

# REVIEW

## Reproducibility

# Designing a rigorous microscopy experiment: Validating methods and avoiding bias

Anna Payne-Tobin Jost  and Jennifer C. Waters 

**Images generated by a microscope are never a perfect representation of the biological specimen. Microscopes and specimen preparation methods are prone to error and can impart images with unintended attributes that might be misconstrued as belonging to the biological specimen. In addition, our brains are wired to quickly interpret what we see, and with an unconscious bias toward that which makes the most sense to us based on our current understanding. Unaddressed errors in microscopy images combined with the bias we bring to visual interpretation of images can lead to false conclusions and irreproducible imaging data. Here we review important aspects of designing a rigorous light microscopy experiment: validation of methods used to prepare samples and of imaging system performance, identification and correction of errors, and strategies for avoiding bias in the acquisition and analysis of images.**

## Introduction

Modern light microscopes are being pushed past their theoretical limits and are now routinely used not just for qualitative viewing but to make quantitative measurements of fluorescently labeled biological specimens. While microscopy is an incredibly powerful and enabling technology, it is important to keep in mind that when we look at microscopy images we are not looking at the specimen. Typically, in cell biology, the biological sample is first manipulated to introduce fluorophores, usually conjugated to a molecule of interest (Allan, 2000; Goldman et al., 2010). A microscope generates a magnified optical image of the spatial distribution of the fluorophores in the specimen (Inoué and Spring, 1997; Murphy and Davidson, 2012). The optical image is recorded by an electronic detector (e.g., a camera) to create a digital image consisting of a grid of pixels with assigned intensity values (Wolf et al., 2013; Stuurman and Vale, 2016). As a fluorescence microscopy image is therefore a digital representation of an optical image of the distribution of fluorophore that was introduced into the specimen, we must consider sources of error in the process and if the image accurately represents the biological phenomena being studied.

For quantitative microscopy, it is critical to validate that a method is capable of detecting the target and providing an accurate measurement. Imaging systems are complex and prone to systematic error that rarely presents in a way that is easily

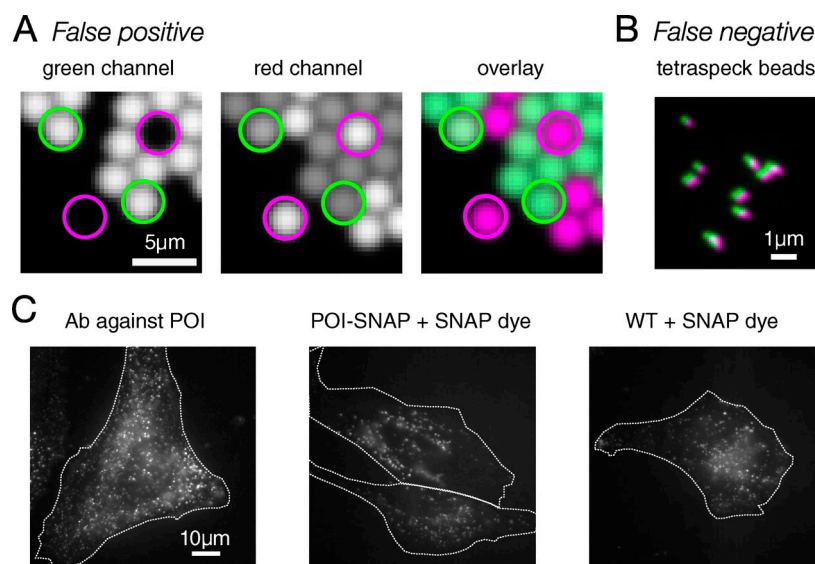
identified during routine use (Fig. 1, A and B; Hibbs et al., 2006; North, 2006; Zucker, 2006; Joglekar et al., 2008; Waters, 2009; Wolf et al., 2013). Therefore, researchers must consider the types of error that might affect interpretation of imaging data, use known samples and control experiments to validate methods and identify errors, and test approaches for error correction. Microscopes marketed or perceived as state-of-the-art or easy to use are still prone to error; there does not exist a microscope, commercial or home-built, that generates images free from all sources of error or that works for every quantitative application. While developments in techniques such as super-resolution and light-sheet microscopy can overcome some previous limitations, they also introduce additional sources of error and require rigorous validation (Hibbs et al., 2006; North, 2006; Joglekar et al., 2008; Waters, 2009; Wolf et al., 2013; Lambert and Waters, 2017).

Microscopy experiments are further complicated by bias inherent to visual assessment of images (Lee et al., 2018). We are all subject to bias in our perception of the meaning of images, patterns within images, and interpretation of data in general (Nickerson, 1998; Russ, 2004; Lazic, 2016). Due to the inherent unconscious nature of bias, we cannot reliably choose to avoid innate biases. Experiments should therefore be designed to minimize the impact of bias, which can accumulate through every step of an experiment that involves visual inspection and decision-making, from image acquisition to analysis.

Department of Cell Biology, Harvard Medical School, Boston, MA.

Correspondence to Jennifer C. Waters: jennifer@hms.harvard.edu.

© 2019 Jost and Waters. This article is distributed under the terms of an Attribution–Noncommercial–Share Alike–No Mirror Sites license for the first six months after the publication date (see <http://www.rupress.org/terms/>). After six months it is available under a Creative Commons License (Attribution–Noncommercial–Share Alike 4.0 International license, as described at <https://creativecommons.org/licenses/by-nc-sa/4.0/>).



**Figure 1. Image errors can lead to incorrect results. (A)** Bleed-through causes a false-positive colocalization result. Green and red beads were mixed and mounted together. There are no beads in these samples/images that are labeled with both green and red dye. With this filter and sample combination, there is significant bleed-through from the green beads into the red channel (see green circles). Pearson's *R* for colocalization is 0.67. Since no pixel that contains green fluorophore also contains red fluorophore, there should be no correlation. **(B)** Channel misregistration causes a false-negative colocalization result. Tetraspeck beads are labeled with four dyes, including dyes imaged in the green and red channels here. Because each bead is labeled with both dyes, there should be complete colocalization between channels, with an expected Pearson's *R* of 1. However, the imaging system has introduced significant misregistration between the channels, leading to a Pearson's *R* of 0.66. **(C)** Nonspecific dye binding leads to a false-positive result. A significant level of nonspecific binding of SNAP dye to cells containing no SNAP tag (WT + SNAP dye) looks qualitatively similar to both cells containing a SNAP tag fused to the POI and immunofluorescence against POI. White dotted lines indicate cell outlines.

Accuracy of imaging data and reproducibility of an experiment are related but distinct (Lazic, 2016). We define accurate imaging data to mean that the image and resulting conclusions correctly represent the specimen and reproducibility as the ability to repeatedly generate the same results using the same materials and methods. An experiment may generate accurate data that cannot be reproduced due to, for example, inadequate validation or methods reporting. Conversely, an experiment may be reproducible but inaccurate if systematic error in the method goes uncorrected. Additionally, sources of error vary between microscopes (Murray et al., 2007; Murray, 2013; Wolf et al., 2013), further contributing to irreproducibility if ignored.

Our goal in this review is to encourage skepticism of the accuracy of microscopy images and appreciation of the effects of bias on microscopy data. Rather than attempt to address the many different types of microscopy experiments used in cell biology research, we provide guidelines that apply to most microscopy experiments and give examples of how they might be implemented. We also provide select educational resources to serve as starting points for those seeking to learn more about specific microscopy techniques (Table 1). We discuss some common sources of systematic error in microscopy images, demonstrate how errors can lead to erroneous conclusions, and review methods of identifying and correcting errors. We discuss steps in microscopy experiments that are subject to bias and the importance of blinding and automation in reducing bias in image acquisition and analysis. This review is focused on issues and misconceptions we commonly encounter when advising and training postdoctoral fellows and graduate students in our microscopy core facility and courses, so principal investigators should take note that the topics presented here represent areas where trainees may need guidance and education.

### Validation and error correction

Validation of methods using known samples and controls demonstrates that a tool or technique can be used to accurately

detect, identify, and measure the intended target. Robust and reproducible microscopy experiments require validation of the imaging protocol, the specific technique used to fluorescently label the sample, and the type of imaging system used to measure the intensity and spatial distribution of fluorescence (Zucker, 2006; Murray et al., 2007; Joglekar et al., 2008; Wolf et al., 2013). Validation may reveal that a given approach will not work for the intended purpose and an alternative is needed; or, validation may demonstrate that image corrections or a change of optics are needed to correct systematic errors.

### Labeling method and sample prep

It cannot be assumed that an antibody, organic dye, or fluorescent protein (FP) will perform as an inert, specific label. All methods of labeling biological specimens with fluorophores have the potential for nonspecificity and for perturbation of the localization or function of the labeled component or associated structures, binding partners, etc. (Couchman, 2009; Burry, 2011; Bosch et al., 2014; Ganini et al., 2017). Immunofluorescence and FP conjugates have become so ubiquitous in cell biology that validation of binding specificity and that the biology is not affected by the label are frequently omitted (Freedman et al., 2016). In our experience, these issues are often more problematic than realized, and there are many published examples of commonly used probes introducing artifacts under specific experimental conditions (Allison and Sattenstall, 2007; Couchman, 2009; Costantini et al., 2012; Landgraf et al., 2012; Schnell et al., 2012; Bosch et al., 2014; Norris et al., 2015; Ganini et al., 2017).

**Autofluorescence.** Many endogenous biological components are fluorescent (e.g., Aubin, 1979; Koziol et al., 2006). Unlabeled controls (samples with no added fluorophore, but otherwise treated identically) are required to validate that fluorescence is specific to the intended target and that autofluorescence is not misinterpreted as the fluorescent label (for a famous example, see De Los Angeles et al., 2015).

