Published 2001 Wiley-Liss, Inc.† Cytometry 44:295–308 (2001)

Statistical Evaluation of Confocal Microscopy Images

Robert M. Zucker* **and Owen T. Price**

Reproductive Toxicology Division, National Health and Environmental Effects Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina

Received 13 December 2000; Revision Received 27 March 2001; Accepted 20 April 2001

Background: The coefficient of variation (CV) is defined as the standard deviation (σ) of the fluorescent intensity of a population of beads or pixels expressed as a proportion or percentage of the mean (μ) intensity ($CV = \sigma/\mu$). The field of flow cytometry has used the CV of a population of bead intensities to determine if the flow cytometer is aligned correctly and performing properly. In a similar manner, the analysis of CV has been applied to the confocal laser scanning microscope (CLSM) to determine machine performance and sensitivity.

Methods: Instead of measuring 10,000 beads using a flow cytometer and determining the CV of this distribution of intensities, thousands of pixels are measured from within one homogeneous Spherotech 10-μm bead. Similar to a typical flow cytometry population that consists of 10,000 beads, a CLSM scanned image consists of a distribution of pixel intensities representing a population of approximately 100,000 pixels. In order to perform this test properly, it is important to have a population of homogeneous particles. A biological particle usually has heterogeneous pixel intensities that correspond to the details in the biological image and thus shows more variability as a test particle.

Results: The bead CV consisting of a population of pixel intensities is dependent on a number of machine variables that include frame averaging, photomultiplier tube (PMT) voltage, PMT noise, and laser power. The relationship among these variables suggests that the machine should be operated with lower PMT values in order to generate superior image quality. If this cannot be achieved, frame averaging will be necessary to reduce the CV and improve image quality. There is more image noise at higher PMT settings, making it is necessary to average more frames to reduce the CV values and improve image quality. The sensitivity of a system is related to system noise, laser light efficiency, and proper system alignment. It is possible to compare different systems for system performance and sensitivity if the laser power is maintained at a constant value. Using this bead CV test, 1 mW of 488 nm laser light measured on the scan head yielded a CV value of 4% with a Leica TCS-SP1 (75-mW argon-krypton laser) and a CV value of 1.3% with a Zeiss 510 (25-mW argon laser). A biological particle shows the same relationship between laser power, averaging, PMT voltage, and CV as do the beads. However, because the biological particle has heterogeneous pixel intensities, there is more particle variability, which does not make as useful as a test particle.

Conclusions: This CV analysis of a 10-μm Spherotech fluorescent bead can help determine the sensitivity in a confocal microscope and the system performance. The relationship among the factors that influence image quality is explained from a statistical endpoint. The data obtained from this test provides a systematic method of reducing noise and increasing image clarity. Many components of a CLSM, including laser power, laser stability, PMT functionality, and alignment, influence the CV and determine if the equipment is performing properly. Preliminary results have shown that the bead CV can be used to compare different confocal microscopy systems with regard to performance and sensitivity. The test appears to be analogous to CV tests made on the flow cytometer to assess instrument performance and sensitivity. Cytometry 44:295–308, 2001. Published 2001 Wiley-Liss, Inc.†

Key terms: confocal microscopy; coefficient of variation; fluorescence; beads; microscope calibration; image statistics

The confocal laser scanning microscope (CLSM) consists of a standard high-end microscope with fine objectives, lasers to excite the sample, fiber optics to deliver the laser light to the stage, acoustical transmission optical filters (AOTF) to regulate the laser light to the stage, filters, dichroics and pinholes to control the light, electronic scanning devices (galvanometers), detection devices to measure photons (i.e., photomultipliers [PMT]), and other electronic components. To operate efficiently and yield high-resolution images, the system must be aligned properly and all components must function correctly. A number of instrument performance tests have been devised to assess laser power, laser stability, field illumination, spectral registration, lateral resolution, axial Z resolution, lens

†This article is a US government work and, as such, is in the public domain in the United States of America.

The research described in this article has been reviewed and approved for publication as an EPA document. Approval does not necessarily signify that the contents reflect the views and policies of the EPA, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

*Correspondence to: Robert M. Zucker, U.S. Environmental Protection Agency, Reproductive Toxicology Division (MD-72), National Health and Environmental Effects Research Laboratory, Research Triangle Park, NC 27711.

E-mail: zucker.robert@epa.gov

cleanliness, lens functionality, and Z-drive reproducibility (1–8). This list is not inclusive and there are other factors to consider to ensure proper function of the instrument (5–7). Unfortunately, manufacturers of the confocal microscopes have not released sufficient specifications on the machines to guarantee proper functioning. Because of this, it is necessary to do a subjective assessment using only a biological reference slide. In our opinion, this is too arbitrary a test when intensity measurements are needed using this optical equipment.

In quantitative fluorescence microscopy, a fluorescence optical microscope is used to acquire fluorescence intensity values emitted from a defined area of the specimen (8). It is generally assumed that the intensity of fluorescence is proportional to the amount of fluorescence present. However, because the fluorescent image is usually weak when compared with other types of microscopy images, it is essential that the system operate at maximum optical efficiency. The sensitivity of the CLSM depends on the brilliance of the light source, the efficiency of the optical system, and the performance of the detection electronics. Therefore, it would be extremely useful if this sensitivity could be maximized (8).

To produce a confocal image, a pinhole is required which decreases the number of photons reaching the detectors. This makes the optical detection system less efficient and less sensitive (7). To compensate for the production of confocality, the PMT voltage is raised to visualize the image, which introduces more PMT noise into the image. To reduce the image noise, frame averaging is used, which may increase specimen bleaching. To create an accurate confocal image, specimen bleaching, system sensitivity, and confocality have to be balanced.

Bleaching is minimized when the instrument produces an image with the least amount of light hitting the specimen. If the specimen fades during the acquisition process, errors in the representation of intensity in the acquired image may occur.

To quantify fluorescence, possible errors in instrument functionality, sample preparation, and mathematical treatment of the three-dimensional (3D) data have to be considered. Sample preparation techniques and instrument stability during operation must be evaluated and kept constant for reliable fluorescent measurements to be made. Specimen factors influencing intensity measurements include the rate of bleaching, the environment of the sample, incorporation of the dye, concentration of dye, mounting media, autofluorescence, energy transfer, and wavelength of excitation and emission. Possible instrument errors include the instability of a light source, in homogeneity of illumination, background fluorescence, light leak from stray room light, instability of photometer detection, and nonlinearity of photometer detection (8–13). These possible instrument errors are applicable to fluorescence optical equipment with cameras and photometers and to confocal microscopes (8–13). Because the CLSM images are digital and made with sophisticated optical equipment, many types of tests need to be made for adequate quality assurance (QA) of this instrument. These tests have the ability to determine if the machine is performing and to test some components in the system for proper functioning. Because the ultimate aim of many of our studies was to acquire data for the quantification of fluorescent probes, the confocal QA tests will help to ensure that the data obtained are accurate. After the data are obtained with a stable machine having known QA parameters, different analysis methods can be applied that adjust for light attenuation with depth, field illumination irregularities, and measurement of objects in 3D space (5,14–16).

The CLSM system is usually evaluated by subjective analysis of biological samples (1–8). Unlike flow cytometry, there is no universal standard with which to evaluate the CLSM or the image quality. Investigators have used beads, spores, pollens, diatoms, fluorescent plastic slides, fluorescence dye slides, silicone chips, and histological slides from plants or animals (1–8). This is by no means a comprehensive list. In most cases, the test sample is of biological origin, which is recommended by the manufacturers of most CLSMs. It would be advantageous to have better methods to measure system performance and evaluate image quality. One aim of this research was to apply similar statistical procedures, used for many years in flow cytometry, as a standard to evaluate CLSM images and system performance (17–22).

One method to assess flow cytometry system performance is to use a population of uniform fluorescent beads and measure the fluorescence and light scatter of approximately 10,000 beads. The measurement of 10,000 beads yields a distribution of fluorescent intensities and sizes, which correlates to the particle variation and system performance factors (17–22). The coefficient of variation (CV) can be measured from this distribution. We applied the CV of a population of beads to evaluate confocal system noise, image quality, and system performance. Instead of using thousands of beads to produce a fluorescent histogram, this novel technique uses thousands of pixels from a single bead to generate a population distribution. The population means and standard deviation, and thus the CV, can be determined from the pixel intensity values. Given the impracticality of imaging tens of thousands of beads to get a distribution of fluorescence intensities or particle sizes with a CLSM, we analyzed a large bead consisting of many pixels. The intensity deviation of these pixels represents the noise in the scanned image of the bead. If the beads are uniform in intensity and size, they can represent a standard for the evaluation of image quality and the performance delivered by a specific manufacturer's system. Generally, it is assumed that the smaller CV represents a system that is aligned properly, is stable, and yields good resolution and system performance. This study was undertaken to evaluate CLSM image quality and system performance with the hope that the subjective methods being used to assess CLSM image

quality and machine performance will be eliminated and replaced with more objective procedures.

