Understanding the strategic importance of BIOMARKERS for the discovery and early development phases

Biomarkers have become an increasingly hot topic in the pharmaceutical industry. How much of this is hype, how much is reality? In the overwhelming majority of cases, biomarker studies can only provide preliminary data. Only about a half dozen biomarkers are considered true surrogates, sufficiently validated to serve as primary endpoints in registration trials. Why bother with biomarkers, if clinical endpoints will be required anyway? Why not just cut to the chase and do a trial recognised by the registration authorities?

The strategic imperative for biomarkers rests largely on the need to bridge that gap between increasingly high-throughput Discovery organisations and the increasingly time- and cost-intensive nature of drug Development. The availability of drug candidates is no longer rate-limiting. There are more than plenty to choose from. Unfortunately, this increased Discovery output has also been associated with increased failure rates in Early Development. 'Plan for success' may be a heartening motto, but is increasingly divorced from the reality of Early Development, where the majority of candidates fail. We must learn to focus development resources on those compounds that are most likely to achieve registration as marketable products. How can we best manage the risk that a compound directed against a novel target, not yet been pharmacologically proven in man, will fail in
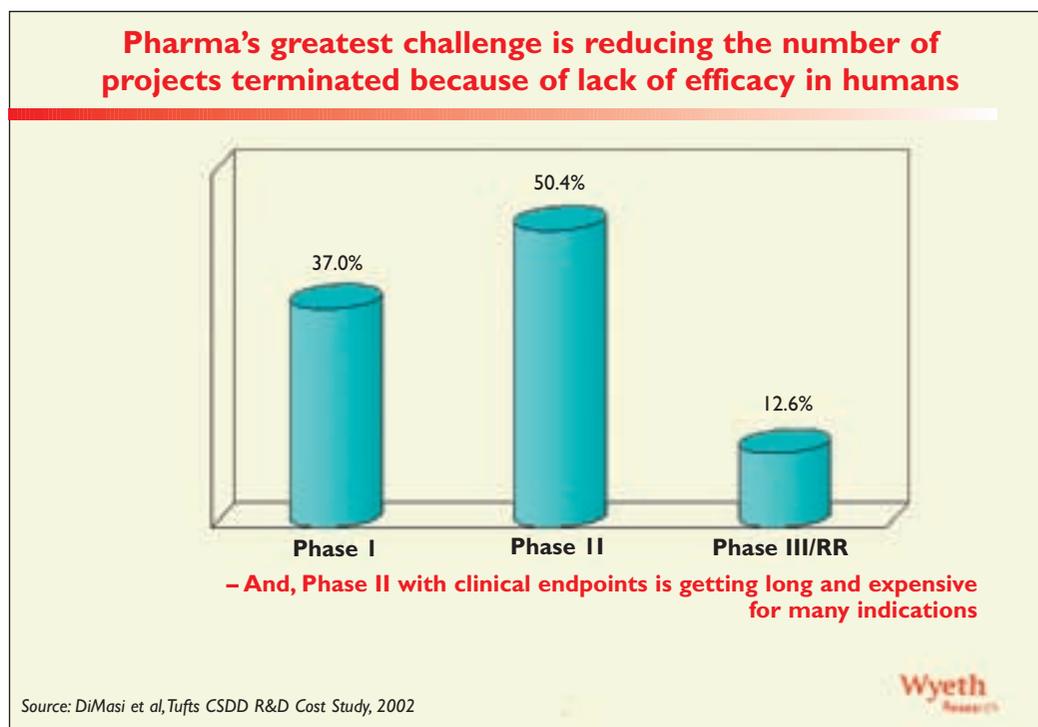
registration trials? In some circumstances, biomarkers may be more efficient than clinical measures in separating the wheat from the chaff.

Let's examine this more closely. The litany of the technical advances that have led to increased discovery output is well known. Complete genome sequences have increased the number of potential drug targets by almost two orders of magnitude. A decade ago, most of the encoded proteins did not even have a name, much less a supporting literature. Little is known about the biology of most of these new targets, much less their role in human disease. Combinatorial chemistry, high throughput screening, monoclonal antibody and other technologies virtually assure a drug candidate for most of these novel targets. Although the path from target to candidate is long, hard and expensive, it has become increasingly predictable and formulaic. We have

By Dr Orest Hurko

Biomarkers

Figure 1

Pharma's greatest challenge is reducing the number of projects terminated because of lack of efficacy in humans

50.4%

37.0%

12.6%

Phase I   **Phase II**   **Phase III/RR**

– And, Phase II with clinical endpoints is getting long and expensive for many indications

Source: DiMasi et al, Tufts CSDD R&D Cost Study, 2002

Wyeth

become very efficient in finding compounds that are safe and effective in laboratory animals, however, for novel targets, we still are not very efficient in identifying drugs that work in people (**Figure 1**). Most do not. The pharmaceutical industry has honed a cost-effective strategy for developing drugs that offer incremental improvements to existing remedies already proven in the clinic. A different strategy may be optimal for unproven targets. Biomarkers may be useful for such a novel strategy.

