Biomarkers and surrogate endpoints: How and when might they impact drug development?

Chetan D. Lathia
Clinical Pharmacology, Bayer Corporation, West Haven, CT 06516, USA
Tel.: +1 203 812 2768; Fax: +1 203 812 5056; E-mail: chetan.lathia.b@bayer.com

Abstract. As the pharmaceutical industry starts developing novel molecules developed based on molecular biology principles and a better understanding of the human genome, it becomes increasingly important to develop early indicators of activity and/or toxicity. Biomarkers are measurements based on molecular pharmacology and/or pathophysiology of the disease being evaluated that may assist with decision-making in various phases of drug development. The utility of biomarkers in the development of drugs is described in this review. Additionally, the utility of pharmacokinetic data in drug development is described. Development of biomarkers may help reduce the cost of drug development by allowing key decisions earlier in the drug development process. Additionally, biomarkers may be used to select patients who have a high likelihood of benefit or they could be used by clinicians to evaluate the potential for efficacy after start of treatment.

1. Introduction

Discovery and development of new molecular entities (NME) or analogs of existing chemical entities is challenging. For the sake of the patients as well as to keep the cost of this complex process as low as possible, pharmaceutical scientists must use every tool at their disposal to increase the speed and efficiency of drug development. By some accounts, approval of new molecules can take up to 10 years from initial discovery to first clinical approval and cost as much as $750 million dollars per compound.

Perhaps one of the reasons for this complexity and cost is the lack of appropriate tools to help guide key decisions and reduce cost during development. For example, the choice of molecules to move forward from early preclinical or early clinical development into full-scale clinical development is currently made based upon pre-clinical experiments whose relevance to the human condition is imperfect. For this reason many agents go through lengthy clinical evaluations before its development is terminated due to lack of clinical benefit. Early indicators of the activity or toxicity associated with a molecule could be potentially utilized to make decisions about further development or early termination.

Once a molecule is in clinical development, methods of selecting patients who might truly benefit from the experimental treatment or for monitoring clinical benefit during treatment may greatly simplify drug development. “Biomarkers” include laboratory assays that are thought to reflect a given molecule’s mechanism of action and/or its pharmacodynamic activity. The hope is that such biomarkers could be used to direct key development decisions such as selection of agents for further development, selection of dose and/or schedule and selection of patients.

Unfortunately, biomarker discovery and development is not risk-free, as the true correlation between a specific biomarker assay and clinical benefit may not be confirmed until after the well-controlled clinical studies are completed. While decisions to use biomarkers in drug development are not without their risks, if they are made using biomarkers with a physiological, pathological, or pharmacological rationale and with supportive in vivo preclinical evidence, biomarkers may sig-

ISSN 0278-0240/02/$8.00 © 2002 – IOS Press. All rights reserved
nificantly enhance the cost- and time-efficiency in drug development. Thus, there are many potential utilities and risks of using biomarkers during drug development. The strengths and liabilities of biomarkers, as well as the process of biomarker discovery and development will be discussed in this review article.

2. The drug development process

In order to understand the utility of molecular markers of a drug’s activity in drug development, it is worthwhile to briefly review the process of drug discovery and drug development. The modern process of drug discovery entails identifying a biological target such as a receptor/protein to which endogenous molecules or xenobiotics could bind to alter the secretion of a protein and/or alter the function of a cell, tissue or organ, thereby altering the progression of disease. Early discovery screening cascades are designed to screen candidate molecules for preclinical research. Using techniques of combinatorial chemistry and high throughput screening, early candidates of preclinical research are selected.

In vitro screening usually includes some type of binding to a specific receptor or cell surface proteins. Subsequently, these candidates are usually evaluated for their activity in preclinical in vivo models of efficacy and simultaneously evaluated for their pharmacokinetic and formulation properties. The preclinical research team, in conjunction with clinical scientists, then chooses a lead candidate, with back-ups, for exploratory development.

