How-to guide on biomarkers: biomarker definitions, validation and applications with examples from cardiovascular disease

V O Puntmann

Correspondence to:
Dr V O Puntmann,
Cardiovascular Section,
Department of Experimental Medicine, Imperial College London, Hammersmith Campus, Du Cane Road, London W12 0NN, UK; v.puntmann@imperial.ac.uk

Received 1 August 2008
Accepted 7 April 2009

ABSTRACT

A biomarker is a characteristic that can be objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention. Many commonly used tests in clinical practice can serve as biomarkers. The majority have been identified on the basis of insight or underlying physiology or biological mechanisms. With increasing knowledge and practical experience, some of these tests have evolved into a measurable end point in clinical research, applied as an indicator of change, for the better or worse. The traditional identification of biomarkers as an observational side product of clinical practice is increasingly turning into an industrialised process of biomarker discovery, supported by standardised paradigms of biomarker validation and translation from bench to bedside. The potential utility of biomarkers in clinical studies, investigating either new treatments or new strategies of clinical management, is capitalising on recent advances in technology, from molecular sciences to powerful imaging, bearing the promise of expediting the discovery of new treatments. In the active search for new biomarkers, many potential candidates can be considered side by side, allowing many failures but a few great winners. Biomarker discovery is an ongoing process, with translation being tested de novo in every single study, providing us with the opportunity to revise our knowledge of the complex scheme of human physiology and pathophysiology. In predicting what Nature has set in place, advances in technology may be only the first step. This review provides an introduction to the field of biomarker discovery and translation. It deals with evolving nomenclature, basic principles of the validation process, and, drawing on examples in cardiovascular medicine, their significance for clinical application.

A biomarker is a characteristic that can be objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention.[1 2] Many commonly used tests in clinical practice are biomarkers; biochemical tests provide soluble biomarkers, whereas physiological assessment and imaging measures provide anatomical and functional biomarkers. The majority have been identified on the basis of biological insight or underlying physiology. With increasing knowledge and practical experience, many of these tests have evolved into a measurable end point in clinical research, applied as an indicator of change, for the better or worse.

The traditional identification of biomarkers as an observational side product of clinical practice is increasingly turning into an industrialised process of biomarker research, supported by standardised paradigms of biomarker validation and translation from bench to bedside. The potential utility of biomarkers in clinical studies, investigating either new treatments or new strategies of clinical management, is capitalising on recent advances in molecular and imaging technologies, bearing the promise of expediting the discovery of new treatments. In the active search for new biomarkers, many potential candidates can be considered side by side, allowing many failures but a few great winners.

Cardiovascular medicine provides many examples of biomarker discovery. The Framingham Study, for instance, clearly underlined the link between everyday clinical assessments, such as blood pressure and serum cholesterol concentrations, and increased cardiovascular risk and helped greatly to develop effective preventive measures and successful treatments.[3] Intensive investigation into the origins of cardiovascular disease and its underlying pathophysiology further contributed a myriad of new biomarkers.[4] These ''measurable and quantifiable parameters'' play diverse roles: they can serve as indices of health compared with disease in physiological assessments, or as an estimation of disease risk, diagnosis and prognosis. Moreover, biomarkers can also give an indication of underlying metabolic or pathophysiological processes, or provide a reflection of substance exposure or interventions.[5] Many of these indices are potentially modifiable and thus provide targeted measures of risk reduction for preventive and therapeutic interventions.

Biomarkers have also gained an important role in the field of clinical management and have established a close link with bedside medicine, by providing metrics of quality in medical care alongside meaningful costing. With effective translation into many clinical guidelines, biomarkers can facilitate the delivery of evidence-based medical care.[6]

This review provides an introduction to the field of biomarker discovery and translation. It deals with evolving nomenclature and basic principles of the validation process and, drawing on examples from cardiovascular medicine, gives practical aspects of the clinical application of biomarkers.

BIOMARKER DEFINITIONS

Many terms with overlapping meanings have been used to describe the nature of disease and treatment effects, including biological markers, biomarkers,

surrogate markers, surrogate end points and intermediate end points, reflecting the multidisciplinary involvement (eg, clinicians, statisticians, industrialists) and variety of potential applications. To take advantage of the vast interest and promote the utility of biomarkers in the assessment of clinical research, a consensus on terminology among the various specialities was initiated in 2001. The National Institutes of Health (NIH) Biomarkers Definition Working Group proposed the first vocabulary framework (box 1).[1] The working group outlined the conceptual role of biological markers or "biomarkers" in describing biological measurements in therapeutic development and assessment. In terms of the proposed practical application, their greatest value is providing "the proof of concept" in early studies, such as efficacy and safety evaluations (in vitro studies in tissue samples, in vivo studies in animal models, and early-phase clinical trials). Biomarkers have many other valuable applications in early disease detection, monitoring of health status (such as staging), assessment of prognosis, and prediction of risk.

