Biomarker-guided trials: challenges in practice

Antoniou, M[1]; Kolamunnage-Dona, R[1]; Wason, J[2]; Bathia, R[3]; Billingham, C[4]; Bliss, JM[5]; Brown, LC[3]; Gillman, A[5]; Paul, J[6]; Jorgensen, AL[1]

1. *University of Liverpool, UK*
2. *Newcastle University and MRC Biostatistics Unit, Cambridge, UK*
3. *MRC Clinical Trials Unit at UCL, UK*
4. *University of Birmingham, UK*
5. *Institute of Cancer Research, London, UK*
6. *University of Glasgow, UK*

Corresponding author: Professor Andrea L Jorgensen, Department of Biostatistics, University of Liverpool, Block F Waterhouse Building, 1-5 Brownlow Street, Liverpool L69 3GS. e-mail: a.l.jorgensen@liverpool.ac.uk

Abstract

Biomarker-guided trials have drawn considerable attention as they promise to lead to improvements in the benefit-risk ratio of treatments and enhanced opportunities for drug development. A variety of such designs have been proposed in the literature, many of which have been adopted in practice.

Implementing such trial designs in practice can be challenging, and identifying those challenges was the main objective of a workshop organised by the MRC Hubs for Trials Methodology Research Network's Stratified Medicine Working Group in

March 2017. Participants reflected on completed and ongoing biomarker-guided trials to identify the practical challenges encountered. Here, the key challenges identified during the workshop including those related to funding, ethical and regulatory issues, recruitment, monitoring of samples and laboratories, biomarker assessment, and data sharing and resources, are discussed.

Despite the complexities often associated with biomarker-guided trials, the workshop concluded that they can play an important role in advancing the field of personalized medicine. Therefore, it is important that the practical challenges surrounding their implementation are acknowledged and addressed.

Keywords: clinical trial, biomarker, challenges, personalized medicine

Introduction

Clinical trials are essential for testing the safety and efficacy of new treatments. Increasingly, biomarkers are becoming an integral part of clinical trials as they are considered key tools in the identification of patient sub-populations most likely to benefit or conversely to incur adverse reactions from a given treatment{Landeck, 2016 #1;Bailey, 2014 #2;La Thangue, 2011 #20;Vargas, 2016 #21}. Hence, so-called biomarker-guided trial designs are pivotal in advancing the field of personalized medicine which aims to give 'the right treatment to the right patient, at the right dose at the right time' (1). Consequently several biomarker-guided trial designs which test the effectiveness of a biomarker-guided approach to treatment have been proposed in the literature, some of which have been adopted in practice.

Detailed reviews of biomarker-guided designs have been published (3-6) and are also available via an online tool "BiGTeD" (http://www.bigted.org/).

A one-day workshop organised by the MRC Hubs for Trials Methodology Research Network's Stratified Medicine Working Group (SMWG) was held in London in March 2017. The aim was to identify and explore the key practical challenges arising when conducting a biomarker-guided clinical trial. The workshop brought together 25 participants with practical experience in conducting biomarker-guided trials from various disciplines including statisticians, trial managers, information systems specialists and clinicians. This workshop was motivated by feedback from trialists and previous literature (3, 4) suggesting that there are substantial challenges associated with undertaking trials adopting these types of designs.

Specific trials were utilised as exemplars to aid discussion and these are the focus of the first part of this paper. The second part provides an overview of the practical challenges raised at the workshop and identified from delegates' experiences, together with some of our own reflections on those issues from our methodology reviews and simulation studies (3, 4, 7). Issues considered include funding, ethical and regulatory issues, recruitment, monitoring of samples and laboratories, biomarker assessment, data sharing, and resourcing. A summary table is also provided of each trial's key characteristics with examples of some of the challenges they faced (Table 1).

Biomarker-guided trials used as exemplars

The majority of trials discussed at the workshop are oncology trials simply because oncology dominates the field of personalised medicine. Many of the challenges identified apply equally to trials in other clinical areas.

i) The National Lung Matrix Trial (NLMT; ongoing trial) **(8):** This is a phase II non-randomized umbrella trial consisting of multiple single arm trials within one protocol. The aim of the trial is to investigate a range of new treatments hypothesized to be of benefit to specific molecularly-defined cohorts of patients with advanced non-small cell lung cancer (NSCLC) and for whom surgery and radiotherapy are not deemed appropriate treatments.

