PEER REVIEWED

Hui Jing Yu, PhD | Colin G. Miller, PhD | Dawn Flitcraft

Metrics in Medical Imaging HS

Changing the Picture

To be confident about making decisions based on medical images for individuals and for clinical trials, medical professionals can use metrics to develop adequate assurance that the images were appropriately acquired and analyzed.

HS Home Study article

LEARNING OBJECTIVE

After reading this article, participants should be able to describe how an imaging core lab partners with sponsors to use metrics to ensure the collection of quality imaging endpoint data for clinical research studies.

DISCLOSURES

Hui Jing Yu, PhD, Colin G. Miller, PhD, and Dawn Flitcraft are employees of and stockholders in BioClinica, Inc.

Medical images such as X-rays, computed tomography (CT) images, positron emission tomography (PET) images, dual energy X-ray absorptiometry (DXA), and magnetic resonance images (MRIs) are essential tools for diagnosing and monitoring diseases and directing treatments. The medical decisions based on these images are vitally important for individual patients and clinical trials as a whole. To be confident about making these decisions, radiologists and other medical professionals must have adequate assurance that the images were appropriately acquired and analyzed.

Medical imaging plays a growing role in clinical trials due to increased use of technology and improved computing power.[1] In clinical trials, medical imaging is used primarily to evaluate efficacy endpoints, and, more and more frequently, for safety evaluations and/or eligibility criteria.

Background

In multicenter clinical trials, the images will be obtained at multiple clinical sites, each with its own standard operating procedures, technologists, procedural protocols, and equipment. The experience of the technologists, the customization of each protocol, and the make and models of the equipment used may vary significantly from one site to another. Additionally, many of these trials take place over periods of time ranging from weeks to years, during which changes of personnel and equipment often occur. Image quality control is required to minimize both inter- and intra-site data variance and to ensure delivery of more precise results.

An imaging core lab (ICL) offers a full suite of medical image management solutions for the lifecycle of a trial and for a wide range of imaging modalities. Table 1 lists typical services provided by ICLs. The goal of an ICL is to unify all the essential image data in a standardized format, in order to expedite the central review of the images and data export[2–4] (see Figure 1 on generic imaging workflow).

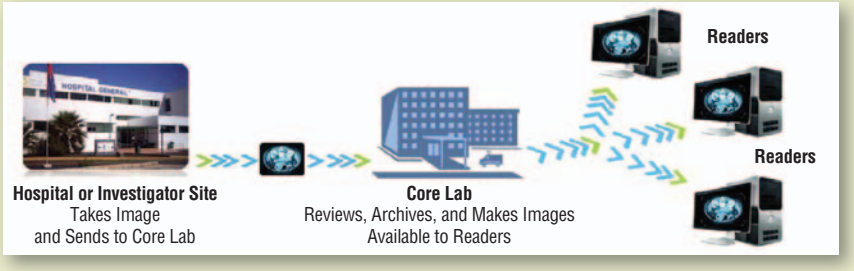
A method to track and uphold rigorous standards, as related to high image quality in a clinical trial context, is required to ensure the endpoints are clearly met. The use of imaging performance metrics to monitor image quality—so that the targets assigned to each metric are met—has therefore allowed appropriate levels of control for both the ICL and sponsors, thereby enhancing trial performance and quality.

Essentially, there are two major types of imaging from a quality control (QC) viewpoint:

Table 1 Comparison of Image Core Lab Services

Study Initiation and Startup	Collection Management	Independent Review
• Identify expert readers and consultants • Assign project team • Engage study startup team • Design imaging protocol • Communication plans • Project-specific work instructions • Develop imaging review charter • Deploy site surveys • Attend investigator meetings • Provide imaging study kits • Perform site visits • Conduct web-based training	• Collect image data • Query sites for missing data • Translate/digitize image data • Image quality assurance/ quality control • Image data query resolution • Archive for long-term storage	• Analyze images • Design independent read system • Develop imaging review charter • Provide reader training • Conduct independent read • Monitor independent read • Monitor inter-/intra-reader variability • Export data

Figure 1 Imaging Core Lab Process Workflow



- two-dimensional (2-D) (e.g., plain film X-ray, DXA, and ultrasound) and
- three-dimensional (3-D) or tomographic techniques (e.g., CT, MRI, and PET).

