Molecular Imaging Research in the Outcomes Era: Measuring Outcomes for Individualized Cancer Therapy

David A. Mankoff, MD, PhD, Finbarr O’Sullivan, PhD, William E. Barlow, PhD, Kenneth A. Krohn, PhD

INDIVIDUALIZED CANCER THERAPY

A recent trend in cancer treatment has been termed individualized or targeted therapy and has the goal of matching cancer treatment to the characteristics of the patient and the biologic features of his or her tumor (1, 2). As cancer treatment moves toward more targeted and individualized therapy, there is an increasing need for tools to guide therapy selection and to evaluate response. In particular, to make appropriate choices for therapy, the cancer physician needs to know the following:

1. How aggressive is the cancer? How likely is it that the cancer will spread and/or cause symptoms or death?
2. What are appropriate targets for cancer therapy? Are the targets expressed by the tumor? Are there resistance factors that may mitigate the success of the treatment? Has the target been suppressed?
3. Is the tumor responding? Can a lack of response for ineffective therapy be identified quickly so that alternative treatment can be tried?

The current approach to patient management relies upon a combination of imaging and biopsy to detect, localize, and confirm tumors and upon in vitro assay of biopsy material to determine tumor biologic features. Relying entirely upon tissue sampling to characterize tumors has two important limitations:

1. Tumors are heterogeneous. Increasing evidence shows that cancer cells are genetically unstable, leading to considerable variability in both cellular genotype and phenotype in an individual tumor and across different tumor sites (3). Therefore, in vitro assay is prone to sampling error, especially with the increasing use of minimally invasive tumor sampling by needle biopsy. An important use of advanced imaging methods is to direct biopsy to appropriately characterize heterogeneous regions of the cancer.
2. In vitro assay does not adequately represent the complex interactions between the tumor, the host tissue, and the selected therapy in vivo. While cancer cells may be sensitive to a particular therapy in vitro, in the patient, factors such as the local microenvironment, drug metabolism, and drug delivery will also affect response and are incompletely reflected by ex vivo testing (4).

Emerging approaches to biochemical and molecular imaging offer the ability to make sophisticated and, importantly, quantitative measurements of in vivo tumor biology. Imaging is therefore ideal for guiding targeted cancer treatment in conjunction with tissue sampling and in vitro assay. Radioisotope imaging using PET is particularly well suited for probing a wide variety of molecular pathways and is the focus of this review.

THE APPROACH TO CANCER IMAGING

Most of cancer imaging thus far has been directed toward cancer detection and staging. The principle guiding
most clinical cancer imaging to date has been to find the cancer and determine where it has spread. This approach has made a significant contribution to cancer treatment and remains an important part of cancer care. In addition to the traditional anatomically-based imaging methods such as CT, emerging biochemical imaging modalities such as $^{18}$F-fluorodeoxyglucose (FDG) PET are increasingly used for cancer detection and especially staging (5).

For biochemical or molecular imaging, which generally relies on imaging probes to generate images, localizing tumor sites requires probes that have higher uptake in tumors than in normal background tissues. An illustrative set of targets for tumor detection is depicted in Figure 1A.

As imaging moves beyond cancer detection to address specific characteristics of the cancer as a tool to guide treatment, the paradigm for tumor imaging must expand (6). For guiding therapy, the absence of a particular tumor feature, for example, an oncogene product, may be as important as its presence. In this regard, cancer imaging must expand to include quantitative assays of tumor biology in addition to simply finding cancer sites. This implies a broader set of imaging targets, depicted in Figure 1B. It also implies the need to simultaneously localize tumors (i.e., the existing paradigm) and measure their biology (i.e., the expanded cancer imaging paradigm). This need increases the importance of recent advances in multimodality imaging such as combined PET/CT or SPECT/CT tomographs and the ability to coregister different images taken at different times (e.g., sequential images of two different PET radiopharmaceuticals) (7).

