The development of rapid high-throughput technologies to analyse DNA, RNA and protein has led to a paradigm shift in our understanding of the molecular basis of disease. Molecules that can be used to differentiate between normal and abnormal biological processes or predict treatment responses are termed “molecular biomarkers” (Box 1). Examples include DNA sequence mutations, epigenetic changes, and levels of messenger RNA or protein expression that are associated with a patient’s risk of disease events, survival time or response to treatment. Use of molecular biomarkers to supplement or replace conventional clinicopathological factors has the potential to transform the practice of medicine by creating new opportunities for developing and tailoring treatments to individual patients.

The clinical value of some molecular biomarkers has been established by trials that have demonstrated that biomarker-guided treatment strategies improve patient outcomes. For example, randomised controlled trials (RCTs) have provided evidence about the value of measuring oestrogen receptor (ER) status in women with breast cancer to identify those who will benefit from targeted treatment with anti-oestrogen therapy and those who will not. However, many molecular biomarkers discovered in observational studies are yet to be adequately evaluated in clinical practice.

In this article, we describe the principles of using biomarkers to individualise treatment and discuss the role of RCTs in assessing their clinical value.

What is individualised treatment?

We rely on RCTs to identify the most effective treatment for patients with a given condition. Classically, these trials randomly assign patients to new or standard treatment and compare the risk of disease events, or time to disease events, to measure the effectiveness of the new treatment as a relative risk or hazard ratio. The absolute risk reduction for individual patients depends on their baseline risk before treatment — that is, their prognosis — and whether their response to treatment varies from the overall effect measured in RCTs — the “treatment prediction” (Box 2). If we have access to this information, we can individualise treatment decisions by weighing up the size of benefits and harms of the new treatment for individual patients.

The full potential of using biomarkers to individualise treatment is not yet well understood. For example, trials have shown that anticoagulant therapy for the prevention of thromboembolic stroke in patients with non-valvular atrial fibrillation is more effective than standard care, but this does not mean that all patients will benefit. On the basis of trial results showing that anticoagulant therapy leads to a 70% relative risk reduction for major stroke, we can estimate that if 100 patients with a baseline risk of stroke of 10% were treated, seven strokes would be avoided. However, some patients would not have had a stroke, regardless of whether they received the anticoagulant therapy, and others would not respond to the anticoagulant therapy and still have a stroke (90 and three patients, respectively), but both groups might still experience the side effects of treatment. Ideally, a biomarker or panel of biomarkers could be used to identify the small subgroup of patients most likely to receive a clinical benefit — those who would otherwise most certainly have a stroke (the perfect prognostic test) and also respond to treatment (the perfect predictive test) — and offer treatment only to this group. This would obviate adverse events and unnecessary costs of treatment for low-risk patients and provide the opportunity to offer an alternative therapy to high-risk patients who would not be responsive to this treatment.

Identifying biomarkers that classify prognosis or predict response

When a biomarker is discovered to be a promising classifier of patient outcome, further studies are needed to validate its reliability.

1 Definitions

Biomarker: a “characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacological responses to a therapeutic intervention”. Molecular biomarkers include genetic factors measured as variations or mutations of DNA sequence, epigenetic factors, and levels of messenger RNA and proteins.

Prognostic biomarker: classifies an individual’s baseline risk of having a clinical event.

Predictive biomarker: classifies the magnitude of an individual’s response to treatment.
2 Role of biomarkers to classify patient prognosis or predict treatment response for individualised treatment decisions*

<table>
<thead>
<tr>
<th>Prognostic biomarkers:</th>
<th>Hypothetical example:</th>
<th>Predictive biomarkers:</th>
</tr>
</thead>
<tbody>
<tr>
<td>use to determine whether further treatment is needed</td>
<td>patients with atrial fibrillation at risk of thromboembolic stroke</td>
<td>Baseline risk (BR) of stroke –10%, estimated from epidemiological studies</td>
</tr>
<tr>
<td>Individual baseline risk</td>
<td>Relative risk reduction (RRR) –70%, estimated from randomised controlled trials comparing new treatment with standard care</td>
<td></td>
</tr>
<tr>
<td>Treatment relative risk reduction</td>
<td>Absolute risk reduction (ARR) –7%, calculated using the formula: ARR = BR x RRR</td>
<td></td>
</tr>
<tr>
<td>Absolute risk reduction</td>
<td>If 100 patients were treated, 7 strokes would be avoided</td>
<td></td>
</tr>
<tr>
<td>Size of benefits and harms</td>
<td>50 patients would not have had a stroke, whether they received the new treatment or not</td>
<td></td>
</tr>
<tr>
<td>Resources and community values</td>
<td>3 patients would not respond to the new treatment and still have a stroke</td>
<td></td>
</tr>
<tr>
<td>Individualised treatment decisions</td>
<td>All patients would be at risk of side effects from the new treatment (risk of intracranial haemorrhage –0.2 per 100 patients treated)</td>
<td></td>
</tr>
</tbody>
</table>


