Bacterial Monitoring in Vials Using a Spectrophotometric Assimilation Method

Leonard J. Galante,[1] Michael A. Brinkley,[1] and Robert A. Lodder[2,3]

Received August 27, 1991; accepted October 17, 1991

Aseptic-filling processes are often used with fragile parenteral products that might be destroyed by terminal autoclaving. However, aseptic filling is not as effective as autoclaving in reducing contamination. As a result, time-consuming microbiological methods and turbidimetry are employed currently as product inspection techniques, but these processes can destroy the product and might not detect low levels of contamination. Thus, near-infrared (IR) light scattering was evaluated in this study as a new method for determining low levels of contamination noninvasively and nondestructively. A new parallel mathematical technique was used in conjunction with near-IR spectrophotometry to detect successfully contamination by several species of bacteria through intact glass vials. Using the near-IR method, products can be evaluated without introducing contamination, preserving the sample vial for dispensing or evaluation by another method.

KEY WORDS: near-infrared; contamination; bacteria; parenteral products; scattering; sterility.

INTRODUCTION

The United States Pharmacopoeial (USP) and Food and Drug Administration (FDA) standards for parenteral products require that these products be free of viable microorganisms. In spite of sterilization procedures such as autoclaving, membrane filtration, and gas sterilization, occasionally contamination of the final product can occur. In a stressed environment, bacteria can grow so slowly that weeks might pass before visible changes indicate product contamination.

Manufacturers must demonstrate that their products are sterile, generally using tests prescribed in the USP. These tests are destructive in nature and require the sampling of a specified number of units of the product. When a product that has already been distributed is determined to be contaminated, it must be recalled, often with great expense and embarrassment to the manufacturer. If only a small number of the units produced are contaminated, a random sampling procedure can easily fail to detect them. In addition, test units can become contaminated during the testing procedure, producing erroneous results and requiring repeat testing.

[1] Glaxo, Inc., Five Moore Drive, Research Triangle Park, North Carolina 27709.

[2] Division of Medicinal Chemistry and Pharmaceutics, College of Pharmacy, University of Kentucky, Lexington, Kentucky 40536-0082.

[3] To whom correspondence should be addressed.

Near-infrared (IR) macroscopic spectrometry using fiber optics was evaluated in this study as a means of detecting bacterial contamination of solutions through the walls of sealed glass vials. Near-IR light scattering from the contents of these vials was compared to bacterial cell counts obtained using the spread-plate method (1). The near-IR technique detects signals obtained through the vial walls at a series of wavelengths and uses these signals to calculate the composition of the vial contents. Corrections for interferences arising from the background and sample matrix can be made using a computerized modeling process that is applied to the near-IR signals. Thus, this spectrophotometric testing method is noninvasive and nondestructive and requires no sample preparation.

Recent work has suggested that the presence of *Candida albicans* at concentrations as low as 3 colony-forming units (cfu) per ml can be detected in PVC intravenous-administration bags with near-IR spectrometry followed by computer processing of spectra (1). *Aspergillus niger* and *Pseudomonas aeruginosa* were also determined in PVC bags in the same report. The study presented in the current report introduces new mathematical transforms, which in conjunction with near-IR macroscopic spectrophotometry, enable quantitative evaluation of solutions in glass vials contaminated with the bacteria *Staphylococcus aureus, Pseudomonas cepacia, Escherichia coli,* and *Pseudomonas aeruginosa.*

THEORY

The detection of bacterial cells present in drug solutions is based on the ability of a spectrophotometer and an associated computer to identify small changes in groups of spectra obtained from a single unit of a product. To identify these changes successfully, spectra recorded at d wavelengths are projected as single points in a d-dimensional hyperspace. Spectral points obtained from sterile vials tend to cluster in a small region of hyperspace. These points are designated the training set T. The spectral points obtained from contaminated vials are designated the test set X. When vials become contaminated, the spectra of the contaminants cause a displacement in the position of the spectral cluster in hyperspace. Spectra obtained from the same vial tend to be less reproducible when the vial is contaminated, resulting in an increase in the volume of the spectral cluster in hyperspace. Simultaneous changes in position and volume of spectral clusters can be determined by comparing two integrals:

(1) the integral of the training set T, from the center of T to the surface of T, and
(2) the integral of T, from the center to the surface, after T is augmented by X.

