

What are biomarkers?

Kyle Strimbu and Jorge A. Tavel

Division of Clinical Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA

Correspondence to Jorge A. Tavel, MD, Division of Clinical Research, National Institute of Allergy and Infectious Diseases, 6700B Rockledge Drive, Room #1127, MSC 7609, Bethesda, MD 20892, USA
Tel: +1 301 402 0564;
e-mail: jtavel@niaid.nih.gov

Current Opinion in HIV and AIDS 2010, 5:463–466

Purpose of review

This article provides working definitions and a conceptual framework to understand the roles of biomarkers in clinical research.

Recent findings

The definitions of the terms discussed in this article – medical signs, symptoms, biomarkers, surrogate endpoints, clinical endpoints, validation – are still under discussion, as are their relationships to each other, but broad consensus has developed in the past decade and a half about the necessity of distinguishing between, in particular, surrogate and clinical endpoints.

Summary

This article outlines the major definitions of the key terms in this field and considers select cases in which misunderstandings about the terms led to flawed research conclusions.

Keywords

biomarkers, clinical endpoints, surrogate endpoints

Curr Opin HIV AIDS 5:463–466
© 2010 Wolters Kluwer Health | Lippincott Williams & Wilkins
1746-630X

Introduction

The use of biomarkers in basic and clinical research as well as in clinical practice has become so commonplace that their presence as primary endpoints in clinical trials is now accepted almost without question. In the case of specific biomarkers that have been well characterized and repeatedly shown to correctly predict relevant clinical outcomes across a variety of treatments and populations, this use is entirely justified and appropriate. In many cases, however, the ‘validity’ of biomarkers is assumed when, in fact, it should continue to be evaluated and reevaluated. This article will consider the current conceptual status of biomarkers as clinical and diagnostic tools and as surrogate endpoints in clinical research with the goal of providing context for interpreting studies that rely heavily on such biological measures.

What is a biomarker?

The term ‘biomarker’, a portmanteau of ‘biological marker’, refers to a broad subcategory of medical signs, that is, objective indications of medical state observed from outside the patient, which can be measured accurately and reproducibly. Medical signs stand in contrast to medical symptoms, which are limited to those indications of health or illness perceived by patients themselves. There are several more precise definitions of biomarkers in the literature, and they fortunately overlap considerably. In 1998, the National Institutes of Health Biomarkers Definitions Working Group defined a biomarker

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’ [1^{••}]. A joint venture on chemical safety, the International Programme on Chemical Safety, led by the World Health Organization (WHO) and in coordination with the United Nations and the International Labour Organization, has defined a biomarker as ‘any substance, structure, or process that can be measured in the body or its products and influence or predict the incidence of outcome or disease’ [2]. An even broader definition takes into account not just incidence and outcome of disease, but also the effects of treatments, interventions, and even unintended environmental exposure, such as to chemicals or nutrients. In their report on the validity of biomarkers in environmental risk assessment, the WHO has stated that a true definition of biomarkers includes ‘almost any measurement reflecting an interaction between a biological system and a potential hazard, which may be chemical, physical, or biological. The measured response may be functional and physiological, biochemical at the cellular level, or a molecular interaction’ [3]. Examples of biomarkers include everything from pulse and blood pressure through basic chemistries to more complex laboratory tests of blood and other tissues. Medical signs have a long history of use in clinical practice – as old as medical practice itself – and biomarkers are merely the most objective, quantifiable medical signs modern laboratory science allows us to measure reproducibly. The use of biomarkers, and in particular laboratory-measured

biomarkers, in clinical research is somewhat newer, and the best approaches to this practice are still being developed and refined. The key issue at hand is determining the relationship between any given measurable biomarker and relevant clinical endpoints.

Biomarkers versus clinical endpoints

Biomarkers are by definition objective, quantifiable characteristics of biological processes. They may, but do not necessarily, correlate with a patient's experience and sense of wellbeing, and it is easy to imagine measurable biological characteristics that do not correspond to patients' clinical state, or whose variations are undetectable and without effect on health. It is also even easier to imagine measurable biological characteristics whose variance among populations is so great as to render them all but useless as reliable predictors of disease or its absence. In contrast, clinical endpoints are variables that reflect or characterize how a patient in a study or clinical trial 'feels, functions, or survives' [1^{••}]. They are, in other words, variables that represent a study patient's health and wellbeing from the patient's perspective.

