

# The FOCUS4 design for biomarker stratified trials

Richard Kaplan

Medical Research Council Clinical Trials Unit, University College London, London, UK

Correspondence to: Prof. Richard Kaplan. Medical Research Council Clinical Trials Unit, University College London, 125 Kingsway, London WC2B 6NH, UK. Email: r.kaplan@ucl.ac.uk.

**Abstract:** Randomised clinical trials (RCTs) remain the gold standard of evidence for the benefit of new therapeutics but standard designs fit awkwardly with key developments in biomarker-stratified drug development. Firstly, the unprecedented number of new agents being developed in oncology (usually with specific targets for which there may be predictive biomarkers) mandates a need for new trial designs that are more efficient in screening out new agents with modest likelihood of benefit, concentrating resources on the most promising ones. The multi-arm multi-stage (MAMS) design developed some years ago addresses this need. Secondly, biomarker-stratified trials, when tackled one biomarker/drug pairing at a time, are inherently highly inefficient. The FOCUS4 trial design was developed to overcome this problem, using a platform that incorporates multiple parallel biomarker-stratified RCTs in individual cohorts, and capable of adapting its design in response to developing evidence.

**Keywords:** Stratified medicine; trials design; colorectal cancer; biomarkers

Submitted Nov 11, 2014. Accepted for publication Jan 25, 2015.

doi: 10.3978/j.issn.2304-3865.2015.02.03

View this article at: <http://dx.doi.org/10.3978/j.issn.2304-3865.2015.02.03>

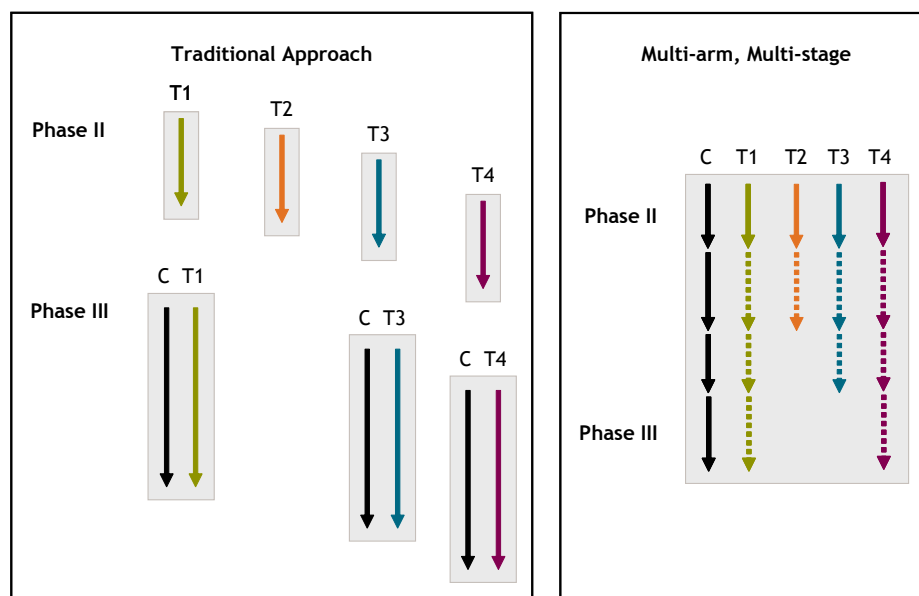
Oncology therapeutics discovery and development is one of the most active research areas for the pharmaceutical and biotechnology industries and likewise enjoys an especially strong academic research base. There has been an explosion of discovery in biomedicine relevant to cancer signalling, genetics, genomics and bioinformatics that has ended the empiricism and serendipity that used to characterise cancer drug discovery—and replaced it with effective high throughput methods for identifying potential drug targets and synthesising or screening agents directed to these targets. There are now many more cancer ‘leads’ and candidate drugs than ever before.

However, despite the recent advances in the development of agents aimed at precisely identifiable molecular targets, the rate of success remains relatively poor and the time scale from early clinical testing to registration is usually a matter of many years. There are exceptions, of course, just as there had been with classical cytotoxic chemotherapy. A handful of cancers are entirely dependent on single driver translocations or mutations for which targeted agents have been dramatically effective and have received marketing approval in record time. However, most solid

tumours have now shown themselves to be dependent on more complex genetic and/or epigenetic mechanisms, and accordingly most new targeted agents have proven no easier to establish as meaningfully or lastingly effective than older chemotherapeutics.

With the fundamental science moving so quickly, lengthy development times for new agents are intensely frustrating for researchers and clinicians, and even more so for patients. Lengthy development is also of course problematic for pharmaceutical companies, but has consequences for society and healthcare resources that are rarely mentioned and probably underappreciated: if a sufficiently effective agent takes the usual 8+ years to get to market (and perhaps longer to establish its optimal relative effectiveness), then there is a very limited period of patent protection remaining in which the pharmaceutical company can recoup its massive investments. If the pathway to confirming the benefit and licensing of novel agents can be shortened, healthcare systems and society should be better able to apply much needed downward pressure on drug pricing.

