Protein biomarker immunoassays opportunities and challenges

With the biopharma industry under heavy pressure, transforming the scientific interchange between the laboratory and the clinic is exceedingly critical. Realising the costly toll of drug attrition in late-stage development, many companies are turning to translational medicine strategies to aid ‘go/no-go’ decisions in preclinical and clinical development. The incorporation of informative biomarkers as predictive tools in clinical trials is increasingly viewed as enabling not only smarter decisions and resource investment but also as potentially yielding new targeted medicines guided by companion diagnostics. Critical for the application of biomarkers is their ‘qualification’, i.e. the demonstration that they yield reliable and meaningful information. While early work has been strongly focused on nucleic acid-based biomarkers (DNA SNPs and mRNA expression profiles), recent experience suggests that the utility of these markers as clinically applicable decision tools may generally be limited. Protein biomarkers, which offer a significantly greater degree of differentiated information content, are likely to close this gap. Immunoassays continue to be the most sensitive, specific and selective technology to interrogate such markers. The development of innovative, suitable and fit-for-purpose protein-based assays represents a significant challenge both with regard to the technical aspects of antibody and assay development and the clinical validation of the analytes targeted. This article will focus on protein biomarkers and antibodies for immunoassay development as we coin the term ‘fit-for-purpose’ antibody.

Personalised healthcare concepts – the role of biomarkers: Most significant advances in medicine have been characterised by an increasingly differentiated approach to patients’ treatment, based on a more sophisticated understanding of disease causation. Thus, the advent of antibiotics allowed the specific treatment of those inflammatory disorders that are related to infection, and the development of subclasses of antibiotics the differentiated, and thus more effective, treatment of infections caused by different microorganisms. Of late, the further refinement of medical intervention guided by an emerging deeper understanding of disease- and patient-heterogeneity has been labelled ‘personalised medicine’, and biomarkers are heralded as the carriers of the
information providing this understanding and differentiation. The personalised medicine concept is shown in Figure 1.

Definition: A biomarker is “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention”. While this article focuses on protein biomarkers, many other types of biomarkers are being used (including RNA, DNA, metabolites, but also modalities such as medical imaging, verbal tests, physical assessment, etc).

Biomarker versus diagnostic: The terms diagnostic or diagnostic testing are often used interchangeably with biomarker and biomarker testing. However, while all diagnostics are biomarkers, the reverse is not true: what is commonly referred to as a ‘biomarker’ most often only has the potential to become a diagnostic. In other words, most so-called biomarkers are still in an exploratory phase, and are thus assayed by research-grade tests, requiring laboratories with highly trained specialists. On the other hand, diagnostic tests are fully developed and highly validated ‘kit’ products that are more straightforward to use. Diagnostic kits can therefore be used in doctor’s offices, hospitals, and even by the patient as over-the-counter products, as is the case with home pregnancy tests. This article focuses specifically on biomarkers in the developmental phase, and in particular on markers where the analyte is a protein.

Value of biomarkers: Biomarkers are increasingly viewed as a key adjuvant to drug discovery and development to mitigate the low success rate and high cost the pharmaceutical industry is experiencing. The role of biomarkers spans all aspects of drug discovery and development. Biomarkers may be used as tools for target discovery, early target assessment, for evaluation of a medicine’s mechanism of action, for dose determination, for prediction of drug effects (efficacy as well as adverse reaction), for patient selection/stratification, for

<table>
<thead>
<tr>
<th>BIOMARKER</th>
<th>TEST</th>
<th>DISEASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prognostic</td>
<td>BRCA1 and BRCA2</td>
<td>Breast Cancer</td>
</tr>
<tr>
<td>Diagnostic</td>
<td>Serum PSA</td>
<td>Prostate Cancer</td>
</tr>
<tr>
<td>Predictive</td>
<td>K-RAS</td>
<td>Colorectal Cancer</td>
</tr>
</tbody>
</table>

Table 1: Personalised medicine for disease management – current examples
therapy monitoring, and for prognostication. While single analytes are the simplest application of biomarkers, the desired specificity and selectivity may sometimes require the examination of panels of protein and other biomarkers.