In actuality, T and X are never precisely known but, rather, are represented by a discrete estimate of spectral points obtained from representative vial samples. In order to increase the reliability of these estimates, all spectra collected are filtered similarly to reduce the effects of noise on the integral analysis.

The spectral filter function $\underline{S}$ is used to represent near-IR spectra in the form of smooth cubic splines that pass near, but not through, the actual spectral data values:

0724-8741/92/0300-0357$06.50/0

$$\mathbf{T} = \underline{\mathbf{S}}(\mathbf{W}_1,\mathbf{Y},\mathbf{W},t_L,\boldsymbol{\delta}) \quad (1)$$

$$\mathbf{X} = \underline{\mathbf{S}}(\mathbf{W}_1,\mathbf{Y},\mathbf{W},t_L,\boldsymbol{\delta}) \quad (2)$$

$\underline{\mathbf{S}}$ fits smooth curves constructed of cubic splines to the spectra Y(W). $\mathbf{W}_1$ specifies the independent variable values (wavelengths) to which interpolation is made. Y consists of the dependent variable values (absorbances or logarithms of reciprocal reflectances) that are interpolated. W contains the independent variable values (wavelengths) corresponding to Y and is generally (though not necessarily) the same as $\mathbf{W}_1$. The scalar t_L is a tolerance value that controls the extent of smoothing and is defined as the acceptable root-mean-square relative deviation of the fitted curve

$$\mathrm{RMS}(\mathbf{Z}) \leq t_L, \qquad \text{where} \qquad z_j = F(w_j - y_j)/\delta_j \quad (3)$$

and $\boldsymbol{\delta}$ is an array of the estimated errors (in standard deviations or SDs) in the absorbance values Y. The smoothing splines are of the form

$$F(\mathbf{W}) = A_j + B_j\Delta + C_j\Delta^2 + D_j\Delta^3 \quad (4)$$

where $w_j \leq w \leq w_{j+1}$ $[1 \leq j \leq m - 1]$, $\Delta = w - w_j$, and m is the number of wavelengths in the W and Y arrays. Note that $\mathbf{W}_1$ may have a different number of wavelengths (columns) than W and Y. The resulting filtered spectral set (either T, X, or both) has as many columns as $\mathbf{W}_1$ and as many rows as the number of spectra to be smoothed (i.e., the number of rows in Y).

The Bootstrap Error-adjusted Single-sample Technique, or BEST, is a flexible clustering procedure that is applied to the smoothed spectra. Extending the method to search for subclusters within a training set requires a filtered training set T and test set X, as well as the calculation of these sets' respective bootstrap distributions, B and $\mathbf{B}_{(X)}$.

A training set of sample spectral values (e.g., reflectance or absorbance) recorded at d wavelengths from n uncontaminated vials is represented by the $n \times d$ matrix T. Often, another $n \times d$ matrix, V, containing validation vials, is also assembled from uncontaminated vials like the training set. The sample set V serves as an indicator of how well the training set T describes the overall population variation of spectral values obtained from uncontaminated vials.

The second step of the basic BEST calls for the calculation of bootstrap distributions. Bootstrap distributions can be calculated by an operation κ; $\kappa(\mathbf{T})$, $\kappa(\mathbf{X})$, and $\kappa(\mathbf{V})$ are calculated in this manner. The results of the κ function are the $m \times d$ arrays B, $\mathbf{B}_{(X)}$, and $\mathbf{B}_{(V)}$, respectively.

The operation $\kappa(\mathbf{T})$ begins by filling a matrix P with sample numbers to be used in bootstrap sample sets $\mathbf{B}_{(s)}$:

$$\mathbf{P} = P_{ij} = \tau \quad (5)$$

The values in P are scaled to the training-set size by

$$\mathbf{P} = [(n - 1)\mathbf{P} + 1] \quad (6)$$

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

$$\mathbf{B}_{(s)} = t_{Kj} \quad (7)$$

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

$$b_{qj} = \sum_{i=1}^{n} b_{(s)ij}/n \quad (8)$$

and the center of the bootstrap distribution is

$$c_j = \sum_{i=1}^{m} b_{ij}/m \quad (9)$$

The operation κ is then repeated using the vial spectra in X and V.

