The pharmaceutical industry’s R&D productivity is hampered by the reality behind one disheartening statistic: over 40 percent of all therapeutic agents deemed successful in Phase II programs subsequently fail in Phase III studies. Given the cost of Phase III trials, companies pay a staggering price for their inability to identify unsuccessful candidates prior to late-stage development.

The situation is even more devastating in oncology. Not only is the attrition rate higher (nearly 60 percent of oncology drugs fail in Phase III), but the stakes are also higher. Lives depend upon the success of anti-cancer therapies, and the industry is particularly heavily invested in this therapeutic area. Currently, nearly 1,000 new cancer drugs are in development, and R&D companies are facing enormous levels of risk in attempting to bring them to market.

One major challenge in the evaluation of novel cancer therapies is that the measures traditionally used to evaluate anti-tumor response during Phase II trials are frequently not accurate enough. The “noise” inherent in measuring disease progression and response using these methods can obscure the true effectiveness of anti-tumor agents, and companies must make their go/no-go decisions based on inconclusive (or in some cases, misleading) information.

Interestingly, it is not the technology used to scan tumors and create images of their macro and micro structures that is insufficient. Computed tomography (CT) and magnetic resonance imaging (MRI) both produce quite elegant views of tumor anatomy. Rather, the difficulty lies in our ability to quantify and interpret the changes captured in those images over the course of therapy.

Fortunately, more sensitive means to measure the anti-tumor response recorded in images are not only available, but are gaining mainstream acceptance for use in Phase II studies. This paper explores the value of using several of these emerging biomarkers—those used to assess a tumor’s macro and micro structure—to predict which compounds will succeed in Phase III trials.

**Oncologics Development By the Numbers**

- Nearly 1000 agents in development
- 60% of compounds that reach Phase III fail (compared to just over 40% for all therapeutic categories)
- About 5% of compounds in development reach approval
- Average cost to develop a drug (including the cost of failures) is $1.2 billion

continued on page 2
The Methodology for Measuring Tumor Burden

After establishing safety and finding the optimal dose, the goal of early-phase clinical trials for oncology therapies is to assess whether the drug under study reached and affected the target. For decades, this assessment has commonly been performed by imaging solid tumors via CT scans and MRIs to determine their size. And companies base their go/no-go decisions at the end of Phase II studies for solid tumor therapies in large part on endpoints around tumor size, as determined through imaging.

The challenge is that even with sophisticated images taken over time, assessing a patient’s tumor burden is a complex judgment call. Within the trial setting, the process needs to be as standardized as possible so that it can be replicated from patient to patient, minimizing bias and variability.

The general methodology for determining tumor response to treatment over the course of a study is well-established, and is embodied in various formal response criteria. The patient is scanned before treatment, and the tumor burden is defined by selecting a subset of lesions to be followed quantitatively, while others are assessed qualitatively. The patient is then scanned at each later visit, until some condition is met (usually tumor progression or some other reason for ending treatment). At each visit, readers assess existing lesions, either quantitatively or qualitatively, and search for new lesions. The responses for individual lesions are combined into an overall response for each visit. At the end of the trial, the series of visit responses is evaluated to yield information that informs an endpoint, such as the date of progression or the best overall response. See Figure 1: Methodology for Measuring Response.

Figure 1. Methodology for Measuring Response

Thus, over the past four decades (See “The Development of Response Criteria”) the industry has adopted formal response criteria, or sets of rules to guide assessments of each tumor type, specifying:

- What is to be measured, e.g., five target lesions
- How it is to be measured—for instance along the single longest diameter using digital calipers
- How responses will be categorized, in other words, what level of response will constitute complete response, partial response, progression and stable disease.

In certain criteria, for example, a 30 percent decrease from baseline in sum of tumor diameters is considered “partial response,” and “progression” is defined as a 20 percent increase from the nadir of the sum of diameters.
Anatomical Measures

Linear Measures of Tumor Size

While different response criteria exist for different tumor types, those used most commonly today for solid tumors involve linear measures—either a single diameter of the tumor or the product of its orthogonal diameters. These criteria were first developed in the late 1970s, when imaging technology was quite primitive by today’s standards, and there was no reliable method for imaging complex three-dimensional structures.