Table 1. **Select educational resources**

Topic	Resources
Microscopy texts	<ul style="list-style-type: none"> <li>• Book: <i>Video Microscopy</i>, Inoué and Spring, 1997</li> <li>• Book: <i>Handbook of Biological Confocal Microscopy</i>, Pawley, 2006b</li> <li>• Book: <i>Fundamentals of Light Microscopy and Electronic Imaging</i>, Murphy and Davidson, 2012</li> </ul>
Microscopy courses	<ul style="list-style-type: none"> <li>• A comprehensive listing of microscopy courses can be found at <a href="https://www.microlist.org">https://www.microlist.org</a></li> <li>• iBiology Microscopy Series (<a href="https://www.ibiology.org/online-biology-courses/microscopy-series/">https://www.ibiology.org/online-biology-courses/microscopy-series/</a>; a free online microscopy course)</li> <li>• Microcourses YouTube channel: Short educational videos on microscopy</li> <li>• Course: Quantitative Imaging: From Acquisition to Analysis, Cold Spring Harbor Laboratory, NY</li> <li>• Courses and workshops offered by the European Molecular Biology Organization in Europe</li> <li>• Courses and workshops offered by the Royal Microscopical Society in the United Kingdom</li> <li>• Course: Bangalore Microscopy Course, National Centre For Biological Sciences, Bangalore, India</li> </ul>
Quantitative microscopy	<ul style="list-style-type: none"> <li>• Review: Waters, 2009. An introduction to sources of error in microscopy</li> <li>• Review: Wolf et al., 2013. Protocols for error detection and correction</li> <li>• Book: <i>Quantitative Imaging in Cell Biology</i>, Waters and Wittmann, 2014</li> </ul>
Live-cell imaging	<ul style="list-style-type: none"> <li>• Review: Ettinger and Wittmann, 2014. A practical introduction to fluorescence live cell imaging</li> <li>• Review: Magidson and Khodjakov, 2013. Strategies for reducing photodamage</li> <li>• Book: <i>Live Cell Imaging: A Laboratory Manual</i>, Goldman et al., 2010</li> </ul>
Colocalization	<ul style="list-style-type: none"> <li>• Review: Bolte and Cordelières, 2006. Validation and error correction in colocalization image acquisition and analysis</li> </ul>
Ratiometric imaging, including Förster Resonance Energy Transfer (FRET)	<ul style="list-style-type: none"> <li>• Reviews: Hodgson et al., 2010; O'Connor and Silver, 2013; Spiering et al., 2013; Grillo-Hill et al., 2014. Validation and error correction in image acquisition and analysis of ratiometric probes</li> </ul>
Fluorescence Recovery After Photobleaching (FRAP)	<ul style="list-style-type: none"> <li>• Reviews: Phair et al., 2003; Bancaud et al., 2010. Protocols for acquisition, correction, and analysis of FRAP data</li> </ul>
Single-molecule imaging	<ul style="list-style-type: none"> <li>• Book: <i>Single-Molecule Techniques: A Laboratory Manual</i>, Selvin and Ha, 2008</li> </ul>
Specimen preparation	<ul style="list-style-type: none"> <li>• Book: <i>Protein Localization by Fluorescence Microscopy: A Practical Approach</i>, Allan, 2000</li> <li>• Review: Shaner, 2014. Protocol for testing FPs in mammalian cells</li> <li>• Review: Rodriguez et al., 2017. Summary of recent developments and new directions in the FP field</li> <li>• Website: <a href="http://fpbase.org">http://fpbase.org</a>. For choosing and comparing FPs</li> </ul>
Image analysis	<ul style="list-style-type: none"> <li>• Review: Eliceiri et al., 2012. A general review of steps in image analysis; also includes options for free and open-source image analysis software</li> <li>• Website: <a href="https://image.sc">https://image.sc</a>. An online forum for questions and discussions on image analysis and open source software</li> <li>• eBook: <i>Analyzing Fluorescence Microscopy Images With ImageJ</i>, Bankhead, 2016</li> </ul>
Experimental design	<ul style="list-style-type: none"> <li>• Book: <i>Experimental Design for Laboratory Biologists: Maximizing Information and Improving Reproducibility</i>, Lazic, 2016</li> <li>• Book: <i>Experimental Design for Biologists</i>, Glass, 2014</li> </ul>

**Antibodies.** There are currently no standardized requirements for validation of specificity of commercial antibodies (Couchman, 2009; Uhlen et al., 2016). Antibody behavior varies in different contexts; an antibody that binds to denatured proteins on a Western blot may not bind to chemically fixed proteins in a biological sample (Willingham, 1999). Even if specificity has been validated by a company or another laboratory, antibody validation should be repeated in each specimen type. Despite its importance, antibody validation is not the norm (Freedman et al., 2016). For well-characterized targets, validation may be as simple as comparison to previous descriptions. When characterizing

localization of a novel epitope, the bar should be much higher. Multiple antibodies against an epitope can be compared for a similar localization pattern. Localization can also be compared with an FP conjugate (Schnell et al., 2012; Stadler et al., 2013). Gene-of-interest knockout and protein expression knock-down provide excellent specificity controls (Willingham, 1999; Burry, 2011).

Secondary antibody specificity must be validated with “secondary-only” controls, in which primary antibody is left out of the labeling protocol (Hibbs et al., 2006; Burry, 2011; Manning et al., 2012). Nonspecific secondary antibody binding may

appear as bright puncta or diffuse fluorescence and can localize to particular compartments or structures that can be misconstrued as biological findings. Secondary antibodies may exhibit nonspecific binding in some cell types but not others, and therefore must be controlled for in every cell type used. Secondary antibodies generated in the same species as the specimen (e.g., using an anti-mouse secondary antibody in a mouse tissue) are highly likely to exhibit nonspecific binding and require specialized techniques (Goodpaster and Randolph-Habecker, 2014). Nonspecific secondary antibody binding is almost always nonuniform, and therefore cannot be accurately “subtracted” computationally. Instead, if detected in secondary-only controls, nonspecific binding should be reduced by using different secondary antibodies or adjusting secondary antibody incubation conditions (concentration, duration, or temperature), blocking (type, temperature, or duration), and/or the number and length of washes (Allan, 2000).

**FPs.** In some cases, FPs have been shown to significantly perturb biological systems (Allison and Sattenstall, 2007; Landgraf et al., 2012; Swulius and Jensen, 2012; Han et al., 2015; Norris et al., 2015; Ganini et al., 2017; Sabuquillo et al., 2017). Fusing an FP to a protein-of-interest (POI) may alter or inhibit function of the POI or its ability to interact with binding partners (Shaner et al., 2005; Snapp, 2009). Expressing an FP-conjugated POI in addition to an endogenous POI may affect biological processes due to higher POI levels. In addition, FP conjugates must compete with endogenous protein for binding sites, etc., which can affect localization and behavior of FP conjugates or endogenous proteins (Han et al., 2015). FPs described as monomeric (often denoted with “m”) may retain affinity, especially if fused to a POI with the ability to oligomerize (Snapp, 2009; Landgraf et al., 2012). FP maturation and the fluorescence reaction both generate reactive oxygen species that can damage biological material (Remington, 2006; Ganini et al., 2017). Despite these potential problems, FPs are often used without adverse effects, but validation of FP conjugates is absolutely required.

The careful selection of an appropriate FP is essential (Snapp, 2009). Online resources like the Fluorescent Protein Database (<http://fpbase.org>) can be used with the primary literature to choose multiple FPs with specifications that match your application (Lambert, 2019). FPs exhibit environmental sensitivities and therefore must be tested in your experimental system (Shaner et al., 2005; Snapp, 2009; Ettinger and Wittmann, 2014; Costantini et al., 2015; Heppert et al., 2016). Comparing multiple options increases chances of identifying FPs that perform well in important parameters for your experiments (e.g., brightness, maturation time, Förster resonance energy transfer efficiency, etc.) and of detecting FP-related artifacts. Ideally, experiments should be performed to completion with multiple FPs, since FP-related artifacts may present at any step in an experimental protocol. Testing FPs before generating CRISPR or mouse lines, or other reagents that require significant time and resources, is particularly advisable.

**Self-labeling proteins (HALO, SNAP, and CLIP).** These tags can be a useful hybrid approach, combining the specificity of genetically encoded fusion proteins with the brightness and

photostability of organic dyes (Keppler et al., 2003; Los et al., 2008). However, all tags and ligands must be validated in your experimental system. Some SNAP ligands, for example, have been shown to exhibit significant nonspecific binding (Bosch et al., 2014). Similar to secondary antibodies, a control with a SNAP ligand in cells lacking any SNAP-tagged POI must be included to ensure the specificity of signal (Fig. 1 C).

**Fixed versus live.** Fixation and extraction/permeabilization artifacts can alter biological structures and localization patterns (Melan and Sluder, 1992; Halpern et al., 2015; Whelan and Bell, 2015). There is no one immunocytochemistry protocol that works well for every specimen and probe. Duration and concentration of multiple fixatives and extraction/permeabilization reagents should be tested for each primary antibody and cell type (Allan, 2000). Fixation can decrease or even eliminate fluorescence of FPs, so protocols should be validated by comparing intensity of FPs in live and fixed cells (Johnson and Straight, 2013). Comparing FP localization before and after fixation may also help identify fixation artifacts. When imaging live samples, phototoxicity can alter biological processes (Icha et al., 2017; Laissue et al., 2017). Live cell imaging controls with no fluorescence illumination should be performed to validate that the image acquisition protocol does not cause detectable cell damage (Magidson and Khodjakov, 2013; Ettinger and Wittmann, 2014).

### Imaging system validation

Once an “expected” biological structure becomes visible in the acquired image, it is common to assume that there is no need for further validation. However, accuracy of microscopy images cannot be assumed and must be validated. Routine imaging system maintenance and performance testing are recommended (Petrak and Waters, 2014), but even well-maintained imaging systems introduce errors into images, which may have profound effects on data interpretation (Fig. 1, A and B). To validate an imaging system, use known samples designed to reveal systematic (i.e., repeatable) errors and then correct the errors in images of biological samples. Validating that the imaging system can detect the target may also be necessary.

Imaging system validation generally consists of the following: (1) collecting images of known samples (Table 2); (2) identifying systematic errors present in images of known samples (Fig. 2); (3) correcting for errors, either computationally or by adjusting specimens, optics, or acquisition parameters (Figs. 2 and 3); and (4) testing correction methods (Figs. 2, 3, and 4). There is no one known sample or correction method that will reveal and correct all possible errors. Instead, researchers should assess whether a particular error may affect interpretation of their experimental results, and test error correction methods. It is critical to validate correction methods; performing inaccurate corrections can make matters even worse (Fig. 4). Errors vary from one microscope to the next, and when different optics or filters are used in the same microscope. Neglecting to correct systematic errors in images can therefore result in inaccurate and irreproducible data (Fig. 1, A and B). Common sources of error in microscopy images are discussed thoroughly elsewhere (Stelzer, 1998; Hibbs et al., 2006; North, 2006; Zucker, 2006; Waters, 2009; Wolf



Table 2. Useful known samples for imaging system validation

Sample	Error	References for correction protocols
(a) Fluorescent microspheres (beads) that are below the diffraction resolution limit of the imaging system	Optical aberrations	Hiraoka et al., 1990; Goodwin, 2013
(b and c) Multi-wavelength beads below the diffraction resolution limit of the imaging system	Channel registration	Hibbs et al., 2006; Spiering et al., 2013; Wolf et al., 2013
(d) Stage micrometer	In magnification/pixel size	Wolf et al., 2013
Flatfield slide	Nonuniform illumination	Model, 2006
Single-labeled biological sample	Bleed-through	Spiering et al., 2013 Bolte and Cordelières, 2006
Stable biological sample	Photobleaching	Bancaud et al., 2010
Unlabeled biological sample	Autofluorescence	Hibbs et al., 2006

We used (a) Molecular Probes FluoSpheres; (b) for high-resolution imaging, Invitrogen TetraSpeck Microspheres, 0.1  $\mu\text{m}$ ; (c) for low-resolution imaging, Invitrogen FocalCheck Beads, 6  $\mu\text{m}$  or 15  $\mu\text{m}$ ; and (d) MicroScope World, 25 mm KR812.

et al., 2013). Here, we discuss examples of validation of detection, illumination, and optics important to consider in many experiments.

**Detection.** Microscopes vary widely in efficiency of photon detection (Swedlow et al., 2002; Pawley, 2006a; Murray et al., 2007). Any given imaging system requires a minimum level of fluorescence intensity in order for the fluorescence to be detectable, and lack of detection may be misinterpreted as absence of fluorophore. Detection of autofluorescence in unlabeled controls can be used to establish optics and acquisition parameters capable of detecting low fluorescence levels and to validate results in which no fluorophore is detected.