As a whole therefore, the fundamental challenges for development of new cancer therapeutics are:



**Figure 1** Schematic of MAMS design.

- How can we speed up development and testing, shortening time to patient access?
- How do we assess activity in early phase trials to improve our success rate in novel agent development?
- How can we predict which patients will respond to a new agent/regimen?

One of the most effective ways to make progress toward accomplishing these goals is to refine our trial methodology to better suit the current environment. Some of the ways this can be done include:

- Identify informative clinical settings (regardless of whether or not also obvious settings for licensing);
- Biomarker-enrich and seek a strong signal of activity in the form of an ambitious hazard ratio (a less strong signal in biomarker-negative patients can be sought later);
- Employ multi-stage trials to eliminate inactive or minimally active agents at as early a point as possible;
- Employ multi-arm trials (test several agents at once);
- Use an ‘umbrella’ or ‘rolling’ trial structure to avoid the long delays in getting new trials up to full speed.

Several of these ideas will be discussed in this paper and have been brought together in the trial design created for the MRC Clinical Trials Unit’s FOCUS4 trial in colorectal cancer.

Oncology is a particularly suitable testing ground for new designs/methodology that may be able to improve drug development success and shorten development time.

The large numbers of candidate agents emerging from early testing are effectively competing for access to clinical trial infrastructures, which means that better methods for prioritisation are sorely needed. Most agents now enter the clinic with a presumed specific molecular target. Pharma sponsors are often committed to co-development of a companion diagnostic biomarker, which they would also market. Targeted agents are rightly seen as having the potential to improve the speed of drug development and registration, but only if the target is correctly defined and genuinely specific; and if there is a reliable way to select patients. In most cases, one or more of these criteria do not apply (or not yet), so phase II and III development has built-in delays to resolve some of these issues, or is forced to accept one or another compromise in design or efficiency. To move forward, stratified (or ‘precision’) medicine needs new trial designs that are inherently more efficient and can accommodate the biomarker selection or validation uncertainties that are usually present, at least at the outset.

The multi-arm multi-stage (MAMS) design (see *Figure 1*) was developed as a first step toward improving the speed and efficiency of large scale randomised clinical trials (RCTs) in oncology (and other fields). The theoretical basis for this approach was described in 2003 (1) and further developed in 2008 (2). Where there is more than one clinically important question to be addressed (which is commonly the case), a multi-arm trial approach

can simultaneously and systematically test each of these approaches against the current standard of care (the control arm). Secondly, a number of pre-planned interim analyses can (with little statistical penalty) examine accumulating information on activity. Recruitment can be terminated early to comparisons that are showing insufficient activity (suggesting that a meaningful improvement in the primary outcome measure is unlikely). With more agents—particularly targeted agents—in development, clinicians and companies should be seeking strong signals, not marginal ones, of activity and clinical benefit.

The efficiency of the multi-stage element is further enhanced if an earlier (or surrogate) outcome measure can be used as the primary outcome measure at the interim analysis, preferably an event that is on the causal pathway to the primary outcome measure. For example, in a setting where most patients die of their cancer, any advantage in overall survival is likely to be preceded by a benefit in disease progression or progression-free survival (PFS); and without a benefit in PFS it is unlikely that a benefit in overall survival will ensue. Therefore, PFS acts as a useful intermediate outcome measure for overall survival. In this example, a benefit in PFS may not necessarily translate to a benefit in survival, so PFS is not sufficient as the primary outcome measure, i.e., it is not a true surrogate. (For many types of cancer it is possible to identify useful intermediate outcomes even if they are not demonstrably true surrogates for overall survival).

The MAMS approach has been built on foundations of implementability (3). The first MRC FOCUS trial in colorectal cancer (4) demonstrated successful recruitment to a 5-arm trial. Large-scale internationally collaborative testing was done trial an ovarian cancer trial, ICON5 (5). Like FOCUS, the ICON5 trial also had multiple investigational arms, but also introduced a single intermediate lack-of-benefit analysis, making it a multi-arm, two-stage trial.

### **STAMPEDE and the MAMS design**

Further methodological work underpinned the important evolution to MAMS design during the development of the STAMPEDE trial in prostate cancer. Recruitment to this trial in patients initiating 1<sup>st</sup>-line hormonal therapy for prostate cancer was likely to take a number of years so the methodology was extended to allow for multiple interim looks at accumulating data at pre-specified intervals (6).

