Beyond survival – what should new cancer drugs have to prove and how?

Anna Wagstaff

Randomised controlled trials have been the gold standard for testing new drugs, and survival is the standard by which they succeed or fail. As our understanding of cancer increases, therapies become more numerous and more complex, and patients live longer, is this still the way to get the best drugs into use most quickly? If not, is there a credible alternative?

Demonstrating in a randomised controlled trial (RCT) that a new treatment keeps patients alive for longer has long been seen as the gold standard evidence for new cancer drugs. That doesn't mean it has been uncontroversial – far from it.

On the plus side, this gold standard answers the key question for patients and doctors: "What is likely to keep me alive for longest?" It also gives payers a strong evidence base to assess the value of the drug. Using 'surrogates' for survival, such as response rate – significant tumour shrinkage – or progression-free survival (PFS) – how long the therapy holds the cancer at bay – are seen as far weaker measures. The notorious ability of cancer cells to find alternative pathways means that early indications of response are often not sustained. Surrogates are also harder to measure than survival, where the endpoint is the finality of death. It can be difficult to interpret what is happening to a tumour even on MRI, giving rise to the phenomenon of pseudoprogressions and pseudoresponses (see e-grandround in this issue for a spectacular example in gliomas). Measurement of progression is also open to variations and depends heavily on how often the patient is evaluated.

However, there is a downside to using survival as the key indicator. It means that researchers must continue a trial until enough patients have died to show a statistical difference in survival. This can be a long and expensive process, especially where the benefit is incremental – which, sadly, is often the case. If the relevant patient population is small and networking is poor, recruitment to the trial can be a very slow process.

A number bad things flow from this:

- potentially beneficial drugs are slow to reach patients who are running out of options
- the cost of the process pushes up the price of the drug, which could restrict patient access
- the time and money used in getting statistical answers could be used for other research
- the longer it takes to answer the questions posed by the RCT, the greater the risk that the question loses relevance, as the standard of care changes or greater insights are gained into the way the disease works.

There are also ethical issues about trials

ILLUSTRATION: FRED VAN DEELEN, WWW.ORGANISART.CO.UK

that require patients in the control arm to die early to prove the superiority of the experimental arm, when their lives might have been extended had they been allowed to cross over to the experimental therapy as soon as it became clear that they were showing a poorer response rate or were progressing earlier.

Times are changing

There are genuine dilemmas here, with no easy answers. But a number of trends in cancer research and cancer care are now changing the terms of this debate:

Better care. For many cancers, survival times are creeping up as the result of improvement in care, including supportive care, and greater specialisation and multidisciplinary working. This is good news for patients, but means that survival endpoints take ever longer to reach.

More therapies. By the end of their lives many cancer patients will have been treated with four, five or six different 'lines' of therapy, moving on to something new each time the previous one ceased to be effective or the side-effects proved too troublesome or a new more promising drug made it to market. For each drug you can measure response rate and PFS, but death happens only once: how can you tell what contribution each drug made to overall survival?

Better organised patients. As patient groups have become more organised and vocal, it is becoming increasingly difficult to justify or recruit to trials that do not allow control patients to cross to the experimental arm once that arm has shown it does better on the PFS measure. The whole purpose of allowing crossover is to minimise the survival gap between the two arms, making it hard if not impossible to demonstrate superior overall survival.

Shrinking patient populations. Randomised controlled trials are all about numbers and statistical proof. As researchers succeed in differentiating the disease into hundreds of biologically distinct entities, the number of patients appropriate for each trial shrinks, making RCTs less of a practical option.

Decreasing toxicity. While cardiac toxicity, for instance, can still be a serious issue with some biologics, in general, newer therapies, including vaccines, have side-effects that are less threatening and less debilitating than traditional cytotoxics. The need to prove beyond doubt the survival benefit of a new drug becomes less pressing where the patients have less to lose.

Intelligent design. RCTs can give answers to specific questions even if you have no understanding of what is driving the disease, or how the drug works or in whom. Now that we understand more about the disease and drug developers invest heavily in extensive preclinical and early clinical studies to build up a clearer picture of their drug, might alternatives to lengthy RCTs be acceptable?

As cancer care and drug development move forward, should proof of overall survival benefit as shown in a randomised controlled trial still be the gold standard for new therapies? How can drug developers provide patients, doctors, regulators, trial participants and payers with the evidence they are looking for? *Cancer World* put the question to some of them.