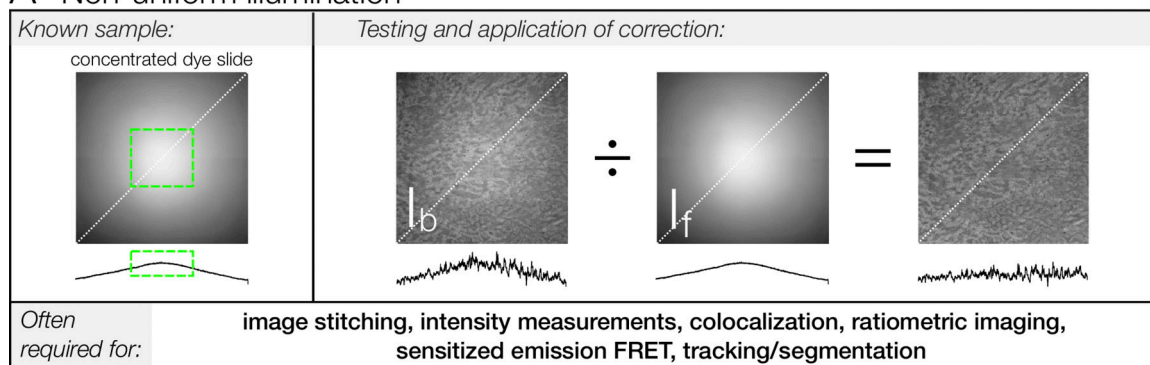
The detection limit is in part determined by the measurement error of the detector, often referred to as detector noise (e.g., read noise; Lambert and Waters, 2014). Noise causes fluctuations in intensity both above and below the ground truth value, meaning that it cannot be subtracted from an image. The magnitude of these fluctuations limits both the minimum intensity and the smallest change in intensity that can be detected. Detector noise properties vary between detector types (e.g., charge-coupled device cameras, sCMOS cameras, and photo-multiplier tubes; Lambert and Waters, 2014; Stuurman and Vale, 2016). Most detectors also add a constant “offset” to the intensity values in the image to prevent data clipping at low intensities. The offset can be measured by acquiring an image with no light directed to the detector and must be subtracted from all intensity measurements.

**Nonuniform illumination.** Fluorescence intensity depends on illumination intensity as well as the fluorophore distribution and inherent brightness, and no microscope has perfectly uniform illumination across the field of view. We therefore cannot assume that differences in fluorescence intensity across images are due to changes in fluorophore without measuring the illumination distribution. The effect of nonuniform illumination may not be obvious by qualitative assessment but can result in inaccurate quantitative measurements of intensity (Fig. 2 A) and make accurate image segmentation (e.g., thresholding) difficult.

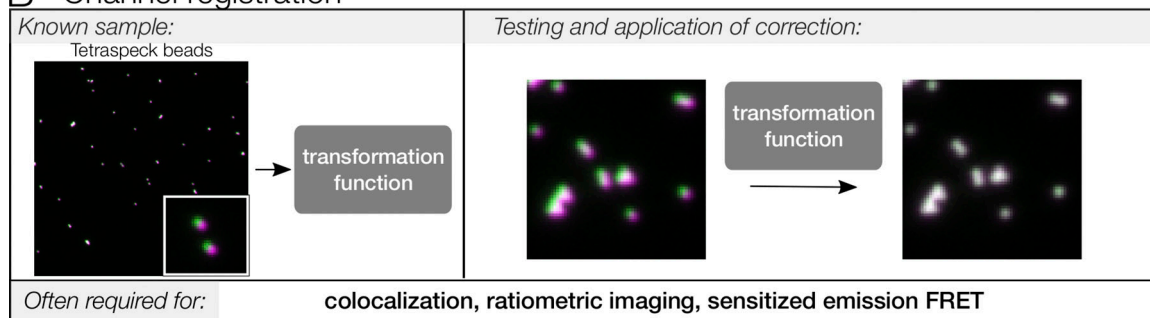
To reveal the illumination pattern, uniformly fluorescent samples are used to generate images in which intensity variations across the field of view are due to illumination variation, not variation in fluorophore concentration (Fig. 2 A and Table 2; Model, 2006; Zucker, 2006; Wolf et al., 2013). Model and Burkhardt (2001) introduced a clever, inexpensive, and easily prepared sample for measuring illumination uniformity that we find performs more consistently and accurately for quantitative applications than other commonly used samples (e.g., dye-infused plastic slides). Images of uniform samples are referred to as “flatfield images” ( $I_f$ ) and are used to correct images of biological specimens ( $I_b$ ) computationally (Fig. 2 A; Model, 2006) or to identify image subregions with minimal illumination variation, to which acquisition or analysis can be restricted. Computational corrections of  $I_b$  are performed using image arithmetic. Image intensity is a product of illumination intensity and fluorophore concentration, so dividing  $I_b$  by  $I_f$  on a pixel-by-pixel basis reveals fluorophore intensity distribution (Bolte and Cordelières, 2006; Model, 2006; Hodgson et al., 2010; Spiering et al., 2013). The flatfield image must be validated; correction with an inaccurate flatfield image can increase error in the corrected image (Fig. 4 C).

**Channel (mis)registration.** Images of different wavelengths (e.g., fluorescence channels) collected from the same sample will not align perfectly in XY or Z, due to differences between fluorescence filters (Waters, 2009) and/or chromatic aberration in lenses (Keller, 2006; Ross et al., 2014). Multicamera and image-splitting systems can introduce additional XY and Z shifts. Channel misregistration due to the imaging system will be repeatable in images acquired with a given set of optics (e.g., combination of filter sets and objective lens) but will vary when optics are changed. A known sample can be used to measure and correct channel misregistration (Fig. 2 B and Table 2; Hodgson et al., 2010; Wolf et al., 2013). For a given imaging system, the extent of the shift may be negligible for some experiments; channel misregistration much smaller than distances being measured may not affect results. When measuring colocalization

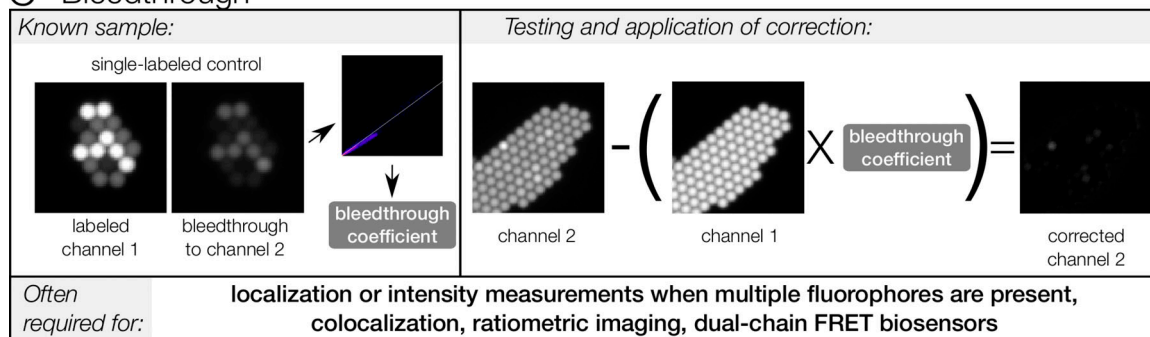
## A Non-uniform illumination



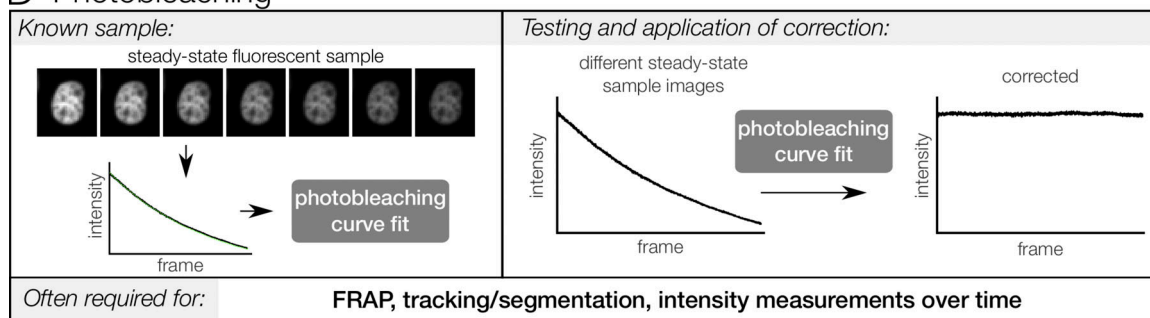
## B Channel registration



## C Bleedthrough



## D Photobleaching



**Figure 2. Measurement and computational correction of image errors.** Known samples are used to measure systematic errors in microscopy images. From the measurement, a correction can be generated, tested, and applied to experimental images. Correction procedures are summarized here and some steps (e.g., background subtraction) have been omitted. Please refer to the main text for references that cover these corrections in more detail. **(A)** Illumination nonuniformity. Concentrated dye is mounted between a coverslip and a slide and sealed. This dye, if sufficiently concentrated, acts as a thin, uniformly fluorescent sample (see Model and Burkhardt, 2001; Model, 2006). This “flat-field image” can be used to determine a region with minimal illumination variation (green box) or can be used to correct experimental images. The correction is tested by applying to a biological sample of roughly uniform intensity across the field of view, here a kidney section labeled with AlexaFluor568 phalloidin. Line scans below each image show intensity along the indicated white dotted line. **(B)** Channel registration. Tetraspeck beads are infused with four fluorescent dyes, including the green and red dyes imaged here (pseudo-colored green and magenta, respectively). Because the images of the beads in each channel should overlay perfectly, they can be used to generate a transformation matrix that

describes the transformation needed to align the images. This matrix is then tested by using it to correct a different image of Tetraspeck beads. Once tested, the matrix can be used to register channels of experimental images. **(C) Bleed-through.** Samples labeled with a single fluorophore are used to measure bleed-through by imaging all channels with the same settings used for acquisition in the experiment. Here, 2.5- $\mu\text{m}$  beads labeled with a dye corresponding to channel 1 are used. The intensity of bleed-through into channel 2 is plotted as a function of intensity of channel 1, and a linear regression of this plot is used to generate a bleed-through coefficient. This coefficient is then tested by applying to a different single-labeled control image and verifying that bleed-through into channel 2 is reduced. Once tested, the bleed-through coefficient can be used to correct for bleed-through in experimental images (provided channels are properly registered, as described above). **(D) Photobleaching.** Samples with steady-state fluorescence are used to generate a photobleaching curve under the planned experimental conditions. This curve is fit to an exponential function, which is then tested by correcting a different set of images of the steady-state sample. Once tested, the correction can be applied to experimental images under similar conditions; that is, if the correction is to be used across multiple days or sessions, it should be validated on images collected on multiple days. FRET, Förster resonance energy transfer.

between objects below the resolution limit, however, even very small shifts may result in inaccurate data (Fig. 1 B).

Channel registration is measured using samples in which the distribution of fluorophore in each channel is known to be identical, such as beads infused with multiple fluorescent dyes (Fig. 2 B and Table 2). Lateral channel registration can be corrected computationally (Fig. 2 B; Hodgson et al., 2010; Wolf et al., 2013). Axial (Z) channel misregistration can be corrected for computationally when a 3D z-series of images is collected or can be compensated for by using a motorized focus motor to move the distance of the measured shift between acquisition of channels.

**Bleed-through.** While it is commonly appreciated that signals from one fluorophore may bleed through filters designed to image another fluorophore, the extent to which this may occur even for commonly used fluorophores is underappreciated (Bolte and Cordelières, 2006; Hodgson et al., 2010; Spiering et al., 2013). Testing and correcting for bleed-through are especially important for colocalization experiments, where it can lead to false positives (Fig. 1 A). Bleed-through is measured using “single-labeled controls”: samples labeled with only one fluorophore (Fig. 2 C). To identify bleed-through, single-labeled controls should be imaged with each fluorescence channel used for the multi-labeled samples, and with the same acquisition settings used for the multi-labeled samples.

Bleed-through is defined as the percentage of fluorophore A signal collected with a fluorophore B filter set. Therefore, the intensity of bleed-through depends on the intensity of fluorophore A (Fig. 3, A and C). Since the intensity of fluorophore A may vary between cells or structures, the detection and intensity of bleed-through can vary throughout the field of view as well.

With a single-labeled control sample, the percentage of fluorophore A bleed-through can be estimated as the slope of a linear regression through a plot of intensity of fluorophore B as a function of intensity of fluorophore A (e.g., using Coloc2 in Fiji; Schindelin et al., 2012; see Fig. 2 C for an example plot). This measured bleed-through coefficient must be validated by performing a correction on a different single-labeled control image. If the coefficient is inaccurate, bleed-through correction may introduce error into the corrected image (Fig. 4 D).