Undeniably, the most reliable way to assess the clinical impact of a therapeutic intervention or management strategy is through its effect on a well-defined clinical end point such as survival, myocardial infarction and stroke. "Clinical end points", framed to separate measurements or analyses of disease characteristics observed in a study or a clinical trial, reflect the effect of a therapeutic intervention. This standard may, however, be excessive for the evaluation of end points that take a long time to achieve and where large numbers of patients are needed to make up for the effect size. This is clearly a time-consuming and expensive undertaking, delivering results with much delay for a population in imminent need, rendering clinical trials geared towards clinical end points a somewhat impractical pursuit of perfection. To facilitate and expedite the assessment of clinical risk and benefits, "surrogate end points" can be used.[1 2 7] Surrogates are biomarkers that are intended to substitute for a clinical end point, based on epidemiological, therapeutic, pathophysiological or other scientific evidence. As only a few biomarkers come with solid and unequivocal evidence to achieve this privileged status, the use of surrogate end points in a clinical trial still requires clear specification of the clinical end point that is being substituted for, the class of therapeutic intervention being applied, and the characteristics of the population and disease state in which the substitution is being made. It is important to preclude the misconception that surrogates can literally replace the clinical end point overall, and not only in a specific and well-defined setting. A conceptual model was developed to show the relation of a biomarker to a clinical end point and the application of the biomarker as surrogate end point in the evaluation of therapeutic interventions (fig 1).[1] The model also shows that biomarkers may be useful in the assessment of safety or efficacy or both. Although narrowly defined for the purpose of their use in a clinical trial, surrogate end points may have broader objectives in clinical practice.

Box 1: National Institutes of Health (NIH) Definition Working Group definitions for biomarkers and clinical end points (modified by NIH[1] and Vassan[4])

- **Biomarker**: a characteristic objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention.
  - **Type 0 biomarker**: a marker of the natural history of a disease that correlates longitudinally with known clinical indices.
  - **Type I biomarker**: a marker that captures the effects of a therapeutic intervention in accordance with its mechanism of action.
- **Clinical end point**: a characteristic or variable that reflects how a patient feels, functions or survives.
  - **Intermediate (non-ultimate) end point**: a true clinical end point (a symptom or measure of function, such as symptoms of angina frequency or exercise tolerance), but not the ultimate end point of the disease.
  - **Ultimate end point**: eg, survival or the rate of other serious and irreversible morbid events.
- **Surrogate end point**: a biomarker intended to substitute for a clinical end point aiming to predict clinical benefit (or harm, or lack of benefit or harm) on the basis of epidemiological, therapeutic, pathophysiological or other scientific evidence.

MULTIPLE ROLES OF BIOMARKERS

Many tools of everyday clinical practice can serve as biomarkers. Some routine blood tests and physiological measurements, such as lipid concentrations and blood pressure, are increasingly being recognised as valuable measures. These very common indices of cardiovascular risk assessment can prove extremely valuable in designing and evaluating new treatments, such as lipid-lowering and antihypertensive therapies.[8–10] Indeed, when evaluating drugs for approval, drug regulatory agencies, such as the American Food and Drug Administration (FDA), consider effects on the level of the risk factor as treatment effects.[11] Risk factor levels can thus serve as surrogate end points for the outcome of primary interest—the incidence of cardiovascular disease. Imaging measures are also increasingly being used to facilitate the early appreciation of benefit or harm and to build on physiological insight.[12] Appreciation of effects on endothelial function, recognised as the earliest stage in atherosclerotic disease, serves as a useful example.[13 14] Measuring changes in the diameter of the brachial artery induced by forceful recovery of blood flow (after a few minutes of occlusion), either by ultrasound or MRI, is a valuable way of assessing whether a drug has helped to recover endogenous stores of nitric oxide that were depleted during vascular inflammation. In more long-term trials, in which appreciation of plaque volumes and risk of cardiac events is in question, imaging of superficially lying carotid arteries can be used.[15] The distance between the intima and media layers, also termed carotid artery intima-media (CIMT), appears to confidently reflect the severity of overall atherosclerosis and risk of future cardiovascular events.[16] Moreover, halting the process of further thickening, or progres-

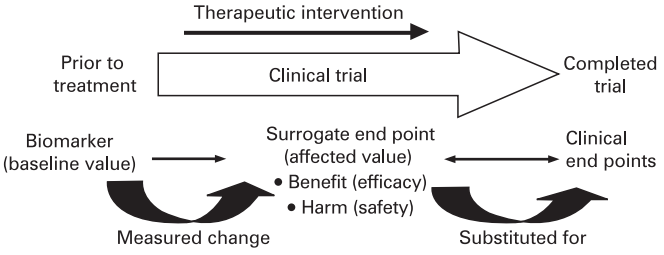

Figure 1 Conceptual model of relationships between biomarkers and surrogate end points in the process of therapeutic evaluation.

sion, or even reversing it, appears a logical ambition, now a routine addition in lipid-lowering trials. Indeed, many clinical investigations may go beyond just recognition of the disease. Imaging for instance is changing from a tool of visual assessment into a quantifiable methodology, allowing measurement of change in time and response to treatment.[12]

Conversely, biomarkers used as surrogate end points in clinical trials can also be applied in a range of clinical roles[2] (box 2):

1. indicators of disease trait (''screening biomarkers''), such as genetic testing and family screening with imaging in hereditary cardiomyopathies;
2. reflection of the disease state or rate of progression (''staging biomarkers''—asymptomatic/preclinical and clinical confirmation of diagnosis), such as the level of impairment of diastolic relaxation in heart failure or echocardiographic surveillance of asymptomatic aortic stenosis;
3. ''antecedent biomarkers'', providing a means of identifying people at high risk of developing disease (risk factors), facilitating early diagnosis with a high degree of accuracy;
4. ''prognostic biomarkers'', enabling clinicians to speculate on the outcome and response to treatment, such as establishing the presence of low-grade inflammation with high-sensitivity C-reactive protein in patients with multiple risk factors.

