

A biologist's guide to the field of quantitative bioimaging

Rebecca A. Senft^{*1}, Barbara Diaz-Rohrer^{*1}, Pina Colarusso², Lucy Swift², Nasim Jamali¹, Helena Jambor³, Thomas Pengo⁴, Craig Brideau², Paula Montero Llopis⁵, Virginie Uhlmann⁶, Jason Kirk⁷, Kevin Andrew Gonzales⁸, Peter Bankhead¹⁰, Edward L. Evans III¹⁰, Kevin W Eliceiri¹⁰, Beth A. Cimini⁺⁺¹

¹ Imaging Platform, Broad Institute of Harvard and MIT

² Live Cell Imaging Laboratory, University of Calgary

³ National Center for Tumor Diseases, University Cancer Center, NCT-UC, Universitätsklinikum Carl Gustav Carus an der Technischen Universität Dresden, Germany

⁴ Informatics Institute, University of Minnesota Twin Cities

⁵ MicRoN Core, Harvard Medical School

⁶ European Bioinformatic Institute, European Molecular Biology Laboratory

⁷ Optical Imaging & Vital Microscopy Core, Baylor College of Medicine

⁸ Mammalian Cell Biology and Development, Rockefeller University

⁹ Institute of Genetics and Cancer, University of Edinburgh

¹⁰ Morgridge Institute and University of Wisconsin-Madison

* Co-first authors RAS and BD contributed equally to this manuscript, and each has the right to list themselves first in author order on their CVs. Remaining author order was determined by relative contribution, with authors having equivalent contributions placed in randomized order

Abstract:

Technological advancements in biology and microscopy have empowered a transition from bioimaging as an observational method to a quantitative one. However, as quantitative bioimaging is being adopted by biologists and these experiments become more complex, researchers need additional expertise to carry out this work in a rigorous and reproducible manner. Here we provide a navigational guide for biologists to understand quantitative bioimaging from sample preparation through image acquisition, image analysis, and data interpretation. We discuss the interconnectedness of these steps and for each, we provide general recommendations, key questions to consider, and links to high quality open access resources for further learning. This synthesis of information will empower biologists to efficiently plan and execute rigorous quantitative bioimaging experiments.

Introduction

The optical microscope is a major cornerstone of biological discovery, yet modern day bioimaging experiments involve much more than looking down a microscope. Proper quantification of imaging data requires planning and decision-making at every step of the imaging experimental workflow. If you are new to microscopy, it can be difficult to know where to start when planning your first experiment. This paper serves as a navigational guide for the important considerations needed in order to best utilize optical microscopy as an effective research tool. Even the most motivated and independent learner cannot master the best practices in microscopy if they do not know *what topics they must fully understand and skills they need to acquire as they work towards proficiency*.

Ideally, you will have convenient local access to training and mentorship as you develop your bioimaging knowledge and skills, but in practice, access to advanced training and support varies widely by site. While a number of hands-on specialized workshops provide excellent opportunities to enhance your training, a wealth of high quality open source resources are also freely available. For each major component of bioimaging experiments (Fig 1; Sample Preparation, Image Acquisition, Image Analysis, and Data Interpretation), we present key considerations for beginners as well as pointers to resources that present further information. We primarily focus on fluorescence microscopy, but many recommendations and resources are generalizable to other kinds of bioimaging.

Throughout these sections, there are a few recurring principles:

- 1) **There is no single correct answer.** In nearly all aspects of a microscopy experiment, you will find many possible options; this is because the optimal choice for a particular set of conditions depends on what you are trying to learn, the details of your system and your sample, and many other factors.
- 2) **Speak to experts before you begin.** These include microscopy hardware experts, imaging scientists, image analysts, statisticians, and other biologists. You may be able to find these locally in core facilities at your institution or online via community-wide forums.
- 3) **The major steps of bioimaging experiments are highly interconnected.** Decisions at one stage affect what is possible at others. It helps to think about experimental design not as a linear process, but as a “reverse workflow” in which the desired end result informs the necessary experimental steps and controls. Begin your experiments with the end in mind.
- 4) **Microscopy is inherently quantitative.** Microscopy produces beautiful images that make it easy to forget that they are inherently quantitative. By designing rigorous, reproducible experiments with proper controls and with optimized workflows, you can generate the highest quality data that will give you the power to derive more meaningful conclusions about biology.

Sample Preparation

General

A good quality imaging experiment begins at the bench. The best performing microscope will not produce good quality data unless you have first optimized sample preparation. Optimal preparation of the sample allows you to mitigate common aberrations in microscopy and enables more accurate measurements. When comparing multiple samples, samples should be processed in parallel whenever possible, using the same reagents to avoid technical variability. Positive and negative controls should always be included to ensure that sample processing went as expected and did not introduce artifacts to the imaging.

Some aspects of sample preparation are universal; a good specific example is the **coverslip thickness** whether using a slide or glass bottom dish. For most high-resolution imaging experiments your coverslip thickness is 0.17 mm (or #1.5 grade). Using a different thickness does not mean that your experiment will not work, but since most objective lenses are corrected for this specific thickness, you will notice that your sample is brighter, appears sharper and has better resolution and contrast with the appropriate coverslip thickness. It is also critical to choose an appropriate **immersion oil** for your coverslip and objective.

For fluorescence imaging, when selecting fluorophores, keep in mind specimen compatibility, available illumination sources and filters, and the overall sample preparation workflow. In general, bright and photostable fluorophores are critical, especially when doing live cell imaging, as their use minimizes phototoxicity and photobleaching, which can affect reproducibility in live cell imaging. In general, organic dyes are brighter and more photostable than fluorescent proteins, but you generally cannot genetically encode them. If doing multi-color imaging, pay attention to whether the spectra of the different fluorophores are overlapping[1].

Live samples

You should always prioritize the health of the sample, taking into account that stress is cumulative. As mentioned above, high quality imaging typically requires glass as the sample interface, but many cell types are miserable when grown directly on glass. You can coat the glass with different coatings (e.g., poly-L-lysine, gelatin, fibronectin) or use dishes made out of **optical polymers** when compatible with your immersion oil[2]. A quick literature search can help you figure out what works for your specific cell type, as the substrate used can affect the biological process you are studying. Please note, standard tissue-culture plastic is autofluorescent, so avoid regular culture dishes.