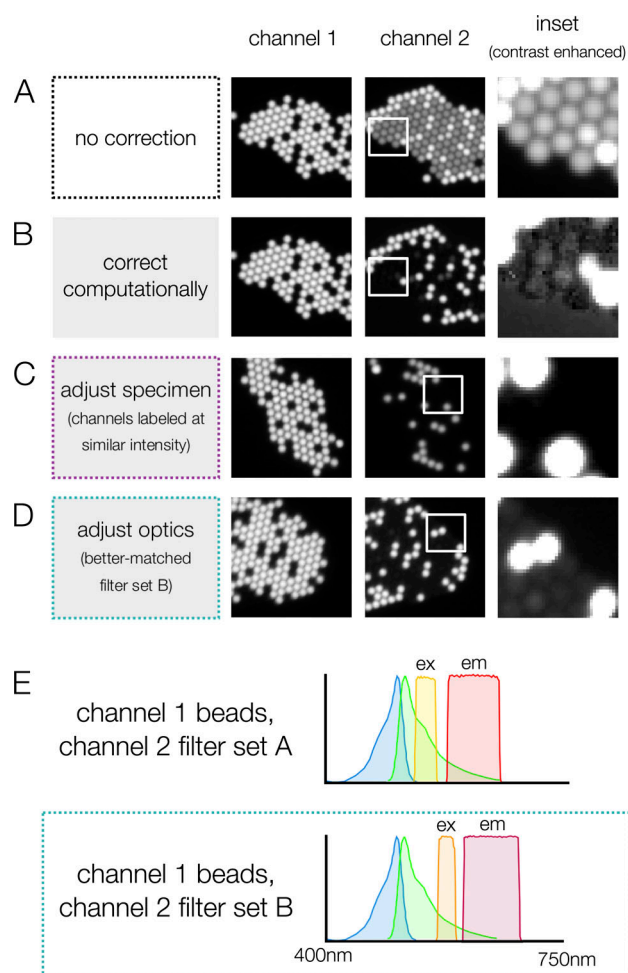
Bleed-through can be corrected for computationally on a pixel-by-pixel basis (Figs. 2 C and 3 B). If a single-labeled control reveals that 10% of the fluorophore A signal bleeds through a fluorophore B filter set, then 10% of the intensity measured in the fluorophore A image can be subtracted from the fluorophore B signal (Fig. 2 C). But if bleed-through is significant, switching

to fluorophores with farther spectral separation or using more spectrally restrictive filter sets may result in more accurate data (Fig. 3 D).

**Photobleaching.** Fluorescent samples photobleach when illuminated. The photobleaching rate varies between different fluorophores and with the local environment of the fluorophore, turnover rate, and intensity of illumination (Diaspro et al., 2006). In quantitative microscopy experiments, acquisition parameters should be carefully adjusted to minimize illumination and therefore photobleaching. Even when photobleaching has been minimized, measuring and correcting for changes in intensity due to photobleaching may be required when making quantitative intensity measurements over time. In addition to photobleaching during acquisition, photobleaching can occur prior to acquiring images, when focusing or selecting a field of view, and may result in fluorescence intensity measurements that do not accurately report the amount of labeled component. To minimize this error, use transmitted light or acquire single fluorescence images when focusing and selecting fields of view.

A photobleaching control sample that contains fluorophore at steady-state can be used to measure and correct for photobleaching that occurs during a time-lapse experiment. The fluorophore intensity in this control must be constant so that any changes in intensity that occur during acquisition can be assumed to be due to photobleaching. Intensity values in the control are measured over time, and the resulting photobleaching curve fit to a single or double exponential (Vicente et al., 2007), which can then be used for correction (Fig. 2 D). The photobleaching control must be as similar as possible to the experimental sample (i.e., the same FP-POI conjugate expressed in the same cell type but treated with a drug that induces steady-state) and identical acquisition conditions must be used for imaging the control and experimental samples. An ideal experimental set-up is to use a multi-well plate and a motorized stage for image acquisition, so images of the photobleaching control and experimental sample can be collected during the same time lapse. Alternatively, a region within an image that should not change in intensity over the course of the experiment can be used as a photobleaching control. This type of correction can be performed with the ImageJ Bleach Correction plugin with the Exponential Fit option selected (Miura and Rietdorf, 2014). In FRAP experiments, intensity normalization using a region of interest outside the bleached region can be used for photobleaching correction (Phair et al., 2003; Bancaud et al., 2010).

Many variables affect photobleaching rate, including the local environment of the fluorophore, turnover rate, position within



**Figure 3. Image errors can be corrected in multiple ways.** (A) Without correction, there is significant bleed-through from channel 1 into channel 2 (dimmer spots in channel 2 image). (B) Bleed-through can be corrected computationally (Fig. 2 C), but the correction can lead to artifacts that skew intensity measurements (see contrast-enhanced inset). Bleed-through can also be reduced by adjusting the specimen (C) or adjusting optics (D) in the microscope. Whether or not bleed-through is a problem for a particular experiment depends on the relative intensity of the fluorophores. In A, the beads in channel 1 are >300× brighter than the beads in channel 2; in C, beads of similar intensity are used, and bleed-through is no longer detectable. In D, a spectrally shifted filter set (E) is used to reduce bleed-through. At a glance, neither of these filter sets appears to have significant overlap with the excitation spectrum of the dye, but the small amount of overlap is exacerbated by the large difference in intensity between the channels.

the field of view, and day-to-day variability in illumination intensity (Diaspro et al., 2006), making accurate and precise correction for measurement of photobleaching challenging and validation of correction methods critical. The correction can be validated by applying to a separate time lapse of the steady-state fluorescent sample (Fig. 2 D) but should also be validated by correcting samples with well-characterized or known intensity changes. Photobleaching correction using an inaccurate photobleaching curve will introduce error into intensity measurements (Fig. 4 E). When performing a photobleaching correction, consider how an inaccurate estimate of photobleaching might alter the results. For example, photobleaching correction of a

dataset intended to measure the degradation rate of a fluorescently labeled POI must be more accurate and precise than photobleaching correction intended only to improve the quality of automated segmentation over the course of a time-lapse.

**Measurement validation.** In addition to identification of and correction for error, quantitative microscopy requires validation that the labeling method and imaging system can be used to make measurements with the accuracy and precision required to test the experimental hypothesis. Measurement validation is usually performed using known samples, manipulation of the biological system (e.g., drug treatments, mutants), and/or positive and negative controls. See Fig. 5 and Dorn et al. (2005), Joglekar et al. (2008), and Wu and Pollard (2005) for examples of measurement validation.

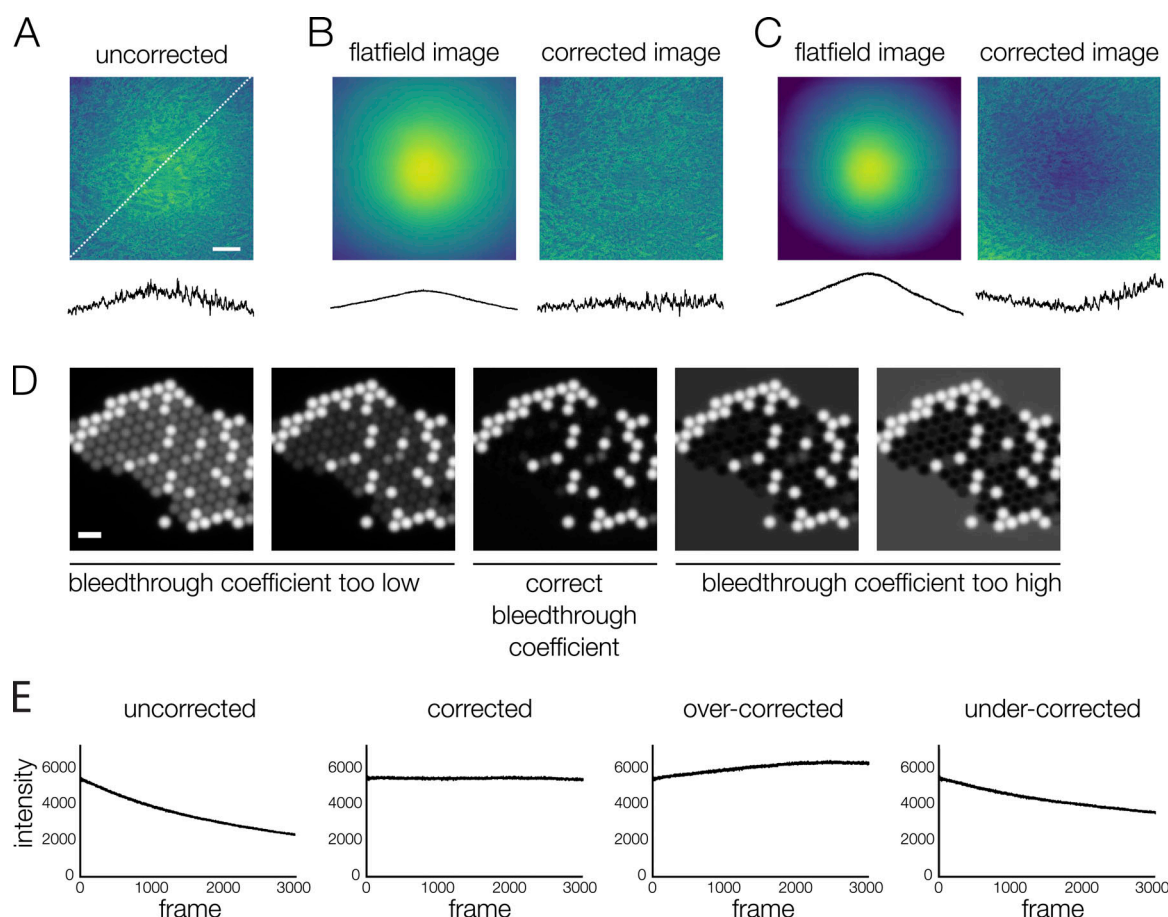
## Bias, blinding, and randomization

### Experimenter bias

Humans are prone to many types of unconscious bias, and scientists are no exception (Nickerson, 1998; Ioannidis, 2014; Holman et al., 2015; Nuzzo, 2015; Lazic, 2016; Munafò et al., 2017). Any subjective decision made in an experiment is an opportunity for bias. Bias is an even larger problem in experiments that generate images; although we trust our eyes to give us accurate information about the world around us, human visual perception is biased toward detection of certain types of features and is not quantitative (Russ, 2004). Qualitative visual assessment of images is subjective and therefore prone to many types of bias, both perceptual and cognitive (Table 3). In particular, conclusions can be influenced by apophenia (the tendency to see patterns in randomness) and confirmation bias (the increased likelihood of seeing a result that fits our current understanding than one that does not; Lazic, 2016; Munafò et al., 2017). Fig. 6 presents an example of images that could easily be misinterpreted due to confirmation bias. Fortunately, bias can be avoided through good experimental design. We present two key methods that are useful in microscopy experiments: blinding and automation.