Committed to showing survival benefit

Uday Bose is European Head of Oncology at Eisai, a Japanese pharmaceutical company that recently entered the cancer field with Halaven [eribulin], a cytotoxic, that was approved by the EMA this March for use in advanced breast cancer.

Bose questions how far overall survival is really seen as the gold standard, citing a study that showed no more than 15 out of 76 phase III studies in metastatic breast cancer published between 1998 and 2007 had overall survival as their primary endpoint, and met that endpoint (*JCO* 28:1958–1962). PFS has become an increasingly common surrogate in this setting. A consequence of this, he argues, is that while women are typically treated with four, five or even more lines of treatment, after the first two lines, doctors and patients have little evidence for survival on which to base further choices.

"I think there is a correlation between the promotion and acceptance of PFS data, because they are the primary data that are being generated. But from the research we did in preparation for our launch, the message that this is the right endpoint to be looking at, rather than overall survival, seems to have been accepted by oncologists as well."

When Eisai presented physicians with a single page showing the profile of Halaven, says Bose, their attention was immediately drawn to the PFS data – the secondary endpoint of the study. They were very interested in the overall survival data when they saw it, he says, but they didn't actually look at it until it was explicitly pointed out. "It's stark how the environment has now evolved in prescriber land that PFS is a valid surrogate, and they are quite convinced that it is a fair and a strong endpoint, even when they are presented with overall survival as primary endpoint."

Bose hopes that Eisai's success in showing overall survival benefit will challenge what he sees as a defeatist acceptance that meeting an overall survival endpoint is an unrealistic expectation in late stages of disease.

He does accept, however, that there are many situations where proving survival benefit may not be possible, and it was a delicate balancing act even in the EMBRACE trial, which demonstrated an extra two months of life for women with metastatic breast cancer who were put on Halaven as a third-line or later treatment. To make the trial more palatable to potential participants, Eisai allowed almost unconstrained 'treatment of physician's choice' (TPC) as the con-

RCTs: GOLD STANDARD OR BLUNT INSTRUMENT?

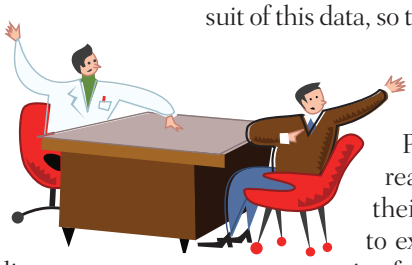
Randomised clinical trials are used to subject hypotheses such as "patients will live longer on drug A than drug B" to a statistical test. They need to recruit enough patients to show, with a high degree of certainty, that the survival difference between the trial arms really does reflect superiority of the treatment rather than having come about by chance. This measure of certainty is represented in the all-important *P*-value; $P<0.001$ being a way of saying that there is a one in a thousand chance that the survival difference shown in the trial, or an even higher difference, does not represent a real and replicable difference. The smaller the difference between the two arms, the more patients must be recruited to reach statistical significance.

Bayesian methodologies, in contrast, make use of all relevant knowledge gained through the multiple studies done during the process of drug development – on the role of the target, the ability of the drug to hit the target, the effect of hitting the target, perhaps the effect of adding a second drug, the dose levels, predictive biomarkers and so on – and can incorporate them mathematically as 'priors' into a model that presents the strength of evidence in terms of 'credible intervals', which are equivalent to the more familiar 'confidence intervals'.

Cancer World will look at Bayesian trial methodologies in greater depth in a future issue.

trol arm – including best supportive care.

As it happens, says Bose, no patients chose best supportive care – something that pleased the patient representatives, he says, "because it challenges the perception that once a woman has gone through first-/second-line treatment they give up and they don't want anything else."



Halaven is currently in a head-to-head study against capecitabine, in an earlier line of disease, after an anthracycline and a taxane. This time Eisai has chosen to use progression-free survival as a co-primary endpoint with overall survival. The company is clear, however, that whatever happens, the trial will continue until there are sufficient overall survival 'events' (i.e. enough deaths) to demonstrate a significant difference in survival. They will not do what so many phase III trials do, and publish an interim report when the number of PFS 'events' (i.e. progressions) has reached a point where they are likely to show a significant difference between the two arms, and then either stop recruiting or allow patients on the control arm to switch over to the experimental treatment.

"If we were to go out with our PFS endpoint, then in our conversations with payers, they may say, 'But you compromised the study. You had a survival endpoint, why didn't you stick with it?'"

