Microsc. Microanal. 17, 598–606, 2011
doi:10.1017/S1431927611000237

Microscopy AND Microanalysis

Quality Assurance Testing for Modern Optical Imaging Systems

Robert F. Stack,[1] Carol J. Bayles,[2] Anne-Marie Girard,[3] Karen Martin,[4] Cynthia Opansky,[5] Katherine Schulz,[5] and Richard W. Cole[1,*]

[1]*Wadsworth Center, New York State Department of Health, P.O. Box 509, Albany, NY 12201, USA*
[2]*Cornell University, Life Sciences Core Laboratory Center, Weill Hall, Ithaca, NY 14853, USA*
[3]*Oregon State University, Center for Genome Research and Biocomputing, 3021 Ag Life Sciences Bldg., Corvallis, OR 97331, USA*
[4]*West Virginia University, Department of Neurobiology and Anatomy, P.O. Box 9128, Morgantown, WV 26506, USA*
[5]*Blood Center of Wisconsin, Blood Research Institute, 8733 Watertown Plank Rd., Milwaukee, WI 53226, USA*

Abstract: The days of being able to ascertain instrument performance by simply peering through the eye pieces at a specimen are gone. However, users and granting agencies need to be confident that data collected on these instruments is uniform and quantifiable both over time and between instruments. Ideally, a LASER should not fluctuate, illumination should be completely uniform, and colors should be perfectly aligned. To check the current performance of imaging equipment, we conducted a worldwide research study utilizing three image-based tests: long-/short-term illumination stability, co-registration of signals across various wavelengths, and field illumination uniformity. To differentiate between "acceptable" and "unacceptable" performance, the deviation in illumination power could not exceed 10% (long term) or 3% (short term), the difference in the center-of-mass of imaged multicolored beads could not exceed >1 pixel between different wavelengths, and field illumination values could not exceed 10% (horizontal) or 20% (diagonal) deviation. This study established the *current state* of microscope performance through simple, efficient, and robust tests, while defining relative standards to assist cores in maintaining their instruments in optimal operating conditions. We developed cross-platform performance standards that will improve the validity of quantitative measurements made using various light microscopes.

Key words: confocal, stability, performance testing, wide field, co-registration, field illumination

Introduction

The light microscope has played a highly influential role in science since its invention in 1595 (Rosenthal, 2009). Modern light microscopes are highly evolved opto-electronic-mechanical devices, costing hundreds of thousands of dollars. The past 35 years have borne witness to an impressive revival of light microscopy. Considerable developments in the field of light microscopy have occurred over this relatively short time frame, ranging from the development of powerful new analytical techniques (i.e., STED, 4 PI, CARS, PALM, STORM) and instrumental capabilities to the striking improvements in modern image processing algorithms. These advances have allowed the light microscope to be used as a powerful quantitative research tool. Improvements in the design of the optical components include, for example, aberration-corrected objective lenses (correction of both chromatic and spherical aberrations), more efficient filters, the common use of LASERs (Light Amplification by Stimulated Emission of Radiation) as an illumination source, the introduction of acousto-optical beam splitters (AOBS) and vastly improved detection devices [i.e., photomultiplier tubes (PMT), charged-coupled device (CCD) cameras, and single photon avalanche diodes]. One of the most important advances has been the development of the confocal microscope, which combines "blur-free" optical sectioning of thick (>50 μm) fluorescent material with resolution similar to that of a light microscope (Pawley, 2006). As a result of these improvements, as well as improved performance and functionality of microscopy systems, there has been a significant increase in costs and complexity of these types of instruments.

The increased cost, coupled with shifting grant support, has resulted in these "high-end" instruments being placed more frequently into multiuser facilities, i.e., imaging cores, with trained personnel. The establishment of imaging cores has shifted responsibility for instrument acquisition, maintenance, and training away from the end user. Imaging cores across the country (and globe) now play a key and increasing role in methods development that has and will continue to aid investigators in advancing their research programs. This ultimately enhances overall scientific knowledge while fostering future collaborations (Guterman, 2010). Among the myriad functions of an imaging core's personnel is to routinely test the performance of their instrumentation. Core users need to be confident that the data collected will be consistent both over time and between specimens. Therefore, the need has arisen to develop good operating

Received October 5, 2010; accepted February 20, 2011
*Corresponding author. E-mail: rcole@wadsworth.org

practice procedures for imaging instrumentation; this would ideally include traceable standards. Until the last decade, simply observing a specimen with the microscope and obtaining what an individual researcher considered to be a "good" image was a sufficient performance test. However, the numerous advances made in light microscopy in recent years necessitate the identification of traceable standards and procedures. Ultimately, these standards would come under the supervision and guidance of the National Institute of Standards and Technology (NIST).