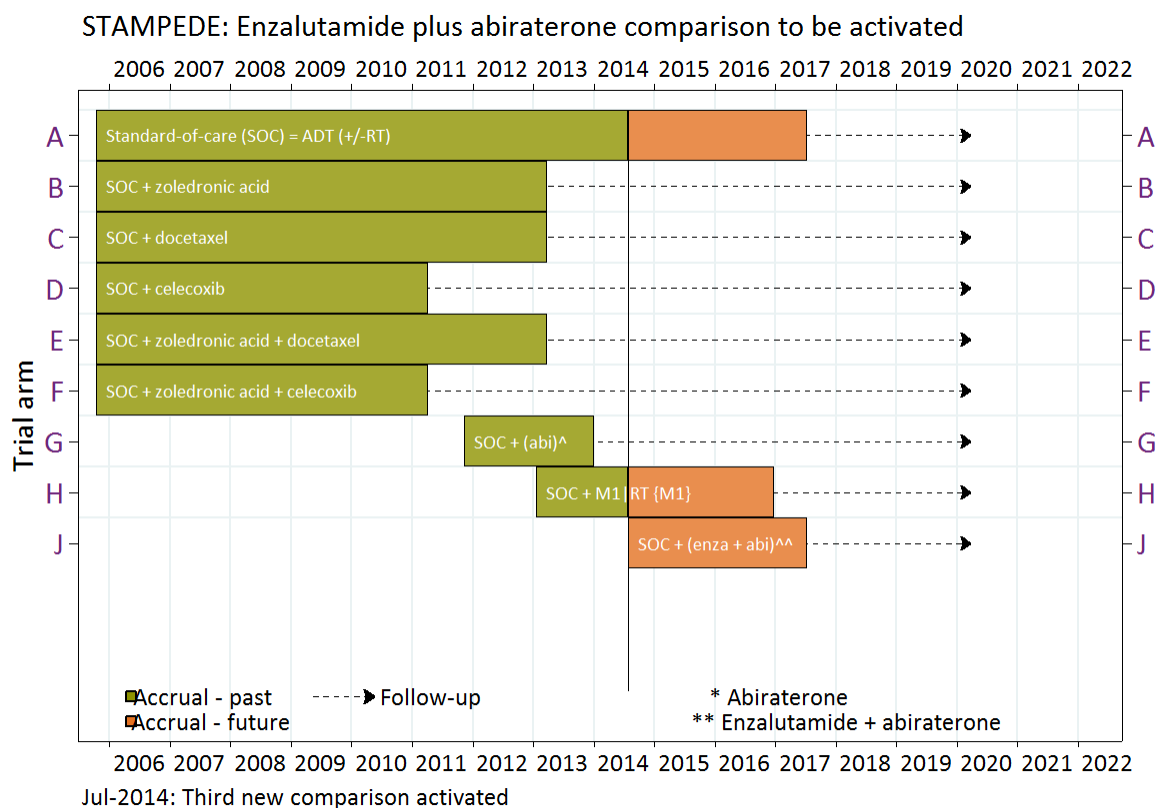
It was of course necessary to familiarise academic collaborators within the NCRI Clinical Studies Group,

NIHR CRN site investigators, funders such as Cancer Research UK, and regulatory agencies with the nature of the MAMS design. A number of implementation challenges were also addressed as the STAMPEDE trial got under way. Eventually, not only was international collaboration established (SAKK, the Swiss Group for Clinical Cancer Research) but, critically, five industry partners joined the trial, each supporting a different arm: Janssen, Astellas, Novartis, Sanofi-Aventis, and Pfizer.

The full implementation of the MAMS approach has proven to be feasible, well received by clinical investigators and potential patients, and acceptable to funders, pharma partners and regulatory authorities. It also proved to have certain additional practical advantages beyond those initially envisaged. During the course of the trial, new developments in prostate cancer therapy introduced both a possible threat to trial completion and a new scientific opportunity. This type of situation is not uncommon in oncology, and our approach to dealing with it is highly generalizable.

RCTs in cancer are long-term undertakings. The approaches that are most worthy of (and ready for) testing in a given setting during the initial development, funding and implementation of a large multi-centre or international trial will usually not be the only ones of such priority during the several years that all such RCTs require for completion. What should researchers do when a seemingly important new therapeutic development arises? What if that development has the potential to interfere with enrollment to an ongoing trial? What if it threatens to make the original trial results less relevant to future clinical practice? For example, while the STAMPEDE trial was looking at three different approaches to treating prostate cancer in the 1<sup>st</sup>-line hormone-naïve setting, a new drug, abiraterone, showed a compelling advantage in survival in the 3<sup>rd</sup>-line setting. There was clear and justified interest from all parties to explore the role of this agent in the 1<sup>st</sup>-line setting.