When thinking about how best to avoid bias in an experiment, it is important to consider whether the experiment is intended for hypothesis generation (“learning”) or hypothesis testing (“confirming”). The two types of experiments have different design requirements, particularly regarding methods to avoid bias (Sheiner, 1997; Lazic, 2016). Hypothesis-generating experiments are about discovery and exploration, and do not necessarily require formal methods to avoid bias, but once a hypothesis is generated it must be tested with an experiment that does include these methods. The same data cannot be used to both generate and support a hypothesis (Kerr, 1998; Lazic, 2016). Humans often detect patterns even in truly random data (Lazic, 2016), leading to hypotheses that may be supported by the original dataset but will not hold up when tested independently. When tested with an independent experiment, true phenomena should continue to occur while overinterpreted patterns will not. Of course, hypotheses may arise from any dataset, including a hypothesis-testing experiment; this is acceptable as long as they are presented clearly as post hoc hypotheses and are subsequently tested independently (Kerr,





**Figure 4. Image corrections must be tested carefully. (A–C)** Flatfield correction. **(B)** When the flatfield image truly represents the illumination distribution, the uniformity of the test image (kidney section labeled with AlexaFluor568 WGA) is improved (see line scans below images, measured at the location indicated by the dotted white line in A). **(C)** When the correction is performed with a flatfield image that does not represent the illumination distribution, or has been normalized incorrectly, the test image is less uniform after correction. Correcting with an inaccurate flatfield image can add error to quantitative intensity measurements. If the flatfield image does not perform well in tests, a better solution is to define a subregion with less variable intensity (see Fig. 2 A). **(D)** Bleed-through correction. If the estimated bleed-through coefficient is inaccurate, bleed-through correction can lead to artifacts in the image that will add error to quantitative intensity measurements. Because these images contain no overlap between channels (mixed beads as in previous bleed-through figures, channel 2 shown), incorrect bleed-through coefficients show obvious artifacts; artifacts will be less obvious in experimental images with some overlap in signal. The bleed-through coefficient should be tested on single-labeled sample images before applying to experimental images. **(E)** Photobleaching correction. The sample in this example is fixed, meaning variation in intensity is due only to photobleaching and detector noise. If the rate of photobleaching is correctly measured, the corrected intensity values remain constant over time. If the rate of photobleaching is over- or underestimated, the corrected intensity values are no longer constant. Inaccurate corrections are obvious when applied to a steady-state sample, but over- or undercorrection may be impossible to detect when applied to a signal that varies over time. Scale bars: (A) 100  $\mu\text{m}$ , (D) 5  $\mu\text{m}$ .

1998). In hypothesis-testing microscopy experiments, every effort must be made to avoid bias through one or more of the methods discussed below.

### Strategies to reduce bias

**Blinding.** Blinding refers to methods that hide the identity of the sample or image from researchers and is a required standard in some fields (MacCoun and Perlmutter, 2015). Despite evidence that blinding reduces the effect of bias (Holman et al., 2015), blinding is rarely reported in microscopy experiments in cell biology.

A common justification to not blind microscopy experiments is that the difference between images from different conditions is obvious, so even relabeled samples will be easy to tell apart. However, the nature of bias suggests that once labels are

removed the difference between conditions may be less obvious. Additional treatments or controls can also be added to datasets to increase variety and make differences between conditions less obvious. Another argument is that blinding may hamper discovery of new phenomena (MacCoun and Perlmutter, 2015). The distinction between hypothesis-generating and hypothesis-testing experiments is important here: the goal of a hypothesis-testing experiment is not discovery. Other interesting phenomena may become apparent before blinding or after results are unblinded, resulting in the generation of new hypotheses that can then be independently tested.

Blinding during image analysis can be addressed multiple ways. It may be as simple as renaming files. Researchers should avoid using experimental images for manual selection of regions/pixels for analysis (i.e., segmentation). Instead, use a

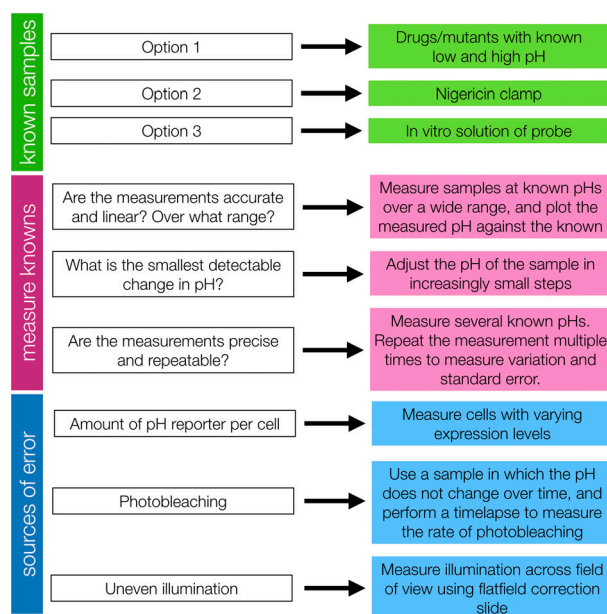


Figure 5. **Measurement validation example: using a fluorescent biosensor to measure subcellular pH.** To validate measurements, known samples (green) are required. These knowns can be used to characterize the dynamic range, linearity, and repeatability of measurements (magenta) and sources of error in the measurements (blue). For more information about pH measurements, see Grillo-Hill et al. (2014) and O'Connor and Silver (2013).

fluorescence channel or transmitted light image from which content of the experimental image cannot be deduced. For example, if an experiment involves measuring changes in localization of a fluorescently labeled POI to an organelle, organelles should be directly labeled with a fluorescent marker that can be used to segment organelles without looking at the channel containing the POI.

**Automation.** Another approach to reducing bias is automation of steps in the acquisition or analysis workflow that could be subject to bias (Lee et al., 2018). Automation reduces bias because it requires researchers to define selection parameters before acquisition or analysis, which are then applied equally to all samples/images.

When acquiring images, researchers often search samples and manually select fields of views to acquire. This can be helpful in reducing acquisition time and data size in some cases—for example, when identifying and restricting acquisition to particular cell types in tissue sections. However, in many cases, visually choosing fields of view to image (and not to image) during acquisition can introduce bias that, if suspected later, cannot be recovered from without repeating experiments. We therefore recommend using microscope automation to acquire images. Microscope acquisition software packages offer tools to randomly select fields of view or to tile an entire coverslip or well. When using a microscope that lacks the necessary motorized components, bias can be reduced by following a strict predetermined selection protocol, just as the computer would in an automated experiment. For example, choose fields of view independently of image content by first aligning the objective lens roughly to the center of the coverslip, then moving a set

number of fields of view before acquisition of each subsequent image. This method works only if the protocol is strictly followed for each experimental condition and repetition; blinding by relabeling and randomizing samples/conditions will aid in ensuring consistency. Criteria to skip a field of view or discard data should be chosen and validated prior to acquiring the dataset.

Manual selection of image features to include in analysis can also introduce bias. Instead of qualitatively choosing features from each image to include in analysis, researchers should define selection criteria that can be reported in methods sections. As with acquisition, selection may be automated or performed manually according to strict criteria. This may be as simple, for example, as including only those pixels with intensity values two standard deviations above background. More often, however, defining sufficient criteria for automated selection of image features is challenging. Machine learning approaches can be very useful for generating selection criteria and are available in free and open-source packages (Caicedo et al., 2017; Kan, 2017; <http://ilastik.org/> [Sommer et al., 2011]; [https://imagej.net/Trainable\\_Weka\\_Segmentation](https://imagej.net/Trainable_Weka_Segmentation) [Arganda-Carreras et al., 2017]). When determining the parameters that will be used for an automated analysis routine, you should not use the exact datasets that you ultimately want to analyze. Regardless of the method used to generate selection criteria, the criteria should be validated. Examples of methods of validating image analysis selection criteria include use of known datasets or comparing the results of multiple researchers tasked with applying selection criteria to the same set of images or of manual versus automated selection of features. Neither the image analysis software nor researchers will select features with 100% accuracy, but automation both reduces bias and allows analysis of larger datasets useful for generating meaningful statistics.

Using automated acquisition or analysis tools to avoid bias does not preclude visual inspection of the data for quality control. Regular checks are necessary to ensure that automated processes are performing as expected, and in most cases require a researcher to interact with the images or data. For example, visual inspection of automated segmentation results is useful to make sure the pipeline works for all the data, not just the data used to design and test the pipeline. However, to avoid bias, quality-control check methods and criteria for adjusting parameters or discarding data should be planned before analysis.

Automation of image acquisition and analysis also aids researchers in reproducing experiments. Clearly defined automated selection procedures can be implemented as new researchers join a laboratory and can be clearly reported in a publication's methods section; the same cannot be said of subjective, unvalidated manual choice of features by a single researcher.

## Methods reporting

For microscopy experiments to be reproducible, microscope hardware and software configurations must be reported accurately and completely. Some journals provide detailed lists of information to include (the *Journal of Cell Biology* included), but many lack specific instructions. In our reading of the literature,

Table 3. Bias in imaging experiments

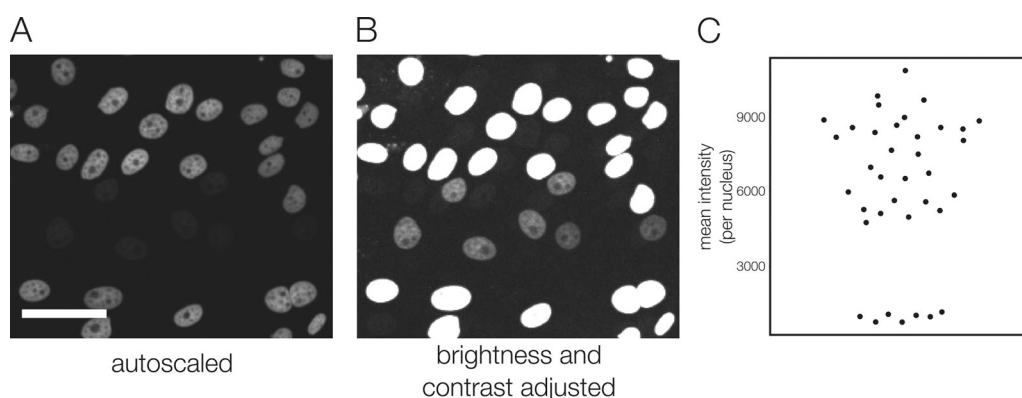
Type of bias	Examples in imaging experiments	Strategies
Selection bias	• Scanning samples for fields of view that “look good” or “worked” based on subjective or undefined criteria (also confirmation bias)	• Use microscope automation to select fields of view or scan the entire well
	• Choosing to image only the brightest cells/samples (e.g., highest expression level)	• Include all data in analysis, or determine criteria to discard a dataset before collecting data
	• Only including data from experiments that “worked” in analysis or publication	
Confirmation bias	• Adjustments to the analysis strategy based on the direction the results are heading	• Validate the analysis strategy using known samples/controls ahead of time
	• Choosing analysis parameters that yield the desired or expected results, rather than choosing through validation with known samples	• Perform analysis blind
	• P-hacking (Head et al., 2015)	
	• Choosing cells or parts of a sample that “make sense” based on the anticipated outcome	
Observer bias/ experimenter effects	• Spending more time focusing by eye (and therefore photobleaching) on one condition than the others	• Perform acquisition and analysis blind
	• Making subjective conclusions based on visual inspection of the image rather than making quantitative measurements	• Make conclusions based on quantitative measurements rather than qualitative visual impressions (measure length/width/aspect ratio, count, measure intensity, etc.)
Asymmetric attention bias/ disconfirmation bias	• Performing image corrections only when result seems wrong or is not as expected	• Consider sources of error, validate, and apply corrections equally to all conditions and experiments

See Lazic (2016), Nuzzo (2015), Nickerson (1998), and Munafò et al. (2017) for more about bias and additional references.

few publications include sufficient information to replicate microscopy experiments. For a checklist of items to include in a microscopy methods section, refer to Lee et al. (2018).