Toxicology, pharmacokinetic and formulation evaluations are conducted in exploratory development to allow the start of Phase I and Phase IIA or proof-of-concept studies in man. If the safety, tolerability and pharmacokinetics of a compound are favorable in early clinical development, proof-of-concept/Phase IIA evaluations are generally started to determine the potential for a full-scale clinical development in expanded Phase IIB and III trials.

Full scale clinical development involves intense efforts from many physicians and scientists. It takes many years and costs many millions of dollars in the conduct of labor-intensive toxicology, clinical pharmacology and clinical trials. A package containing the appropriate clinical and preclinical data as well as the chemistry and manufacturing data is submitted for regulatory approval. Special Phase IV post-marketing studies are occasionally conducted to support additional indications or to provide additional safety or pharmacokinetic data for the product.

Various stages of development, from the pre-clinical discovery to the clinical trials in patients, are typically conducted with only limited attention to the direct biological activity of the compound. Often the choice of dose and schedule for further development is made based upon limited clinical data. As such, it is not uncommon for clinical investigators to restart Phase II and III studies when additional clinical data instigate revision of the dosing strategies. Furthermore, the primary endpoint of the animal and patient trials is to quantify and characterize the clinical benefit (e.g., pain reduction for analgesics, survival prolongation for anti-cancer agents) and not to quantitate the impact of a given agent on it’s direct biological target. As such, clinical studies are large (e.g., up to several thousand patients) and long (e.g., several years). The true clinical benefit of a given experimental agent is fully understood and full regulatory approval is granted after many years.

This review will examine the utility of biomarkers of biological activity as potential tools for streamlining the various phases of drug development. An excellent review of the use of biomarkers and surrogate endpoints in drug development and regulatory decision-making was recently published [11].

3. Definitions

The Biomarkers Definitions Working Group recently published a paper to provide a framework for conducting future research [2]. These definitions are not provided here verbatim. Instead, working definitions are provided and in some instances elaborated with examples.

3.1. Biomarker

Biomarker is a measurable property that reflects the mechanism of action of the molecule based on its pharmacology, pathophysiology of the disease or an interaction between the two (Fig. 1). A biomarker may or may not correlate perfectly with clinical efficacy/toxicity but could be used for internal decision-making within a pharmaceutical company.
3.2. Predictive assays

Predictive assays are a category of biomarkers that help identify potential responders to therapy. These are laboratory-based evaluations that predict the activity of a NME in a given disease for a given set of patient characteristics. For example, breast cancer patients with elevated levels of serum HER-2 concentrations are more likely to derive clinical benefit from treatment with trastuzumab, than are those patients who do not. Trastuzumab is a monoclonal antibody that putatively blocks the HER-2 neu receptor [13]. It is interesting to note that only 20% of patients identified by this biomarker have responded [15].

3.3. Clinical endpoint

A clinical endpoint is the true measure of how the patient benefits. Clinical endpoints are measured in a sufficient number of patients in prospectively designed, controlled clinical trials to reflect the effect of the pharmacological agent or New Molecular Entity (NME) on the disease process. Using clinical efficacy and safety data collected in these clinical trials, and, if appropriate, comparing these to other approved agents to treat the same disease, regulatory agencies determine the ratio of risk to benefit. This along with other regulatory considerations generally serves as the basis of drug approval. Mortality, morbidity, and survival are some of the definitive clinical endpoints used to assess response to therapeutic interventions. For example, survival is considered as the gold-standard clinical endpoint for anticancer agents and cardiovascular morbidity is considered a desirable for several antihypertensive and anti-atherosclerotic agents. Quality of life instruments could evaluate how patients feel and, are gaining utility in late-stage clinical trials. These instruments may also help with procuring a desirable price with a Health Maintenance Organization (HMO) or a regulatory agency.