The QC for 2-D imaging is more critical on positioning, since slight rotation or incorrect positioning may hide important anatomic features. The 3-D techniques tend to need more QC on the acquisition settings and review of patient motion, since the acquisition times are longer. Image QC primarily consists of checks on correct positioning, complete anatomical positioning, lack of patient motion, and a check on the correct acquisition or instrument settings, such as the scan mode (e.g., T1 or T2, etc.) in MRI or scan thickness and coning in CT.

Image Quality Metrics

Within the lifecycle of an imaging trial, trial performance can be tracked using four types of metrics: cycle time, timeliness, quality, and efficiency/cost (Figure 2 and Table 2). Quality metrics can be further tracked as image quality (metrics as determined by reader or independent reviewer, although images are checked for quality at the technologists' level before sending to the reader), image queries sent to sites, missing imaging visits, and adherence to acquisition protocol.

A key first step to ensure that high-quality imaging endpoint data are collected for studies is to have standardization of image acquisition between sites. This can usually be accomplished by providing training to each site via a group location, telephone, Web conference, etc. Occasionally, site visits (visits to educate the technologists) are performed if the protocol is deemed to be more challenging than the standard-of-care procedure, or if the sites are not adhering to the imaging guidelines. Imaging guidelines are provided to the site simply to communicate and document the image-related expectations and requirements for a trial.

On an ongoing basis, data arrive at the ICL and are inspected for image quality, usually by radiological technologists, prior being sent for the radiological evaluation or read. The reader can then determine the presence or absence of necessary imaging and the associated image quality.

Image quality metrics can be calculated based on the percentage of images that are readable (evaluable), suboptimal (readable but not optimal),

Figure 2 Imaging Performance Flowchart

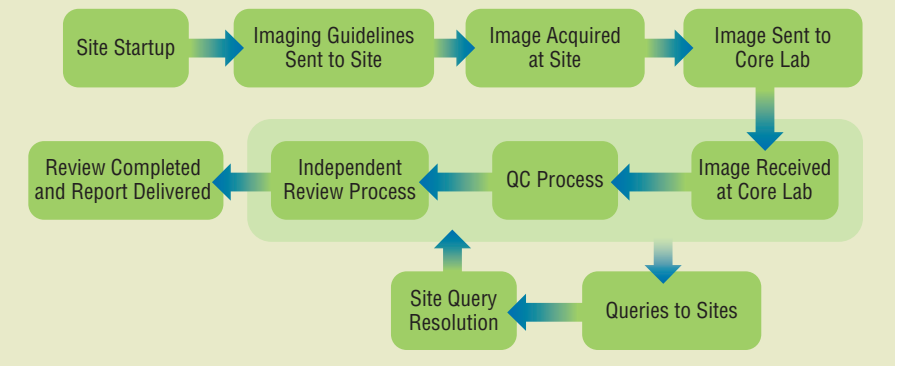


Table 2 Imaging Core Lab Performance Metrics

Metric	Category	Metric Title	Unit of Measure	Reporting Frequency	Target
1	Project Startup	Average number of days study kit sent	Turnaround time - Days	Monthly	3
2	Project Startup	Completion of site qualification/training	Percentage (%)	Monthly	95%
3	Project Startup	Independent review charter	Turnaround time - Days	Monthly	5 days from receipt of latest draft/final protocol
4a	Image Acquisition	Average number of days from image time point acquisition to receipt	Turnaround time - Days	Monthly	3 (electronic transfer)
4b	Image Acquisition	Average number of days from image time point acquisition to receipt	Turnaround time - Days	Monthly	7 (traditional transfer)
5a	Image Acquisition	Average number of days from image receipt to initial feedback sent to site	Turnaround time - Hours (eligibility/safety)	Monthly	24 hours
5b	Image Acquisition	Average number of days from image receipt to initial feedback sent to site	Turnaround time - Days (standard study)	Monthly	3 days
6	Image Processing	Average number of days from image receipt to ready for independent review	Turnaround time - Days (standard study)	Monthly	3 days
7	Image Processing	Average number of days from when the image is designated for review to completion of the review, excluding images which have outstanding queries	Turnaround time - Days (standard study)	Monthly	Variable
8	Quality	Percentage of non-evaluable images vs. total images received	Percentage (%)	Monthly	≤ 3%
9	Quality	Percentage of non-evaluable/missing baseline images	Percentage (%)	Monthly	≤ 2%
10	Quality	Quality of data export	Percentage (%)	Monthly	99%
11	On-Time Delivery	On-time delivery of read report(s)	Percentage (%)	Monthly	98%
12	On-Time Delivery	On-time delivery of data export(s)	Percentage (%)	Monthly	98%
13	On-Time Delivery	On-time delivery of FINAL data export	Percentage (%)	Monthly	99.9%
14	Image Queries	Percentage of Queries	Percentage (%)	Monthly	< 10%
15	Image Queries	Average number of days queries outstanding	Turnaround time - Days	Monthly	7

