected to be observed from cells) is then selected in the spectra of each bag. The region in this work encompasses one-third of the recorded wavelength spectrum, leading to $s_{(t)} = [w/3]$. A separate set of derivative spectra is then calculated for the region of interest

$$\mathbf{I}_{(\mathrm{d2})i} = \left|\frac{\mathrm{d}\mathbf{I}_{(1)i\{s_t, s_t+1, s_t+2, ..., d\}}}{\mathrm{d}\mathbf{N}_{(m)\{s_t, s_t+1, s_t+2, ..., d\}}}\right| \tag{3}$$

The two spectra from each bag that show the most distinguishing spectral features are selected by

$$\mathbf{I}_{(\mathrm{s1})i} = \sum_{j=1}^{d} i_{(\mathrm{d1})ij} \tag{4}$$

$$\mathbf{I}_{(\mathrm{s2})i} = \sum_{j=1}^{d} i_{(\mathrm{d2})ij} \tag{5}$$

$$p_1 = \underline{\mathrm{M}}(\mathbf{I}_{(\mathrm{s1})}, \{1,2,...,u\}) \tag{6}$$

$$p_2 = \underline{\mathrm{M}}(\mathbf{I}_{(\mathrm{s2})}, \{1,2,...,u\}) \tag{7}$$

$$\mathbf{H}_{(1)j} = \mathbf{I}_{(1)p_1 j} \tag{8}$$

$$\mathbf{H}_{(2)j} = \mathbf{I}_{(1)p_2 j} \tag{9}$$

The distinctive spectra, $\mathrm{H}_{(1)}$ and $\mathrm{H}_{(2)}$, are combined by

$$\mathbf{H}_{(1)j} = \mathbf{H}_{(1)\{w, w-1, w-2, ..., 1\}} \tag{10}$$

$$\phi = \mathbf{H}_{(1)u} - \mathbf{H}_{(2)1} \tag{11}$$

$$\mathbf{H}_{(2)j} = \mathbf{H}_{(2)j} + \phi \tag{12}$$

$$\mathbf{T}_{i\{1,2,...,u\}} = \mathbf{H}_{(1)j} \tag{13}$$

$$\mathbf{T}_{i\{u+1, u+2, ..., 2u\}} = \mathbf{H}_{(2)j} \tag{14}$$

$$\mathbf{T}_i = \mathrm{W}(\mathbf{T}_i) \tag{15}$$

to form an augmented spectral matrix that is useful in quantitative and qualitative analysis. The augmented space $\mathbf{T}$ thus has $d = 2u$ dimensions (columns) with one row for each sample bag.

Generally, another $n$-by-$d$ matrix $\mathbf{V}$, containing validation samples, is also assembled from the same source as the training set and is likewise treated in accordance to eq 1–15. The sample set $\mathbf{V}$ serves as an indicator of how well the training set describes its overall population variation. New spectra of sample bags under test are denoted $\mathbf{X}$ and are also treated in accordance to eq 1–15 before quantitative or qualitative analysis.

Bootstrap distributions are calculated by an operation $\kappa$, and $\kappa(\mathbf{T})$, $\kappa(\mathbf{X})$, and $\kappa(\mathbf{V})$ are each calculated in this manner. The results are the $m$-by-$d$ arrays $\mathbf{B}$, $\mathbf{B}_{(\mathrm{X})}$, and $\mathbf{B}_{(\mathrm{V})}$. The operation $\kappa(\mathbf{T})$, for example, begins by filling a matrix $\mathbf{P}$ with sample numbers to be used in bootstrap sample sets $\mathbf{B}_{(s)}$

$$\mathbf{P} = p_{ij} = r \tag{16}$$

The values in $\mathbf{P}$ are scaled to the training-set size by

$$\mathbf{P} = [(n)\mathbf{P} + 1] \tag{17}$$

A bootstrap sample $\mathbf{B}_{(s)}$ is then created for each row $i$ of the $m$-by-$d$ bootstrap distribution $\mathbf{B}$ by

$$\mathbf{B}_{(s)} = t_{\mathrm{K}j} \tag{18}$$

where $\mathbf{K}$ are the elements of the $i$th rows of $\mathbf{P}$. The $q$th row of $\mathbf{B}$ is filled by the center of the $q$th bootstrap sample

$$b_{qj} = \sum_{i=1}^{n} b_{(s)ij}/n \tag{19}$$

and the center of the bootstrap distribution is

$$c_j = \sum_{i=1}^{m} b_{ij}/m \tag{20}$$

The operation $\kappa$ is then repeated by using $\mathbf{X}$ and $\mathbf{V}$.

The multivariate data in the bootstrap distributions are then reduced to a univariate form

$$s_{(\mathrm{T})i} = \left(\sum_{j=1}^{d} (b_{ij} - c_j)^2\right)^{1/2} \tag{21}$$

$$s_{(\mathrm{X})i} = \left(\sum_{j=1}^{d} (b_{(\mathrm{X})ij} - c_j)^2\right)^{1/2} \tag{22}$$

$$s_{(\mathrm{V})i} = \left(\sum_{j=1}^{d} (b_{(\mathrm{V})ij} - c_j)^2\right)^{1/2} \tag{23}$$

and these distances are ordered and trimmed according to a trimming-index set

$$\mathbf{P}_{(\mathrm{T})} = \{mp + 1, mp + 2, mp + 3, ..., m - mp\} \tag{24}$$

to reduce the leverage effects of isolated selections at the extremes of the bootstrap distributions.

Cumulative distribution functions (CDFs) for quantile–quantile (QQ) plotting are formed by

$$\mathbf{C}_{(\mathrm{T})} = \partial(\mathbf{S}_{(\mathrm{T})\mathbf{P}_{(\mathrm{T})}}, \mathbf{S}_{(\mathrm{T})\mathbf{P}_{(\mathrm{T})}}) \tag{25}$$

$$\mathbf{C}_{(\mathrm{X})} = \partial(\mathbf{S}_{(\mathrm{T})\mathbf{P}_{(\mathrm{T})}}, \mathbf{S}_{(\mathrm{X})\mathbf{P}_{(\mathrm{T})}}) \tag{26}$$

$$\mathbf{C}_{(\mathrm{V})} = \partial(\mathbf{S}_{(\mathrm{T})\mathbf{P}_{(\mathrm{T})}}, \mathbf{S}_{(\mathrm{V})\mathbf{P}_{(\mathrm{T})}}) \tag{27}$$

Graphing either $\mathbf{C}_{(\mathrm{X})}$ or $\mathbf{C}_{(\mathrm{V})}$ on the ordinate vs $\mathbf{C}_{(\mathrm{T})}$ on the abscissa produces a standard QQ plot. Patterns in the QQ plot can be used to analyze structure in the spectral data, and the significance of the correlation between $\mathbf{C}_{(\mathrm{T})}$ and $\mathbf{C}_{(\mathrm{X})}$ can be used as an indication of the existence of subclusters in the spectral data. In the plot, a straight line with unit slope and an intercept of 0 indicates that the two CDFs are essentially identical (this line should be observed when $\mathbf{C}_{(\mathrm{V})}$ is on the ordinate and $\mathbf{C}_{(\mathrm{T})}$ is on the abscissa). The presence of breaks in the line indicates that the CDF on the ordinate is multimodal (i.e., that the test set and training set of samples are not the same). Sharp bends in the QQ line also indicate the presence of more than one distribution in the CDF on the ordinate.