There has long been broad consensus that clinical endpoints are the primary, and to some the only relevant, endpoints of all clinical research and ultimately of all biomedical research. The goal of clinical practice is to improve morbidity and mortality, not to change quantifiable features of patients' innate biochemistry, for instance, with no outward clinical effect. Similarly, patients seek treatment for their diseases, not for the numerical measures that frequently but not perfectly correlate with their illnesses.

Survival is considered by many the gold-standard clinical endpoint for most HIV trials, but other well defined, unambiguous clinical variables, such as stroke, myocardial infarction, and occurrence of predefined opportunistic infections, have also been used as endpoints in appropriate circumstances; they offer clear, unambiguous data with the potential to show definitively whether interventions are effective or ineffective, as well as safe or unsafe. Not all clinical endpoints are created equal, however. Examples of clinical data elements that provide less reliable, less quantifiable information include breath sounds, pain, and alleviation of 'symptoms' in cases in which symptoms are not defined in advance.

Biomarkers as surrogate endpoints

When used as outcomes in clinical trials, biomarkers are considered to be surrogate endpoints; that is, they act as surrogates or substitutes for clinically meaningful endpoints. But, not all biomarkers are surrogate endpoints, nor are they all intended to be. Surrogate endpoints are a

small subset of well characterized biomarkers with well evaluated clinical relevance [1^{••}]. To be considered a surrogate endpoint, there must be solid scientific evidence (e.g., epidemiological, therapeutic, and/or pathophysiological) that a biomarker consistently and accurately predicts a clinical outcome, either a benefit or harm. In this sense, a surrogate endpoint is a biomarker that can be trusted to serve as a stand-in for, but not as a replacement of, a clinical endpoint.

Even biomarkers that are statistically validated to be surrogates for a given clinical endpoint may not actually be part of the pathophysiological pathway that results in that endpoint. In some cases, there may be evidence that the biomarkers measure a process or product of a key pathway stage, but assuming this relationship in all cases risks mistaking correlation for causation. Many possible explanations exist for biomarkers that correlate with clinical endpoints under only limited circumstances. For example, multiple interrelated disease pathways may be involved, or the biomarkers might be indirect signs of a pathway that are not fundamental to the key disease processes.

There are a number of advantages of using biomarkers as surrogate endpoints in trials. Primary clinical endpoints, such as survival, can occur so infrequently that their use in clinical trials can be highly impractical or even unethical. For many diseases, clear clinical endpoints such as survival or recurrence of, for instance, a cardiovascular event may occur only after many years of treatment. Biomarkers can provide researchers interim evidence about the safety and efficacy of such treatments while more definitive clinical data is collected. In some cases, it may be preferable to use established biomarkers as surrogate endpoints to reduce the risk of harm to patients: the early data provided by biomarkers can allow researchers the opportunity to stop interventions potentially harmful to patients before the associated clinical data would be available. In other cases, biomarkers may simply allow researchers to design smaller, more efficient studies, reducing the number of patients exposed to a given experimental treatment. By shortening the time to approval of new treatments, more efficient trials could speed the overall drug development process, allowing effective treatments to reach their target patient populations sooner, while conserving both material and human resources for other research projects, even other clinical trials.

Of note, the term 'surrogate marker' – in contrast to the term 'surrogate endpoint' – has been criticized for inaccurately suggesting that the surrogate, in this case a biomarker, is a stand-in for another marker [1^{••}]. In fact, biomarkers as surrogate endpoints should be stand-ins only for clear clinical endpoints, and their use as surrogate

endpoints is only appropriate under limited circumstances as discussed below.

Characterization and evaluation of biomarkers

To identify biomarkers as surrogate endpoints requires the determination of relevance and validity. Relevance refers to a biomarker's ability to appropriately provide clinically relevant information on questions of interest to the public, healthcare providers, or health policy officials. Validity refers to the need to characterize a biomarker's effectiveness or utility as a surrogate endpoint. Unfortunately, validity is not typically black or white, but instead a spectrum. Some researchers have, in fact, rejected the term validation as 'unsuitable' to the study of biomarkers because it suggests that there can be a complete biological understanding of the relationship between a given biomarker and a clinical endpoint, an assumption they reject [1^{••}]. Instead, an alternate term that has been offered is 'evaluation' to refer to the ongoing process of studying biomarkers' success at acting as surrogates for individual clinical endpoints. According to this stringent view, treatment development using biomarkers does not ever prove conclusively that treatments result in particular clinical endpoints. Rather, the clinical trial process is one of progressively reducing uncertainty about the relationship between an intervention, a biomarker, and a clinical endpoint.