3.4. Surrogate endpoint

A surrogate endpoint is a biomarker that correlates very well with the activity and/or toxicity of the molecule. It is generally an acceptable endpoint for registration of a molecule with regulatory agencies. A biomarker qualifies as a surrogate endpoint for a clinical endpoint after controlled clinical trials show a statistically and clinically significant correlation between the two. Surrogate endpoints must reliably predict both the safety and efficacy associated with a pharmacological agent. Surrogate endpoints are sometimes obtained from studies evaluating the natural history of the disease or epidemiology. There are only a few examples of a biomarker becoming a qualified surrogate endpoint. Blood pressure is an accepted surrogate endpoint for antihypertensive agents as it predicts cardiovascular disease, heart failure, stroke and kidney failure [4]. On the other hand, insufficient evidence of a strong correlation between a biomarker and an ultimate clinical endpoint, would disqualify its use as a surrogate endpoint. For example, for several classes of agents bone mineral density has shown to have a good correlation with fracture rates. However, this has not turned out to be the case for fluoride [16].

Atkinson [1] has compiled a table of biomarkers which have found varying degrees of clinical utility. Tables 1 and 2 present selected examples of biomarkers and surrogate endpoints.

4. Applications of biomarkers in drug discovery and development

Biomarkers may be utilized in several phases of drug development. They are useful for in vitro evaluations of hundreds of candidates that are typically screened during drug development. Biomarkers can be useful in lead candidate selection during preclinical re-
Table 1

<table>
<thead>
<tr>
<th>Biomarker/surrogate endpoint</th>
<th>Type of drug</th>
<th>Clinical endpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood pressure</td>
<td>Antihypertensives</td>
<td>Stroke, atherosclerosis, heart failure</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>LDL lowering statins</td>
<td>Coronary artery disease, heart attacks</td>
</tr>
<tr>
<td>Viral RNA</td>
<td>Antiretroviral agents</td>
<td>Survival, decrease in infections</td>
</tr>
<tr>
<td>HbA1c, glucose</td>
<td>Antidiabetic agents</td>
<td>Diabetic neuropathy</td>
</tr>
<tr>
<td>CD4+ T cells</td>
<td>Antiretroviral agents, cytokines</td>
<td>Sustained reduction in viral RNA</td>
</tr>
<tr>
<td>Intraocular pressure</td>
<td>Antiglaucoma agents</td>
<td>Preservation of peripheral vision</td>
</tr>
<tr>
<td>Bone mineral density (BMD)</td>
<td>Antosteoporotic agents</td>
<td>Fracture rate</td>
</tr>
<tr>
<td>MRI scans</td>
<td>Agents for treatment of MS</td>
<td>Decrease in rate of progression disease (check PDR)</td>
</tr>
<tr>
<td>CT scans for tumor size</td>
<td>Anticancer agents</td>
<td>Survival</td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>Biomarker/predictive assay</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intraocular pressure</td>
<td>Glaucoma</td>
</tr>
<tr>
<td>Electrocardiogram</td>
<td>Arrhythmic cardiovascular disease</td>
</tr>
<tr>
<td>Prostate specific antigen</td>
<td>Prostate cancer</td>
</tr>
<tr>
<td>CT scans/X-rays</td>
<td>Cancer</td>
</tr>
<tr>
<td>Carcinoembryogenic antigen</td>
<td>Ovarian cancer</td>
</tr>
<tr>
<td>Increased serum alpha fetoprotein</td>
<td>Hepatocellular carcinoma</td>
</tr>
<tr>
<td>Increased viral load</td>
<td>HIV infection</td>
</tr>
<tr>
<td>Increased blood glucose</td>
<td>Diabetes</td>
</tr>
<tr>
<td>Reduced bone mineral density</td>
<td>Osteoporotic disease</td>
</tr>
<tr>
<td>Fever</td>
<td>Infection</td>
</tr>
<tr>
<td>Reduced absolute neutrophil count</td>
<td>Risk for infection</td>
</tr>
<tr>
<td>Reduced platelet count</td>
<td>Risk of bleeding</td>
</tr>
<tr>
<td>Increased PT/PTT</td>
<td>Risk of bleeding</td>
</tr>
<tr>
<td>Increased AST/ALT/Alkaline phosphatase</td>
<td>Damage to the liver</td>
</tr>
<tr>
<td>Increased serum creatinine</td>
<td>Damage to the kidney</td>
</tr>
<tr>
<td>Increased serum creatinine phophokinase</td>
<td>Rhabdomyolysis</td>
</tr>
</tbody>
</table>

There is significant value in measuring appropriately selected biomarkers in early phase development of the NME to ensure that it alters the pharmacology/physiology/pathology as desired. A NME that does not demonstrate a change in the pathway it is supposed to affect, may have a lower likelihood of success. The utility of biomarkers will be described with examples in the following sections.