The imaging media will affect both sample health and signal. Some components such as **antioxidants** and reactive oxygen species **scavengers** can reduce phototoxicity and photobleaching, but you must test whether these compounds affect the health of your sample [3,4]. Typical media components such as phenol red, fetal bovine serum, riboflavins and

vitamins can produce a high fluorescence background; thus you must balance the components you add to your media so that your sample is healthy but also so that you can detect your signal [5–7].

Choosing dyes and fluorescent proteins with longer excitation wavelengths (e.g., NucRed instead of Hoechst) and using gentler illumination settings (lower irradiation, shorter exposure times, etc) will allow you to reduce phototoxicity and observe your sample for longer periods. It is always a good idea to monitor your sample in brightfield mode, as certain morphological changes (e.g., blebs, condensed nuclei, rounded cells...) can indicate cell stress.

Fixed samples

The first step you should consider is the **fixation procedure**. The type of fixative, its concentration, and the fixation conditions (e.g., temperature, buffer components and fixation length) have a great effect on the preservation of the structures you want to observe and the general morphology of your sample. It is generally a good idea to assess whether fixing specimens changes their morphology. As with many other aspects, you should optimize the fixation conditions for each sample and staining you use. If carrying out immunofluorescence, in which antibodies are used to label proteins of interest, you will need to **permeabilize** the sample so that your antibodies can access the intracellular space. Please note, soluble proteins and lipids that are not cross-linked will be removed during permeabilization. Quality of your labeling depends on both the proper fixation and permeabilization.

The key to successful immunolabeling is to use the appropriate **controls**. Do not assume that the vendor has done the appropriate validation steps to ensure the antibody is specific to your protein of interest. Additionally, secondary antibodies can bind nonspecifically to different cellular structures. It is important to take the time to validate your antibody, as the issues with antibody integrity are well known[8]. Appropriate **blocking** during immunostaining will reduce nonspecific binding, but secondary-only controls are a must; isotype controls are also highly recommended and when possible knockdown and knockout controls provide the gold standard for testing the specificity of the primary antibody.

In addition to immunolabeling, you can use dyes that are specific to cellular components (e.g., MitoTracker and DAPI). When using such dyes, follow the manufacturer's recommendations as some do not withstand fixation. If detecting genetically expressed fluorescent proteins in a fixed sample, please note that the fixative can affect the structure of the fluorescent proteins and reduce their brightness. If the signal is no longer detectable after fixation, you can do immunolabeling against the fluorescent protein.

The last consideration for fixed samples is the **mounting medium**. Mounting media are not alike and are more or less compatible with different fluorophores. Mounting media maintain the appropriate pH for your fluorophore, have **photobleaching-reducing** agents and have a high **refractive index** (RI) meant to reduce aberrations due to a mismatch in RI between the sample, the coverslip and the immersion medium. Many mounting media come with DAPI, but these

may produce high background or uneven staining in tissue sections and thus should be validated in your system before use [9,10].

Which type of sample should I choose?

Cultured cells (i.e., stable cell lines, stem cell-derived or dissociated cells) are one of the most widespread sample types to address biological processes in image-based science, especially because they are easy to maintain and observe under a microscope. However, this is an artificial system as cells usually exist within the context of a tissue or organism. Thus, for some biological questions you may want to look at a more complex system like whole tissues or 3D organoids.

Whole tissue is more difficult to stain and image, as the different structures heavily scatter light which reduces image quality. **Tissue sectioning** may be an appropriate solution if you want to image tissues. However, imaging thin sections limits three-dimensional information and therefore, sometimes, imaging intact tissue is necessary. In this case, **tissue clearing** may be the key to address your biological question, as it enables you to make your sample transparent. Tissue clearing comes in many flavors and new methods are constantly being published [11], thus, it is a good idea to do a literature search to pick a method that has been used with your particular sample and/or target. Additionally, it is critical that you understand what instrumentation you have available, as most clearing techniques require special objectives and/or microscope modalities.

For live tissue, you can use **ex-vivo techniques** [9] where an excised piece of tissue is kept alive in a chamber providing oxygenated media and flow. However, this imaging requires heating, flow, and oxygenation equipment in addition to the microscope, and must be carefully planned.

Alternatively, you may need to do **intravital imaging** [12]. In most cases you would need to perform different surgical techniques to generate “optical windows” to access the specific organ or feature you want to image. Like ex-vivo imaging, you will need specific environmental conditions and equipment to maintain the optimal imaging conditions and specialized instrumentation like a multiphoton microscope. Please note, you will require skilled veterinary technique and ethical approval by your institute’s ethical committee.

Most of the time sample preparation is determined by the target biological question (e.g., an experiment measuring cell motility can’t be done on fixed samples), but when there are multiple options, consider techniques that will best preserve the signal you want to measure, ideally with minimal changes to cell biology.

Image Acquisition

Fluorescence microscopes are available with a dizzying variety of features and capabilities[13]. Before you can decide which type of microscope will best suit your experiments, ask a few simple questions to help guide your decision.

How thick is your sample? How transparent is it?

One of the first things to consider is the type of specimen you are working with. As a starting point, it is important to consider the sample in terms of its transparency and thickness.

Relatively thin specimens such as tissue culture cells and ultrathin cryosections (<10 μm thick) do not scatter light significantly, so they lend themselves nicely to standard **widefield**[14] and **super-resolution**[15,16] microscopy.

Tissue sections thicker than 10-20 μm start to introduce significant out-of-focus fluorescence leading to haziness and blur in your image, so here you may want to consider **confocal**[17] and/or computational **deconvolution**[18] techniques. These approaches allow you to acquire “optical sections” that enhance contrast and resolution compared to widefield techniques. If you do not have access to a confocal microscope, combining widefield with deconvolution is a good option for samples with thicknesses up to about 20 μm . From 20-150 μm , confocal microscopes are the gold standard for optical sectioning.