Can't this mismatch be remedied most simply by further polishing in Discovery? We all know that there is a limited predictive power of animal *in vitro* studies. Understandably there is greater predictive power of animal models for established targets than for novel ones. Using known clinical results as a gold standard, one can then back track to the animal models to quantify their predictiveness for future 'me too' drugs. What to do for targets that do not have such a gold standard? Multiple surveys confirm the logical *a priori* expectation that compounds directed against novel targets fail more frequently than do those against targets conjectured, but not yet proven, to be effective in human disease. Novel targets bring new opportunities, but also new risks. I refer to the risk that a compound may be effective in preclinical models, but may nevertheless fail in later registration trials.

Some of this risk reduction must be done in human subjects. Animal models may share some

similarities with a human disease, but they are never identical. Such models may share an aspect of the pathophysiology, but rarely, if ever, the disease in its entirety. Furthermore, our science is not yet sufficiently robust to allow prediction of behaviour in intact patients from observations with isolated human cells or macromolecules. Finally, we must not forget that significant metabolic pathways in laboratory species may be minor or redundant in humans. Biomarkers, though not identified as such, have always been an integral part of drug discovery. The consequence of these differences is that only so much risk reduction can be accomplished pre-clinically. In certain circumstances, biomarkers can be useful for risk reduction in human subjects. Drug discovery biologists have always used direct biological measures to guide the development of compounds. Clinicians conducting registration trials have had to rely on very indirect, inherently noisy clinical measures: Does the patient function or feel better or live longer? How efficient would discovery scientists be if these were the only measures at their disposal? In the words of Tachi Yamada, what we need is a "Discovery Laboratory in Humans." Only targets that have proven themselves in this laboratory or in the clinic warrant the massive investment required for traditional registration studies.

Why do compounds effective in laboratory animals fail in the clinic? (**Figure 2**) One can consider

five major categories, the first four of which may be mitigated in some circumstances by judicious use of biomarkers:

1) The drug is given to the wrong subjects, an admixture of responders and non-responders.
2) The drug is not given at the right dose, either failing to achieve adequate receptor occupancy or exceeding concentrations that saturate the receptor, and thus unnecessarily increasing the chances for off-target effects.
3) The indirect efficacy signals provided by clinical measures are either too noisy or too late to provide efficient initial testing of unproven targets.
4) Some patients get sick from drug exposure.
5) The drug only works in animals, not in humans.

For certain compounds, biomarkers can be used to used to reduce failures from the first four causes, and for ill-fated compounds at least provide a cost-effective, timely and definitive confirmation of the fifth. Precious development resources can thus be devoted to only those drugs that have been shown to work in humans.