The multivariate data in the bootstrap distributions are then reduced to univariate forms:

$$s_{(T)i} = \left[\sum_{j=1}^{d} (b_{ij} - c_j)^2\right]^{1/2} \quad (10)$$

$$s_{(X)i} = \left[\sum_{j=1}^{d} (b_{(X)ij} - c_j)^2\right]^{1/2} \quad (11)$$

$$s_{(V)i} = \left[\sum_{j=1}^{d} (b_{(V)ij} - c_j)^2\right]^{1/2} \quad (12)$$

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

$$\mathbf{P}_{(T)} = \{mp + 1, mp + 2, mp + 3, \ldots, m - mp\} \quad (13)$$

to reduce the leverage effects of isolated selections at the extremes of the bootstrap distributions. A hypercylinder can be formed about the line connecting C to the center of $\mathbf{B}_{(X)}$ or $\mathbf{B}_{(V)}$, giving directional selectivity to the information in $\mathbf{S}_{(T)}$, $\mathbf{S}_{(X)}$, and $\mathbf{S}_{(V)}$ if desired in large hyperspaces. $\mathbf{S}_{(T)}$, $\mathbf{S}_{(X)}$, and $\mathbf{S}_{(V)}$ then have n_h elements instead of m elements.

The validation of the spectrometric method for analysis of vials is complicated by the introduction of the hypercylinder. In addition, the results of the method do not appear to be superior when the hypercylinder construct is applied to spectra obtained from the vials. Thus, the hypercylinder implementation of the BEST is not employed in the analysis of the vials.

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

$$\mathbf{C}_{(t)} = \partial[\mathbf{S}_{(T)\mathbf{P}_{(T)}}', \mathbf{S}_{(T)\mathbf{P}_{(T)}}] \quad (14)$$

$$\mathbf{C}_{(X)} = \partial[\mathbf{S}_{(T)\mathbf{P}_{(T)}}', \mathbf{S}_{(X)\mathbf{P}_{(T)}}] \quad (15)$$

$$\mathbf{C}_{(V)} = \partial[\mathbf{S}_{(T)\mathbf{P}_{(T)}}', \mathbf{S}_{(V)\mathbf{P}_{(T)}}] \quad (16)$$

Plotting the elements of $\mathbf{C}_{(T)}$ on the abscissa versus the elements of either $\mathbf{C}_{(X)}$ or $\mathbf{C}_{(V)}$ on the ordinate produces a quantile–quantile (QQ) plot (2). Patterns in a QQ plot can be used to analyze structure in the spectral data obtained from the vials. The correlation between $\mathbf{C}_{(T)}$ and $\mathbf{C}_{(X)}$ can be used as an indication of the existence of subclusters in the spectral data that reflect contamination in the vials. In the QQ plot, a straight line with unit slope and an intercept of O indicates that the two cumulative distribution functions are essentially identical [as should be observed when $\mathbf{C}_{(V)}$ is on the ordinate or when the test vial is uncontaminated]. In the extended BEST QQ plot, the presence of breaks in the line indicates that the CDF on the ordinate is multimodal and that the test spectra and training spectra are not the same (i.e., the test

vial is contaminated). Sharp bends in the QQ line are also indicative of the presence of more than one distribution in the CDF on the ordinate (i.e., the presence of contamination).

MATERIALS AND METHODS

The spectrometric determination of microorganisms in parenteral products is based upon the proposed scattering of near-IR light by solid objects inside sealed glass vials. Monochromatic, near-IR light is directed into the sample and then scattered back by the solid material in the sample into an integrating sphere for collection and detection. Changes in the spectra can also result from processes other than the scattering of light by cells in the vial, such as the transmission of light through the vial and absorption by molecules in solution. The fiber-optic diffuse-reflectance probe shown in Fig. 1 (Bran + Luebbe, Inc., Elmsford, NY) was used to collect spectral data over a wavelength range from 1100 to 1360 nm. In this macroscopic detection procedure, light is directed into the sample vial 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 bundle and source-intensity variations. A pair of aluminum covers is used to reduce stray light that can enter the macroscope through the vial. 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 was transmitted through the optical fibers using an InfraAlyzer 500 scanning spectrophotometer (Bran + Luebbe). Reflectance values were collected on a MicroVAX II computer system (Digital Equipment Corp., Maynard, MA) and an IBM 3090-600J vector supercomputer. Spectral data were processed in Speakeasy IV Epsilon and Zeta (Speakeasy Computing Corp., Chicago, IL) programs written specifically for this purpose.