Analogous to the current state-of-performance testing in light microscopy, the state-of-performance testing in mass spectrometry was in its infancy in the not-so-distant past. What started with minimal manufacturer chosen tuning and calibration standards (two to three decades ago) eventually became the NIST Standard Reference Data Program, which currently contains mass spectra for over 15,000 compounds. Certified performance standard testing materials for many types of mass spectrometry are now available through the NIST Standard Reference Materials program. Lab-specific acceptance criteria are now common during the purchasing process in a mass spectrometry core, and proteomics data acceptance criteria (for publication) have become routine (Latterich, 2006; Nesvizhskii et al., 2007). In the field of quantitative real-time polymerase chain reaction, guidelines have been established recently related to the minimum information for publication of experimental data (Bustin et al., 2009). These guidelines seek to establish a universal baseline for the validity of results *prior* to publication, and primarily focus on aspects such as sample preparation, reagent quality, and data analysis. The fact that the guidelines do not focus on *instrumental* quality assurance does not detract from the reality that yet another crucial analytical field in biomedical research is undergoing an evolution of sorts related directly to quality assurance. Given that various federal government agencies, including the U.S. Food and Drug Administration, NIST, and the National Institutes of Health have already started to move forward with the development of imaging performance standards (Reiss, 2010), one should expect it will only be a matter of time before "standard" images and reference materials become not only available, but integrated with the evaluation and purchasing process for light microscopes. As evidence of this trend, ASTM International (formerly the American Society for Testing and Materials) recently released ASTM E2719, the *Guide for Fluorescence-Instrument Calibration and Qualification*. The guide is primarily designed for steady-state spectrofluorometers but, as pointed out in a recent review, contains many aspects that may be adaptable to other fluorescence applications (DeRose & Resch-Genger, 2010). Finally, the trend toward standardization in imaging is branching out in many ways. For instance, there have been recent calls for the use of reproducible imaging units reported by cameras; away from the arbitrary units commonly known as grayscale intensities to a more quantitatively sound unit such as photoelectrons (Sharma, 2010).

Methods and Materials

Study Design

The rationale for this study was to determine the current "state" of performance for optical microscope systems worldwide. To reach the largest possible audience (Fig. 1), the study was posted on the confocal list server (http://lists.umn.edu/cgi-bin/wa?A0=confocalmicroscopy) as well as announced at several meetings focused on optical microscopy. A total of 23 laboratories across 8 countries participated in the study, submitting data from a total of 36 microscopes. Most of the data from responding laboratories was from confocal microscopes representing most major manufacturers (Fig. 2). The acceptance criteria used for the study were based on the industry standard for instrument performance. All interested laboratories received the necessary test material free-of-charge [courtesy of The Association of Biomolecular Resource Facilities (ABRF)]. The individual laboratories reported back data from their microscopes; we then collated and tabulated the results found here.

This study was not intended to compare the performance of different instrument manufacturers, or to ascertain which brand had better performance in a given area. Instead, our purpose was to provide a quality assurance basis for current and future optical microscopy platforms.

LASER, Stage, and PMT Stability

This test was designed to measure the stability of the LASER(s) and ultimately the illumination system of confocal laser scanning microscopes or wide-field microscopes. The first step of the protocol was to warm up the LASER(s) or illumination source for a minimum of 1 h. Next, an image of an appropriate fluorescently-doped plastic slide (Chroma, Bellows Falls, VT, USA) was obtained using a 10× or 20× [low numerical aperture (N.A.)] microscope objective, focusing approximately 20 μm into the plastic. Because several different LASER lines may be imaged with one slide, we recommend using the *red* slide for most LASER lines because it has the largest excitation/emission range. We recommend focusing a scratched surface on the slide first, and then moving the stage slightly aside before focusing into the slide. The general acquisition parameters were as follows: detector gain and offset for each PMT detector were set such that the mean gray level within the image (with an 8 bit setting) had a value of approximately 150 (out of 255) with no saturated pixels (values of 255). The LASER power, gain, and offset values varied depending on which LASER line(s) were tested. For long-term stability tests, images were collected every 30 s for a total of 3 h and for short-term stability images were collected every 0.5 s for a total of 5 min. These time intervals were chosen to represent relevant imaging paradigms, i.e., the acquisition of a single Z stack (5 min) or, in the case of the 3 h test, the acquisition of multiple samples in an experiment. Sequential acquisition mode was used to collect as many LASER lines as possible within one experiment. Ideally, a different PMT should be used for each LASER line. No frame or line averaging was

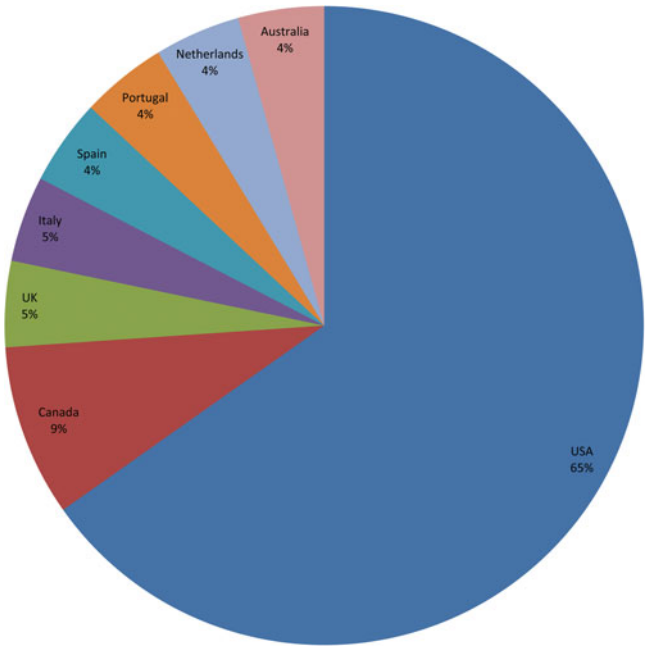

Figure 1. Pie chart indicating the countries that provided data for this study. A total of 23 laboratories from 8 different countries participated in the study.

