

CHAPTER

1

INTRODUCTION AND BACKGROUND

Robert J. Nordstrom and Janet F. Eary

Nordstrom, R. J. and Eary, J. F., "Introduction and background," in *Quantitative Imaging in Medicine: Background and Basics*, edited by R. J. Nordstrom (AIP Publishing, Melville, New York, 2021), pp. 1-1-1-8.

GETTING STARTED

Today, disease characteristics visualized in an image are used to create a clinical differential diagnosis and suggest a prognosis. Many choose to trace the earliest beginnings of clinical imaging to the x-ray picture Wilhelm Roentgen took of his wife's hand in 1895, using the newly discovered x rays. The discovery of x rays was a remarkable feat, worthy of the first Nobel Prize in physics, and many present-day medical imaging devices are based on the use of x rays to diagnose abnormalities within the human body. Medical imaging, however, has moved well beyond the operation of placing the patient between an x-ray source and a film plate for a prescribed period of time. This simple procedure continues to have an important role in the clinical setting, but much of clinical imaging now involves a dynamic image acquisition process with complex data collection, sophisticated data processing, and image display. Each of these steps must be performed with proper precision before image data are generated. In addition, each step is an opportunity for information noise to enter the process, degrading the quality of the final image.

The history of medical imaging has been the topic of many texts (Bradley, 2008; Hendee and Ritenour, 2014; Kevles, 1997; Scatliff and Morris, 2017). The second half of the 20th century has seen a steady march toward more sophisticated and complex imaging methods capable of providing 3D images of organs and tissue, functional images of metabolic processes, and insights into tissue diffusion and perfusion processes. Our understanding of disease complexity has grown, in large part, due to advances in imaging.

Through its development, imaging interpretation has remained largely qualitative. The medical practice of the radiologist is to interpret image information. Therefore, high visual quality in images is essential, so manufacturers of imaging devices such as computed tomography (CT), positron emission

tomography (PET), magnetic resonance, and ultrasound devices strive to provide high-quality medical images for radiologists to examine and interpret visually.

The advent of digital imaging has helped to usher in the era of quantitative imaging. Although certain quantitative functions can be performed on analog images, availability of digital image data has enabled a wealth of quantitative methods to extract image information. The chapters in this book and *Applications and Clinical Translation* (Nordstrom, 2021) demonstrate the depth and diversity of quantitative imaging today.

1.1 WHAT IS QUANTITATIVE IMAGING?

Many definitions of quantitative imaging have been created. The Radiological Society of North America (RSNA) organized an opportunity to bring together a group of scientists, radiologists, and clinicians for the purpose of advancing quantitative imaging in the clinical environment. The result is the Quantitative Imaging Biomarker Alliance (QIBA), organized in 2007 and committed to transforming patient care by making radiology a more quantitative science. This involves creating methods and standards that will achieve accurate and reproducible quantitative results from medical images. QIBA has defined quantitative imaging as “the extraction of quantifiable features from medical images for the assessment of normal or the severity, degree of change, or status of a disease, injury, or chronic condition relative to normal” (Sullivan *et al.*, 2015, p. 814). The important feature of this definition is to recognize that quantitative imaging is a *process*. It is not simply a digital image. It is the complete process of collecting the data with standardized methods from qualified imaging devices, processing the data with validated algorithms, and displaying the final image in a standard format. It forms the basis for the fields of artificial intelligence (AI) and deep-learning, statistics, standards, quality control and assurance (QC/QA), and many more.

Quantitative imaging is more than a single process. It is often a group of procedures, each with specialized functions and variables. Quantitative imaging procedures are often referred to by the developer community as “tools.” These tools can be used separately to extract specific information from a medical image, or they can be used in a pipeline, with each tool building on the information extracted from the output of the previous tool. Just as different imaging devices have specific functions that qualify that modality for certain imaging functions, quantitative imaging tools have different purposes for extracting image information.

