Translational Oncology
www.transonc.com

Volume 2 Number 4 December 2009 pp. 195–197 **195**

Quantitative Imaging for Evaluation of Response to Cancer Therapy

Laurence P. Clarke, Barbara S. Croft, Robert Nordstrom, Huiming Zhang, Gary Kelloff and J. Tatum

Cancer Imaging Program, NCI, NIH, Bethesda, MD, USA

Advances in molecular medicine offer the potential to move cancer therapy beyond traditional cytotoxic treatments to safer and more effective targeted therapies based on molecular characteristics of a patient's tumor. Within this context, the role of quantitative imaging as an *in vivo* biomarker has received considerable attention as a means to predict and measure the response to therapy. For example, the ability to predict the response to therapy quantitatively, early in the drug or radiation therapy regime, would facilitate adaptive therapy trial strategies, that is, that permit alternative treatment regimens in cases where initial therapy response was ineffective. Similarly, the ability to measure the response to therapy should provide a more robust means for both therapy dose management and correlation of imaging results with other laboratory biomarkers. The latter is required for clinical decision making in the clinical setting. The National Cancer Institute (NCI) in collaboration with the Food and Drug Administration (FDA) has therefore promoted a number of initiatives supporting the role of molecular imaging in drug trials. The major goal of these initiatives is the "qualification" of the proposed molecular imaging protocol(s) that can be incorporated into current or future drug trials submitted to the FDA. Clinical research strategies that will help achieve these goals are described in the published literature [1–5].

The physical uncertainty in the measurement of response to therapy as applied to biomarker trials is also important because it may negatively influence the ability to detect an early response. While we expect biologic variability in the response rate of individuals, we also expect the physical uncertainty resulting from data collection and analysis methods to be quantified such that we can power a trial to detect a given biologic effect size across individuals. Such sources of error are often dependent on the selected imaging platform and imaging protocol. As a result, there is a critical *a priori* need to characterize the physical measurement uncertainty across different commercial imaging platforms before the implementation of multisite, multiplatform clinical trials. This characterization is necessary to determine the level of random (noise) and systematic (bias) error in the measurement system that impacts the minimum change that can be measured. In other words, if we use appropriate, quality-controlled imaging platforms and optimal low-noise, unbiased data analysis methods, physical measurement uncertainty can be decreased, thus allowing the required number of individuals used in trials to be decreased while maintaining the same power to detect a given biologic response. This information is especially critical if imaging is to be used to support adaptive trial designs. In addition, it may be possible, depending on the commercial imaging platforms involved and clinical protocol, to harmonize data collection and analysis across the different imaging platforms used in multicenter clinical trials, reducing, for example, the uncertainty in data correlation with other laboratory biomarkers, as required for clinical decision making [6–9].

NCI is currently addressing the physical measurement uncertainty problem for imaging as a biomarker using several synergistic and complementary initiatives. The goal is to engage all of the important stakeholders in the imaging field (academia, imaging and pharmaceutical industries, scientific imaging societies, and the different agencies of the federal government) to first explore the development of a broad consensus on how to develop quantitative imaging methods and then to encourage the adoption of more standardized methods for quality assurance and quantitative imaging by the academic and industry communities as a long-term goal. These initiatives include the following:

1. Reference Image Database to Evaluate Response to Therapy (RIDER): The RIDER project, established as a pilot project in 2005, refers to the development of Web-accessible public resources of image data, designed to permit the comparison of different imaging and analysis methods for the measurement of therapy response, archived on the NCI imaging archive, described below [10,11].
2. CaBIG Imaging Workspace and NCI Image Archive: The goal of this initiative, supported by NCI caBIG, is to develop an open computer architecture that supports interoperability of database and software tools and to provide a comprehensive image archive to store and query these data, meeting all DICOM and caBIG requirements. A secondary goal is to permit an environment where plug and play software tools can be used to help develop common approaches for data collection and analysis, thus supporting quantitative imaging standards for imaging as a biomarker [12,13]
3. Quantitative Imaging Network (QIN found in PAR-08-225): The QIN project was recently published in 2008 as a program announcement. The goal of this initiative is to support multidisciplinary research teams to develop quantitative imaging methods to measure the response to therapy, using current commercial imaging platforms. The teams will then optimize the performance of their quantitative imaging tools within ongoing phase 1 to 3 clinical therapy trials. A secondary goal is to provide

Address all correspondence to: Laurence P. Clarke, Suite 6066, EPN 6130 Executive Blvd, Bethesda, MD 20892-741. E-mail: lclarke@mail.nih.gov
Received 12 August 2009; Revised 12 August 2009; Accepted 12 August 2009

DOI 10.1593/tlo.09217

the image, meta data, clinical outcome data, and measurement results as a public resource using the NCI image archive [14].