Even though this is no longer the case, linear measures of tumor size—such as those used in the Response Evaluation Criteria in Solid Tumors (RECIST), the Response Assessment in Neuro-Oncology (RANO) criteria, the Response Criteria for Malignant Lymphoma (RCML), and others—still have their place and offer several benefits.

First, linear measures are relatively simple to apply, and they work well as an approximation of tumor size when tumors have simple shapes that grow uniformly in three dimensions. Clinical radiologists, however, still need to be trained on how to take these measures from scans; doing so in a formalized way as required by the defined criteria is not normally part of daily clinical practice.

Second, the industry has gained ample experience with applying linear measures over the years, and regulators have accepted them as valid assessments of anti-tumor activity in a variety of settings. Indeed, a reduction in a tumor’s longest diameter does correlate with survival in early-phase studies and RECIST-determined progression-free survival correlates with overall survival.

Limitations of Linear Measures

Yet, linear measures of tumor size do have some significant limitations as biomarkers of treatment response. And, given how imaging technology has changed in the past 40 years, it is time to ask if we can do better.

Unfortunately, identifying and measuring the single longest diameter is not always as clear cut as it sounds and requires readers to exercise judgment. As a result, studies have shown that there is considerable variability from reader to reader in using linear measures of tumor size. This leads to a high rate of mis-classification of tumor response in studies involving multiple sites and numerous local readers—a problem that can be mitigated to some extent by using a small pool of centralized readers, and an adjudication process to resolve disagreements.

However, a more fundamental issue is the fact that most tumors are not perfect spheres; they often have irregular shapes marked by involutions and convolutions, not to mention internal heterogeneity. Plus, tumors may grow or shrink without changing along the single longest dimension, as illustrated in Figure 2.

Figure 2. Tumor Diameters Can Be Misleading

A tumor can shrink or undergo necrosis without changing its longest diameter.
Thus, measuring along one or two dimensions oversimplifies the situation with a one-size-fits-all approach that is not sensitive enough to detect changes in the size of complex shapes.

Volumetric Measures of Tumor Size

Nearly 25 years ago, tumor volume was demonstrated to be a more sensitive measure of treatment response than changes to the tumor’s single longest diameter. However, at the time, CT scans were slow and gave poor spatial resolution, resulting in crude estimates of tumor volumes. Also, defining the complete extent of each tumor (a process called segmentation) had to be done manually, with a radiologist outlining the tumor on every slice on which it appeared. This process was time-consuming and expensive.

In the years since, CT scanners have become more widely available, faster, and able to produce higher resolution images. Vastly improved computing power allows for three-dimensional reconstruction of scanned images. Tumor segmentation can now be done using semi-automated tools that define the boundaries of irregular tumors with minimal reader effort and, in some cases, are able to distinguish viable tumor from areas of necrosis. Thus, our ability to measure tumor volumes quickly and accurately has improved dramatically. Volume measurements are best performed at a central image analysis facility, where every image is analyzed by highly experienced readers using the same software.

Conservative estimates indicate that the variance in measuring tumor volumes this way is such that if the measured volume of a tumor changes by 30 percent, there is a 95 percent probability that this is a true change in tumor biology and not noise in the measurement (if the scanning and measurement is done according to certain guidelines). A 30 percent change in volume corresponds to only a 9 percent change in diameter.

The sensitivity of volumetric assessment means that it can detect responses that are not detectable using linear measures. In one trial, in which 253 patients were treated for non-small cell lung carcinoma (NSCLC), 10 patients with geometrically complex tumors were selected for analysis by both volumetric and RECIST 1.1 criteria. Based on the measure of the single-longest diameter, only three patients met the criteria for partial response, compared to eight using volumetric measures.

Furthermore, tumor progression can be spotted far earlier (months earlier, in ICON’s experience) using volumes than with linear measures. Figure 3 compares sequential measures of a tumor’s longest diameter with measures of volume in one patient. The volume measurements show both a clearer response to the therapy and detection of progression (growth after initial shrinkage) 26 weeks before progression is seen by linear measurement.

Such early detection of progression is, of course, of clinical benefit to the patient, who can stop an ineffective treatment and try another therapy. For the sponsor of an early-phase clinical trial, the sensitivity of volumetric measures to changes in tumor size yields greater statistical power per subject, allowing a given trial to be performed using fewer subjects.