Quantitative imaging procedures are designed to fit clinical imaging needs. For example, lung nodules or liver lesions whose response to therapy is a reduction in size can be measured using a simple linear measurement across a tumor mass diameter in an image. The simple linear tumor mass diameter measure is currently the input into the RECIST schema (Response Evaluation Criteria for Solid Tumors), which has been used in clinical trials for patient treatment response assessment since 2000. Recently, radiomics approaches have begun to play a role in image analysis for tumor characterization

and treatment response assessment. Image-derived parameters with biological significance, such as tumor size, location, and heterogeneity, and parameters derived from image feature content, such as spatial wavelet coefficients, texture, and voxel information density gradients, are extracted from images and compared with patient outcome to determine prognosis in many applications. In another example, the extent of brain tumor infiltration into surrounding normal tissue is an important contributor to treatment decision. Interpretation of magnetic resonance spectroscopy data in brain tumors can be used to determine the spatial extent of tumor margin infiltration, with the use of tumor metabolite ratio measures, such as choline and NAA (N-acetyl aspartate).

Image-based quantitative measures qualify as tumor biomarkers when properly validated for specific applications. A biomarker is defined as a “characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathologic processes, or biological responses of a therapeutic intervention” ([Biomarkers Definitions Working Group, 2001](#), p. 91). The value of quantitative image data in the clinical environment is based on its reliability and reproducibility. For a quantitative biomarker, reliability is the degree to which the result of a measurement, calculation, or specification is accurate. It is measured as “the ratio of results variance between subject measures of total variance based on the observed measurement” ([Kessler et al., 2015](#), p. 13). Reproducibility is the extent to which consistent results are obtained when an experiment or measurement is repeated. This value represents the “measurement precision under a set of repeatability conditions of measurement” ([Kessler et al., 2015](#)). There is a difference between these two biomarker characteristics: a measurement can be highly reproducible without being reliable. For example, consider making a length measurement with a ruler that is not the correct length. The measurement can be consistently the same (reproducible), but it will not be the correct length (reliable). This kind of bias in the measurement can also be found in quantitative imaging output measurements.

Simply extracting measurable information from medical images does not guarantee a reliable and reproducible biomarker. Even with the utmost precision taken in a measurement, reliability and reproducibility can be elusive. Quantitative imaging requires dedication to standards and protocols during all aspects of data collection and processing. The final measurement from any medical image can be compromised by lack of attention to detail during the data collection and processing steps. As a result, a great deal of quantitative imaging effort must be placed in the process of actually collecting and processing the final image, before actual information extraction begins. In addition to being both reliable and reproducible, a quantitative imaging biomarker must be sufficiently sensitive to the pathophysiological conditions being imaged. For example, if the imaging is included in a clinical trial to measure disease response to therapy, then the biomarker must be sensitive to the range of parameter change anticipated from the therapy. This requirement places constraints on the quality of the biomarker and on the biomarker measurement methods. Methods with excess bias (systematic error) or variance (random error) will not be sufficient to yield a reliable and reproducible biomarker. A biomarker, according to the definition, must be a measurable quantity, as free from hardware errors as possible, and also free from observer bias and judgment. Therefore, specific requirements for imaging biomarker validation must be determined and applied to the method or methods.

1.2 STANDARDS

A biomarker is a surrogate to an underlying disease process. Not only must its measurement be reliable and reproducible on any specific imaging device, but it must also be universally valid. That is, it must be invariant across different devices and operators. Variability in the measurement of a biomarker must be reduced as much as possible to ensure this invariance. Variability in the form of biological variability among patients, data collection variability from different devices and operators, and variability in interpretation of the final images must be considered. Petrick *et al.* in Chap. 12 of Nordstrom (2021) discuss these sources of variability in detail.