Materials. Forty clear glass vials (10 ml, 20 mm, Type I borosilicate; West Company, Lancaster PA) containing a nutrient medium were divided into five groups of eight vials each. The nutrient medium was trypticase soy broth (TSB) (3). Vials were filled with TSB, capped with a butyl rubber stopper and aluminum-crimp top, and then sterilized in an autoclave for 18 min at 121°C. The vials in the first four groups were injected with one of the four species of bacteria. The last group of eight vials served as control vials and were injected with only nutrient medium. Vials were injected through the rubber stopper with a sterile disposable syringe and 21 G × 1.5-in. needle (Becton Dickinson, Rutherford, NJ).

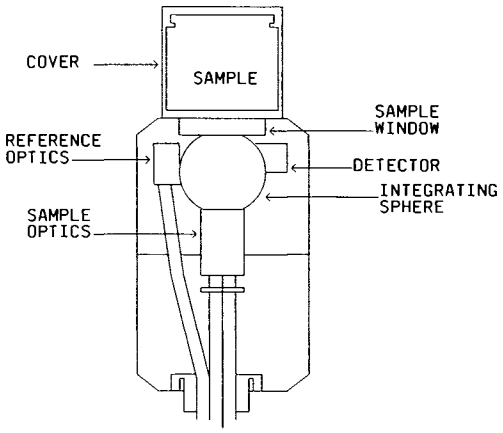

Fig. 1. The fiber-optic diffuse-reflectance probe used to collect spectral data through the walls of sealed vials.

The four microorganisms injected were *Staphylococcus aureus* (American Type Culture Collection No. 6538), *Pseudomonas cepacia* (ATCC No. 25416), *Escherichia coli* (ATCC No. 8677), and *Pseudomonas aeruginosa* (ATCC No. 9027). These microorganisms were chosen to represent the variety of bacteria that must be monitored in order to meet USP and FDA requirements.

Sample Preparation. Inoculum was prepared by transferring the respective microorganism from a lyophilized culture onto a trypticase soy agar medium and incubating for 18–24 hr for sufficient growth. The agar and incubation time are consistent with harvesting procedures for pharmaceutical microbiological assays (3).

Cell concentrations of each inoculum were adjusted to a target range of 1–50 colony-forming units (cfu) per 0.10 ml by diluting with TSB. This range was selected to produce a starting concentration of about 1 cfu/ml/vial, which represents a reasonable contaminant load for a sterility violation. The number of cfu per milliliter in the inoculum was determined in quadruplicate by the spread-plate method. The averages from four plates were 12, 120, 444, and 12 cfu/ml for *Staphylococcus aureus, Pseudomonas aeruginosa, Escherichia coli,* and *Pseudomonas cepacia,* respectively.

Each set of eight vials was injected with 0.10 ml of inoculum from one of the four microorganisms (except for control vials). The vials were inverted several times to distribute the cells throughout the vials.

Data Analysis. All eight vials containing the same microorganism in each group were inoculated sequentially prior to collecting training-set scans. Six of the vials in each group were scanned. The remaining two vials (vials 7 and 8) from each group were used to measure microorganism growth by microbiological methods. The time lag between the scanning of the first and that of the sixth vial was approximately 30 min. Ten repetitive scans of the wavelength range from 1100 to 1360 nm were taken from each vial at slightly different positions on the vial. During spectral analysis, a spectrum recorded at 130 wavelengths within the 1100 to 1360-nm region was projected as a single point in a 130-dimensional hyperspace. The analytical procedure located the center of the training set from the spectra recorded at time = 0 in a 130-dimensional space and integrated outward steadily in all directions in hyperspace 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 formed a function that was compared to a second integral, which was 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 vials at a later time, such as 6, 12, or

18, 24, or 48 hr after the injection of microorganisms). A plot of the first integral versus the second integral was 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 resulting from the projection of a test set, which was slightly smaller in volume than the training set, into an augmented-set space.