Volumetric measures show great promise for their sensitivity and are being used successfully in Phase II studies. Before they will be accepted by the U.S. Food & Drug Administration for use in Phase III trials, however, they must be validated as being predictive of clinical response. Several academic groups are working on this problem, as well as collaborative bodies such as the Quantitative Imaging Biomarkers Alliance (QIBA), a group sponsored by the Radiological Society of North America which includes radiologists, scientists, and other participants from academia, industry, and regulatory bodies (such as the FDA).

Physiological Measures

Though more accurate than linear measurements, volumetric assessment still only evaluates tumors at the level of gross anatomy. Imaging can also be used to measure changes in a tumor’s microscopic anatomy. Here we’ll cover two emerging biomarkers of anti-tumor response that are available for use in early-phase trials and that can inform go/no-go decisions.

Diffusion-Weighted MRI

Diffusion-weighted MRI (DWI-MRI) measures the random movement of water molecules in the spaces between cells to detect changes within tumors that previously could only be seen histologically. Tightly packed tumor cells restrict the motion of water. As tumor cells die, water molecules can move about more freely, increasing the “apparent diffusion coefficient” (ADC) of water molecules, which is measured by DWI-MRI.
Over the past decade, evidence has been mounting that an increase in ADC following therapy indicates tumor cell death. This has been the case in studies of breast cancer, brain tumors, and selected liver and bone tumors. ADC is also useful in differentiating between true disease progression and pseudo-progression in certain tumor types. For instance, in glioblastoma a tumor may actually increase in size or intensity of contrast enhancement following treatment, even if the treatment is working and tumor cells are dying.

The challenge with DWI-MRI is that the technology is not yet standardized across equipment vendors. This means that measurements taken with one vendor’s system may not be comparable to those taken on another, increasing the variability of measurements performed in a multi-center clinical trial.

Although QIBA is working to standardize how ADC measures are taken and calculated across equipment manufacturers, sponsors need not wait until their work is done to take advantage of DWI-MRI technology. It is possible to use ADC values as a biomarker in early phase trials, provided that 1) sites can be trained on strict protocols designed to minimize variability and 2) procedures are in place to ensure compliance with these protocols. This can be achieved by a facility that conducts rapid quality control and gives sites feedback on their compliance with the trial protocol.

Texture Analysis Using CT

CT scans can be used to measure lesion attenuation by assessing the X-ray absorbance of the tissue, pixel by pixel. These measures are expressed in Hounsfield units and relate to the density of the tumor tissue. Response criteria using such information are well-established for GIST, and the methodology is being explored with certain liver tumors and with soft tissue sarcomas. In addition to density, there is evidence that heterogeneity within a mass on CT provides valuable information on treatment response in some tumor types.

Other Imaging Biomarkers

In addition to volumetrics for imaging macroscopic structure and the modalities described above for microstructural analysis, there are imaging biomarkers that show functional properties of tissue. These include positron emission tomography (PET), using well established tracers such as fluorodeoxyglucose (FDG) and more experimental tracers to show aspects of tissue metabolism. Both CT and MRI also allow dynamic contrast techniques to be used to show changes in tissue vascularity, which can be particularly useful in evaluating the effects of cancer drugs with anti-angiogenic properties. These imaging biomarkers are also the subject of intense investigation, and have been used in early phase trials, but are beyond the scope of this paper.

Ultimately, the optimal approach to imaging tumor in early phase trials is a combination of structural and functional imaging modalities that give a broad range of information about tumor biology and treatment effects. The specific combination of imaging biomarkers that is optimal for any...
Anti-Tumor Response as a Continuous Variable

In all existing formal tumor response criteria, the status of a subject at each trial visit is determined by broad categories such as Complete Response, Partial Response, Stable Disease or Progressive Disease. While the parameters that define each category are adjusted for different types of tumors, they nonetheless may be too broad to use as the basis of Phase II go/no-go decisions.

It may well be that new imaging biomarkers should be used to look at response as a continuous variable—an ability that has significant implications for trial costs as well as outcomes. Measuring continuous variables, such as volumes or ADC distribution metrics, yields greater statistical power per subject, and thus more information that can be extracted from any trial. Proper application of continuous variables to a study would require sophisticated statistical planning and close partnering with service providers who can deliver high-quality measurements of such variables.

study must be determined for each combination of tumor type and treatment and worked out in collaboration between the trial team and experts in radiology and in imaging statistics.