A number of other research reports and books have described other tests that are used to evaluate microscope performance (9–13). This manuscript deals with a new test based on the CV concept, which was devised to measure primarily the sensitivity of a confocal microscope. The sensitivity of any fluorescence optical system depends on the intensity of the light source, the efficiency of the optical system, and the quality of the detection system (7). For confocal microscopes specifically, the sensitivity comprises variables that include PMT noise, laser noise, alignment, and system efficiency. It would be extremely useful if there was a test that could assess sensitivity in this optical equipment. We believe that we have developed a fluorescent bead test that can be used to measure sensitivity over time, so an assessment can be made on how the machine is performing over time. The test can also be used to compare the sensitivity of two machines from one manufacturer or compare the machines from different manufactures with regard to sensitivity and performance. We have shown that the Bead CV test confirms principles of noise reduction by averaging sequential frames. The noise is reduced inversely as the square root of the number of frames averaged (12,23). We hope that this test will be used in conjunction with other tests to help replace the subjectivity in measurements in evaluating confocal microscopes for system performance.

MATERIALS AND METHODS

Beads

The beads were obtained from Spherotech (Libertyville, IL). They included the 10-μm Rainbow fluorescent particles (FPS-10057) and the 6.2-μm Rainbow three different intensity beads (FPS-6057-3). The polystyrene 10-μm beads (refractive index [RI] =1.59) were mounted with optical cement (RI = 1.56) on a slide using a 1.5- size coverglass. The Leica immersion oil has an RI of 1.51.

Biological Test Slides

FluoCells (F-14780, Molecular Probes, Eugene, OR) were stained with three fluorochomes (Mitotracker Red CMXROS, BODIPY FL Phallacidin, and DAPI) and were used as biological test slides. Additional slides were made in our laboratory with cells growing on coverslips, fixed with paraformaldehyde, and stained with DAPI and other fluorochromes.

Confocal Microscope

The Leica TCS-SP1 and Leica TCS4D (Heidelberg, Germany) confocal microscope systems contained an argon-krypton laser (Melles Griot, Omnichrome) emitting 488, 568, and 647 nm lines and a Coherent Enterprise (Auburn, CA) laser emitting 365 nm lines. The TCS-SP1 had an AOTF to regulate the light bands. The results should be applicable to all point scanning systems featuring different laser configurations. Some of the Leica statistical software features may not be present in the other confocal machines, which may affect the ease of analysis. For comparison, a Zeiss 510 system was used. It contained an argon laser (25 mW) and two helium-neon (HeNe) lasers (543 nm, 1 mW; 633 nm, 5 mW) with an AOTF and a merge module.

Power Meter

The power meter used to measure light on the microscope stage was a Laser mate/Q (Coherent) with a visible detector (LN36). A power meter (1830C) from Newport Corporation with an SL 818 visible wand detector can also be used for power measurements. On most confocal systems, there is a 10× lens: Zeiss uses a 10× Plan Neofluar (numerical aperture [NA] 0.3) and a Leica has a 10× Plan Fluortar (NA 0.3). The test was made using a 10× (NA 0.3) objective. The lens is raised to its maximum specified height. The detector is secured on the stage and centered grossly using either laser light or mercury fluorescent light. The detector position is then adjusted more accurately to achieve maximum signal intensity by using the microscope's x/y joystick. The CLSM zoom factor is set from 8 to 32 to reduce the beam scan and to focus it into the "sweet spot" of the detector. The scanner is set at bidirectional slow speed to reduce the time period that the power meter is reading "0." The power derived from this measurement is dependent on the magnification and NA of the lens used. Each lens will have a unique set of values, which is dependent on the objective's NA and other transmission factors. The power meter diode in the scan head was not reliable and could only be used as a crude estimate of the functioning of the laser.

Software Analysis

The Leica software was used to evaluate most of the images. Sometimes, the TIFF images acquired with the TCS-SP1 software were imported into Image Pro Plus (Media Cybernetics, Silver Springs, MD) for subsequent measurement and analysis.

Definitions

The CV is defined as the standard deviation (SD; σ) of the population of beads or pixels expressed as a proportion or percentage of the mean (μ). In this study, the CV is preferred over SD as a measure of variability. SD is often correlated positively with the population's mean. Also, the CV is independent of the unit of measurement, unlike the SD. This makes the CV a measure of the relative magnitude of variation, whereas SD is an absolute measure of variation.

RESULTS

In our flow cytometry core facility, a homogeneous population of fluorescent beads with a small CV (Molecular Probes 2.5-μm alignment beads [A-7302, EX 488 EM 515-660]) is used for alignment, system performance, and reliability of a BD FACSCalibur flow cytometer. The test is made by adjusting the mode of the bead histogram into

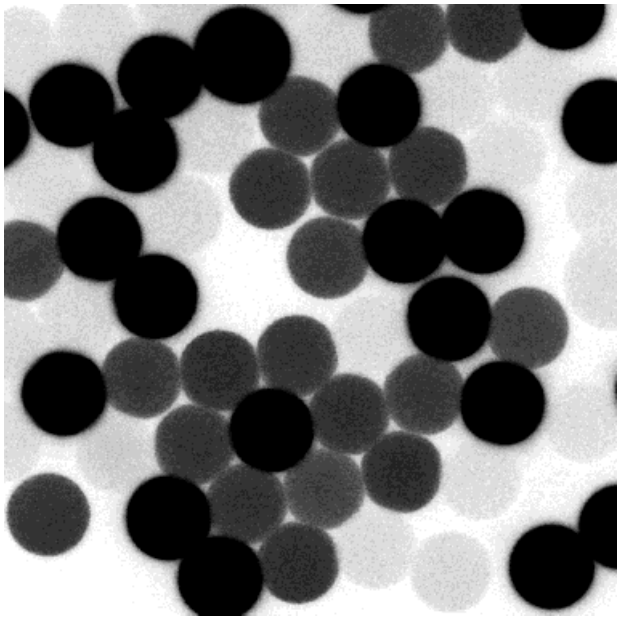

FIG. 1. TIFF image (512 × 512) of 6.2-μm Spherotech three intensity beads acquired with a 100× Plan APO objective (NA 1.4). The three different intensity levels of pixels of the 6.2-μm beads were observed within the linear 256 gray scale levels. The GSVs in the image have been inverted for publication clarity.

channel 400 and measuring the resulting CV of the particle distribution. This was usually found to be between 1% and 2.5% for the two scatter and three fluorescent signals. Variations from the expected CV and designated channel parameters at specific PMT settings indicate that the system is not performing properly and is probably blocked. This results in altered fluidic flow and a wider CV. In order to check the electronic linearity and sensitivity of the flow cytometer, many laboratories measure routinely a population beads of varying intensity, which allows for the determination of system linearity, noise, and performance (17,18). We have tried to adapt these two tests (alignment and linearity/sensitivity), which are used routinely for instrument standardization and calibration in flow cytometry, to perform a QA a confocal microscope.

In the first test, a population of uniformly sized beads (6.2 μm) of varying intensity was imbedded in a single focal plane using optical cement (Spherotech, FPS 6057-3). Many beads of the three different intensity levels were contained in any microscope field using a 100× Plan Apo objective (NA 1.4; Fig. 1). In order to perform this multi-intensity bead test using a CLSM, it is preferable to have uniform field illumination and all beads in the same focal plane. The image is acquired at the bead's maximum diameter (center of the bead). If the beads in the region of interest (ROI) are not in the same focal plane, then a stack of images has to be obtained followed by a maximum projection of the image to correct for beads residing in different focal planes. This is a time-consuming but necessary process because beads that reside outside of the focal plane have been observed to exhibit different intensity levels than beads observed at their maximum diameter. An examination of the bead image confirmed that the population of multi-intensity beads consisted of pixels displaying three major levels of intensity. The three different pixel intensity levels of the 6.2-μm beads could be observed within the linear 256 gray scale levels of the 512 × 512 image produced by the CLSM (Fig. 1). This population of intensity beads should yield a distribution with a histogram displaying three peaks of distinct pixel intensities (Fig. 2). The relationship of these peaks to each other and their individual CVs should yield information on the functionality, performance, and reliability of the machine. Under optimal conditions of machine operation, the CVs of the bead populations could be decreased such that the mean bead intensities of the three subpopulations do not change but the CVs of the subpopulations are reduced. The optimal conditions that delivered the most significant reduction in the CV were increased frame averaging, reduced PMT voltages, and increased laser power (Fig. 2). If the beads are relatively homogeneous with respect to their intensity, the CVs of each population will be a measure of particle variability, variations in field illumination, and machine variability (i.e., laser power, stability, and PMT). The proposed measurement of CV among a population of beads is similar to that conducted in similar tests using a flow cytometer.

