What imaging are biomarkers and how are they used?

John C Waterton[1,2]
[1]AstraZeneca, Alderley Park, MACCLESFIELD, Cheshire SK10 4TG UK
[1,2]University of Manchester, MANCHESTER M13 9PT UK
john.waterton@astrazeneca.com
john.waterton@manchester.ac.uk

Imaging measurements have been made in biomedical research and in medical practice for over a century: until recently, however, they were rarely described as biomarkers. During the last ten years a new biomarker vocabulary of “validation”, “qualification” and “evaluation” has emerged. This vocabulary incorporates imaging measures naturally alongside other biomedical measurements. This presentation will introduce four different dimensions in biomarker thinking, with examples of how each applies to imaging.

Following Atkinson et al, 2001(1) a biomarker is defined as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention". In contrast, a clinical endpoint is "a characteristic or variable that reflects how a patient feels, functions, or survives". A surrogate endpoint is a special class of biomarker that regulatory authorities accept can validly substitute for a clinical endpoint in drug approvals.

Dimension one: biomarker strength of inference

Few imaging (or other) biomarkers are surrogate endpoints: examples are time-to-progression in oncology (essentially an imaging biomarker), which has been used to support the approval of certain anti-cancer therapies; or DEXA, where absence of change in bone mineral density following drug therapy predicts lack of harm, i.e. no increased risk of osteoporotic fracture. The key insight, however, from the new biomarker vocabulary, is that biomarkers which are not validated may still be useful. In this case, we evaluate, not validate, the biomarker. When our confidence in the biomarker is sufficient for a particular purpose, we can regard it as qualified to support a particular decision. Biomarker qualification should be fit-for-purpose, and a biomarker that is qualified for one purpose may be insufficiently qualified for a different purpose. Although the term qualification was originally introduced in a regulatory context, it is useful to broaden the concept to describe the process of gaining scientific understanding of the biomarker performance to support any decision taken on biomarker data. Evaluation is essentially risk-management: the connection between the biomarker and underlying biology is assessed in animal and human studies (e.g. by imaging-pathology correlation); reproducibility, and effect size assessed, and correlations with clinical endpoints established, to understand the risks of false-positive or false-negative biomarker signals.

Dimension two: biomarker modality

There are four principal approaches to providing biomarkers in medicine. These are biofluids, tissue biomarkers, physiologic biomarkers, and imaging biomarkers.

- Biofluid biomarkers are obtained from fluids such as blood, urine, or saliva. They have the advantage that they are relatively non-invasive and can be measured repeatedly using robust protocols, and are inexpensive. It is from biofluid biomarkers that the vast armamentarium of genetic, genomic and proteomic biomarkers so familiar in medical modern research has been developed. A common disadvantage of soluble biomarkers is that they integrate information from the entire body and therefore can be poor at identifying biochemical abnormalities in tiny disease foci.
- Tissue biomarkers are obtained by removing samples of solid tissue from the body,

for example at biopsy. Sensitive and specific biomarkers can be obtained via histopathology and histochemistry. In contrast to biofluid biomarkers, tissue biomarkers are generally invasive, difficult to measure repeatedly, and expensive, but are highly suitable for focal disease assessment.

- Physiologic and biophysical biomarkers use a range of clinical tests such as blood pressure, lung function, or electrocardiography. Like biofluid biomarkers, they tend to be non-invasive, easy to measure repeatedly, and inexpensive.
- Imaging biomarkers have a unique profile in that they are relatively non-invasive, easy to repeat, and of course highly suitable for the study of focal disease. Their principal weakness is that they are often expensive to measure and it can be difficult to devise and disseminate robust and reproducible imaging protocols.

Dimension three: biomarker inference

We can distinguish four different kinds of biomarker, based on the inference drawn from the biomarker measurement. In all cases biomarkers are only useful if they allow us to predict a patient's future clinical status before that is clinically evident.

- A prognostic biomarker measured at a single baseline time point, in combination with clinical data, better predicts future outcome than does the clinical assessment alone. In this context, diagnostic and screening biomarkers can be included with prognostic biomarkers, since diagnosis or screening have little value without some prognostic content. From this perspective, most of diagnostic radiology exists to provide prognostic biomarkers. Prognostic imaging biomarkers from screening include measurements made from foetal ultrasound, mammographic evidence of tumour, and chest radiography for tuberculosis.
- A predictive biomarker measured at a single baseline time point in combination with clinical data permits a better choice of therapy, than the clinical data alone. Predictive biomarkers are sometimes called companion diagnostics, and the use of a specific biomarker to support a specific choice of therapy is sometimes called personalised healthcare. Examples of regulator-approved predictive imaging biomarkers include CT exclusion of haemorrhage to predict lack of harm from alteplase (Activase), and ^{131}I-tositumomab imaging to predict efficacy of tositumomab (Bexxar).
- A monitoring biomarker is one that is monitored after a treatment is given. Better decisions about treatment adjustment, alteration, or withdrawal can be made monitoring biomarker together with clinical data, than can be made with clinical data alone. Examples of monitoring biomarkers include whole-body FDG PET to monitor for relapse after successful cancer therapy, or cardiac ultrasound to monitor for potential harm caused by cardiotoxicity of certain anti-cancer agents.
- A response biomarker is one where the change in biomarker value following a therapeutic intervention, together with clinical data, predicts outcome better than the clinical data alone. Response biomarkers are mainly used in population studies in clinical trials in medical research, in particular in drug development. Examples of response biomarkers include multiple sclerosis enhancing plaque count, K^{trans} in dynamic contrast enhanced MRI in tumours, cartilage thickness in osteoarthritis, carotid intima media thickness, or tumour volume.

Dimensional four: biomarkers of benefit, harm, and the pharmacological audit trail.

A biomarker may be useful because it predicts efficacy, or lack of efficacy, or harm or lack of harm. For response biomarkers of drug efficacy, imaging (and other) biomarkers can be useful to build the link between molecular pharmacology and disease modification:

- Does the drug reach its target intact (drug metabolism and pharmacokinetics)? Example: ^{19}F MRS of 5-fluorouracil in metastatic colorectal cancer

- Does the drug modulate its molecular target? Example: ^{13}C MRS of glycolytic flux in diabetes
- Does the drug elicit a pharmacologic response in its target tissue (sometimes called pharmacodynamic biomarkers)? Example: K^{trans} in solid tumours.
- Does the drug ameliorate or reverse the structural changes in disease? Example: cartilage thickness in osteoarthritis.

If a drug fails to show clinical efficacy, this "pharmacological audit trail" is useful to identify the cause of hypothesis failure (e.g. wrong target).

Imaging biomarkers of harm (toxicity) may be less familiar but are not uncommon. They may be used to monitor anticipated toxicities (for example cardiotoxicity of anticancer drugs), and abnormal values of efficacy biomarkers sometimes can be used as biomarkers of toxicity.

In later stages of drug development efficacy biomarkers of response can help select doses, schedule, and perhaps combinations, or inform the discussion between the drug developer and the regulatory authority about the design of the phase III programme including doses and safety assessment. Occasionally a change in a biomarker in Phase III trials may supports the regulator's decision to approve the drug where the clinical efficacy data are compelling but less extensive than would otherwise be demanded, or even support the regulator's decision to approve the drug in the absence of clinical efficacy data (i.e. surrogacy). More commonly, for a drug that has previously been approved, change in the biomarker may support a regulator's decision to approve the drug for an additional indication.

References

Atkinson Jr, A., *et al*. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework Clinical Pharmacology & Therapeutics (2001) 81, 104-107