NLMT runs alongside the Cancer Research UK Stratified Medicine Programme (SMP2), where a next generation sequencing 28 gene panel test is used to assess the genetic profile of trial participants which then determines which single arm trial (strata), and hence drug, they are assigned to. The trial adopts a Bayesian adaptive design with an interim analysis at 15 patients for each strata and final analysis of a target group of 30 patients per strata. The trial was designed to evaluate a common set of outcome measures with primary outcome measures chosen specifically for each treatment arm. A clinically relevant signal of efficacy is defined: for cytostatic agents as median progression-free survival greater than 3 months; for other agents as rates of objective response{Eisenhauer, 2009 #19} or durable clinical benefit (defined as remaining free of disease progression at a CT or MRI scan approximately 24 weeks after starting treatment, or thereafter) with a

critical cut-off greater than 30% for single agent and 40% for combination therapy arms.

ii) Phase II trial of olaparib in patients with advanced castration resistant prostate cancer (TOPARP) (ongoing at time of workshop, now closed to recruitment) (9): This is an open label, phase II, single arm, 2 part adaptive design trial for biomarker-driven selection based on response rate. It aims to evaluate the anti-tumour activity of the Poly (ADP-ribose) polymerase (PARP) inhibitor, olaparib, in metastatic castration resistant prostate cancer (mCRPC) and to identify molecular signatures for PARP inhibitor sensitivity with a pre-planned analysis to identify a biomarker-defined sensitive subgroup. In the first part unselected (i.e. without biomarker-guided patient selection) mCRPC patients are all treated with olaparib. If during the first part the response rate is high (i.e., ≥ 50% responding) the trial will close and a randomized placebo controlled clinical trial to evaluate the efficacy and safety of olaparib in these unselected mCRPC patients is undertaken. If the response rate is low (i.e. response rate < 10%), the trial is stopped. If in the intermediate range (10-50% responding), potential biomarkers of response are investigated and if a potential biomarker is identified, with those positive for the biomarker having a high response rate (≥50%), the trial continues to the second part where only biomarker selected patients are included.

iii) Adaptive multi-arm phase II trial of maintenance targeted therapy after chemotherapy in metastatic urothelial cancer (ATLANTIS) (ongoing) (10): This is an adaptive multi-arm randomized phase II trial which aims to explore whether maintenance targeted therapy after chemotherapy, with treatment randomisation

based on biomarker profile, delays time to progression and increases overall survival for patients with advanced urothelial cancer. The initially planned biomarker is androgen receptor status with patients who are androgen receptor positive randomised between enzalutamide and placebo. The "adaptive" element of ATLANTIS is the ability to add comparisons in other biomarker selected subgroups (for example a comparison of rucaparib v placebo is planned in patients who have BRCA1 or BRCA2 mutations, either as a somatic or germline event, or with evidence of homologous recombination deficiency)

iv) PRIMUS001 (ongoing) (11): This is an adaptive phase II trial, with biomarker evaluation integrated into the trial which aims to assess the efficacy of FOLFOX-A (FOLFOX and nab-paclitaxel) when compared to AG (nab-paclitaxel and gemcitabine) in patients with metastatic pancreatic cancer, both in a biomarker-positive group and in biomarker-unselected patients. PRIMUS001 will determine whether there is a benefit from FOLFOX-A compared to AG, and if there is a benefit whether this is in all patients or in biomarker +ive patients only. As the study proceeds there are a number of interim analyses following which subsequent recruitment may be restricted to biomarker +ve patients if there is no evidence of benefit of FOLFOX-A compared to AG in biomarker unselected patients.

v) SALONICA (planned trial): This is a stratified adaptive trial in ovarian cancer aiming not only to detect the key genomic determinants of response and resistance to neoadjuvant platinum-base chemotherapy in high-grade serous ovarian cancer but also to identify and validate putative biomarkers as well as test

several novel drugs and corresponding putative biomarkers in women with poor response to neoadjuvant platinum chemotherapy through a phase II trial platform. SALONICA is initially based on a sequence of single-arm biomarker unselected phase II designs, but as information on the mutational changes and associated biomarkers in ovarian cancer accumulates the ambition is to move to a design based on Bayesian Adaptive Randomisation (BAR).

vi) TASTER (planned trial): This trial aims to identify predictors of response to novel combination therapies in Chronic Myeloid Leukaemia (CML) patients who do not respond to tyrosine kinase inhibitor therapy. Both in vivo models of drug response and clinical data will be used to identify molecular signatures of stem cell resistance and build and validate predictive models of drug response from which the best treatment for a patient can be selected. The success of the predictive model will be assessed in standard single arm phase II design for each candidate novel combination.

vii) POETIC (Peri-Operative Endocrine Therapy for Individualizing Care) (ongoing trial) (12, 13): This is a randomized, multicentre phase III trial which aims to investigate whether having perioperative aromatase inhibitor (AI) therapy for postmenopausal women with ER+/PgR+ positive invasive breast cancer is more effective than having standard care alone. 4,476 patients were recruited from 130 UK centres. Patients received either AI therapy for 4 weeks (two weeks before and two weeks after surgery) or no peri-operative AI therapy. Whilst ER is a well established biomarker it is not usually used to direct therapy so early in the patient

pathway, thus new procedures had to be established for the trial to ensure its measurement was available at the time of diagnosis based on a core biopsy.