The Pearson product moment correlation coefficient between the two integrals or CDFs is used as a means of quantifying the differences between the test set and training set. The correlation between the two integrals decreases steadily with time when an IV bag is contaminated. The correlation coefficient can be used to provide both an indication of the number of microorganisms present in a sealed container as well as how long the microorganisms have been present in the container and what kind of microorganisms are present in the container. The identification of microorganisms

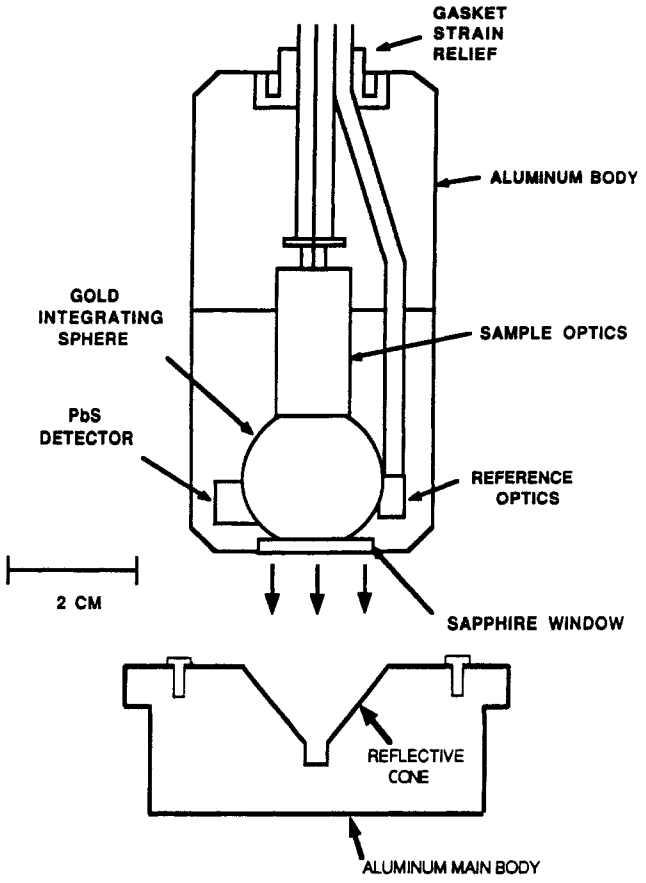


**Figure 1.** Fiber-optic diffuse-reflectance probe used to collect spectral data from a near-IR beam within a wavelength range from 1100 to 1360 nm.

is accomplished by preparing training sets of each type of microorganism expected in the bag and projecting test spectra into a training set space or library. Overlap should occur between the test group and one of the groups in the training set library if the test bags are contaminated with one of the microorganisms used to develop the training set library.

## EXPERIMENTAL SECTION

This determination of microorganisms in parenteral products is based upon the proposed scattering of near-IR light by solid objects inside sealed IV bags. Monochromatic, near-IR light is directed into the sample, and the solid material in the sample appears to scatter light back into an integrating sphere for collection and detection. (It is possible that the spectra result from a process other than scattering from cells in the sample, such as transmission through the bag and absorption by molecules, or other processes.) The fiber-optic diffuse-reflectance probe shown in Figure 1 (Bran+Luebbe, Inc., Elmsford, NY) was used to collect spectral data from a near-IR beam within a wavelength range from 1100 to 1360 nm (total collection time = 22 s). Light is directed into the sample from a fiber-optic bundle that is placed in the 1-in. gold integrating sphere directly opposite the sample window (or beam port). A reference fiber-optic bundle directs near-IR light onto the wall of the integrating sphere. The dual sample and reference bundle configuration simulates a double-beam spectrometric system to compensate for noise caused by bending of the fiber and source-intensity variations. Signal values are recorded as a ratio of intensities between the sample and reference beams. The logarithm of the reciprocal of the reflectance intensity is transmitted to a computer for analysis.

**Equipment.** Near-IR energy is transmitted through the optical fibers using an InfraAlyzer 500 scanning spectrophotometer (Bran+Luebbe). Data are collected on an IBM PS/2 Model 50 computer (IBM Corp., Armonk, NY) running IDAS software (Bran+Luebbe). Collected reflectance values are then transferred to a MicroVAX II computer system (Digital Equipment Corp., Maynard, MA) and an IBM 3090-300E vector supercomputer. Spectral data are processed in Speakeasy IV Epsilon (Speakeasy Computing Corp., Chicago, IL) programs that were written specifically for this purpose.

**Materials.** Thirty poly(vinyl chloride) (PVC) IV bags containing 5% Dextrose Injection USP (Viaflex 150-mL containers, lot no. CO92445, Baxter Healthcare Corp., Deerfield, IL) were injected with 3 mL of Zantac Injection (lot no. B2997LB, 25 mg/mL ranitidine as the hydrochloride, Glaxo, Inc., Research Triangle Park, NC). Bags were injected through the additive port with a sterile disposable syringe and 21 gauge × 1.5 in. needle (Becton Dickinson, Rutherford, NJ). The nominal ranitidine concentration in each bag was approximately 0.5 mg/mL.

The microorganisms injected were *Candida albicans* (American Type Culture Collection number 10231), *Aspergillus niger* (ATCC no. 16404), and *Pseudomonas aeruginosa* (ATCC no. 9027). These microorganisms were chosen to include a species of yeast, mold, and bacteria, respectively, which are typically tested to meet USP and FDA requirements.

