Overview of biomarkers and surrogate endpoints in drug development

John A. Wagner*
Department of Clinical Pharmacology, Merck Research Laboratories, Rahway, NJ, USA

Abstract. There are numerous factors that recommend the use of biomarkers in drug development including the ability to provide a rational basis for selection of lead compounds, as an aid in determining or refining mechanism of action or pathophysiology, and the ability to work towards qualification and use of a biomarker as a surrogate endpoint. Examples of biomarkers come from many different means of clinical and laboratory measurement. Total cholesterol is an example of a clinically useful biomarker that was successfully qualified for use as a surrogate endpoint. Biomarkers require validation in most circumstances. Validation of biomarker assays is a necessary component to delivery of high-quality research data necessary for effective use of biomarkers. Qualification is necessary for use of a biomarker as a surrogate endpoint. Putative biomarkers are typically identified because of a relationship to known or hypothetical steps in a pathophysiologic cascade. Biomarker discovery can also be effected by expression profiling experiment using a variety of array technologies and related methods. For example, expression profiling experiments enabled the discovery of adipocyte related complement protein of 30 kD (Acrp30 or adiponectin) as a biomarker for in vivo activation of peroxisome proliferator-activated receptors (PPAR) \( \gamma \) activity.

Keywords: Acrp30, adiponectin, biomarkers, clinical endpoint, pharmacodynamic marker, PPAR\( \gamma \), surrogate endpoint, type 2 diabetes

1. Introduction

Many current therapeutics were developed without well-defined molecular targets. Completion of the human genome project promises not only a more advanced understanding of the pathogenesis of human disease, but also better definition of molecular targets for potential new treatments. In fact, modern genomics has created new challenges for drug development by increasing the number of new targets and molecules that require clinical evaluation. In turn, these new challenges can be met, at least in part, by use of new and existing biomarker technologies.

There are numerous ways that investigation and use of biomarkers can aid drug development and the practice of medicine. In the setting of drug development, one critical role played by the use of biomarkers is to provide a rational basis for choosing lead compounds [3,15]. Importantly, use of biomarkers can also aid in dose selection of new therapies [2,10]. Investigation of biomarkers in the setting of drug development may aid in determining or refining mechanism of action of a new or existing therapeutics. Along similar lines, investigation of biomarkers may help determine or refine pathophysiology. Finally, if a particular biomarker qualifies for use as a surrogate endpoint, then such use can aid in interactions with regulatory agencies for review and approval of new therapeutics, and may ultimately benefit medical practice by allowing use of new diagnostic tests.

2. Definitions

Definitions of biomarkers and related terms have been recently refined by a Biomarkers Definitions

* Address for correspondence: John A. Wagner, MD, PhD, Department of Clinical Pharmacology, Merck Research Laboratories, PO Box 2000, RY34-A548, 126 East Lincoln Avenue, Rahway, NJ 07065, USA. Tel.: +1 732 594 0274; Fax: +1 732 594 5405; E-mail: wagner@merck.com.
Working Group with members from FDA, NIH, academicians, and industry [2]. Table 1 reviews the definitions of biomarkers, pharmacodynamic markers, surrogate endpoints, and clinical endpoints. The term biomarker is the most general case; it refers to any useful characteristic that can be measured and used as an indicator of a normal biologic process, a pathogenic process, or a pharmacologic response to a therapeutic agent [2]. A pharmacodynamic (PD) marker specifically refers to a biomarker of pharmacologic response. A clinical endpoint actually quantifies a characteristic related to how a patient feels, functions, or survives, and a surrogate endpoint is a biomarker that is meant to substitute for a clinical endpoint. PD markers and surrogate endpoints are subsets of biomarkers (Fig. 1). There are relatively few biomarkers that qualify for the evidentiary status of surrogate endpoints. The primary examples of surrogate endpoints are also PD markers, but it is important to note this is not necessarily the case. Surrogate endpoints are also referred to as surrogate markers in the biomarker literature. The Biomarkers Definitions Working Group has pointed out that the term surrogate endpoint is preferred because the use of this term properly connotes that the biomarker is being used to substitute for a clinical endpoint [2].