4. Network for Translational Research (NTR found in RFA-CA-08-002): The goal of NTR is to develop, validate, and translate the next generation of multimodality and molecular imaging platforms for targeted cancer problems. Although the NTR goals are not specifically directed at the role of imaging as a biomarker, the research goals include the harmonization and standardization of quantitative imaging methods that may serve as a combined platform to detect, diagnose, and treat cancer. The NTR is also tasked to support image data sharing as a public resource using the NCI image archive [15]
5. Academic Industry Partnerships for Translational Research (R01: PAR-07-214). The goal of this program announcement is the support of academic and industry partnerships directed at translational research and to encourage more rigorous methods to translate and validate imaging methods. For example, this initiative encourages the submission of investigator-initiated R01 applications to develop databases and quantitative imaging methods for targeted cancer problems with an option to be linked to any of NCI research networks as previously described above [16].

This issue of *Translational Oncology* includes four reports [17–20] that reflect the progress made during the last 18 months by the RIDER resource. NCI supported a total of six contracts during 2008 for data collection (five academic sites) that have included phantom measurements and/or patient data collections from clinical trials or other investigations supported by the academic institutions. Each of these academic investigators agreed to share their data as a public resource including their results of analysis. This resource will thus serve the imaging research community by permitting comparative analysis of data sets by other investigators. This public resource and all related documentation on the data collections are Web-accessible and can be found at the National Cancer Imaging Archive. The National Cancer Imaging Archive is now referred to as the National Biomedical Imaging Archive (NBIA) because it is archiving data from several National Institutes of Health (NIH) institutes and other agencies of the federal government [12]. It includes repeat and longitudinal data using either quality assurance phantoms or patient studies. Imaging modalities include x-ray computed tomography (CT), positron emission tomography (PET)–CT, dynamic contrast-enhanced (DCE)–magnetic resonance imaging (MRI), and diffusion-weighted (DW) MRI, and different organ systems (lung, breast, neuro) are represented in the database. In addition several scientific societies such as the Society of Nuclear Medicine, American Association of Physicists in Medicine, International Society for Magnetic Resonance in Medicine (ISMRM), and Radiological Society of North America (RSNA) have agreed to participate in these efforts. The FDA Center for Devices and Radiological Health with partial contract support from NCI and the National Institute of Biomedical Imaging and Bioengineering, has generated a wide range of image data collections using an x-ray CT lung phantom with repeat and longitudinal studies using a array of simulated lung nodules and should serve as a resource to evaluate software tools and Response Criteria for Solid Tumors measurements, where spatial ground truth is known. In addition, National Institute of Standards and Technology (NIST) has initiated strategic plans to address standards for biomedical research where imaging is one of the areas that are being addressed. Two key initiatives have been initiated: a Ge-68 reference standard for longitudinal PET quantitative imaging studies [21] and a Biochange Project that provides a means to benchmark the performance of software tools for measurement of change over time during the course of therapy [22].

Each of the four publications in this issue of *Translational Oncology* was submitted specifically as a consensus publication, with scientific input from all investigators involved in the NCI RIDER initiative. The goal of this work was to explore the different sources of uncertainty, identify those can be minimized, and establish a very practical means to characterize measurement uncertainty. The first report by Meyer et al. [17] provides a solid scientific rationale and research strategy for the design of the data collections that address a means to estimate measurement uncertainty for both idealized phantom measurements and patient data. This publication should serve as an excellent reference resource for investigators developing reference databases resources or clinical trial protocols, where physical measurement uncertainty is important to characterize. The next three publications by McNitt-Gray et al. [18], Kinahan et al. [19], and Jackson et al. [20] address the specific measurements and methods for analysis for three modalities that are emerging for imaging as a biomarker, namely x-ray CT, PET-CT, and DCE and DW MRI. They clearly pose different challenges that are modality and imaging protocol specific and likely require differing complexities of associated modeling and image analysis tools. The intent of these publications was to provide an initial level of consensus for how to characterize the physical measurement uncertainty for these imaging methods in terms of quantity assurance and data analysis. This preliminary work by all should serve to help develop a consensus for how to physically characterize the quality assurance and quantitative imaging methods that will be explored in the NCI-planned QIN [14] and potentially other clinical trial networks. As a related but parallel effort, Barnhart and Barboriak [23] have reviewed how statistical analysis of repeatability derived from repeat patient data sets can be used to interpret parameter changes in individual patients, to evaluate interchangeability of devices, and to power studies to more precisely determine repeatability.