viii) FOCUS4 trial (ongoing trial) (14): This is an umbrella clinical trial consisting of parallel, molecularly stratified randomized comparisons in patients with metastatic colorectal cancer (mCRC). Patients with newly diagnosed mCRC are registered into the trial and commence their standard first line chemotherapy which typically lasts for approximately 16 weeks. During this time, a sample of their tumour is sent away to one of two dedicated FOCUS4 laboratories who perform genomic and molecular tests on the tumour. This enables stratification of the patients into one of a number of pre-specified molecular subgroups (called cohorts). Patients are then offered entry into a randomized trial (called comparison) testing a specific targeted therapy for their subtype of cancer. All these comparisons are randomized and controlled and wherever possible use a placebo in the control group.

ix) EU-PACT trial (completed trial) (15): This was a pragmatic, single-blind, randomised controlled trial to determine whether genotype-guided dosing of the anticoagulant warfarin is superior to standard dosing. Patients commencing warfarin were randomised to one of two trial arms. Those randomised to the genotype-guided dosing arm had their genotype tested at three genetic variants using a point of care test, with results available within two hours. Their genotype was fed into a computer based loading dose algorithm, together with demographic and clinical information, and a personalised loading dose recommended for the first three days. Similar information was then fed into a maintenance dose algorithm to

determine dose on days 4 and 5 of treatment. From day 6 dosing was according to standard clinical care. Those randomised to the control arm had their loading and subsequent doses calculated according to standard approaches, with no reference to genotype. All patients were followed up for three months and their anticoagulation control assessed.

Challenges

Funding issues

Biomarker-guided trials often have a complex design – both scientifically and logistically and it is therefore not surprising that the resources required are typically considerably higher than for trials with more simple designs. Nonetheless, funders show substantial enthusiasm for supporting biomarker-guided trials, since it is recognized that despite increased costs the trial may well be more efficient in demonstrating patient benefit. When considering the additional resources required, the increased administrative burden is a major factor: for instance, in umbrella type designs necessary documentation and multiple approvals need to be repeated for each treatment group of the trial. How those amendments are handled (e.g. the addition of a new trial group), can depend on cost. For example, it is typical for changes that don't require additional funding from charitable or public bodies (generally where funding is provided by a pharmaceutical partner) to be implemented quickly without additional approvals, but if the amendment is likely to require additional funding support then it is necessary for it to go through the more classic route of peer-review and research grant approval.

Further, despite the attractive flexibility they bring, by virtue of their design the overall costs of adaptive trials in particular are difficult to predict at the outset due to the uncertainty surrounding their future direction, for example the number of additional/discontinued groups and final sample size. In addition, with science evolving at such a fast pace, biomarker assessment costs may change with changing technologies. Open communication between those involved in planning such a trial and funders is important from the outset to determine the best way to handle applying for funding. Such open communication will also help inform funders on the implications of using such designs for their funding streams.

So that overall costs can be considered, and to avoid triggering further full processes for committee approval with the addition of each new trial group, it can be advantageous for applicants to provide details on potential additions at the outset to allow funders to forecast and earmark the foreseeable additional budget and provide approval in principle. A similar agreement is already in place with CRUK for NLMT. So, a researcher putting an application for an umbrella trial, for example to include initial trial groups A to D, would be required to also estimate how much it would cost to add groups E, F, and G at time points X,Y and Z. Understandably, providing such predictions of future costs can be difficult as it requires knowledge not only of the approximate size of the trial groups to be added (or indeed removed) within those changes, but also the time point at which they will be added and the approximate end date.

Further, including additional forecasted costs could easily make a trial unattractive to funders, with projected total costs for a large trial using up the entire

budget for a funding call. Funders will always be faced with many competing funding requests, many of which will have simpler and easier to understand designs and more transparent budgets.

It is anticipated however that once many of the currently ongoing trials are completed, there will be a better understanding of the value for money offered by such trial designs. However it is important to note that this could be misleading in itself since it is widely recognised in the trial community that many such trials may have been significantly under-resourced. Quite often, it is the Clinical Trial Units (CTU) costs (e.g. trial management, trial monitoring, statistical analysis and oversight) that are compromised.

One possible model is to fund the molecular screening platform as a separate venture from the trial itself and run them as two interrelated studies. This can be seen in NLMT where the Stratified Medicine Programme 2 (SMP2) provides a comprehensive screening programme funded by Cancer Research UK in collaboration with pharmaceutical partners and the NLMT is funded as a separate Cancer Research UK trial grant. SMP2 provides the patients for NLMT so the success of the trial is entirely predicated on the success of SMP2 and clearly close interaction between the two separate projects is essential. Such a funding model can be appropriate if the stratifying biomarkers involved are novel and outside of routine testing and provides transparency in terms of the costs for the two key major elements in such a trial.

