Standardization of High Content Imaging and Informatics

O. Joseph Trask, Jr.,¹ and Paul A. Johnston²

¹Cellular Imaging Core, The Hamner Institutes for Health Sciences, Research Triangle Park, North Carolina.

²Department of Pharmaceutical Sciences, School of Pharmacy, University of Pittsburgh, Pittsburgh, Pennsylvania.

INTRODUCTION

There is an old folk saying popularized by the late Bert Lance that reads “if it ain’t broke, don’t fix it.” The application of this philosophy to scientific endeavors creates major problems if scientists fail to appreciate when the instruments used to collect and analyze data are neither properly operated nor calibrated. Such behaviors undermine the widely held expectation that scientists do everything possible to ensure the quality and reproducibility of their data. Recently, the Society of Biomolecular Imaging and Informatics (SBI²) emerged to become engaged in a technology discipline that has grown exponentially over the past 20 years without adopting either standards or guidelines. The procedures to validate the integrity and performance of automated imaging platforms and the image analysis methods used to produce data for publication or drive business decisions are not routinely available to the end user. Rather, the majority of laboratories using automated imaging platforms for high content screening (HCS) typically employ reference control compounds (activators and/or inhibitors) to demonstrate the performance of their assays. Although this practice may currently be considered acceptable by most programs and facilities in academia, biopharmaceutical discovery, and research institutions in nonclinical settings, there are some major underlying assumptions that may not always be valid: the instrument has been calibrated correctly and is being operated optimally to acquire images; the image analysis algorithm has been appropriately optimized and applied in an unbiased manner; and that the data are suitably annotated, stored, and retrievable following archive. The truth is, “any image is data, good or bad,” and therefore it is critical that the imager be set up to acquire the best images possible and they be stored, analyzed, and annotated appropriately.

The advantage of implementing standards in HCS is to provide a mechanism for the scientific community to directly compare data generated across multiple HCS platforms. These standards not only provide a means to calibrate instrumentation, assays, and results, but also can be applied to cross-reference data generated in one laboratory against other laboratories to verify the reproducibility of HCS data independently. Adoption of a standardized annotation and nomenclature to describe data analysis methods and outputs will facilitate understanding and comparisons within the scientific community. Standardized biological data is not a new concept, and the urgent need and support for standards is becoming even more evident in research as scientists are challenged with a plethora of data and the inability to reproduce published work.^{1,2} Leonard Freedman from the Global Biological Standards and colleagues recently published a commentary in *PLoS Biology* about the cost of irreproducibility in preclinical sciences.³ The fact that scientific research is not always reproducible is not surprising, and to a certain extent this idea has become accepted by many that have worked in the field for years. What makes the commentary of Freedman *et al.* of greater importance is the mind-boggling estimate of \$28 billion in economic losses resulting from only 50% of scientific work being reproducible. With the global shrinkage in sponsored research funds from both private and government agencies, the scientific community needs to minimize the waste in spending and improve the reproducibility of published research to address the urgent need to bring new therapies to market.

Clearly, one approach to enhance the reproducibility of research findings would be to implement strategies that improve data integrity. Systematically, these should include thoughtful experimental design, validation of starting materials such as cells and growth nutrients, implementation of appropriate assay development strategies, and statistical validation of experimental results. The use of known standards and reference compounds can also help to validate assay performance. Examples of how to implement standardized bioassays practically for high throughput screening and drug discovery are provided in the NIH Assay Guidance Manual, originally donated by Eli Lilly and Company more than

10 years ago. Guidelines in this manual have become an excellent resource for researchers developing biochemical and cell-based assays.⁴ Included in the guidelines is a section on HCS describing the basic principles and concepts, the selection of a cell model, and methods to develop, optimize, and validate HCS assays. There are also two case studies on NF- κ B translocation and neurite outgrowth, and an advanced HCS methods section. Resources such as these provide insight and guidance to help newcomers to HCS develop robust and reproducible assays.

As HCS continues to expand into new scientific areas and it adapts to new reagents, cell models, and biological applications, the need for standardization becomes increasingly apparent. The successful outcome to forge new standards requires the partnership between manufacturers, end users, and scientific organizations. An example of this success was the 1995–1996 proposal by the leadership and members of the Society for Biomolecular Screening (SBS) for a standardized microtiter plate (microplate) with a rectangular footprint dimensional size of 127.76 mm $\times$ 85.48 mm (ANSI/SLAS 1-2004), better known today as the ANSI/SLAS plate standard.