Bose has seen exactly this happen with some other cancer drugs, and he doesn't want to repeat the mistake. There's no great ethical achievement, he points out, in stopping a trial early or allowing patients to cross over to the experimental arm on the basis of promising PFS data, if the consequence is that payers then refuse to reimburse the drug on the grounds of insufficient evidence on survival.

There are three things he would like to see happen. One is that drug developers and oncologists raise their ambitions and go the extra mile to try to get overall survival data wherever they can. The second is for a regulatory and payer environment that encourages the pursuit of this data, so that companies won't feel they could end up penalised if they have strong PFS data but fail to reach significance on their survival data, due to explicable confounding factors.

The third, which he believes is crucial to being able to give payers what they want, is the introduction of a value-based system of pricing that recognises that any given cancer drug can give different levels of benefit depending on what cancer, what stage, and what line of treatment it is used in.

Identifying your target group is the key

Oliver Kisker is vice president of global clinical development for the oncology unit at Merck-Serono, where he works with a wide variety of cancer therapies including Erbitux [cetuximab], the *EGFR* inhibitor approved for use in some colorectal cancers and squamous cell head and neck cancers, and for which Merck is now seeking approval for use in certain non-small-cell lung cancers.

Kisker shares the view that being able to show your drug improves survival is always desirable, but in some indications it is difficult to achieve: "In some areas, where few treatment options are available, it is important to demonstrate that overall survival is really better compared to the competitor. But if you have an indication where treatments are much more available, like for colorectal cancer, it will be much more difficult."

How can you prove the overall survival advantage of a drug used first line in colorectal cancer, he asks, without dictating in the protocol what patients should get not only in the first line, but also in second, third and fourth lines – which is not something physicians or patients would be likely to accept. And if you don't, then how can you stop differences in overall survival being confounded by differences in the subsequent therapies?

The answer, he suggests, is to ensure the survival benefit is sufficiently strong to show through despite the confounding impact of therapies taken after the trial. And the way to do that is to identify the patient group that derives a real benefit from the new treatment.

This is how Merck showed the survival benefit for adding Erbitux to the FOLFIRI regimen for first-line treatment of patients with metastatic colo-rectal cancer. In an undifferentiated patient population, the combined treatment showed a significant improvement in PFS, with a hazard ratio of 0.85, but the difference in overall survival failed to reach significance. However, Merck had taken tissue from the majority of its trial patients, and was able to reanalyse the data after stripping out the results from patients with a mutated *KRAS* gene. This effectively doubled the response rate figures for the wild-type (normal) *KRAS* patients; it strengthened the difference in PFS data from a hazard ratio of 0.85 to 0.696; and, importantly, the difference in overall survival became statistically significant, showing an additional 3.5 months over the control arm.

"This demonstrates you can do it," says Kisker. "It's not just a question of overall survival as a primary endpoint; it's a question of how to develop our products. We have to address, even in preclinical models, how drugs might work, what is the mode of action and what might be potential biomarkers. You then go for a phase I study, which should be used to identify patients who might

"There will always be situations where proving overall survival benefit simply isn't possible"

benefit, by including even at that stage a marker that could identify the right patients. You then use expansion cohorts [add in patients with the marker of interest] where you can see if these patients really do benefit.

"Then we have a combination of expansion cohort and biomarker, and then you go to phase II, which gives a much clearer picture of how patients might benefit based on molecular profile. Then you do additional analysis here with further markers, identify them, the right profile, the right patients, and then you go to phase III."

This is the strategy Merck is following now with all its cancer drug developments, says Kisker. In lung cancer they are looking at high levels of *EGFR* expression as a marker for response to Erbitux. And they are investigating the *MGMT* biomarker, associated with DNA repair function, as a possible predictor of response to temozolomide, which is combined with cilengitide, their experimental integrin inhibitor for first-line use in glioblastoma.

Trialling the drug only in the population that responds well not only increases the benefit you can show, but as Kisker points out, it also decreases the number of 'events' needed to prove this bigger benefit, which means trial sizes are smaller.

But no matter how well you do this work, he adds, there will always be situations where proving overall survival benefit simply isn't possible. A classic example is where you are trialling a drug for use in first line, when it has already been approved in a later setting. It is not only unethical but also impossible to deny a patient in the control arm access to the treatment once they have progressed to the point where that drug has been approved and is freely available for use. "If you have crossover you need to think about it and discuss with the regulators about the crossover effect of a drug already approved for a later stage indication."