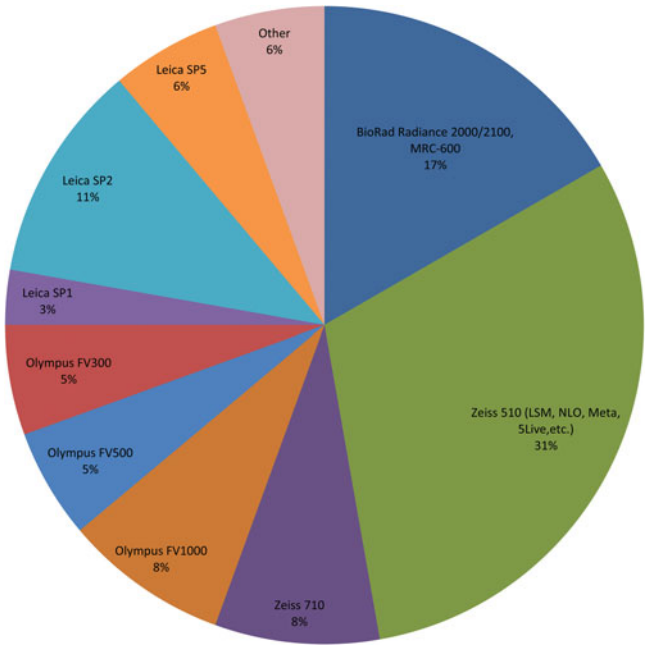

Figure 2. Pie chart illustrating the different imaging workstations that provided data. Data from a total of 36 microscopes were submitted.

used during image acquisition. Finally, as a way to check for photobleaching during image acquisition, the fluorescent slide was shifted laterally by approximately one-half of the field of view and an additional image was taken for all wavelengths imaged. While modern LASERs are very stable ($< \pm 1\%$ drift after warm-up), the overall observed (at the specimen) stability will be successively lower as a result of instability in optical/electronic elements within the microscope. The LASER stability metrics were based on nominal microscope manufacturer acceptance criteria and routinely obtainable values.

PMT Co-Registration

This test measured to what extent different detectors (PMTs or CCDs) will co-register (superimpose). Similar to the stability tests, the first step of the protocol was to warm up the LASER(s) or illumination source for 1 h. Next, TetraSpeck^{TM} 4 μm fluorescent microspheres (Invitrogen, Inc., Carlsbad, CA, USA) were imaged in three dimensions with a 40$\times$ or higher objective lens with a high N.A. ($\geq$1.2). A Z-series of images was collected using sequential scanning to allow as many detectors/wavelengths as possible to be used. For wide-field systems, different filter cubes or filter wheel settings were used in place of sequential scanning. The purpose of this test was to measure the extent of co-registration, or superimposition, of different detectors. The TetraSpeck^{TM} slides were prepared in the lab using approximately 4 μL of the 4 μm microsphere solution (#T-7284) and 8 μL of ProLong Gold® antifade reagent (#P36930, Invitrogen, Inc.). We encouraged performing this test on more than one bead to help separate results from aberrant beads from that of systematic co-registration issues. The general acquisition parameters were as follows: the pixel size was set to approximately one-half the resolution of the objective lens (Pawley, 2006). A zoom of approximately 10 was used for confocal laser scanning microscopes, and the images were collected using a standard three or four color protocol. For confocal microscopes equipped with a multiphoton (MP) LASER, the non-descanned detectors (NDDs) were used to detect the MP channel(s). PMT settings were similar to those for the laser stability tests with average intensities of beads of $\sim$150 (gray level) and no saturated pixels. Since modern imaging systems are designed to produce quantifiable data, there needs to be high fidelity in the registration between imaging channels. This will ensure that separate channels co-register to a single pixel. The PMT co-registration metric was based on nominal microscope manufacturer acceptance criteria and routinely obtainable values.

Field Illumination

This test measured the uniformity of illumination across the scanned image field. Similar to the two previous tests, the first step of the protocol was to warm up the LASER(s) or illumination source for 1 h. This test was performed on all objective lenses available and was initially conducted at the manufacturers minimum zoom specification (this varies between manufacturers, generally between 0.7–1.25). Images were collected of a cover-slipped area from the green/orange fluorescently-doped plastic slide (Chroma), using a two to four line average per frame. The 488 nm or 543/561 nm LASER combination was used at powers set such that the PMT gain on the mean pixel value was approximately one-half the maximum, i.e., 125 out of 255. Once again, modern imaging systems are designed to produce

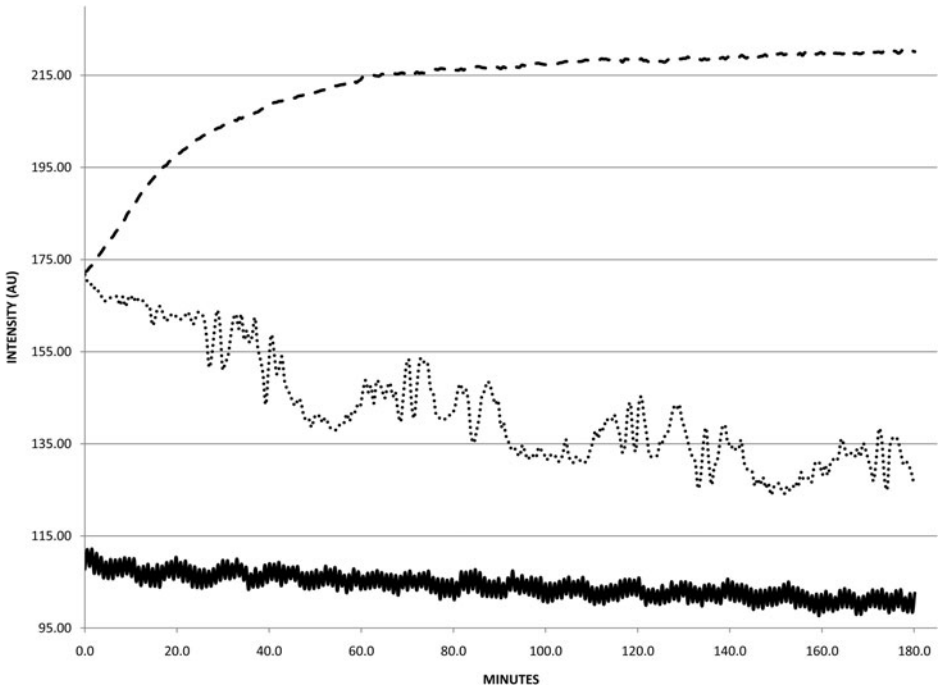

Figure 3. LASER stability (intensity) versus time (3 h) demonstrating both an acceptable (solid line, 488 nm) and two unacceptable confocal systems; improper warm-up (dashed line, 633 nm), many possible causes (dotted line, 543 nm). The acceptance criteria for illumination stability could not exceed 10% (long term). This test is designed to simulate stability during the capture of multiple specimens.