The presence of these sources of error creates the requirement for standards and quality control in all aspects of image collection, processing, and display. Performance and reference standards are concepts central to quantitative imaging. Without standards in every aspect of image acquisition, reconstruction, results generation, and image display in cancer imaging, it would be difficult to determine if changes in tumor characteristics are caused by therapy effects, or variations in image hardware or software performance. When measuring tumor response to therapy in a clinical trial, the observation of a small difference in an image-derived parameter, such as tumor size or perfusion characteristics, can be an early indicator of response, indicating a better patient treatment outcome. In adaptive treatment clinical trials, small changes in tumor parameters can be the deciding factor for whether the patient remains on the current therapy regimen or is switched to a different course. Without image processing standards, system bias and variance effects could be larger than the tumor changes from treatment, jeopardizing response assessment.

It may be idealistic to state that quantitative imaging is applied to enable clinical imaging devices to perform as calibrated measuring instruments. By reducing the sources of error in imaging device output and creating harmonization among identical imaging devices in a modality, the goal of achieving calibrated output performance is many steps closer to this ideal. Therefore, standards in all aspects of imaging, from data collection to final image presentation, are important for clinical decision making.

1.2.1 Standard operating procedures

Standard operating procedures (SOPs) are important aspects of clinical studies. Their purpose is to guide clinical patient care and research to comply with good clinical practice (GCP). This is an international guideline for ethics and scientific quality for designing, conducting, and reporting clinical trials. Published by the International Conference on Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH), these guidelines cover every aspect of clinical trials. Specific SOPs are created by organizations to define the procedures by which GCP will be followed. These can define procedures for instrument use, how clinical trial subjects are selected to achieve statistically significant results, the time between imaging procedures, and dozens of other parameters.

Device SOPs are written to guide personnel in the operation of specific devices for specific purposes. For example, a clinical CT device will have specific procedures for calibration, operation, and maintenance specified. If the scanner is selected for use in a multicenter clinical trial, additional procedures might be imposed to ensure that the output from this particular device complies with output from other devices in the trial. These additional procedures must be carried out under GCP.

In clinical practice, constant performance assessment of the imaging devices is also important. While image acquisition and interpretation methods may differ among imaging facilities, this variability may have little or no impact on patient diagnosis or medical practice. In a clinical trial, however, variability in imaging procedures may result in increased output data variability that can adversely affect the ability of the trial to achieve its desired endpoints. Trial-specific device and procedure performance standards might be adopted in the clinical trial design to minimize this problem ([Kennan *et al.*, 2016](#)).

1.2.2 Site qualification, phantoms, and digital reference objects

Site qualification is an important step in any clinical trial. In trials using imaging data, site qualification usually involves two parts. The first involves imaging device qualification for the imaging tasks required in the trial. This can be performed as a part of site selection for a multisite trial or performed in advance of any trials, so the clinical trials organization can catalog sites for future trials participation. An example of the latter is provided in [Rosen *et al.* \(2017\)](#). The National Cancer Institute (NCI) has supported efforts to establish a clinical trial-ready imaging resource within the NCI-designated Cancer Centers capable of conducting quantitative imaging in treatment trials. Three different clinical imaging modalities have been qualified: PET, CT, and magnetic resonance imaging (MRI).

Clinical imaging devices use complex reconstruction algorithms to create images from a stream of raw data. The raw data and process must be checked and calibrated regularly according to manufacturer specifications and operating procedures specified by hospital and clinic oversight organizations, such as JCAHO (Joint Commission on Accreditation of Hospitals Organization). This encompasses the second part of clinical trial site qualification. Most of these procedures are sufficient for an imaging institution to qualify as a clinical trial site; however, use of imaging phantoms may be a part of site qualification. These are physical objects used as substitutes for human subjects in clinical imaging devices. They are created in a variety of sizes and shapes, depending on the use to which they are intended, and they are used to calibrate the output characteristics of imaging devices for harmonization between devices and sites. Different phantoms are customized with known imaging characteristics, to calibrate specific imaging systems. They are available commercially from a number of vendors. Site SOPs dictate which phantoms and how often each should be used to keep the imaging devices and software in calibration. Often, traceability to a national or international technical standard is required. The National Institute of Standards and Technology (NIST) develops and disseminates imaging phantoms for all clinical scanner modalities as a part of its standards mission.

