Biomarkers (or, more correctly, biological markers) are believed by many to be an integral component of modern antineoplastic drug development (1). An enhanced molecular understanding of neoplasia, accompanied by more sensitive and higher throughput diagnostic tools, has facilitated an explosion of biopharmaceutical and government-sponsored investigation in this area in recent years, culminated by the launching this past fall of a biomarker consortium encompassing the NIH and the Food and Drug Administration as well as several large drug manufacturers (2). All of the large cancer medical meetings prominently include biomarker discussions as part of the agenda and the key cancer journals have accelerated the publication of biomarker-related articles. The sentiment around this subject among industry and academic experts seems to be quite enthusiastic, despite a dearth of evidence supporting the utility of these markers relative to the capital and resources used to develop them.

As shown by Goulart et al. (3), there has been an increasing investment in these often invasive and costly biomarker studies, specifically in phase I trials, primarily based on the belief that these studies will lead to optimization of dose and schedule or alternatively lead to a decision to curtail development. The study by Goulart and colleagues is of great importance because it now allows us to take a critical look at the scientific value of these biomarker studies to assess whether the putative importance of biomarkers in phase I oncology trials can be supported by scientific results that influence drug development.

However, many investigators do not understand the definition, and, more importantly, the limitations of biomarkers, defined as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention” (4). A subset of biomarkers can potentially substitute for clinical end points, and in this scenario can be denoted as surrogate end points. Examples of surrogate end points used in other therapeutic areas include serum cholesterol (lipid-lowering agents) or viral load (antiretroviral agents). One term that has become widely used in oncology is “surrogate marker,” which has no generally accepted regulatory meaning and should be avoided, with either “biomarker” or “biological marker” generally being the intended meaning.

Goulart and colleagues provide a meta-analysis of the use of biomarkers in phase I oncology trials over a recent 12-year period. They conclude that this is a relatively recent phenomenon that it has been more common for National Cancer Institute–sponsored studies and for noncytotoxic agents, and that it has “made a limited and primarily supportive contribution to dose selection, the primary end point of phase I studies,” although biomarker studies often “provided evidence supporting the proposed mechanism of action.” Notably, Goulart and colleagues used a nonstandard definition of biomarker, requiring that such a marker be measured by molecular, biochemical, or imaging techniques, and explicitly excluded common clinical end points of toxicity (e.g., blood counts). Thus, standard clinical observations, which would support the proposed mechanism of action (e.g., acneiform skin rash, increase in blood pressure, neutropenia), are excluded from their definition and were not considered in this analysis.

Goulart and colleagues focused on the contributions of biomarker studies to three distinct aims of phase I studies: (a) dose selection, (b) schedule selection, and (c) support for the proposed mechanism of action of the drug. The data set used included 87 published single-agent phase I studies, all of which had previously been submitted as an abstract to an annual meeting of the American Society of Clinical Oncology (during the period 1991-2002).

Dose selection has traditionally been the primary objective of phase I studies (5) based on estimation of the maximum tolerated dose, usually determined on the basis of the observed blood count nadir. It should be noted that in this maximum tolerated dose paradigm, the blood count is actually a biomarker for toxicity, as a low blood count without clinical sequelae is inconsequential. Furthermore, myelosuppression after administration of a cytotoxic agent provides evidence for mechanism of action. Thus, the concept of using biomarkers to define dose is not novel. On the other hand, our experience with blood count nadirs should educate us that a biomarker effect is not equivalent to a clinical effect. Although a higher dose of drug (e.g., paclitaxel) will lead to greater toxicity, it does not always lead to greater efficacy (6). The optimal dose can only be defined through conduct of a randomized trial, which is a standard part of phase II development in all therapeutic areas other than oncology. However, biomarkers can be quite useful to define a range of doses to be studied in phase II trials, and, in this context, can supplement standard clinical data about toxicity, efficacy, and practical aspects of dosing such as bioavailability, infusion volume, or tablet number (7).
Not surprisingly, biomarker studies (as defined by Goulart and colleagues) did not have a major effect in selecting a dose. Although such studies were often interpreted as supporting the selection of dose, only one study was identified by Goulart and colleagues in which the biomarker results alone led to dose selection. In taking a close look at this study of Allovectin-7, a lipid-associated gene transfer product under development by Vical (8), as well as the follow-up studies of this agent, it seems that the exception (an apparent relationship between a biomarker result and dose selection) proves the rule: biomarkers do not identify an optimal dose. Rubin and colleagues concluded that the proper dose for phase II testing was "the lowest effective dose," which was 10 μg per injection, with "effective" defined as transfection of HLA-B7 expression. A phase II trial was conducted based on this conclusion (9), which did not provide sufficient evidence to proceed to phase III testing at this dose. Thus, further studies of this novel agent in melanoma used high-dose Allovectin-7 (10) leading to phase III testing using the 2-mg dose, which is 200 times higher than recommended by Rubin and colleagues. Thus, this single article identified by Goulart and colleagues exemplifying the utility of biomarkers to define the optimal dose did not further the development of this agent, and in fact may have impeded its development, as it led to further development at a dose subsequently deemed to be too low by the sponsor.