A population of single-intensity beads could provide more accurate information than multi-intensity beads. After examining images of a population of multi-intensity beads, two distinct sources of variation were confirmed. The first source of variation is the difference in intensity among the three subpopulations of beads. The second source of variation is the difference in intensity within a single bead. The pixels residing within a bead can vary in intensity due to the limits of resolution, variations between the different parts of the bead, and the inherent noise in the system. Limiting the evaluation to a population of single-intensity beads removes the first source of variability. By defining an ROI in the center of the bead, the variation due to imperfect resolution (pixel intensity decreases near the edge of the bead) may be minimized. This second bead test yields a population of approximately 100,000 pixel intensity values that can be analogous to a flow cytometry population of thousands of fluorescent bead intensity values. Both CLSM and flow data yield a histogram population from which a mean intensity (μ), SD (σ), and CV (σ/μ) may be obtained.

A series of homogeneous beads exhibiting uniform intensities at three different sizes (5, 10, and 15 μm) was obtained from Spherotech and tested for applicability to a single-bead test sample. The beads were analyzed with a 100× Plan Apo objective (NA 1.4). The 5-μm beads were too small and the pixels at the edge of the bead effected greatly the distribution. The 15-μm beads were too large. When using an Airy disk of 1, there was a dark region in the middle of the bead, which indicates that the system confocality eliminated these fluorescence pixels. The

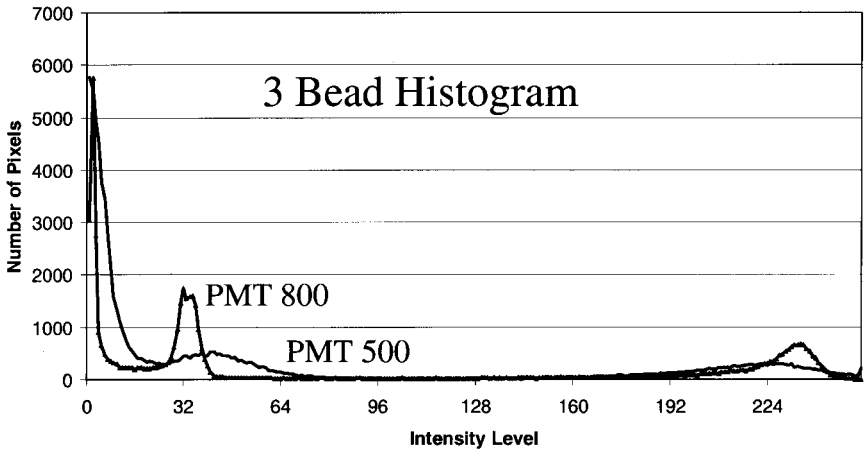

FIG. 2. Two images of three intensity beads were acquired with PMT voltages of 500 and 800. The lower PMT voltage (PMT = 500) was obtained by setting the laser to a near maximum value and having the AOTF at maximum values. The higher PMT voltage (PMT = 800) was obtained by decreasing the amount of laser light with the AOTF adjustment. The histogram of pixel intensities displays peaks with smaller CVs at the lower PMT setting.

10-μm beads appeared to be the correct size using an Airy disk of 1. The image that was captured contained relatively homogeneous pixels throughout the bead area using two different microscope systems (Zeiss 510 and Leica TCS-SP1). The area inside the bead was of sufficient size to allow a uniform ROI to be defined within it. It was helpful to zoom the bead four times to increase the quantity of pixels contained within the ROI. Repeated sampling of the same bead resulted in minimal bleaching and the CV did not change significantly during subsequent scans. The measured CV of fluorescence intensity of this 10-μm bead population on a BD FACSCalibur flow cytometer was 5%.

Figure 3 illustrates the pixel distribution of a 10-μm bead that was measured with a PMT voltage setting of 400 and 600 and a zoom of 4×. These two bead images were obtained in the following manner. The mean intensity value in the ROI within the bead was set at channel 150 by adjusting the AOTF manipulation instead of actually lowering/raising the laser power. The higher PMT voltage yielded a broader histogram, which translated into more pixel intensity variations. Because the CV (σ)/(μ) is defined as the SD (σ) divided by the mean (μ), the quality of the images can be compared using this technique. As the quality (less noise) of the images increases, the CV of the population of pixel intensities within the bead decreases. In order to compare images between machines, it is critical that as many variables as possible be kept constant (8).

Images of beads were acquired at various PMT settings (400, 499, 601, 700) and the pixel distribution was determined by measuring the identical ROI inside each bead's image. As the PMT voltage increased, the range of the pixel intensities and CV also increased (Fig. 4). It is best to

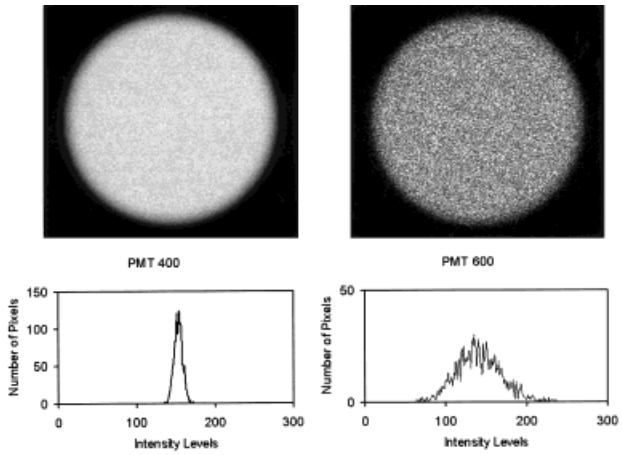

FIG. 3. TIFF images of a 10-μm Spherotech bead were obtained with two PMT settings (PMT = 400, PMT = 600) with a zoom of 4 and no frame averaging using a 100× Plan Apo lens (NA 1.4). An ROI was drawn in the interior of the bead and the histogram of the population of pixel intensities is displayed in the bottom panels. The mean pixel intensity in both images was approximately 150 intensity levels and was obtained by keeping the PMT at 400 or 600 and adjusting the laser power with the AOTF.

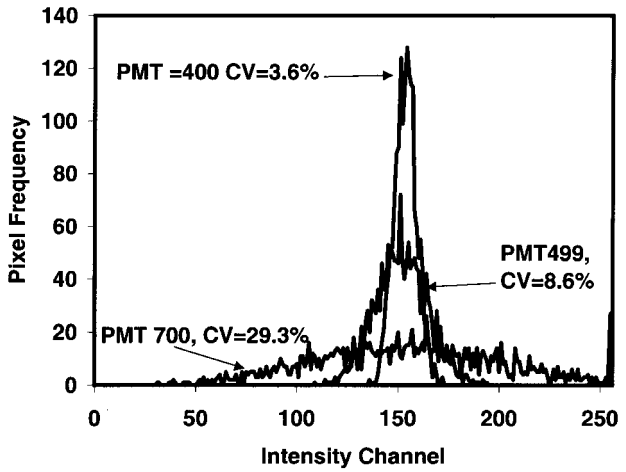

FIG. 4. Noise and PMT voltage. The PMT voltage was increased by adjusting the AOTF to ensure that the bead exhibited the same intensity (mean GSV = 150) level in all images. The CV increased from 3.6% to 29 % as the PMT voltage increased from 400 to 700. Only three PMT settings of 400 (CV = 3.6%), 499 (CV = 8.6%), and 700 (CV = 29.3%) are represented.