Breaking down the barriers between Discovery and Development
Novel biomarkers suitable for use in humans must begin Development while the candidates are still in the Discovery phase, or from previous clinical studies with antecedent compounds. Some direct biological readouts used in laboratory animals are not available in the clinic, for both ethical and practical reasons. By the time a particular compound has advanced into the clinic, it is too late to come up with a novel biomarker. One can then only rely on biomarkers that have already been developed in previous studies. This recommends a rethinking of the prevalent practice of isolated planning and funding of individual projects devoted to the development of a single compound, as well as the relationships between Discovery and Development. There are both organisational and cultural differences to be considered. There are separate reporting structures, separate deliverables, separate budgets. Why should Discovery pay for biomarkers that are only going to be in Development? There are also methodological differences. For the most part, Discovery works in small iterative experiments, trial and error, each experiment informing design of the next. In contrast, the development organisation needs to rely on a few large trials, meticulously planned in advance, with little opportunity for midcourse corrections, if any. Discovery picks optimal experimental conditions with inbred animals, which are treated according to precise temporal protocols, housed and fed
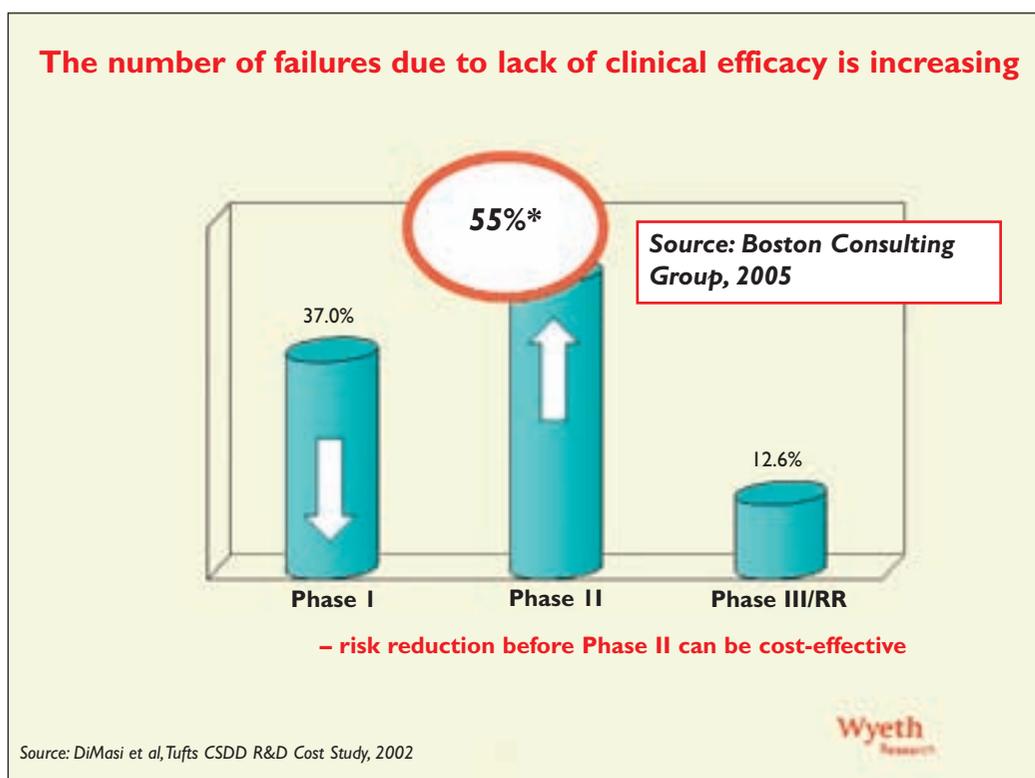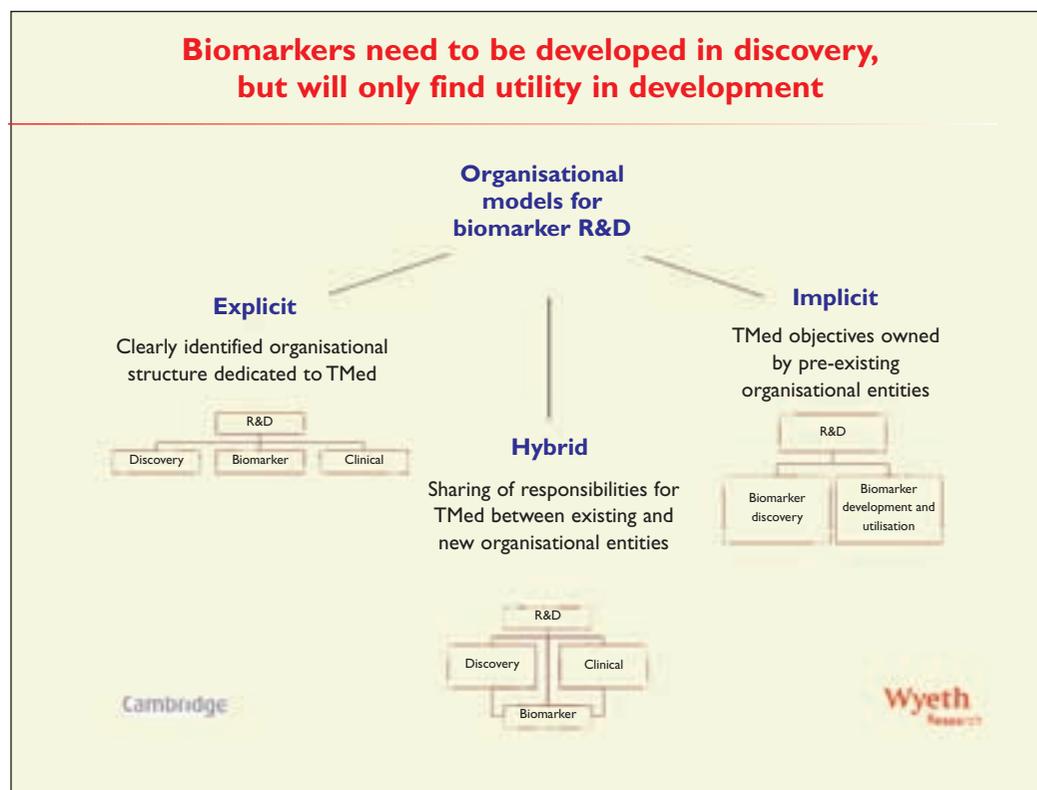
The number of failures due to lack of clinical efficacy is increasing