Confocal microscopes reach their limit when samples reach about 150 μm in thickness, and from here **multiphoton microscopy**[19] starts to have advantages for imaging thicker sections and whole mounts. If you are working with fixed tissue, an **optical clearing method**[11,20] combined with **light-sheet**[21] microscopy can also improve tissue penetration. Tissue transparency will be very important, so you may want to consider tissue clearing strategies when imaging opaque specimens thicker than 500 μm for both multiphoton and light-sheet techniques. Light-sheet is also useful for inherently clear samples such as embryos from various species, and is useful for observing development of live specimens in early developmental stages.

Not every lab will have access to all types of microscopes. Here it's important to remember that the best microscope for your project is the microscope you have! You can go far with a basic widefield microscope, and local institutions often have reciprocal use agreements if the microscope you wish to use is available nearby. If needed, there are also several global opportunities to access microscopes through collaborative initiatives that are increasing accessibility (Table 1).

What Settings Do I Need to Adjust?

Once you have chosen your microscope, it is time to work out the optimal conditions for image acquisition. When acquiring images, no matter how simple or complex the microscope, we

recommend that you take the time to understand the microscope light path from the light source to the detector(s). Knowing which components in the light path impact the image quality will enable you to accurately adjust their settings for your application, as well as troubleshoot any minor issues with the equipment such as incorrect filters, prisms, etc. being in the path when they should not be.

Pay particular attention to the major components that will impact the quality of the images you collect. The following is a summary of key features and their typical adjustments you should look out for. Note that while these parameters are necessary, one must also check that the microscope is in good working order (e.g. cleaned per manufacturer instructions, objectives attached securely, brightfield light path is aligned correctly for Köhler illumination if brightfield imaging is being used); the maintainer of your microscope will typically do this on a regular basis but speak with them about any known issues for a given microscope and how to detect and solve them.

Standard microscopes will contain the following components critical to good image quality:

1. Light source(s)
2. Objective(s)
3. Filters
4. Detector: Camera, or other light detector.

Light source(s): On a widefield microscope, the light source used for visual observation is often the same as for data acquisition. Advanced systems may have separate light sources for visual observation versus data acquisition. While the light path is fixed for fluorescence on most microscopes, it is often composed of several independently movable components for brightfield and must be properly aligned before use.

Wavelength and intensity: It is important to match the excitation wavelength to the **absorption** spectrum of the fluorophore and adjust the intensity so that you can detect a bright enough signal while minimizing photobleaching/cell viability.

Objective(s): It is important to choose an objective that gives you both the magnification and resolving power to suit your experiment. While magnification is important, the resolving power of the objective is the most critical feature and is defined by the **numerical aperture (or NA)** and the wavelengths used for imaging. The NA is the measure of the light-gathering ability of the objective and determines how well fine features in a specimen will be resolved. Increasing NA will improve both resolution and how efficiently fluorescence emission is collected.

The **color correction** of the objective lens (usually labeled as: Fluor, Apo, or Super Apo on the lens) is critical for multicolor imaging as it allows focusing on different colors at the same plane allowing you to bring multiple different-colored objects into focus at the same time. “Fluor” typically focuses 2 colors at once, while “Apo” and “Super Apo” can handle 3 to 6 different colors at once depending on the lens.

Working distance (WD in millimeters) determines how far a lens can focus into a sample. This also includes the thickness of the coverslip and any mounting media layer, so be sure your working distance is far enough to reach! If you have trouble focusing on your sample and seem to “run out” of focus, it is often because the working distance of your lens is too short.

Fluorescence Filter sets: For fluorescence, match the filter sets to the fluorophore **absorption** and **emission spectra**. For quick reference on the spectral characteristics of different fluorophore combinations, consult a spectral viewer such as fpbase.org. Using a mis-matched filter can lead you to flawed conclusions, or even prevent you from imaging at all.

Detectors: Most instruments use either a camera (widefield/super resolution/lightsheet) or Photo-Multiplier Tube (PMT) (confocal/2-photon) for the detection of fluorescence emission.

For cameras, adjusting the **exposure** controls the amount of time the camera is collecting emission light. The optimal exposure time will depend on the brightness of the sample relative to the intensity of the excitation light. Adjust the excitation intensity to expose your specimen to the lowest dose of excitation light possible to minimize **photobleaching** and **phototoxicity**. For fluorescence, fluorophores often photobleach easily so it is better to use a lower illumination power and longer camera exposure when using camera-based systems. For scanning systems that use PMTs, consider line (or frame) **averaging** to improve the signal-to-noise ratio. Over-exposure, also known as **saturation**, can be caused by too much excitation or too long an exposure on the camera and will render your images unsuitable for quantification.

General Guidelines

The image is more than a picture: It is a data set that must be as high quality and reproducible as possible (for why, see the next section). You often need to balance different needs for spatial resolution, number of targets, and when examining dynamic processes, acquisition speed as well. Optimization by definition is iterative, and you will need to test different settings before finalizing. Review and analyze your preliminary data before finalizing your approach for your whole project. The imaging parameters of the microscope used must be recorded and reproduced within each experiment[7].

Image Analysis

Common image analysis methods

Microscopy images are inherently quantitative, which makes them a very powerful data source if used correctly. Images are grids of pixels, each with a brightness (intensity) represented by a number. As a biologist, image analysis allows you to translate these numbers into insights that answer biological questions. For example:

1. Are cells in this treatment expressing more/less protein than controls?

2. Does the subcellular localization of this protein change in different treatments?
3. How many spots (e.g., mRNA molecules) are in my cells?
4. How do my cells move, divide, or change *over time*?

Many biological questions, like all the above, require measuring individual cells within your images. **Segmentation** is assigning the pixels in your image to different 'objects'. Objects can be cells, or anything you care about measuring individually in your image: tissue sections on a slide, spots within cells, fluorescent beads, etc. In other words, segmentation is about finding the things in your image you want to measure. Segmentation commonly involves **thresholding** your image, or deciding that pixels with intensities above a certain level belong to our objects of interest and the rest are parts of the image we do not wish to measure, for example, the background.

In practice, segmentation can be challenging and just thresholding the original image is rarely sufficient. **Image processing** is a category of operations that filter or transform an image to either correct for aberrations or make segmentation easier by suppressing noise in the image. A common source of confusion for beginners is what is 'allowed' when processing images. In general, anything that produces more accurate segmentation is fair game. The goal of this processing is to create a mask of objects, a map where pixels are labeled according to which object they belong. There is some interpretation involved in defining 'accurate' segmentation, but as a good practice, you should overlay outlines of your objects *on your original images* to help you assess segmentation accuracy.