NCI is currently supporting two new contracts for 2009. For example, in addition to harmonization of methods for data analysis, there is a critical need to explore how to improve harmonization of data collection across different commercial imaging platforms. An excellent example is DW MRI, where there is a diverse array of different vendor-supported methods for data acquisition and analysis. This emerging method is proving to be very valuable for providing early indication of response to therapy. NCI organized a workshop on this topic at the ISMRM meeting in 2008 that resulted in a comprehensive consensus report [24]. A second workshop was held at the ISMRM in 2009, where the imaging industry agreed to explore strategies to optimize and harmonize DW MRI data collection, with a goal of reducing platform dependence on DW measurements. This is the first reported effort for harmonizing data collection across commercial platforms for imaging as applied to oncology. NCI expects that under the QIN initiative that additional efforts will be explored to harmonize data collection for other modalities and imaging protocols such as PET-CT, DCE MRI, and MR spectroscopy. The second contract addresses the collection of PET-CT phantom data across multiple sites that are representative of the environment of clinical trials sites. The goals are to optimize data analysis methods that reflect time-related drift in instrument performance for both the dose calibrator and PET systems across all sites, using the NIST reference standard for Ge-68. This work is being performed in consultation with the Society of Nuclear Medicine and American Association of Physicists in Medicine,

and the American College of Radiology Imaging Network (ACRIN) PET Core Laboratory.

In summary, the initial attempts for data and results sharing under the RIDER project, as reflected the publications reported in this issue, has stimulated the research community to volunteer to create similar resources, several of which are archived on NBIA. The scale of this interest was initially evident in 2006 by the attendance of all imaging stake holders at an interagency workshop held at NIST that addressed the physical measurement uncertainty for imaging as a biomarker [6]. The leadership of all the major imaging society attended this workshop and lead to a decision by the RSNA to take the initiative to organize all stakeholders in this area including industry to consider the development of physical imaging standards. As a result of this and several other NIH workshops and workshops organized by several imaging societies on the topic of imaging as a biomarker, the RSNA in 2008 formulated the Quantitative Imaging Biomarker Alliance (QIBA) [25]. QIBA has already organized a series of meetings to engage academia and the imaging and pharmaceutical industries to consider quantitative imaging standards for oncology and other disease entities. One example of the outcome of QIBA efforts has resulted in a recent publication on the potential role of Volumetric CT as a biomarker for therapy response that has included participation of industry stakeholders in the consensus process [25]. Because the principal government agency responsible for standard development and establishment is NIST, it would be appropriate for NIST to play a key role in leading a consortium of imaging societies that will address standards for biomedical measurements. This activity should be informed by NIH-supported research in the evaluation of biomarkers and validation of other measurements. As a result of the broad engagement of all major stakeholders, imaging is now poised to one of the first biomarker methods that may be (physically) standardized within a reasonable time line, both nationally and ideally internationally.

References

[1] Kelloff GJ, Hoffman JM, Johnson B, Scher HI, Siegel BA, Cheng E, Cheson BD, O'Shaughnessy J, Guyton KZ, Mankoff DA, et al. (2005). Progress and promise of FDG-PET imaging for cancer patient management and oncologic drug development. *Clin Cancer Res* **11**, 2785–2808.

[2] Kelloff GJ, Sullivan DC, Wilson W, Cheson B, Juweid M, Mills GQ, Zelenetz AD, Horning SJ, Weber W, Sargent DJ, et al. (2007). FDG-PET Lymphoma Demonstration Project Invitational Workshop. *Acad Radiol* **14**, 330–339.

[3] Kelloff GJ, Krohn KA, Larson SM, Weissleder R, Mankoff DA, Hoffman JM, Link JM, Guyton KZ, Eckelman WC, Scher HI, et al. (2005). The progress and promise of molecular imaging probes in oncologic drug development. *Clin Cancer Res* **11**, 7967–7985.

[4] Weber WA (2006). Positron emission tomography as an imaging biomarker. *J Clin Oncol* **24**, 3282–3292.

[5] Wahl RL, Jacene H, Kasamon Y, and Lodge MA (2009). From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J Nucl Med* **50**, 122s–150s.

[6] Clarke LP, Sriram RD, and Schilling LB (2008). Imaging as a biomarker: standards for change measurements in therapy: workshop summary. *Acad Radiol* **15** (4), 501–530.