FOCUS4 provides an example of how exploring alternative funding arrangements led them to successfully securing funding for their trial. Joint funding was applied for between CRUK and the Efficacy and Mechanism Evaluation (MRC/NIHR EME) Programme. Every time a new cohort is added an EME Sub Board meeting is held with representatives from CRUK and EME. A scientific rationale and funding model has to be presented by the investigators to this Sub Board for scrutiny. This approach has worked well and could be a viable option for similar trials as long as the funding bodies are encouraged by the efficiencies and opportunities of joint long-term commitments.

There has also been some confusion amongst researchers in the UK about who should fund the additional biomarker tests within a trial. It has previously been suggested that this is a National Health Service (NHS) cost since it is used to direct treatment, however since the test is often unavailable on the NHS, it could be considered to be a research cost. Another viewpoint is that, in the case of a test not yet implemented in practice, if the cost of the test in a research setting exceeds the hypothetical cost of using the test in routine practice, the additional cost should be covered by research funding, with the hypothetical costs associated with using the test in practice, being classed as (potentially excess) treatment costs. This situation may be slowly changing, however, as we move into an era where more biomarker tests are routinely undertaken in practice.

Finally, an additional funding issue relates to whether the trial uses previously untested biomarkers or more established and validated ones; the former may incur additional costs for the development, validation and standardization of

appropriate assays, and delays in the expected start date. Further, issues with sample quality can cause problems for recruitment in situations where results are required with a tight turnaround, as can problems with the assay e.g. its sensitivity.

In summary, detailed and early planning with clear communication between researchers and funders is vitally important to ensure that future trials can be fairly considered and appropriately funded. There is also room for learning, with those with practical experience of such trials sharing their knowledge and experiences with funding bodies as well as funding bodies, with their broader oversight across a spectrum of trials that they fund, sharing the same with researchers. These trials can appear overwhelming if viewed within the classical approvals paradigm but are not as complicated as is often believed. With some designs, they can be considered as a collection of individual separate trials with some additional biomarker analyses. If their benefits and limitations are communicated effectively, then they should be embraced rather than feared.

Ethical and Regulatory Issues

A key issue here is the different ways in which regulatory bodies choose to classify a biomarker-guided trial. For instance, there is an expectation that when adding a new Investigational Medicinal Product (IMP) to an umbrella trial there should also be a new CTA (Clinical Trial Authorization), which may not necessarily be required. Consideration needs to be given to the subtleties of adding a new IMP, for example if it comes from a different class of drugs than existing IMPs and with a different safety profile and changes the scientific intent of the trial a new CTA may

be entirely appropriate, whilst unnecessary with more similar IMPs. It may also be believed that from a commercial perspective the trial will be testing, developing and marketing a companion diagnostic alongside the therapeutic, which is not always the case. Early discussion with the competent authority is strongly advised.

Although the general consensus is that research ethics committees view these types of trials very positively, many ongoing administrative issues need to be addressed. Whilst an ethics committee might give overall ethics approval at the outset, it is often not clear how the addition of new trial groups will be approved later. Depending on local practice, amendments may not be reviewed, discussed and approved by a sub-committee; or may even come through simply as a chairman's action. Consequently, the trial documents are perhaps not checked in the same way as the original application and the amendments may not receive the same level of scrutiny. In addition, there is inconsistency in what documentation ethics committees request for amendment approval, with some requesting a new submission and others seeking a major amendment. It is important that a collaborative relationship is maintained with the Health Research Authority (HRA) to ensure that administrative systems, paperwork, and version control are adapted to adequately deal with these types of amendments. Researchers with experience of running such trials are well placed to advise in this regard.

Similar collaborative relationships also need to be maintained with the relevant regulatory authority (e.g. the Medicines and Healthcare products Regulatory Agency (MHRA) or EMA or FDA). For example, the name of a trial's CTA is based on the initial treatment drugs included in the trial; however, these

may not be part of the trial for the whole duration of the trial (e.g. due to ineffective treatment groups being dropped and other promising ones being added) which can lead to confusion in terms of terminology.

From the perspective of patients, some issues need to be considered relating to the informed consent process. There are examples of having to consent patients into the trial on the same day of diagnosis, for example so that a sample can be sent immediately for biomarker testing to avoid delaying treatment down the line, which clearly requires both careful and appropriate communication. Another issue could arise, particularly in oncology trials, due to the possibility that biomarker screening might fail requiring a second biopsy. Obtaining a second biopsy can be painful, have associated risks and be difficult be obtain if patients are not well enough. In such cases, it may or may not be appropriate to include a trial option for non-stratified patients (including those with failed biopsies), particularly if biomarker screening is invasive or has a high failure rate.