When you measure the intensities of pixels belonging to your objects, be sure to do so *on your original images*, possibly after background or illumination correction, and not on the processed images used for segmentation. However, often the important measurements are based upon the shape and number of objects alone, and do not require using pixel intensities. When this is the case, the main thing is to ensure that the spatial calibration (pixel size) is correct.

The image processing, segmentation, and measurement are combined in an **image analysis pipeline**. A typical pipeline might have the following general steps. Outputs of each step are presented after the '→' and are used as input to the next step unless otherwise noted:

1. Illumination or background correction (if used) → corrected images
2. Image processing to enhance our objects of interest → processed images
3. Segment your objects of interest *in the processed images* → objects
4. Measure those objects using intensities *in the corrected images* → measurements

Software

Once you have defined your biological question and the type of analysis you want to perform in your images, the next step is to select appropriate image analysis software. There are many options to pick from, we have narrowed down the list based on the following criteria:

- File types - software interoperability
- Open access - access to everyone at no cost

- Automation - ability to process large quantity of images
- Stability - history of maintenance and user support
- Documentation - manuals, tutorials, and examples
- Pipeline creation - ability to save workflows to increase reproducibility
- Knowledge base - ease of use by beginners

ImageJ a tried and true basic image analysis software that can be extended further with the use of plugins or the “batteries-included” distribution Fiji [22–24].

CellProfiler allows users to create image analysis pipelines that are easily scalable to use in thousands of images[25].

Icy is not only an image analysis software but also a platform for exchange of protocols within the image analysis community[26].

QuPath is a great tool for handling large 2D images, like those generated in histology[27].

Napari allows non-programming users to use data analysis tools written in Python[28].

You should make sure to cite any software tools you use in your work, including version numbers where available. You should also check for how the authors wish their software to be referenced, which often includes citing a specific publication. This is particularly important for open-source tools, because the authors typically depend on grant funding to continue maintaining and advancing the software and citations are a critical metric for obtaining funding[29].

Data management and sharing

Recent advances in microscopy such as whole mouse brain imaging, or sophisticated computer-controlled imaging experiments generate an amount of data in the TB range. Traditional methods of storing and retrieving data are no longer a viable option and data needs to be stored once and not moved. Furthermore, funding agencies are now integrating data management plans into the requirements for a successful grant application. The reproducibility crisis has further spawned community-lead initiatives to tackle the data deluge. FAIR (Findable, Accessible, Interoperable, and Reproducible) principles are no longer a fringe initiative, but are increasingly part of the research vocabulary.

Thanks to long-standing community engagement from a number of institutions around the world, we have valuable tools at our disposal to tackle data management, especially in the microscopy space. Two notable actively maintained platforms are OMERO (<https://openmicroscopy.org>) and OpenBIS (<https://openbis.ch/>). OMERO has a stronger emphasis on image data management and sharing, while OpenBIS aspires to be a comprehensive platform for experimental data, from inventory, lab notebook and data repository, at the expense of a steeper learning and adoption curve.

Images ideally present a long-term resource for the scientific community and may be re-interpreted, re-analysed or reproduced in future. To this end one should consider data accessibility to peers. Image data in papers at the moment is often not freely available but instead ‘available upon request’. Several repositories (Zenodo, figshare) and dedicated image archives (EMBL-EBI BioImage Archive), as well as a number of government agencies (e.g. NDA from the NIMH, TGCA NCI), or Image Data Resource (IDR) allow scientists to make image data available long term; in addition, dedicated annotated (added-value) databases are beginning to emerge that contain curated data for images to be further analyzed [30,31].

As an individual investigator, using version control is critical for large datasets and complex analysis pipelines. Version control is the ability to keep track of different versions of a script, or document, and switch to any of these at any time without losing the others. This is used extensively in the software industry and open source repositories such as GitHub, but can also be found in cloud document editing (e.g. Google Docs), where versions are kept automatically. An interesting tool that is emerging, especially in the neuroimaging world, is Datalad. Built on top of established version control software (Git), it provides an easy means to track changes to entire datasets, rather than only code.

Data Interpretation

Image data can support scientific observations both qualitatively (figures) and quantitatively (number tables and plots), and at times the same image data may be reused for either direction. Users should familiarize themselves with the distinction between quantitative and qualitative data analysis, as requirements for image formats and image handling differ. Hereafter, we discuss important aspects about the presentation and interpretation of these two kinds of output.

Qualitative data: image figures

Image figures show a representative observation, a qualitative finding, or a range of possible states. Image figures primarily serve for a qualitative assessment of the data, and as such no quantifications should be performed on images prepared for presentation purposes. To this end, the images must be presented in “adequate resolution” that is suitable to support the claim. Precise pixel-sizes necessarily depend on the image acquisition method, object size, and experimental objective. Image figures require a detailed method section (sample preparation, imaging setup, [32]) and a figure legend briefly explaining image object (specimen, tissue, cell type), colors (fluorophores, stainings, dyes) and all annotations.

When **adjusting images** for presentation, it is acceptable to adjust the image size, orientation and the frame (crop). Note that depending on the software used, pixel information may be redistributed when images are turned. All formatting adjustments that do not change the conclusions are allowed, but authors are advised to become familiar with instances that would be considered misleading[33]. Similarly, adjustments of the brightness and contrast are permitted when they are required to display observations faithfully - adjustments that result in

disappearance of image details and non-linear adjustments are however considered misleading[33].

To allow broad accessibility authors should carefully choose the **colors used in images**. Fluorescence microscope image data is usually collected in grayscale and may be presented as such to best present image details. Often fluorescent channels are shown in the color representing the capturing wavelength (blue, green, red, far-red), however authors should be aware that color-coding intensity values on a black background reduces the level of detail (though this varies by the color in question [34]). When multi-color images are shown, authors should consider if color combinations are accessible to color-blind audiences (e.g. not combine red with green, but rather magenta and green, see Jambor et al for examples) and possibly additionally show individual channels in grayscale for maximizing accessibility and detail. Tools for color blindness simulation of the images exist in image processing software (ImageJ/Fiji) and visibility of colors in final image figures can be tested with applications such as ColorOracle. Understandable images must contain at minimum an indication of the **physical size** of the object shown; this is most often done with a scale bar, however currently scale information is missing or incomplete in almost 50% of the publications. Further, any **annotations** added to images must be explained (in figure or figure legend). Annotations include symbols, arrows, letter codes, regions of interest/zoomed views and authors must also explain and annotate the colors used in images (staining, dye, labeled structure).