Schedule selection is rarely a focus of phase I studies because most phase I studies use a single schedule. As with dose, identification of the optimal schedule requires a randomized phase II trial with appropriate end points. Not surprisingly, biomarker studies (as defined by Goulart and colleagues) infrequently contributed to the selection of a schedule for phase II.

An often stated additional goal of biomarker studies is to provide evidence of "hitting the target" (1), dubbed by Workman as the "pharmacologic audit trail" (11). The implication is that if this cannot be shown, then the drug should not proceed into phase II trials. However, Goulart and colleagues found no evidence that biomarkers have ever been used for such purpose, at least in oncology, and thus this goal only has hypothetical value. Furthermore, expected mechanism-related toxicities [e.g., acniform skin rash for an epidermal growth factor receptor (EGFR) inhibitor, hypertension for a vascular endothelial growth factor receptor inhibitor, and neutropenia for a kinin kinese protein inhibitor; refs. 12–14] all provide evidence of target inhibition, albeit not evidence for target inhibition in the tumor. However, it is likely that target inhibition in the tumor, even if reliably measurable, would vary greatly between tumors, and may or may not correlate with efficacy. In addition, our assumptions about the target may be wrong, as illustrated by the development of two marketed agents, estramustine (originally developed as an estrogen receptor-targeted alkylating agent) and sorafenib (originally developed as a Raf kinase inhibitor). Biomarkers, which were not used, would only have obfuscated the development of these agents.

Unproven predictive biomarkers (i.e., those that associate with treatment effect) are similarly problematic. Whereas "wins" have occurred here, notably with the use of Her-2/neu expression to determine trastuzumab response, with bcr-abl detection to predict response to imatinib in chronic myelogenous leukemia and with positron emission tomography scanning to predict drug response in lymphoma, most attempts to identify such biomarkers have been nothing more than expensive fishing expeditions. Drug response is multifactorial; patient populations are heterogeneous; potential markers are innumerable; and scientific underpinnings to marker development are imperfect. Randomly, 5% of markers will associate with drug response and 5% with drug nonresponse at a significance level of 0.05, and thus results of studies evaluating a multitude of markers are hypothesis-generating at best and require prospective confirmatory studies at additional cost of time and money.

For example, despite being seemingly obvious predictors, neither vascular endothelial growth factor receptor nor EGFR expression (15) correlates with response to the relevant antagonists. The fact that cetuximab was developed only in EGFR-positive tumors (16) shows how use of biomarkers can lead to drugs not being tested where they can lead to the broadest utility. Similarly, had investigators tested erlotinib only in patients who had EGFR mutations, the drug would be underused today as survival benefit occurs in nonresponders to the drug (17). Incorrectly applied biomarker studies run risks of both exposing patients to ineffective drugs (in the case of false positives) and, perhaps more importantly, discarding useful drugs (in the case of false negatives). Overall, the evidence suggests that making "go/no-go" decisions on flimsy biomarker data is foolhardy.