quantifiable data; uniformity of the illumination across the imaging field is paramount. While zero deviation would be ideal, the metric was based on nominal microscope manufacturer acceptance criteria and routinely obtainable values.

Data Analysis

Data analysis was performed for the LASER/PMT stability tests, and measurements of the range (maximum and minimum), the standard deviation, the mean, the spread, the spread/mean, the percent change, and the variance were calculated. The PMT co-registration data analysis was performed on single bead image stacks. For each color image of the beads, a line scan function was used to plot the intensity across a single bead for each slice in the Z-series. Using ImageJ's measurement function, the center-of-mass for the "most in focus" slice (i.e., brightest) was determined and compared for all PMTs to measure the co-registration. Finally, the field illumination data analysis was performed by using three distinct line scan profiles (horizontal and two diagonals; all intersecting the image center) to check for intensity drop off near the edges of the field. The percent change in intensity from the center of the image to the edge or corner was calculated for each line profile.

Results

LASER Stability

In addition to instability within the LASER's themselves, temporal measurements of image intensity data can be influenced by many factors. For example, problems with the detection system (PMTs), acousto-optical tuning filter (AOTF), stage (*x*, *y*, or *z* axis drift), and photobleaching of the test slide, among others, can all contribute to poor overall performance (Pawley, 2000; Zucker & Price, 2001; Zucker, 2006*a*). Figure 3 illustrates long-term stability curves from three different microscopes. The (solid) line shows acceptable 3 h stability, with variations in intensity that fall within the study criteria. However, other systems show unacceptable long-term stabilities, for different reasons. Data shown from the diode LASER (dotted line) could be due to variability in the optical system and not just the LASER itself because one would expect a diode LASER to be more stable than observed here. The instability seen in another case (dashed line) is most likely due to insufficient warm-up of the LASER, a condition that is thankfully simple to correct. The long-term stability test was designed to mimic experimental variations that would be seen over the acquisition of multiple specimens. Operating conditions that would result in the stability profiles represented by the diode laser and LASERS that are not warmed up properly would result in unacceptable experimental data, regardless of the specimens being imaged.

Typically, variations in LASER intensity that differ by no more than 5% over the short time span (5 min) are acceptable (Fig. 4, solid line). However, many systems tested showed unacceptable short-term stabilities (Fig. 4, dashed and dotted lines). It is important to remember that the short-term test was designed to mimic experimental variations that would be observed over the acquisition of a single Z-series. Therefore, any quantification from data collected

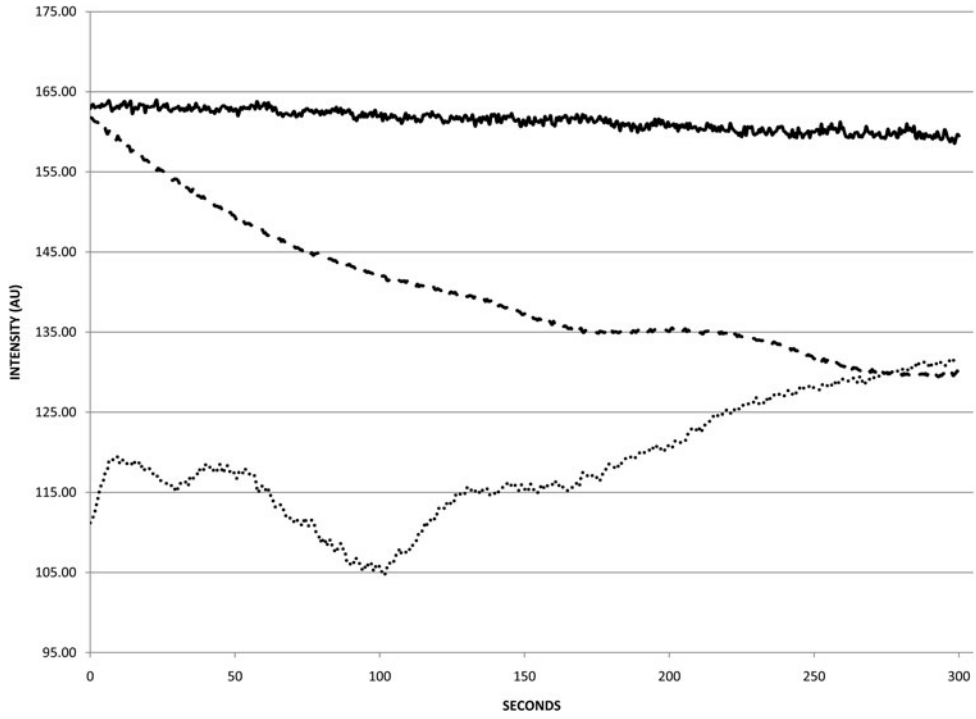

Figure 4. LASER stability versus time (5 min) demonstrating both acceptable (solid line, 594 nm) and two unacceptable confocal systems [dashed line (561 nm) and dotted line (532 nm)]. The acceptance criteria for illumination stability could not exceed 3% (short term). This test is designed to simulate stability during the capture of a single Z-series.

on unacceptable systems would be skewed; calculating probe concentrations or concentration differences across the sample would be extremely difficult with so much variability over such a short time frame. Finally, it is extremely significant to note that slightly more than one-half of the microscopes tested did not meet the acceptance criteria for the long- or short-term stability tests.