Quantitative data: measurements and statistics

Before attempting to extract meaning from a collection of measurements, regardless of their exact nature, one must first ponder how to compare them meaningfully. Image processing operations bring up the question of whether measurements should be extracted from raw or processed images, since processing methods may alter image content in ways that irreversibly distorted measurements (see image analysis section). Since imaging experiments rarely involve a single image file, another important aspect to think about is whether measurements can meaningfully be compared across images, and what is the true quantity getting measured. Intensity-based measurements, as well as indirect “second order” measurements like shapes or counts are affected by non-biological experimental variations that confound measurements referred to as batch effects [35]. Mitigating batch effects starts with keeping track of metadata about the experimental and imaging conditions [36].

Statistics provide a formal way to utilize measurements extracted from images in order to answer a specific question. The path towards identifying an appropriate statistical analysis is composed of three steps: 1) formulate a clear question, 2) identify a statistical tool (metric, test) that can answer that question, 3) understand the underlying assumptions of that tool, and whether the measurements considered satisfy these assumptions. While digging into the details of statistics is beyond the scope of this paper, we hereafter provide a non-exhaustive list of recurring questions.

- **What metric best answers my biological question?** It can be complicated to decide which metric(s) are best for any biological question; the more specific the question, the greater the chance of success. "Did expression of marker X change in my cells?", for

example, could be answered by looking at the total intensity of a given fluorophore, its mean intensity, and/or its *distribution* across certain parts of the cell; "Did the total amount of marker X increase within the nucleus?" is easier to use for metric selection.

- **Is my data normally distributed?** Many statistical tools assume that the data/measurements considered follow a Gaussian distribution, but this is rarely the case in biological data. Microscopy images are most often composed of different “groups” (e.g., background and foreground intensities, positive and negative phenotype), resulting in multimodal distributions of measurements. The validity of the assumption of normality can easily be verified relying on tests such as the Kolmogorov–Smirnov or the Shapiro–Wilk tests. The most famous example of the impact that normality may have on the choice of statistical method is perhaps the mean and the median: both aim to return the “central” value in a collection, but only one of them (the median) matches our intuition whether the data is normally distributed or not.
- **What is the right sample size for my study?** The amount of samples required to detect an effect of interest depends on the statistical analysis being performed and on an assumed statistical model of the studied dataset. To complexify things further, what is considered as a sample depends on the question: individual samples may be biological replicate (e.g., animal), individual petri dishes or wells, or individual cells. Consider displaying your quantitative data using approaches that display summary statistics (e.g., the mean) alongside the individual data points, such as with SuperPlots[37].

Resources in the table below dig further into these questions and provide in-depth insights into the most commonly used statistical tools in the context of biological research and/or bioimaging.

Community-wide resources

In addition to the specific recommendations and resources we’ve highlighted here, we also want to welcome you to the bioimaging and bioimage analysis community, which you can access via the forums forum.image.sc and forum.microlist.org. On these forums you can ask questions to imaging and image analysis experts, from issues with specific software to general experimental design questions and questions about controls. We aim to be an inclusive, friendly environment for beginners to experts alike. We hope the forum can make expert advice more accessible, especially for scientists at institutions without a dedicated imaging or image analysis core facility

Sample Preparation		
Resource Name	Link	Brief description
FPbase [1]	https://www.fpbases.org/	Database for identifying fluorophores by brightness, spectra and assessing compatibility with other fluorophores and with microscope filters
Bio-protocol	https://bio-protocol.org/en/about	Website to search for protocols across biological disciplines, including protocols associated with work published elsewhere. All protocols are available under an open access license (CC BY or CC BY-NC)
Image Acquisition		
Resource Name	Link	Brief description
Nikon MicroscopyU[38]	https://www.microscopyu.com/microscopy-basics	Fundamentals of microscopy explained for beginners with lots of images and plain language descriptions of terms used in microscopy
Fluorescence microscopy - avoiding the pitfalls [39]	https://journals.biologists.com/jcs/article/120/10/1703/29404/Fluorescence-microscopy-avoiding-the-pitfalls	Short overview of some of the most common pitfalls for beginners to fluorescence microscopy
Image Analysis		
Resource Name	Link	Brief description
Peter Bankhead's <i>Intro to Bioimage Analysis</i> [40]	https://bioimagebook.github.io/	Guide for absolute beginners to image analysis, including embedded questions/answers, exercises with Python and ImageJ, and videos to check your understanding
Reproducible image handling and analysis [41]	https://www.ebiopress.org/doi/full/10.15252/embj.2020105889	An article reviewing major pitfalls in image handling and how to avoid them and create reproducible analysis workflows
iBiology Bioimage Analysis video series	https://www.youtube.com/watch?v=1x04vi6Ub4I&list=RDCMUCimxROEIINsiRqT9X5tg2XQ&index=4	Video series that introduces Bioimage analysis, including overviews of image processing, segmentation, tracking, making and interpreting measurements, tips and pitfalls
Data Interpretation		

Resource Name	Link	Brief description
Creating clear and informative image-based figures for scientific publications [34]	https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3001161	Review article on how to create accessible, fair scientific figures, including guidelines for microscopy images
Modern Statistics for Modern Biology [42]	https://www.huber.embl.de/msmb/	Online statistics for biologists textbook with code examples (in R)
Community-wide Resources		
Resource Name	Link	Brief description
Advanced Imaging Center	https://www.aicjanelia.org/apply	Access to the state of the art microscopy instruments and imaging experts
Africa Microscopy Initiative	https://www.microscopy.africa/	Access to advanced microscopes, molecular biology and cell culture equipment for scientists in Africa
Euro-Bioimaging	https://www.eurobioimaging.eu	Access to microscopy instruments and training for scientists in Europe
Image.sc [43]	https://forum.image.sc/	Discussion forum for bioimage analysis software
Bioimage ANalysis Desktop (BAND)	https://band.embl.de	Access to virtual desktops so you can use bioimage analysis software from a browser

Table 1. Beginner-friendly resource list. A collection of beginner-friendly, online, open access resources for each stage of a quantitative bioimaging experiment as well as community-wide resources.