Validation and qualification are other key terms used for discussion of biomarkers. Validation is the assessment of the assay or measurement performance characteristics including sensitivity, specificity, and reproducibility. Qualification, or evaluation, is the evidentiary process of linking a biomarker with a clinical endpoint such that it can be used as a surrogate endpoint. The biomarker literature occasionally uses validation and qualification or evaluation synonymously; however, this should be avoided because the validation and qualification processes must be distinguished and the term validation does not adequately describe the qualification process [2].

3. Principles of biomarker development and use

There are numerous factors that recommend the use of biomarkers in drug development including the ability to provide a rational basis for selection of lead compounds, as an aid in determining or refining mechanism of action or pathophysiology, and the ability to work towards qualification of a biomarker as a surrogate endpoint. Typically, an understanding of pathophysiology governs development of biomarkers. As illustrated in Fig. 2, pathophysiology leading to a specific disease outcome is typically a multistep process. A putative biomarker may be (i) involved in the pathophysiology of a disease outcome, (ii) related, but not directly involved in the pathophysiology of a disease outcome, or (iii) not involved in the pathophysiology of a disease outcome. Thus, putative biomarker A is identical to one of the pathophysiologic steps leading to the disease outcome. Putative biomarker B is not directly involved in the pathophysiology of the disease outcome, but is directly correlated with one of the steps leading to the disease outcome, and is, in turn, correlated with the disease outcome. Putative biomarker C is not involved in the pathophysiology of the disease outcome and is not correlated with the disease outcome. Putative biomarkers A and B rationally fit into the pathophysiologic cascade of a disease outcome; confirmatory studies may demonstrate these measures as appropriate biomarkers and they may, in turn, qualify as surrogate endpoints of the disease outcome. On the other hand, a rational basis for putative biomarker C is relatively lacking. Despite a rational basis for biomarker selection, a biomarker must be qualified as a surrogate endpoint through well-controlled clinical studies.

High serum total cholesterol and LDL-cholesterol are regarded as key pathophysiologic steps leading to coronary atherosclerosis. High serum lipids are thought to lead to an excessive lipid burden in the coronary vessel wall, ultimately resulting in atherosclerosis and associated sequelae. Thus, a rational basis exists for recommending total cholesterol as a biomarker of coronary atherosclerosis; in the scheme depicted in Fig. 2, total cholesterol could be labeled Biomarker A. Furthermore, a rational basis exists for use of total cholesterol measurements as a pharmacodynamic marker for the efficacy of lipid lowering therapeutics such as hydroxy-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors. However, it was critical to acquire experience in well-controlled clinical trials to qualify the use of total cholesterol as a surrogate endpoint for atherosclerosis. The 4444 patient Scandinavian Simvastatin Survival Study (4S) provided the critical evidence to qualify total cholesterol as a surrogate endpoint for atherosclerosis. The 4S established that simvastatin decreased overall mortality in association with decreased serum total cholesterol and LDL-cholesterol. Not only
Table 1 Definitions

Biomarker:
A characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic response(s) to a therapeutic intervention

Pharmacodynamic (PD) marker:
A biomarker of pharmacologic response

Surrogate endpoint:
A biomarker that is intended to substitute for a clinical endpoint and is expected to predict clinical benefit (or harm or lack of benefit or harm) based on epidemiologic, therapeutic, pathophysiologic, or other scientific evidence

Clinical endpoint:
A characteristic or variable that reflects how patient feels, functions, or survives

Validation:
Assessing the assay or measurement performance characteristics including sensitivity, specificity, and reproducibility

Qualification or evaluation:
The evidentiary process of linking a putative surrogate endpoint with a clinical endpoint

Fig. 1. Pharmacodynamic markers and surrogate endpoints are a subset of biomarkers.

did this study qualify total cholesterol as a surrogate endpoint for atherosclerosis events and overall mortality, but also laid the groundwork for use of total cholesterol in general medical practice. Biomarkers are also clearly useful prior to qualification as surrogate endpoints. Thus, total cholesterol was useful in the development of HMG-CoA reductase inhibitors prior to its qualification as a surrogate endpoint. For example, total cholesterol measurements were successfully used as a pharmacodynamic marker in a Phase II dose-ranging study of simvastatin to aid in dose selection [12].