After scanning vials 1 through 6 from each group at 0, 6, 12, or 18, 24, and 48 hr, microorganism concentrations in vials 7 and 8 were measured by removing 1.0 ml of solution from each vial with a 3.0-ml sterile syringe. A 0.50-ml aliquot (or diluted aliquot at high concentrations) was transferred to each of two plates to determine the average cell concentration (cfu/ml). Trypticase soy agar was used as a growth medium for all bacteria, and colonies were counted after 48 hr at 30–35°C. The results for each microorganism (cfu/ml) are shown in Table I.

RESULTS AND DISCUSSION

A total of 40 vials was prepared containing sterile nutrient medium. Eight of these vials were injected with additional sterile medium and served as control vials over the 48-hr period. Eight vials were injected with *Staphylococcus aureus,* eight with *Pseudomonas cepacia,* eight with *Pseudomonas aeruginosa,* and eight with *Escherichia coli.* The vials from each set were scanned 10 times immediately following injection, and these spectra are depicted in Figs. 4–8 as the time = 0 correlation value.

The vials were allowed to incubate at room temperature (about 24°C) for 2 days. During incubation 10 spectra were collected from each of the vials at 6-, 12-, or 18-, 24-, and 48-hr increments following injection. Integrals calculated from the center of a cluster of spectral points in hyperspace at time = 0 were correlated to integrals calculated from the center of a respective cluster of spectral points collected at a later time. The results of these correlations are presented in Figs. 4–8.

The solid line in each of these figures represents the average correlation of six replicate analyses of vials containing a certain type of injection (either control injection or bacterial injection). The dashed lines represent ±1 standard error bar on the average correlation. The dot-dashed line signifies a 98% confidence limit on the time = 0 spectral

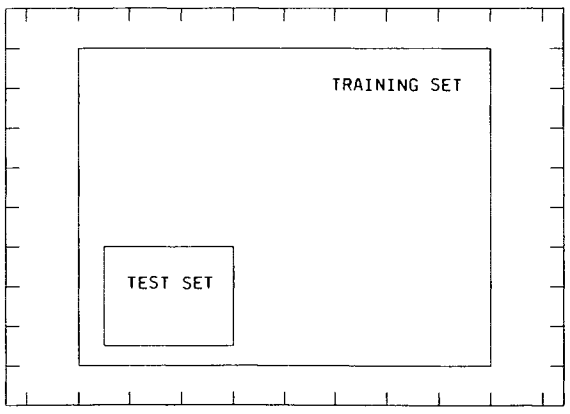

Fig. 2. A schematic diagram of the training-set and test-set projection process.

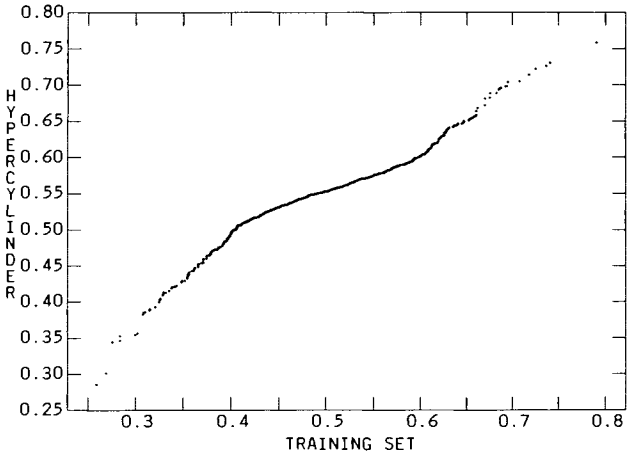

Fig. 3. Quantile–quantile plot of two slightly different integrals, the result of the test set being slightly smaller than that of the training set.

cluster of a given injection type (i.e., in 98% of observations, the correlation between two sets of spectra obtained from time = 0 spectra of a given injection is higher than this value).

Figure 4 depicts the change in spectra of the control vials with time. The spectra of the control vials do not change significantly (at the 98% level) over the 48-hr period.

Figure 5 shows the change in the spectra of the vials injected with *Staphylococcus aureus* with respect to time. The average correlation determined from these vials crosses the 98% confidence limit between 11 and 12 hr after injection of 10–100 cfu.