Conclusion

Each imaging experiment is different, here we aimed to present general guidelines and basic concepts that can guide your decision-making process when designing your own experiment. Designing your experiments so that they can be quantitative and reproducible not only benefits you but the community around you, to that end we reiterate to make your data, analysis pipelines and other resources open to the public. The information presented is not exhaustive and we recommend talking to the experts and looking at resources here and on the companion website as well as published literature that is relevant to your experiment.

Acknowledgements

The authors thank Anne Carpenter, Ellen Dobson and Pearl Ryder for helpful discussions around important resources to highlight.

For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising from this submission.

Funding

Funding was provided by the National Institutes of Health (NIH COBA P41 GM135019 to BAC and KWE). This project has been made possible in part by grant number 2020-225720 to BAC from the Chan Zuckerberg Initiative DAF, an advised fund of the Silicon Valley Community Foundation. VU is supported by EMBL internal funding. ELE is supported by a Morgridge Postdoctoral Fellowship.

Checklist:

The following list provides questions to consider before embarking on a quantitative bioimaging experiment.

Sample Preparation

- ☐ What is the most appropriate sample type for me to use - cultured cells, tissue sections, whole organisms?
- ☐ What samples and/or conditions will I test, including positive and negative controls?
- ☐ What dyes or stains (if any) will I use?
- ☐ Is my fluorophore combination appropriate?
- ☐ What controls must I perform to assess the health of my sample?
- ☐ What controls must I perform to address the specificity of my reagents?

Image Acquisition

- ☐ What microscope modality should I use? Do I need confocal?
- ☐ Is the microscope modality appropriate for my sample type?
- ☐ Am I using the right objective for my sample?
- ☐ Are the filter sets appropriate for my fluorophore combination?
- ☐ Are my acquisition settings appropriate and consistent?

Image Analysis

- ☐ Have I determined the correct metric for what I want to measure in my images?
- ☐ Have I set up my analysis correctly to generate quantitative data?
- ☐ Are my measurements made equivalently for controls and experimental samples?
- ☐ Are all code, scripts, pipelines etc. included with my publication materials?
- ☐ Is all of the raw data accessible or included with the publication?

Data Interpretation

- ☐ Do qualitative figures comply with best practices on colors used, annotations, and other adjustments?
- ☐ Do all microscopy images include a scale bar?
- ☐ Have all adjustments made to qualitative figures been disclosed in compliance with best practices and journal policies?
- ☐ Have appropriate summary statistics been used for quantitative figures?
- ☐ Do my quantitative figures allow the reader to assess the shape of the distribution, including individual plot points wherever possible?

Figures

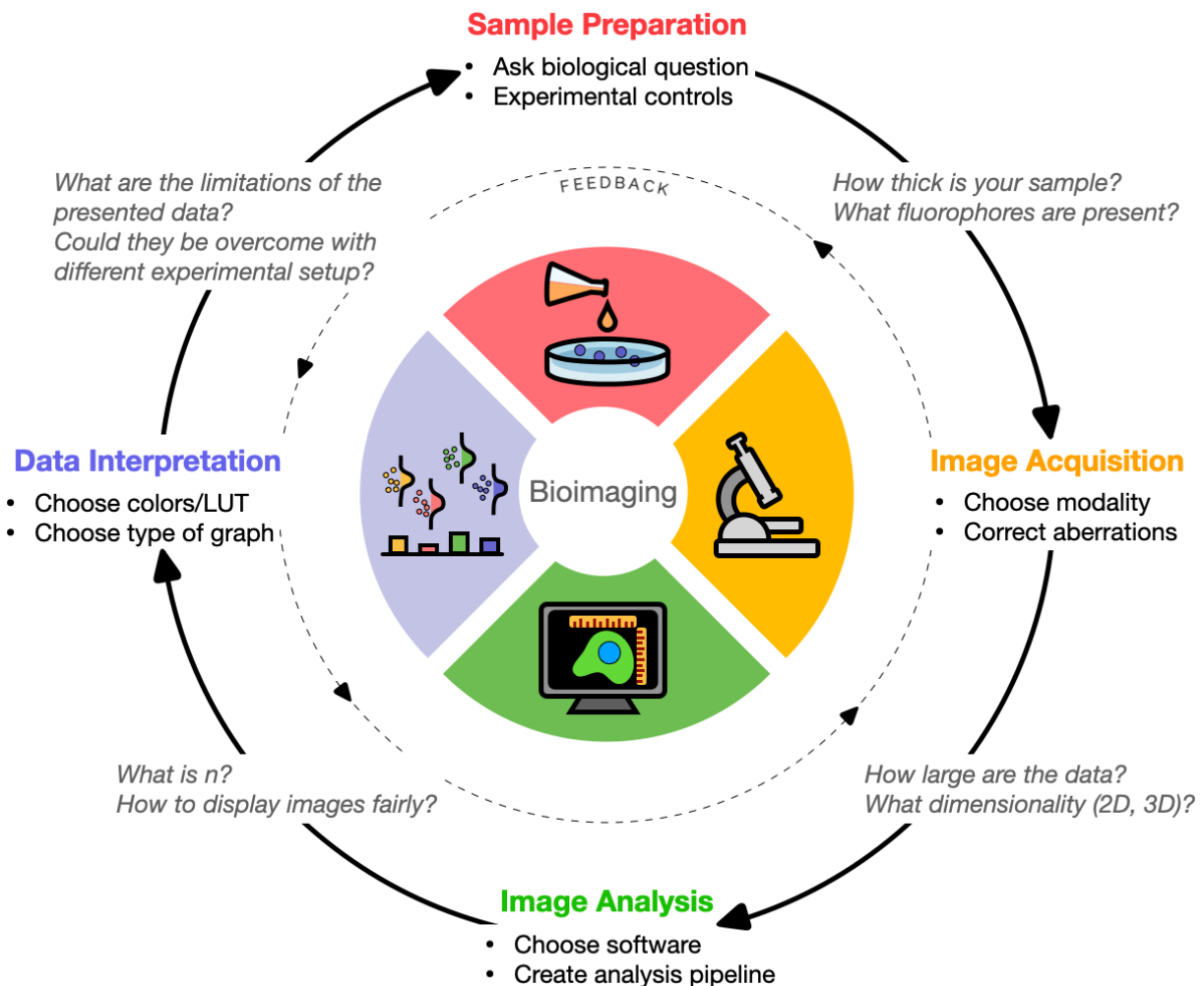


Figure 1. The steps of any quantitative bioimaging experiment are mutually interconnected. Each quadrant represents one of the four key elements to any bioimaging experiment, from sample preparation, to image acquisition, image analysis, and data interpretation. Between quadrants, example questions demonstrate how decisions made at once level feed forward to affect subsequent steps. Analogously, desired endpoints inform how earlier steps should be performed.