A rational basis for recommending the use of a putative biomarker does not guarantee the utility of the biomarker or its qualification as a surrogate endpoint. The rationale exists for recommending maintenance of a normal sinus rhythm as a surrogate endpoint for reduced mortality in the setting of suppression of ventricular arrhythmia following myocardial infarction; however, experience acquired in a well-controlled clinical trial did not support this rationale. In fact, the results of the Cardiac Arrhythmia Suppression Trial (CAST) revealed that mortality was increased by anti-arrhythmic therapy following myocardial infarction [7]. Although there was a clear rationale for use of maintenance of normal sinus rhythm as pharmacodynamic marker and surrogate endpoint for reduced mortality, it was clear from CAST that maintenance of normal sinus rhythm did not qualify as a surrogate endpoint. In the scheme presented in Fig. 2, maintenance of normal sinus rhythm could be labeled biomarker C – not directly related to disease outcome. This example highlights the crucial role for qualification studies in biomarker development.

Commonly, biomarkers and surrogate endpoints are conceptualized with reference to measurement of efficacy, but biomarkers of safety and tolerability are equally important. For example, prolongation of the electrocardiogram (ECG) QT interval is widely accepted as a biomarker for the potential of a drug to produce torsades de pointes and sudden death. Altered cardiac ventricular repolarization may progress to torsades de pointes and sudden death. Drug-related torsades de pointes and sudden death are unpredictable and rare. Thus, these phenomena are usually only observed after significant post-marketing experience with a new therapeutic is acquired. Altered cardiac ventricular repolarization is associated with prolongation of the QT interval. Thus, prolongation of the QT interval may be considered a biomarker for the potential of a drug to produce torsades de pointes and sudden death (see the review in this issue by Sides, pages 57–62, for further discussion of the QT interval).

Biomarkers require validation in most circumstances. Validation is the assessment of the assay or measurement performance characteristics including sensitivity, specificity, and reproducibility. Validation of biomarker assays is a necessary component to delivery of high-quality research data necessary for effective use of biomarkers (see the review in this issue by Swanson, pages 47–56, for further discussion of biomarker validation). Qualification is necessary for use of a biomarker as a surrogate endpoint.
Fig. 2. Pathophysiology is typically a multistep process. A putative biomarker may be (i) involved in one of the steps of the pathophysiology of a disease outcome (Biomarker A), (ii) related, but not directly involved in the pathophysiology of a disease outcome (Biomarker B), or (iii) not involved in the pathophysiology of a disease outcome (Biomarker C).

Biomarker discovery is a graded process by which evidence is acquired linking a biomarker with a clinical endpoint such that it can be used as a surrogate endpoint; it is not necessarily an all or none characterization.

Examples of biomarkers come from many different means of clinical and laboratory measurement. In the clinical setting, measurements made in blood samples are especially useful because of their convenience. Examples include measurements of routine analytes like total cholesterol as well as more complicated measurements such as plasma HIV-1 RNA as a pharmacodynamic marker of antiretroviral therapy in HIV-related disease [11] or tumor antigens as cancer biomarkers (see the review in this issue by Kiviat and Critchlow, pages 73–81, and by Kirkwood and Hockett, pages 63–71, for further discussion of profiling technologies). Areas of complex biology or where the pathophysiology is incompletely understood may particularly benefit from this approach.

Complex and incompletely understood biology hampers efforts in drug development in the area of type 2 diabetes. Thus, expression profiling is especially relevant for therapeutics such as antidiabetic peroxisome proliferator-activated receptor (PPAR) γ agonists that regulate gene expression. At present, there are no clearly defined human biomarkers that are specific for in vivo activation of PPARγ as well as useful in healthy subjects and patients with type 2 diabetes. In patients with type 2 diabetes, measures of glucose metabolism, including fasting plasma glucose, are useful biomarkers, but these measures are not specific to activation of PPARγ.