Setting aside that semantic debate, there are several levels of validity – or levels of surrogacy success – that must be considered and studied. At the simplest level is measurement validity: is the biomarker proposed as a surrogate capable of being measured objectively and reproducibly in a given case, and does it measure an objective, quantifiable characteristic successfully? One step beyond this, the internal or study validity of the surrogate must be evaluated: within the study, can the biomarker be measured not just with precision and reproducibility, but also with accuracy? In other words, within this study population and situation, does the biomarker correlate strongly with the clinical endpoint for which it is serving as a surrogate? The next level of validation is external validity: can this surrogate be shown to have similar predictive power in other populations or in other related treatment studies? If so, the biomarker can be considered to be a useful surrogate marker in studies that are closely related to the studies establishing its conditional 'validity'.

The final level of surrogacy success that must be considered has not, some commentators have observed, been given enough attention. Once biomarkers become established surrogate markers for predicting the effects of a given class of treatments on one clinical endpoint, can

they be safely relied upon to serve as surrogates for other related clinical endpoints? Or, can they be used as surrogate markers in evaluating other classes of treatments? The assumption has frequently been made in study design that biomarkers can be used broadly, once they become established in narrow research contexts. However, this scientifically unsound approach to trial design has in past decades led to flawed research conclusions, several of which have been considered in greater depth in review articles on the topic [4[•],5,6]. For years, researchers used suppression of arrhythmias as a surrogate endpoint for decreased morbidity due to cardiovascular disease, resulting in the approval of antiarrhythmia drugs (e.g., encainide, flecainide, moricizine) that later trials actually found to increase mortality in certain patient populations [4[•]]. More recently, a large and well publicized trial of the combination of two cholesterol-lowering drugs, ezetimibe and simvastatin, highlighted the risk of relying too much on biomarkers; although the combination treatment lowered patients' cholesterol levels more than simvastatin alone, it did not lead to any improvement in atherosclerosis or overall mortality, calling into question a great deal previous research that depended on the assumption that lowering cholesterol necessarily lowered morbidity and mortality [7]. In both these cases, as in many others, despite the best biological and statistical evidence, biomarkers that were 'validated' even in a series of previous trials were found to be poor predictors of clinical outcomes.

Biomarkers as always provisional

Biomarkers have been approved by the U.S Food and Drug Administration (FDA) regulation for use as surrogate endpoints in the treatment development process [8]. The FDA allows provisional intervention approval with surrogate marker-defined efficacy but requires further phase-IV follow-up studies that prove that relevant clinical endpoint correlation exists. Some cautious researchers and commentators have suggested that biomarkers are most effective in and best left for use as endpoints in phase I and phase II trials. Their use can help determine what potential treatments are worth the effort and resources of large, well powered phase III trials. There is, of course, the risk that researchers may inappropriately abandon large-scale research on treatments that, although actually effective at improving clinical outcomes, do not appear to be effective on the basis of biomarker analysis. This type of false negative has, in fact, already occurred in at least one case involving a trial of chronic granulomatous disease in children [9]. The trial, which was designed to measure both surrogate outcomes and true clinical endpoints, ultimately showed that the treatment, interferon gamma, was effective at reducing mortality, but there was no associated improvement on the surrogate outcome, production of

superoxide, which was expected to increase the patients' ability to kill bacteria. The biological process that led to improved clinical outcomes, in other words, was not captured by the biological mechanism proposed and predicted by the researchers. Had the trial been designed to rely first on interim analysis of the biomarkers alone, the true and clinically relevant effects of the treatment might never have been discovered. The lesson here is the same as in the cases of false positives: treatments' effects on biomarkers used as surrogate outcomes do not necessarily predict true clinical outcomes. Without confirmatory clinical endpoint analysis, the overreliance on biomarkers, even ones previously considered validated in particular treatment contexts, presents a serious and persistent risk of producing misleading, and in some cases dangerous, erroneous conclusions.