Digital reference objects (DROs) are digital phantoms. They are software constructs of specific information and are used as input for simulation studies. They test the abilities of image output software packages to provide the correct information. Objects in the DRO, such as spheres or voids, are described mathematically, and noise of different amounts and frequency distribution can be added if necessary. Partial volume effects can be simulated within PET or CT reference objects. In PET DROs, for example, standard uptake value (SUV) levels can be set to determine the ability of various software packages to generate the correct values.

1.2.3 Testing standards

The example of bias in length measurement given earlier demonstrates that there is no way to know that the bias in an image dataset exists except by imposing a set of standard quality control procedures on the measurement instrument (in this case, the ruler) to discover and correct the bias. For example, the ruler could be measured against the International System (SI) of Units (where the meter is defined as the distance light travels in a vacuum in $1/299\,792\,458$ seconds). For medical imaging devices, phantoms with NIST traceability can discover bias in imaging systems.

However, variance (random errors) also exists. To measure the variance, test–retest studies are usually performed. Use of phantoms (NIST traceable or not) in multiple identical image acquisitions can discover variance in device output. Test–retest studies determine if the device can reliably replicate a result more than once. In addition, test–retest studies can determine the robustness of algorithms used to construct or analyze medical images, and are useful in determining inter-reader reproducibility in clinical trials. Exact methods and statistical analyses are beyond the scope of this book, but the importance of test–retest as a method to demonstrate reduction of error in quantitative imaging methods as standards are applied must be noted.

There is also a source of error introduced by the biological variability between human subjects. Not only are there significant inter-subject differences, such as weight, gender, age, alcohol use, smoking history, but studies show that there can be intra-subject differences depending on the interval length between scans for an individual. Proper SOPs and selection criteria for subjects in clinical trials can help to control the variations that might occur due to these biological differences. The procedures associated with quantitative imaging will not minimize or eliminate potential biological variability, but they will permit those variations to be measured without the addition of device variability.

1.3 IS QUANTITATIVE IMAGING NEEDED?

With the additional effort in maintaining adherence to SOPs and any other standards that might apply, it is reasonable to ask if quantitative imaging adds value to a clinical trial. For the answer, the features and benefits of quantitative imaging must be weighed against the costs of additional SOP and standards applications. The features of quantitative imaging are the topics of the chapters this book and *Applications and Clinical Translation* (Nordstrom, 2021). They involve the algorithms for segmenting

regions of interest in the image, combining images to create 3D versions of disease locations, extracting information from these regions for comparison with patient outcome, etc., that form the basis for predictive models, methods of harmonization of an imaging modality, and many more. The benefits provided by quantitative imaging lie in the contribution this image output provides to patient care and the insights offered in new treatment evaluations.

The addition of quantitative image-based measures to qualitative interpretations can contribute precise information on disease characteristics. This information, along with patient data, can be stored in a database for future reference. It can also be used to help resolve differences in image interpretation and add the value of quantitative imaging biomarker results to patient management.

New treatment clinical trials require extensive validation of the safety and efficacy of a drug before FDA approval. For image data contributions to the new treatment applications, high standards for imaging method sensitivity and specificity must be demonstrated. The number of subjects required to complete a clinical study may be affected when the measurement of a critical image-based biomarker is influenced by statistical or instrument noise. This adds time and cost to the treatment testing and approval process. By reducing the variation in the image biomarker measurement process through quantitative imaging procedures, the effort and cost of new treatment approval can be reduced considerably.