55%*

Source: Boston Consulting Group, 2005

37.0%

12.6%

Phase I   **Phase II**   **Phase III/RR**

– risk reduction before Phase II can be cost-effective

Source: DiMasi et al, Tufts CSDD R&D Cost Study, 2002

Wyeth

Biomarkers

Figure 3

Biomarkers need to be developed in discovery, but will only find utility in development

under identical conditions. In contrast, Development seeks the broadest possible label by testing compounds on a wide variety of ages, with different concurrent medications and lifestyles. Biomarkers used in the context of a 'learn and confirm' model[1] may provide a suitable alternative for early development of compounds directed against high-risk novel targets.

Most large pharmaceutical companies and also the regulatory agencies have accepted the need for discovery and development of biomarkers. However, there is no widespread agreement on the optimal organisational models for biomarker research and development. Many companies have formed Translational Medicine, Discovery Medicine, Experimental Medicine, Investigational Medicine or other similarly named units. Some have chosen an explicit model with clearly identified organisational structure dedicated to the discovery and development of biomarkers. In these companies, R&D has added a third, separate biomarker unit to stand alongside the two traditional Discovery and Clinical Development organisations. Unfortunately, this replaces one interface with two. Others have adopted an implicit model, with biomarker objectives owned by pre-existing organisational entities, without an explicit biomarker unit. A third is a hybrid model that allows

for sharing of responsibility between a biomarker discovery unit in Discovery closely partnered with a biomarker laboratory in Development (**Figure 3**).

Specific biomarkers for specific purposes

There are several things one can do to bridge the gap between Discovery and Development. Biomarkers can fashion early Development trials along the lines of Discovery experiments. That is to say, they can set up initial trials in such a way as to optimise the chance of demonstrating some potentially useful effect in human subjects by identifying the subjects that are most likely to respond; ensuring that receptor occupancy by the drug is optimal; and measuring the most sensitive efficacy biomarker. If there is no efficacy seen under these optimal circumstances, there has to be the institutional resolve to walk away from the compound and redistribute resources to others. If, however, the compound shows some promise under these optimal – but admittedly very artificial – circumstances, one can test in a broader, commercially viable population with standardised dosing and clinical endpoints. That is what Discovery has been doing all along. Only in laboratory animals.

For this strategy to be successful, it is important to utilise specific biomarkers to make specific

Biomarkers

decisions. There are very few biomarkers, if any, suitable for all purposes. Referring back to the four remediable causes of clinical failures[2], one can envision four potential uses for biomarkers

1) Biomarkers that permit the identification of optimally responsive subjects for initial human testing, before making attempts to broaden the label. This is particularly important for high-risk targets that have never been pharmacologically proven in human beings.

2) Biomarkers that directly quantify drug-target interactions, guiding choice of initial dosing regimens. These will ensure that the dose is adequate but not excessive; particularly for early development before the drug-response relationship for clinical effects is established.

3) Efficacy biomarkers that provide a cost-effective method for the demonstration of the relevance of an unprecedented drug target to the pathophysiology of the diseases(s) for which a label is sought, before financial commitment to traditional large registration studies with clinical endpoints.

4) Toxicity biomarkers that would ideally permit exclusion of subjects likely to fall ill after drug exposure, or, at the very least identify, after dosing, an unwanted side reaction in its earliest stages, permitting withdrawal of subjects before they became clinically ill.