Despite the extensive early trial work involved, however, Kisker feels that new drugs still need to prove themselves in standalone phase III RCTs. "You try to answer questions you have raised in phase I and early phase II, but in the end you need to show it in a phase III, because this is a requirement by regulatory agencies."

He doesn't rule out the possibility of extending phase IIs into the phase IIIs in the future, though it's not a design Merck currently uses. "I think it is an interesting approach, that you could carry on and reduce time to approval. But I think we would need to have further discussion with agencies, because designing studies in this way is not always accepted by agencies. This is one of the things where we need to interact with agencies to speed approval of drugs by using these kinds of designs."

Better interaction is also his solution to the question of how to satisfy the demands of payers. Kisker makes time to talk to the people who have a say over reimbursement, to discuss the issues he faces in developing a particular drug, and to find out from them what sort of data they need. "Payers are becoming more and more important, and both sides need to understand one another. They need to understand where we are, because they want to see patients benefit from therapy. We on the other hand need to see what are the points that we have to address."

He emphasises again the importance of finding the right patient group. "You need to include as early in the trial as possible personalised medicine through stratifying patients using biomarkers. If you see benefit for these patients, payers will have nothing against it."

One size does not fit all

Rafael Amado is Head of Oncology Development at GlaxoSmithKline, where current strategies include starting development with a tightly defined patient population – patients with known cancer promoting molecular alterations, such as *BRAF* mutations, for example – and developing drugs, or combinations of drugs, that will be effective in the small population of patients whose cancers are driven by those alterations.

"It's fair to insist on survival data when you are using broad-spectrum, toxic treatments which afford only small incremental benefit, as in the case of cytotoxics in most advanced cancers," says Amado, "but when you are using drugs targeted to molecularly characterised populations, which can drive large effects in surrogate endpoints, and are less toxic, I do not think that overall survival needs to remain the gold standard against which we measure new drugs. I understand that we have to show at least reasonable likelihood of clinical benefit. But if we continue to think of overall survival as the gold standard, the development process will continue to be long and cumbersome, and it will become more and more difficult to obtain it as an

endpoint. Indeed, as diseases become more chronic, waiting for survival will continue to tie up investment and resources and delay innovation."

He mentions the controversy over the use of overall survival as an endpoint in melanoma with innovative drugs such as *BRAF* and *MEK* inhibitors. These are drugs with understood mechanisms of action that have shown impressive response rates and progression-free survival in advanced melanoma. Trials are being conducted against dacarbazine – an ancient and largely ineffective cytotoxic. Carrying on the trial until enough patients die to reach statistically significant data on overall survival would rule out the possibility that PFS gains might fail to translate into longer survival, as the cancer finds alternatives to the blocked pathway – which turns out to be a real concern. But Amado says GSK would not be prepared to go down that road.

"We are developing a *MEK* and a *BRAF* inhibitor. Our randomised trials use crossover from the control to either *MEK* or *BRAF* after disease progression, or a control arm that includes an active targeted therapy, as we feel all patients in these trials should have the opportunity to access these drugs."

Amado concedes that progression-free survival is not a foolproof surrogate for overall survival. "In the field of angiogenesis for example, there have been preclinical studies showing that a rebound pro-angiogenic effect can occur after withdrawal of antiangiogenic therapy, suggesting that while a patient can respond during treatment, when one withdraws the drug the tumour may come back with a more aggressive phenotype. So I think it is incumbent on the investigators and sponsors to show that a phenomenon like this doesn't occur."

One way to do this, says Amado, is to look at whether the early difference between the overall survival (OS) data for the two arms wanes over time. "If you see OS hazard ratios that are trending progressively in the wrong direction after drug discontinuation, that should raise concerns. There are also analyses one can do looking at time to death from disease progression between test and control arms, which can help rule out a potential rebound effect," he says.

The question that should be asked, he suggests, is whether failure to meet a survival endpoint was due lack of statistical power, lack of a treatment effect, excess toxicity of the treatment arm, compromise of delivery of standard therapy, or tumour promotions such as directly or via a rebound effect. "For instance, we recently learned that the use of *EGFR* inhibiting antibodies in patients with colorectal cancer harbouring *KRAS* mutations seem to indeed shorten survival."