The following sections provide an overview of the current landscape for the development and incorporation of standards in three areas of HCS technology: Reference Material Standard, Image File Format Standard, and Image Analysis Data Standard.

TERMINOLOGY

Even the terminology used to describe the technology discipline has not been standardized. Since the introduction of computer-assisted automated microscopy imaging, the scientific community has created several acronyms for the technology, including cellular analysis (CA), cellular imaging (CI), cellular imaging analysis (CIA), high content screening (HCS), high content cellular screening (HCCS), high content analysis (HCA), high content cellular analysis (HCCA), high content imaging (HCI), fluorescent image analysis (FIA), microscopy image analysis (MIA), image cytometry (IC), or even the non-acronym description of automated confocal high throughput microscopy. None of these acronyms or descriptions effectively captures the entire technology process; rather, they describe specific applications. For simplicity and consistency, “HCS” is used to describe the technology below.

REFERENCE MATERIAL STANDARD FOR INSTRUMENT PERFORMANCE

In laboratories that are not required to operate under a GLP or GMP environment, sometimes only passing attention is given to the instruments used to generate research data. In

reality, however, all data should be considered “suspect” until it is known whether the instruments or devices being used to capture images are calibrated and performing optimally. The automated HCS platform should be able to capture the same quality image any day after it was installed and commissioned by the manufacturer. While it is critical to verify instrument performance, it is just as important to have a method to compare and reproduce data acquired on HCS platforms from different manufacturers in the same laboratory, or in a laboratory in a different hemisphere.

Automated HCS imaging platforms are complicated instruments with lots of moving parts and components that all need to be calibrated and operated in the correct manner: the stage; the light source; the autofocus system; and the light path, including objectives, filters, mirrors, and the camera or photomultiplier tube. A critical component of the research workflow from cells to data should be a process to determine if the HCS platform is fully optimized and validated to acquire high-quality images on a daily basis. Typically, a service engineer from the manufacturer will set up and calibrate the imager during the installation, but generally the end user does not receive specific training on either the calibration process or specific tests that could be run to verify instrument performance on a routine basis. Instead, the end user only receives training on how to acquire and analyze images. There is a critical need for the HCS community to require manufacturers to define a calibration process with specific tests and provide a material standard so that end users can verify the proper operation of the HCS platform. These issues are often overlooked by those just learning about HCS, but also by very experienced practitioners of the technology. Some may question if this is really a problem: “the manufacturer of my instrument calibrated and set up the imager” or “my instrument works, it takes great pictures and the data look great, so what is the problem?” However, good or bad, any image is data, and it is important to ensure that the HCS instrument is taking the best possible image and to record the instrument settings so that the results can be reproduced. In addition to improving overall instrument performance, the use of proper calibration protocols and material standards can help to normalize the instrument baseline, sensitivity, or limits of detection. The adoption of these procedures and material standards will greatly improve the likelihood that HCS experimental data can be reproduced both within and between laboratories.

In an effort to understand the current concerns and needs in the community for qualifying instrument performance, the SBI² Southeast Regional Symposium held in Research Triangle Park, North Carolina, in April 2015 gathered information

through a questionnaire provided to attendees. The survey asked a few very basic questions about quality control and standards used to measure HCS instrument performance in an attempt to understand laboratory practices better. Only 30% of attendees responded, resulting in a small sample size of 33. Nevertheless, although the survey was not statistically powerful, it provided some insights into practices being applied at universities, institutes, government agencies, and manufacturing and biopharmaceutical companies. The majority of respondents were scientists, technicians, and graduate students, but there were also a number of management or professor-level responses that helped shape the results. Approximately 50% of the respondents work with or in a shared core facility setting. When asked how often their laboratory or core perform quality control tests on a high content imager to measure overall performance, the most chosen answer was “only after instrument repair or service.” It was not daily, weekly, monthly, or even yearly; it was only when service to the instrument was required. Does this mean the instruments in our laboratories are producing quality images on a daily basis that are acceptable and the need for a standard is really not that important? Are scientists dependent on the instrument “as is” or do they just trust that the instrument is performing under optimal conditions? Users dependent on core facilities expect the service they are paying for to be optimal. They rarely ask to see the quality control reports from the facility, so they are placing their trust in something not under their control. To help further explore attitudes, the last question in the survey was, “Does having a material standard matter to your research?” The question did not ask if we need a standard but instead whether it is relevant in research. Does it matter? Eighty-five percent of the survey respondents answered “Yes, having a material standard is important to my research.” The survey also questioned the type of reference material being used to calibrate HCS instrumentation and found that most groups used colored slides and/or microplates containing fluorescently labeled cells (discussed further below). Most respondents believed that if the HCS imager was producing images and associated data, then it must be performing adequately. Importantly though, there were many respondents who expressed their concerns that having a material standard is critical to their research. This survey itself is not a comprehensive representation of the entire community, but it has provided a broad snapshot of the philosophy of scientists in the community.