We will discuss each of these biomarker applications in turn

Identification of optimal subjects
For initial early development studies, it may be best to begin testing in subjects selected with much narrower inclusion criteria than would be practicable for a commercially viable label. A commercially viable label is always the ultimate goal. However, before the efficacy of a target or compound is established in humans, there is value in initial trials that give the compound the best possible opportunity to demonstrate efficacy. "If it's going to work in anybody, it ought to work in these subjects." The time-delay and extra expense required for the insertion of such an intermediate translational study before a traditional Phase II trial may be cost effective in certain circumstances. The financial trade-off of delay to market with risk reduction should be estimated explicitly using standard risk assumptions and forecasting techniques.

Biomarkers may identify patients most likely to respond to the drug. More than half a century ago, gram staining of sputum was recognised as a useful biomarker for penicillin trials by distinguishing pneumonia subjects with responsive pneumococcal disease from those non-responsive pneumonia patients with gram-negative, viral or tuberculous pneumonia. Using such a patient

Figure 4

We have used biomarkers to understand patient response in our Alzheimer's vaccine trial

● The Aβ 1-42 aa peptide (AN1792) is associated with plaques in the brains of AD patients

● Immunisation with Aβ 1-42 has shown efficacy in animal models of AD

● 48% of AD patients mounted an antibody response to immunisation with AN1792

● Only subjects that mount an antibody response would be expected to benefit from treatment
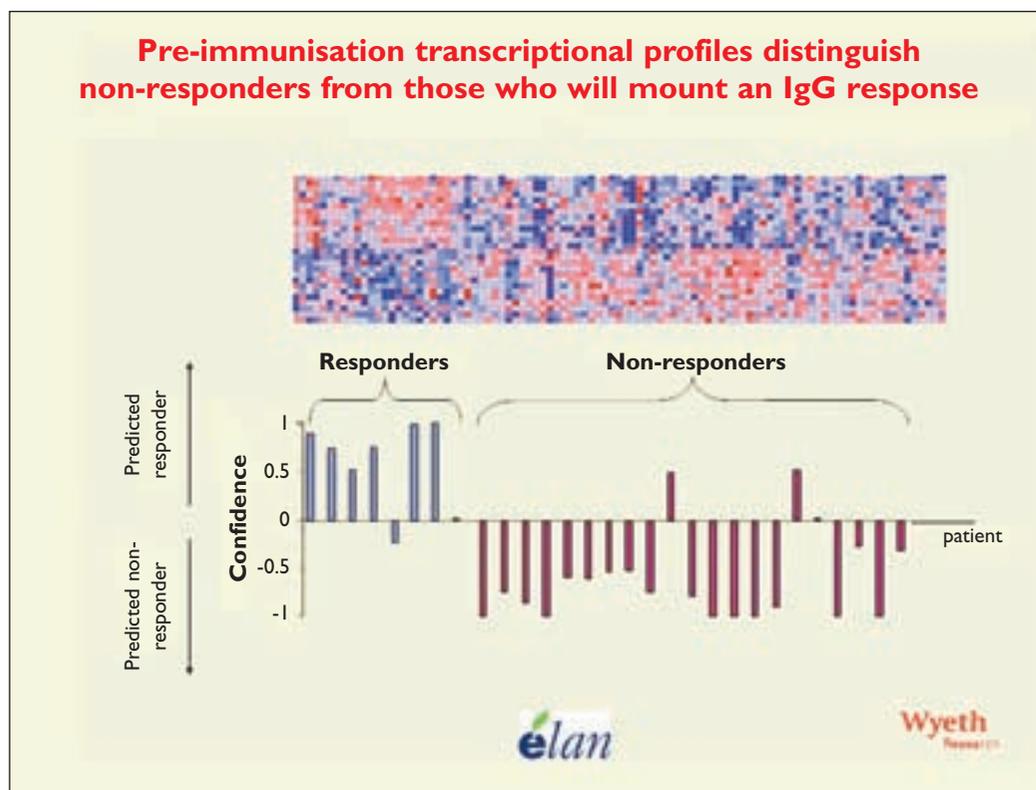
● **What biomarkers exist to predict antibody response?**

élan   Wyeth Research

Pre-immunisation transcriptional profiles distinguish non-responders from those who will mount an IgG response

selection biomarker, one only enrolls subjects that are probable responders. A smaller number of subjects need to be enrolled to obtain a statistically significant efficacy signal. One also reduces the ethical concern that some subjects will be exposed without any reasonable expectation of benefit.