... used to explore the relationship between biomarker test results as a continuous variable and treatment effects.\textsuperscript{10,11}

The Cholesterol Treatment Trials’ prospective meta-analysis of cholesterol-lowering trials confirmed the prognostic value of conventional biomarkers (total serum cholesterol, age, presence of hypertension, etc) but found that none of these biomarkers provided additional information to predict treatment response (ie, no biomarker–treatment interaction).\textsuperscript{9} For example, patients with the highest cholesterol level (>6.5 mmol/L) had the highest risk of major vascular events, but all patients, including those at the lowest risk (cholesterol level, <5.2 mmol/L), benefited from cholesterol-lowering treatment, a relative risk reduction for major vascular events of around 20% in each risk group was reported (Box 3). Given similar treatment effects across different risk groups, patients at the highest risk will gain the greatest absolute benefit from treatment in terms of the total number of events avoided. Thus, prognostic information can be used to target treatment to patients at sufficiently high risk of events that the size of the absolute benefits justifies the costs and potential harms of the treatment.

Some biomarkers can provide both prognostic and predictive information, and this is often the case in cancer. An example is the ER status in breast cancer. Trial data from patients assigned to surgery without systemic adjuvant therapy have indicated that ER-positive tumours are associated with better early survival than ER-negative tumours; however, the difference in prognosis by ER status is relatively small and time-dependent.\textsuperscript{13,14} It was the discovery, through RCTs, that ER status predicts response to tamoxifen that transformed the use of anti-oestrogen therapy in the management of breast cancer. The Early Breast Cancer Trials’ Collaborative Group’s meta-analysis of tamoxifen trials showed that tamoxifen reduced the risk of death among women with ER-positive tumours but not among women with ER-negative tumours (test for interaction, $P < 0.00001$; Box 3).\textsuperscript{12}

**Trial design for biomarker evaluation**

When the potential role of a biomarker to guide treatment has been defined, its clinical value can be assessed by following the same rules of evidence that apply to the assessment of new treatments. Ideally, an RCT would be undertaken to test the hypothesis that the proposed biomarker-guided treatment strategy improves patient outcomes compared with the conventional testing strategy. Sometimes this information is available from existing RCTs of treatment. For example, tamoxifen trials have provided evidence about the difference in patient outcomes if tamoxifen is reserved for patients with ER-positive breast tumours compared with the scenario without ER testing where treatment may be given to all women (or none).\textsuperscript{12} One exception where an RCT may not be justified is when a prognostic biomarker can accurately identify patients at low risk of disease who can safely avoid treatment. In this situation, single-arm studies demonstrating the prognostic accuracy of the biomarker versus conventional tests may suffice.

**Classical trials**

Using a classical RCT design, the impact of treatment with the availability of a novel biomarker can be tested by randomly assigning patients to either the biomarker test with or without conventional tests, or conventional tests only (Box 4). This design evaluates the effectiveness of the combination of the new bio-
marker test and the subsequent treatment intervention. However, it is an inefficient design as a large number of patients are needed but only patients with a biomarker test result that disagrees with the conventional test result contribute to the expected outcome from the treatment intervention. All patients testing positive by both investigations will receive the same treatment intervention, hence the biomarker does not discriminate outcomes for these patients.

**Targeted trials**

An alternative approach is a targeted RCT, also referred to as an enrichment design, which involves upfront testing of all patients for the biomarker of interest and selecting only patients with biomarker results that will lead to a change in outcome with the proposed treatment. This is more efficient than a classical design as those without the biomarker of interest are excluded from the study, hence fewer patients are enrolled (Box 4). A positive trial result provides evidence about the effectiveness of using the biomarker to guide treatment. The design is generally used when there is robust a priori biological evidence that treatment is effective or more effective in the biomarker-defined population.

As targeted trials recruit patients with the same biomarker status, they can be used to assess the efficacy of a new treatment in a biomarker-defined population, or the efficacy of an existing treatment in a subgroup of patients newly defined by a biomarker. Recent oncology trials provide examples of these scenarios.

**Testing a new treatment in a biomarker-defined population.**

The pivotal trial for trastuzumab, a monoclonal antibody against the human epidermal growth factor receptor 2 (HER2) protein, was undertaken in a targeted population of women with metastatic breast cancer who were positive for HER2 (Box 5). This trial showed the efficacy of the combined strategy of testing for HER2 status and treating women with HER2-positive tumours with trastuzumab. Subsequent research has focused on improving assay methods and the cut-off level of HER2 amplification or expression to optimise the predictive performance of HER2, to identify women who will benefit from treatment with trastuzumab.