Effective communication with patients is also fundamental to ensure a clear understanding of the purpose of biomarker trials, and whilst they are often about targeting treatments to patients most likely to benefit, they can also be about trying to avoid treatments in patients who are unlikely to benefit from them. This may aid acceptance by those not being offered an experimental treatment based on their biomarker profile. Whilst on the surface personalizing treatment may sound like the optimal solution, it should not be communicated as if a treatment will definitely work in a patient with given biomarker status, but rather is an approach that will mean it is potentially more likely to work. It is also essential that patients

understand that being screened for a biomarker does not guarantee they will be eligible for the trial, since they will often have to meet additional eligibility criteria. Further, there may be a delay in meeting eligibility criteria such that the trial is closed before the patient can be recruited.

An additional ethical challenge can arise in trials, where genetic markers are being assessed, and susceptibility to certain other diseases are uncovered – so-called 'incidental' findings, and this is a subject of much debate(16, 17). From the patient's perspective, in theory this issue can be covered in the informed consent process by allowing them to opt-in or opt-out of information on incidental findings. In reality however, the issue is far more complex since making a truly informed decision would require the patient to have an extensive amount of specialist genetic counselling for numerous conditions unrelated to the primary reason for the genetic test. Further, it can pose a moral dilemma to those involved in conducting the trial. Whilst there are clear advantages arising from incidental findings which can be actioned medically, there is a risk of false positive findings, and knowledge of future disease risk and the anxiety it brings can do more harm than good in asymptomatic patients(17). Additionally, from the patients' perspective, they can often mistakenly assume that having certain mutations in their tumour means an increased risk of disease in relatives. Hence, more careful communication is needed in order to clarify the difference between mutations in a tumour and germline mutations, and which type they are being tested for.

In summary, several ethical and regulatory challenges can arise ranging from a lack of consistency surrounding administrative procedures to issues relating to

communications with patients. It is essential that accurate information about biomarker-guided trials is communicated to all relevant stakeholders so that they are aware of the characteristics and advantages of such trials.

Recruitment

Uncertainty in recruitment rates, especially in trials that include rare biomarker groups, can be a major dilemma. The prediction of recruitment rate into umbrella trials can be difficult due to several factors. One of these factors is uncertainty surrounding the estimated prevalence of each biomarker since this might not be accurately known at the design stage. The uncertainty is greater in the case of trials that evaluate multiple biomarkers as overlapping groups can occur (i.e. a single patient positive for multiple biomarkers). Another contributing factor is that the failure rate of laboratory diagnostic biopsies in the technology hubs is difficult to predict. The funders and sponsors regularly question whether the achieved recruitment rate is close to that projected. Recalculations and protocol amendments may be required, which can often be more complex for biomarker-guided trials than for a traditional trial. Hence, a more flexible methodology is needed for predicting recruitment rate for these trials. Indeed, a more sophisticated statistical approach to prediction that incorporates the uncertainties could be considered in order to provide a realistic range for expected recruitment into each biomarker group.

Recruitment issues can be patient related or researcher related. From the patients' perspective, if there is considerable time between a patient undergoing

molecular screening and being approached about a treatment trial they may be fatigued or experiencing toxicity symptoms after first line treatment, or their disease may have progressed, or they may simply not be interested in the new drug and would like to take a break from treatment. In addition, having complex tissue sampling (mandatory fresh biopsies) is always a challenge for recruitment since some patients would prefer to not have such invasive testing, and in addition due to the complexity sampling may take some time in which case a patient's status and ability to participate may have changed.

From the researchers' perspective, slow trial set up due to the necessary sample collection and processing procedures that need to be established not only delays recruitment but might also lead to study sites losing their enthusiasm. In turn, this may affect the motivation of commercial partners to get involved. Further problems can arise when it is difficult to predict recruitment timelines, as seen in the TOPARP trial. It was difficult to accurately predict recruitment as there was a "lead site" effect at the Chief Investigator's site. Due to a change in the formulation of the novel agent it was necessary to initially only recruit patients at the lead site for safety reasons. Complex sampling collection and processing requirements took external sites longer to establish and as a consequence the management of collaborators' expectations (funders, sites, investigator and commercial partners) was challenging. Additionally a higher screening failure rate was noted at external sites, potentially due to the population of patients seen. Therefore, screening activity at site was increased along with activity at the central lab/CTU. Due to the time to deliver biomarker results to sites, sites tested patients for the biomarker

earlier in the patient pathway than anticipated leading to a pool of biomarker patients waiting to be eligible for the trial and increased activity at the central laboratory and CTU.