A current example is given by molecular biomarkers that distinguish otherwise histologically identical tumours. For instance, an EGFR mutation biomarker that allows identification of those lung cancer patients that are exquisitely responsive to treatment with anti-EGFR (Gefitinib, Iressa)[3]. In about 10% of non-small cell lung cancer there are mutations in the EGFR receptor gene, which encodes a protein that drives lung cancer growth. Iressa preferentially kills cancer cells bearing such mutations. Patients identified by this biomarker were predicted *a priori* to be likely responders. Experience has proven this correct. As it happens, fears of market fragmentation were unfounded. Other, much larger studies demonstrated that some patients without the mutation occasionally respond to Iressa, but with a much lower frequency. So, after the first initial demonstration of efficacy in a very tightly defined patient group, it may be possible to explore a broader label. One could also argue that fragmentation was further offset by readier acceptance of this novel compound in a

crowded field, and by ethical as well as other considerations. Although plausible, these latter assertions are harder to quantify.

An experimental patient selection biomarker has emerged during our Wyeth-Elan collaboration on active immunisation against an amyloid fragment for treatment of Alzheimer disease[4]. The aβ1-42 amino acid peptide is associated with plaque in the brains of Alzheimer patients and has been hypothesised to be pathogenic. Immunisation with aβ peptide has shown efficacy in transgenic animal models of Alzheimer disease. However, Alzheimer patients are elderly. With old age comes a decreased ability to mount an antibody response. Indeed, in our trial, only 48% of Alzheimer patients mounted an antibody response to immunisation with aβ peptide (**Figure 4**). Only subjects that mounted an antibody response would have been expected to benefit from immunisation. We asked: "What biomarkers could identify those patients that are capable of mounting an antibody response?" Margot O'Toole, Ron Black, Andy Dorner and their colleagues collected baseline blood samples from subjects before they had been immunised. In a post-hoc analysis, they identified a pre-immunisation transcriptional profile in peripheral blood monocytes that distinguished those individuals who subsequently mounted an antibody

Biomarkers

response, from those that did not (**Figure 5**). Should this finding prove robust, we may have a way forward. We could enroll only predicted antibody-responders, excluding likely non-responders from subsequent trials.

Quantification of drug-target interaction

Although much attention has been focused on biomarkers for patient selection or for efficacy, less appreciated are biomarkers used to guide dose selection. And yet it is this class of biomarkers that has already quietly established itself as a proven workhorse in early Development. Discovery scientists often underestimate the challenge of identifying appropriate doses in initial clinical studies. However, it is well documented that many clinical failures have resulted from the inability of the drug to reach the target in sufficient quantity or duration[5]. Measurement of blood levels provides little assurance of adequate receptor occupancy, particularly for those compounds destined for targets in the central nervous system or in poorly vascularised tumours. Although there are multiple indirect estimates of receptor occupancy provided by pharmacodynamic biomarkers, the most direct estimate is often provided by ligand-displacement PET or SPECT. These techniques are well established, albeit tedious and expensive. The rate-limiting step to such displacement studies is the paucity of suitable ligands – only several dozen – available. It is almost as difficult to develop a novel ligand, as it is to develop a drug. For novel targets, ligand development must begin at risk while the lead series is still in early Discovery phase. Discovery scientists have cheaper and more direct ways of quantifying drug-target interactions. And yet it is they who must first develop the ligands in laboratory animals for use by their Development colleagues in human subjects, where there is often no suitable alternative. By the time a compound enters the Development phase, it is too late to start thinking about a novel ligand.

Early and sensitive indicators of efficacy

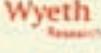The third general class of biomarkers consists of those that indicate efficacy – demonstration that the

drug is favourably altering the pathophysiology of the disease. Clinical proof of efficacy may take many years for chronic diseases. For example demonstration of disease-modification in chronic neurodegenerative disorders – rather than simple amelioration of symptoms – may take upwards of 10 years if one were to use clinical criteria alone. One wonders how ethical it is to test subjects for this long with non-efficacious or non-proven drugs. Furthermore, it is not financially possible for most organisations to run decade-long registration trials with each of their promising candidates. Efficacy biomarkers that offer a quick read on the ability of novel drugs to alter the pathophysiology of the disorder could provide a useful intermediate step in Early Development. Only those compounds that looked promising in exploratory human trials would receive the investment required for registration studies with primary clinical endpoints.

For compounds designed to remove amyloid from symptomatic or presymptomatic Alzheimer patients, PET imaging of amyloid deposits in the brain affords a useful biomarker approach. There are available a number of PET ligands that quantify the amyloid deposits in the brain, using a minimally invasive, ethically acceptable procedure. In a translational study of several months duration, amyloid deposits would be quantified at enrollment, and again after several months of treatment.