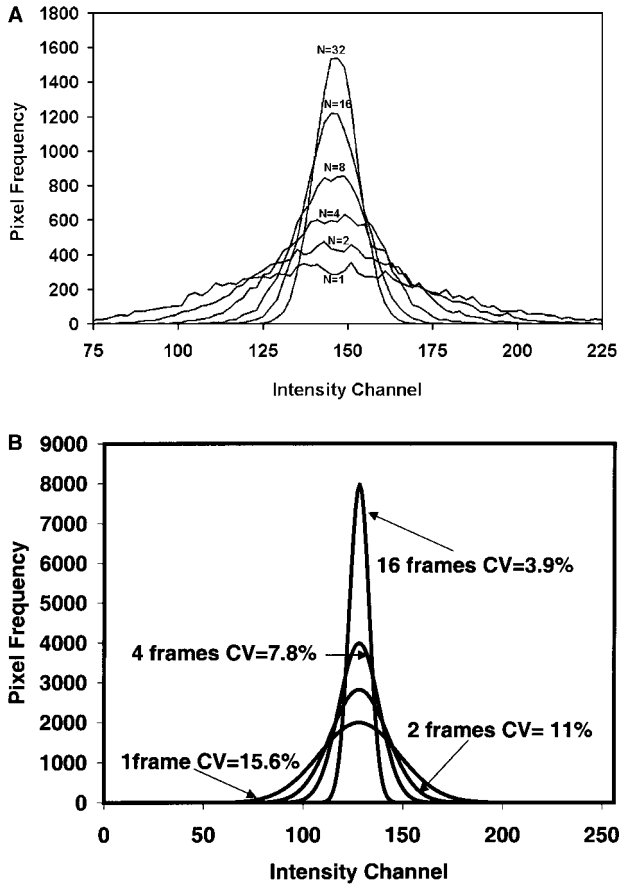

FIG. 5. **A:** Effects of frame averaging. A 10-μm bead was frame averaged 2, 4, 8, 16, and 32 times to obtain distributions of pixel intensities. Bleaching was minimal during the experiment. The mean intensity channel was kept constant at channel 146. An increase in averaging from 1 to 32 decreased the CV in the following manner: 21.96% (1), 15.53% (2), 11.13% (4), 7.99% (8), 5.84% (16), 4.32% (32). The CV in the distribution decreased by the square root of "n" times the bead was averaged. **B:** Theoretical effects of frame averaging. A hypothetical Gaussian distribution was chosen, simulating an actual distribution in Figure 5A. As the averaging increases from 1 frame to 16 frames, the CV of the histogram decreases from 15.6% to 3.9%. The improvement in image quality is proportional to the square root of the number of frames averaged (13). The averaging decreases the pixel variation, which lowers the SD and decreases the CV. Theoretically, by averaging the image n times, the CV and SD are decreased by the square root of n. If the mean is assumed to be constant (channel 128), the histogram distributions generated by averaging can be produced and the CV and SD calculated.

operate the confocal with conditions that yield a minimal CV, which will translate into good image quality.

Video microscopy studies have shown that the noise is reduced inversely as the square root of the number of frames averaged is increased (9,12,23). A major gain in noise reduction is obtained after averaging only a few frames. By continuously averaging additional frames, the signal-to-noise value is only affected slightly. As reported in video microscopy studies, we found that the quality of the image is increased by averaging confocal TIFF images together. A distribution of acquired data (Fig. 5A) and a theoretical distribution of this relationship (Fig. 5B) illustrate the effects of frame averaging on CV. The improvement in quality is proportional to the square root of the number of frames averaged in video microscopy and confocal microscopy images (12,23).

To obtain the theoretical distribution, a bell-shaped Gaussian distribution of intensity values was made (23). The mean was assumed to be constant. By averaging the distribution 2, 4, 8, 16, and 32 times, the shape of the distribution would change as the peak became successively higher and the width successively smaller. Theoretically, this averaging of the image "n" times decreases the noise by the square root of "n." This is also equivalent to decreasing the CV and SD of the distribution by the square root of n. The theoretical curves can thus be generated as both the CV and SD are decreased by the square root of n. As shown in Figure 5B, the number of averaged frames increased from 1 to 16 frames and the CV decreased from 15.6% to 3.9%.

Because both PMT voltages (Figs. 3, 4) and frame averaging (Fig. 5) influenced the CV value, an experiment was designed to test this relationship. The PMT voltage was decreased gradually from 1,000 to 450. For each PMT setting, the frame averaging was increased from 1 to 32 (Fig. 6). The mean intensity (channel 150) in an ROI for each setting of PMT voltage and frame averaging was measured. The laser power was kept constant and the power was adjusted with the AOTF. The CV was determined by recording the mean and the SD of the pixel intensities using the Leica statistical program built into its analysis package. Either increased frame averaging or lower PMT voltages could decrease the CV value. At

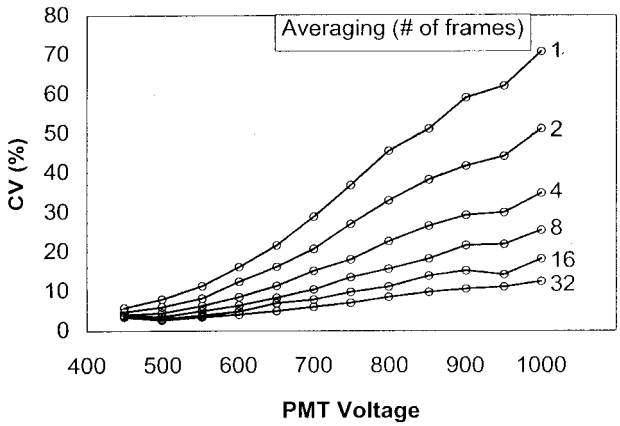

FIG. 6. Effects of averaging and PMT on CV. The noise present in the system was evaluated using a 10-μm bead with a 100× Plan Apo objective (NA 1.4). The excitation laser wavelength was 488 nm and emission was a 50-nm band pass filter (505–555 nm). The test was made by decreasing the PMT value and adjusting the laser power with the AOTF to ensure that the mean pixel intensity was at a value of 150. The higher PMT values were taken to minimize possible bleaching. Images corresponding to 1, 2, 4, 8, 16, and 32 were obtained at each PMT setting. The noise at a specific setting can be reduced if frame averaging is increased. The CV is defined as the SD (σ) of the fluorescent intensity of a population of beads or pixels expressed as a proportion or percentage of the mean (μ) intensity. ($CV = \sigma/\mu$).

Table 1
*PMT Comparison and Noise**

Excitation	Emission	PMT	PMT voltage	CV (%)	Relative CV
488 nm	505–555 nm	1	474	6.06	100
		2	428	6.58	108.65
		3	425	6.23	102.86
488 nm	555–600 nm	1	471	6.02	100
		2	432	7.00	116.25
		3	421	6.46	107.17
568 nm	580–630 nm	1	439	4.00	100
		2	411	4.88	122
		3	393	4.49	112.11
647 nm	665–765 nm	1	802	20.30	100
		2	732	22.70	111.68
		3	675	20.30	100.12

*The noise of the system was evaluated using a 10-μm bead (Spherotech) and a 100× Plan Apo (NA 1.4) objective. The intensity of a 10-μm bead was determined at a constant laser power, a zoom of 4, and no averaging using various PMT settings. The emitted light was measured in each of the three PMTS. The pixels in each ROI were set to a mean of approximately 150 and the SD of pixel distribution was measured to determine the CV. The CV of the pixel intensity within the bead was measured at each PMT setting. PMT 1 is low noise blue sensitive whereas PMT 2,3 are far-red sensitive. The quality and the performance of each PMT can be measured with this test.

higher PMT settings, it is necessary to frame average to reduce the CV (Fig. 6). However, lower CVs can be achieved by using an efficient, low- noise PMT that is operated at low voltages.

This bead CV test also illustrates a method to access the operation and quality of the PMTs in the system. The use of the Leica SP system easily allowed for pairing different PMTs with different excitation wavelengths. In effect, any PMT could be used in conjunction with any of the four excitation wavelengths. Although the PMT position will affect the CV, it is not considered to be a major contributor and in this assessment all the PMTs were considered equivalent regardless of their location in the scanhead. Two types of PMTs are used in the Leica system: PMT1 is considered low noise and PMT2 and 3 have high efficiency and sensitivity in the far-red wavelength regions. The system was set up with a triple dichroic (TD) using 488, 568, and 647 nm wavelength excitation. The three PMTs were adjusted to allow the mean pixel intensities at channel 150. The relative intensities were measured with the three PMTs for all conditions (Table 1). Due to the physical location of the three PMTs, the most efficient one was PMT1 because it has the least reflected light. However, this test showed that the least noise was derived from PMT1 under most excitation wavelengths and emission detection conditions. PMT2 is usually chosen in this system to detect emitted fluorescence derived from 568 nm excitation. However, in this test, PMT2 exhibited over 20% more noise than PMT1 when using 568 nm excitation. PMT3 was superior (smaller CV) to PMT2 in all conditions tested. Clearly, with 488 or 568 excitation using only one parameter detection, PMT1 should be used in preference to PMT2. Lower CVs will require less frame averaging to produce better image quality. The PMTs, and thus CV and image quality, will deteriorate with time. Therefore, it is important to measure the initial quality of the PMT and then to measure periodically changes in PMT performance over time. This test is useful to determine system quality and to identify a possible problem in PMT performance prior to a hard failure.