The use of imaging to direct cancer therapy also increases the importance of reliable, reproducible, and quantitative data from the imaging studies. This implies a critical need for tomograph users and especially tomograph manufacturers to use uniform imaging acquisition and analysis protocols and to incorporate standardized quantitative calibration measures into routine scanner quality control. Such measures will ensure consistent quantitative data across different sites and across different tomographs. Recent NCI consensus publications for DCE-MRI (8) and FDG PET (9) are good first steps in this direction, but further progress will require considerable efforts by manufacturers and users to make reliable and reproducible data across multiple sites a clinical reality.

**Figure 1.** Illustration of cancer cell targets for cancer imaging. Targets for tumor detection (A) must be processes present in high levels in the tumor, but absent or present in low levels in normal tissue. There are a greater number of possible targets for cancer imaging to guide tumor therapy (B), where measuring both increased and decreased levels is important in choosing and monitoring treatment.

In the imaging paradigm for cancer detection and localization, the typical study design is as follows:

1. Perform imaging test in patients with suspected (detection) or known (staging) cancer.
2. Compare imaging results to biopsy and/or clinical follow-up to determine the presence or absence of tumor at the time of imaging.

For this study design, appropriate metrics are the sensitivity and specificity of the test, or more generally, ROC analysis to determine the ability of the imaging study to correctly classify patients with or without cancer. Multi-center studies have been successfully designed and performed using this approach; some good recent examples are studies testing breast MRI for cancer detection in high-risk women (10, 11).

As cancer imaging moves beyond the existing paradigm of detection and staging to help guide targeted cancer therapy, different study designs and different metrics are needed. Rather than only a detection tool, imaging
also becomes a cancer biomarker, akin to tissue or blood biomarkers, with a different set of requirements than for cancer detection (12). To match the clinical questions that need to be answered for cancer therapy, imaging should be able to perform three types of tasks: prognosis, prediction, and outcome assessment.

- **Prognosis:** Prognostic assays measure cancer aggressiveness and the likelihood that the given cancer will spread and cause patient death. These assays predict disease-free or overall survival (DFS or OS, respectively) for a group of patients with a particular cancer. Prognostic measures identify those cancers most likely to cause trouble for the patient. They help match the aggressiveness (and thus morbidity) of the treatment to the aggressiveness of the cancer. An example of a tissue prognostic assay is the MIB1 or Ki-67 index of cellular proliferation, which has been prognostic for a number of cancers (13).

- **Prediction:** Predictive assays measure the likelihood of response to a particular therapy for a given cancer. These assays predict whether or not the patients will have an objective response to a particular treatment, response (R) or no response (NR), and also the likelihood that the patient will remain progression free during or after the treatment, progression-free survival (PFS). Predictive assays help in making specific therapeutic choices. An example of a tissue-based predictive assay is estrogen receptor (ER) expression in breast cancer as a predictor of response to endocrine therapy, where the absence of ER expression or the progesterone receptor (PR) virtually eliminates the chance of response to endocrine therapy and directs breast cancer patients to alternative treatments, such as chemotherapy (14).

- **Response:** A response assay measures whether or not a tumor has responded to treatment. The measurement may be on a continuous scale—for example, the diameter of a tumor—but is frequently divided into categories such as complete response (absence of tumor after treatment [CR]), partial response (tumor responding but residual tumor present after therapy [PR]), stable disease (SD), and progressive disease (PD). Anatomic imaging has traditionally played an important role in response assessment using size change measures such as the RECIST criteria (15); however, molecular assays are increasingly used to make earlier and more predictive measures of tumor responses. An example is the serum marker PSA, which is frequently used as a measure of prostate cancer disease burden (16).

It is important to note some practical difficulties associated with imaging biomarker studies that are distinct from standard cancer therapy trials. Unlike therapy studies, it can be challenging from both a practical and ethical standpoint to perform large-scale, fully randomized studies of a diagnostic biomarker. Limited availability and considerable cost make large-scale multicenter molecular imaging trials difficult. The consideration of possible patient safety issues—for example, discovering a previously unknown site of disease during an imaging study—makes it difficult to conduct blinded studies. As a result, most clinical molecular imaging studies are based on smaller local studies with nonrandomized designs. While every effort must be taken to appropriately stratify and randomize patients in these smaller studies, the practical difficulties associated with imaging studies imply a need to incorporate relevant covariates into data analysis to enhance the statistical power of designs. Thus, multivariate analysis tools (Cox modeling and logistic regression [discussed later]) are often essential.