In addition, the dropout rate from trials can be significant, particularly where trials involve patients with rapidly progressing disease. For example, in the NLMT trial where patients with advanced lung cancer were considered, genetic profiling was undertaken on diagnostic samples whilst the patients were undergoing standard first line treatment and by the time they were ready to enter the trial after progression from first line treatment, the condition of many had deteriorated too much for them to participate. It is not uncommon for a patient to have died before the results are available. Even if they are still alive, the patient's condition may have deteriorated or they may have decided they no longer want to be involved in the trial. Risk of dropout is further increased since once someone has been recruited, the schedule of trial assessments can be too demanding, and the patient may decide to take the simpler option of not partaking in the trial. Further, receiving a novel therapy may require travel to a more distant location and those with advanced disease may find it challenging to do so. The likelihood of dropping out can be reduced by ensuring rapid turnaround times for biomarker test results, which allow treatment to begin more quickly.

To summarize, given the multiple factors impacting how likely patients will be identified, recruited and retained in a biomarker-guided trial, estimating an accurate rate of recruitment will always be difficult. It is suggested, therefore, that well-designed pilot and feasibility studies are undertaken prior to trial

commencement to ensure a more accurate understanding of recruitment rate as well as a smoother and more rapid process of site set-up. It is often more attractive to incorporate a feasibility study into the main trial, with in-built go/no go criteria, so that starting the trial itself is not unduly delayed. Laboratories should also be sufficiently equipped and efficient to deal with rapid biomarker analysis turnaround. It is important to note that the trials discussed here, and their associated recruitment challenges are some of the first of their kind, and that experience of working on these and other similar trials will also guide us in predicting more accurate and achieving better recruitment rates in future, as well as ease the process of site set-up.

Monitoring samples and laboratories

It is expected that good internal audit trails are in place within laboratories undertaking biomarker assessment for clinical trials, however logistical problems can occur in the transfer of results from laboratory to CTU. Often, results for exploratory biomarkers are batched with hundreds, or thousands of biomarker results transferred at a time, so it is important to agree on procedures for transferring the data accurately before trial commencement. Problems can arise when not all laboratory staff are trained in trials related GCP (Good Clinical Practice) and this is important to ensure that there is a sufficient audit trial and no breaches in confidentiality of biomarker data. Therefore, it is important that laboratories have a good understanding of GCP requirements.

Tracking patient samples requires a significant amount of work and coordination, and is often more complex than what is typically required from laboratory information management systems. For example, a first sample may be received and there might be insufficient tumour, meaning that another sample has to be requested. A full audit trail is required to ensure that the correct biomarker test result is used in the analysis. A significant amount of data cleaning is also typically required.

In terms of the handling and tracking of samples, local research nurses, pathologists, laboratory staff as well as the CTU will be involved. In FOCUS4, a challenge arose with sample management in that patients could be registered up to 12 weeks after starting their first line chemotherapy meaning that a fast turnaround was required at the laboratories to ensure biomarker results were received before the patients had ended their 16 weeks of first line therapy. Further difficulties arise when the tissue obtained was inadequate or is not viable and further requests for samples need to be made back to the original hospital pathology departments.

To ensure optimal efficiency, it is recommended that lots of samples are batched up to be sent all at once instead of using additional resources on several small runs. However, this can lead to problems when lower than anticipated recruitment leads to further delays as labs wait for enough samples to justify running a batch. It is also not practical in smaller trials where a quick turnaround or 'fresh' samples are essential.

Examples of additional problems with sample management were observed in FOCUS4, including resolution of pathology number discrepancies and failure to send GCP compliant documents to the CTU. Further, in NLMT, a particular challenge arose with the lab reports. Here, the genetic result reports did not state which strata in the trial the patient was eligible for, and thus the CTU personnel were required to read the complex reports and determine the appropriate treatment allocation for the patient. Not only was this an additional burden on CTU staff but required rigorous procedures to minimise the risk of error, including sign off of all allocations by the Chief Investigator.

Another challenge associated with biomarker analysis is that science is advancing rapidly with many new opportunities arising in biomarker assessment. It is recommended that a separate lab manual is used outside the protocol in order to minimize any associated protocol amendments.

In terms of ensuring completeness and quality of tissue samples received, communication and collaboration between clinicians and laboratory staff should be strengthened to ensure that samples are taken, stored and sent off in accordance with the protocol. In addition, the CTU's central trial monitoring capabilities should be utilized to ensure efficient sample tracking. Strong collaboration between the CTU and the laboratory staff is essential given how dependent the success of a biomarker-guided trial is on accurate and timely delivery of lab results.

Biomarker assessment

One major challenge during biomarker assessment can arise when samples are heterogeneous. This misclassification problem is less of an issue when patients are randomized but it can lead to a dilution of any observed treatment effects. Biomarker misclassification therefore represents a challenge within biomarker-guided trials, and further sensitivity analyses may be needed to address its effect on the trial result.