Co-Registration

Co-registration is a principal determining factor in the image registration of specimens labeled with multiple fluorophores. If the alignment of the system is not identical (both laterally and axially) for all of the PMT channels or filter cubes/filter wheel positions used in an imaging experiment, the overlay of these channels will result in a substandard "registered" image. This would mean that even for perfectly co-registered samples ideal co-registration would not be observed from image data collected on these systems. An "acceptable" co-registration is seen when the intensity of a multicolored 4 μm fluorescent bead that is imaged sequentially with different laser excitation wavelengths in five separate PMT detectors shows nearly perfect overlap along the lateral (x, y) axis (Fig. 5). There is near-perfect overlap of the five separate curves; i.e., none of the curves are shifted to the left or right, which indicates that the system has excellent lateral co-registration. A simple way of determining axial co-registration is to see if the bead size is similar in different color channels. If the z-axis focus is shifted for different color channels, then for a fixed image plane in space the bead will have a different axial location and will appear to have a different size for the various colors. Thus, from the same plotted lateral dataset, the axial registration can be measured full-width at half-maximum (FWHM) of each curve. For the example shown here, the FWHM of each individual curve is nearly identical, indicating a system that also has excellent axial co-registration (Fig. 5). Ideally, for high quality, high N.A. lenses, co-registration values measured should exhibit no more than a *one pixel* shift. One wide-field microscope tested by imaging the same multicolored 4 μm fluorescent beads using three separate filter cubes showed "unacceptable" levels of co-registration (Fig. 6). The intensity profiles of the beads from the different colored images do not overlap, especially the "TRITC" curve (solid line), indicating a problem with the lateral co-registration. The different apparent sizes also indicate poor axial registration. It is important to note that slightly more than one-third of the microscopes tested did not meet the acceptance criteria for the center-of-mass displacement.

Field Illumination

Field illumination plays a pivotal role in determining the validity of quantitative measurements. A varied photon count as a result of an uneven illumination will essentially render such measurements less and less useful as a function of the degree of nonuniform illumination (Zucker, 2006*b*). Line scan data were taken from three separate intensity profiles across an image: a horizontal line scan across the center of the image and two diagonal line scans from corner to corner. An example is shown of the variations in intensity levels for all three line profiles that fall within our accep-

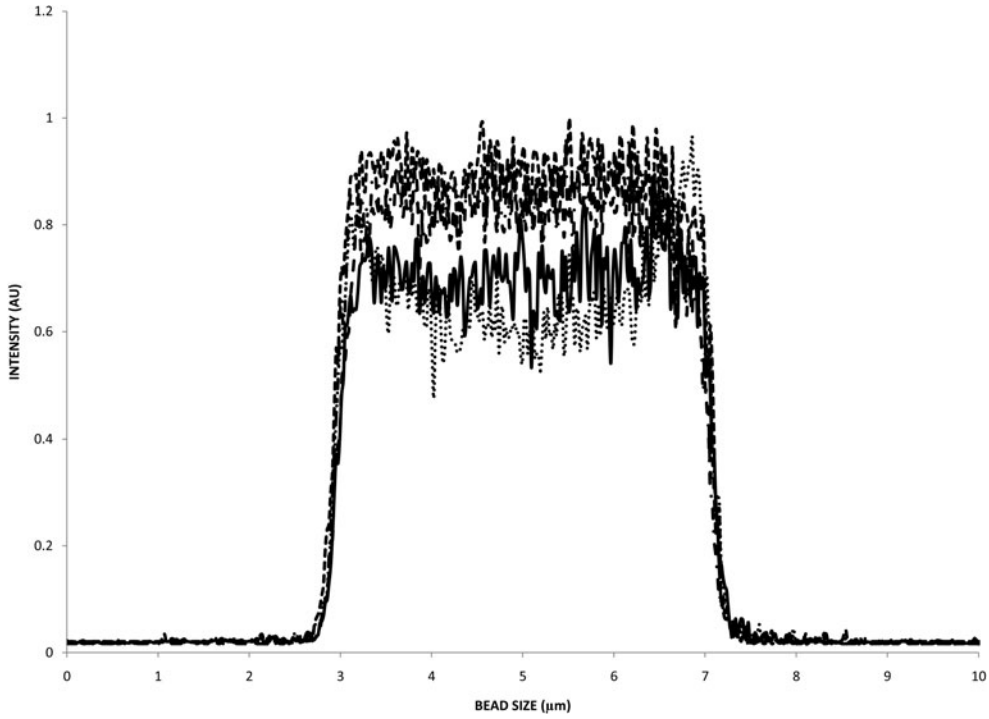

Figure 5. Acceptable co-registration. Line scan data from a 4 μM bead imaged with five different PMTs. The near superimposition of the traces indicates a high degree of registration between the five detectors; multichannel data such as this could be interpreted with a high level of confidence.