In the remaining section, we outline how cancer imaging clinical studies should be designed to address these clinical needs for targeted therapy. We provide selected examples of imaging tests that have been used to assess prognosis, prediction, and early response, relying heavily on experience using PET imaging at our center.
bility of specific markers, cohorts with differing values of the prognostic marker are compared against each other for the outcome to determine if the marker is significantly associated with survival. Because follow-up time can vary when patients are accrued over a number of years, methods employing censoring, such as the Kaplan-Meier method (illustrated in Fig. 2) or Cox proportional hazards method, are frequently used. The Cox method is particularly advantageous for the continuous data resulting from quantitative imaging in that the method can probe for a quantitative relationship between the intensity of the prognostic marker and outcome. The resulting output from this analysis is a hazard ratio, which describes the relative risk of the outcome (for example, death) per unit increase in the magnitude of the prognostic variable. Multivariate Cox analysis can adjust for other known prognostic markers, and test whether a new marker provides independent prognostic capability. It is our method of choice, for example, for evaluating whether new PET probes provide incremental information to FDG PET.

Examples of established tissue-based prognostic markers are indices of proliferation for a number of tumors (13), the estrogen receptor for breast cancer (14), and the expression of oncogenes such as the epidermal growth factor receptor (EGFR) (17). Some molecular imaging prognostic markers are also emerging. Elevated glycolysis, measured by FDG PET, appears to be associated with more aggressive and lethal cancer in a variety of different tumor types (18–23). The association can be quite striking. For example, recent data from MSKCC by Robins et al. (24) suggest that FDG is an exquisitely prognostic marker for iodine-refractory thyroid cancer. While patients with FDG-negative thyroid cancer almost never die of their disease, the mortality rate of FDG-positive disease, especially when present in larger volume, approaches 60%, quite striking for a typically indolent disease such as thyroid cancer. Other prognostic markers measured by molecular imaging are also emerging. Recent studies using PET to measure tumor hypoxia have shown that the presence of hypoxia by PET predicts, as expected, a worse outcome for cervical and head and neck cancers (25, 26).

**Prediction**

Predictive assays indicate the likelihood of response to specific anticancer therapies. In practice, a variety of factors impact response to therapy so that predictive assays are most useful for determining which patients are unlikely to respond to a particular therapy or may benefit from an alternative therapy. Study design is illustrated in Figure 3. Results are often expressed as the fraction of patients responding, or response rate, versus the expression of the putative predictive marker. Other statistical methods, such as logistic regression, test the ability of the marker to predict a particular result, in this case as response to treatment. For survival outcomes as part of therapeutic efficacy evaluation, Kaplan-Meier and Cox regression may be used. In these analyses, it is particularly useful to test for an interaction of the marker with efficacy of each type of treatment tested as a measure of the predictive capability of the marker.

Examples of tissue-based assays include the expression of ER and progesterone receptor (PR) as predictors of breast cancer endocrine therapy response (27), thymidine synthetase (TS) expression as a predictor of response to 5-fluorouracil (5-FU) (28), and HER2 expression as a
predictor of response to trastuzumab (29). Note that markers can be both prognostic and predictive, as is the case, for ER and HER2 expression. Examples of molecular imaging paralleling these tissue-based assays are emerging using PET ER imaging (30, 31), PET or MR imaging of 5-FU (32, 33), and SPECT and PET imaging of HER2 expression (34, 35). Recent data from our center illustrate the use of 18F-fluoroestradiol (FES) PET to image ER expression in metastatic breast cancer as a predictor of response to breast cancer to endocrine therapy. (FES PET) can yield insights beyond those provided by the individual procedures.

The model of using molecular imaging as a predictive assay becomes complex as molecular pharmacology identifies an increasing array of therapeutic targets. In particular, the rate-limiting technology for imaging becomes the development and especially validation of a library of imaging probes that can be used to measure target expression. It is unlikely that imaging will be able to follow the example of in vitro analysis using microarrays (36), where thousands of targets can be probed at once. It will therefore be important to recognize common targets or perhaps common downstream pathways that can serve as indicators of pathway activation. This is a practical and realistic approach to probe development that matches well with the emerging trend toward drugs with broader substrate specificity, such as multi-targeted or “dirty” kinase inhibitors (37).