Inoculum was prepared by transferring the respective microorganism from a lyophilized culture onto a solid agar medium and incubating at suitable temperatures for sufficient growth. For *Pseudomonas aeruginosa*, Trypticase Soy Agar was used, and the incubation time was 18–24 h. Sabouraud Dextrose Agar was used for *Candida albicans* and *Aspergillus niger* with incubation times of 40–48 h and 7 days, respectively. These agars and incubation times are consistent with harvesting procedures for pharmaceutical microbiological assays (*9*).

Cells were harvested into a sterile conical tube with 5% Dextrose Injection USP instead of sterile saline TS (*9*) to be consistent with the diluent used in the IV bags. Cell concentrations for each species were adjusted to a target range of 10–100 colony-forming units (cfu) per 0.10 mL using 5% Dextrose Injection USP. This range was selected to give a starting target concentration of approximately 1 cfu/mL per bag, which represents a reasonable contaminant load for a sterility violation. The number of colony-forming units per milliliter in the inoculum of each species was determined in quadruplicate by the spread-plate method. The average inoculum concentrations from four plates were 1650, 100, and 120 cfu/mL for *Pseudomonas aeruginosa*, *Candida albicans*, and *Aspergillus niger*, respectively.

The additive port of each of the 30 bags was injected with 0.10 mL of inoculum from one of the three microorganisms (10 bags of each type). The bags were inverted several times to distribute the cells throughout the bag.

**Data Analysis.** A spectral training set was constructed for each group of 10 bags containing a single variety of microorganism. The spectral training set was collected immediately after injection of the microorganisms. Spectra were also obtained from the bags before injection of the microorganisms. These spectra, however, were not used as the training set because the near-IR method appeared to detect the injection of medium and microorganisms, which results in a large disturbance in the spectra. All ten bags containing the same organism were inoculated sequentially prior to the training-set scans. The time lag between the scanning of the first bag and tenth bag was approximately 1 h. Furthermore, 12 scans of the wavelength range from 1100 to 1360 nm were taken from each bag at different portions of the bag. Therefore, each training set consisted of 10 IV bags containing one of three microorganisms in a 5% dextrose solution with drug. Twelve scans were taken from each bag so that each of the three training sets contained 120 spectral scans. During spectral analysis a spectrum recorded at 130 wavelengths in the 1100–1360 nm region was projected as a single point in a 260-dimensional hyperspace. Thus, each training set was composed of a cluster of 120 points in a 260-dimensional hyperspace. The analytical procedure located the center of the training set from the spectra recorded at time zero in a 260-dimensional space and integrated outward steadily in all directions in space from the center of this training set to the "edges" of the training-set cluster (the edges are defined typically as being three multidimensional standard deviations away from the center). This integral forms a function that is compared to a second integral, which is determined by integrating from the center of the combined training set and test set of spectra (where the test set is the spectra of the same group of IV bags at a later time, such as 6, 12, 24, or 48 h after the injection of microorganisms). The 260-dimensional points from these test-set

**Table I. Concentrations (cfu/mL)[a] of Microorganisms at 22 °C and 32 °C in 150-mL[b] Viaflex IV Bags Containing 5% Dextrose and 3 mL of Zantac Injection**

| | *Aspergillus niger* | | | *Candida albicans* | | | *Pseudomonas aeruginosa* | | |
|---|---|---|---|---|---|---|---|---|---|
| time, h | 22 °C | 32 °C | 32 °C[c] | 22 °C | 32 °C | 32 °C[c] | 22 °C | 32 °C | 32 °C[c] |
| | 100[d] | 100 | 120 | 100 | 100 | 100 | 1000 | 1000 | 1650 |
| 0[e] | 0 | 0 | 2 | 0 | 0 | 0.3 | 1 | 1 | 3 |
| 6 | 3 | 2 | 3 | 4 | 3 | 3 | 14 | 27 | 57 |
| 12 | 12 | 10 | 12 | 13 | 15 | 13 | 88 | 135 | 199 |
| 18 | | | 65 | | | 80 | | | 725 |
| 24 | 174 | 150 | 145 | 610 | 475 | 553 | 954 | 1430 | 2325 |
| 48 | 530 | 465 | | 1850 | 1250 | | 22150 | 36500 | |

[a] Each value represents an average determined from two bags. [b] Commercial IV bags typically contain an overfill volume of 10%. The actual average volume of the bags, including diluent and drug, was 170 mL. [c] Values determined from bags 1 and 10 scanned by the near-IR spectrometer. [d] Initial inoculum. [e] Concentration measured just after bags were injected with 0.1 mL of above inoculum.

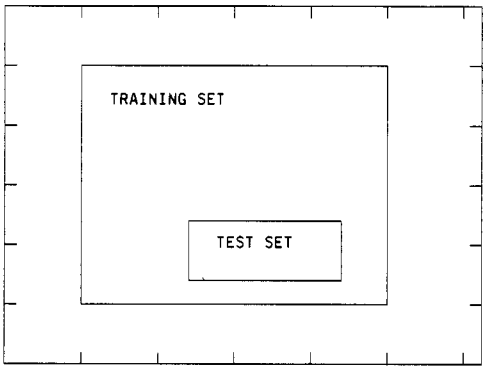


**Figure 2.** Schematic diagram of the training set and test set projection process, showing simultaneous differences in location and scale.

bags scanned at the later times are projected into the same space as the training-set spectra to form an augmented spectral cluster. Integrating from the center of the augmented set out in all directions at a constant rate produces a second integral. A plot of the first integral versus the second integral is used to form a QQ plot. Figure 2 is a schematic diagram of the training set and test set projection process. Figure 3 is a QQ plot of two slightly different integrals that result from the projection of a test set into an augmented-set space when the test set is slightly larger than the training set.

Microorganism concentrations in bags 1 and 10 were measured at 0, 6, 12, 18, and 24 h by removing 0.40 mL of solution from each bag with a 0.5-mL sterile syringe. A 0.10-mL aliquot (or diluted aliquot at high microbiological concentrations) was transferred to each of four plates to determine the average cell concentration in cfu/mL. Trypticase Soy Agar was used as a growth medium for the *Pseudomonas aeruginosa* aliquots, and colonies were counted for 48 h at 30–35 °C. Sabouraud Dextrose Agar was used as the growth medium for the *Candida albicans* and *Aspergillus niger* aliquots, and colonies were counted after 72 h and 7 days, respectively, at 20–25 °C. The results for each microorganism (cfu/mL) are shown in Table I.