Irrespective of their role, the realistic applicability of biomarkers in a clinical setting involves features such as high sensitivity and specificity of an assay or investigation, providing the result with a clear message of either confirmation of disease (high positive predictive value) or exclusion (negative predictive value).[16] Moreover, in diagnoses in an acute clinical setting such as acute coronary syndrome, attributes such as a rapid sustained increase to allow early detection and in the critical time window, the amount of release proportional to disease extent and speed of assay conducive to point-of-care testing are critical.[17] Biomarkers used to monitor a chronic disease may resemble those used for diagnosis. However, the measurable effect of progression or response to treatment takes place longitudinally within the same patient, between the baseline values and those obtained at follow-up. Here, narrow intraindividual variation and tracking with disease outcome or therapy become crucial. To reflect therapeutic intervention, biomarker concentrations should relate to the underlying pathophysiology, usually substantiated in the context of preclinical research and prospective clinical trials. Review of clinical practice and opportunity cost analysis require biomarkers with a clearly applicable clinical role, where disease management affects the patient's outcome, with proof of cost-effectiveness.

MODERN BIOMARKER DISCOVERY

The discovery of biomarkers as an observational side product of clinical practice is giving way to active and industrialised research in which many potential candidates can be compared side by side. Modern investigational paradigms in biomarker discovery resemble those used in the assessment of drugs and therapeutics, borrowing much of the terminology and validation structure.[1–4] Biomarker discovery involves two complementary approaches: the ''knowledge-based'' (deductive method) and the ''unbiased'' (inductive strategy).[6] The knowledge-based strategy relies on a direct understanding of the underlying processes, such as the course of atherosclerosis and the evolution of its sequelae. It may consist of improving existing biomarkers to enhance their performance or designing assays for attractive new candidate markers, informed by the biology of the disease process. The unbiased approach involves examining tens of thousands of molecules using current technological advances to characterise the biomolecular signature of a stage of the disease.

Box 2: Multiple roles of biomarkers

- Antecedent: identifying the risk of developing an illness
- Screening: screening for subclinical disease
- Diagnostic: recognising overt disease
- Staging: categorising disease severity
- Prognostic: predicting future disease course/response to therapy

The standardised approach of validation includes comparison with a ''gold standard'' of practice of an adequate sample of patients within a defined spectrum of a certain condition. A sound validation process typically requires two sets of tests: an initial preliminary training set and a validation or test set.[6]

It is a common finding that the performance of biomarkers in subsequent studies differs from the original study, and solid two-level evidence of the original study is therefore important. To justify and appraise the role of biomarkers in clinical practice, evidence is needed that clinical management indeed varies with biomarker outcome, such as implementation in risk-reduction strategies.[16] A biomarker-guided approach needs to be over and above the current practice, clearly translating into better patient outcomes and reduced cost. This is usually assessed in the context of a randomised controlled trial (RCT), testing the risk-reducing interventions. Such RCTs aim to prove the effectiveness of screening and provide valuable data for cost-effectiveness analyses.

The results of an experiment performed by a particular researcher or group of researchers are generally evaluated by other independent researchers who repeat the experiment themselves, based on the original experimental description. Achieving high interobserver and intraobserver reproducibility of the measurements is important and relates to the agreement of test results with the same or different operators, test apparatus and laboratory locations. This agreement is often reported as a standard deviation (SD); a small SD reflects highly reproducible results and reliability of the method. The result values are said to be commensurate if they are obtained (in distinct experimental trials) according to the same reproducible experimental description and procedure. Despite the unequivocal utility and widespread availability of echocardiography the most commonly performed non-invasive investigation), echocardiographic measurements have historically suffered from problems of reproducibility. These measurements are highly dependent on the imaging skills of the operator and availability of acquisition windows, with often unfavourable anatomy. Some highly accurate but technically demanding measurements, such as assessment of the ejection fraction by M-Mode in parasternal long-axis views, are being actively discouraged in clinical use because of high levels of intraobserver and interobserver disagreement. MRI, on the other hand, with highly reproducible and accurate imaging quality, appears virtually free of technical limitations in the hands of any dedicated cardiac radiographer.

A GOOD BIOMARKER FOLLOWS THE RULES

Interpretation of biomarker results with a meaningful outcome is another crucial step in the process of validation. The accuracy

of a biomarker test is best expressed in terms of its sensitivity (SN: detection of disease when disease is truly present, ie, identifying true positives) and its specificity (SP: recognition of the true absence of disease).[17] The relationship between the two features, as well as the performance of the classifier in identifying the disease, can be visualised and studied using receiver-operating characteristic curves, where the performance of each biomarker is plotted against that of the clinical "gold standard".[18] Biomarker performance can also be represented by plotting the fraction of true positives (TPR = true positive rate) against the fraction of false positives (FPR = false positive rate), indicating the conditional probability of a positive test result for a random patient exceeding that for a random non-affected person (table 1).