With the many benefits of quantitative imaging, what are the impediments to its acceptance? This question is very difficult to answer because, as this book and *Applications and Clinical Translation* (Nordstrom, 2021) will show, quantitative imaging is a host of standards applications, algorithms, image features, AI processes, and more. It is safe to say that today, all of these aspects of quantitative imaging are not a part of the current clinical radiology procedures. It is equally safe to state that almost every clinical trial center in the United States uses some form of quantitative imaging to determine disease progression or treatment response. RECIST is probably the most frequently used method of extracting a measurable quantity from images. The uses of radiomics and AI are increasingly being described in the research, but few, if any, have yet to be shown to have clinical trial utility.

The reluctance of the imaging community to adopt quantitative imaging processes in clinical practice is discussed by [Miles et al. \(2018\)](#). They conclude that, in the case of the assessment of tumor heterogeneity, imaging practitioner engagement is a barrier to effective translation of quantitative imaging assessment of this tumor characteristic. This finding is not unexpected, but as new imaging specialists become familiar with the benefits of quantitative imaging, translation of quantitative imaging tools into clinical utility will increase.

REFERENCES

Biomarkers Definitions Working Group, "Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework," [Clin. Pharmacol. Ther.](#) **69**, 89–95 (2001).

- Bradley, W. G., “History of medical imaging,” *Proc. Am. Philos. Soc.* **152**(3), 349–361 (2008), see www.jstor.org/stable/40541591
- Hendee, W. R. and Ritenour, E. R., *Medical Imaging Physics*, 4th ed. (Wiley-Liss, 2002).
- Kennan, K. E., Peskin, A. P., Wilmes, L. J., Aliu, S. O., Jones, E. F., Li, W., Kornak, J., Newitt, D. C., and Hylton, N. M., “Variability and bias assessment in breast ADC measurement across multiple systems,” *J. Magn. Reson. Imaging* **44**(4), 846–855 (2016).
- Kessler, L. G., Barnhart, H. X., Buckler, A. J., Choudhury, K. R., Kondratovich, M. V., Toledano, A. Y., Guimaraes, A. R., Filice, R., Zhang, Z., Sullivan, D. C., and QIBA Terminology Working, “The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submission,” *Stat. Methods Med. Res.* **24**(1), 9–26 (2015).
- Kevles, B. H., *Naked to the Bone: Medical Imaging in the Twentieth Century* (Rutgers University Press, New Brunswick, NJ, 1997).
- Miles, K. A., Squires, J., and Murphy, M., “Potential barrier to the clinical translation of quantitative imaging for the assessment of tumor heterogeneity,” *Acad. Radiol.* **25**(7), 935–942 (2018).
- Nordstrom, R. J., *Quantitative Imaging in Medicine: Applications and Clinical Translation* (Melville, New York: AIP Publishing, 2021).
- Petrick, N., Li, Q., Gavrielides, M. A., and Delfino, J., “Analytical and clinical validation,” in *Quantitative Imaging in Medicine: Applications and Clinical Translation*, edited by R. J. Nordstrom (Melville, New York: AIP Publishing, 2021), pp. 12-1–12-34.
- Rosen, M., Kinahan, P. E., Gimpel, J. F., Opanowski, A., Siegel, B. A., Hill, G. C., Weiss, L., and Shankar, L., “Performance observations of scanner qualification of NCI-designated cancer centers: Results from the centers of quantitative imaging excellence (CQIE) program,” *Acad. Radiol.* **24**(2), 232–245 (2017).
- Scatliff, J. H. and Morris, P. J., “From röntgen to magnetic resonance imaging: The history of medical imaging,” *NC Med. J.* **75**(2), 111–113 (2014).
- Sullivan, D. C., Obuchowski, N. A., Kessler, L. G., Raunig, D. L., Gatsonis, C., Huang, E. P., Kondratovich, M., McShane, L. M., Reeves, A. P., Barboriak, D. P., Guimaraes, A. R., Wahl, R. L., and RSNA-QIBA Metrology Working Group, “Metrology standards for quantitative imaging biomarkers,” *Radiology* **277**(5), 813–825 (2015).