Figure 6 graphs the change in spectra of the *Pseudomonas aeruginosa* vials with time. The correlation for these vials crosses the 98% level about 19 hr following injection with 10–100 cfu. Interestingly, the spectra begin to resemble the time = 0 spectra again after approximately 43 hr. The rise in *Pseudomonas aeruginosa* correlations is similar to that reported for *Pseudomonas aeruginosa* in PVC administration bags (1). It appears that the growth rate of *Pseudomonas aeruginosa* in vials is different from that observed in PVC bags. However, this apparent difference in growth is probably due to differences in the medium used in the vial and bag studies as well as to differences in the optical collection efficiency.

Figure 7 shows the change in the spectra of *Pseudomonas cepacia* vials with time. The average correlation for the vials crosses the 98% level initially about 8 hr after injection of 10–100 cfu. The correlation does not continue to fall with time but rises and begins to resemble that of the time = 0 spectra after 20 hr. However, after 34 hr the spectral correlation drops across the 98% confidence level again. This "double-crossing" of the confidence limit was also observed in *E. coli* vials (see Fig. 8). The correlation for *E. coli* crossed the 98% confidence level initially at approximately 12 hr after injection of inoculum. The vials resembled the time = 0 spectra for a short period from 22 to 26 hr after the injection of bacteria. Beyond 26 hr after injection, the vials apparently dropped below the 98% correlation confidence level continuously.

In a related study of *Candida albicans* in PVC admin-

Table I. Concentrations (cfu/ml)[a] of Bacteria Grown at 24°C in vials[b] Containing TSB Growth Medium[c]

	Pseudomonas aeruginosa			*Escherichia coli*			*Staphylococcus aureus*			*Pseudomonas cepacia*		
Time (hr)	Vial 7	Vial 8	Mean	Vial 7	Vial 8	Mean	Vial 7	Vial 8	Mean	Vial 7	Vial 8	Mean
Initial inoculum[d]	120	120	—	444	444	—	12	12	—	12	12	—
0[e]	1.0	0.75	0.88	3.8	3.5	3.7	0	0	0	0	0	0
6	52	53	53	29	31	30	2.0	3.0	2.5	6.0	8.0	7.0
12	—	—	—	—	—	—	10×10^2	11×10^2	11×10^2	19×10^3	20×10^3	20×10^3
18	69×10^3	73×10^3	71×10^3	52×10^4	56×10^4	54×10^4	—	—	—	—	—	—
24	60×10^6	85×10^6	73×10^6	17×10^7	16×10^7	17×10^7	18×10^3	21×10^3	20×10^3	10×10^4	12×10^4	11×10^4
48	12×10^8	18×10^8	15×10^8	20×10^9	22×10^9	21×10^9	53×10^4	62×10^4	58×10^4	95×10^5	10×10^6	98×10^5

[a] Concentrations listed are an average determined from the counts of two spread plates.
[b] The 10-ml glass vials were filled to their maximum volume, which was actually 12 ml.
[c] Vials 7 and 8 of the "control" set were measured at 0, 6, 18, 24, and 48 hr. These vials were found to be free of microbial contamination.
[d] The inoculum concentration for each microorganism was determined from the counts for four spread plates.
[e] Time zero concentrations were measured just after vials were injected with 0.1 ml of the above inoculum.

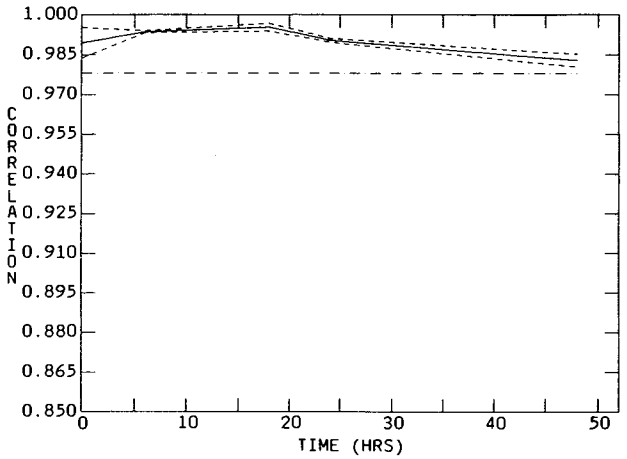

Fig. 4. The correlation between the hyperspace integrals from spectra of the control vials at time = 0 and the spectra of the same vials at time = t.