4.1. Drug discovery

Early in drug discovery, hundreds of molecules are evaluated in a series of experiments to determine if they are good candidates for further research. The rational drug discovery process screens these molecules against appropriate targets/receptors/proteins. For example, LDL lowering statins are screened based on their activity against HMG Co A reductase enzyme. The rationale is based on clinical epidemiologic information showing that the incidence of coronary heart disease is linked with LDL cholesterol and inhibiting the HMG Co A reductase enzyme helps as it is a rate-limiting step in the synthesis of LDL cholesterol [9].

ACE inhibitors are evaluated in vitro by their inhibition of angiotensin converting enzyme (ACE). Generally, compounds with the highest in vitro potency and selectivity for the intended pathway are chosen for further research. Showing that a purported biochemical pathway is modified in selecting a molecule for further research makes intuitive sense as it builds confidence in the mechanism of action of the drug and provides tools for biomarker development in preclinical and clinical in vivo evaluations.

4.2. Lead candidate selection

In vivo animal models are frequently used to further characterize late-stage research compounds before bringing them forward into clinical development. For example, blood pressure is good marker of cardiovascular morbidity and mortality. Compounds which lower blood pressure in appropriate animal models are likely to show similar antihypertensive activity in clinical trials. From a toxicologic standpoint, a molecule which shows prolongation of QT interval in appropriate animal models is also likely to do so in man. ACAT
inhibitors act by inhibiting the ACAT enzyme thereby reducing the formation of atherosclerotic lesions in arteries [17]. Preclinical proof-of-concept data were generated by evaluation of ACAT inhibition in monocytes turned macrophages in arterial lesions in rabbits dosed with an ACAT inhibitor [3]. Similarly, angiotensin II inhibitors have been evaluated in preclinical in vivo models by measurement of angiotensin II and aldosterone in animals in addition to the measurement of blood pressure.

4.3. Early clinical development

To increase the efficiency and quality of clinical development, pharmaceutical companies seek to improve their screening processes for NMEs by making critical go/no go decisions early in drug development. If the NME does not demonstrate the proof-of-concept in an appropriately designed early clinical trial, development should be terminated. This is particularly useful in therapeutic areas where biomarkers may have found some predictive value for efficacy or patient benefit. For example, successful lipid lowering HMG Co-A reductase inhibitors such as statins may show early evidence of cholesterol lowering in otherwise normal, mild to moderate hypercholesterolemic patients in early Phase I or II trials in a 3- to 4-week multiple dose study. If such a study is designed as a dose-ranging, placebo controlled, randomized, double-blind trial it may be able to help in the selection of doses for future Phase II trials. Similarly, a Phase I multiple dose study with an anticancer agent in a focussed patient population which shows antitumor activity may help with dose-selection of an anticancer agent in Phase II trials. For evaluation of an agent to treat asthma, one may measure FEV1 (Forced Ejection Volume in 1 second) and/or FVC (Forced Vital Capacity) in asthmatic patients to determine the appropriate Phase II dose. In the absence of these more predictive markers, pharmaceutical companies may choose to measure other biomarkers, including plasma concentration-time profiles, to assist with go/no go decisions. Small and mid-size biotechnology companies often develop molecules by altering a few amino acids in a peptide/protein molecule to block a receptor/pathway in a known pathophysiological pathway. Measurement of a protein molecule whose concentration in serum/plasma changes downstream from the blocked receptor/pathway may serve as a useful tool in determining if the molecule works through the putative mechanism of action. In their proof-of-concept studies, these companies may choose to look at that downstream protein or its mRNA to make a go/no go decision with regards to its mechanism of action. The extent of ACE inhibition, even though it may not have clear and distinct relationship with clinical outcome, has been used by pharmaceutical companies to evaluate their NME with that of other pharmaceutical companies before moving their compound further in drug development. It is well known, for example, that pharmaceutical companies may terminate the development of a poorly absorbed compound if, with their best dosage form, exposures comparable to those leading to preclinical efficacy are not observed.