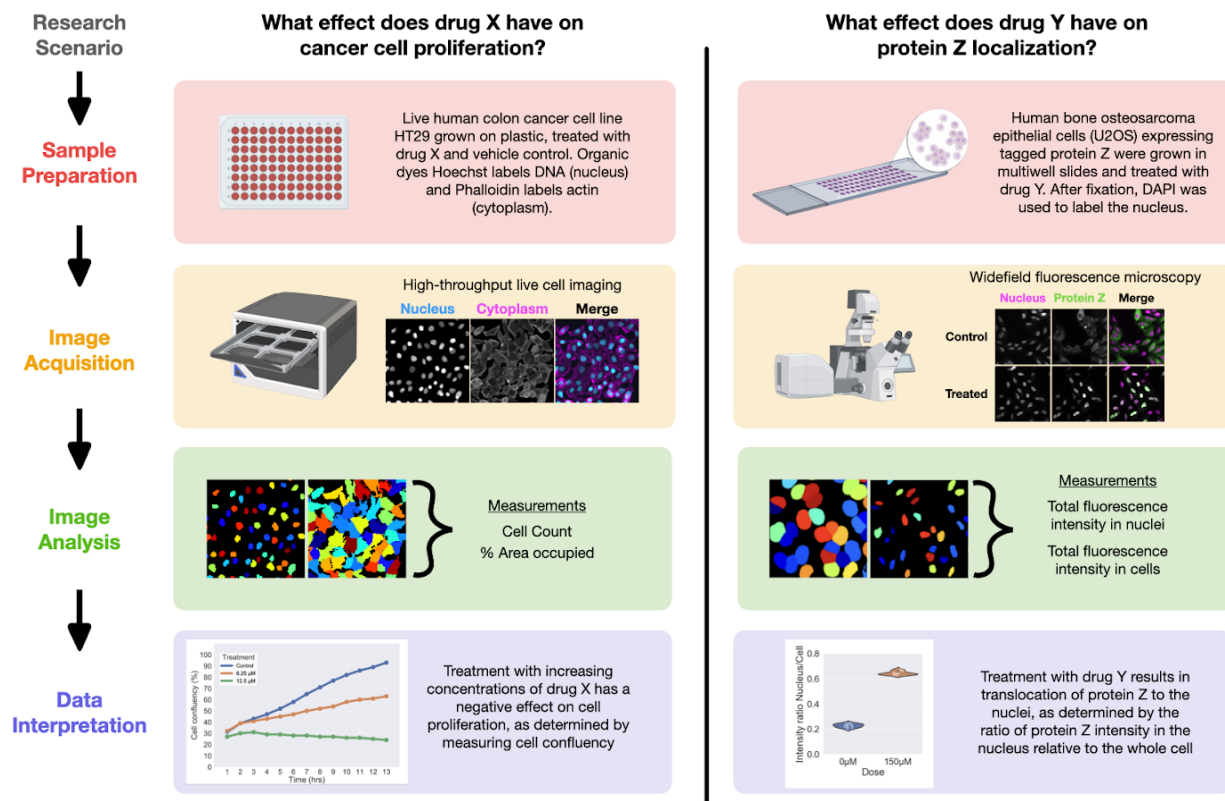


Figure 2. Two quantitative bioimaging example experiments. Each hypothetical scenario is broken down into the four major stages of quantitative bioimaging and demonstrates how a biological question is translated into an experiment that yields interpretable data. Images are sourced from BBBC008v1 [44] and BBBC013v1 (provided by Ilya Ravkin) available from the Broad Bioimage Benchmark Collection [44,45]. Created with BioRender.com.

References

1. Lambert TJ. FPbase: a community-editable fluorescent protein database. *Nat Methods*. 2019;16: 277–278.
2. Compatible Immersion Oil for Polymer Coverslip- FAQ - ibidi. [cited 6 Dec 2022]. Available: <https://ibidi.com/content/551-immersion-oils-compatible-with-ibidi-labware-products>
3. Icha J, Weber M, Waters JC, Norden C. Phototoxicity in live fluorescence microscopy, and how to avoid it. *Bioessays*. 2017;39. doi:10.1002/bies.201700003
4. Tosheva KL, Yuan Y, Matos Pereira P, Culley S, Henriques R. Between life and death: strategies to reduce phototoxicity in super-resolution microscopy. *J Phys D Appl Phys*. 2020;53: 163001.
5. Frigault MM, Lacoste J, Swift JL, Brown CM. Live-cell microscopy – tips and tools. *Journal of Cell Science*. 2009. pp. 753–767. doi:10.1242/jcs.033837