## RESULTS AND DISCUSSION

**Microorganism Growth.** Each bag contained approximately 0.1 mg/mL phenol because this preservative is present in the drug formulation of Zantac Injection. Although this phenol level is insufficient to preserve the bags, it is high enough to decrease organism growth rates. The slightly elevated ambient temperature of the laboratory (30–35 °C) needed to facilitate operation of the near-IR spectrometer also had an effect on microorganism growth. The high ambient temperature increased the growth rate of *Pseudomonas aeruginosa* but decreased the growth rates of *Aspergillus niger* and *Candida albicans*. These results were determined by storing two duplicate bags for each microorganism at 22 °C and 32 °C and comparing their growth-rate profiles for a 48-h period. These supportive data are also shown in Table I. In all cases, the solutions remained clear throughout the course of the experiment with no visible signs of product contamination.

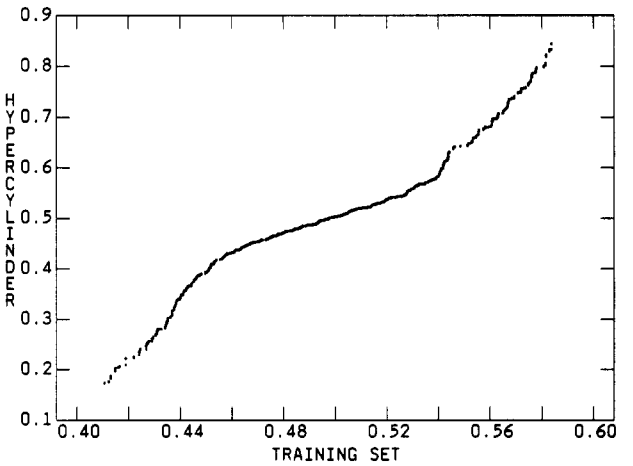


**Figure 3.** Quantile–quantile plot of two slightly different integrals that result from the projection of a test set into an augmented-set space when the test set is slightly different from the training set (here, the test set is slightly larger in volume than the training set).

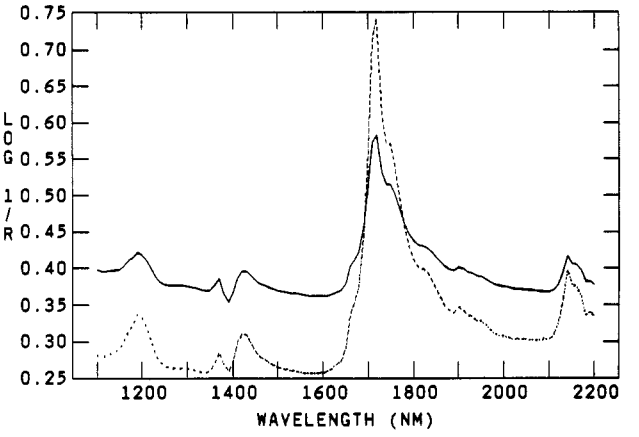


**Figure 4.** Near-IR spectra of the polymer forming the Viaflex bags. The spectra were obtained at different locations on a single bag.

The discrepancy between the inoculum concentrations and the microbial concentrations measured in the bags just after inoculation (cf. Table I) appears to be caused by differences in sampling procedure. Inoculum concentrations were determined from 1.0-mL aliquots removed from the inoculum with a volumetric pipet. Microbial concentrations in IV bags were determined from 0.10-mL aliquots removed with a syringe. A 0.10-mL sampling size per test was used to minimize the volume of liquid removed from the bags. Given the goals of the experiment and the inherent variability associated with microbiological testing, these discrepancies do not pose a major concern.