Response

Even with increasing sophistication in the use of prognostic and predictive markers to guide cancer therapy, it is unlikely that we will be able to predict completely how any individual responds to a particular treatment. With an increasing array of alternative treatments, it will be important to make early assessments of response in order to identify, in particular, ineffective therapy and to move on to alternative treatment. Both oncologists and radiologists have relied upon tumor size measurements to assess response (15), but this approach has a number of limitations, especially for measuring early response. Most notably, a reduction in tumor size is a late marker of successful cancer therapy and is preceded by a number of important biologic steps, illustrated in Figure 5. Although
there has been some success using serum biomarkers to measure tumor response (38), the regional in vivo information provided by imaging is key to accurate response assessment and complementary to blood markers. Molecular imaging is ideally suited to measuring early response by quantifying processes such as cellular proliferation and cell death that are altered early in the course of treatment.

The study design to test measures of response is illustrated in Figure 6. Serial measures over the course of therapy are compared to a standard response measure, such as size criteria or preferably histopathology, to estimate the accuracy of response classification. Appropriate metrics for the accuracy of the response measure are the sensitivity and specificity for measuring response versus the gold standard, or the ability to classify patient response assessed using ROC analysis. Equally important, the value of a response measure as a surrogate outcome, namely its ability to predict patient outcomes (DFS, or OS), should also be tested using a design similar to prognostic markers. Response measures that correlate with subsequent patient outcome will be the most valuable in terms of evaluating treatment efficacy. The study design illustrated in Figure 6 also offers the opportunity for multivariate analysis of response patterns to infer factors associated with response and resistance to particular forms of therapy.

An illustrative example from our center is shown in Figure 7, using serial 99mTc-sestamibi (MIBI) uptake as a measure of breast cancer response to neoadjuvant chemotherapy. We and others have shown that this method tracks breast cancer response, presumably as a measure of tumor blood flow and vascularity, akin to contrast MRI (39). Comparison of the change in MIBI uptake to pathologic response showed the ability of MIBI to classify patients with a complete pathologic response (CR) versus a partial response (PR) (ROC A_z = 0.96) (40). Perhaps more striking, the residual tumor MIBI uptake post-therapy predicted both DFS and OS, providing prognostic information that was independent of other markers (41).

Molecular imaging approaches offer the potential for very early measures of response. Studies have suggested that certain targeted therapies, for example, imatinib, cause a very early and profound decline in FDG uptake, and FDG PET depicts response within 48 hours (42). Imaging other processes such as cellular proliferation and apoptosis, where changes occur early in the course of cancer response, offer other novel approaches to measuring early cancer response (43, 44).

**SUMMARY**

Advances in molecular imaging, combined with the goal of personalized cancer therapy, call for new approaches to clinical study design for trials testing imaging to guide therapy. The role of cancer imaging must expand and move beyond tumor detection and localization to incorporate quantitative evaluation of regional tumor phenotype. Imaging study design and outcome analysis must move beyond metrics designed to measure the performance for detection to include measures of prognosis, prediction of therapeutic success, and early therapy re-
This implies a need for different approaches to cancer imaging clinical trials and changes in their regulatory oversight. Demonstration that a biochemical or molecular imaging method correctly and accurately measures a specific biologic feature should be sufficient for approval for clinical trials. It may be possible that a combination of imaging procedures known to accurately depict tumor phenotype may be prognostic, even if the individual study cannot be directly validated against patient outcomes. Therefore, it will be important to be able to apply a range of possible imaging studies to different targeted cancer therapy trials. Academia and industry must work together with regulatory agencies and payers to facilitate well designed clinical studies, with appropriate outcome measures, to test the effectiveness of imaging in helping to direct cancer therapy. These will ensure the appropriate use of imaging to direct treatment and make an important step towards individualized cancer therapy.

REFERENCES