Biological Samples

FluoCells (F-14780, Molecular Probes) were excited with a 568 nm laser line and detected with a 580–630 band pass filter in PMT2. Figure 7 shows the difference in

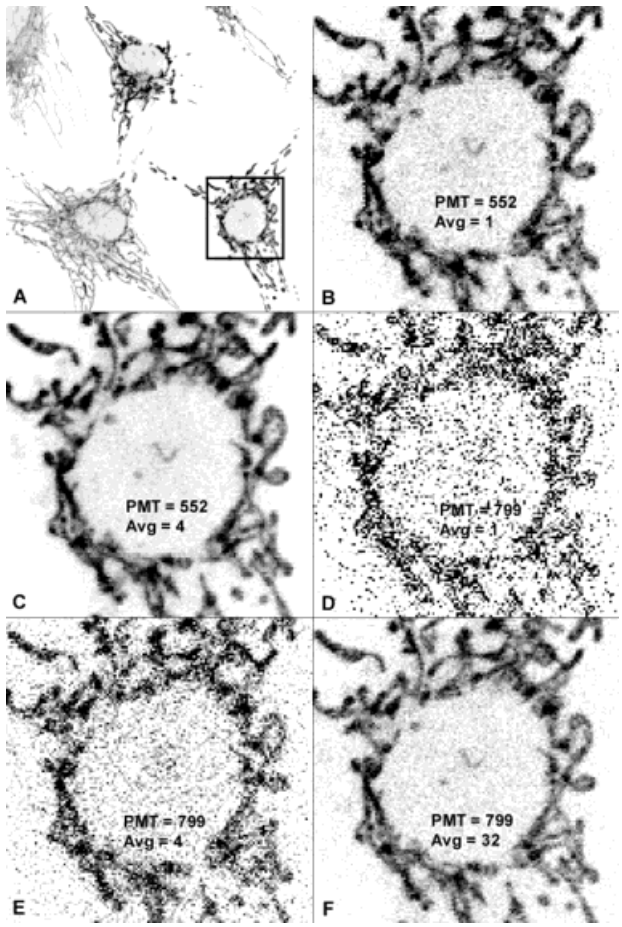

FIG. 7. PMT and averaging of FluoCells. FluoCells (F-14780, Molecular Probes) were excited with a 568 laser line and detected with a 580–630-band pass in PMT2. The resolution was measured by averaging (AV) 1, 4, or 32 times at two PMT settings (552 or 799). **A:** Distribution of three cells at normal magnification. **B–F:** One cell located in the box in Figure 7A was zoomed 4× with Image Pro Plus. The settings in the panels are as follows: A control (PMT = 552, AV = 1), B (PMT = 552, AV = 1), C (PMT = 552, AV = 4), D (PMT = 799, AV = 1), E (PMT = 799, AV = 4), and F (PMT = 799, AV = 32). Note the difference in pixel variations in the six panels of the same cell acquired at different PMT/averaging settings. The CVs of an ROI in the nucleus of the various panelsare as follows: B, 49%; C, 40%; D, 212%; E, 109%; F, 49%. This figure of a biological cell demonstrates a similar relationship between PMT values and averaging as shown for beads in Figures 4–6 for beads.

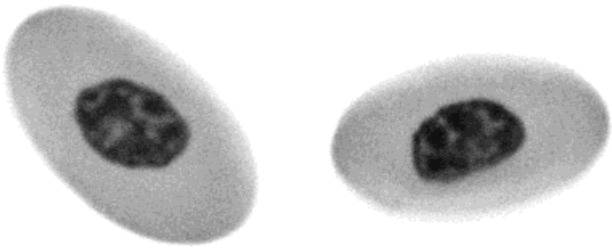

FIG. 8. CRBCs stained with AO. The images were obtained with a 100× Plan Apo lens (NA 1.4) using 488 laser light excitation and a band pass of 505–555 for emission. The PMT was set to 799 and averaging was 32 frames with a zoom of 4. Images of these cells were obtained using different averaging and PMT values and the CV values are reported in Table 2. The GSVs in the image have been inverted for publication clarity. The nucleus has the most intense fluorescence whereas the cytoplasm is less intense.

image quality when averaging 1, 4, or 32 times at two PMT settings (552 or 799). The images are zoomed 4× using Image Pro Plus to illustrate the individual pixels (Fig. 7A–F). The CVs of a selected ROI in the nucleus varied with the number of frames averaged and the PMT voltage used. The best image quality (low CV) consisted of either low PMT voltages (Figs. 7B,C) with minimum frame averaging or high PMT voltages with 32 frames averaged (Fig. 7F). High PMT settings (Figs. D,E) with minimum frame averaging (1 or 4) demonstrated high CVs and poor image quality. In all cases, the increase in averaging resulted in a decrease in the CV and a corresponding increase in image quality.

In contrast, raising the PMT voltages increased the CV and decreased image quality. The higher PMT settings necessitated the use of more frame averaging to increase image quality. Figure 7 shows that the relationship among PMT voltage, frame averaging, and CV on image quality on cells was similar to that described with beads in Figures 4–6. The noise in Figure 7 is also reduced as the square root of the frames averaged (12,23). The CV will decrease by two when samples averaged 4× and will decrease by 4 when samples averaged 16×.

Acridine orange (AO)-stained chicken red cells (CRBC) consist of two definite regions, a homogeneous cytoplasm without structural detail and a heterogeneous nucleus containing detail (Fig. 8). CRBCs are used widely as test particles to evaluate flow cytometry machine alignment and staining applications. They were the biological particles chosen to study the relationships among PMT voltage, averaging, laser power, and noise. The experiment with CRBC was performed similarly to that using the Spherotech bead (Fig. 6) and FluoCells (Fig. 7). Briefly, images were taken at two to three PMT voltages and frames were averaged between 1 and 16 times. The cytoplasm and the nucleus showed that an increase in frame averaging decreased the CV (Table 2). The CV of the heterogeneous nucleus representing a broad distribution of pixels was greater than the homogeneous cytoplasm representing a more narrow distribution of pixels. As expected, the CV showed a greater decrease when averaging was used at higher PMT settings than at lower PMT settings (12). In the two fields representing three cells, the cytoplasm and nucleus showed a similar decrease in CV values as the averaging was increased. In certain cases, the dynamic range of intensities in the CRBCs did not allow adequate readings from the cytoplasm due to its very low intensity values and the absence of Gaussian distributed pixel intensities.

The CV technique developed on beads was applied to biological specimens (FluoCells, AO-stained chicken cells) to observe if the same principles are applicable to beads and biological cells. The biological specimens exhibit details and structure that help to create a good image. However, they are not as reproducible as beads, they bleach more readily, they degenerate over time, the initial CVs are larger, and they have more variability in fluorescence staining. The details in a biological image generate good contrast but also create a larger CV, making it less effective as a test particle. The biological samples provide a subjective assessment of the CLSMs performance and they are not as effective as the10-μm bead in determining objectives and statistical values that can be used as reference points to compare data from one machine or between different machines.

Table 2
*CV of CRBC Nucleus and Cytoplasm**

Cell/field PMT	Averaging	Nucleus-CV%	Cytoplasm-CV%
Cell 1 F1			
569	1	19.2	23.8
569	4	13.6	23.8
799	1	51.2	80.1
799	4	24.7	37.2
799	16	15.7	18.1
Cell 2 F1			
569	1	20.5	24.9
569	4	15.7	24.3
799	1	42.3	69.9
799	4	21.8	35.6
799	16	14.1	17
Cell 3 F2			
529	1	14.2	31.6
529	4	12.3	24.1
594	1	12	54.5
594	4	13.2	33.7
799	1	36.7	126.3
799	4	20.1	86.3
799	8	15.6	63.7
799	32	11.6	36.5

*The pixel distribution in an ROI in three representative CRBCs illustrated in Figure 8 is described under different PMT settings, lasers power, and averaging. The laser power in the system was decreased by the AOTF. In the two fields displayed, representing three cells, the cytoplasm and nucleus showed a similar decrease in CV values as the averaging increased or the PMT decreased. The ROI in the CRBC cytoplasm or nucleus demonstrated a decrease in CV with increased averaging. This followed a relationship that was similar to that of the beads (Figs. 4–6) in which a twofold increase in averaging decreased the CV by the square root of the CV.

CRBCs and FluoCells (Figs. 7, 8) appear to generate the same relationship among PMT voltage, frame averaging, and CV expressed with beads in Figure 6. However, the bead is preferred as a test sample because it is more homogeneous with less staining variability, reduced bleaching, and greater reproducibility.

Sensitivity

The sensitivity of a confocal microscope is an important parameter to measure as the value influences PMT voltage, laser power, and frame averaging. The values also relate to alignment and performance of the CLSM. Table 3 compares a Leica TCS-SP1 containing one argon-krypton laser emitting three laser lines with a Zeiss 510 system that contains three individual lasers with a merge module. It is important that the acquisition parameters be as equivalent as possible when comparing different systems. Every effort was made to ensure that the acquisition conditions (e.g., pinhole size, scan speed, pixel size) were consistent between the two machines. The test particle was a 10-μm Spherotech bead and measurements were made using a 100× Plan Apo objective (NA 1.4) with a zoom factor of 4. The laser power in both systems was measured on the stage using a 10× (NA 0.3) objective and a power meter detector located firmly on the stage. The sensitivity of the two machines was compared by maintaining the laser power at a constant value of 1 mW for 488 light and 0.2 mW for 568/543 light. Using the Leica TCS-SP1, 1 mW of 488 power measured on the stage yielded a CV value of 4%, whereas the CV value was 1.3% with the Zeiss 510. Comparable power readings showed the CV to be almost three times lower with the 488 and 568 lines with the Zeiss system as with the Leica system. To get equivalent CVs on a Leica machine, the samples will have to be frame averaged or the laser power will have to be increased. Increasing the laser power to maximum power resulted in the CV being lowered with both the Zeiss 510 and Leica TCS-SP1 systems (Table 3). However, to reduce sample bleaching, it is important to operate the CLSM at lower laser power values and have a higher sensitivity (efficiency) in the optical system. These maximum power measurements on the stage are related to the system alignment and the functionality of different components. This CV sensitivity data may be considered an initial reference point that can be used by other investigators to assess the performance of their CLSMs. Using this approach, it is possible to compare the sensitivity of systems in different laboratories.