**Personalised medicine:** Biomarkers, by virtue of their potential to differentiate among disease states as well as patient characteristics, are essential for the realisation of personalised medicine and provide a critical link in the bench-to-bedside research effort which translational medicine represents. While oncology is the area of medicine where we have seen most progress to date in the implementation of biomarkers, and thus already a clear move towards personalised medicine, pharmaceutical companies are gearing up to extend the concept of personalised therapies to many disease areas. Two major challenges – selection of subpopulations of patients likely to show a favourable treatment response, and monitoring therapy efficacy – can both be addressed with the help of suitable biomarkers. Although biomarkers will play an essential role in realising the envisioned future potential of personalised medicine, their utility requires a substantive effort directed at their clinical qualification in appropriately designed clinical trials with well-defined, relevant endpoints and reference cohorts.

**DNA/RNA vs protein biomarkers:** While platforms for high-throughput analysis of DNA variants and mRNA expression profiles have been developed and are in broad use, the results, with regard to the discovery of clinically applicable biomarkers, have so far been somewhat disappointing, owing perhaps to the inability of nucleic-acid based markers to integrate information related to the downstream processing of proteins that may significantly impact the information content of a marker. Also, environmental and lifestyle confounders are not at all or less likely to be integrated into and reflected by signals obtained from nucleic-acid-based assays.

**Stages of biomarker development:** The development of a clinically applicable biomarker is a complex process requiring sequential and iterative steps. Often overlooked yet critical to this process are proper preclinical and clinical study design and sample collection. Therefore, these will be given special attention here. Thus, the development stages for biomarkers are: 1) study design, 2) sample collection, 3) biomarker discovery, 4) biomarker assay development and initial clinical validation, and 5) confirmation and validation of the biomarker in additional, independent studies. Each stage is dependent on the previous stage, and at each step a close interaction between drug- and

---

**Figure 2**
The diagnostic biomarker immunoassay workflow depicts the different phases of diagnostics biomarker development, the interrelationship between the different components during development and the nature of the iterative processes in development to commercial launch.
Biomarkers

Fig. 3
The diagram illustrates the concept of ‘fit-for-purpose’ antibody through use in various applications. For instance the requirement for an ELISA is substantially different from that of IHC or diagnostic assay vs laboratory assay.

assay-developers, and between pharmaceutical and diagnostic experts will be needed in a hand-in-glove fashion. The stages of companion diagnostic biomarker development are diagramed in Figure 2.

Importantly, biomarker R&D needs to be carried out in parallel with the discovery and development process of a new drug – all too often it still happens as an afterthought once the medicine has entered the stage of clinical trials. While a biomarker will ultimately have to be validated in the context of human pathobiology, there are important opportunities of using biomarkers, and of characterising their potential and attributes, already in the preclinical phase, principally in animal models. This raises additional challenges, as it requires affinity reagents useable in animals as well.

Protein biomarker technology platforms: hypothesis-driven vs hypothesis-free approaches

Protein biomarker assay platforms: Two types of protein assay platforms are currently applied to discover protein biomarkers and to measure them quantitatively and qualitatively (i.e., to determine the isoform state of a protein such as phosphorylation). It is instructive to point out here that an antibody use is a ‘fit-for-purpose’ antibody. For example, the requirement for an ELISA is substantially different from that of IHC or diagnostic assay vs laboratory assay. Figure 3 illustrates the concept of ‘fit-for-purpose’ antibody through use in various applications.