**Testing an existing treatment in a subgroup of patients newly defined by a biomarker.** Oncotype DX is a 21-gene prognostic assay developed to classify women with node-negative, ER-positive breast cancer into three categories according to their risk of developing recurrent disease (low, intermediate and high risk). It has been proposed to guide treatment decisions by sparing women who are at low risk unnecessary chemotherapy, and identifying those who are at high risk and need treatment. Oncotype DX is currently being prospectively assessed in the TAILORx trial (Trial Assigning Individualized Options for Treatment [Rx]). The primary objective of the trial is to investigate the efficacy of chemotherapy as an addition to hormone therapy in women who are at intermediate risk (recurrence score, 11–25) (Box 5). The working premise is that patients in the intermediate-risk group will do no worse with hormone therapy alone than they would with hormone therapy plus chemotherapy. This study assumes that chemotherapy does not improve outcomes in patients at low risk (recurrence score, <11) but will be beneficial in patients at high risk (recurrence score, >25).

Targeted RCTs offer an efficient method to show proof of concept for the efficacy of treatment in biomarker-selected patient groups. They can be used where there is a biological assumption that only a biomarker-defined patient group will benefit from the new treatment; hence restricting the trial to this group (eg, biomarker-positive patients) is more efficient and potentially more ethical. However, molecular pathways of tumour pathogenesis are complex and often not clearly understood, so questions about the efficacy of treatment in biomarker-negative patients may also warrant consideration. For example, trastuzumab was recently reported to have efficacy in HER2-negative patients in a retrospective analysis of a broader patient population that included both HER2-positive and HER2-negative patients.

**Non-targeted trials**

In contrast to targeted trials, non-targeted (or unselected design) trials do not restrict recruitment to a single biomarker-defined subgroup of patients. This design is needed to assess the predictive value of a biomarker to distinguish between patients who will...
respond differently to the same treatment (Box 4), or compare the efficacy of a new biomarker-guided treatment strategy with current best practice using conventional testing.

The International Breast Cancer Study Group Trial IX, which randomly assigned postmenopausal node-negative early-stage breast cancer patients, stratified according to ER status, to chemotherapy followed by tamoxifen versus tamoxifen alone, is an example of the first situation. An example of the second situation is the ongoing MINDACT (Microarray In Node-negative Disease may Avoid Chemotherapy) trial, which is comparing the prognostic value of a 70-gene signature for breast cancer with established clinicopathological criteria to identify women with node-negative early-stage breast cancer who can safely avoid adjuvant chemotherapy (Box 5). Unlike the TAILORx trial, the MINDACT trial does not restrict randomisation to one biomarker-defined patient subgroup. All women are assessed by using both the new biomarker gene signature and conventional criteria to classify their risk of disease recurrence. Only women with discordant results (biomarker-negative and conventional criteria-positive, or vice versa) are randomly assigned to receive chemotherapy or no chemotherapy; women with concordant results are not randomised but treated according to the standard of care (women at low risk with observation, and women at high risk with chemotherapy). This trial will provide data that compare patient outcomes when chemotherapy is selected according to the biomarker versus conventional criteria.

Non-targeted trials generally require a large sample size, particularly if the incremental effects of offering treatment to patients reclassified using a biomarker are small. Hence, their main disadvantage is that they may not be feasible if a small proportion of patients tested with the target condition are eligible — for example, if biomarker prevalence is low. Design variations to optimise the efficiency of non-targeted trial designs are discussed elsewhere and are beyond the scope of this article.17

Biomarker analysis within an existing RCT

A pre-specified analysis of prospective data collected from an existing RCT of treatment can sometimes be used to test a hypothesis about the prognostic or predictive value of a biomarker (Box 4). This study design was used to assess K-ras status as a predictor of response to panitumumab, an epidermal growth factor receptor inhibitor, in patients with metastatic, chemotherapy-refractory colorectal cancer. Investigators used archival tissue from an earlier RCT, which demonstrated the efficacy of panitumumab compared with best supportive care in a non-targeted population, to test the hypothesis that K-ras status predicts treatment response (Box 5).22,25

This study design has the advantage of time and cost efficiency. Its major disadvantage is the potential for selection bias if archival tissue or serum data are not available for all patients. Other disadvantages include measurement bias if measurements of the biomarker or statistical analyses are not blinded, and chance false-positive findings due to multiple comparisons, particularly with trial data used to explore multiple candidate biomarkers. These problems reduce the validity of study results, so analyses should only be considered if the biomarker can be measured on all or a large representative sample of all patients. Furthermore, the design needs a biologically plausible hypothesis, a prospectively defined protocol for the biomarker assays, and statistical analysis plans.