Amado questions the need for randomised controlled trials as the gold standard for all drug approvals, and points out that Bayesian designs are often used by sponsors and the US National Cancer Institute to do proof-of-concept trials, and are endorsed by the FDA for use in device approvals. "Traditional statistical designs have the potential to slow down drug development particularly in disease settings with small patient populations, such as for instance non-small-cell lung cancer with *ALK* translocations or *BRAF* mutations. To compound the problem, effective inhibition of some genetic aberrations may depend on blocking more than one target to ensure efficacy or prevention of resistance. The use of Bayesian statistics can also model not just the overall treatment of targeted drugs alone or in combination, but how their effects vary depending on a variety of factors, including biological heterogeneity".

The RCT approach can only really answer one or two questions at a time, says Amado, and is simply too blunt an instrument for these sorts of developments, because there are too many variables: in what type of tumour does a given drug work best, in which molecular alteration, which pathways do you target in the setting of combinations (and which plays the primary role) and what doses do you use. Many companies are therefore already relying heavily on Bayesian approaches (using modelling and probability methodology – see box p 24) to guide their proof-of-concept development, says Amado, and he expects that trend to continue.

"When you have so many variables and are trying to test a combination against a given standard, you end up with multiple-arm studies. And if you are not incorporating the knowledge that you get from every patient you end up with a large proof-of-concept trial that is often very difficult to interpret beyond the primary endpoint and safety. So eventually proof of concept is going to be more and

"The RCT is too blunt an instrument for these sorts of developments, because there are too many variables"

more iterations of trials in which arms get added on and arms get graduated or removed. We are now often using these adaptive and Bayesian designs."

Neither the EMA nor the FDA has ruled out approving a drug on the basis of evidence generated using Bayesian methodologies, and Amado hopes that the field will evolve in that direction, particularly in situations with highly biologically segmented populations where the use of novel agents result in large treatment effects. "In oncology we still have to wait for the first example of a full approval to come out of a Bayesian design. But I think that if regulators are willing to accept proof of concept based on Bayesian design as end of phase II data, it is only one step removed from accepting these designs for approval. Consider that large effects in OS mediated by a novel agent in a phase III RCT require a much higher and faster rate of death events occurring in the control arm than in the novel therapeutic arm. The question is whether allowing excess patients to die at a higher and faster rate in the control arm is appropriate when we know that the novel drug is highly active from phase II studies. For instance, when phase II studies already suggested that PFS with a novel agent is substantially longer than PFS or even OS observed with traditional chemotherapy, one could argue there is a loss of equipoise in randomising patients to the control arm".

GSK has been in discussions with regulators in US and EU to reach agreement on the design of RCTs for use in registering new combination therapies. "When using combination therapies you have to supply proof of the contribution of each compound to the benefit of the combination, and to do that one needs relatively large trials involving at least three arms. One way to decrease the size of the trial is by using a surrogate endpoint (e.g. PFS instead of OS) in one of the comparisons. Another step to simplify the development of two unapproved drugs is to use one of them as a comparator, rather than including a fourth arm for an approved standard. This can be done if the drug has significant activity as a single agent in phase II; although such a trial, if successful, would result in approval of the combination alone, it would likely not support approval of each of the agents as monotherapy."

For the payers, says Amado, the big issue may become whether paying for both drugs up front offers better value for money than using the two drugs in succession. "It is incumbent on us to demonstrate that the value of the combination goes beyond an endpoint such as progression-free survival or even overall survival, because these comparisons are done to single agents and not to sequencing multiagents." We will have to demonstrate that using combinations upfront is superior to the sequential use of each drug in terms of clinical outcomes and cost-effectiveness. In the case of *BRAF* and *MEK* it is possible that sequential use may be of no value as drugs may be cross resistant; in that case the only possible use of them would indeed be in combination.

Tailor-made trials

Hilary Calvert is director of Anticancer Drug Discovery at the University College London Partners, where he is involved in many phase I trials. He confesses to an ambivalent attitude on the need always to demonstrate overall survival benefit. Whatever else cancer patients may want from a drug, says Calvert, we can be pretty confident that they want it to make them live longer, so we do need to show that can happen,

WHAT DO PATIENTS WANT?

Cancer patients do want to live. But at what cost? As survival times increase, issues of quality of life become increasingly important, and drug developers are now encouraged, by regulators and payers, to incorporate quality of life measures into their trial designs.
How best to do this remains a problem.