Table 3
*Comparison of CLSM**

Laser type	Wavelength (mW)	Power (mW)	CV-bead % (SD/M)
Fixed power comparison			
Argon-krypton	488	1	4
(75 mW, Leica)	568	0.2	4.6
Argon 25 mW (Zeiss)	488	1	1.3
HeNe 1 mW (Zeiss)	543	0.2	1.9
Maximum power comparison			
Argon-krypton	488	1.1	3.8
(75 mW, Leica)	568	1.45	2.6
Argon 25 mW (Zeiss)	488	3.2	1
HeNe 1 mW (Zeiss)	543	0.23	1.9

*A Leica TCS-SP1 containing one argon-krypton laser emitting three laser lines is compared against a Zeiss 510 containing three individual lasers and a merge module. The CVs were obtained from a 10-μm-bead using a 100× Plan Apo objective (NA 1.4). The laser power was derived by using a 10× (NA 0.3) objective and a power meter situated on the stage. By setting the power to a fixed value of 1 mW, 488 nm laser light, or 0.2 mW, 568 nm laser light on the stage, the sensitivity of two machines was compared. The CV of the bead was almost three times lower with the 488-nm and 568-nm laser lines using the Zeiss 510 system compared with the Leica TCS-SP1 system. By increasing the lasers to their maximum power, the CV values were decreased. These maximum power measurements are useful to indicate alignment of the system and functionality of different components.

DISCUSSION

This study was undertaken to evaluate CLSM machine performance by developing new tests and improving establised ones. We anticipated that the data derived from these tests would be used for QA. The data would not only be useful to compare machines from one manufacturer, but it would be able to compare machines and data from different manufacturers. We used CV data to show that various components (lasers, PMTs, excitation) in different confocal microscopes were operating at suboptimal levels, resulting in poor performance and the eventual replacement of components to correct the problems. It should be emphasized that the CV test is only one of many tests that must be used to evaluate system performance. Other important and useful tests monitor field illumination, spectral registration, axial registration, laser power, laser noise, alignment, and lens cleanliness (1–8). Unfortunately, with a confocal microscope, one test cannot be used to assess complete machine functionality. Other methods will continue to be developed to measure QA performance.

The bead CV test can be used to assess the performance and sensitivity of a confocal microscope. Three examples demonstrate the usefulness of the CV test: PMT functionality (Table 1), system sensitivity comparison (Table 3), and power efficiency/throughput (Table 3). PMTs vary in performance criteria. By using the bead CV test, the functionality of different PMTs at different wavelengths was assessed (Table 1). The sensitivity of a confocal microscopy system relates to its optical efficiency, alignment, and components within the system. Because every system is unique in its operation, the CV test provides a way to compare and contrast units from one manufacturer or between different manufacturers. The example presented in Table 3 uses a Leica TCS-SP1 system containing an argon-krypton laser emitting three wavelengths of light and a Zeiss 510 system containing three lasers with a

merge module. Both systems have an AOTF and deliver the light to the stage using fiber optics.

Different machines can be compared using the CV test if the conditions of acquisition are constant. The CV comparison test was made on different confocal machines using fixed power (1 mW of 488 nm and 0.2 mW of 568/543 nm light) on a Rainbow Spherotech bead with a 100× Plan Apo objective (NA 1.4). Scanning speed, detection pinhole size, and pixel size were kept at relatively similar values. We found that the three laser systems contained in the Zeiss 510 had over three times the sensitivity as the Leica TCS-SP1 system, which contained only one argon-krypton laser emitting three lines. Leica's newer systems (TCS-SP2) have a redesigned scan head that uses a similar laser configuration as the Zeiss 510 unit. The TCS-SP2 will probably show increased sensitivity but this has not been tested yet. By using the CV comparison, the relative sensitivity of each system at a fixed laser power was determined.

The second sensitivity test measured the maximum power on the stage that can be derived from each CLSM system and assessed the corresponding CV of a 10-μm bead. By measuring the maximum power for each wavelength in a CLSM, the optical efficiency, alignment, and functionality of different components can be assessed. The higher laser power decreased the PMT voltages, which are necessary to obtain an image of a bead with the mean being located at channel 150. For example, raising the power of the 568 nm laser line from 0.2 to 1.45mW in the Leica TCS-SP1 decreased the CV from 4.6% to 2.6%. Similarly, raising the power of the 488 nm laser line in the Zeiss 510 unit from 1 to 3.1mW decreased the CV from 1.3% to 1%. In both cases, the increased laser power increased the system's sensitivity, yielding less noise than at the lower fixed power values. However, the use of higher laser power also comes with the disadvantage of increased specimen bleaching. Because the goal of the instrument is to produce an image with the least amount of light hitting the specimen, it is important to have an efficient optical device to excite the sample and detect the emitted fluorescence (3,7,8).

If the power is fixed to a given value that the lasers in both machines can achieve (i.e.,1 mW 488 nm; 0.2 mW 568 nm), the bead CV test shows that the Zeiss 510 system is more sensitive and will deliver better resolution at less power than the Leica TCS-SP1 system. This does not mean that the Leica system (TCS-SP1) is not capable of yielding "pretty pictures" and good, high-quality data. However, it does mean that in order to achieve an equivalent level of sensitivity and resolution with the Leica TCS-SP1 unit, there will be slightly more bleaching in the sample and the PMT will be operated at higher voltages, necessitating additional averaging. This difference may be due to the amount of light being absorbed by the prism in the Leica TCS-SP1 system or to the fact that the optical configuration is not as efficient as the Zeiss 510 system because of some attenuation of light. These factors enable the Zeiss unit to produce an image with less light hitting the specimen.

Other parameters should also be taken into consideration when comparison shopping. For instance, the Leica system uses a filterless spectrophotometer, which eliminates the need for barrier filters to reject bandwidths of light and provides a very accurate way of acquiring the desired emitted light. The Zeiss system uses individual detection pinholes in front of the PMTs, which provide additional degrees of adjustment to colocalize different wavelengths of light on the PMTs. However, the advantages of extra detection pinholes and barrier filters come with the disadvantage of making the system more difficult to align. Another factor to consider is that the Leica TCS-SP units are designed to deliver better than 350 nm values in the Z-axial resolution reflecting mirror test (extremely important in determining system resolution), which is superior to the values provided by other manufacturers (1,2). We achieved a Z resolution of 185 nm on the TCS-SP1 unit using a Leica 100× Plan Apo lens. As noted previously, Leica's new TCS-SP2 system has a redesigned scan head and a three-laser configuration similar to that of the Zeiss 510. This system should deliver better sensitivity and CV values than its predecessor, the TCS-SP1 unit. Many other factors should influence the purchase decision, including QA test parameters, software, operating system, lasers, optics, filter functionality, pinhole design, ease of alignment, upgrade policy, financial stability of the company, and service issues (e.g., customer support, training, frequency of repair, service personnel competency, and repair downtime). Finally, a good long-term working relationship with the manufacturer is essential (1–8, 24).

Extreme care must be taken when extending our data to compare models of confocal microscopes among manufacturers. All factors being equivalent among microscopes, the bead test may be used in a side-by-side comparison. However, there are many CLSM variables that must be controlled when performing machine comparisons. These include scan speed, pixel dwell time, pixel size, lens quality, filters, detection pinhole size, illumination pinhole existence, scan field size, zoom, objective, field illumination, and fluorescence transmission (1–8). These factors will affect the amount of photons entering the detection PMT and thus the CV of the bead intensity population. The test will be invaluable in determining machine sensitivity and performance among CLSMs that are operated in an equivalent manner.

Another example of the usefulness of the CV data was illustrated using the UV Coherent Enterprise laser (60-mW argon laser, 365 nm). The measurement of UV light with a power meter detector can be made with objectives that range between 5× and 20×. With high NA objectives, the bead CV test system can be used not only as a sensitivity test but also as a relative power test indicator. The confocal UV system had insufficient power, which was demonstrated using a 10-μm Spherotech Rainbow bead. The laser was operated at maximum power for a short while

and the UV fluorescence from the bead was visualized at channel 150 by increasing the PMT to relatively high voltages, which resulted in a high CV value. At maximum laser power for UV (365 nm) and visible (568 nm) light, the following PMT and CV values were obtained: PMT = 679 V, CV = 19% for UV and PMT = 382 V, CV = 5% for 568 nm excitation. Explanations for the difference in CV between the two excitation wavelengths include the following: the bead may not fluoresce as much with UV light as with visible 568 nm light, the optical system may be more efficient and less attenuated using visible excitation relative to UV excitation, or there may be insufficient UV laser power but adequate 568 nm excitation power. In summary, the high bead CV value that occurs with UV excitation may be due to a lack of laser power, a bad PMT, an alignment problem yielding low laser power, or the bead does not excite well at that wavelength.