The most commonly used reference materials for HCS instrument performance are colored slides and microbeads. Colored slides from Chroma Corporation have been a great asset to the microscopy community and have been used for years to measure flat-field illumination correction. These

colored slides come with the caveat that oxidation occurs upon interrogation of the sample area with subsequent decay of its calibration properties. To address this, Chroma recommends end users use a map of the slide so that no area is probed twice. MultiSpeck™ multispectral microbeads originally developed by Molecular Probes (Life Technologies) and used in flow cytometry and confocal microscopy applications have also been used to monitor HCS imager performance. In fact, BD BioSciences provided end users with a single 96-well microplate embedded with MultiSpeck microbeads. These plates were exclusively made for BD BioSciences Pathway HCS instruments to monitor light source and optics performance. However, BD is phasing out manufacturing of HCS instruments, and these plates will no longer be available. Like the colored slides from Chroma Corporation, the MultiSpeck microbeads and other fluorescent beads can oxidize and degrade when exposed to light and therefore have a limited lifespan for measuring performance over time. Likewise, standard reference materials should have an expiration date for quality control purposes. Another challenge is whether the reference materials are representative of the dynamic range of the biological samples being acquired on the HCS platform with endpoints chosen to match their fluorescent saturation limits.

The material reference standards should not to be confused with two other essential methods for instrument calibration. The first is correction for field illumination in the optical path. In this case, fluorescent dye solutions can be selected to match end user probe sets to correct for unbalanced fluorescent illumination in the image. Vignetting is a common observation, in which the corners of image background are dimmer than the center part of the image field due to limitations of the optical path, camera sensor, electronic instability, or noise. By applying a process known as flat-field illumination correction, the uneven variations in fluorescence saturation across the captured field area are equilibrated, resulting in a more uniform homogeneous background. The second very critical aspect of instrument performance to be measured and monitored is the axial resolution in *xyz* dimensions; without appropriate calibration, determination of co-localization measurements can be significantly affected.

To address the issues discussed above and to drive change, SBI² has recently formed a Material Standards Committee that is being spearheaded by Steve Titus at NIH/NCATS, Robert Zucker at US-EPA, and Michael Halter at the National Institutes of Standards and Technology (NIST). Michael Halter *et al.* recently published a paper with colleagues in *Cytometry* about an automated protocol for performance benchmarking a widefield fluorescent microscope system.⁵ In this paper, they describe two potential standard material candidates—uranyl-ion-doped glass

and Schott 475 GG filter glass—both of which exhibit stable and homogeneous fluorescence when excited on a widefield fluorescence microscope. These materials are under investigation to determine their utility as a material standard. Steve Titus has begun to qualify this approach on automated HCS platforms at the NCATS screening facility, and Robert Zucker brings decades of experience to determine calibration and alignment of confocal microscope devices properly.⁶ Together, this committee will make recommendations to the community for the procedures and reference materials required to calibrate HCS systems properly. Once announced, we hope that these methods will be adopted by the society and by the HCS instrument manufacturers, perhaps leading to more routine methods to track instrument performance within part of the graphical user interface of commercial software platform packages, similar to the approach used in flow cytometry. The purpose of having calibration by the end user is not to increase unnecessary service calls to the instrument manufacturers but to monitor the overall reliability of the instrument to capture quality images for image analysis processing. To this end, it is the intent of the SBI² scientific community to work with manufacturers to advance the use of a universal material standard to measure instrument performance.