The commonly used mnemonic "SPin/SNout" indicates that positive results from specific tests confirm the presence of a disease (SPin: test with high true negative rate (TNR)), whereas negative results from sensitive tests exclude it (SNout: test with high true positive rate (TPR)), as a "best of both" approach. In reality, few tests have the power to effectively "rule in" or "rule out" disease at all times. A typical example of a diagnostic test with good predictive values is dobutamine-stress wall-motion (DSMR) and perfusion imaging (MRP) with MRI. Merging the high-fidelity protocols, originally developed for use in echocardiography and perfusion scintigraphy, respectively, along with the high predictive values within a single and superior imaging modality, this procedure is excellent for assessment of high-risk patients: myocardial ischaemia detected by MRP, and DSMR can be safely used to identify patients at high risk of subsequent cardiac death or non-fatal myocardial infarction, thus providing high positive predictive value. An important goal of any stress modality is not only to identify patients at high risk, but also to distinguish those with a low cardiac event rate, providing negative predictive value. Indeed, normal MRI stress test results are associated with a very low annual cardiac event rate of ~1%, effectively providing a 3-year event-free warranty.[19]

As illustrated in the above example, we can apply the concept of probability as "a measure of a state of knowledge" or Bayesian probability and look at results as altering disease probability.[20] This requires an estimation of a pretest probability that will be adjusted up or down by the test result for a factor, called the likelihood ratio (LR). LR intends to portray the likelihood of a positive test result in someone with a disease compared with someone without the disease (LR+), and the likelihood of a negative result in someone with the disease compared with someone without the disease (LR−). In effect, LR converts a pretest probability into a post-test probability: the upward adjustment of the probability after a positive result is called the LR(+) and is a number >1, and the downward adjustment after a negative result is the LR(−) and is a fraction <1. Thus, ruling a disease in or out, or considering the likely outcome in order to guide management, depends on a comparison of post-test probability with thresholds for further action such as the probable presence of severe disease in someone with multiple risk factors, availability for further investigations (in-house availability of catheter laboratories or outside referral), or side effects of treatment (potentially teratogenic treatment of hypertension in a woman of child-bearing potential).

Table 1 Clinical decision-making toolbox

Result	Decision
True positive (TP)	Correct hit
True negative (TN)	Correct rejection
False positive (FP)	Type I error (false alarm)
False negative (FN)	Type II error (true miss)
True positive rate (TPR) (Sensitivity - SN)	TPR = TP/P = TP/(TP + FN)
False positive rate (FPR)	FPR = FP/N = FP/(FP + TN)
Accuracy (ACC)	ACC = (TP + TN)/(P + N)
Specificity (SP)	SP = TN/(FP + TN) = 1 − FPR
Positive predictive value (PPV)	PPV = TP/(TP + FP)
Negative predictive value (NPV)	NPV = TN/(TN + FN)

The key feature of the LR approach is that it incorporates both sensitivity and specificity.[21] Although theoretically independent, there is often a trade-off between the two features in reality, and it is important to recognise the situations when one can be sacrificed at the expense of the other. Imagine a biomarker indicating the diagnosis of a rare disease in asymptomatic individuals. In practice, it may be far more consequential to mislabel a healthy person than to miss an individual in whom disease is actually present. Here, we take the rule that LR+ (usually more than 10) is required to outweigh the burden of missing a person with a rare condition. Yet, the diagnosis of a rare disease remains an important goal to fulfil. Instead of applying expensive tests in the community, such as echocardiography, we can use a series of low-cost screening tests in "AND" fashion: the disease is considered likely when all tests are positive. Thus, although the sensitivity of the individual tests is lower, the specificity of the final results when used together is enhanced. A typical case is the investigation of causes of shortness of breath in a young person in whom ECG may be unreliable and an abundance of true negatives is a reasonable expectation. Here, a sequential strategy can be used with an intermediate step between ECG and echocardiography in the form of a reliable soluble biomarker, such as brain natriuretic factor, indicating involvement of the heart, where raised troponin confirms damage to the muscle, and negative D-dimer rules out increased R-sided pressures due to thromboembolic disease. The alternative is the use of multiple tests in parallel, where disease is considered likely when any of the tests are positive ("OR rule"), and sensitivity is increased at the expense of specificity.

When used in the diagnosis of a potentially life-threatening condition in a symptomatic patient, a biomarker should have the ability to exclude disease with high sensitivity.[21 22] Therefore, a test with high sensitivity is crucial, and low LR− (typically <0.10) is necessary. The cost and harm resulting from missing the diagnosis (predicting health when disease is likely) is greatly outweighed by the burden of any additional testing or treatment of the sequelae of the missed acute event. In a typical case of a patient with undiagnosed ischaemic chest pain, subsequent damage to the heart muscle leading to heart failure represents preventable clinical harm and a tremendous burden to the rest of the healthcare system. Biomarkers of timely risk stratification and prevention are therefore in great demand.

BIOMARKERS AND SURROGATE END POINTS IN DRUG DISCOVERY

Biomarkers serve a wide range of purposes in drug development, clinical trials and therapeutic assessment strategies.[5] They can provide a basis for the selection of lead candidates for clinical trials, for contribution to the understanding of the pharmacology of candidates, and for characterisation of the subtypes of disease for which a therapeutic intervention is most appropriate. When the measurement of a biomarker is considered for evaluating the response to a therapeutic intervention, it is

important to identify the purpose of the biomarker in the drug development and evaluation process. If it agrees closely with the clinical outcome of interest, such as morbidity and mortality, biomarkers may serve as surrogate end points for prediction of clinical outcomes.[23] Only a subset of biomarkers may achieve surrogate end-point status. By definition, characterisation of a biomarker as a surrogate end point requires it to "reasonably likely predict clinical benefit, based on epidemiologic, therapeutic, pathophysiologic or other evidence."[1 24 25] Again, a surrogate marker is only a useful measure if it is accurate, with high sensitivity and specificity for the expected outcome, highly reproducible in standardised assays, preferably minimally invasive, and acceptable to the patient. Likewise, it must be associated with an outcome message that is easily interpretable for the purpose of everyday clinical practice, supported by solid evidence that a knowledge of biomarker concentrations clearly affects the course of clinical management.