or not readable both by the technologist and the readers. This can obviously be evaluated on the study level, but also on country- and site-specific levels.

If there are issues, a query can be generated and sent to the site for immediate resolution. The percentage of site queries is a performance metric that captures rate of issue as an indication of whether or not the site training addressed the necessary key points for acquisition and how closely the protocol was being followed.

When a query is unresolved, or imaging cannot or could not be performed, or protocol is not followed, the result is missing imaging data for either the baseline visits or nonbaseline visits. Such metrics can be defined and tracked throughout the trial, allowing for early escalation of potential site performance or study protocol design issues.

Lastly, the number of image acquisition technique–related amendments, upon the agreement between the ICL and sponsors, could be incorporated as a metric as an indirect measurement of image quality (e.g., the greater the number of amendments, the lesser the robustness of acquisition protocol and quality).

Two Case Studies

What follows are two case studies presented as examples of how ICLs use metrics to ensure that high-quality imaging data are collecting for studies (see Table 3 for a summary).

Case 1

Many ICLs have started to include imaging performance metrics as part of their standard reports. Implementing a set of standardized metrics can allow the early escalation of potential core lab or site performance issues that require immediate remediation and identification of any need to retrain

Table 3 Case Study and Metrics Summary

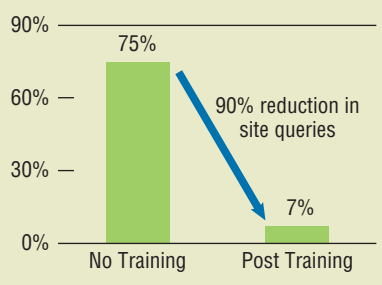
Case Study	Category	Metric Title	Target
1	Image Queries	Percentage of site queries	< 15%
2	Image Quality	Percentage of non-evaluable baseline images	< 2%
2	Image Quality	Percentage of non-evaluable images (non-baseline)	< 2%

sites or redesign image acquisition guidelines.

One example of this approach is the case where a client considered an X-ray procedure to be so simple and straightforward that site training was not requested, and it was assumed that a paper manual would suffice.[5] Unfortunately, that decision resulted in a data clarification form (DCF) rate (site queries) of approximately 75%, which caused both a lack of precision and loss of time due to requests for repeat procedures, and directly translated into poor data and increased costs.

The sponsor quickly produced a CD-based training program[6] for this study, including a test to ensure understanding of the material, and requested that the sites have the appropriate personnel take the program. The result after sites completed the CD training was a 90% decrease in the DCF rate to less than 7%. This training format provides an excellent, cost-effective way to ensure protocol compliance while improving the precision of study data. This, in turn, either improves overall statistics and/or shortens the time required to detect significant change, thus reducing overall cost for sponsors.

Figure 3 Case 1 Exhibit: Post-Training Result in Reduction of Site Queries



Case 2

The integration of the ICL as part of clinical trials in all therapeutic areas where medical images are collected is particularly important to harmonize data quality across sites. For example, in an oncology trial involving more than 30 countries and 200 sites across the globe, it was challenging to obtain high-quality, standardized data due to varying technical capabilities at hospitals and imaging facilities. Other sources of variations included study duration and the imaging modalities involved.

In this case, the study lasted for six years, with CT and MRI data being collected for all time points. At screening (baseline), all subjects were required to have a bone scan (nuclear medicine image), which was to be repeated at follow-up time points if disease was present at baseline or if clinically indicated.