If a putative amyloid-lowering agent did not lower amyloid, it would not be expected to alter disease progression in a definitive clinical trial of much longer duration. Although it seems reasonable that amyloid lowering should precede clinical improvement, it is important to remember that the amyloid hypothesis is as yet unproven, and that the relative role of fibrillar, monomeric or oligomeric species is not well established. For this reason, PET demonstration of amyloid lowering is certainly not sufficient to guarantee registration of the drug, nor should it be. This is still somewhat speculative. Nevertheless, if, at a given dose, a drug does not lower amyloid, a company may well decide to shift its resources to another compound. Decisions have to be made on the basis of the best available evidence, even if not definitive. Indeed, I would argue that the most cogent argument for the use of efficacy biomarkers in lieu of registrable clinical endpoints can be made for drugs designed to alter the natural history of chronic diseases.

Presymptomatic detection of toxicity
The fourth question that could be addressed by biomarkers is whether s drug is likely to be harmful and, if so, to whom. This question may be ultimately addressed with toxicity biomarkers. Another specific example is again provided by our Wyeth-Elan collaboration of active immunisation

with aβ, for the treatment of Alzheimer disease[4]. Unexpectedly, our Phase IIa trial of AN1792 demonstrated encephalitis in 6% of immunised subjects. In this trial, there were 18 cases of encephalitis, all in the AN1792 treated group (**Figure 6**). The onset typically occurred after the second injection. In 16 of the 18 cases, the clinical presentation was confusion, headache, lethargy, mononuclear pleocytosis. Twelve of these individuals returned to baseline or near baseline status within weeks. However, this was a serious enough side-effect that it halted ongoing clinical research with this specific immunogen.

In the same set of experiments described above, Margot O'Toole and her colleagues discovered a transcriptional profile in the baseline pre-immunisation blood samples that may distinguish those patients who went on to develop encephalitis from those that did not. If we are able to verify these results, we could consider using this profile to eliminate subjects at risk for developing encephalitis from subsequent clinical trials.

Biomarkers – their role for the future

There is an emerging consensus that biomarkers have utility in drug development. However, it is important not only to understand the current uses of biomarkers as well as their future prospects, but equally important to dissect the hyperbole promulgated by some enthusiasts. In considering the future, it is important to remember that biomarkers can only approach the gold standard of clinical measures, which, by definition, they can only approximate. Drugs will only be registered if they allow patients to live better or longer lives, not because of the alternation of a molecular or physiological parameter. Our science is not sufficiently robust to predict desired clinical outcomes from first principles. Empiric observation is required. Furthermore, in this era, it is the validation of a biomarker against the gold standard of clinical measures – not the technology – that is most rate limiting.

Having said that, I hasten to emphasise that several biomarkers that have not succeeded in getting the high level of validation required for true surrogacy, have already abundantly demonstrated their utility for internal decision-making and for mechanistic studies. This is true even though they are not yet substitutes for registrable endpoints – and, in most cases, probably never will be. Non-surrogate biomarker studies are thus a prelude, not a substitute, for clinical studies. Because of this, biomarker studies are only cost-effective if they are quicker and cheaper than standard clinical studies (or if

they inform mechanistic questions for future Discovery efforts). The cost, including time, must be quantified and balanced against expected risk reduction. Because of this, biomarkers may not be cost-effective for some acute disorders. For many acute indications, it may be cheaper and faster just to do a clinical study. However, even in selected cases, there may be utility of biomarkers for patient- and/or dose-selection, as well as for mechanistic studies

With our consultants from Cambridge Pharma, we conducted a survey of our competitors to assess their current views of biomarkers. We found that new technologies for biomarker discovery have been adopted widely. We surveyed our competitors on their use of biochemical assays, expression profiling, proteomics, proteomics, non-invasive imaging, as well as pharmacogenomics. Our consultants asked them to assess not only their present utility, but also their expected utilities in five years' time. Biochemical assays and imaging are believed by most companies to have the greatest current utility. Some technologies were less favoured than others. Most companies, when surveyed last year, thought that metabonomics was mostly 'a lot of noise'. In five years most agreed that non-invasive imaging and expression profiling will have increasing importance. There is widespread agreement that maximum utility of these technologies has not been realised. The major issue to be resolved is in the statistical analysis, both of 'Omic' technologies and of imaging data. The problems with analysing chips with 20,000 variables surveying only 100 patients are, at present, not fully developed.