A useful example of the biomarker discovery process is a putative specific PPARγ biomarker that was highlighted by results from expression profiling experiments [4]. This example reveals the utility of expression profiling in biomarker discovery and also outlines the transition from preclinical to clinical studies during early biomarker use. The cDNA encoding adipocyte related complement protein of 30 kD (Acrp30 or adiponectin) was originally identified by subtractive cloning or mRNA differential display from cultured adipocytes versus preadipocytes [9,17]. Acrp30 is a protein of 247 amino acids with substantial homology to complement factor C1q. The protein contains both an amino terminal collagenous domain and a C-terminal globular domain, and x-ray crystallography reveals that the globular domain is structurally similar to TNFα [18]. The distribution of Acrp30 mRNA in mouse, rat, and human is confined almost exclusively to adipose tissue, its expression is increased during differentiation of cultured adipocytes, and its expression is decreased in white adipose tissue derived from obese mice (ob/ob, db/db) or humans versus lean controls [4,9,17]. In addition, Acrp30 is a secreted protein that circulates in plasma at high concentrations (5–10 µg/mL) [17]. These observations suggest that
Acrp30 may be a circulating adipose-derived factor that, like leptin or resistin, could influence energy balance and/or metabolism, possibly associated with obesity.

Acrp30 may be functionally related to insulin sensitization [1,4,8]. Acute administration of recombinant Acrp30 lowered postprandial levels of glucose, triglycerides, and free fatty acids in mice fed with high-fat and sucrose [8]. Both the full length protein and a truncated 27 kDa version of the protein containing the globular head domain had efficacy in reducing postprandial glucose levels and the 27 kDa protein was shown to induce fatty acid oxidation in cultured skeletal muscle cells. Acrp30 increased the expression of enzymes associated with fatty acid oxidation and energy dissipation in muscle. Acrp30 in combination with leptin resulted in insulin sensitization in lipopolysaccharide [19]. Single injections of full length recombinant Acrp30 produced lowering of postprandial glucose levels in both normal and ob/ob mice [1]. Recombinant Acrp30 could also enhance the effect of insulin to suppress glucose output from cultured hepatocytes suggesting that a primary effect of this protein may be to augment hepatic insulin action [1,19].

Expression profiling experiments revealed that PPARγ agonists of multiple structural classes could strongly induce Acrp30 mRNA expression in cultured 3T3-L1 preadipocytes [4]. Quantitative PCR demonstrated that in vivo treatment with rosiglitazone increased Acrp30 mRNA levels in white adipose tissue of obese, insulin resistant, db/db mice [4]. Increased Acrp30 was observed at doses of rosiglitazone that were also associated with decreases in glucose and triglycerides. Quantitative Western analysis revealed that circulating Acrp30 levels could also be increased in vivo by PPARγ agonist treatment [4]. In db/db mice, rosiglitazone increased circulating Acrp30 levels, and these changes were more pronounced (4-6-fold) than the changes observed at the RNA level in adipose tissue (2-3-fold). The PPARγ agonist associated increases in Acrp30 are specific to PPARγ because no Acrp30 increases were observed following administration of the PPARα agonist fenofibrate to db/db mice, nor following administration of the non-PPAR associated anti-hyperglycemic agent metformin to HFD/STZ mice, despite appropriate levels of lipid altering and glucose lowering efficacy [4].

Circulating Acrp30 levels were also increased in vivo by PPARγ agonist treatment in humans [4,21]. In healthy subjects, a two week treatment with rosiglitazone (4 mg twice daily) produced a marked increase in Acrp30 levels in a placebo-controlled clinical study [4]. Similar to observations in mice, the PPARγ agonist fenofibrate (200 mg once daily for two weeks) did not alter circulating Acrp30 levels in humans, supporting the specificity of Acrp30 as a biomarker for PPARγ activity. Importantly, increased Acrp30 levels were also observed after a 6 month rosiglitazone treatment in patients with type 2 diabetes in a placebo-controlled clinical study [21].