Because an ICL was involved in the study, the sites received standardized instructions (image guidelines) at the start of the study and, for the most part, the imaging quality was comparable across sites. However, many imaging protocols, including the one for this study, provide high-level requirements for imaging and do not include the necessary level of detail that is realistically needed.

In this study, 95% of total data submitted for the study was digital, with only 5% of film data submission. Of the film data submitted, 90% were bone scans, because it was challenging for some sites to provide technically adequate digital images in the correct format. Instead, images submitted digitally were in JPEG format with improper leveling and windowing, resulting in images with a lack of details.

Considering that bone scans were required for all subjects at screening and these images were used to determine subject progression, the problems encountered with film data could have resulted in a much lower than ideal rate of readable images for follow-up imaging time points. However, through tracking of the relevant metrics, the ICL was able identify these issues early and, together with the sponsor, worked with the regional monitors and sites to find locations where subjects could receive bone scans that were acceptable and usable for the study. Furthermore, the ICL provided extra training to optimize contrast for screenshot images.

Such implementation resulted in excellent submission turnouts as measured by image quality metrics. Out of 789 baseline time points, only one was not readable (i.e., 99.87% baseline images readable), and out of 4,810 follow-up time points, only 40 were not readable (i.e., 99.17% non-baseline images readable). Most of the scans that were considered not readable were caused by missing anatomy.

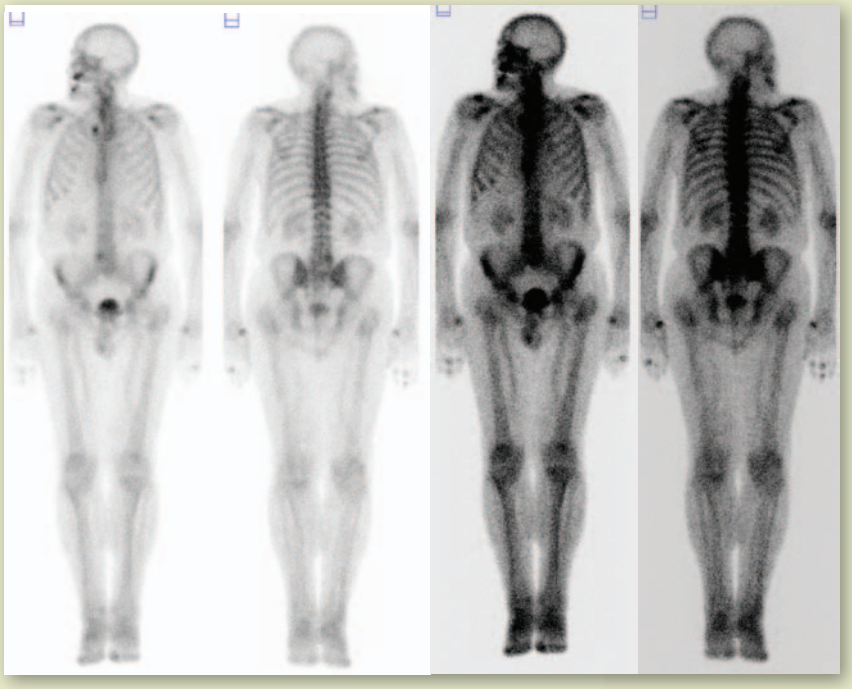
Thus, if the ICL is not involved at the start and standardized guidelines are not provided, then studies can run into data quality issues that might otherwise be avoided.

Electronic Image Submission

Overall, submission of images via electronic means reduces the transit time from the site by greater than 80% from traditional means (courier). This is achieved by avoiding customs involvement when moving the package in and out of countries, as well as any other human or weather involvement, which could delay the shipment from the site to the ICL.

Electronic submission is the quickest way to submit images to the ICL and enables the site to mask the image data

Figure 4 Case 2 Exhibit: Good Quality Bone Scan (Left Set) vs. Poor Quality Bone Scan (Right Set)—Lesions Not Visible on Poor Quality Scan



before transmitting to the ICL. Also, electronic submission is spotlighted in the training materials and at the investigator meetings because this is a great solution to a constant challenge for all clinical trials. Even with minimal setup required for electronic image transmission programs, some sites continue to send image data via courier. This may be out of habit because firewalls and other technology hurdles are typically not an issue at the sites. The ICLs should work very closely with the CRO to encourage sites to use electronic image transmission programs.