A word has to be said about surrogate markers, which are obviously the ultimate goal. Most biomarkers are for internal decision-making only prior to registration trials with clinical endpoints. However, the rare achievement of surrogate status enormously simplifies drug registration programmes. For example, both blood pressure and serum cholesterol levels[6] have been accepted as surrogate efficacy biomarkers for the prevention of stroke and of myocardial infarction. Using these surrogate markers, it is possible to register drugs once it is demonstrated the lowering of blood pressure or cholesterol without the necessity of going on to decade-long trials directly demonstrating reduction of strokes or myocardial infarctions. However, validation to the level of surrogacy – use as primary endpoints in registration studies – requires correlation of clinical outcomes with multiple converging drug mechanisms as well as demonstration of an established mecha-

References

1 Sheiner, L (1997). Learning vs confirming in clinical drug development. Clin Pharm Ther. 61: 275-291.

2 Dimasi, JA (2002). The value of improving the productivity of the drug development process: faster times and better decisions. Pharmacoeconomics.; 20 Suppl 3:1-10.

3 Paez, JG, Janne, PA, Lee, JC, Tracy, S, Greulich, H, Gabriel, S, Herman, Kaye, FJ, Lindeman, N, Boggon, TJ, Naoki, K, Sasaki, H, Fujii, Y, Eck, MJ, Sellers, WR, Johnson, BE, Meyerson, M (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science. 304:1497- 1500.

4 O'Toole, M, Janszen, DB, Slonim, DK, Reddy, PS, Ellis, DK, Legault, HM, Hill, AA, Whitley, MZ, Mounts, WM, Zuberek, K, Immermann, FW, Black, RS, Dorner, AJ. (2005). Risk factors associated with beta-amyloid (1-42) immunotherapy in preimmunization gene expression patterns of blood cells. Arch Neurol.; 62: 1531-1536.

5 Frank, R, Hargreaves, R (2003). Clinical biomarkers in drug discovery and development. Nat Rev Drug Discov. 2: 566-80.

6 Tobert, JA (2003). Case history: Lovastatin and beyond: the history of the HMG-CoA reductase inhibitors. Nat Rev Drug Discov. 2: 517-526.

7 http://adni-info.org/index.php?option=com_content&task=blogcategory &id=0&Itemid =43.

Biomarkers

nistic link. A major undertaking that exceeds the resources of any single institution. Correlation does not prove causality.

Understandably, only a small handful of surrogates exist currently. The level of validation required demands a decade(s)-long undertaking, requiring resources of multiple partners. Pre-competitive partnerships to develop such partnerships are already in place. A good example is provided by the ongoing Alzheimer's disease imaging initiative (ADNI) sponsored by the Foundation for the National Institutes of Health, in collaboration with more than 40 academic and industrial partners[7]. The primary goal of this project is identification of biomarkers of disease progression for use as endpoints in clinical trials for the prevention and treatment of Alzheimer's disease. Under investigation as possible biomarkers are volumetric MRI, resting glucose PET as well as a neuropsychological battery and the analysis multiple biological specimens. However, even this massive effort does not guarantee that a surrogate will emerge.

It is likely that the handful of surrogates that will emerge will be from consortia such as this. Furthermore, it is likely that the majority of biomarkers will not achieve the extensive validation required for surrogacy. They will be useful, nevertheless. Biomarkers are already beginning to improve the efficiency of Early Development, much as they have been a cornerstone of drug Discovery efforts for over a century. **DDW**

*After graduation from Harvard College, graduate studies at the Massachusetts Institute of Technology, **Dr Orest Hurko** was awarded an MD from the Harvard Medical School. After a fellowship in the laboratory of Biochemical Genetics at the NIH and a year at the National Hospital (Queen Square) as the William O. Mosely*

Travelling Fellow of Harvard, Dr Hurko joined the faculty of the Johns Hopkins University School of Medicine in Neurology, with joint appointments in Medicine, Pediatrics and Neurological Surgery. He directed the Neurological and Neurosurgical Consultation Service, and attended at the Moore Genetics Clinic. His research focused on the molecular and clinical aspects of heritable neurological and skeletal disorders. He joined Wyeth in January 2003 where he is now Assistant Vice-President, Translational Medicine. Prior to that he was Head of Investigational Medicine in the Neurology CEDD at GSK, in Harlow, UK.