Immunooassays: direct use of antibodies

Immunofluorescence-based assays are the mainstay of testing for proteins. They use antibodies directed against the protein or isoform of interest. Detection of the antibody-antigen (protein) complex provides the quantitative measurement of the amount of antigen present in the sample. A variety of methods are used that vary both by how the antibody and antigen come into proximity of each other to form a complex (based on what the antibody or antigen is fixed to) and by the detection method used to monitor the amount of complex. Western blots are the simplest and most widely used immunoassay method in biomedical research; ELISA (enzyme-linked immune-sorbent assay) is the method most often used in clinical settings (e.g., PSA test). A number of platform technologies offer methods for multiplexed and miniaturised immunoassays (e.g., Luminex, MesoScale Discovery and PerkinElmer). Development of antibody-based assays is a time-consuming, resource-intensive effort and frequently hampered by cross-reactivity to other antigens. Moreover, results from immunoassays often do not discriminate among
Biomarkers

closely related forms of specific proteins. Currently, commercial tests are available for more than 200 different proteins (using various methodologies); custom-immunoaffinity assay services for others are provided by a number of specialty-providers.

Mass spectrometry assays: hidden importance of antibodies

While mass spectrometry has been widely used over the years for hypothesis-free detection of protein biomarkers, its application has been impeded by lack of sensitivity and the non-quantitative nature of the tests. More recently, a variant of the technology commonly referred to as peptide MRM (multiple reaction monitoring) is gaining importance as a more quantitative variant for protein biomarker measurements of this platform. Peptide MRM can combine the high selectivity and specificity of MS for the protein of interest with impressive quantitative accuracy and dynamic range. Quantitation obtained by this method is based on the peak area for the mass spectra data of the analyte relative to a known quantity of an isotope-labelled standard. The peak area can be used to provide relative quantitation (similar to most immunoaffinity assays) or absolute quantitation (protein concentration)\(^2\). Proteins at low abundance levels in samples will require a method to enrich for the protein(s) of interest. One such enrichment method is immuno-enrichment with antibodies to the protein or the peptides\(^3\), another one is the immune-adsorption-based depletion of the sample of abundant protein species. Thus, even MS is highly dependent on antibody technology. Such is the case of the MRM assay, which uses the stable isotope standards and capture by antigen-antibody reactions. In the case of parathyroid hormone (PTH), MSIA enables them to quantify protein variants. In the case of parathyroid hormone (PTH), MSIA enables them to quantify protein variants. In the case of parathyroid hormone (PTH), MSIA enables them to quantify protein variants.

Steps towards a meaningful protein immunoassay

Antibody availability and quality: A number of public domain initiatives to develop antibodies against all human proteins (eg HUPO Antibody Initiative [www.hupo.org]; FP7 Program ‘Affomics’) are currently being undertaken. Both immuno- and MS-based protein assay platforms will benefit from these initiatives as they will provide critical assay reagents. Antibodies are essential for immuno-affinity-based assays, where the selectivity of the antibody needs to be as high as possible. Mass spectrometry-based assays use antibodies to enrich a complex sample (eg plasma) for the proteins of interest; here, the specificity of the antibody does not have to be as high since the antibody is not used to characterise/identify the protein of interest. Therefore, a ‘fit-for-purpose’ antibody is viewed based on application and utility.

Analytical accuracy: Analytical accuracy refers to a number of metrological parameters that determine the performance of an assay on the level of technical specification. The term represents a composite assessment that comprises both random
Biomarkers

and systematic influences, ie, both precision and trueness. Its numerical value is the total error of measurement. Other metrics of analytical performance include a number of different ‘detection limits’ that define different properties of the assay, including the Lower Limit of Detection (LLOD), the Instrument Detection Limit (IDL), the Method Detection Limit (MDL), the (Lower) Limit of Quantitation (LOQ or LLOQ), and the Practical Quantitation Limit (PQL), as well as the Coefficient of Variation (CV) as a normalised measure of dispersion of a probability distribution (defined as the ratio of the standard deviation to the mean). In addition, parameters such as analytical range, analyte stability, standard stability and reagent stability need to be tested and described as part of analytical performance characterisation.

Diagnostic accuracy and clinical validation: Diagnostic accuracy, determined by a process commonly referred to as clinical validation, refers to the degree with which results of a test concur with what would be considered the current gold standard of clinical assessment of the interrogated question. This may be another biomarker (ie an established reference test) or – the real gold standard – a clinical outcome or endpoint, such as survival or death. Accuracy can be expressed through sensitivity and specificity, positive and negative predictive values, or positive and negative diagnostic likelihood ratios. Each measure of accuracy should be used in combination with its complementary measure: sensitivity complements specificity, positive predictive value complements negative predictive value, and positive diagnostic likelihood ratio complements negative diagnostic likelihood ratio. All of these parameters are not intrinsic to the test and are determined by the clinical context in which the test is employed. A summary of the characteristics, and the strengths and weaknesses of these metrics is presented in Table 2.