For the pharma company sponsor, launching such a trial would take time, and the large numbers of patients required (more patients are needed for the required number of events in 1<sup>st</sup>-line treatment) would have necessitated competing directly with ongoing STAMPEDE recruitment (in the UK). Meanwhile, the STAMPEDE investigators did not wish to abandon STAMPEDE, nor to miss the opportunity to learn more about abiraterone in an important disease setting. Therefore, MRC CTU methodologists and the STAMPEDE management team considered whether the trial could be amended to draw the new drug into the STAMPEDE study itself as a new, separate comparison.



**Figure 2** Schema of STAMPEDE.

There were non-trivial practical and statistical considerations that had to be addressed in such a development, but this approach was ultimately taken very successfully, and the new comparison was incorporated into STAMPEDE via protocol amendment. The new research question was able to get underway in the UK dramatically more quickly than could have happened in any other scenario. Recruitment to the new comparison (along with the others remaining) was incredibly rapid (7). As of November 2014, the trial has enrolled 6,000 patients and is the largest ever RCT of treatment for prostate cancer. Data from five different treatment approaches will mature this year and be reported, while a further three other treatment comparisons are well underway. Indeed, this has provoked the introduction of two further new comparisons into the trial, including the newly-activated combination of enzalutamide (Astellas) plus abiraterone (Janssen) with support from both companies.

New arms introduced after trial initiation are tested only in comparison to concurrently randomised patients in the

control arm, as the patients enrolled in later time periods (with a different array of experimental arms available to them) would be expected to differ in significant ways from the original cohort (Figure 2). As a logical corollary of this, should the standard of care (the control arm for this trial) change, this can also be accommodated by amendment and the newly added experimental arms continued with revamped statistical considerations (8).

This newly recognised capability to incorporate new treatment developments within a running MAMS type trial is now viewed as a further key advantage of the MAMS design (9). In the case of STAMPEDE and prostate cancer, we can be confident that, whatever the treatment outcomes, the addition of abiraterone and the other new arms since has guaranteed that the therapeutic approaches tested are amongst those with the highest current level of clinical interest, even 7 or more years after the trial was originally designed. The reported results for at least some of these are likely to make a large impact on the field of advanced prostate cancer therapy.

### **FOCUS4 as a prototype biomarker-stratified 'umbrella' trial**

The STAMPEDE experience was a key influence on the design of the biomarker-stratified FOCUS4 trial programme in colorectal cancer. Specifically, we wished to retain STAMPEDE's desirable features such as the ability to add and drop agents based on planned interim analyses, to provide a running trial platform within which multiple pharma companies can test their agents expeditiously and relatively cheaply against a control arm (but not against each other), and its 'umbrella' nature, that is, its capability of including nearly all patients at a given stage, and opening very widely to expedite accrual. But FOCUS4 was developed with a quite distinct purpose, to make stratified (or 'precision') medicine much more practical.

The rapid progress in detailed molecular characterisation of cancers, together with the ability to use this information to develop specifically targeted candidate therapeutics, has led to a multitude of possible predictive biomarkers and an unprecedented array of potential new treatments. Most research clinicians are confident that a biomarker-stratified approach to trials (and eventually practice) is the most promising way forward. Certain practical issues present themselves, such as biomarker turnaround times and assay validation across multiple labs, but trials organisations are experienced with addressing such requirements [as our team did in the pilot FOCUS3 trial (10)].

However, there remains a challenge that is fundamental to the development of large scale trials of stratified medicine: despite the explosion of promising associations between potential biomarkers and expected targets of novel agents, very few of these associations are mature and fully validated as predictive (not merely prognostic), and thus ready to serve as a sound basis for allocation of treatments. Ideally validation would itself be the result of prospective large scale trials, and there also would be known satisfactory levels of sensitivity and specificity and a high negative predictive value NPV. However, in most cases there will be large gaps in the evidence until quite late in the development of a particular therapeutic agent (or biomarker).

As a result, investigators and pharma companies are faced with a dilemma when considering conventional designs for trials: they must accept one or another type of compromise in their design. One approach is to include all patients with the relevant disease, even though the expected sensitive cohort may make up a small minority. This is not only

expensive in terms of resources and patient numbers, but also inevitably competes directly with any other ongoing broadly inclusive studies. Alternatively, sponsors could restrict the trial to include only patients with the putative biomarker, anticipating that the marker will eventually be validated, and screening many patients for each one who turns out to be eligible. This is also inefficient and leaves unanswered the question of possible drug benefit in biomarker-negative cases.