An additional problem that permeates all the aforementioned potential uses of biomarkers is that the analytics surrounding them are often imprecise and not reproducible, making data interpretation difficult (18). This is further compounded by the fact that tissue fixation and preservation limit the feasibility of many analyses (19).

The reason biomarkers have received so much attention is that cancer drug development is long, risky, and expensive. Taking into account typical costs of pharmaceutical research and development, along with a 95% attrition rate from investigative new drug to approval (20), it costs over a billion dollars (probability adjusted) to bring a cancer drug from early discovery to approval (21). Thus, the use of molecular and other markers has been proposed to enrich patient populations to make trials more efficient and maximize therapeutic index (22). However, there are several costs to performing biomarker analyses. For one, adding additional secondary end points in a trial adds an average of $6,675 per patient in variable costs (23). Second, a false-positive result will be particularly expensive, given the costs of the requisite confirmatory trial. Finally, if results from such studies are misinterpreted or inaccurately used, as exemplified by the requirement for EGFR testing for cetuximab, they will lead to poor clinical decision making, erroneous reimbursement decisions, and harm to patients. Therefore, weakly founded biomarker studies only make pharmacoeconomic sense if they can reliably lead to less expensive drug development and/or clinical benefit to patients. These outcomes have yet to be achieved, even with molecular profiling of cancer (24).

The enthusiasm for biomarkers as a panacea for go/no-go decision-making as well as optimization of dose and schedule comes from a historical oncology perspective on drug development. Whereas other therapeutic areas emphasize well-controlled dose-ranging phase II trials (including placebos as ethically feasible) as a cornerstone for decision-making about activity and dose (and occasionally schedule), oncology has generally eschewed standard principles of drug development.
in favor of single-arm uncontrolled phase II trials, which have a low positive predictive value for success (25–28). It thus should come as no surprise, as Goulart and colleagues point out, that biomarker usage is rare in early-phase trials in therapeutic areas outside of oncology. We propose that those engaged in antineoplastic drug development adopt these standard principles and propose the general clinical development plan detailed in Table 1, which de-emphasizes the role of tumor biopsies (and other biomarker studies).

Alternatively, we can continue to follow our current path, which would hypothetically lead to the misdevelopment of 5-fluorouracil because it would “obviously” be developed only in patients whose tumors overexpress thymidylate synthase, the target of the drug. We would also attempt serial tumor biopsies to carefully measure target inhibition and might convince ourselves that we had an effective dose, even if we had no evidence of mechanism-related toxicity (i.e., mucositis and myelosuppression). Of course, it is an untestable hypothesis to suggest that we would fail with its development today.

In contrast, sorafenib was successfully developed for metastatic renal cell cancer without incorporation of biomarkers, which could have doomed the drug if the development had focused on its putative target, Raf kinase (29–31). It is a testable hypothesis that the use of biomarkers, particularly tumor biopsies, is inversely correlated with success in drug development.

Despite the limitations noted previously, we believe that the authors do as good a job as is reasonably feasible (given the constraints on data access and interpretation) in cataloging how biomarkers are used, or perhaps more aptly said, misused, in early phase oncology trials today. Their conclusion that biomarkers rarely, if ever, affect dose and schedule selection corroborates work by others (32) who have studied this even in agents that have no maximum tolerated dose. Given that biomarker support of mechanism, or lack thereof, has not contributed to go/no-go decisions in practice, sponsors should reconsider the value of including any biomarker evaluations in phase I oncology studies. Furthermore, those studies that carry more than minimal incremental risk without a strong scientific basis and a testable hypothesis could be considered unethical, especially research-specific biopsies of solid tumors for exploration of “secondary objectives.”

References
18. Liu ET. Molecular oncodiagnostics: where we are...
Biomarkers in Phase I Oncology Trials: Signal, Noise, or Expensive Distraction?

Mark J. Ratain and Robert H. Glassman


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/13/22/6545

Cited articles
This article cites 29 articles, 18 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/13/22/6545.full.html#ref-list-1

Citing articles
This article has been cited by 12 HighWire-hosted articles. Access the articles at:
/content/13/22/6545.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.