The value of establishing the quantitative accuracy of PET/CT imaging

Paul E Kinahan¹, David A Mankoff², Hannah M Linden¹
¹ University of Washington
² University of Pennsylvania

A total of 1,665,540 new cancer cases and 585,720 deaths from cancer are projected to have occurred for the United States in 2014, and one in four deaths in the United States will have been from cancer (1). The reasons for the increasing burden of cancer to society and the meager progress are complex, but one reason is that many cancer standard therapies are ineffective. This is compounded by the increasing difficulty in conducting clinical trials to evaluate new therapies (2).

At the same time, there has been an explosive growth of knowledge of the interactions between many of the genetic drivers of cancer, including oncogenes and tumor suppressors and insights into the genetic heterogeneity of human tumors that are enabled by exponential improvements in DNA sequencing methods (3). The use of targeted small molecule anticancer treatments and immunotherapies will likely require diagnostic tests to optimize efficacy (4). For these reasons molecular imaging characterization of tumors and their habitats are likely to become standard clinical practice in targeted therapies, but for this we need improved tools to help guide drug development, identify critical targets, and determine whether the target or pathway is 'hit'. Such tools will catalyze efforts in precision medicine, and help clinicians evaluate choices sooner for individual patients.

Quantitative imaging with positron emission tomography (PET) combined with x-ray computed tomography (CT) is a valuable tool for assessment of a tumor's response to therapy and for clinical trials of novel cancer therapies because it can measure functional and molecular changes at multiple tumor sites, with faster and more specific indicators of response than anatomical size changes (5). Success with this approach has been demonstrated using 18F-fluorodeoxyglucose (FDG) for evaluation of therapy-induced changes in glucose metabolism in lung cancer (6) and other tumors. Results from the NeoALTTO (Neoadjuvant Lapatinib and/or Trastuzumab Treatment Optimization) trial in breast cancer showed an early predictive value for HER2-directed targeted therapy (7). Similarly, an early FDG PET scan predicts response to tyrosine kinase inhibitors in GIST tumors treated with imatinib (8), and lung cancers treated with gefitinib (9). While these results and others support the use of FDG PET as both an integrated and an integral marker in trials of targeted therapy, well-designed prospective studies are needed to firmly establish this role for FDG PET (10). A key step in the validation of any biomarker is the evaluation of analytic validity - in this case, the quantitative accuracy of the FDG uptake measures. Specifically, what are the error bars in measurements from the PET images?

The study Weber et al. (11) presents needed data for PET to be usable in a clinical trial of new lung cancer studies. By pooling together prospective multi-center test-
retest data from the ACRIN 6678 and Merck MK-0646-008 trials (74 patients total) in a rigorous analysis, the authors were able to demonstrate the limits of normal variability of tracer update in multi-center imaging of advanced NSCLC (non-small cell lung cancer stages III-IV). These limits provide the framework for determining if there has been a response to therapy, or if a measured change is due to variations to be expected in practice. There have been previous studies measuring the variability from test-retest studies (as noted by the authors) and while the variability observed in this study is slightly higher than in previous single-center studies, it is similar to the results of a previous carefully controlled multi-center test-retest study in 62 patients with gastrointestinal malignancies (12).

The test-retest results such studies are needed to determine parameters for the use of quantitative PET/CT imaging for clinical trials and clinical care. Two other important components of such a framework are the definition of response criteria and comprehensive standards for equipment, protocols, and QA/QC procedures. The former is provided by the continued development of the PERCIST criteria (13), which are supported by the results of this and previous studies. The standards are being provided by several international efforts, including the Quantitative Imaging Biomarker Alliance (QIBA) (14), which in cooperation with other groups (SNMMI, EANM, RSNA, FDA, NIST, NIH, manufacturers, pharma, and others) has developed the Uniform Protocol for Imaging in Clinical Trials (UPICT) (15) and the FDG-PET/CT Profile. The 'Profile' is a new type of document that affirms the measurement bias and/or precision under specific conditions, and then specifies what is required to meet the level of measurement accuracy. This is includes a subset of the UPICT protocol, as well as the entire instrumentation chain. Careful evaluation has revealed that the entire imaging chain needs QA/QC, and that even the display stations used for analysis are subject to errors if not checked, and tools for quantitative imaging 'challenges' developed through the NCI Quantitative Imaging Network (QIN) (16) and QIBA are now being used for this.

Given the challenges with reproducibility and the pooling of data for meta-analysis in biomedical research (17), the sharing of data by Merck and ACRIN is a welcome model of progress in medical imaging research, and a foundation for further progress in the development of the precision medicine paradigm. The study Weber et al. (11) is both an exemplary illustration and a contribution to the necessary data needed to take full advantage of what quantitative molecular imaging can offer for cancer clinical trials, oncologists and precision medicine for cancer patients (18).

Supported by NIH grant U01CA148131.