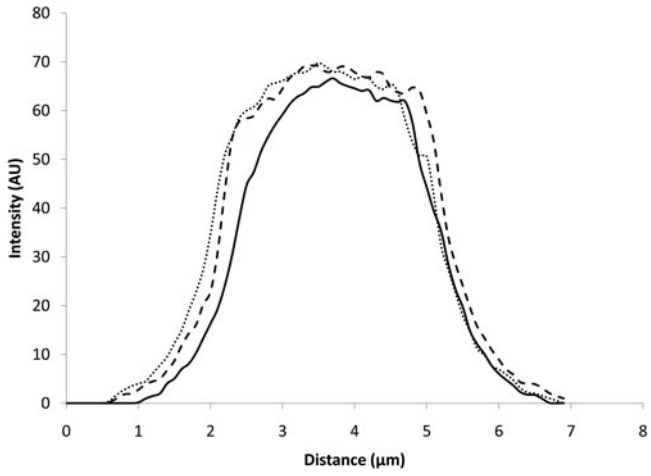

Figure 6. Unacceptable co-registration. Line scan data from a 4 μM bead imaged with three different filter cubes (wide-field system). The three channels imaged were: DAPI (dashed line), FITC (dotted line), and TRITC (solid line). The lateral displacement of the traces indicates a poor degree of registration between the three filter cubes and extreme caution would be needed when interpreting multichannel data.

tance criteria, which indicates a well-aligned microscope (Fig. 7). However, some systems showed acceptable field illumination along the horizontal line (Fig. 8, solid black line), while both diagonal profiles showed unacceptable field illumination uniformity (Fig. 8, dashed and dotted lines). The diagonal line scans indicate a nonuniform illumination pattern in the lower left and right quadrants of the image. In this particular case, increasing the zoom factor used will not significantly improve the quality of the field illumination. To regain acceptable illumination uniformity, this microscope may need to be realigned or the objective lens may require service. Surprisingly, less than 5% of the 93 submitted datasets were at or below the acceptable deviation along both the horizontal and diagonal axes (acceptable deviation was set to 10% intensity variation along the horizontal axis and 20% along the diagonal axis). No significant improvement in these results was achieved after applying a three point rolling average to reduce noise within the images. Even after utilizing the rolling average, the mean deviation for the horizontal line scans was 37% and 45% for the diagonal line scans. There are several possible causes for these results: very low intensity pixels at the edges of the image (usually due to scan or raster errors), a large amount of noise in the image (gain of PMT too high), and most concerning, poor alignment of the laser inputs or the pinholes (usually occurs on the illumination side). To address the first two causes for poor field illumination, we are currently developing a mathematical ranking model that will weigh specified factors and produce a single field uniformity value rather than looking at the individual intensity profiles. See future directions for a more complete description of the model.

Discussion

LASER, Stage, and PMT Stability

Based on the poor overall performance (all three tests) of the imaging systems tested from across the world, the im-

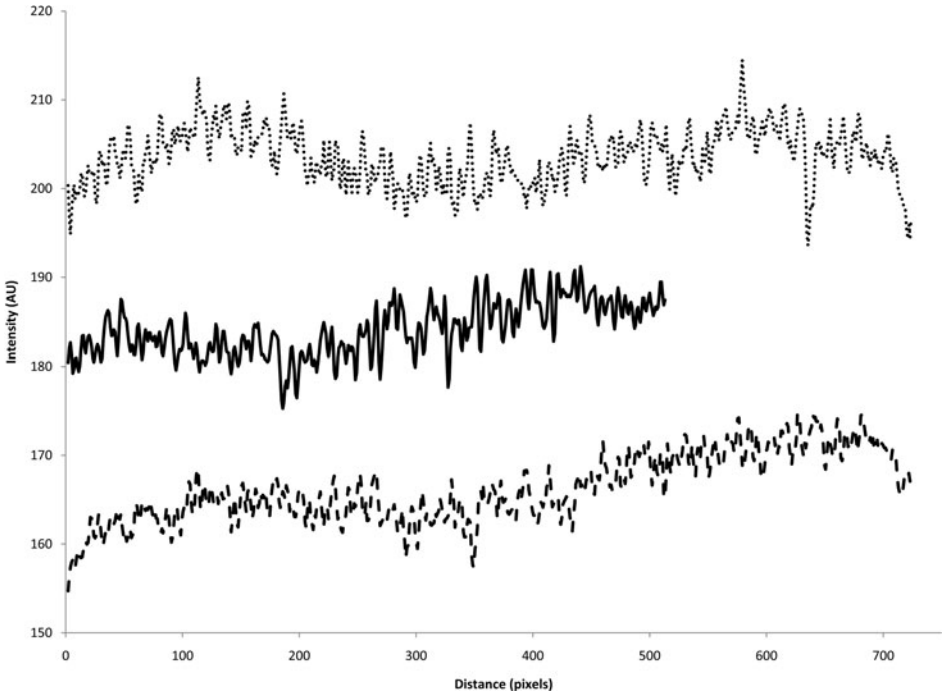

Figure 7. Acceptable field illumination. Intensity plot profiles (one horizontal and two diagonal axes) from an image of a homogeneous fluorescent slide. The solid line represents a 3-point rolling average of the horizontal line scan while the dashed and dotted lines represent a 3-point rolling average of the ULLR (upper left to lower right) and LLUR (lower left to upper right) diagonal line scans, respectively. The total deviation is 10.3%, 12.0%, and 8.7%, which meets the acceptance criteria of 20%, 20%, and 10%, respectively. Comparison of pixel values within an image generated with such illumination uniformity could be interpreted with a high level of confidence. Some of the individual plot profiles were scaled for illustrative purposes.