Implications for future trials

The current availability of genomic, transcriptomic, proteomic, metabolic and other similar technologies provides unprecedented opportunities for individualised treatment through the discovery of molecular biomarkers and the development of molecularly targeted therapies. The conventional rules of evidence for evaluating these new technologies have not changed, but the need for more efficient RCTs has led to innovations in the design of targeted and non-targeted trials. If an existing treatment trial has archived specimens from all, or most, participants, then an efficient and reliable evaluation of a new biomarker may be achievable by further analysis of these specimens. Future trials should be designed to anticipate these data needs.

Acknowledgement

We thank Rhana Pike for assistance in preparation of this article.
## 5 Purposes of randomised controlled trials (RCTs) used to evaluate biomarkers: recent examples in oncology

<table>
<thead>
<tr>
<th>Trial</th>
<th>Target condition</th>
<th>Inclusion criteria</th>
<th>Treatment comparison</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Targeted trials</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HER2 trastuzumab trial</td>
<td>Progressive metastatic breast cancer</td>
<td>HER2-positive tumour</td>
<td>Trastuzumab + chemotherapy versus</td>
<td>Test for effect of new treatment in a subgroup of patients defined by a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>chemotherapy alone</td>
<td>existing biomarker</td>
</tr>
<tr>
<td>TAILORx</td>
<td>Operable node-negative, ER-positive, HER2-negative</td>
<td>Oncotype DX score 11–25 (intermediate risk of recurrence)</td>
<td>Chemotherapy + tamoxifen versus</td>
<td>Test for effect of existing treatment in a subgroup of patients defined</td>
</tr>
<tr>
<td></td>
<td>early-stage breast cancer</td>
<td></td>
<td>tamoxifen alone</td>
<td>by a new biomarker</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Non-targeted trials</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBCSG Trial IX</td>
<td>Operable node-negative, postmenopausal early-stage</td>
<td>ER-positive, ER-negative and ER-unknown</td>
<td>Chemotherapy + tamoxifen versus</td>
<td>Test for differences in treatment effect according to biomarker status</td>
</tr>
<tr>
<td></td>
<td>breast cancer</td>
<td></td>
<td>tamoxifen alone</td>
<td></td>
</tr>
<tr>
<td>MINDACT</td>
<td>Operable node-negative, ER-positive early-stage</td>
<td>Genomic test result discordant with clinical criteria</td>
<td>Chemotherapy versus no chemotherapy</td>
<td>Test for differences in treatment effect between biomarker-guided- and</td>
</tr>
<tr>
<td></td>
<td>breast cancer</td>
<td>for classification of risk of recurrence: genomic low</td>
<td></td>
<td>conventional test-guided treatment strategies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>risk, clinical high risk; or genomic high risk, clinical</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>criteria for classification of risk of recurrence:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>genomic high risk, clinical low risk</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>risk of recurrence: genomic low risk, clinical low risk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biomarker analysis within</td>
<td>K-ras panitumumab trial</td>
<td>Subset of trial participants with measurable K-ras</td>
<td>Panitumumab + best supportive care versus</td>
<td>Test for differences in treatment effect according to biomarker status</td>
</tr>
<tr>
<td>an existing RCT</td>
<td></td>
<td>mutation status of tumour (mutant versus wild-type)</td>
<td>best supportive care alone</td>
<td></td>
</tr>
</tbody>
</table>

EGFR = epidermal growth factor receptor. ER = oestrogen receptor. HER2 = human epidermal growth factor receptor 2. IBCSG = International Breast Cancer Study Group. MINDACT = Microarray In Node-negative Disease may Avoid ChemoTherapy. TAILORx = Trial Assigning Individualized Options for Treatment (Rx).

### Competing interests
None identified.

### Author details
Chee K Lee, MB BS, MMSci, FRACP, Oncology Research Fellow
Sarah J Lord, MB BS, MS(Epi), Epidemiologist
Alan S Coates, AM, MD, FRACP, Clinical Professor
R John Simes, MB BS, FRACP, SM, Director
1 NHMRC Clinical Trials Centre, University of Sydney, Sydney, NSW.
2 Screening and Test Evaluation Program, University of Sydney, Sydney, NSW.
3 School of Public Health, University of Sydney, Sydney, NSW.
4 International Breast Cancer Study Group, Bern, Switzerland.
Correspondence: chee.lee@ctc.usyd.edu.au

### References
FROM BENCH TO BEDSIDE


(Received 21 Nov 2008, accepted 16 Feb 2009)