IMAGE FILE FORMAT STANDARD

Historically, the scientific community has learned many lessons in developing and adopting standards that have worked well for other cellular technologies; most notable and similar in many respects to HCS is flow cytometry. Since it was introduced commercially in the early 1970s, flow cytometry underwent years of growth, with each manufacturer creating unique associated priority data file formats before a universal data file standard called Flow Cytometry Standard (FCS) was introduced and adopted in 1984.⁷ FCS binary files could be imported and analyzed on any compatible software platform. The FCS format was adopted and accepted by not only the scientific community but also the instrument manufacturers and the International Society for Analytical Cytology (ISAC), since rebranded as the International Society for the Advancement of Cytometry. ISAC has lobbied for the adoption of an open universal image file standard since the introduction of commercial HCS platforms in the mid to late 1990s. Similar to the early days of flow cytometry, commercial HCS instruments generate unique priority image file formats that are not compatible other HCS platforms or with third-party image analysis tools.

In 2010, ISAC held a special workshop for image cytometry standards that propelled early discussions about image file standards at the Society for Biomolecular Screening (SBS) and

other HCS meetings. ISAC in collaboration with key personnel and industry partners led the development of the Image Cytometry Experiment (ICE) format in 2011.⁸ The ICE format is now available as an export option for some of the commercial HCS vendors. Likewise, the open microscopy environment (OME) file format described by Jason Swedlow and colleagues has provided the HCS community with another option for exporting image files.^{9,10} The OME format and associated metadata can be exported from some of the commercial HCS instruments and is compatible with third-party image analysis tools.

Another file format that is commonly employed by almost all of the HCS imaging software platforms is the Tagged Image File Format (TIFF).¹¹ The TIFF file is a universal image file format that has “lossy” and “lossless” compression and has been widely used to register electronic image signatures for more than three decades. However, the TIFF format lacks the necessary associated header information about the instrument settings recorded during image capture, a necessary component to identify how the HCS imager captures the image. For these reasons, while the TIFF file alone can be used across multiple HCS software platforms, it is not considered suitable as a standardized HCS file format.

Although both the ICE and OME image file formats and their associated metadata are compatible with open source and commercial image analysis software products, neither ICE nor OME have been universally adopted as the HCS “image standard” by all of the commercial HCS platform manufacturers. The purpose of having the HCS standard file is uniformity, which will allow opening a single file to obtain all of the necessary information associated with how the image was captured and processed. This includes type of instrument, light source, filters, exposure times, plate parameters, well position, segmentation masks, overlays, and so on without direct dependence on associated files that can be unlinked or lost. Perhaps SBI² and ISAC in partnership with key leaders in the field and manufacturers of HCS technologies can find a solution. Nonetheless, the jury is still out as to whether ICE, OME, or another file format will be fully adopted by all of the manufacturers of HCS instrumentation and image analysis software as a means to create a single image format standard.

IMAGE ANALYSIS DATA STANDARD

Often overlooked is how imaging data are analyzed, named, and interpreted. Digital images are composed of pixels, and pixels are assigned to objects through marker-based, object-based, or contour-based segmentation. Information about objects in an image is condensed into features: intensity, color, size, shape, texture, or dynamics. Objects and regions of an image are classified into user-defined categories based on

their features or properties, or other embedded recordings extracted from the image.

Image analysis algorithms derive qualitative and quantitative measures of features. In the current state of HCS, each automated HCS platform, open-source or third-party image analysis software has a unique set of terms or descriptors for how the data are extracted from digital images and for the data features output by modules, algorithms, rule-sets, or scripts. Each HCS instrument company or developer of open-source software platform uses unique sets of data features or classifiers that are not annotated with the same terms. Measuring a cell-based fluorescent protein translocation event with image analysis algorithms from different sources may often result in different IC50 or EC50 values, even from the same image set. The problem exists because the translocation event may be measured and defined differently by distinct image analysis software. Even when multiple HCS platforms generate essentially equivalent fluorescent intensity values in different cellular compartments, one algorithm may use a ratio measurement, while another may use a difference calculation to represent translocation. In addition, the nomenclature used for objects, features, and regions of the cell have not been standardized. A compounding issue is the inability to scale HCS data at the pixel level. The lack of HCS standardization makes it difficult to both interpret and cross-reference HCS data from PubChem, published manuscripts, or other sources because it is unclear if the measurements were derived and calculated using the same methodology.

It is absolutely critical to have standardized naming conventions for objects, features, and regions of the cell, and for the procedures used to segment images to be able to understand and interpret scientific results from published work across multiple HCS platforms. The BioAssay Ontology (BAO) project is a recently published example of concepts and challenges in the standardization process.^{12,13} With the guidance and leadership of one of the principle investigators of the BAO project, Vance Lemmon at the University of Miami, SBI² will undertake to advance the development of standardized data analysis descriptors and naming conventions for HCA methods.