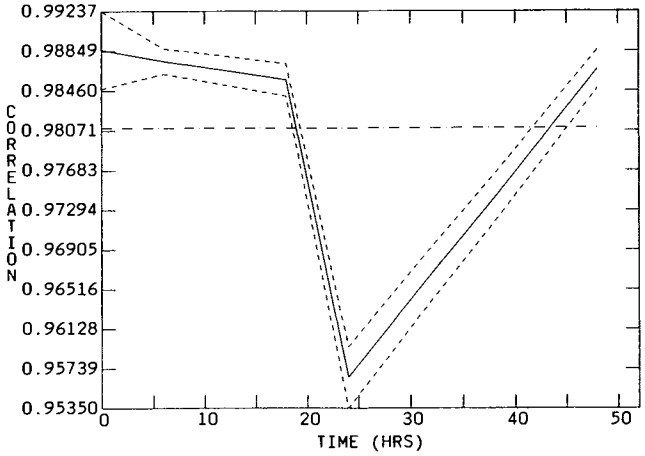

Fig. 6. The correlation between the hyperspace integrals from spectra of the *Pseudomonas aeruginosa* vials at time = 0 and the spectra of the same vials at time = t.

istration bags, the correlation of spectra to time 0 values was monitored continuously. The first time the correlation was observed to drop below the 98% confidence level, the medium was withdrawn from the bags through a 0.2-μm filter and deposited into new bags. Fresh medium was then flushed back into the original bags, carrying with it all of the solid material from the filter. In this way, the solid material (cells, etc.) from the original bags was separated from the solution. When the spectra from the bags containing the solid material were compared to the time = 0 bag spectra, the bags containing the cells and solid material were identical to the time = 0 bags at the 98% level. The bags containing the solution from the original bags, however, were different from the time = 0 bag spectra at the 98% level, indicating that the earliest changes observed in bag spectra are due to changes in the composition of the solution in the bags, and not to an increase in scattered light from cells in solution.

The double crossing of the average correlations across the 98% confidence limit may therefore represent a change in the physical mechanism at the source of the spectral changes observed by the near-IR macroscopic detection technique. The first crossing may occur when changes in the composition of the solutions alter the transmission spectra of the bags. As the number of cells in the bags increases, the number of scattering events per unit volume also increases, shortening the average pathlength and reducing the effect that solution changes have on the observed spectra. Hence, the spectra recorded at later times may begin to resemble the spectra recorded at time = 0. Finally, as cells continue to proliferate, the spectral contribution of light scattered by cells may become so significant in the observed spectra that a drop in correlation between time = 0 spectra and later spectra begins again.

CONCLUSIONS

The near-IR method of vial inspection appears to be a viable method in detecting changes in the composition of sealed glass vials for parenteral use. The use of simultaneous

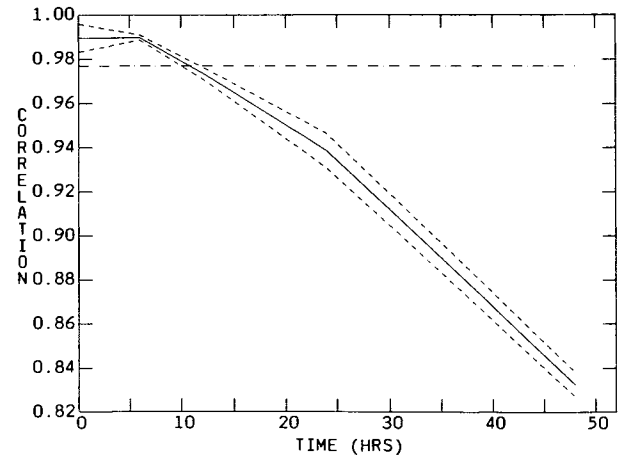

Fig. 5. The correlation between the hyperspace integrals from spectra of the *Staphylococcus aureus* vials at time = 0 and the spectra of the same vials at time = t.