- Studies show that doctors consistently rate side-effects as less significant than they are rated by patients – and it tends to be doctors rather than patients who fill in the trial forms.
- Even where patients are asked to rate side-effects on a scale, the frequency or severity may say little about how much it matters to the patient – diarrhoea may be less debilitating if you don't have to be out and about a lot; loss of feeling in the fingers, or disfiguring rashes affect people different ways. Even indicators such as whether the patient can continue working depend to some extent on what options they have.
- Patients may also have reasons to hide from their doctors the severity of side-effects if they think that telling the truth may lead to them being taken off a treatment they want to keep.
- Evidence on how patients see the trade-off between extra months of life and quality of life is scant and somewhat contradictory. A study done in 1990 showed that patients are prepared to take a greater hit on their quality of life for some extra time than their doctors (or the general public) would consider acceptable (Slevin et al., *BMJ* 300:1458–1460). A more recent study presented at the 7th European Breast Cancer Conference (Sheik-Youssouf et al., *EJC* Suppl 8:77), which looked exclusively at patients with metastatic breast cancer, suggests doctors require the offer of an additional two to six months of life as enough to consider trying a new therapy rather than the standard options, while almost two-thirds of patients want the promise of at least 10 months' additional survival.

"Cancer is just too complex and varied for golden rules or one-size-fits-all gold standards"

especially where the drug is fairly toxic. But then he cites the history of development of AIDS therapies, which took place without any of the stringent regulatory controls imposed in cancer.

"We've seen an absolute revolution in survival in HIV and they're now saying that it maybe takes five years off your expected lifespan rather than killing you within a few years. With all the enthusiasm and emphasis put on AIDS research, there's now about five different targets in the HIV system and about five different drugs available for each one. Physicians don't have any restrictions on prescribing them and nor have they ever had to prove they are value for money."

Progress in AIDS therapies was all about finding the combinations that work best, which makes it an interesting analogy to current approaches in cancer. "What they do is they look at the viral load, and they measure the changes very quickly until they find the combination that works." If AIDS research had been forced to jump through the sorts of hoops still required of cancer therapies, where you have to prove each individual drug with a set of trials for overall survival, says Calvert, they would never have progressed as fast – if at all.

He concedes, however, that AIDS, like heart disease and many other conditions, has something that cancer lacks: a good surrogate endpoint. Cancer has no equivalent of viral load or cholesterol level, which have been shown to correlate closely with survival. "Maybe the closest analogy in cancer is chronic myeloid leukaemia, where you can look for the BCR-ABL fusion, so you have a quick marker, and consequently it has been possible to develop a number of different drugs. I'm not saying it would work for all forms of cancer, but this is why I am a bit ambivalent about the very rigorous approach."

Calvert doesn't deny that approving a drug on progression-free survival can lead to wrong decisions. Iressa (gefitinib) for non-small-cell lung cancer was a case in point – it was approved by the FDA on the basis of its PFS figures, but it failed to show significant improvement in survival. This might never have become clear if there had been no requirement to show overall survival benefit. Avastin (bevacizumab) was another similar case. Calvert suggests that, given what we know about the fiendish ability of cancer cells to mutate in order to keep multiplying, the possibility that their response to being deprived of vascularisation would be to become more invasive in the search for alternative sources of blood might have been anticipated by both developers and regulators.

Which is all very well to say in hindsight, but is there any way to say in advance when disease-free or progression-free survival may be an acceptable surrogate and when it is not? "I can't think of a rule that would tell you that. Looking at a particular drug and its mechanism of action I could give you an opinion – it might well be wrong. It's a good question but a tough one."

What Calvert is saying, in effect, is that cancer is just too complex and varied for golden rules or one-size-fits-all gold standards. By the same logic, he agrees that Bayesian trial designs should replace the gold standard RCT in certain settings in the future, despite its quite formidable complexity. "I think traditional hypothesis testing methodology may well get very clumsy for things where we have a rationale for selecting quite small subsets of patients and giving them different things. We do need a mathematical logical approach that will take our subjectivity out of whether we think something is working or not. But the classical RCT with a 0.05 *P*-value and one hypothesis that you accept or reject on the basis of it may be too blunt an instrument for that."

That then leaves the question of how drug developers are going to convince regulators and payers of the risk–benefit and value of their products, without any gold standards for approval? Just like the development process, says Calvert, you have to do it on a drug-by-drug (or combination-by-combination) basis. "People need to take on board getting the right expertise onto the committee that makes the judgement [on approval or reimbursement], and really engage in a lot of detail about each drug, its mechanism, and the best way to evaluate it. I don't think there is a global solution, but if you know in enough detail how things are happening, you can come up with a good plan for each individual drug."