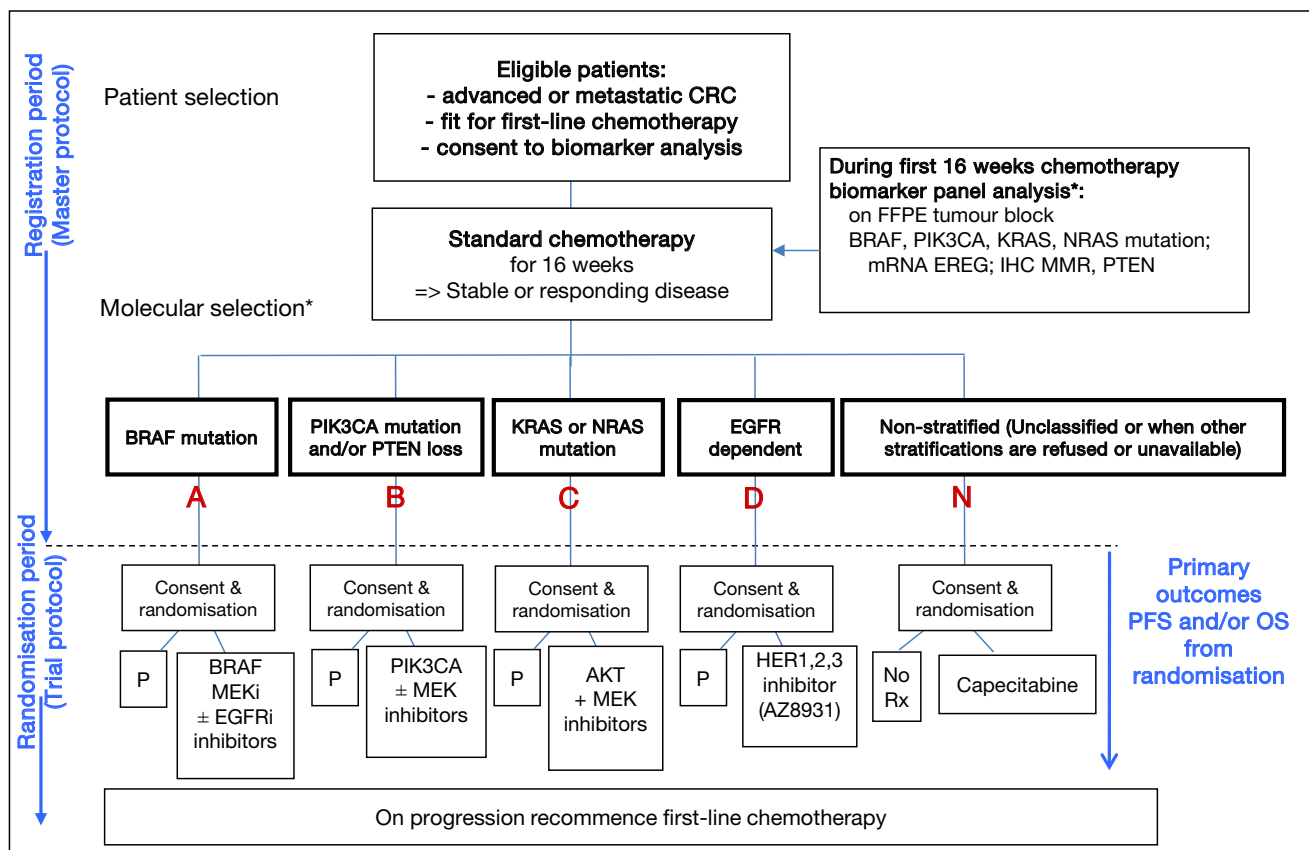
These 'marker by treatment interaction' or 'biomarker stratified' designs both require quite large sample sizes (the included patients or the screened-but-eliminated patients) because they need to size the trial either on the difference between the effect of the treatment in biomarker-positive and -negative patients (an interaction), or on the effect in all patients, which is likely to be modest. Both approaches are potentially highly inefficient and therefore inhibit progress towards identifying the genuine breakthrough therapies from amongst all the available candidates.

The FOCUS4 design was developed to provide a more efficient framework for trials of biomarkers as predictors of response to new agents (or combinations) and to build in the structure for including promising but as yet unvalidated agents and biomarkers, adapting to developing evidence as it arises (11). FOCUS4 is an integrated trial programme of parallel, molecularly stratified, and randomised comparisons of maintenance therapies for patients with advanced or metastatic colorectal cancer after receiving 1<sup>st</sup>-line chemotherapy.