A bridging strategy to move a compound from Phase I to limited Phase II development in Japan is often required. For a cytotoxic anticancer agent under development exposure-toxicity relationship can be used. For example, if thrombocytopenia is the dose-limiting toxicity and occurs in a predefined range of exposures in clinical trials in North America or Europe, and a similar exposure-toxicity relationship is observed in Caucasian and Japanese patients, it may provide assurances to the pharmaceutical company and to the regulatory authority about similarity between Caucasian and Japanese patients. This may allow start of Phase II/III clinical trials with limited number of patients in Japan for evaluating similarity in efficacy and toxicity in these patients.

By developing a hypothesis and a clear decision-tree based on the outcome of these early development clinical trials, pharmaceutical companies may be able to make good decisions with regards to taking their best NME candidates forward in development. Early proof-of-concept trials not only help in making go/no go decisions but could also help in selecting an appropriate dose and exposure to move forward in future clinical trials.

4.4. Full-scale clinical development

A large amount of human resources as well as monies are spent on full-scale development clinical trials. Dose-ranging Phase II studies as well as full-fledged Phase III trials could be conducted and potentially bridged with the use of biomarkers. HIV-I RNA concentrations and CD4+ T-lymphocyte counts have been shown to be independent prognostic markers of clinical progression in patients receiving antiretroviral therapy for HIV related disease [12]. A sufficient change in these markers from baseline has led to accelerated approval [6] of a number of antiretroviral agents. However, the FDA and other regulatory agen-
cies would subsequently ask the pharmaceutical company to provide data demonstrating clinical benefit in future clinical trials [6].

Bone mineral density (BMD) as measured by Dual Energy X-ray Absorptiometry (DEXA) may allow bridging Phase III trials from one indication (hip fracture) to another indication (vertebral fracture) for a molecule under development if BMD is correlated with fracture rate for the first indication. Demonstration of a clinically significant increase in BMD in vertebral bones over a period of sufficient duration may be adequate for regulatory approval of the agent for the indication of vertebral fracture. With Fosamax, the FDA allowed approval of a sustained release once-weekly formulation based primarily on BMD data. The therapeutic equivalence of once-weekly Fosamax 70 mg and Fosamax 10 mg daily was demonstrated in a one-year study of postmenopausal women with osteoporosis. The two treatment groups were also similar with regard to BMD increases at other skeletal sites [14]. However, if for a particular molecule, increases in BMD do not correlate with fracture rate, changes in BMD may not serve as a useful measure to evaluate its efficacy. BMD measurements in clinical studies with fluoride did not reflect decreases in patient fracture rates [16].

Magnetic resonance imaging, along with appropriate clinical evaluations (disease progression based on a clinical score or annual exacerbation rate), is a putative secondary endpoint for the evaluation of multiple sclerosis (MS). It is conceivable that, if a good correlation is shown between the clinical endpoint and MRI in patients with secondary progressive MS, MRI may be used to explore the clinical benefit of that NME in a different MS population.

Investigators have evaluated compounds which putatively change the morphology of an atherosclerotic plaque using MRI in rabbit models of atherosclerosis [10]. MRI could be a sensitive tool to evaluate the change in plaque morphology from baseline in patients with atherosclerotic plaque [5]. If preclinical evaluations are conducted to show that change in plaque morphology from baseline is measurable using MRI, a pharmaceutical company may choose to evaluate MRI as a biomarker for atherosclerotic disease progression to evaluate further development.