Another issue is that whilst it is relatively straightforward to look for the presence or absence of a particular mutation in a particular gene, it is much more difficult to be able to say with confidence that a gene is normal in order to be able to classify a patient. Therefore, the analytical validity of a biomarker in terms of sensitivity and specificity is a challenging but very important issue and understanding the accuracy of an assay is a necessary consideration. If a sample fails completely, it is easy to class it as failed; if there is a partial fail, this represents a difficult result to handle and in the case of partial failure it can be difficult to classify a patient based upon the result.

Challenges can arise when a laboratory is required to change the staining machine and assay during the course of the trial. In this case it is likely that measurements taken prior to the change may need to be repeated using the newer technology or at a minimum calibration of the results investigated. Apart from the significant cost implications, it is also necessary to appropriately consider cases where the new result differs from the previous one. Conducting an analysis that is

stratified by the date the assay measurement changed may be an alternative way of handling these sorts of biomarker adaptations during the trial.

Data sharing issues

When a pharmaceutical company is involved in a trial, along with the clinical study report it may be expected that the trial data will be shared with the company at the end of the trial, within a data sharing framework, and this will be detailed in the contract. However, in early phase trials companies may want data to be shared in real time or at least at periodic intervals (e.g., to guide business decisions) during the trial. Current consensus suggests that this is not good practice for phase III trials. For single arm phase II trials which are more exploratory in nature opinions differ as to its merits, especially when treatments are being evaluated using a response endpoint. One argument against this type of data sharing during the trial is that historically, if you questioned why a phase II trial had failed, one reason was that the clinicians or chief investigators were too selective in picking their patients when they had a fixed threshold of responders to reach to call it a success (e.g., selecting patients more likely to respond creating a distorted cohort of patients in the latter part of the trial). Sharing data during the trial could result in such situations arising again.

Further, whilst decisions in terms of the closure of strata are the responsibility of the trial oversight committees, pharmaceutical companies may wish to be involved in the decision making process.

Whilst data sharing requests from pharmaceutical companies are likely to be common in biomarker-guided trials, differing viewpoints in terms of how and when data should be shared can be particularly challenging for the trial management team. To ensure that good relations are maintained with all interested parties, it is recommended that a clear data sharing policy and common data standards are developed and agreed at the point of contract negotiation, prior to trial outset, with all aspects of decision-making explicitly stated as the remit of the independent trial oversight committee.

Resources

In terms of clinical trials unit (CTU) management ensuring the availability of appropriate resources is a challenge. Biomarker-guided trials require adequate funding for dedicated personnel. The complexity of the required IT support is frequently underestimated and essential for biomarker-guided trials. The complexity of the Case Report Forms (CRF) is a particular challenge, since the data required often varies between strata. Hence, there is a need for several different case report forms equivalent to having many separate trials but with the additional burden of needing a more sophisticated over-arching database structure. Protocol amendments lead to additional problems due to the fact that for just one amendment (e.g. an additional medical assessment), all CRFs require modification. The consequences of needing separate CRFs for different comparisons were observed in the FOCUS4 trial. The trial uses electronic data capture (eDC) where local site staff enter data directly into the database. When the trial was first set-up, it opened with only one molecular comparison and a single non-stratified

comparison. The aim was to have all cohorts eventually in one main database, however, this has proven difficult and it was decided to include future comparisons in separate databases, meaning that sites have to open a number of different databases to enter data for their FOCUS4 patients rather than just one.

Furthermore, administrative support for tasks such as preparing site packs is often underestimated, and the need for collaboration between a clinical trials unit and biomarker labs adds further pressure onto resources. Challenges associated with such collaboration relate to laboratory agreements (e.g., impact on data sharing) and the processes for tracking, blinding and pseudo-anonymization of samples. In addition, specialist biomarker expertise is required, something which is not typically available within a trials unit.

More complex work is also needed when adding new treatment groups to platform trials or making other adaptations to a trial. Several issues need to be considered at that time; in essence incorporating a new treatment group in a master protocol is equivalent to setting up a new trial, including protocol writing and case report forms development, database development, setting up of contracts, drugs supply etc. while recruitment, co-ordination and data collection for existing treatment groups continues. Further, whilst existing systems and processes may work with an initial small number of groups, they may not work as well with a much larger number of groups, and it is therefore difficult to predict level of resource upfront, leading to inefficiencies down the line.

POETIC faced several challenging issues, many of those relating to the need to extend the clinical trials culture across multidisciplinary teams involved at cancer diagnosis, and the integration of research protocols into busy clinics. For these reasons, a variety of pathways (different fresh tissue collection options) as well as different types of tissue (availability of biological and non-biological centres) were considered.