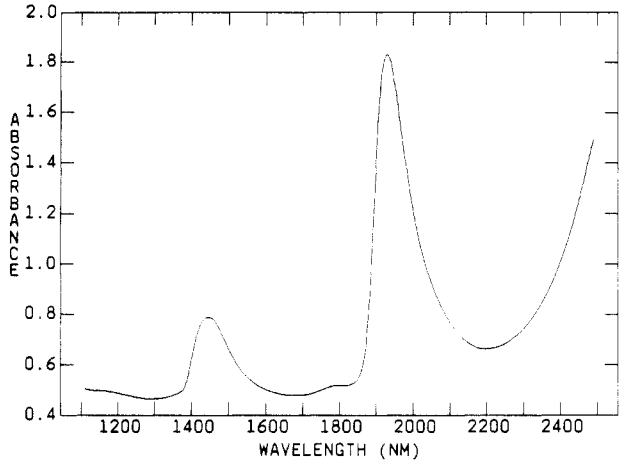


**Figure 5.** Near-IR spectrum of water in a similar spectral region to that shown in Figure 4.

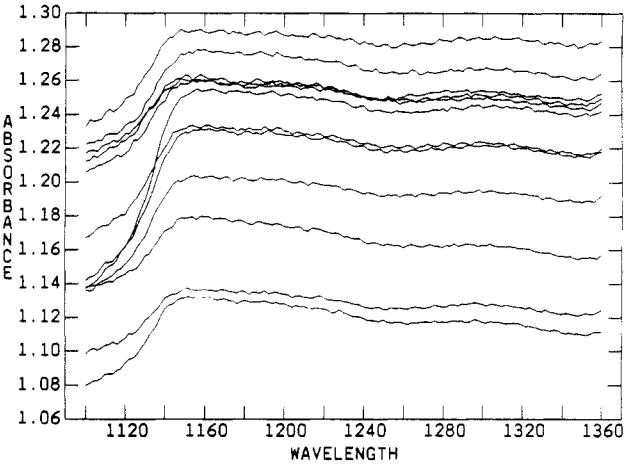


**Figure 6.** Twelve raw spectra (not processed by any filtering method) taken from a single uncontaminated bag.

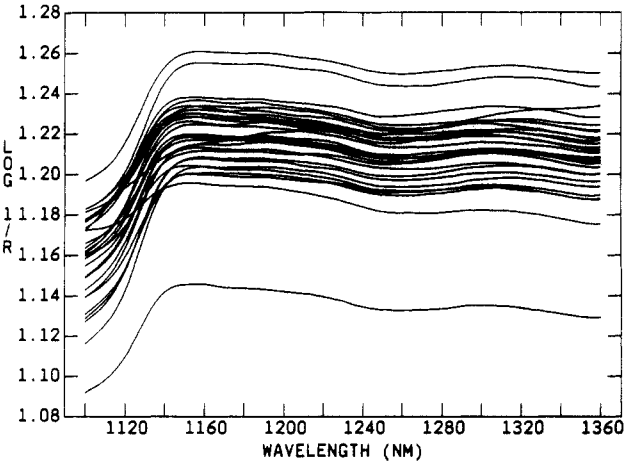


**Figure 7.** Smoothed spectra from 30 uncontaminated, 150-mL, 5% dextrose bags containing drug.

**Near-IR Results.** Figure 4 shows two near-IR spectra of the polymer forming the Viaflex bags. The two baselines and peak heights differ because the readings were taken at different locations on the Viaflex bag that had different thicknesses. Figure 5 is the near-IR spectrum in the same region for water. It is apparent from an examination of these two figures that there are relatively narrow spectral regions that may show high sensitivity to scattered light from cells inside vials or bags. The regions around 1450 and 1940 nm are effectively obscured by intense water absorbance. Absorbance of scattered light by the IV bag polymer occurs around 1720 nm. Therefore, measurements of light scattered from cells and returning through the water and the bag into the detector in the fiber optic probe should be easiest in the 1100–1360 nm region, in a small region around 1600 nm, and in the 2000–2200 nm region. Unfortunately, the background absorbance of water in the 2000–2200 nm region is quite high, and consequently this region is virtually useless unless all of the material one wishes to examine is adhered to the bag wall, a situation which minimizes the amount of water that the signal must pass through. The fact that water absorbs more strongly in the "windows" around 1600 and 2200 nm than in the window from 1100 to 1360 nm means that one should be able to localize the growth of microorganisms in a kind of depth profiling by looking at spectral absorbances near 1100, 1600, and 2200 nm. Light scattering from free-floating microorganisms should appear mainly in the 1100–1360 nm region. However, microorganisms adhering to the walls of the container should appear at the 1600 and 2200 nm regions as well as in the 1100–1360 nm region. In fact, one might expect them to ppear more strongly in the 1600 and 2200 nm regions than the 1100–1360 nm region because their absorption coefficients should be higher at the higher near-IR wavelengths than at the lower near-IR wavelengths. Therefore, because the path length for materials adhered to the wall would be very limited, signals for microorganisms adhering to the walls would be expected to be more intense in the 1600 and 2200 nm wavelength regions. Spatial profiling can also be accomplished with the near-IR method. A three-dimensional picture of the contents of the bag can be roughly obtained if the bag is held motionless and multiple scans are obtained by moving the fiber-optic probe. In this way, if spectral changes begin to appear routinely at certain locations, such as near an injection port, the problem is already nearly isolated.

In a preliminary study, full spectral scans from 1100 to 2200 nm were obtained from two innoculated bags (*pseudomonas* injection of 10–100 cfu). No significant absorbances were observed in the 1600 and 2200 nm regions, and it was believed that microorganisms injected into the bags were floating freely in solution. These experiments were therefore confined to the 1100–1360 nm region. Figure 6 shows 12 spectra taken from a single bag. These spectra are raw spectra and are not processed by any filtering methods. They cover the entire spectral range from 1100 to 1360 nm. The spectra appear to be relatively noisy because very little light is actually reflected back into the probe by scattering from the solution. During the first few hours of cell incubation, few cells are in the solution and very little contamination exists.

Spectra are filtered digitally and the 12 spectra taken from a single bag are processed to combine them into a single spectrum. The process of digitally filtering the spectra produces smooth and relatively noise-free curves such as those seen in Figure 7. The spectra in Figure 7 come from 30 clean (i.e., uncontaminated), 150 mL, 5% dextrose bags containing drug. In the early stages of contamination, it is important to move the probe around the bag because scattered light will appear only in a few locations. The observation of scattering in the first few hours of incubation appears to be somewhat of a statistical phenomenon.

The mathematical problem is first one of identifying which of the 12 spectral scans from a bag actually show backscattered light. Identifying and quantifying the cells are then accomplished by using these particular scans. The first derivative is calculated for each of the 12 spectra and the absolute values of the first derivative in the regions near 1100 nm and 1260 nm are examined closely. The sum of the absolute values of the first derivatives in these two regions is calculated for each of the 12 spectra, and the spectra showing the maximum sums are used to create a new spectrum. If the same spectrum has

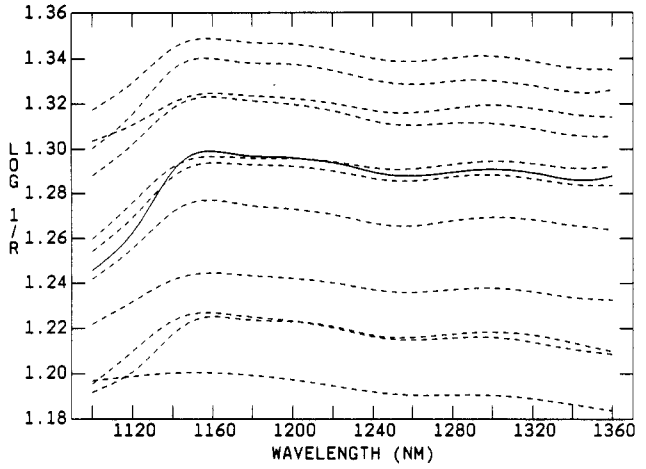


**Figure 8.** Twelve scans taken from an uncontaminated bag. The solid line shows the curve with the maximum absolute value of the first derivative.

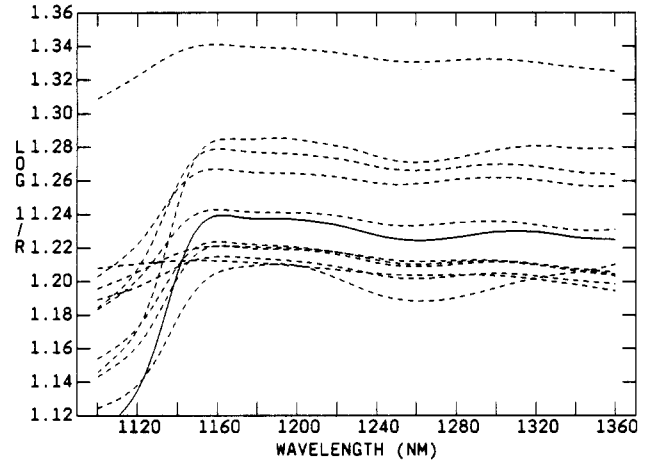


**Figure 9.** Twelve scans taken from a contaminated bag containing bacteria. The solid line shows the curve with the maximum absolute value of the first derivative.