portance of standardizing and conducting these tests on a regular basis cannot be overstated. During quantitative light microscopy image analysis, the performance of the microscope is crucial for obtaining accurate quantitative data. These tests not only reveal issues with LASER power stabilities, they will also reveal problems with other components of the imaging system such as instabilities in the stage, PMTs, and AOTF. Poor stability may be a consequence of room temperature fluctuations and/or insufficient heat dissipation in the LASER cooling system. Moreover, improper venting of the LASER cooling/exhaust system will cause room temperature fluctuations. In addition, insufficient warm-up of LASERs and/or photobleaching of the specimen can also cause apparent instability in the data from this test. Stage drift will cause a slow decrease (or increase) in *all* LASER lines. If stage drift is suspected as a cause of instability, removing the slide and using the transmitted detector (if available) will eliminate axial changes. Another method to determine stage stability is to use a mirror slide that allows small axial changes to be easily detected. LASER power can also be measured at distinct points along the optical path to determine which component(s) may be the root cause of the observed instability. Examples of possible inspection points would be: the LASER source, the entrance and exit of the fiber optics, after filter cube/filter wheels, and the exit of the AOTF (if so equipped). *Please note: accessing internal optical components may expose personnel to a risk of injury from both high voltages associated with LASER power supplies as well as possibly damaging radiation (both visible and in-visible). Therefore, consult a trained service engineer before attempting these measurements/alignments.* PMT stability can be easily determined by imaging with more than one PMT simultaneously. If the instability is similar with all PMTs, then they are most likely not the cause. In addition, room temperature fluctuations can be tracked with the use of temperature probes. There are many different types available; certain types can log temperature though a web browser interface via an ethernet connection. Ultimately, the stability data should be within $\pm 10\%$ of the initial value obtained for the 3 h test and within $\pm 3\%$ of the initial value obtained for the 5 min test.

PMT Co-Registration

Biological sciences rely on co-localization measurements to make conclusions about the biology they are studying. If systems are not performing up to a minimum standard, then these studies could be missing important interactions simply because the imaging systems are generating erroneous data. Image co-registration determines the quality of image overlays when using multiple fluorophores and detection channels. As evidenced by the intensity versus distance (i.e., bead size) data shown, issues related to *lateral* co-registration were observed as a lateral shift in the bead's center-of-mass. Issues related to *axial* co-registration were

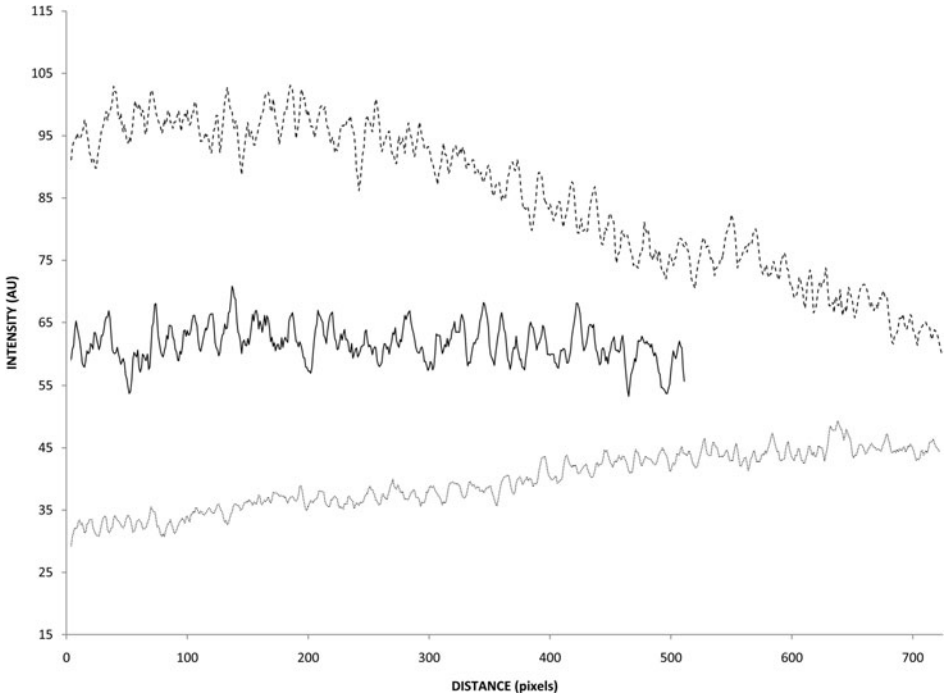

Figure 8. Unacceptable field illumination. Intensity plot profiles (one horizontal and two diagonal axes) from an image of a homogeneous fluorescent slide. The solid line represents a 5-point moving average of the horizontal line scan while the dashed and dotted lines represent a 5-point moving average of the ULLR and LLUR diagonal line scans, respectively. The total deviation is 64.9%, 70.0%, and 41.6%, which exceeds the acceptance criteria of 20%, 20%, and 10%, respectively. This amount of change across the image would make comparison of intensity within the image nearly impossible. Some of the individual plot profiles were scaled for illustrative purposes.

only observed as a "smaller" bead size in the intensity plots—in essence, a line with a smaller distance or pixel number width. In cases where both the lateral and axial co-registrations were skewed, the data revealed plot lines that were not only shifted to the left or right, but also had varying widths. The system would need to be realigned to correct the problems. For lateral co-registration deficiencies, the alignment of the fiber optic coupler (if more than one fiber is used for delivery) would need to be checked and aligned if necessary. The optical dichroic filters would also need to be checked and aligned if necessary. In systems utilizing acousto-optical crystals (AOTF/AOBS), they may need to be retuned or replaced. In the case of wide-field microscopes, the use of filter wheels is preferable (because no motion of optical elements occurs) or commercially available "zero-shift" filter cubes would be recommended. When a multiphoton LASER was tested, the NDDs were used with the following reasons in mind: because there is no "descan" as with internal PMTs, there is no emission pinhole, and the light path to the NDDs is shorter. All of these factors enhance the detection sensitivity of the NDDs. This performance test gives end users of microscopes a straightforward measure to determine if test data indicate a deficiency in the lateral and/or axial co-registration.