Conclusion

The use of medical imaging in clinical trials has developed from the early days of passively collecting images and having them evaluated on light (film) boxes by radiologists. Improvements in the related technology over time have greatly increased the ability of medical experts to use imaging as a critical biomarker, whether for eligibility, safety, or efficacy. The practice now stands as its own major scientific pursuit as well as a focus for operational logistics management.[7]

The critical use of metrics has helped empower this progress; metrics in their own right are of little value, unless they can effect change to a process. The examples provided here have demonstrated the value of metrics for contributing to the ICL operational capabilities and, ultimately, for providing improved study outcomes through decreased variability in data, leading to greater statistical confidence in study findings. Greater statistical confidence will ultimately lead to a decreased number of patients in future trials.

Finally, the ethical and financial implications of using appropriate metrics should not be underestimated.

References

1. Reiber H, van Kuijk C, Schwarz L. 2005. Medical imaging and its use in clinical trials. *European Pharmaceutical Contractor* Autumn 2005: 80–4.
2. Miller CG, Noever K. 2003. Taking care of your subject's image: the role of medical imaging core laboratories. *Good Clinical Practices Journal* 10(9).
3. Miller CG. 2005. Medical imaging core laboratories. *Applied Clinical Trials* October 2005.
4. Miller C, Noever K. 2003. Scanning data for clinical information. *Good Clinical Practices Journal* 10(9): 21–4.
5. Pearson D, Miller CG. 2007. *Clinical Trials in Osteoporosis*. Springer Verlag.
6. Miller CG, Wertz K. 2007. Education alternatives for imaging techniques in clinical trials. *World Pharma Network* July: 42–3.
7. van Meurs P, Miller CG. 2011. Image conscious: new FDA guidance. *Samedan Ltd* December: 24–8. **ACRP**

Hui Jing Yu, PhD, is a medical affairs scientist at BioClinica, Inc., where she provides scientific and medical support to the pharmaceutical industry on the use of imaging biomarkers in clinical trials. She also provides support to internal business development, marketing, and operations teams. She holds a bachelor's degree in biomedical and electrical engineering, an MSc degree focused on physiology and biophysics research, and a PhD in biomedical engineering from Stony Brook University in New York. She has written and coauthored several scientific publications and, as the primary author, she drafted and reviewed this manuscript. She can be reached at huijing.yu@bioclinica.com.

Colin G. Miller, PhD, is senior vice president for medical affairs at BioClinica, Inc., where he is responsible for medical and scientific consulting. He joined BioClinica (formerly Bio-Imaging Technologies, Inc.) in 1999 as vice president of business development. He has also served as director of clinical services at Bona Fide (a company he started in 1994 that was later acquired by Bio-Imaging Technologies, Inc.); and as the head of the physical measurements team for Europe at Procter & Gamble Pharmaceuticals. A fellow of the Institute of Clinical Research, he also is an associate member of the Radiological Society of North America, a member of the American Society of Bone and Mineral Research, and a member of the Metrics Champion Consortium. He has written and coauthored more than 40 scientific publications. He received his bachelor's degree in physiology and zoology from the University of Sheffield and a PhD from the University of Hull, both in the U.K. As a coauthor, he critically reviewed this manuscript.

Dawn Flitcraft is senior vice president for client services at BioClinica, Inc., where she oversees the project management, imaging core lab, and clinical operations departments of the Medical Imaging Solutions Division and is responsible for overall client relations. She joined BioClinica as director of project management when it acquired Quintiles Intelligent Imaging in 2001. She held several positions at Quintiles, including image processing specialist, senior manager, and finally director of clinical research and development. She holds a bachelor's degree in biology and nuclear medicine from Cedar Crest College in Allentown, Pa.; has certifications from the Nuclear Medicine Technology Certification Board, the American Registry of Radiologic Technologists (Nuclear), and the American Registry for Diagnostic Medical Sonography; and is a member of the Steering Committee for the Metrics Champion Consortium. As a coauthor, she critically reviewed this manuscript.