Measures of change in validated surrogate end points are more and more often considered as a basis for the regulatory approval of pharmaceutical agents.[26] RCTs can evaluate the effect of a therapeutic intervention on a surrogate end point and provide evidence for its efficacy or safety. When regulatory standards are met, a provisional approval of a new drug application may be granted. When approval is based on a surrogate end point or on an effect on a clinical end point other than survival or irreversible morbidity, this is linked with an expectation that this evaluation still needs to address the effects of the intervention on clinical outcomes at a later stage. The applicant is expected to study the drug further and to verify and describe its clinical benefit where there is uncertainty as to the relation of the surrogate end point to clinical benefit or of the observed clinical benefit to ultimate outcome.[26]

Typical surrogate end points used to assess the clinical efficacy of cardiovascular drugs include concentrations of low-density lipoprotein cholesterol and blood pressure.[8] Lowering cholesterol concentrations is at the heart of primary and secondary prevention strategies, and pharmaceutical interventions targeting cholesterol concentrations appeared to be in line with this. The objective of lipid-altering therapy is not merely to alter serum lipids but to reduce the morbidity and mortality from cardiovascular disease associated with abnormal serum lipid concentrations.[27] Yet, it is becoming clear that biochemical optimisation may only be partially responsible for reduction of the cardiovascular burden, and that the strategies to reduce vascular inflammation and stabilise plaques may be central to successful prevention. Trials in which the addition of single-target agents, such as ezetimibe (the only effect of which is lipid lowering), do not provide equivalent clinical results, such as halting progression in CIMT,[28] but another surrogate end point. Although the results of the ezetimibe clinical end point trial are awaited, the recent results of the JUPITER Study lend further support to targeting inflammation in primary prevention strategies.[29] In this study, intervention with rosuvastatin in apparently healthy people with acceptable concentrations of low-density lipoprotein (3.5 mmol/l), but slightly raised C-reactive protein (>2 mg/l), provides a significant clinical benefit of cardiovascular end point reduction. The dual role of statin therapy thus unfolds the meaning of the two surrogate end points, where lipid lowering becomes secondary to the evidence of reducing inflammation and plaque stabilisation.

LINKING THE BIOMARKER TO THE CLINICAL END POINT

Linking the biomarker to clinical outcome requires a firm promise of reliability for the surrogate end points through a structured approach and a substantial body of evidence. The original framework for the validation of biomarkers in atherosclerotic disease was proposed by Boissel[26] and subsequently adapted by Espeland *et al*.[30] There are three clinical characteristics of a successful surrogate marker: efficiency, linkage and congruence (box 3). To be considered efficient, the surrogate marker should display superior accessibility in terms of technical and temporal acquisition allowing accurate information acquisition within shorter time frames and clinical trials with fewer resources and less subject participation. By linkage, it is meant that a plausible underlying relationship between the surrogate marker and the clinical end point must be demonstrated and substantiated by comprehensive scientific evidence. For congruence, the surrogate should produce parallel estimates of risk and benefit as end points. Moreover, there should be a clear difference in surrogate marker measurements between individuals with and without the disease. In intervention studies, the expected clinical benefits should be deducible from the observed changes in the surrogate marker.

The strength of the biomarker is measured by estimating the proportion of treatment effect that is accounted for by the surrogate end point, where excellent performance equates with a valid surrogate end point accounting for all of the effects of the intervention on the clinical end point. Yet, the major concern about the use of biomarkers as surrogates for clinical end points is that, in most circumstances, not all treatment effects are fully accounted for by a single biomarker.[24 31 32] If the clinical outcome of interest is influenced by numerous factors, residual confounding factors may reduce the validity of the surrogate end point.[23] For practicality and to circumvent the shortcomings of many potentially useful candidates, a surrogate end point is regarded as useful if it explains at least 50% of the effect of an exposure or intervention on the outcome of interest.

Surrogate end points offer the advantage that they may be gathered in a shorter time frame and with less expense than the usual clinical end points in comparison with large RCTs, involving thousands of participants and substantial follow-up of clinical events. Linking the underlying pathophysiology and exposure/intervention of interest is an important advantage of many surrogates, facilitating the assumptions of causality more readily than the distant clinical events.[24 25] Biomarker discovery and validation of new surrogate end points undeniably appear to be at the centre of advances in cardiovascular medicine. Government agencies, academic research institutions and healthcare-associated industries all recognise the importance of biomarkers, not only in unravelling the underlying pathophysiology of disease, but also in promoting the discovery of novel

Box 3: Clinical criteria for surrogacy,[26] modified by Espeland *et al*[30]

- Efficiency
  - easy to measure (preferably non-invasively)
  - changes in the surrogate marker precede the clinical end points
- Linkage
  - supported by epidemiological and clinical studies
  - pathophysiological explanation
  - expression of risk
- Congruence
  - estimates of risk and benefits parallel the clinical end points

Key learning points

- A biomarker is a measurable characteristic of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention.
- A clinical end point is a characteristic that reflects how a patient feels, functions or survives, and a surrogate end point is a biomarker that is intended to reliably substitute for the former, within a narrowly defined clinical setting (such as a clinical trial) and a shorter time span.
- Only a few biomarkers attain the privileged status of a surrogate end point, through a structured approach and a substantial body of evidence.
- Biomarker discovery has evolved from coincidental side products of clinical practice into active research in which many new candidates can be compared side by side.
- Evolution of biomarkers represents a well-coordinated effort in a multidisciplinary environment allowing effective translation from the bench to the bedside.

drug targets. In the early stages of drug development, biomarkers are a well-established tool in evaluating early signals of efficacy and safety, selecting dosage, and identifying the target population. Advances in cellular and molecular pathophysiology and mechanism-driven pharmacology, alongside the growing epidemic of cardiovascular disease, highlight the need for accelerated drug development, in which surrogate markers may play important role. Moreover, the increasing costs and lengthy process of conducting RCTs, and the inevitable regulatory reviews, all reduce efficient throughput in drug development. To allow timely access to new therapies, there is potentially tremendous value to public health in accelerating the discovery and development processes for cardiovascular therapeutics through smaller, shorter studies, using well-chosen and validated surrogate end points. Providing the evidence of ongoing clinical end point trials to verify these results against mortality and irreversible morbidity in retrospect may become the acceptable norm.