During acquisition and analysis, keep a detailed record of all protocols, hardware configuration, and software versions and settings used. Some acquisition information is recorded in image metadata, but beware: metadata may be inaccurate if the

configuration of the microscope was altered without changing the software. Whenever possible, use image management tools that combine image storage and annotation in the same place (Allan et al., 2012). Minimally, make a data management plan before you start a project so you can easily track down raw data, acquisition settings, image processing steps, etc., when putting together a publication (Lee et al., 2018). Reproducibility of



**Figure 6. Visual inspection of images is prone to confirmation bias. (A and B)** In this example, cells labeled with a fluorescent nuclear marker exist in two populations, one with very bright nuclear labeling and the other with much dimmer labeling. If the image is autoscaled (A), the dimmer population is invisible, but brightness and contrast adjustments show that there is also a population of cells with lower intensity labeling (B). Making conclusions based on images displayed using autoscale (the most common default display in image acquisition programs), rather than measuring image intensity values, could lead to inaccurate conclusions. A researcher who is convinced by the image display because it represents the expected result, and therefore makes the decision not to complete a full quantitative analysis, is subject to confirmation bias. Scale bar: 50  $\mu$ m. **(C)** Measured intensity of the nuclei in the images. Each dot represents the mean intensity of one nucleus.

microscopy experiments will almost certainly be enhanced through free and open sharing of detailed imaging protocols (through platforms such as <https://bio-protocol.org> and <https://www.protocols.io>) and primary data (through the Image Data Resource [Williams et al., 2017] and open data formats [Linkert et al., 2010]).

Even microscopy methods sections that include detailed information about acquisition and analysis often neglect to describe method validation. We encourage researchers to report any validation steps they have taken in the development of quantitative microscopy assays. It is a common misconception that images can be taken at face value, and this is exacerbated by lack of reporting of validation steps. Quantitative microscopy papers often read as simple experiments, because the complex validation steps are not described (or perhaps were not performed). This not only reinforces misconceptions that microscopy is simple and “seeing is believing,” it makes it impossible to assess quality of results.

## Conclusion

The importance of quantitative imaging methods in cell biology research is becoming widely recognized. Quantification has many benefits over qualitative presentation of a handful of images from a dataset, including strict definition of selection criteria and ease of inclusion of larger datasets and statistical tests. However, in our experience, researchers often underestimate the difficulty of rigorous quantitative microscopy experiments. When advising researchers on microscopy experimental design, we routinely suggest procedures for validation of methods and control experiments. We sometimes get the response that a particular control or validation is not possible because, for example, required samples cannot be prepared. It is important to understand that omitting method validation and control experiments restricts accurate conclusions that can be made from an experiment, often to an extent that makes the experiment not worth the time and effort. Early and thorough planning, including investigating the necessary known and control samples before experimental samples are prepared, can help to circumvent this situation. Whenever possible, it is advisable to consult or seek collaboration with expert microscopists and image analysts; quantitative imaging experiments can be difficult to design and validate, and even for experts often require multiple rounds of testing and optimization to achieve confidence in the reproducibility of a protocol. Unfortunately, researchers cannot assume that an imaging protocol described in a published methods section will contain all validations and controls required to interpret results, so it is highly advisable to consult primary literature on methods, reviews by experts, and educational resources (Table 1) when adapting a published protocol for use in your own experiments.

## Materials and methods

### Fig. 1 methods

(A) Images of 2.5- $\mu$ m Inspeck beads (100% intensity green, 0.3% intensity red; ThermoFisher) mounted in glycerol were collected with a Nikon Ti2 microscope, using a Plan-Apochromat 20 $\times$  0.75-NA objective lens (Nikon) and Orca-Flash 4.0 LT camera

(Hamamatsu Photonics) controlled by NIS Elements (Nikon). Green channel images were acquired with a Semrock filter set (466/40 excitation, 495 dichroic, and 525/50 emission). Red channel images were collected with a Chroma Cy3 filter set (see Fig. 3 E, filter set A, for spectra). Green circles denote selected green beads and magenta circles denote selected red beads. Images are contrast stretched (brightness and contrast “reset”) using Fiji (Schindelin et al., 2012). (B) 0.2- $\mu$ m Tetraspeck beads were imaged on a DeltaVision OMX Blaze microscope (GE) using a Plan Apo 60 $\times$  1.42-NA objective lens (Olympus) and the FITC and mCherry presets (488 laser, 528/58 emission filter, and 568 laser, 609/37 emission filter). Channels were imaged on separate cameras (pco.edge; PCO). Images were pseudo-colored, overlaid, and brightness and contrast adjusted in Fiji (Schindelin et al., 2012). (C) HeLa cells were either fixed and immunofluorescence labeled against POI or imaged live after treatment with SNAP-Surface AlexaFluor647 (New England BioLabs) and washing. POI-SNAP cells contain a POI-SNAP tag fusion; WT cells contain no SNAPtag. All cells were imaged on a Nikon Ti TIRF microscope with an Apo TIRF 100 $\times$  1.49-NA oil objective lens (Nikon) and an Imagem EMCCD camera (Hamamatsu Photonics) controlled by Metamorph (Molecular Devices). Acquisition settings for SNAP dye conditions are identical, but not for immunofluorescence images due to different dyes used. Gamma 0.5 was applied to images in C using Fiji.

### Fig. 2 methods

(A) A concentrated fluorescein slide was prepared according to Model (2006). Images were collected with a Nikon Ti2 microscope equipped with a Plan Fluor 10 $\times$  0.3-NA Ph1 DLL objective lens (Nikon) and ORCA Flash 4.0 LT camera (Hamamatsu Photonics) controlled by NIS Elements (Nikon). FluoCells prepared slide #3 (ThermoFisher) was used for test images. Fiji was used to generate line scans, subtract camera offset, perform correction (Image Calculator function), and adjust brightness and contrast. (B) 0.2- $\mu$ m Tetraspeck beads were imaged on a DeltaVision OMX Blaze microscope (GE) as described in Fig. 1 B. Images were pseudo-colored and overlaid, and brightness and contrast was adjusted in Fiji (Schindelin et al., 2012). Transformation matrix was generated and applied using the *imregtform* and *imwarp* functions in MATLAB (Image Processing Toolbox). (C) 2.5- $\mu$ m Inspeck beads (100% intensity green, 0.3% intensity red; ThermoFisher) were mounted in glycerol and imaged on the same setup as Fig. 1 A. The bleed-through coefficient was estimated using slope of the correlation regression line generated by the Coloc2 plugin in Fiji. The correction was performed using Fiji (Math > Multiply and Image Calculator functions). (D) LLC-PK1 cells expressing H2B-mCherry were plated on a #1.5 coverslip-bottom 35-mm dish (MatTek), fixed with formaldehyde, and imaged on a Nikon Ti-Eclipse microscope equipped with a Yokogawa CSU-X1 spinning disk confocal head, controlled by Metamorph (Molecular Devices). Images were acquired with a Hamamatsu Flash 4.0 V3 sCMOS camera using a Plan Apo 20 $\times$  0.75-NA objective lens. A 561-nm laser was used for illumination. Emission was selected with a 620/60-nm filter (Chroma). Brightness and contrast were adjusted in Fiji (Schindelin et al., 2012). Intensity of a manually selected ROI within one nucleus



was measured in Fiji and plotted in Microsoft Excel. Single exponential fit with offset was performed using the curve fitting tool in Fiji. Correction was performed in Microsoft Excel by dividing the measured intensity value at each time point by the value of the normalized fitted exponential function.

### Fig. 3 methods

Images were collected with the same setup as in Fig. 1 A, with the exception of panel D, which used filter set B illustrated in the spectrum (545/30 excitation, 570 dichroic, and 610/75 emission; Chroma). Exposure times are the same for all channel 2 images. A longer exposure time was used for channel 1 in panel C, because the beads used were much dimmer. Computational bleed-through correction in B was performed as described in Fig. 2. Images were adjusted for brightness and contrast in Fiji. Spectra were downloaded from ThermoFisher (beads) and FPbase.org (filters), and then plotted in Excel.

### Fig. 4 methods

(A–C) Images were collected and corrected as described in Fig. 2 A. Images are displayed using the mpl-iridis LUT (Fiji). (D) Images were collected and corrected as described in Fig. 2 C. (E) Images were collected and corrected as described in Fig. 2 D. Over- and undercorrected plots were generated using exponential functions that were altered to reflect a higher or lower photobleaching rate.

### Fig. 6 methods

Images were collected as described in Fig. 2 D. Nuclei were segmented in Fiji by thresholding, binary morphological operations and the Analyze Particles tool. Mean intensity measurements per nucleus were plotted using PlotsOfData (<https://huygens.science.uva.nl/PlotsOfData/>; Postma and Goedhart, 2018 Preprint).

## Acknowledgments

The authors thank Talley Lambert for helpful discussions and comments and Ruth Brignall for sharing her SNAP-tag control images. Additional thanks to the Departments of Cell Biology, Systems Biology, and Biological Chemistry and Molecular Pharmacology at Harvard Medical School for support.

The authors declare no competing financial interests.

Author Contributions: Both authors contributed to manuscript writing and editing. A.P.-T. Jost performed experiments and created the figures.

Submitted: 19 December 2018

Revised: 26 February 2019

Accepted: 27 February 2019

## References

Allan, V.J., editor. 2000. *Protein localization by fluorescent microscopy: a practical approach*. Oxford University Press, Oxford.

Allan, C., J.M. Burel, J. Moore, C. Blackburn, M. Linkert, S. Loynton, D. Macdonald, W.J. Moore, C. Neves, A. Patterson, et al. 2012. OMERO: flexible, model-driven data management for experimental biology. *Nat. Methods*. 9:245–253. <https://doi.org/10.1038/nmeth.1896>

Allison, D.G., and M.A. Sattenstall. 2007. The influence of green fluorescent protein incorporation on bacterial physiology: a note of caution. *J. Appl. Microbiol.* 103:318–324. <https://doi.org/10.1111/j.1365-2672.2006.03243.x>

Arganda-Carreras, I., V. Kaynig, C. Rueden, K.W. Eliceiri, J. Schindelin, A. Cardona, and H. Sebastian Seung. 2017. Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics*. 33:2424–2426. <https://doi.org/10.1093/bioinformatics/btx180>

Aubin, J.E. 1979. Autofluorescence of viable cultured mammalian cells. *J. Histochem. Cytochem.* 27:36–43. <https://doi.org/10.1177/27.1.220325>

Bancaud, A., S. Huet, G. Rabut, and J. Ellenberg. 2010. Fluorescence perturbation techniques to study mobility and molecular dynamics of proteins in live cells: FRAP, photoactivation, photoconversion, and FLIP. *Cold Spring Harb. Protoc.* 2010.pdb.top90. <https://doi.org/10.1101/pdb.top90>

Bankhead, P. 2016. Analyzing Fluorescence Microscopy Images With ImageJ. <https://petebankhead.gitbooks.io/imagej-intro/content/> (accessed March 15, 2019)

Bolte, S., and F.P. Cordelières. 2006. A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* 224:213–232. <https://doi.org/10.1111/j.1365-2818.2006.01706.x>

Bosch, P.J., I.R. Corrêa Jr., M.H. Sonntag, J. Ibach, L. Brunsvel, J.S. Kanger, and V. Subramaniam. 2014. Evaluation of fluorophores to label SNAP-tag fused proteins for multicolor single-molecule tracking microscopy in live cells. *Biophys. J.* 107:803–814. <https://doi.org/10.1016/j.bpj.2014.06.040>

Burry, R.W. 2011. Controls for immunocytochemistry: an update. *J. Histochem. Cytochem.* 59:6–12. <https://doi.org/10.1369/jhc.2010.956920>

Caicedo, J.C., S. Cooper, F. Heigwer, S. Warchal, P. Qiu, C. Molnar, A.S. Vasilievich, J.D. Barry, H.S. Bansal, O. Kraus, et al. 2017. Data-analysis strategies for image-based cell profiling. *Nat. Methods*. 14:849–863. <https://doi.org/10.1038/nmeth.4397>