5. Use of plasma concentration-time data as a biomarker for efficacy [8]

If the exposure-response relationship is adequately evaluated in placebo-controlled dose-ranging clinical trials and the relationship is characterized well, this relationship may be used to support new dosing regimens, new or modified dosage forms or formulations, or a different route of administration. If a dosing regimen that has not been evaluated in clinical trials is evaluated in a pharmacokinetic (PK) trial and found to be advantageous, and if the relationship between exposure and response is well-characterized, this information could be used to extrapolate from previous clinical results. Similarly, if the pharmaceutical company chooses to develop a modified release dosage form, a known exposure-response relationship could be used to seek approval. An example of approval of Fosamax 70 mg was provided earlier. Procardia XL tablets appear to have received regulatory approval based on a known exposure-response relationship between nifedipine concentrations and blood pressure/angina and additional supportive clinical pharmacology data for these modified-release tablets. In limited circumstances, if it can be demonstrated that changing the route of administration does not alter the metabolic profile or the exposure-response relationship of a compound, further development of the approved agent, using a new route of administration may be warranted. Exposure-response relationship data can also be used to support a change in an existing formulation which may in turn change its pharmacokinetics. Bioequivalence testing is usually sufficient to demonstrate that a new formulation is equivalent to that used to generate the primary efficacy and safety data. In such a situation, the 90% confidence interval of ratio of the geometric means of Cmax and AUC values of the new formulation to the reference formulation is required to be within 80–125%. If the ratio of new formulation falls outside this confidence interval and if the pharmaceutical company has data showing that a wider confidence interval is acceptable based on available clinical data, the FDA may consider these two products to be bioequivalent. In other cases where bioavailability of the new formulation is altered, the FDA may allow altering the strength of the product so that the bioavailability of the altered formulation is similar to that of the marketed product.

6. Risks of using biomarkers

As with any new tool, the use of biomarkers may be associated with risks. These risks have been described by Lesko and Atkinson [11] and could possibly arise because of the following reasons:
1. The NME affects the biomarker but does not affect the clinical outcome. In this case, the biomarker is non-specific. If such a biomarker is chosen in early phase clinical development, the pharmaceutical company could end up wasting a lot of money on clinical development relying on an inappropriate biomarker.

2. The NME affects the biomarker and clinical outcome to a different extent. In this case, there will be some correlation between biomarker and clinical outcome but the biomarker will not be able to fully account for the effect on clinical outcome. If the chosen biomarker accounts for a small portion of the clinical benefit, the pharmaceutical company could make a wrong decision to discontinue the development of a good drug.

3. The biomarker may be associated with only a portion of the effects on clinical outcome. For example, quinidine was found to suppress cardiac arrhythmias leading to normalization of sinus rhythm. However, it also caused premature deaths [7].

7. Conclusions

It is expensive and time consuming to tackle the challenges of discovering and developing new molecular entities without the use of well-defined biomarkers. The use of biomarkers could not only help with reducing the cost of drug development but also develop a better understanding of the disease process and possibly improve patient care. Many validated and unvalidated biomarkers are currently in clinical use. However, there are many, many untapped opportunities in enhancing drug development. Drug development could substantially benefit by making judicious decisions using biomarkers. Biomarkers have their utility in all phases of a drug’s life-cycle. Perhaps, they are most valuable in terminating development of molecules, before a vast amount of resources are expended in expensive Phase 3 clinical trials, which don’t exhibit their proof-of-principle based on the putative mechanism of action. From a clinician’s standpoint, validated biomarkers which preselect a patient population or track the course of a disease (or lack thereof) may add significant value. Clearly, the use of biomarkers is not without risks. If these risks are adequately understood and quantified, decision-making in modern drug development may be significantly accelerated and enhanced by the use of carefully selected biomarkers.

Acknowledgements

The author thanks Rachel Humphrey, MD for a thorough review of this document and Pavur Sundaresan, MD, PhD, for his encouragement and support and to Geraldine Tomassi for helping format this document.

References