Biomarker validation requirements typically progress from the early to the later stages of drug discovery and development. Research-grade (‘home-brew’) assays are sufficient in the early, exploratory stage where biomarkers are discovered or interrogated as putative markers. As evidence is accumulated and the likelihood of clinical validation rises, assays that fulfill somewhat higher quality requirements (such as ‘analyte-specific reagents’) will be used, in a setting that is more controlled (eg CLIA accreditation). In yet later stages of clinical development, ultimately fully approved in vitro device status for regulatory authority approved tests fit for commercialisation is required. If a biomarker is to be used as a

Figure 4
The drawing shows the utility of immunoassay and mass spectrometry technology platforms. The role, application as well as the potential benefit derived from each platform is mapped to the development stage. [Sources: Adapted from MIT Broad Institute and NCI/CPTC websites and E.A. Dalnasso, GEN, Planning for Success in Biomarker Discovery Appropriate Proteomics Platform and Careful Study Design Can Improve Positive Results, Jun 15 2008 (Vol. 28, No. 12)]
diagnostic on the label for a marketed therapy then typically samples from Phase III trials are used to demonstrate the value of the biomarker5.

Clinical utility: The term ‘clinical utility’, while widely used, is ill-defined. It is commonly used as a synonym for studies of clinical effectiveness and/or economic evaluations. The most basic definition of clinical utility refers to an estimation of the respective benefits and risks resulting from test use. Risks and benefits are, at this higher level perspective, to be seen as encompassing both medical and economic connotations and considerations, even though the discussion of benefits and risks is often restricted to the former.

As healthcare payers are becoming increasingly cost-conscious, and reimbursement decisions are being more commonly influenced by medical-economic considerations, clinical utility is quickly becoming the overriding consideration with regard to the introduction of a companion or other diagnostic. This adds significantly to the burden of research and development expenses for diagnostic companies since reliable estimates of clinical utility will usually require prospective, controlled studies in which clinical end-points are reached and where interventions are not guided by testing for the biomarker of interest.

On the level of medical considerations, in the narrowest sense of the term, clinical utility refers to the ability of a screening or diagnostic test to prevent or ameliorate adverse health outcomes such as mortality, morbidity or disability through the adoption of efficacious treatments conditioned on test results. A screening or diagnostic test in isolation does not have inherent utility; because it is the adoption of therapeutic or preventive interventions that influence health outcomes, the clinical utility of a test depends on effective access to appropriate interventions, or to the way it beneficially can affect the choice of an intervention. This use of the term ‘utility’ is consistent with standard practice in evidence-based medicine, which focuses on objective measures of health status to evaluate interventions. Clinical utility can more broadly refer to any use of test results to inform clinical decision-making. Finally, in its broadest sense, the medical interpretation of clinical utility can refer to any outcome considered important to individuals, their families, as well as to other societal strata.