The bead CV is dependent on the excitation/emission characteristics of the bead and the amount of photons hitting the bead. However, the same noisy effects that were observed with UV-excited 10-μm Rainbow beads were also observed with a Molecular Probes test slide (FluoCells) stained with DAPI. These noisy images were obtained at high PMT settings and frame averaging was needed to make the image acceptable. A CLSM with insufficient laser power will not be efficient and the noise problems that are generated must be addressed by either frame averaging or sample preparation techniques with higher fluorochrome staining concentration. Caution must be used when increasing fluorchrome concentrations as the chemistry of the sample could be changed, resulting in possible energy transfer or quenching. Excessive averaging will also effect bleaching. The best option is to use an efficient optical system.

The confocal image is constructed sequentially from the output of a point detector. The digitized value at each pixel should reflect the average detector response during the time the beam dwells on each point in the field. The goal of the instrument is to produce an image with the least amount of light hitting the specimen. However, the general nature of the pinhole is to exclude light. This prevents the majority of photons from reaching the detector, creating a poor image and making the system less efficient. To compensate for this confocality, the PMT voltage is raised, which introduces more noise into the image. The functionality of a PMT is a critical element in obtaining good images. The quantum efficiency, spectral response, inherent noise, response time, and linearity affect PMT performance. Other factors include PMT dark current noise, amplifier electronic noise, and random variations of the digital output signals derived from identical photons. A number of reports have alluded to this PMT problem and have suggested that cooled PMTs, newer designed PMTs, and other low noise detection devices will increase performance (3,7,9,25,26). Using the bead CV method, we observed differences in the performance of the three PMTs in our system. Because placing these devices in small, tight quarters in the scan head can generate heat, temperature will be a factor in the performance as it strongly influences the PMT dark current. The design of more efficient and reliable units will be a critical factor in increasing the accuracy, sensitivity, and performance of a confocal microscope.

We have shown that the CV of a bead or cell image varies with regard to PMT settings (Figs. 3, 4, 6–8). The data support the principle that a reduction of noise is related to the number of frames averaged (Fig. 5). Noise is reduced inversely as the square root of the number of frames averaged (12,23). An inefficient machine (i.e., one that suffers from insufficient laser power, misalignment, fiber-coupling transmission, and noisy PMTs) will need to be run at high PMT values. This will yield a high CV and a poor-performing system. Obtaining bad CV data with a fluorescence bead informed us that a serious problem existed with our CLSM's UV power. This prevented the machine from providing meaningful data and measurements. Lower bead CVs suggest that there was more laser power, which would allow us to operate the PMTs at reduced values and obtain better system performance. In this case, the bead is used as a power indicator for a 100× objective, in addition to being used as a statistical CV indicator in our system. If there is a decline in power due to misalignment or improper fiber coupling, it will be represented as less power. Power measurements derived using low magnification objectives can also suggest problems with the UV laser power (Table 3) in the system. Low UV power measurements and high CV noise measurements are indications of an incorrectly aligned system that needs the attention of service personnel. This example can also be applied to visible excitation systems as insufficient power is a strong indicator of a failing laser or a misaligned system.

We have demonstrated the relationship among frame averaging, PMT voltage, laser power, and image noise using a bead and biological particles. Generally, it is best to lower the PMT voltage. If this is not possible, frame averaging will also increase image quality. PMT noise is an essential component that must be addressed to achieve satisfactory performance in a CLSM. (7,25,26). If the optical efficiency of the system is not optimum either due to misalignment or insufficient laser power, high PMT settings will be required. However, the high PMT settings will necessitate increasing the frame averaging in order to deliver a quality image (12). Other ways to increase image quality include opening the pinhole to let in more light and using band pass filters or slits that allow more fluorescent light to be detected by the PMT. Both of these solutions decrease resolution or allow irrelevant light into the image. In the future, advancements in PMT quality or the use of cooled PMTs will help to achieve better confocal performance (7,9,25,26).

The CV test is a fairly sensitive marker of detecting problems within instruments. However, there are a number of issues to consider when applying this test. The test measures total instrument performance and may not be able to determine where a specific problem occurs. For

instance, any problems in alignment, laser power, fiber, or PMT will necessitate running the machine at higher PMT values, thus yielding greater CVs. It is not possible to determine initially what factor contributed to the high CV only that there is a problem with the system resulting in bad resolution. However, once problems have been identified, the CV test starts a process to find the cause and resolve the issue. Like many tests designed for the confocal microscope, we found some limitations. All the beads on the slide need to be uniform in size and intensity or multiple measurements will suffer from sampling differences. Our beads were shown by flow cytometry to have a CV of 5% in intensity measurements. Bleaching of the sample was minor and should not affect pixel distribution. Sequential measurements should be made when bleaching is minimal (i.e., the bead should be first measured at low power and low averaging). We tested 5, 10, and 15-μm beads as reference particles. The 15-μm bead had a hole in the center of the image. This was due to taking the image at the center of the bead and Airy disk 1 (pinhole) discriminating out of focus light due to the commonality of the large bead. The 5-μm particles were too small and were zoomed 8×, which increased bleaching. The 10-μm bead appeared to be the correct size using a pinhole Airy disk 1. Smaller pinhole settings will affect the CV by raising the pinhole and possibly forming a black hole in the center (as did the 15-μm particle). The 10-μm bead should be zoomed 2–4× with a 100× objective to encompass more pixels in the image. The bead can be used with a 63× objective. However, with the lower power objective, the beads were considered too small. These beads, which are embedded in a slide and kept at room temperature, have worked successfully for 2 years in our laboratory. It is anticipated that commercial companies will make fluorescence substrate slides to replace the beads, allowing them to be used with all objectives and other types of optical equipment.

The lenses and optical configurations between machines may be slightly different, which could lead to optical aberrations (5–8,27,28). This test worked with 100× and 63× lenses using a 10-μm Spherotech bead. Larger or smaller beads were unacceptable for reasons described previously. Optical aberrations, sampling density, and bead brightness have been considered. Chromatic or spherical aberrations did not affect this test as we used one wavelength at a given time with a very large particle. The 10-μm beads are embedded in optical cement, which has an RI similar to that of the immersion oil used so optical aberrations were minimized and did not to affect the quality of pixel distributions (27). Any optical aberrations that do occur should also be minimal as they are maintained the same throughout all the tests. Sampling was done at 1×, 2×, and 4×. The same CVs occurred, indicating that oversampling or undersampling was not a factor. Normally, the sampling density of a 10-μm bead was zoomed 4×. The higher zoom allowed a larger ROI and a greater number of pixels to be measured. The ROI in the bead is large. Over 100,000 pixels are sampled from the middle of each bead, eliminating the decrease in intensity at the edges. This is a sufficient number of pixels to allow the statistics to be highly significant. However, smaller particles require zooming, which may involve the oversampling and bleaching factors to be considered. In our experience, the smaller particles are unsatisfactory due to the excessive zoom resulting in bleaching.

The bead test sample consists of relatively bright beads embedded in a slide. To measure the performance of the PMT using the CV bead test, the laser intensity is reduced through the AOTF to a set value allowing the PMT voltages to vary between 500 and 700. Lowering the laser power with the AOTF resulted in PMT settings that were equivalent to those observed with a dim particle or with a biological particle. The 10-μm beads are homogeneous, they do not bleach quickly, and they can be used in PMT voltage ranges that test the functionality of the PMT. A dimmer test bead may indeed be better than a bight particle if it has uniform fluorescence. A dim particle would allow the PMT to be run at higher values and the AOTF to be run in the middle range instead of the high attenuation ranges. We were fortunate to obtain a population of homogeneous 10-μm particles that had a CV of only 5%. It is significant that biological samples (FluoCells and chicken cells [Figs. 7, 8]) with relatively dim fluorescence had CV measurements that maintained a similar relationship as with beads shown in Figures 4–6.

A CV test comparison between a 15-mW quiet air-cooled argon laser in a Zeiss 510 unit and our noisy 75-mW Omnichrome argon-krypton laser in a Leica TCS-SP1 revealed that less noise was produced using a 15 mW argon laser. The CV test showed that operating our TCS-SP1 unit with slow scans increased the noise compared with medium or rapid scans, which are presumably due to laser fluctuations of the argon-krypton laser during scanning. As the dwell time on each pixel was longer, more laser fluctuations could be detected by each pixel measurement. The peak-to-peak noise (4%–5% Omnichrome argon-krypton) of a laser was being detected. We anticipate that this observation will be expanded into a useful laser noise test in future QA studies.