SUMMARY

Undoubtedly, the advancement of HCS technology with improved light sources, cameras and detectors, optics, and new reagents and optical microplates has helped generate better-quality images. This progress has reduced but not eliminated the dependence on flat-field illumination correction and background subtraction. As advancements in microscopy technology has occurred, the scientific community has embraced newer cell-based models such as 3D scaffolds, mixed culture systems,

and new fluorescent probes that require an even greater emphasis on image quality to make accurate and quantitative measurements. This is most evident when processing z-stack images captured on HCS platforms. For these reasons, it is absolutely critical that microscope-based HCS platforms are properly calibrated and operated to reduce unnecessary noise, and to optimize axial alignment of *x* and *y* horizontal dimensions, and the vertical *z* position.

The scientific community must embrace the urgent need for standardization from cells to HCS data to enhance reproducibility and reduce unnecessary economic loss. Together, the scientists and end users partnering with the manufacturing companies, colleagues, and government agencies can strive to build the necessary toolboxes and protocols to ensure the highest quality of all aspects of HCS data. While standards alone may not entirely solve the problem of irreproducibility, they most certainly will enhance the likelihood that published HCS data can be duplicated, and minimize the experimental variables that require adjustment, thus saving much time and money.

In keeping with the mission of SBI², we and the HCS community will embark on the development and implementation of the standards (Image File Format Standard, Data Integrity Standard, and the Reference Material Standard) discussed above. Moreover, there is a requirement for additional standards and guidelines for HCS best practices to assure both data quality and reproducibility. These include standardization of the nomenclature and terminology used in the “algorithms” used in image analysis processing, and of the parameters generated by these analyses, and even guidelines for data mining and the statistical analysis of HCS data. To get involved or if you would like to learn more about how SBI² is working to bring standards forward in the community, please contact us at info@sbi2.org.

DISCLOSURE STATEMENT

No competing financial interests exist.

REFERENCES

1. Editorial: *Nat Cell Biol* 2008;10:1123–1124.
2. Brazma A: Standardizing biological data. *Bioinformatics* 2001;17:113–114.
3. Freedman LP, Cockburn IM, Simcoe TS: The economics of reproducibility in preclinical research. *PLoS Biol* 2015;13:e1002165.
4. Assay Guidance Manual. Sittampalam GS, Coussens NP, Nelson H, et al. (eds.) Assay Guidance Manual [Internet]. Eli Lilly & Company and the National Center for Advancing Translational Sciences, Bethesda, MD, 2004–. www.ncbi.nlm.nih.gov/books/NBK53196/ (Last accessed on July 30, 2015).
5. Halter M, Bier E, DeRose P, et al: An automated protocol for performance benchmarking a widefield fluorescence microscope. *Cytometry* 2014;85:978–985.
6. Zucker RM, Price OT: Practical confocal microscopy and the evaluation of system performance. *Methods* 1999;18:447–458.

7. Murphy RF, Chused TM: A proposal for a flow cytometric data file standard. *Cytometry* 1984;5:553–555.
8. Image Cytometry Experiment Format. <http://flowcyt.sourceforge.net/ice/att1.pdf> (Last accessed on July 30, 2015).
9. Platani M, Goldberg I, Lamond AI, Swedlow JR: Cajal body dynamics and association with chromatin are ATP-dependent. *Nat Cell Biol* 2002;4:502–508.
10. Swedlow J, Goldberg I, Brauner E, Sorger P: Informatics and quantitative analysis in biological imaging. *Science* 2003;300:100–102.
11. Adobe Systems Inc. Tagged Image File Format (TIFF) Developer Resources. <http://partners.adobe.com/public/developer/tiff/index.html> (Last accessed on July 30, 2015).
12. Visor U, Abeyruwan S, Vempati U, Smith R, Lemmon V, Schurer S: BioAssay Ontology (BAO): a semantic description of bioassays and high-throughput screening results. *BMC Bioinformatics* 2011;12:257.
13. Abeyruwan S, Vempati UD, Küçük-McGinty H, et al: *J Biomed Semantics* 2014;5:S5.

Address correspondence to:

O. Joseph Trask, Jr.

Head, Cellular Imaging Core

The Hamner Institutes for Health Sciences

6 Davis Drive

Research Triangle Park

Durham, NC 27709

E-mail: jtrask@thehamner.com