### Table 2: Summary of characteristics for various parameters used in describing clinical test performance

<table>
<thead>
<tr>
<th>METRIC</th>
<th>DEFINITION</th>
<th>STRENGTHS</th>
<th>WEAKNESSES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>(TP+ TN)/N</td>
<td>Intuitive</td>
<td>Depends on prevalence</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>TP/(TP + FN)</td>
<td>Does not depend on prevalence</td>
<td>Applies only to diseased persons</td>
</tr>
<tr>
<td>Specificity</td>
<td>TN/(TN + FP)</td>
<td>Does not depend on prevalence</td>
<td>Applies only to non-diseased persons</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>TP/ (TP + FP)</td>
<td>Clinical relevance</td>
<td>Depends on prevalence</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>TN/ (TN + FN)</td>
<td>Clinical relevance</td>
<td>Depends on prevalence</td>
</tr>
<tr>
<td>Positive likelihood ratio</td>
<td>(TP/(TP + FN))/((FP/ (TN + FP))</td>
<td>Does not depend on prevalence</td>
<td>Applies only to positive tests</td>
</tr>
<tr>
<td>Negative likelihood ratio</td>
<td>(FN/(TP + FN))/((TN/ (TN + FP))</td>
<td>Does not depend on prevalence</td>
<td>Applies only to negative tests</td>
</tr>
<tr>
<td>Odds ratio</td>
<td>TP X TN/FN X FP</td>
<td>Does not depend on prevalence; combines sensitivity and specificity</td>
<td>Values FP and FN errors equally; not intuitive</td>
</tr>
<tr>
<td>Area under curve</td>
<td>Area under ROC curve</td>
<td>Does not depend on prevalence; combines sensitivity and specificity</td>
<td>Lack of clinical interpretation</td>
</tr>
</tbody>
</table>

FN = false-negative; FP = false-positive; N = sample size; ROC = characteristic; TN = true-negative; TP = true-positive.
Clinical applications for protein biomarkers: Examples of clinical utility are discussed below.

**Patient selection.** Using appropriate biomarkers it will be increasingly possible to target the enrolment into clinical trials to substrata of patients within a (conventional classification of) disease. The challenge at the level of clinical development for a new medical entity is that often the validation of a presumed biomarker as a predictor of drug efficacy or safety is still at the level of plausibility, but not of clinical evidence. So more often than not both marker-positive and marker-negative patients will need to be enrolled, and the marker interrogated as one of the parameters according to which the trial will be analysed. Still, there may be cases where the information already known about the biomarker is sufficiently compelling to use it right away for study-subject stratification, such as was the case in the clinical development of trastuzumab (Herceptin®), where only women with Her2-positive malignancies were included.

**Treatment monitoring.** Historically, 80% of the compounds in Phase II clinical trials fail for lack of efficacy. In the last five years, pharmaceutical companies have increased biomarker efforts early in drug discovery programmes to ensure sufficient biomarkers are in place before programmes enter Phase II.

**Adverse events.** Biomarkers may also allow the recognition of individuals that carry a high risk of encountering serious adverse events, and thus avoiding exposure. Given the rare occurrence of serious adverse events, and the need for statistically reliable data, the successful detection of markers for serious adverse events may be particularly difficult.

**Extension of indications.** Compounds that are successfully launched as new medicines addressing a particular disease based on a recognised molecular mechanism may be found effective in additional indications where the information already known about the biomarker is sufficiently compelling to use it right away for study-subject stratification, such as the case in the clinical development of trastuzumab (Herceptin®), where only women with Her2-positive malignancies were included.

**Statistical significance vs magnitude of effect: requirements for clinically applicable biomarkers**

Academic biomarker research tends to be concerned primarily with the statistically determined reliability of a finding. Thus, if the association of a biomarker with a particular phenotype is reproducible, it is interpreted to convey real biological findings that may be of great interest to fundamental understanding of biological mechanisms – even if the magnitude of effect that can be seen under a particular set of experimental constraints is small. Likewise, the Geoffrey Rose paradox explains that true, reproducible observations may be of great importance from the perspective of public health even if the overall magnitude of effect is modest. Thus, targeted modulation of variables that affect disease risk by small amounts may have important effects if applied to large populations, and are thus commonly used to guide health policy. Academic epidemiological research evaluating associations of predictors for outcomes often uses odds ratios or relative risk as its primary parameters for reporting results, particularly in the field of genetic association studies. In complex polygenic diseases, the magnitude of effects commonly found for associations for individual genetic variants generally ranges between ORs of 1-2. Demonstrating statistical significance for such modest-sized effects usually requires fairly large studies, particularly if multiple comparisons are made (as is the case with genome-wide association studies).