Costantini, L.M., M. Fossati, M. Francolini, and E.L. Snapp. 2012. Assessing the tendency of fluorescent proteins to oligomerize under physiologic conditions. *Traffic*. 13:643–649. <https://doi.org/10.1111/j.1600-0854.2012.01336.x>

Costantini, L.M., M. Balaban, M.L. Markwardt, M. Rizzo, F. Guo, V.V. Verkhusha, and E.L. Snapp. 2015. A palette of fluorescent proteins optimized for diverse cellular environments. *Nat. Commun.* 6:7670. <https://doi.org/10.1038/ncomms8670>

Couchman, J.R. 2009. Commercial antibodies: the good, bad, and really ugly. *J. Histochem. Cytochem.* 57:7–8. <https://doi.org/10.1369/jhc.2008.952820>

De Los Angeles, A., F. Ferrari, Y. Fujiwara, R. Mathieu, S. Lee, S. Lee, H.-C. Tu, S. Ross, S. Chou, M. Nguyen, et al. 2015. Failure to replicate the STAP cell phenomenon. *Nature*. 525:E6–E9. <https://doi.org/10.1038/nature15513>

Diaspro, A., G. Chirico, C. Usai, P. Ramoino, and J. Dobrucki. 2006. Photobleaching. In *Handbook Of Biological Confocal Microscopy*. J. Pawley, editor. Springer, Boston, MA. 690–702. [https://doi.org/10.1007/978-0-387-45524-2\\_39](https://doi.org/10.1007/978-0-387-45524-2_39)

Dorn, J.F., K. Jaqaman, D.R. Rines, G.S. Jelson, P.K. Sorger, and G. Danuser. 2005. Yeast kinetochore microtubule dynamics analyzed by high-resolution three-dimensional microscopy. *Biophys. J.* 89:2835–2854. <https://doi.org/10.1529/biophysj.104.058461>

Eliceiri, K.W., M.R. Berthold, I.G. Goldberg, L. Ibáñez, B.S. Manjunath, M.E. Martone, R.F. Murphy, H. Peng, A.L. Plant, B. Roysam, et al. 2012. Biological imaging software tools. *Nat. Methods*. 9:697–710. <https://doi.org/10.1038/nmeth.2084>

Ettinger, A., and T. Wittmann. 2014. Fluorescence live cell imaging. *Methods Cell Biol.* 123:77–94. <https://doi.org/10.1016/B978-0-12-420138-5.00005-7>

Freedman, L.P., M.C. Gibson, A.R.M. Bradbury, A.M. Buchberg, D. Davis, M.P. Dolled-Filhart, F. Lund-Johansen, and D.L. Rimm. 2016. [Letter to the Editor] The need for improved education and training in research antibody usage and validation practices. *Biotechniques*. 61:16–18. <https://doi.org/10.2144/000114431>

Ganini, D., F. Leinisch, A. Kumar, J. Jiang, E.J. Tokar, C.C. Malone, R.M. Petrovich, and R.P. Mason. 2017. Fluorescent proteins such as eGFP lead to catalytic oxidative stress in cells. *Redox Biol.* 12:462–468. <https://doi.org/10.1016/j.redox.2017.03.002>

Glass, D.J. 2014. *Experimental design for biologists*. Second edition. Cold Spring Harbor Laboratory Press. 289 pp.

Goldman, R.D., J.R. Swedlow, and D.L. Spector, editors. 2010. *Live Cell Imaging: A Laboratory Manual*. Second edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 631 pp.

Goodpaster, T., and J. Randolph-Habecker. 2014. A flexible mouse-on-mouse immunohistochemical staining technique adaptable to biotin-free reagents, immunofluorescence, and multiple antibody staining. *J. Histochem. Cytochem.* 62:197–204. <https://doi.org/10.1369/0022155413511620>

- Goodwin, P.C. 2013. Evaluating optical aberrations using fluorescent microspheres: methods, analysis, and corrective actions. *Methods Cell Biol.* 114: 369–385. <https://doi.org/10.1016/B978-0-12-407761-4.00015-4>
- Grillo-Hill, B.K., B.A. Webb, and D.L. Barber. 2014. Ratiometric imaging of pH probes. *Methods Cell Biol.* 123:429–448. <https://doi.org/10.1016/B978-0-12-420138-5.00023-9>
- Halpern, A.R., M.D. Howard, and J.C. Vaughan. 2015. Point by point: an introductory guide to sample preparation for single-molecule, super-resolution fluorescence microscopy. *Curr. Protoc. Chem. Biol.* 7: 103–120. <https://doi.org/10.1002/9780470559277.ch140241>
- Han, B., A. Tiwari, and A.K. Kenworthy. 2015. Tagging strategies strongly affect the fate of overexpressed caveolin-1. *Traffic.* 16:417–438. <https://doi.org/10.1111/tra.12254>
- Head, M.L., L. Holman, R. Lanfear, A.T. Kahn, and M.D. Jennions. 2015. The extent and consequences of p-hacking in science. *PLoS Biol.* 13: e1002106. <https://doi.org/10.1371/journal.pbio.1002106>
- Heppert, J.K., D.J. Dickinson, A.M. Pani, C.D. Higgins, A. Steward, J. Ahninger, J.R. Kuhn, and B. Goldstein. 2016. Comparative assessment of fluorescent proteins for in vivo imaging in an animal model system. *Mol. Biol. Cell.* 27:3385–3394. <https://doi.org/10.1091/mbc.e16-01-0063>
- Hibbs, A.R., G. MacDonald, and K. Garsha. 2006. Practical confocal microscopy. In *Handbook of Biological Confocal Microscopy*. J. Pawley, editor. Springer, Boston. 650–671. [https://doi.org/10.1007/978-0-387-45524-2\\_36](https://doi.org/10.1007/978-0-387-45524-2_36)
- Hiraoka, Y., J.W. Sedat, and D.A. Agard. 1990. Determination of three-dimensional imaging properties of a light microscope system. Partial confocal behavior in epifluorescence microscopy. *Biophys. J.* 57:325–333. [https://doi.org/10.1016/S0006-3495\(90\)82534-0](https://doi.org/10.1016/S0006-3495(90)82534-0)
- Hodgson, L., F. Shen, and K. Hahn. 2010. Biosensors for characterizing the dynamics of rho family GTPases in living cells. *Curr. Protoc. Cell Biol.* Chapter 14:1–26.
- Holman, L., M.L. Head, R. Lanfear, and M.D. Jennions. 2015. Evidence of experimental bias in the life sciences: Why we need blind data recording. *PLoS Biol.* 13:e1002190. <https://doi.org/10.1371/journal.pbio.1002190>
- Icha, J., M. Weber, J.C. Waters, and C. Norden. 2017. Phototoxicity in live fluorescence microscopy, and how to avoid it. *BioEssays.* 39:1700003. <https://doi.org/10.1002/bies.201700003>
- Inoué, S., and K. Spring. 1997. *Video Microscopy*. Second edition. Springer, New York. 742 pp.
- Ioannidis, J.P.A. 2014. How to make more published research true. *PLoS Med.* 11:e1001747. <https://doi.org/10.1371/journal.pmed.1001747>
- Joglekar, A.P., E.D. Salmon, and K.S. Bloom. 2008. Counting kinetochore protein numbers in budding yeast using genetically encoded fluorescent proteins. *Methods Cell Biol.* 85:127–151. [https://doi.org/10.1016/S0091-679X\(08\)85007-8](https://doi.org/10.1016/S0091-679X(08)85007-8)
- Johnson, W.L., and A.F. Straight. 2013. Fluorescent protein applications in microscopy. *Methods Cell Biol.* 114:99–123. <https://doi.org/10.1016/B978-0-12-407761-4.00005-1>
- Kan, A. 2017. Machine learning applications in cell image analysis. *Immunol. Cell Biol.* 95:525–530. <https://doi.org/10.1038/icb.2017.16>
- Keller, H.E. 2006. Objective lenses for confocal microscopy. In *Handbook of Biological Confocal Microscopy*. J.B. Pawley, editor. Springer, Boston. 145–161. [https://doi.org/10.1007/978-0-387-45524-2\\_7](https://doi.org/10.1007/978-0-387-45524-2_7)
- Kepler, A., S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel, and K. Johnsson. 2003. A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.* 21:86–89. <https://doi.org/10.1038/nbt765>
- Kerr, N.L. 1998. HARKing: hypothesizing after the results are known. *Pers. Soc. Psychol. Rev.* 2:196–217. [https://doi.org/10.1207/s15327957pspr0203\\_4](https://doi.org/10.1207/s15327957pspr0203_4)
- Kozioł, B., M. Markowicz, J. Kruk, and B. Plytycz. 2006. Riboflavin as a source of autofluorescence in *Eisenia fetida* coelomocytes. *Photochem. Photobiol.* 82:570–573. <https://doi.org/10.1562/2005-11-23-RA-738>
- Laisue, P.P., R.A. Alghamdi, P. Tomancak, E.G. Reynaud, and H. Shroff. 2017. Assessing phototoxicity in live fluorescence imaging. *Nat. Methods.* 14: 657–661. <https://doi.org/10.1038/nmeth.4344>
- Lambert, T.J. 2019. FPbase: A community-editable fluorescent protein database. *Nat. Methods.* In press.
- Lambert, T.J., and J.C. Waters. 2014. Assessing camera performance for quantitative microscopy. *Methods Cell Biol.* 123:35–53. <https://doi.org/10.1016/B978-0-12-420138-5.00003-3>
- Lambert, T.J., and J.C. Waters. 2017. Navigating challenges in the application of superresolution microscopy. *J. Cell Biol.* 216:53–63. <https://doi.org/10.1083/jcb.201610011>
- Landgraf, D., B. Okumus, P. Chien, T.A. Baker, and J. Paulsson. 2012. Segregation of molecules at cell division reveals native protein localization. *Nat. Methods.* 9:480–482. <https://doi.org/10.1038/nmeth.1955>
- Lazic, S.E. 2016. *Experimental design for laboratory biologists: Maximising information and improving reproducibility*. Cambridge University Press, Cambridge. 412 pp. <https://doi.org/10.1017/9781139696647>
- Lee, J.-Y., M. Kitaoka, and D.G. Drubin. 2018. A beginner's guide to rigor and reproducibility in fluorescence imaging experiments. *Mol. Biol. Cell.* 29: 1519–1525. <https://doi.org/10.1091/mbc.E17-05-0276>
- Linkert, M., C.T. Rueden, C. Allan, J.M. Burel, W. Moore, A. Patterson, B. Loranger, J. Moore, C. Neves, D. Macdonald, et al. 2010. Metadata matters: access to image data in the real world. *J. Cell Biol.* 189:777–782. <https://doi.org/10.1083/jcb.201004104>
- Los, G.V., L.P. Encell, M.G. McDougall, D.D. Hartzell, N. Karassina, C. Zimprich, M.G. Wood, R. Learish, R.F. Ohana, M. Urh, et al. 2008. HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* 3:373–382. <https://doi.org/10.1021/cb800025k>
- MacCoun, R., and S. Perlmutter. 2015. Blind analysis: Hide results to seek the truth. *Nature.* 526:187–189. <https://doi.org/10.1038/526187a>
- Magidson, V., and A. Khodjakov. 2013. Circumventing photodamage in live-cell microscopy. *Methods Cell Biol.* 114:545–560. <https://doi.org/10.1016/B978-0-12-407761-4.00023-3>
- Manning, C.F., A.M. Bundros, and J.S. Trimmer. 2012. Benefits and pitfalls of secondary antibodies: why choosing the right secondary is of primary importance. *PLoS One.* 7:e38313. <https://doi.org/10.1371/journal.pone.0038313>
- Melan, M.A., and G. Sluder. 1992. Redistribution and differential extraction of soluble proteins in permeabilized cultured cells. Implications for immunofluorescence microscopy. *J. Cell Sci.* 101:731–743.
- Miura, K., and J. Rietdorf. 2014. ImageJ Plugin CorrectBleach V2.0.2. Zenodo. Available at: <https://zenodo.org/record/30769#.XiflKyhKg2w> (accessed February 2, 2019).
- Model, M.A. 2006. Intensity calibration and shading correction for fluorescence microscopes. *Curr. Protoc. Cytom.* Chapter 10:Unit0.14.
- Model, M.A., and J.K. Burkhardt. 2001. A standard for calibration and shading correction of a fluorescence microscope. *Cytometry.* 44:309–316. [https://doi.org/10.1002/1097-0320\(20010801\)44:4<309::AID-CYTO1122>3.0.CO;2-3](https://doi.org/10.1002/1097-0320(20010801)44:4<309::AID-CYTO1122>3.0.CO;2-3)
- Munafò, M.R., B.A. Nosek, D.V.M. Bishop, K.S. Button, C.D. Chambers, N. Percie Du Sert, U. Simonsohn, E.J. Wagenmakers, J.J. Ware, and J.P.A. Ioannidis. 2017. A manifesto for reproducible science. *Nat. Hum. Behav.* 1:0021. <https://doi.org/10.1038/s41562-016-0021>
- Murphy, D.B., and M.W. Davidson. 2012. *Fundamentals of Light Microscopy and Electronic Imaging*. Second edition. Wiley-Blackwell, Hoboken, NJ.
- Murray, J.M. 2013. Practical aspects of quantitative confocal microscopy. *Methods Cell Biol.* 114:427–440. <https://doi.org/10.1016/B978-0-12-407761-4.00018-X>
- Murray, J.M., P.L. Appleton, J.R. Swedlow, and J.C. Waters. 2007. Evaluating performance in three-dimensional fluorescence microscopy. *J. Microsc.* 228:390–405. <https://doi.org/10.1111/j.1365-2818.2007.01861.x>
- Nickerson, R.S. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.* 2:175–220. <https://doi.org/10.1037/1089-2680.2.2.175>
- Norris, S.R., M.F. Núñez, and K.J. Verhey. 2015. Influence of fluorescent tag on the motility properties of kinesin-1 in single-molecule assays. *Biophys. J.* 108:1133–1143. <https://doi.org/10.1016/j.bpj.2015.01.031>
- North, A.J. 2006. Seeing is believing? A beginners' guide to practical pitfalls in image acquisition. *J. Cell Biol.* 172:9–18. <https://doi.org/10.1083/jcb.200507103>
- Nuzzo, R. 2015. How scientists fool themselves - and how they can stop. *Nature.* 526:182–185. <https://doi.org/10.1038/526182a>
- O'Connor, N., and R.B. Silver. 2013. Ratio imaging: practical considerations for measuring intracellular Ca<sup>2+</sup> and pH in living cells. *Methods Cell Biol.* 114:387–406. <https://doi.org/10.1016/B978-0-12-407761-4.00016-6>
- Pawley, J.B. 2006a. Fundamental Limits in Confocal Microscopy. In *Handbook of Biological Confocal Microscopy*. J.B. Pawley, editor. Springer US, Boston, MA. 20–42.
- Pawley, J.B., editor. 2006b. *Handbook of Biological Confocal Microscopy*. Third edition. Springer US, Boston, MA.
- Petrak, L.J., and J.C. Waters. 2014. A practical guide to microscope care and maintenance. *Methods Cell Biol.* 123:55–76. <https://doi.org/10.1016/B978-0-12-420138-5.00004-5>
- Phair, R.D., S.A. Gorski, and T. Misteli. 2003. Measurement of dynamic protein binding to chromatin in vivo, using photobleaching microscopy. *Methods Enzymol.* 375:393–414. [https://doi.org/10.1016/S0076-6879\(03\)75025-3](https://doi.org/10.1016/S0076-6879(03)75025-3)
- Postma, M., and J. Goedhart. 2018. PlotsOfData—a web app for visualizing data together with its summaries. *bioRxiv*. doi: (Preprint posted November 13, 2018) <https://doi.org/10.1101/426767>