The trial design exploits a 'window of opportunity', maintenance after response to 1<sup>st</sup>-line chemotherapy in advanced or metastatic colorectal cancer. This setting allows us to test for initial signals of clinical efficacy in pre-specified biomarker defined subgroups of targeted novel agent(s) before resistance to standard agents occurs (*Figure 3*). FOCUS4 employs the multi-stage methodology and ability, derived from the STAMPEDE experience, to remove or add agents in order to achieve cost and time efficiencies. It was also set up to be highly adaptable to new biomarker and clinical data as the trial proceeds. The key features and principles of the design are as follows:

- (I) Allows for efficient screening and inclusion of nearly all patients in a particular setting (1<sup>st</sup>-line colorectal cancer in this case) into a single trial that includes multiple biomarker-defined cohorts;
- (II) Allows for adaptive changes in response to relevant new developments, including addition of new agents;

## FOCUS4 Trial Schema



\* The molecular cohorts are arranged in a hierarchy from left to right. For example a patient with both a PIK3CA mutation and a KRAS mutation will be classified into the PIK3CA mutation cohort.

**Figure 3** FOCUS4 schema.

(III) Allows for the inclusion of new biomarker cohorts as developing evidence for them warrants.

FOCUS4 systematically employs the multi-stage element of the MAMS design, with the potential to employ the multi-arm element when appropriate. In current parlance, it has an “umbrella design” i.e., a stratified trial design with nested, virtually separate, parallel RCTs for biomarker-defined subgroups of patients, each with its own appropriate control. Each of these is actually a separate randomised phase II/III trial that can stop early for lack of benefit or continue to its final stages.

Because of the ‘umbrella’ design, promising predictive biomarkers (often promising because they are closely linked to the molecular target of a new drug) can be included prior to full validation and preliminarily tested for association with drug benefit. As and when new data from within or

outside of the FOCUS4 trial suggests that the biomarker selection needs to be refined or adjusted, the appropriate decision can be made as to whether the change can be accommodated with adjustment of sample size or, in more extreme situations, the change made and the 4-stage design re-started in that cohort. Even if biomarker refinement adds or removes patients from other biomarker cohorts, the trial as a whole can continue.

The inclusion of biomarker-matched control patients for each cohort separately is a key feature because it allows us to separate prognostic biomarker effects from those that are predictive of a response to the particular treatment. Several other ‘umbrella’-type trial designs that have been developed have not had this feature and they are not well suited to distinguishing whether any benefit seen is genuinely related to the biomarker selection criteria employed.

**Table 1** Summary of generic operating characteristics and timelines for each cohort

Molecular cohort	Randomised allocation ratio	Phase	Outcome/stage	Target HR	Max number of events required: total [control arm]	Estimated cumulative analysis time (months)	Max number of patients required
BRAF mutation	2:1	2	PFS/I	0.5	41 [16]	20.4	61
			PFS/II	0.5	76 [28]	32.5	97
		3	PFS/III	0.5	118 [42]	46.5	139
			OS/IV (potential)	0.65	217 [79]	100.4	301
PIK3CA mutation and/or PTEN loss	2:1	2	PFS/I	0.65	107 [40]	17.0	170
			PFS/II	0.65	197 [71]	26.5	264
		3	PFS/III	0.65	303 [107]	37.2	373
			OS/IV (potential)	0.7	289 [109]	54.6	546
KRAS or NRAS mutation	2:1	2	PFS/I	0.65	109 [41]	16.1	177
			PFS/II	0.65	198 [72]	22.8	273
		3	PFS/III	0.65	302 [107]	31.4	378
			OS/IV (potential)	0.7	287 [109]	50.6	574
EGFR dependent	2:1	2	PFS/I	0.65	109 [41]	20.0	180
			PFS/II	0.65	198 [72]	30.6	275
		3	PFS/III	0.65	301 [107]	42.3	381
			OS/IV (potential)	0.7	289 [109]	60.8	547

**Table 2** Anticipated recruitment for screening and randomisation across cohorts

Molecular cohort	Prevalence (%)	70 screened patients per month		
		Randomised patients (months)	Total over 4 years	Total over 5 years
Total	100	32	1,536	1,920
BRAF mutation	8	2.6	125	156
PIK3CA mutation and/or PTEN loss	30	9.6	461	576
KRAS or NRAS mutation	33	10.6	509	636
EGFR dependent	27	8.6	413	516
Unclassified	2	0.6	28	36