[7] McLennan G, Clarke LP, and Hohl R (2008). Imaging as a biomarker for therapy response: cancer as a prototype for the creation of research resources. *Clin Pharmacol Ther* **84** (4), 433–436.

[8] Armato S III, Meyer C, McNitt-Gray M, McLennan G, Reeves A, Croft B, and Clarke L (2008). The Reference Image Database to Evaluate Response to Therapy in Lung Cancer (RIDER) Project: a resource for the development of change-analysis software. *Clin Pharmacol Ther* **84** (4), 448–456.

[9] Petrick N, Brown DG, Suleiman O, and Myers KJ (2008). Imaging as a tumor biomarker in oncology drug trials for lung cancer: the FDA perspective. *Clin Pharmacol Ther* **84** (4), 523–525.

[10] RIDER. Available at: http://imaging.cancer.gov/images/Documents/eb7bee83-9e9f-47f8-ab39-2e998bf8a95d/RIDER%20white%20paper_071306.pdf. Accessed October 12, 2009.

[11] RIDER. Available at: https://wiki.nci.nih.gov/display/Imaging/RIDER. Accessed October 12, 2009.

[12] NCBI. Available at: http://ncia.nci.nih.gov/. Accessed October 12, 2009.

[13] Imaging Workspace. Available at: https://cabig.nci.nih.gov/workspaces/Imaging. Accessed October 12, 2009.

[14] QIN. Available at: http://grants.nih.gov/grants/guide/pa-files/PAR-08-225.html. Accessed October 12, 2009.

[15] NTR. Available at: http://grants.nih.gov/grants/guide/rfa-files/RFA-CA-08-002.html. Accessed October 12, 2009.

[16] NCI PAR. Available at: http://grants.nih.gov/grants/guide/pa-files/PAR-07-214.html. Accessed October 12, 2009.

[17] Meyer CR, Armato SG III, Fenimore CP, McLennan G, Bidaut LM, Barboriak DP, Gavrielides MA, Jackson EF, McNitt-Gray MF, Kinahan PE, et al. (2009). Quantitative imaging to assess tumor response to therapy: common themes of measurement, truth data, and error sources. *Transl Oncol* **2** (4), 198–210.

[18] McNitt-Gray MF, Bidaut LM, Armato SG III, Meyer CR, Gavrielides MA, McLennan G, Petrick N, Zhao B, Reeves AP, Beichel R, et al. (2009). Computed tomography assessment of response to therapy: tumor volume change measurement, truth data, and error. *Transl Oncol* **2** (4), 216–222.

[19] Kinahan PE, Doot RK, Wanner-Roybal M, Bidaut LM, Armato SG III, Meyer CR, and McLennan G (2009). PET/CT assessment of response to therapy: tumor change measurement, truth data, and error. *Transl Oncol* **2** (4), 223–230.

[20] Jackson EF, Barboriak DP, Bidaut LM, and Meyer CR (2009). Magnetic resonance assessment of response to therapy: tumor change measurement, truth data and error sources. *Transl Oncol* **2** (4), 211–215.

[21] Zimmerman BE, Cassna JT, and Fitzgerald R (2008). Standardization of Ge-68/Ga-68 using three liquid scintillation counting based methods. *J Res Natl Inst Stand Technol* **113** (5), 265–280.

[22] NIST Biochange 2008 Pilot. Available at: http://www.itl.nist.gov/iad/894.05/biochange2008/Biochange2008—webpage.htm. Accessed October 12, 2009.

[23] Barnhart HX and Barboriak DP (2009). Applications of the repeatability of quantitative imaging biomarkers: a review of statistical analysis of repeat data sets. *Transl Oncol* **2** (4), 231–235.

[24] Padhani A, Liu G, Mu-Koh D, Chenevert TL, Thoeny H, Takahara T, Dzik-Jurasz A, Rose B, Van Cauteren M, Collins D, et al. (2009). Diffusion-weighted magnetic resonance imaging (DW-MRI) as a cancer biomarker: consensus recommendations. *Neoplasia* **11** (2), 102–125.

[25] RSNA QIBA. Available at: http://www.rsna.org/Research/qiba.cfm. Accessed October 12, 2009.

[26] Buckler AJ, Mozley PD, Schwartz L, Petrick N, McNitt-Gray M, Fenimore C, O'Donnell K, Hayes W, Kim HJ, Clarke L, et al. (in press). Volumetric CT in lung cancer: an example for the qualification of imaging as a biomarker, *Acad Radiol*.