Conclusion

Biomarkers play a critical role in improving the drug development process, as well as in the larger biomedical research enterprise. Understanding the relationship between measurable biological processes and clinical outcomes is vital to expanding our arsenal of treatments for all diseases, and for deepening our understanding of normal, healthy physiology. Since at least the 1980s, the necessity of using biomarkers as surrogate outcomes in large trials of major diseases, such as cancer [10] and heart disease [11], has been widely discussed. The FDA continues to promote the use of biomarkers in basic and clinical research, as well as research on potential new biomarkers to use as surrogates in future trials [12]. However, for all their potential to do good – to speed drug development, to reduce exposure to ineffective experimental treatments, and so on – biomarkers present substantial risks when trial designers confuse them with clinical endpoints.

Biomarkers could only serve as true replacements for clinically relevant endpoints if we completely understood the normal physiology of a biological process, the pathophysiology of that process in the disease state, and effects of an intervention – pharmacological, device, or otherwise – on these processes. As we rarely, if ever, have the full picture of those types of processes, as there are always more details we do not know or understand, biomarkers as surrogate endpoints need constant reevaluation. Studies using biomarkers should always have, as ultimate

measures, clinical outcomes, at least for retrospective analysis of biomarker correlation success. Without continual reevaluation of the relationship between surrogate endpoints and true clinical endpoints, we risk again approving whole classes of drugs that either have no additional benefit or, worse, that harm patients.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

Additional references related to this topic can also be found in the Current World Literature section in this issue (p. 560).

- 1 Biomarkers Definition Working Group. Biomarkers and surrogate endpoints: • preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001; 69:89–95.

This paper was produced by the Biomarkers Definition Working Group, convened by the National Institutes of Health Director's Initiative on Biomarkers and Surrogate Endpoints. The working group brought together leaders in the field of clinical trials and biostatistics from the U.S. NIH, the FDA, academia, and industry to develop consistent, comprehensive definitions of terms relating to the use of biomarkers as surrogate outcomes, and these definitions are frequently cited as the standard in the literature.

- 2 WHO International Programme on Chemical Safety. Biomarkers in risk assessment: validity and validation; 2001. <http://www.inchem.org/documents/ehc/ehc/ehc222.htm>.
- 3 WHO International Programme on Chemical Safety. Biomarkers and risk assessment: concepts and principles; 1993. <http://www.inchem.org/documents/ehc/ehc/ehc155.htm>.
- 4 Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996; 125:605–613.
This widely cited article looks carefully at the past successes and failures of biomarkers as surrogate outcomes in clinical trials, across a variety of research fields. It also explores in some depth the biological circumstances under which 'validated' biomarkers can appear to be part of a well characterized biological process but, in fact, because of oversimplification of the processes involved, still mislead investigators.
- 5 D'Agostino RB Jr. Debate: the slippery slope of surrogate outcomes. *Curr Control Trials Cardiovasc Med* 2000; 1:76–78.
- 6 DeGruttola D, Fleming TR, Lin DY, Coombs R. Perspective: Validating surrogate markers: Are we being naïve? *J Infect Dis* 1997; 127:237–246.
- 7 Mitka M. Controversies surround heart drug study: questions about Vytorin and trial sponsors' conduct. *JAMA* 2008; 299:885–887.
- 8 Food and Drug Administration Modernization Act of 1997, 21 CFR314 <http://www.fda.gov/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticAct/FDCA/SignificantAmendments/totheFDCA/FDAMA/FullTextofFDAMALaw/default.htm>.
- 9 A controlled trial of interferon gamma to prevent infection in chronic granulomatous disease. The International Chronic Granulomatous Disease Cooperative Study Group. *N Engl J Med* 1991; 324:509–516.
- 10 Ellenberg SS, Hamilton JM. Surrogate endpoints in clinical trials: cancer. *Stat Med* 1989; 8:405–413.
- 11 Wittes J, Lakatos E, PRobstfield J. Surrogate endpoints in clinical trials: cardiovascular diseases. *Stat Med* 1989; 8:415–425.
- 12 Food and Drug Administration. Innovation and stagnation: challenges and opportunity on the critical path to new medical products (white paper); 2004, updated 2010. <http://www.fda.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/ucm204289.htm>.