DISCOVERING BIOMARKERS OF CARDIOVASCULAR DISEASE

It is widely agreed that inflammation within the vessel wall is present at all stages of atherothrombotic disease, from the development of early lesions to the occurrence of clinical events. Many locally and systemically released biological mediators are involved in a dynamic and delicate balance through different stages of disease.[33–35] Accordingly, biomarkers may have a variety of functions, corresponding to different stages in disease development. The bio-signature of a particular stage of a disease is potentially valuable in the differentiation of pathophysiological responses in the presence of disease, and their features may vary with their intended use.[35] An improved understanding of cardiovascular pathophysiology, and the interrelated roles of vascular inflammation, oxidative stress in cell/tissue injury and responses, has increased insight into the role of various signalling pathways and mediators. Indeed, efforts to ascertain the role of many of these substances as biomarkers are greater than ever.

The development of cardiovascular biomarkers can be challenging for several reasons. Atherosclerosis is a generalised systemic disease and is inherently so complex that a simple set of biomarkers is unlikely to capture most of the individual predisposition to develop cardiovascular disease or herald its sequelae. Despite its complexity, a recently proposed triad of vulnerable systems—vulnerable plaque, vulnerable blood and vulnerable myocardium—designate possibilities for biomarker discovery and development of new treatment targets.[36 37] In terms of developing biomarkers, all three components are becoming available for appreciation within a window of opportunity, available for biomarkers to emerge and allow intervention. The recent advent of emerging technologies, multidisciplinary life-science approaches, and advances in imaging have increased recognition that atherosclerosis arises out of the dynamic dysregulation of several gene regulatory networks, proteins and metabolic alterations, reflected in complex pathophysiological perturbations. Leading to particular phenotypes of pathology, this is likely to be amenable to morphological characterisation in vivo, such as with imaging.

Integrating molecular biology with clinical physiology may thus increase the potential to decipher the complex bio-signature in time for intervention. Indeed, ''systems biology'' of soluble biomarkers, coupled with ''systems physiology'' by means of imaging and tissue characterisation, is gaining ground and applicability in the estimation of risk and detection of subclinical disease in high-risk patients. In addition to the aforementioned examples on the use of CIMT and assessment of endothelial dysfunction, detailed plaque characterisation via

Current research questions

- Biomarker discovery has entered a new promising stage of development. Expediting the process through simultaneous discovery may yield more potential candidates.
- Discovery of reliable biomarkers through robust validation methodologies remains an important priority. This, however, needs to reflect the variety of possible applications and demands flexible, purpose-driven processes in well-defined clinical settings where the role of biomarkers is being addressed.
- Single biomarkers are unlikely to capture the complex process of human pathophysiology. Research may need to be geared towards sets of biomarkers, reflecting different, but intercalated, processes describing the human condition.

Key references

1. Biomarkers Definitions Working Group. Biomarkers and surrogate end points: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001;**69**:89–95.
2. Chambless LE, Folsom AR, Davis V, *et al*. Risk factors for progression of common carotid atherosclerosis: the Atherosclerosis Risk in Communities Study, 1987–1998. *Am J Epidemiol* 2002;**155**:38–47.
3. Prentice RL. Surrogate end points in clinical trials: definition and operational criteria. *Stat Med* 1989;**8**:431–40.
4. Kastelein JJ, de Groot E. Ultrasound imaging techniques for the evaluation of cardiovascular therapies. *Eur Heart J* 2008;**29**:849–58.
5. Neubauer S. Cardiac physiology investigated by new methods of imaging. *Clin Med* 2007;**7**:189–91.

ultrasound or tomographic imaging is now possible, supporting the estimation of vascular instability and associated risk of events.[38] Indeed, various imaging techniques can contribute important information because of their inherent advantages: whereas MRI provides broad access to plane visualisation and detailed and morphological characterisation with no radiation,[39] CT coronary angiography may be able to provide fast and non-invasive luminal information, once the problems of radiation exposure and resolution comparable to invasive coronary angiography are resolved.[40]

Molecular imaging—biology and physiology hand-in-hand—continues to hold great promise for direct quantification of the causative molecular constituents of disease in time and space. Potential future tools in cardiovascular disease are targeted contrast agents for vascular inflammation and plaque characterisation, identifying expression of adhesion molecules, activation of matrix metalloproteinases, and early angiogenic expansion of the vasa vasorum that supports plaque development. Up to the point of clinical applicability, intravascular ultrasound and optical coherence tomography continue to be important (but invasive) tools of plaque characterisation.[41 42]