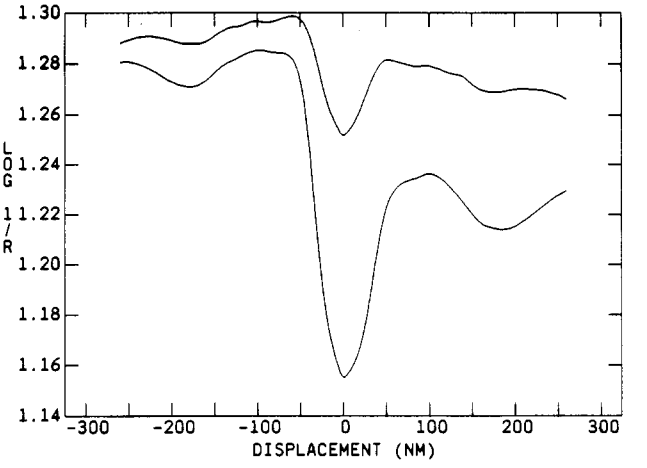


**Figure 10.** Processed scans for both uncontaminated and contaminated IV bags. The lower curve is obtained from a contaminated bag while the upper curve is obtained from an uncontaminated bag.

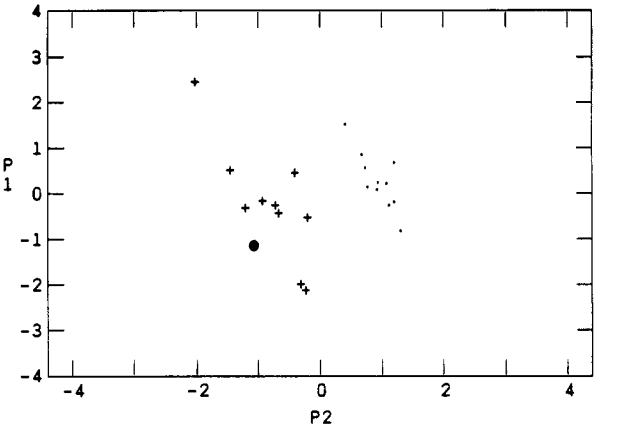


**Figure 11.** The projection of spectra of a clean bag (given by points) and a contaminated bag (given by pluses) on a plane in multidimensional hyperspace. The contaminated bag produces a larger spectral cluster in hyperspace, presumably because its spectra are more variable.

the maximum absolute value of the first derivative at both wavelengths, then it is the only spectrum selected, and the resulting normalized curve is symmetrical around the zero wavelength displacement point.

Figures 8 and 9 demonstrate why this spectral preprocessing is necessary. Figure 8 shows 12 scans taken from an uncontaminated bag. The solid line shows the curve with the maximum absolute value of the first derivative. Figure 9 shows 12 scans taken from a contaminated bag containing bacteria. The solid line again shows the spectrum with the maximum absolute value of the first derivative. It is evident that in clean bags the major source of spectral variation is a baseline variation that is predominantly path length dependent. In contaminated bags, however, certain spectra will show large backscattering peaks that appear as dips in the spectra near 1100 and 1260 nm. Other scans on contaminated bags will show no backscattering at all. The value of the preprocessing technique for IV bag spectra becomes apparent when one examines Figure 10, which shows scans for both uncontaminated and contaminated IV bags. In Figure 10, the displacement value of zero represents the backscattering observed at 1100 nm. The displacements that appear at 160 nm (both positive and negative) represent scattered light observed as 1260 nm in the original spectra. In Figure 10, the lower curve is obtained from a contaminated bag while the upper curve is obtained from a clean uncontaminated bag. The preprocessing and filtering procedure is used to select the spectra that show the most backscattering of light, and these spectra are transformed to principal axes and used in the hyperspace integration method.

When the test set spectra are different from the training set spectra, bends or breaks appear in the line of the QQ plot. The presence of bends (where two lines appear in the QQ plot with different slopes) indicates the presence of two groups in hyperspace with different sizes. The presence of a break in the QQ plot line indicates two groups in space centered at different locations. The presence of both a bend and a break indicates that two groups have different sizes and locations in multidimensional hyperspace. Applying linear regression to the points on the QQ plot produces an equation whose linear coefficients have particular significance. When the spatial volume of the test set is smaller than that of the training set, the slope and intercept of the linear equation, determined by regression, have values between 0 and 1. However, when the volume of the test set is larger than the training set, the coefficients of the straight line through the QQ plot tend to have a large positive slope and a large negative intercept. Confidence limits are set on the correlation between the two CDFs in the QQ plot. The confidence limits are set through a bootstrap process similar to that used in eqs 16–20.

Figure 11 depicts the projection of spectra of a clean bag (given by points) and a contaminated bag (given by pluses) on a plane in multidimensional hyperspace. The plane corresponds to that defined by the first and second principal axes. Figure 11 demonstrates that contaminated bags produce spectral points in hyperspace that are more widely scattered

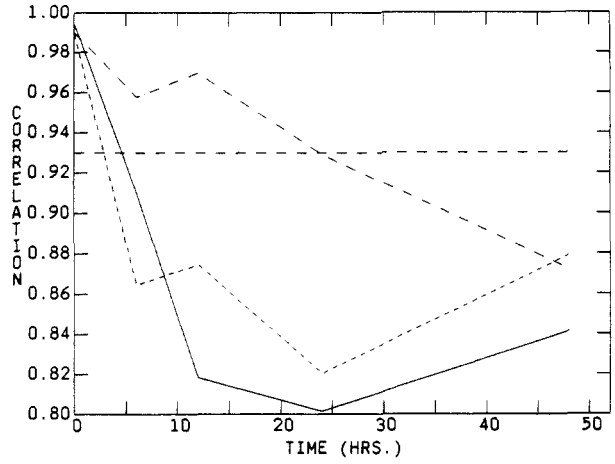


**Figure 12.** Growth curve measured by backscattering of near-IR light from the bags. The solid line represents the bacterium, the short-dashed line represents the yeast, and the dashed–dotted line represents the mold. The horizontal long-dashed line represents the 98% confidence limit.