6. Bogdanov AM, Kudryavtseva EI, Lukyanov KA. Anti-fading media for live cell GFP imaging. *PLoS One*. 2012;7: e53004.
7. Montero Llopis P, Senft RA, Ross-Elliott TJ, Stephansky R, Keeley DP, Koshar P, et al. Best practices and tools for reporting reproducible fluorescence microscopy methods. *Nat Methods*. 2021;18: 1463–1476.
8. Uhlen M, Bandrowski A, Carr S, Edwards A, Ellenberg J, Lundberg E, et al. A proposal for validation of antibodies. *Nat Methods*. 2016;13: 823–827.
9. Krishnamurthy S, Brown JQ, Iftimia N, Levenson RM, Rajadhyaksha M. Ex Vivo Microscopy: A Promising Next-Generation Digital Microscopy Tool for Surgical Pathology Practice. *Arch Pathol Lab Med*. 2019;143: 1058–1068.
10. Quantitative Imaging in Cell Biology. *Methods in Cell Biology*. 2014. doi:10.1016/c2013-0-09942-2
11. Richardson DS, Guan W, Matsumoto K, Pan C, Chung K, Ertürk A, et al. TISSUE CLEARING. *Nat Rev Methods Primers*. 2021;1. doi:10.1038/s43586-021-00080-9
12. Pittet MJ, Weissleder R. Intravital imaging. *Cell*. 2011;147: 983–991.
13. Sanderson MJ, Smith I, Parker I, Bootman MD. Fluorescence microscopy. *Cold Spring Harb Protoc*. 2014;2014: db.top071795.
14. Swedlow JR, Platani M. Live cell imaging using wide-field microscopy and deconvolution. *Cell Struct Funct*. 2002;27: 335–341.
15. Verdaasdonk JS, Stephens AD, Haase J, Bloom K. Bending the rules: widefield microscopy and the Abbe limit of resolution. *J Cell Physiol*. 2014;229: 132–138.
16. Stemmer A, Beck M, Fiolka R. Widefield fluorescence microscopy with extended resolution. *Histochem Cell Biol*. 2008;130: 807–817.
17. Jonkman J, Brown CM, Wright GD, Anderson KI, North AJ. Tutorial: guidance for quantitative confocal microscopy. *Nat Protoc*. 2020;15: 1585–1611.
18. Swedlow JR. Quantitative Fluorescence Microscopy and Image Deconvolution. *Methods in Cell Biology*. Academic Press; 2007. pp. 447–465.
19. Ustione A, Piston DW. A simple introduction to multiphoton microscopy. *J Microsc*. 2011;243: 221–226.
20. Gómez-Gavito MV, Sanderson D, Ripoll J, Desco M. Biomedical Applications of Tissue Clearing and Three-Dimensional Imaging in Health and Disease. *iScience*. 2020;23: 101432.
21. Wan Y, McDole K, Keller PJ. Light-Sheet Microscopy and Its Potential for Understanding Developmental Processes. *Annu Rev Cell Dev Biol*. 2019;35: 655–681.
22. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012;9: 671–675.

23. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012;9: 676–682.
24. Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET, et al. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*. 2017;18: 529.
25. Stirling DR, Swain-Bowden MJ, Lucas AM, Carpenter AE, Cimini BA, Goodman A. CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics*. 2021;22: 433.
26. de Chaumont F, Dallongeville S, Chenouard N, Hervé N, Pop S, Provoost T, et al. Icy: an open bioimage informatics platform for extended reproducible research. *Nat Methods*. 2012;9: 690–696.
27. Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: Open source software for digital pathology image analysis. *Sci Rep*. 2017;7: 16878.
28. Sofroniew N, Lambert T, Evans K, Nunez-Iglesias J, Bokota G, Winston P, et al. napari: a multi-dimensional image viewer for Python. 2022. doi:10.5281/zenodo.7276432
29. Levet F, Carpenter AE, Eliceiri KW, Kreshuk A, Bankhead P, Haase R. Developing open-source software for bioimage analysis: opportunities and challenges. *F1000Res*. 2021;10: 302.
30. Ellenberg J, Swedlow JR, Barlow M, Cook CE, Sarkans U, Patwardhan A, et al. A call for public archives for biological image data. *Nat Methods*. 2018;15: 849–854.
31. OSF. [cited 14 Dec 2022]. Available: <https://osf.io/>
32. Marqués G, Pengo T, Sanders MA. Imaging methods are vastly underreported in biomedical research. *Elife*. 2020;9. doi:10.7554/eLife.55133
33. Crome DW. Avoiding twisted pixels: ethical guidelines for the appropriate use and manipulation of scientific digital images. *Sci Eng Ethics*. 2010;16: 639–667.
34. Jambor H, Antonietti A, Alicea B, Audisio TL, Auer S, Bhardwaj V, et al. Creating clear and informative image-based figures for scientific publications. *PLoS Biol*. 2021;19: e3001161.
35. Sonesson C, Gerster S, Delorenzi M. Batch effect confounding leads to strong bias in performance estimates obtained by cross-validation. *PLoS One*. 2014;9: e100335.
36. Caicedo JC, Cooper S, Heigwer F, Warchal S, Qiu P, Molnar C, et al. Data-analysis strategies for image-based cell profiling. *Nat Methods*. 2017;14: 849–863.
37. Lord SJ, Velle KB, Mullins RD, Fritz-Laylin LK. SuperPlots: Communicating reproducibility and variability in cell biology. *J Cell Biol*. 2020;219. doi:10.1083/jcb.202001064
38. Kenneth R. Spring, Hiroshi Komatsu, Martin L. Scott, Stanley A. Schwartz, Thomas J. Fellers, Kathleen E. Carr, Matthew Parry-Hill, Michael W. Davidson. Basic Concepts and Formulas in Microscopy. In: Nikon's MicroscopyU [Internet]. Available: <https://www.microscopyu.com/microscopy-basics>

39. Brown CM. Fluorescence microscopy--avoiding the pitfalls. *J Cell Sci.* 2007;120: 1703–1705.
40. Bankhead P. Introduction to Bioimage Analysis. 2022. Available: <https://bioimagebook.github.io/README.html>
41. Miura K, Nørrelykke SF. Reproducible image handling and analysis. *EMBO J.* 2021;40: e105889.
42. Holmes S, Huber W. Modern Statistics for Modern Biology. Feb 2019. Available: <https://www.huber.embl.de/msmb/>
43. Rueden CT, Ackerman J, Arena ET, Eglinger J, Cimini BA, Goodman A, et al. Scientific Community Image Forum: A discussion forum for scientific image software. *PLoS Biol.* 2019;17: e3000340.
44. Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang I, Friman O, et al. Genome Biology. 2006. p. R100. doi:10.1186/gb-2006-7-10-r100
45. Ljosa V, Sokolnicki KL, Carpenter AE. Annotated high-throughput microscopy image sets for validation. *Nat Methods.* 2012;9: 637.