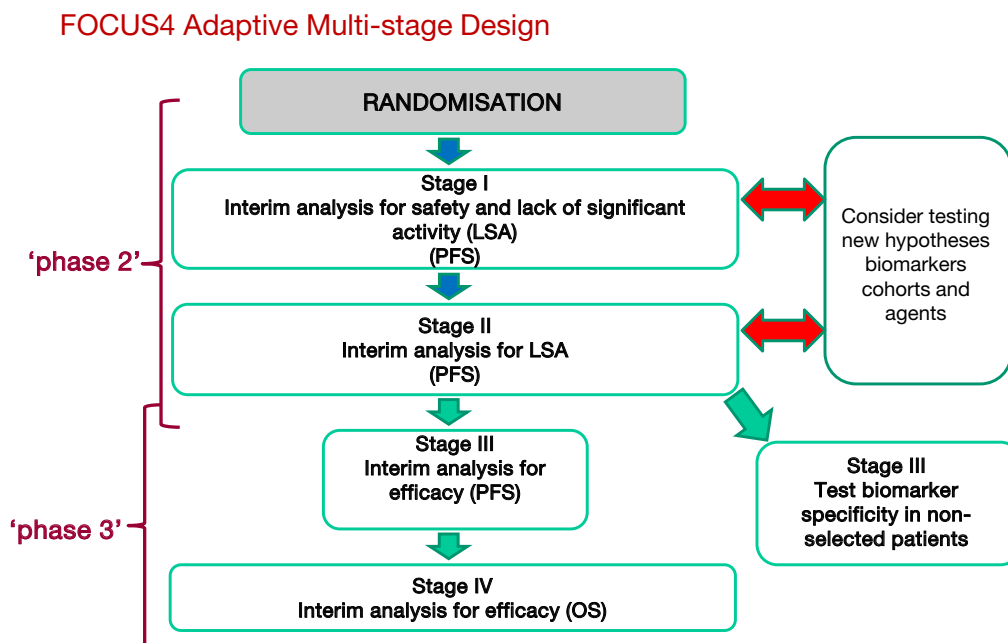
Each biomarker/treatment comparison has 4 stages: two lack of activity/signal-seeking stages (effectively a two-stage phase II trial), and two efficacy stages, the first a phase III trial with a PFS primary endpoint, and if pre-specified criteria are met, OS as the definitive endpoint for the final stage. A general plan as a guide to sample sizes and number of events required is presented in *Table 1*. The expected frequency of the various cohorts is given in *Table 2*.

Although in certain settings PFS benefit may support licensing approval, especially perhaps in small (rare) biomarker cohorts, the final OS stage is a feature of the design that is available as needed to serve as the basis for registration within some biomarker cohorts.

Recruitment will be stopped early to any treatments that

do not meet the progressively more stringent criteria at the three interim analyses. However, that does not mean that subsequent patients in those biomarker subgroups will be left without a research option within the trial. Instead, the FOCUS4 design will accommodate to a lack of PFS benefit from one of the test regimens in one of several ways, depending on accumulating data and the status of various other cohorts:

- (I) If there are available alternative high priority novel agents with mechanisms that match the biomarker subgroup, the new agent(s) will be substituted and will be tested, starting again at the first stage of a new multi-stage analysis plan.
- (II) If a drug being tested within FOCUS4 for a different



**Figure 4** ‘Stages’ of FOCUS4 design.

biomarker subgroup has progressed through its first two interim analyses (which indicates roughly 90% likelihood of finding activity at the level sought), then patients from other subgroups may be tested with that drug—to ascertain whether the biomarker selection is or is not definitive in predicting sensitivity. This approach—testing initially for activity in an enriched population expressing the putative biomarker, and only if adequate activity is found, then look for any activity in patients with different biomarker patterns—is more efficient than testing the agent a priori in unselected patients (see *Figure 4*).

- (III) When neither of the above situations applies, FOCUS4 has one randomised comparison that is not biomarker selected and will run throughout the length of the trial. This was incorporated into the design in order to have a research option available when biomarker-driven cohorts are closed [either temporarily, as between analysis stages, pending availability of a replacement agent for that cohort, etc., or permanently when neither (I) or (II) above applies], and for individual potential subjects who may be unwilling to travel to the major cancer centres equipped to administer new agents/combinations at a still relatively early stage of testing.

This biomarker non-specific arm addresses a standard but as yet unresolved question about the utility of maintenance fluoropyrimidine chemotherapy during a treatment break from more intensive combination chemotherapy. This remains an important clinical issue, with an incomplete body of evidence from other trials.

Thus FOCUS4 arm N is a safety valve of a sort, assuring that recruitment momentum can be sustained despite the many interim analysis points of the trial and the resulting unpredictable trial modifications, and that the trial can accept virtually any patient who does not progress through standard 1<sup>st</sup>-line chemotherapy for colorectal cancer. This ability to recruit ‘everyone’ makes the trial attractive to research teams, and to potential subjects, who do not have to face being asked to consider a trial and then be informed that they cannot join it because of their biomarker profile. It also helps assure the recruitment momentum that is attractive to pharma collaborators. Even if this particular clinical question had not been one of great interest (or if this question is resolved mid-trial), it would/will remain of strategic importance to have another research question not dependent on biomarker selection to serve these logistical needs.