FUTURE PERSPECTIVES

Biomarkers and surrogates have a clear role in integrated clinical physiology, drug discovery and device development. It is timely for the biomarker industry to expand, and evaluation of potential candidates should be encouraged alongside clinical end-point-driven trials. Successful candidates, performing efficiently and cost-effectively in rapid ''rule out'' or ''rule in'' strategies and facilitating early triage of patients into low-risk and high-risk treatment groups, will be quickly integrated into clinical decision-making protocols. Biomarkers that facilitate choice of the most appropriate drug and enable titration of drug dose to avoid side effects or to screen for potential drug interactions are likely to be attractive to clinicians. In the NHS, screening biomarkers will compete for limited healthcare budgets, and only those with excellent performance characteristics will find support for their clinical utility. An increasingly self-informed public may access some of the biomarker assays on an over-the-counter basis or choose to undergo self-screening investigations for a fee. Here, consumer-safety governance plays an important role in ensuring quality control and avoiding over-interpretation. Biomarkers that guide routes of prevention or diagnosis of life-threatening events will probably continue to be clinically useful.

It is important to emphasise that biomarker discovery is an ongoing process, with translation of established biomarkers being tested de novo in every study, providing us with the opportunity to revise our knowledge of the complex systems of human physiology and pathophysiology. In predicting what Nature has set in place, advances in technology may only be a minor step.

CONCLUSIONS

Biomarkers represent an exciting field, with a variety of potential applications at the crossroads of scientific achievements, discovery of new drugs and clinical approaches. The introduction of formal terminology with a conceptual framework has facilitated communication in multidisciplinary environments. Biomarkers, which evolved as unintended products of clinical practice, have now moved into an era of industrialised active research and parallel comparisons of many candidates. Biomarkers can play diverse roles in the clinical setting, facilitating the management of clinical conditions from diagnosis to prognosis and guidance of treatment. In the narrowly defined setting of clinical trails, they can be used as surrogate end points, allowing shorter drug testing periods and expediting development of new treatments. Taken together, the evolution of biomarkers represents the coordinated and concerted effort of basic scientists, clinicians, technology experts, epidemiologists, statisticians, academic and industrial sponsors, and regulatory agencies within a cooperative framework.

MULTIPLE CHOICE QUESTIONS (TRUE (T)/FALSE (F); ANSWERS AFTER THE REFERENCES)

1. Biomarkers

A. Feature a unique and case-specific characteristic
B. Are limited to the serological tests used
C. Are increasingly being developed as a side product of common practice
D. Many represent a reliable measure of clinical outcome

2. Randomised controlled trials

A. Represent a cornerstone of clinical research
B. Aim to deliver exclusively clinical end points
C. Cost-effectiveness is not regarded as a trial end point
D. Provide evidence applicable to clinical practice

3. Surrogate end points

A. Can reliably replace clinical end points in a variety of clinical scenarios
B. Can substantially shorten the drug development process
C. Represent a separate category from biomarkers
D. Are only useful when easy to obtain and reproducible

4. Cardiovascular disease

A. Is easy to predict with a single biomarker
B. Imaging biomarkers of risk, such as CIMT, is overused in everyday clinical practice
C. Cholesterol concentrations reliably reflect the risk of future events
D. The presence of subclinical inflammation predicts less risk of cardiac events

5. Biomarker development

A. Is purely coincidental and does not require expertise
B. Is likely to see a decline within a decade
C. Requires a single-discipline approach in appraisal processes
D. Real-life application is tested de novo in every new study

Acknowledgements: The editorial contribution of Mr Chris Harris, of Imperial College London, is greatly appreciated.

Competing interests: None.

Provenance and peer review: Not commissioned; externally peer reviewed.

REFERENCES

1. **Biomarkers Definitions Working Group**. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001;**69**:89–95.
2. **Tardif JC,** Heinonen D, Orloff D, *et al*. Vascular biomarkers and surrogates in cardiovascular disease. *Circulation* 2006;**113**:2936–42.
3. **Anderson KM,** Odell PM, Wilson PW, *et al*. Cardiovascular disease risk profiles. *Am Heart J* 1991;**121**:293–8.
4. **Vasan RS.** Biomarkers of cardiovascular disease: molecular basis and practical considerations. *Circulation* 2006;**113**:2335–62.