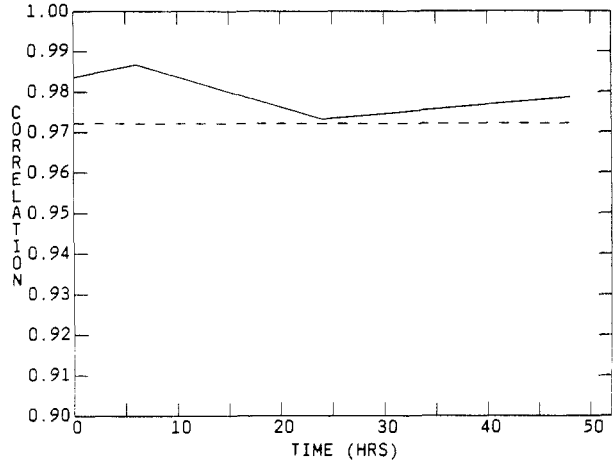


**Figure 13.** Spectra from two control bags over time. These bags are exactly like the experimental bags, except that they were injected with a sterile medium.

than clean bags. The spectra in Figure 11 represent 12 scans taken at various locations on each of the two bags (one clean bag spectrum with an A/D (analog/digital) spike is eliminated). The fact that the pluses from the contaminated bag do not overlap the cluster formed by the points from the clean bag indicates that a spectral difference exists between the clean bag and the dirty bag. The distance between the cluster of points formed from spectra of the clean bag and the cluster formed from the contaminated bag provides an indication of the amount of material that is responsible for the contamination of the dirty bag. The direction of the displacement from the center of the clean bag provides an identifying spectrum of the material responsible for the contamination. Thus, distance gives an indication of the number of microorganisms that are present in the bag, while direction identifies the microorganisms present in the bag that are responsible for the contamination.

Table I shows the microorganism concentrations in IV bags for the yeast, *Candida albicans*; the bacteria, *Pseudomonas aeruginosa*; and the mold, *Aspergillus niger*. These concentrations were obtained with standard microbiological assays. Figure 12 is calculated from scans that have been averaged for each of the 10 bags used to form Table I. Figure 12 is, in effect, a "growth curve" of sorts, measured by transmission and backscattering of near-IR light from the bags and/or retroreflector. The 98% confidence limit on backscattering from a clean training set is given as the horizontal short-dashed line in Figure 12 (the level is slightly above 0.93). The solid line represents the bacteria, the dotted line represents the yeast, and the dashed–dotted line represents the mold. After inoculation with 165 cfu per IV bag, the bacteria (*Pseudomonas aeruginosa*) were observed at about 6 h at an average concentration of 57 ± 4 cfu/mL per bag). After injection of 10 cfu yeast (*Candida albicans*) per bag, growth was detected at about 4 h with less than an average concentration of 3 ± 1 cfu/mL per bag. Injection of the mold, *Aspergillus niger*, at a level of 12 cfu per bag allowed detection of growth by near-IR spectrometry at about 24 h at an average concentration of 145 ± 20 cfu/mL per bag. Figure 13 shows spectra from two control bags over time. These bags are exactly like the experimental bags, except that they were injected with a sterile medium instead of cell cultures. The spectra show no significant change over time.

In Figure 12, it appears that *Candida albicans* is the fastest growing species in the bags. However, yeast cells range in size from 3 to 14 μm and are larger than the bacterial cells (about 0.5–2 μm). At least initially, this size advantage might make yeast a better source of light-scattering material than the bacteria, which actually grows faster. Eventually the bacterial growth appears to overtake the size advantage enjoyed by the yeast, and the bacteria then provide the strongest light backscattering signal. Mold is intermediate in size (approximately 3–8 μm) and the slowest growing species as indicated by its continued correlation change up to nearly 50 h, while the correlation of the yeast and bacteria has begun to level off, presumably due to the presence of the preservative (phenol) in the drug. What is most noticeable in Figure 12, however, is that even at 6 h and below, the near-IR method is able to detect contamination at a 98% confidence limit for yeast and bacteria. Table I shows that neither the yeast nor the bacteria have grown to anywhere near their final concentrations after this short period of time. Nevertheless, the near-IR method appears most sensitive during this period and is able to detect this contamination. The correlation between spectral clusters at 6 h is poor for *Candida albicans*, *Pseudomonas aeruginosa*, and *Aspergillus niger*. The poor correlation between the spectral clusters at all times (at 6 h and beyond) suggests that the near-IR method is able to differentiate between these cell types as well as possibly to provide an indication of the extent of their growth. Table II shows correlations between spectral groups 46 h after injection. The correlations of each of the groups at 48 h to their respective groups at injection (0 h) are approximately the same. Thus there is less likelihood that the different intergroup correlations would reflect only a difference in the correlations of each group to their initial spectra. At 48 h, the near-IR method appears able to differentiate between the mold and yeast and between the mold and bacteria, but cannot differentiate between yeast and bacteria.

**Table II. Correlations between Spectral Groups in Hyperspace at 48 h after Injection**

| | training group | | |
|---|---|---|---|
| test group | bacteria | yeast | mold |
| bacteria | 0.985 78 | 0.986 56 | 0.930 45 |
| yeast | 0.979 28 | 0.980 25 | 0.917 91 |
| mold | 0.930 64 | 0.942 30 | 0.996 47 |

## CONCLUSIONS

These data suggest that changes in near-IR spectra, taken through the IV bags with a fiber-optic probe and without product destruction, correlate to organism growth. Moreover, spectra also appear to distinguish between bags contaminated with different classes of microorganisms. With further de-

velopment, it should be possible to implement this technique as an on-line sterility assurance method in parenteral-production facilities, particularly for filling processes (e.g., aseptic fill) that require very careful control and monitoring because of less than desirable assurance levels. The inability to test all parenteral units in an automated fashion is a serious limitation to conventional microbiologic testing, particularly in cases where microbial contamination is not distributed uniformly throughout a batch (*1*, *5*). It is very difficult, if not impossible, to detect a small percentage of contaminated units within a large batch. Near-IR spectrometry with a fiber-optic probe can be used as an alternative or adjunct method to conventional microbiologic testing in quality assurance and other applications where large quantities of cells must be identified and quantified in a relatively short period of time.

The variation in replicate spectra taken from the same IV bag is larger than that observed with other containers because of the flexibility of the PVC plastic and poor near-IR transparency. To reduce the number of replicate scans needed and to improve confidence statistics, further optimization and improvements are needed in the sampling procedure (i.e., configuration of the optical probe, sample container, and wavelength range scanned).

Glass containers positioned directly in contact with the optical probe should provide even better results because glass is more rigid and offers greater transparency in the near-IR spectral region. Experiments are underway to determine the reliability of using near-IR spectrometry for distinguishing sterile and contaminated solutions in glass containers and differentiating microbial contaminants in those containers. This work should also provide more information on the source of the analytical signal (scattering from cells or absorption by solution) that causes the drop in correlation observed in contaminated bags.