The mechanism by which PPARγ agonists increase Acrp30 mRNA expression is unclear because there are no consensus PPRE sites in the murine Acrp30 promoter region [21]; however, potential sites for C/EBPβ suggest a potential mechanism for induction of Acrp30 during adipocyte differentiation and/or as a secondary mechanism by which PPARγ might induce Acrp30 gene expression. The magnitude of the induction as well as the observation that elevated levels are sustained for the duration of the treatment suggests that a solely post-translational mechanism is unlikely [4]. The rapid induction of Acrp30 levels after initiation of PPARγ agonist treatment also suggests that an increase in the number of new adipocytes could not fully account for this phenomenon. In addition, increased adiposity generally does not lead to an increase in Acrp30 levels in serum [4,6].

Regulation of Acrp30 mRNA and circulating levels by PPARγ agonists in preclinical animal models, healthy human subjects, and patients with type 2 diabetes suggests a role for this protein as a novel PPARγ-specific biomarker. In addition, a mechanistic role for Acrp30 related to insulin sensitization [1,8,20] strengthens the validity of this approach as a PPARγ biomarker. The data from preclinical models and healthy human subjects suggests that increased Acrp30 is a relatively early, specific response to activation of PPARγ. A detailed time course experiment in healthy subjects and patients with type 2 diabetes will be necessary to demonstrate that Acrp30 increases precede overt changes in glycemic control or other measures of the metabolism. Further studies will also be required to assess whether improvements in insulin sensitivity correlate with Acrp30 induction and to define a PPARγ agonist dose response relationship.

Use of Acrp30 as a putative biomarker highlights many of the potential roles biomakers can play in drug development and in general medical practice. Acrp30 may prove very useful in relatively short term clinical studies in healthy subjects or in patients with type 2 diabetes to assess whether a new potential PPARγ agonist is efficacious in humans, possibly reducing the re-
requirement for longer duration “proof-of-concept” clinical studies in patients with type 2 diabetes. Acrp30 levels may aid in selecting doses of novel PPARγ agonists. In addition, Acrp30, if affected in short durations of treatment in patients with type 2 diabetes, could prove valuable for the determination of whether individual patients were responding to treatment with PPARγ agonists in general medical practice. Given that suppressed Acrp30 levels were observed in patients with dominant-negative PPARgamma mutations [4], it is also tempting to speculate Acrp30 deficiency has an important role in the pathogenesis of severe insulin resistance and type 2 diabetes which is present in these patients.

Expression profiling experiments may be useful in the discovery of other potential biomarkers for type 2 diabetes and a variety of other therapeutic areas. Clearly, the understanding of pathophysiology will continue to govern the development of new biomarkers; however, expression profiling will have an increasing role in biomarker discovery. Biomarkers may prove useful not only in drug development, but possibly in refining the understanding of the pathophysiology and optimizing treatment in medical practice.