Conclusion

Today’s imaging technology offers a number of ways to assess treatment response. While linear measurements are well-established, and there is a great deal of experience with their use in Phase III studies, they are not always sensitive enough to be the basis of go/no-go decisions in earlier research phases. By using newer, more sensitive measures of tumor volume and density, sponsors can gather more precise information on tumor response and generally observe changes sooner. This is beneficial to patients, and also to trial sponsors, since it allows them to conduct trials more quickly and with fewer subjects.

These measures must be further studied and validated to be acceptable as clinical endpoints in late-stage research. However, they have already proven themselves as informative—and often superior—measures of anti-tumor response in early-phase research. They can hold the key to companies improving their R&D productivity by helping them advance to Phase III with only those agents that have a high chance of success.

References

15. Sinkus, Ralph, et al., "Apparent diffusion coefficient from magnetic resonance imaging as a biomarker in oncology drug development"
## The Development of Response Criteria

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1950s</td>
<td>Decrease in tumor mass detected via physical palpation or observable in X-ray; no new lesions or no lesions increasing in size; OR physician quorum voting in favor of interpreted benefit to the patient.</td>
</tr>
<tr>
<td>1960</td>
<td>C. Gordon Zubrod and the Eastern Cooperative Oncology Group argue for agreed standards.</td>
</tr>
<tr>
<td>1976</td>
<td>Drs. Moertel and Hanley tested 16 oncologists’ ability to detect changes in the size of solid spheres beneath a thin mattress. The bidimensional diameters of the sphere had to be at least 50 percent smaller from one tumor to the next for the oncologist to detect it with an error rate of less than 10 percent. Thus, was born the two-dimensional tumor burden assessment in which a 50 percent reduction in tumor size was necessary to declare tumor response. This was carried forward and arbitrarily translated to new modalities such as cross-sectional imaging.</td>
</tr>
<tr>
<td>1979</td>
<td>World Health Organization criteria for X-ray and early CT = bidimensional measurement.</td>
</tr>
<tr>
<td>2000</td>
<td>Response Evaluation Criteria in Solid Tumors (RECIST) introduced by the US National Cancer Institute and the European Organization of Research and Treatment of Cancer (revised in 2009) replaced bidimensional measures with the single longest lesion diameter and added rules on the minimum size of tumors that can be tracked.</td>
</tr>
</tbody>
</table>

continued on page 8
Jamal Gasmi, MD, PhD, Vice President & Global Leader Oncology & Hematology Drug Development, ICON Clinical Research

Dr. Gasmi is a board certified Medical Oncologist, with more than 20 years experience in clinical research and development of oncology/haematology drugs. This includes more than fifteen years of global senior leadership positions in the pharmaceutical & biotechnology industry.

Dr. Gasmi has a distinguished track record in drug development, and has a wealth of experience of producing and implementing drug development plans. He has directed and overseen oncology drug development programs that resulted in marketing approval. As a medical oncologist, Dr Gasmi held senior academic positions in France and the UK. He is an active member of ASCO, ASH and ESMO for 20 years and author of extensive scientific publications in oncology and Physiology from Bath University, UK.
Email: jamal.gasmi@iconplc.com

James Conklin, MD, MS, Senior Vice President Medical & Scientific Affairs, ICON Clinical Research

James Conklin, MD, MS is the Senior Vice President for Medical & Scientific Affairs at IMI. Dr. Conklin is a serial entrepreneur, who pioneered the first commercial imaging core laboratory twenty years ago. He is a Johns Hopkins trained research physician with undergraduate and graduate training in electrical engineering. Earlier in his career he was a Professor of Radiology at several medical schools and has published extensively on medical imaging technologies editing Imaging Techniques in Biology and Medicine for Academic Press in 1988. He is board certified in internal and nuclear medicine. Dr. Conklin has been involved in the development and approval of numerous new drugs in oncology and been using quantitative imaging techniques for over thirty years.
Email: james.conklin@iconplc.com

Gregory Goldmacher, MD, PhD, Senior Director, Medical & Scientific Affairs, Director Oncology, ICON Clinical Research

Dr. Goldmacher is a radiologist and provides medical, scientific and regulatory leadership for trials in every phase of clinical development, as well as training physicians and study staff worldwide in clinical trial imaging methods.
Email: gregory.goldmacher@iconplc.com