Field Illumination

Field illumination is a crucial instrument parameter, particularly if the images produced are intended to be used *quantitatively*. Nonuniform field illumination results in a different number of photons being delivered across the field of view. As a simple example, if you are trying to determine the level of expression of a fluorescent protein within your population of cells, the cells in the center of the image will appear to be brighter than those at the edges of the image. Therefore, when quantifying the cellular intensities, the intensity distribution will be much broader than expected simply due to field nonuniformity. Ultimately, gray level values should be linear with protein levels. An important consideration in setting up proper field illumination is that not all manufacturers have the same *default* zoom setting; therefore, it is critical that researchers operate, at least initially, using the manufacturers' recommended zoom settings. Utilizing the manufacturers recommended zoom setting, variations in pixel values should not exceed 10% in the horizontal plane and should not exceed 20% corner-to-corner. A well-aligned microscope will show a drop-off at these points that is symmetrical and within the aforementioned ranges. If the variation (drop-off) exceeds the acceptable range but is still uniform in nature, an increase in zoom may equate to an improvement in the evenness of the illumination. However, if the drop-off exceeds the acceptable range and is not symmetrical in nature, the microscope is in need of realignment. In this particular case, no amount of increased zoom will help improve the data. Potential remediation of field illumination deficiencies would include testing the pinhole alignment and LASER couplings. If a

misaligned pinhole is suspected, fully opening and retesting is advised (adjust LASER power and PMT gain appropriately). LASER couplings should be checked point-by-point to determine if/where the problem exists. Once we began to review the data to determine how many objective/zoom combinations would actually meet the acceptance criteria, we realized the testing procedure needed to be refined to reduce noise. Therefore, we decided to include a four-to-five frame average. Currently, we are collaborating with a biostatistician using the field illumination data from this study to create an algorithm capable of producing a "flatness factor" value based on field illumination line scan data. This will quickly and accurately give an imager a reference value of the uniformity of their particular system.

Future Directions

The next round of proposed tests will complement the three procedures described herein by adding new procedures. These new experiments will include: (1) a test of the system's overall performance and degree of blurring of a subresolution point object by measuring the point spread function, and (2) spectral wavelength testing and calibration of a system through the use of either a multi-ion discharge lamp as the illumination source, or a mirror slide to redirect the LASER output to the detector(s). Additionally, plans are underway to develop a robust nanoscale test specimen with a known fluorescent spacing and output. This test specimen will be a fluorescently-doped plastic with extremely fine, repetitive detail, which will aid in gauging not only a system's limit of resolution but also its overall optical efficiency. As we discovered, there are many factors that will contribute to nonuniform illumination. Among these are improper PMT gain and/or LASER settings that result in excessive noise, misaligned illumination sources, scan errors, and photobleaching. Not all of these errors are equally detrimental. For example, a few rows of dim pixels at the edge of an image will not impact the quality of the image. Conversely, a misaligned illumination system can severely affect image quality. As mentioned previously, we are currently developing an algorithm capable of interpreting field illumination data that will be far more useful than simple deviation along two axes. The algorithm will be not only user-friendly but will also be a more appropriate way for researchers to gauge the uniformity of their field illumination than simply calculating deviations in intensity from the center of the field. Ideally, this algorithm will at some point be available as a plug-in for ImageJ.

Our main goal is to have well-defined test specimens and protocols so that the average research lab can conduct these tests efficiently and measure the performance of their microscope. These results could be used as an evaluation tool in purchasing new equipment and also as a maintenance tool to ensure optimal and consistent instrument performance over the long term within labs or core facilities. This will be essential in ensuring that the quality of quantitative data is available to researchers for both short- and long-term research projects.

Acknowledgments

The authors wish to thank the ABRF for their continued and enthusiastic interest and for graciously providing funding for this important undertaking. In addition, the authors wish to thank Robert M. Zucker, George McNamara, and Claire M. Brown for their *ad hoc* participation on the light microscopy research group and their willingness to share their vast experience in the field of light microscopy QA. Finally, the authors acknowledge the use of the Wadsworth Center's Advanced Light Microscopy & Image Analysis Core Facility for a portion of the work presented.

References

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J. & Wittwer, C.T. (2009). The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* **55**(4), 611–622.

DeRose, P.C. & Resch-Genger, U. (2010). Recommendations for fluorescence instrument qualification: The new ASTM Standard Guide. *Anal Chem* **82**(5), 2129–2133.

Guterman, L. (2010). Access sparks action. *NCRR Reporter* **34**(1), 4–8.

Latterich, M. (2006). Publishing proteomic data. *Proteome Sci* **4**, 8.

Nesvizhskii, A.I., Vitek, O. & Aebersold, R. (2007). Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nat Methods* **4**(10), 787–797.

Pawley, J. (2000). The 39 steps: A cautionary tale of quantitative 3-D fluorescence microscopy. *BioTechniques* **28**(5), 884–886.

Pawley, J. (2006). *Handbook of Biological Confocal Microscopy*. New York: Plenum.

Reiss, S.M. (2010). Quality and standards: Making bioimaging measure up. *BioOptics World* **3**(1), 14–18.

Rosenthal, C.K. (2009). Light microscopy: Contrast by interference. *Nature Milestones* | Milestone **8**.

Sharma, D. (2010). Standardized units for reproducible imaging experiments. *Biophotonics* (September), 32–35.

Zucker, R.M. (2006*a*). Quality assessment of confocal microscopy slide-based systems: Instability. *Cytometry A* **69**(7), 677–690.

Zucker, R.M. (2006*b*). Quality assessment of confocal microscopy slide based systems: Performance. *Cytometry A* **69**(7), 659–676.

Zucker, R.M. & Price, O. (2001). Evaluation of confocal microscopy system performance. *Cytometry* **44**(4), 273–294.