Flow Cytometry Comparison

Our approach was to use a methodology similar to that used in flow cytometry to evaluate system performance. Both alignment beads and fluorescent intensity beads are used routinely to evaluate the performance of the flow cytometer. The alignment beads measure background noise, electronic noise, laser instability, fluidic instability, and system alignment. Changes in PMT settings and CV values over days indicate a change in flow cytometer performance. Most variations found in flow cytometer are due to fluidic problems. The tests can also observe system-wide problems. In our laboratory, the increase of the CV of a population of alignment beads is indicative of the flow cytometer not performing optimally.

Can a test, similar to that used by flow cytometrists, be developed in order to test confocal microscope perfor-

mance? The flow cytometer uses thousands of beads to produce a distribution of fluorescence intensity values. It is impossible to count that many beads and deal with the issue of multiple focal planes to get a meaningful distribution of fluorescence intensity. Multi-intensity beads were used initially. This test resembled closely the test system used in a flow cytometer to check linearity and sensitivity (17–21). Although the test provided an indication of the system's performance and also allowed many beads to be examined in a single scan (Figs. 1, 2), it exhibited inherent errors due to inconsistent field illumination, the existence of beads at different focal planes, and unequal illumination, due to the limits of resolution, at the edges of the bead. These factors will increase the CV of any population of bead intensities measured using a microscope. Because this test measured only the pixels in the image field, we substituted one large bead in the place of the multi-intensity beads. Using one large bead and measuring the pixels contained within a large ROI, we analyzed the inherent sources of noise. By zooming the bead 4×, problems with field illumination, focal excitation plane, and the presence of unequal illumination at the edges of the bead were eliminated.

This single-intensity bead test was used to determine the sensitivity and performance of the CLSM and to assess PMT functionality. The CRBC sample and Fluo slide were used to demonstrate how a biological sample displaying structural detail cannot be used as effectively as a statistical test (beads) to evaluate confocal system performance. They also demonstrated that the same relationships between PMT voltage, laser power, and average are maintained with beads and biological particles. A user-defined biological sample is still useful and worthwhile to test system performance. However, it is subjective in nature and not as effective in deriving reproducible statistical results as those gained with bead tests. However, biological samples should continue to be used as a relatively quick, easy, subjective assessment of machine performance. The biological sample, by its very nature, yields a population of pixel intensities with many different gray scale values (GSV). For instrument standardization (i.e., flow cytometry), it is necessary to have a sample that is relatively uniform in intensity (homogeneous). It appears that the fluorescent intensity bead is superior to a fluorescent biological particle in this respect. As always, the instability of the dyes over time and the variability of staining must be considered if daily tests between machines are compared. However, in our experience, daily tests on a confocal microscope do not appear to be as useful, simple, or as efficient as they are on a flow cytometer.

In summary, the data derived from CV bead analysis suggests that a relationship exists among image quality, optical efficiency, laser power, and PMT voltage. This study illustrates clearly the relationship between noise and image quality. It reiterates that efficient PMTs should be used and operated at low voltages when possible. The CV measures image quality and is affected by many variables, including PMT voltage, frame averaging, pinhole size, laser power, PMT type, scan speed, and optical efficiency. The image quality of any biological sample is affected by the same variables that are necessary to achieve a low CV in a bead sample. The best way to achieve good image quality is to have an optically efficient excitation/emission system with an efficient detection system. In order to achieve better image quality, it may be necessary to increase the laser power, decrease the scan speed, increase the pinhole size, and increase frame averaging. It must be emphasized that bleaching must be controlled when running the system with high laser powers and excessive frame averaging. Antifading agents may have to be added to fixed samples to reduce the expected bleaching.

ACKNOWLEDGMENTS

We thank Jeff Wang (Spherotech) for providing us with a series of test particles to evaluate and finally produce a particle of uniform size and fluorescence intensity that did not bleach readily with repeated samplings. We also thank Earl Puckett (USEPA/NHEERL machine shop) for making the power meter detector holders to fit on the microscope stage and Keith Tarpley (Computer Sciences Corp., Morrisville, NC) for his help in figure presentation.

LITERATURE CITED

1. Zucker RM, Price OT. Practical confocal microscopy and the evaluation of system performance. Methods 1999;18:447–458.
2. Zucker RM, Price OT. Evaluation of confocal microscopy system performance. Cytometry 2001;44(4):273–294.
3. Centroze V, Pawley J. Tutorial on practical confocal microscopy and use of the confocal test specimen. In: Pawley J, editor. Handbook of biological confocal microscopy, 2nd edition. New York: Plenum Press; 1995. p 559–567.
4. Sheppard CJR, Shotton DM. Confocal laser scanning microscopy. New York: Bios Scientific Publishing; 1997.
5. Marjlof L, Forsgren PO. Accurate imaging in confocal microscopy. In: Matsumato B, editor. Methods of cell biology. San Diego: Academic Press; 1993. p 79–95.
6. Carter D. Practical considerations for collecting confocal images. Methods Mol Biol 1999;122:35–57.
7. Pawley J. Fundamental limits in confocal microscopy. In: Pawley J, editor. Handbook of biological confocal microscopy, 2nd edition. New York: Plenum Press; 1995. p 19–36.
8. Pawley J. The 39 steps: a cautionary tale of quantitative 3-D fluorescence microscopy. Biotechniques 2000;28:884–886.
9. Rost FWD. Quantitative fluorescence microscopy. Cambridge: Cambridge University Press; 1991. p 1–62.
10. Jericevic Z, Wiese B, Bryan J, Smith LC. Validation of an imaging system: steps to evaluate and validate a microscope imaging system for quantitative studies. Methods Cell Biol 1989;30:47–83.
11. Van den Doel LR, Klein AD, Ellenberger SL, Netten H, Boddeke FR, van Vliet LJ, Young IT. Quantitative evaluation of light microscope based on image processing techniques. Bioimaging 1998;6:138-149.
12. Cardullo RA, Alm EJ. Introduction to image processing. Methods Cell Biol 1998;56:99–115.
13. Wolf DE. Introduction to image processing. Methods Cell Biol 1998; 56:117–134.
14. Rigaut JP, Vassy J, Herlin P, Duigou F, Masson E, Briane D, Foucrier J, Carvajal-Gonzalez S, Downs AM, Mandard AM. Three-dimensional DNA image cytometry by confocal scanning laser microscopy in thick tissue blocks. Cytometry 1991;12:511–524.
15. Czader M, Liljeborg A, Auer G, Porwit A. Confocal 3-dimensional DNA image cytometry in thick tissue sections. Cytometry 1996;25:246–253.
16. Liljeborg A, Czader M, Porwit A. A method to compensate for light attenuation with depth in three-dimensional DNA image cytometry

using a confocal scanning laser microscope. J Microsc 1995;177 (Pt 2):108–114.
17. Hoffman RA. Standardization, calibration and control in flow cytometry. In: Robinson P, editor. Current protocols in cytometry. New York: John Wiley & Sons; 2000. p 1.3.1–1.3.18.
18. Shapiro H. Practical flow cytometry, 3rd edition. New York: Wiley-Liss; 1995. p 190–198.
19. Watson JV. Introduction to flow cytometry. Cambridge: Cambridge University Press; 1991. p 150–185.
20. Watson JV. Flow cytometry data analysis: basic concepts and statistics. Cambridge: Cambridge University Press; 1992. p 4–125.
21. Muirhead K. Quality control for clinical flow cytometry. In: Bauer KD, Duque RE, Shankey TV, editors. Clinical flow cytometry principles and applications. Baltimore: Williams and Wilkins; 1993. p 177-199.
22. Zucker RM, Elstein KH, Gershey EL, Massaro EJ. Increased sensitivity of the Ortho analytical cytoflurograph by modifying the fluidic system. Cytometry 1990;11:848–851.
23. Russ JC. Image processing handbook, 3rd edition. Boca Roton: CRC Press; 1998.
24. Steyger P. Assessing confocal microscopy systems for purchase. Methods 1999;18:435–446.
25. Art J. Photon detectors for confocal microscopy. In: Pawley J, editor. Handbook of biological confocal microscopy, 2nd edition. New York: Plenum Press; 1995. p 183–195.
26. Pawley JB. Sources of noise in three dimensional microscope data sets in three dimensional confocal microscopy: volume investigations of biological specimens. New York: Academic Press; 1994.
27. White NS, Errington RJ, Fricker MD, Wood JL. Aberration control in quantitative imaging of botanical specimens by multi dimensional fluorescence microscopy. J Microsc 181:99–116.
28. Piston DW. Choosing objective lenses: the importance of numerical aperture and magnification in digital optical microscopy. Biol Bull 1998;195:1–4.