5. **Colburn WA.** Optimizing the use of biomarkers, surrogate endpoints, and clinical endpoints for more efficient drug development. *J Clin Pharmacol* 2000;**40**:1419–27.
6. **Wade A.** Derivation versus validation. *Arch Dis Child* 2000;**83**:459–60.
7. **Fox N,** Growdon JH. Biomarkers and surrogates. *Neuro Rx* 2004;**1**:181.
8. **Ridker PM,** Rifai N, Rose L, *et al*. Comparison of C-reactive protein and low-density lipoprotein cholesterol levels in the prediction of first cardiovascular events. *N Engl J Med* 2002;**347**:1557–65.
9. **Kannel WB.** Epidemiology of essential hypertension: the Framingham experience. *Proc R Coll Physicians Edinb* 1991;**21**:273–87.
10. **Lewington S,** Clarke R, Qizilbash N, *et al*. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet* 2002;**360**:1903–13.
11. **The Food and Drug Modernization Act of 1997**. Title 21 Code of Federal Regulations Part 314 Subpart H Section 314.500.
12. **Smith JJ,** Sorensen AG, Thrall JH. Biomarkers in imaging: realizing radiology's future. *Radiology* 2003;**227**:633–8.
13. **Anderson TJ,** Uehata A, Gerhard MD, *et al*. Close relation of endothelial function in the human coronary artery and peripheral circulations. *J Am Coll Cardiol* 1995;**26**:1235–41.
14. **Leeson CP,** Robinson M, Francis JM, *et al*. Cardiovascular magnetic resonance imaging for non-invasive assessment of vascular function: validation against ultrasound. *J Cardiovasc Magn Reson* 2006;**8**:381–7.
15. **Chambless LE,** Folsom AR, Davis V, *et al*. Risk factors for progression of common carotid atherosclerosis: the Atherosclerosis Risk in Communities Study, 1987–1998. *Am J Epidemiol* 2002;**155**:38–47.
16. **Kattan MW.** Evaluating a new marker's predictive contribution. *Clin Cancer Res* 2004;**10**:822–4.
17. **Pepe MS,** Thompson ML. Combining diagnostic test results to increase accuracy. *Biostatistics* 2000;**1**:123–40.
18. **Hanley JA,** McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;**143**:29–36.
19. **Jahnke C,** Nagel E, Gebker R, *et al*. Prognostic value of cardiac magnetic resonance stress tests: adenosine stress perfusion and dobutamine stress wall motion imaging. *Circulation* 2007;**115**:1769–76.
20. **Jaynes ET.** *Probability theory: the logic of science*. Cambridge: Cambridge University Press, 2003.
21. **Sackett DL,** Haynes RB, Guyatt GH, *et al*. *The interpretation of diagnostic data: clinical epidemiology, a basic science for clinical medicine*. Boston, MA: Little, Brown, 1991:69–152.
22. **Batstone G.** *Practising by the evidence: the role of pathology*. *J Clin Pathol* 1997;**50**:447–8.
23. **Prentice RL.** Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med* 1989;**8**:431–40.
24. **De Gruttola VG,** Clax P, DeMets DL, *et al*. Considerations in the evaluation of surrogate endpoints in clinical trials: summary of a National Institutes of Health workshop. *Control Clin Trials* 2001;**22**:485–502.
25. **Freedman LS,** Graubard BI, Schatzkin A. Statistical validation of intermediate endpoints for chronic diseases. *Stat Med* 1992;**11**:167–78.
26. **Boissel JP.**Standards for reporting clinical trials. *Clin Trials Metaanal* 1992;**28**:5–7.
27. **Visser ME,** Jakulj L, Kastelein JJ, *et al*. LDL-C-lowering therapy: current and future therapeutic targets. *Curr Cardiol Rep* 2008;**10**:512–20.
28. **Kastelein JJ,** Akdim F, Stroes ES, *et al*. ENHANCE Investigators. Simvastatin with or without ezetimibe in familial hypercholesterolemia. *N Engl J Med* 2008;**358**:1431–43.
29. **Ridker PM,** Danielson E, Fonseca FA, *et al*. Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. *N Engl J Med* 2008;**359**:2195–207.
30. **Espeland MA,** O'leary DH, Terry JG, *et al*. Carotid intimal-media thickness as a surrogate for cardiovascular disease events in trials of HMG-CoA reductase inhibitors. *Curr Control Trials Cardiovasc Med* 2005;**6**:3.
31. **Lin DY,** Fleming TR, DeGruttola V. Estimating the proportion of treatment effect explained by a surrogate marker. *Stat Med* 1997;**16**:1515–27.
32. **Fleming TR,** DeGruttola V, DeMets DL. Surrogate endpoints. In: Armitage P, Colton T, eds. *Encyclopedia of biostatistics*. New York: John Wiley & Sons, 1998;**6**:4425–31.
33. **Libby P.** Lesion versus lumen. *Nat Med* 1995;**1**:17–18.
34. **Libby P.** Inflammation in atherosclerosis. *Nature* 2002;**420**:868–74.
35. **Libby P,** Theroux P. Pathophysiology of coronary artery disease. *Circulation* 2005;**111**:3481–8.
36. **Naghavi M,** Libby P, Falk E, *et al*. From vulnerable plaque to vulnerable patient: a call for new definitions and risk assessment strategies: part I. *Circulation* 2003;**108**:1664–72.
37. **Naghavi M,** Libby P, Falk E, *et al*. From vulnerable plaque to vulnerable patient: a call for new definitions and risk assessment strategies: part II. *Circulation* 2003;**108**:1772–8.
38. **Kastelein JJ,** de Groot E. Ultrasound imaging techniques for the evaluation of cardiovascular therapies. *Eur Heart J* 2008;**29**:849–58.
39. **Neubauer S.** Cardiac physiology investigated by new methods of imaging. *Clin Med* 2007;**7**:189–91.
40. **Miller JM,** Rochitte CE, Dewey M, *et al*. Diagnostic performance of coronary angiography by 64-row CT. *N Engl J Med* 2008;**359**:2324–36.
41. **von Birgelen C,** Hartmann M, Mintz GS, *et al*. Relationship between cardiovascular risk as predicted by established risk scores versus plaque progression as measured by serial intravascular ultrasound in left main coronary arteries. *Circulation* 2004;**110**:1579–85.
42. **Rieber J,** Meissner O, Babaryka G, *et al*. Diagnostic accuracy of optical coherence tomography and intravascular ultrasound for the detection and characterization of atherosclerotic plaque composition in ex-vivo coronary specimens: a comparison with histology. *Coron Artery Dis* 2006;**17**:425–30.

Answers

1. (A) T; (B) F; (C) F; (D) F
2. (A) T; (B) F; (C) F; (D) T
3. (A) F; (B) T; (C) F; (D) T
4. (A) F; (B) F; (C) F; (D) F
5. (A) F; (B) F; (C) F; (D) T