On the other hand, when tests are to be applied to clinical decision-making in individual patients, information content with regard to the magnitude of association becomes critically important. Only reliable tests – with an acceptably low rate of false positive and/or false negative results – can responsibly be used in this setting. These performance parameters tend to be most directly gleaned from stating sensitivity and specificity (or positive predictive value [PPV] and negative predictive value [NPV]), and these are therefore the parameters commonly referred to in clinical diagnostics. In general, tests are not considered as particularly useful if the area under the ROC does not exceed 0.8, which translates into balanced sensitivities and specificities of about 0.75. On the other hand, a balanced sensitivity and specificity of 0.75, translates into an odds ratio of about 9; if an effect of this magnitude is indeed present, it will be readily recognisable even in quite modestly-sized studies.

**Companion diagnostics: co-development of diagnostic test and therapeutic moiety**

The definition of a companion diagnostic is a new diagnostic test developed in combination with a new therapeutic where the diagnostic test is essential to select those patients that should receive the drug. So, if the label of a new medicine is to include the obligatory use of a test, then the manufacturer needs to ensure that drug and test are developed in tandem, and receive regulatory approval in such a fashion that they can reach the market simultaneously.
Authorities are paying increasing attention to the challenges that this process may present. The US Food and Drug Administration has issued a concept paper on ‘Drug-Diagnostic Co-Development’ (www.fda.gov/Cder/genomics/pharmacoconceptfn.pdf) that specifically addresses this issue. A number of working groups, including the International Conference on Harmonisation are also currently addressing aspects related to this issue, from coming unified terminology and language around biomarkers and companion diagnostics, to formal issues relating to regulatory submissions and the evidentiary standards that will be required to accept a companion diagnostic as valid and useful (Figure 1).

Conclusion: information content counts – proteins as the richest source

The challenges for biomarker science are clear and rather daunting. Those markers that ultimately will provide the kind of information content that will qualify them in the clinical arena as diagnostic or prognostic tools will need to satisfy high standards both with regard to sensitivity/specificity and prediction of health outcomes on a prospective timeframe. Importantly, these requirements will have to be judged and adjusted in a context-specific way – different clinical applications will require different qualities of biomarkers. In considering the attributes that will render a biomarker truly useful, it may be argued that protein markers are likely to dominate the field – with the associated cost and complexity of assaying these most demanding markers.

Successfully realising the promise of biomarker-driven personalised healthcare will require a highly collaborative and integrated, multidisciplinary and multifaceted coalition of basic scientists, clinical investigators, the pharmaceutical and diagnostic/device industry, healthcare professionals and payers. Last, but by no means least, patients will have to play a seminal role in this effort as members of the coalition, since improving their lives is where, as always, ultimate proof for the value of biomedical innovation rests.

Dr M. Walid Qoronfleh held appointments with increasing responsibilities at SmithKline Beecham (now GSK), Sterling-Winthrop (now Sanofi-Aventis), the National Cancer Institute, AntexPharma and ThermoFisher. Dr Qoronfleh is the founder of three biotechnology companies and is the Founder and Managing Director of the boutique consulting company Q3CG. He obtained his PhD from The University of Louisville Medical School and received his MBA from Penn State University.

Dr Klaus Lindpaintner has been Vice-President of Research & Development and Chief Scientific Officer of Strategic Diagnostics Inc since February 2010. Prior to joining SDIX, Dr Lindpaintner was with F. Hoffmann-La Roche Ltd for 13 years. He most recently served as Director of the ‘Roche Molecular Medicine Laboratories’ at the company’s headquarters in Basel, Switzerland, and as Roche’s Global Head, Molecular Medicines Policy and External Affairs, co-ordinating the company’s efforts and activities in implementing biomarker research based on genetics, genomics, proteomics, and associated disciplines from early discovery to late-stage clinical trials. Dr Lindpaintner graduated from the Innsbruck University Medical School with a degree in Medicine and from Harvard University with a degree in Public Health.

References

Dr M. Walid Qoronfleh is currently the Executive Director of Life Science business at SDIX, serving the diagnostics (IVD) and immunoassay markets where he has general management accountability.

Dr Qoronfleh has more than 20 years’ of scientific and business experience. Most recently, he served as the Vice-President of Business Development at NextGen Sciences, a commercially-driven biotech-