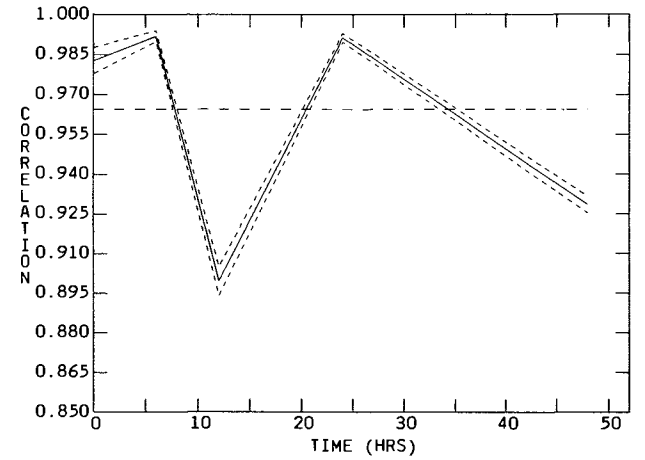

Fig. 7. The correlation between the hyperspace integrals from spectra of the *Pseudomonas cepacia* vials at time = 0 and the spectra of the same vials at time = t.

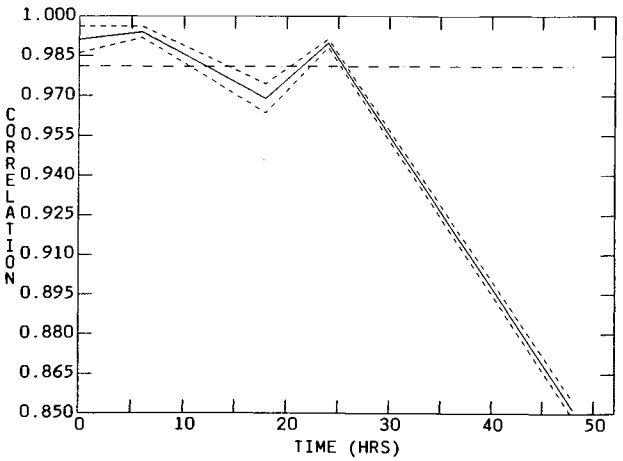

Fig. 8. The correlation between the hyperspace integrals from spectra of the *Escherichia coli* vials at time = 0 and the spectra of the same vials at time = t.

transmission and reflection optics [similar to the optical system employed in the near-IR inspection of IV bags (1)] seems to offer earlier and more precise detection of microorganisms. The addition of transmission optics to the scattering optics depicted in Fig. 1 will soon enable simultaneous recording of transmission and reflectance spectra, rather than reflectance spectra alone, and thus will help to clarify the physical mechanism by which bacterial growth is detected.

ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation through Grant RII-8610671, the Commonwealth of Kentucky through the University of Kentucky's Center for Computational Sciences, Glaxo, Inc., and BRSG S07 RR05857-09, awarded by the Biomedical Research Support Grant Program, Division of Research Resources, National Institutes of Health.

NOMENCLATURE

Special Defined Operations

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

$\kappa(Z)$ 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

n_h The number of replicate spectral points falling inside a hypercylinder

p Proportion of a distance distribution to trim from each end of the distribution

Matrices, Vectors, and Arrays

$\mathbf{B} = (b_{ij})_{m,d}$ $m \times d$ 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 B

$\mathbf{P} = (p_{ij})_{m,n}$ Training-set sample numbers selected for the bootstrap-sample sets used to calculate the 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 [B, $\mathbf{B}_{(X)}$, or $\mathbf{B}_{(V)}$]

$\mathbf{S}_{(T)} = [s_{(T)i}]_m$ Euclidean distances of training-set replicates from 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 C

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

REFERENCES

1. L. J. Galante, M. A. Brinkley, J. K. Drennen, and R. A. Lodder. Near-infrared spectrometry of microorganisms in liquid pharmaceuticals. *Anal. Chem.* **62**:2514–2521 (1990).
2. R. A. Lodder and G. M. Hieftje. Quantile BEAST attacks the false-sample problem in near-infrared reflectance analysis. *Appl. Spectrosc.* **42**:1351–1365 (1988).
3. R. A. Lodder and G. M. Hieftje. Detection of subpopulations in near-infrared reflectance analysis. *Appl. Spectrosc.* **42**:1500–1512 (1988).
4. R. A. Lodder and G. M. Hieftje. Quantile analysis: A method of characterizing data distributions. *Appl. Spectrosc.* **42**:1512–1520 (1988).
5. United States Pharmacopoeia (USP XXII) and National Formulary (NF XVII), United States Pharmacopoeial Convention, 1989, Section ⟨51⟩.