Considered as a whole, the FOCUS4 trial and design have a number of advantages (*Table 3*) and are adaptive in three ways:



**Table 3** Advantages of FOCUS4 design

---

Allows for trial when biomarkers are incompletely characterised and/or not fully validated
'Umbrella' structure allows for efficient inclusion of less common biomarker cohorts
Tests more than one treatment at the same time, each against its own appropriate control
But does not test cohorts/agents against each other
Molecularly enriched cohorts maximise likelihood of detecting promising new treatments
Employs ambitious hazard ratio targets (accepting risk of rejecting minimally effective agents)
Multi-stage design analyses at pre-specified time points for early detection of insufficient activity
Initial emphasis is phase II in intention (signal seeking)
But can continue efficiently (seamlessly) into phase III
Efficient design for ascertaining the specificity of any positive results in relation to the biomarker selection used
Adaptive: allows for efficient incorporation of new information on drugs or biomarkers into a large ongoing trial

---

- (I) The predictive biomarkers can be refined as either trial data or external data evolve;
- (II) New candidate treatments (drugs or combinations) can be introduced, either in an additional newly defined biomarker cohort or in cohorts for whom the initial candidate agent was not found to have sufficient activity to proceed to the next stage;
- (III) For treatments that do appear to be showing sufficient activity (against control) in the enriched (biomarker-selected) patients, there is a framework for preliminarily testing the predictive association by opening cohorts of patients without the biomarker ('off-target' efficacy).

The trial is a major effort of course, and the multiple collaborations and agreements necessary required a great deal more effort than the arrangements for most trials. We have had to assemble a steering group and an Independent Data Monitoring Committee to deal with the frequency of, and rapid turn-around needed for, multiple interim analyses on one arm or another (and the protocol amendments required).

We took the strategic decision that there should be specific chief investigator for each of the comparisons as well as co-Chief Investigators (Tim Maughan and Richard Wilson) for the programme. Further key collaborators have developed important translational research projects that will be tightly integrated. We were fortunate too that two major trial funders in the UK (Cancer Research UK the NIHR Efficacy and Mechanism Evaluation Programme) were interested in the novel design and willing to support its assembly. Counterbalancing these complexities is a set of compelling benefits, most importantly the opportunity to test simultaneously the use of a number of biomarkers and

the benefit of a number of agents and to adapt the trial to emerging developments in the field.

### Acknowledgements

None.

### Footnote

*Conflicts of Interest:* The author has no conflicts of interest to declare.

### References

1. Royston P, Parmar MK, Qian W. Novel designs for multi-arm clinical trials with survival outcomes with an application in ovarian cancer. *Stat Med* 2003;22:2239-56.
2. Parmar MK, Barthel FM, Sydes M, et al. Speeding up the evaluation of new agents in cancer. *J Natl Cancer Inst* 2008;100:1204-14.
3. Seymour MT, Maughan TS, Ledermann JA, et al. Different strategies of sequential and combination chemotherapy for patients with poor prognosis advanced colorectal cancer (MRC FOCUS): a randomised controlled trial. *Lancet* 2007;370:143-52.
4. Barthel FM, Parmar MK, Royston P. How do multi-stage, multi-arm trials compare to the traditional two-arm parallel group design--a reanalysis of 4 trials. *Trials* 2009;10:21.
5. Bookman MA, Brady MF, McGuire WP, et al. Evaluation of new platinum-based treatment regimens in advanced-stage ovarian cancer: a Phase III Trial of the Gynecologic Cancer Intergroup. *J Clin Oncol* 2009;27:1419-25.

6. Sydes MR, Parmar MK, James ND, et al. Issues in applying multi-arm multi-stage methodology to a clinical trial in prostate cancer: the MRC STAMPEDE trial. *Trials* 2009;10:39.
7. Sydes MR, Parmar MK, Mason MD, et al. Flexible trial design in practice - stopping arms for lack-of-benefit and adding research arms mid-trial in STAMPEDE: a multi-arm multi-stage randomized controlled trial. *Trials* 2012;13:168.
8. Attard G, Sydes MR, Mason MD, et al. Combining Enzalutamide with Abiraterone, Prednisone, and Androgen Deprivation Therapy in the STAMPEDE Trial. *Eur Urol* 2014;66:799-802.
9. Sydes MR, Parmar MK. The need for a cultural shift from two-arm to multi-arm RCTs. *Trials* 2013;14:O3.
10. Maughan TS, Meade AM, Adams RA, et al. A feasibility study testing four hypotheses with phase II outcomes in advanced colorectal cancer (MRC FOCUS3): a model for randomised controlled trials in the era of personalised medicine? *Br J Cancer* 2014;110:2178-86.
11. Kaplan R, Maughan T, Crook A, et al. Evaluating many treatments and biomarkers in oncology: a new design. *J Clin Oncol* 2013;31:4562-8.

**Cite this article as:** Kaplan R. The FOCUS4 design for biomarker stratified trials. *Chin Clin Oncol* 2015;4(3):35. doi: 10.3978/j.issn.2304-3865.2015.02.03