References

[1] A.H. Berg, T.C. Combatisiaris, X. Du, M. Brownlee and P.E. Scherer, 2001 The adipocyte-secreted protein Acrp30 enhances hepatic insulin action, Nat Med 7 (2001), 947–953.
[2] Biomarkers Definitions Working Group, Biomarkers and surrogate endpoints: preferred definitions and conceptual framework, Clin Pharmacol Therapeutics 69 (2001), 89–95.
[3] J.W. Blue and W.A. Colburn, Efficacy measures: surrogates or clinical outcomes, J Clin Pharmacol 36 (1996), 767–770.
[4] T.P. Combs, J.A. Wagner, J. Berger, T.W. Doebber, W.-J. Wang, B.B. Zhang, M. Tanen, A.H. Berg, S. O’Rahtilly, D.B. Savage, K. Chatterjee, S. Weiss, P.J. Larson, K.M. Gottsediener, B.J. Gertz, M.J. Charron, P.E. Scherer and D.E. Moller, Induction of Acrp30 levels by PPARgamma agonists: a potential mechanism of insulin sensitization, Endocrinology 143 (2002), 998–1007.
[5] S.G. Darvey and J. Pekkanen, Should there be a moratorium on the use of cholesterol lowering drugs? BMJ 304 (1992), 431–434.
[6] K. Das, Y. Lin, E.Y.Z. Widen and P.E. Scherer, Chromosomal localization, expression pattern and promoter analysis of the mouse gene encoding adipocyte-specific secretory protein Acrp30, Biochem. Biophys. Res. Comm. 280 (2001), 1120–1129.
[7] D.S. Echt, P.R. Liebon, B. Mitchell, R.W. Peters and D. Obias-Manno, Mortality and morbidity of patients receiving encaidine, flecainide or placebo: the cardiac arrhythmia suppression trial, N Engl J Med 324 (1991), 781–788.
[8] J. Fruebis, T.S. Tsao, S. Javorschi, D. Ebbets-Reed, M.R. Erickson, F.T. Yen, B.E. Bihain and H.F. Lodish, Proteolytic cleavage product of 30-kDa adipocyte complement-related protein increases fatty acid oxidation in muscle and causes weight loss in mice, Proc Natl Acad Sci USA 98 (2001), 2005–2010.
[9] E. Hu, P. Liang and B.M. Spiegelman, AdipoQ is a novel adipose-specific gene dysregulated in obesity, J Biol Chem 271 (1996), 10697–10703.
[10] L.J. Lesko and A.J. Atkinson, Use of biomarkers and surrogate endpoints in drug development and regulatory decision making: criteria, validation, strategies, Annu Rev Pharmacol Toxicol 41 (2001), 347–366.
[11] I.C. Marschner, A.C. Collier, R.W. Coombs, R.T. D’Aquila and V. D’Gurrutola, Use of changes in plasma levels of human immunodeficiency virus type 1 RNA to assess the clinical benefit of antiretroviral therapy, J Infect Disease 177 (1998), 40–47.
[12] M.J.T.M. Mol, D.W. Erkelens, J.A. Gevers Leuven, J.A. Schouten and A.F.H. Stalenhoef, Effects of simvastatin (MK-735) on plasma lipids in familial hypercholesterolaemia, Lancet 2 (1986), 936–939.
[13] M.F. Muldoon, S.B. Manuck and K.A. Matthews, Lowering cholesterol concentrations and mortality: a quantitative review of primary prevention trials, BMJ 301 (1990), 309–314.
[14] M.F. Oliver, Doubts about preventing coronary heart disease: multiple interventions in middle aged men may do more harm than good, BMJ 304 (1992), 393–394.
[15] P. Rolan, The contribution of clinical pharmacology surrogates and models to drug development: a critical appraisal, Br J Clin Pharmacol 44 (1997), 219–225.
[16] Scandinavian Simvastatin Survival Study Group, Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S), Lancet 344 (1994), 1383–1389.
[17] P.E. Scherer, S. Williams, M. Fogliano, G. Baldini and H.F. Lodish, A novel serum protein similar to C1q, produced exclusively in adipocytes, J Biol Chem 270 (1995), 26746–26749.
[18] L. Shapiro and P.E. Scherer, The crystal structure of a complement-lq family protein suggests an evolutionary link to tumor necrosis factor, Curr Biol 8 (1998), 335–338.
[19] Y. Wang, A. Xu, C. Knight, L.Y. Xu and G.J.S. Cooper, Hydroxylation and glycosylation of the four conserved lysine residues in the collagenous domain of adiponectin. Potential role in the modulation of its insulin-sensitizing activity, J Biol Chem. (2002), (in press).
[20] T. Yamauchi, J. Kamon, H. Waki, Y. Terauchi, N. Kubota, K. Hara, Y. Mori, T. Ide, K. Murakami and N. Tsuboyama-Kasaoka et al., The fat-derived hormone adiponectin reverses insulin resistance associated with both lipoatrophy and obesity, Nature Med. 7 (2001), 941–946.
[21] W.S. Yang, C.Y. Jeng, T.J. Wu, S. Tanaka, T. Funahashi, Y. Matsuzawa, J.P. Wang, C.L. Chen, T.Y. Tai and L.M. Chuang, Synthetic peroxisome proliferator-activated receptor-gamma agonist, rosiglitazone, increases plasma levels of adiponectin in type 2 diabetic patients, Diabetes Care 25 (2002), 376–380.