## LIST OF SYMBOLS

Special defined operations:

W, linear ("moving average") smoothing

$d(f(x))/dx$, derivative of $f(x)$

$\underline{M}(f(x),x)$, $x|(d(f(x))/dx) = 0 \wedge (d^2(f(x))/dx^2) < 0$

r, random number on $0 < x < 1$; Monte Carlo integration of a continuous uniform distribution

$\kappa(\mathbf{Z})$, creates a bootstrap distribution containing $m$ elements for a set of real samples, and finds the center of this bootstrap distribution

$[x]$, the greatest-integer function of a scalar, matrix, or array

$\partial(x)$, ordered elements of $x$ ($x$ is a matrix or array)

=, equals, or "is replaced by" when the same variable appears on both sides of =

Scalars:

$n$, the training-set, test-set, and validation-set size, i.e., the number of samples that the set contains

$d$, the number of wavelengths and the dimensionality of the analytical space

$m$, the number of sample-set replications forming a bootstrap distribution (user determined)

$i$, an index for counting rows in a matrix or array

$j$, an index for counting columns in a matrix or array

$p$, proportion of a distance distribution to trim from each end of the distribution

$u$, the number of spectra collected from a single sample bag

$w$, the number of wavelengths collected from a single sample bag

$s_t$, an index marker for a wavelength region of interest

$p_1$, the index number of a spectrum showing the greatest overall signal in a set of $u$ spectra

$p_2$, the index number of a spectrum showing the greatest analytical signal over a wavelength region of interest in a set of $u$ spectra

$\phi$, the difference between the absorbances of two spectra from a single bag at the lowest wavelength

Matrices, vectors, and arrays:

$\mathbf{N}_{(m)} = (n_{(m)j})_w$, wavelength vector recorded by spectrometer

$\mathbf{I}_{(d1)} = (i_{(d1)ij})_{u,w}$, first derivatives of all spectra collected from a single bag

$\mathbf{I}_{(1)} = (i_{(1)ij})_{u,w}$, smoothed set of $u$ spectra collected from a single bag

$\mathbf{I} = (i_{ij})_{u,w}$, set of $u$ spectra as collected from a single bag

$\mathbf{I}_{(d2)} = (i_{(d2)ij})_{u,w-s_t}$, first derivative of region of interest

$\mathbf{I}_{(s1)} = (i_{(s1)i})_u$, sum of absolute value of first derivative of full spectra

$\mathbf{I}_{(s2)} = (i_{(s2)i})_u$, sum of absolute value of first derivative of wavelength region of spectral interest

$\mathbf{H}_{(1)} = (h_j)_w$, spectrum selected by the index $p_1$

$\mathbf{H}_{(2)} = (h_j)_w$, spectrum selected by the index $p_2$

$\mathbf{B} = (b_{ij})_{m,d}$, $m$-by-$d$ bootstrap distribution of training-set sample spectra

$\mathbf{B}_{(X)} = (b_{ij})_{m,d}$, bootstrap distribution of test-set sample spectra

$\mathbf{B}_{(V)} = (b_{ij})_{m,d}$, bootstrap distribution of validation-set sample spectra

$\mathbf{C} = (c_j)_d$, center of the bootstrap distribution $\mathbf{B}$

$\mathbf{P} = (p_{ij})_{m,n}$, training-set sample numbers selected for the bootstrap-sample sets used to calculate bootstrap distribution

$\mathbf{T} = (t_{ij})_{n,d}$, training-set sample spectra

$\mathbf{X} = (x_{ij})_{n,d}$, test-set sample spectra

$\mathbf{V} = (v_{ij})_{n,d}$, validation-set sample spectra

$\mathbf{K} = (k_j)n$, training-set sample numbers selected for a particular bootstrap sample

$\mathbf{B}_{(s)} = (b_{(s)ij})_{n,d}$, bootstrap sample set used to calculate single rows of a bootstrap distribution ($\mathbf{B}$, $\mathbf{B}_{(X)}$, or $\mathbf{B}_{(V)}$)

$\mathbf{S}_{(T)} = (s_{(T)i})_m$, Euclidean distances of training-set replicates from $\mathbf{C}$, the center of the bootstrap distribution of the training set

$\mathbf{S}_{(X)} = (s_{(X)i})_m$, Euclidean distances of test-set replicates from $\mathbf{C}$

$\mathbf{S}_{(V)} = (s_{(V)i})_m$, Euclidean distances of validation-set replicates from $\mathbf{C}$

$\mathbf{P}_{(T)} = (p_i)_{m-2pm}$, set of $(m-2pm)$ indices used for trimming distance distributions

$\mathbf{C}_{(t)} = (c_{(t)i})_{2m-4pm}$, cumulative distribution function (CDF) formed by the trimmed and ordered elements of the training-set bootstrap distribution; CDF has $(2m-4pm)$ elements

$\mathbf{C}_{(X)} = (c_{(X)i})_{2m-4pm}$, CDF formed by the trimmed and ordered elements of the test-set and training-set bootstrap distributions

$\mathbf{C}_{(V)} = (c_{(V)i})_{2m-4pm}$, CDF formed by the trimmed and ordered elements of the validation-set and training-set bootstrap distributions

## LITERATURE CITED


(1) Avallone, Henry L. *J. Parenter. Sci. Technol.* **1985**, *39* (2), 75–79.

(2) *Validation of Steam Sterilization Cycles*; Technical Monograph No. 1; and *Validation of Aseptic Filling for Solution Drug Products*; Technical Monograph No. 2; Parenteral Drug Association, Inc.: Philadelphia, PA, 1980.

(3) Lee, John Y. *Pharm. Technol.* **1989**, *13* (2), 66–72.

(4) "Quality Control Reports (The Gold Sheet)". In *F-D-C Rep.* **1988**, *22* (3), 1–6.

(5) Avallone, Henry L. *J. Parenter. Sci. Technol.* **1986**, *40* (2), 56–57.

(6) *FDA Guideline on Sterile Drug Products Produced by Aseptic Processing*; Food and Drug Administration: Rockville, MD, 1987.

(7) Lodder, Robert A.; Selby, Mark; Hieftje, Gary M. *Anal. Chem.* **1987**, *59*, 1921–1930.

(8) Lodder, Robert A.; Hieftje, Gary M. *Appl. Spectrosc.* **1988**, *42*, 556–558.

(9) *Preparation of Inoculum*, United States Pharmacopeia (USP XXII) and National Formulary (NF XVII), United States Pharmacopeial Convention; United States Pharmacopeia: Rockville, MD, 1989; Section <51>, pp 1478–1479.


RECEIVED for review May 22, 1990. Accepted August 23, 1990. This research was supported in part by grants from Glaxo, Inc., Research Triangle Park, NC; Bran+Luebbe, Inc., Elmsford, NY; the University of Kentucky Center for Computational Sciences; the National Science Foundation through Grant No. RII-8610671; and BRSG S07 RR05857-09 from the National Institutes of Health.