- Remington, S.J. 2006. Fluorescent proteins: maturation, photochemistry and photophysics. *Curr. Opin. Struct. Biol.* 16:714–721. <https://doi.org/10.1016/j.sbi.2006.10.001>
- Rodriguez, E.A., R.E. Campbell, J.Y. Lin, M.Z. Lin, A. Miyawaki, A.E. Palmer, X. Shu, J. Zhang, and R.Y. Tsien. 2017. The growing and glowing toolbox of fluorescent and photoactive proteins. *Trends Biochem. Sci.* 42:111–129. <https://doi.org/10.1016/j.tibs.2016.09.010>
- Ross, S.T., J.R. Allen, and M.W. Davidson. 2014. Practical considerations of objective lenses for application in cell biology. *Methods Cell Biol.* 123: 19–34. <https://doi.org/10.1016/B978-0-12-420138-5.00002-1>
- Russ, J.C. 2004. Seeing the Scientific Image. *Proc. RMS.* 39:1–15.
- Sabuquillo, P., A. Gea, I.M. Matas, C. Ramos, and J. Cubero. 2017. The use of stable and unstable green fluorescent proteins for studies in two bacterial models: *Agrobacterium tumefaciens* and *Xanthomonas campestris* pv. *campestris*. *Arch. Microbiol.* 199:581–590. <https://doi.org/10.1007/s00203-016-1327-0>
- Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, et al. 2012. Fiji: an open-source platform for biological-image analysis. *Nat. Methods.* 9: 676–682. <https://doi.org/10.1038/nmeth.2019>
- Schnell, U., F. Dijk, K.A. Sjollem, and B.N.G. Giepmans. 2012. Immunolabeling artifacts and the need for live-cell imaging. *Nat. Methods.* 9: 152–158. <https://doi.org/10.1038/nmeth.1855>
- Selvin, P.R., and T. Ha, editors. 2008. *Single-Molecule Techniques: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 507 pp.
- Shaner, N.C. 2014. Fluorescent proteins for quantitative microscopy: important properties and practical evaluation. *Methods Cell Biol.* 123:95–111. <https://doi.org/10.1016/B978-0-12-420138-5.00006-9>
- Shaner, N.C., P.A. Steinbach, and R.Y. Tsien. 2005. A guide to choosing fluorescent proteins. *Nat. Methods.* 2:905–909. <https://doi.org/10.1038/nmeth819>
- Sheiner, L.B. 1997. Learning versus confirming in clinical drug development. *Clin. Pharmacol. Ther.* 61:275–291. [https://doi.org/10.1016/S0009-9236\(97\)90160-0](https://doi.org/10.1016/S0009-9236(97)90160-0)
- Snapp, E.L. 2009. Fluorescent proteins: a cell biologist's user guide. *Trends Cell Biol.* 19:649–655. <https://doi.org/10.1016/j.tcb.2009.08.002>
- Sommer, C., C. Straehle, U. Kothe, and F.A. Hamprecht. 2011. Ilastik: Interactive learning and segmentation toolkit. In *Proceedings—International Symposium on Biomedical Imaging*. IEEE. 230–233. doi: <https://doi.org/10.1109/ISBI.2011.5872394>
- Spiering, D., J.J. Bravo-Cordero, Y. Moshfegh, V. Miskolci, and L. Hodgson. 2013. Quantitative ratiometric imaging of FRET-biosensors in living cells. *Methods Cell Biol.* 114:593–609. <https://doi.org/10.1016/B978-0-12-407761-4.00025-7>
- Stadler, C., E. Rexhepaj, V.R. Singan, R.F. Murphy, R. Pepperkok, M. Uhlen, J.C. Simpson, and E. Lundberg. 2013. Immunofluorescence and fluorescent-protein tagging show high correlation for protein localization in mammalian cells. *Nat. Methods.* 10:315–323. <https://doi.org/10.1038/nmeth.2377>
- Stelzer, E.H.K. 1998. Contrast, resolution, pixelation, dynamic range and signal-to-noise ratio: Fundamental limits to resolution in fluorescence light microscopy. *J. Microsc.* 189:15–24. <https://doi.org/10.1046/j.1365-2818.1998.00290.x>
- Stuurman, N., and R.D. Vale. 2016. Impact of New Camera Technologies on Discoveries in Cell Biology. *Biol. Bull.* 231:5–13. <https://doi.org/10.1086/689587>
- Swedlow, J.R., K. Hu, P.D. Andrews, D.S. Roos, and J.M. Murray. 2002. Measuring tubulin content in *Toxoplasma gondii*: a comparison of laser-scanning confocal and wide-field fluorescence microscopy. *Proc. Natl. Acad. Sci. USA.* 99:2014–2019. <https://doi.org/10.1073/pnas.022554999>
- Swulius, M.T., and G.J. Jensen. 2012. The helical MreB cytoskeleton in *Escherichia coli* MC1000/pLE7 is an artifact of the N-Terminal yellow fluorescent protein tag. *J. Bacteriol.* 194:6382–6386. <https://doi.org/10.1128/JB.00505-12>
- Uhlen, M., A. Bandrowski, S. Carr, A. Edwards, J. Ellenberg, E. Lundberg, D.L. Rimm, H. Rodriguez, T. Hiltke, M. Snyder, and T. Yamamoto. 2016. A proposal for validation of antibodies. *Nat. Methods.* 13:823–827. <https://doi.org/10.1038/nmeth.3995>
- Vicente, N.B., J.E.D. Zamboni, J.F. Adur, E.V. Paravani, and V.H. Casco. 2007. Photobleaching correction in fluorescence microscopy images. *J. Phys. Conf. Ser.* 90:012068. <https://doi.org/10.1088/1742-6596/90/1/012068>
- Waters, J.C. 2009. Accuracy and precision in quantitative fluorescence microscopy. *J. Cell Biol.* 185:1135–1148. <https://doi.org/10.1083/jcb.200903097>
- Waters, J.C., and T. Wittmann, editors. 2014. *Methods in Cell Biology: Quantitative Imaging in Cell Biology*. Vol. 123. Academic Press, Waltham, MA.
- Whelan, D.R., and T.D.M. Bell. 2015. Image artifacts in single molecule localization microscopy: why optimization of sample preparation protocols matters. *Sci. Rep.* 5:7924. <https://doi.org/10.1038/srep07924>
- Williams, E., J. Moore, S.W. Li, G. Rustici, A. Tarkowska, A. Chessel, S. Leo, B. Antal, R.K. Ferguson, U. Sarkans, et al. 2017. Image Data Resource: A bioimage data integration and publication platform. *Nat. Methods.* 14: 775–781. <https://doi.org/10.1038/nmeth.4326>
- Willingham, M.C. 1999. Conditional epitopes: is your antibody always specific? *J. Histochem. Cytochem.* 47:1233–1236. <https://doi.org/10.1177/002215549904701002>
- Wolf, D.E., C. Samarasekera, and J.R. Swedlow. 2013. Quantitative analysis of digital microscope images. *Methods Cell Biol.* 114:337–367. <https://doi.org/10.1016/B978-0-12-407761-4.00014-2>
- Wu, J.-Q., and T.D. Pollard. 2005. Counting cytokinesis proteins globally and locally in fission yeast. *Science.* 310:310–314. <https://doi.org/10.1126/science.1113230>
- Zucker, R.M. 2006. Evaluation of Confocal Microscopy System Performance. In *Cell Imaging Techniques*. Vol. 319. D.J. Taatjes, and B.T. Mossman, editors. Humana Press, New York. *Methods in Molecular Biology*.