To summarize, the resources required for efficient management of a biomarker-guided trial should not be under-estimated and clinical trial units in particular need to ensure that they are prepared in particular for the administrative burdens that come with such trials, and adequately cost them into any funding applications.

Discussion

At our workshop '*Biomarker-guided trials: challenges in practice*' several practical challenges were considered:

- Funding issues, including higher resources due to typical complexity of biomarker-guided trials, difficulties in making accurate cost predictions at the outset, confusion over who should meet biomarker testing costs and the need for sharing of knowledge and experience between researchers and funders' regarding the implications of using such designs.

- Ethical and regulatory issues, including uncertainty about whether amendments require new approvals, the need to maintain a collaborative relationship and

effective communication with HRA and regulators and issues relating to communications with patients.

- Recruitment issues such as the difficulty in predicting an accurate recruitment rate, delays in setting up sites, and unknown patient dropout rates.

- Issues arising in the tracking and monitoring of samples and laboratories when not all laboratory staff are GCP trained, the need for efficient sample processing and tracking, dealing with changing technologies, challenges of biomarker misclassification and the need to establish effective communication and collaboration between clinicians and laboratory staff to address current challenges.

- Issues regarding data sharing agreements, particularly when working in collaboration with pharmaceutical companies.

- Resourcing issues, including underestimation of the extent of IT and administrative support required, and of the complexity of databases and CRFs.

Although many of the challenges discussed relate to the more complex biomarker-guided trials such as umbrella trials, similar challenges can appear in biomarker-guided clinical trials more generally. Likewise, we acknowledge that some of the challenges identified are equally relevant to more complex trials irrespective of whether they incorporate biomarkers or not (eg predicting the cost of adding/removing arms, approving amendments, CTU resource issues), whilst others are specific to biomarker-guided trials (eg predicting recruitment rates when

biomarker prevalence unknown, ethical issues from communicating results of biomarker tests, sample processing and laboratory challenges) .

Despite the aforementioned challenges, the biomarker-guided trials discussed within this report represent successful research projects using novel designs, which will hopefully inform future practice. NLMT, for example, provides a great opportunity for widespread national collaboration with leaders from the lung cancer community within academia, the health service and the pharmaceutical industry, and direct collaboration with CRUK. It promises to make a real contribution to the knowledge on precision medicine by testing new drugs tailored to a specific biomarker-defined subgroup.

Further, the FOCUS4 trial, uses an efficient Multi-Arm, Multi-Stage (MAMS) design which has proved to be particularly efficient in the mCRC disease setting where the progression-free survival (PFS) event rate is high and interim analyses are triggered quickly. Its other successes include having a strong collaborative trial management group with a very engaged overall Chief Investigator (CI), the use of different CIs for each comparison, early engagement with the Research Network, clear protocol structure and nomenclature as well as the single regulatory and ethics approvals.

TOPARP demonstrated anti-tumour activity of olaparib in patients with advanced CRPC (Mateo J et al, N Engl J Med 373(18):1697-708) and was the first molecular treatment stratification in metastatic castrate resistant prostate cancer (mCRPC). A successful collaboration between ICR, AstraZeneca and Cancer

Research UK (National Cancer Research Network Collaboration) led to the FDA granting olaparib breakthrough therapy designation based largely on the results of TOPARP-A.

EU-PACT provides an example of an international, multi-site trial which, due to its pragmatic approach and adoption of the biomarker-strategy design allowed the improved treatment outcomes from using biomarker-guided approach to prescribing warfarin to be demonstrated. This has led to a subsequent matched-cohort study which demonstrated the successful implementation of the biomarker-guided approach into clinical practice(18), and a trial based on EU-PACT is currently being planned in Africa to test the clinical utility of a personalised approach to warfarin dosing in low-resource settings.

To conclude, the examples of biomarker-guided trials considered here demonstrate the real benefits of adopting such designs, despite the teething problems resulting from using such novel methodologies. However, the significant investments required to successfully conduct such trials should not be underestimated, and it is imperative that the practical challenges they bring for clinicians, laboratories, regulators, academia, industry and patients as outlined above should be acknowledged and addressed at the outset. As the need for trials in stratified medicine increases however, it is anticipated that through experience stakeholders will become more familiar with the designs and the procedures involved in conducting and managing them will evolve and adapt accordingly. It is important therefore that the knowledge gained by those with experience of biomarker-guided trials is communicated to the wider research community such

that all stakeholders are educated about the complex issues that biomarker-guided clinical trials face and recommendations for how they may be overcome.

Conflict of interest

The authors do not have any competing interests to declare.

Funding

The workshop that led to this publication was funded by the MRC Network of Hubs for Trials Methodology: grant MR/L004933/1.

Acknowledgements

We wish to express our sincere appreciation to all the people who shared their expertise and experience with us during the preparation of this report. Our special thanks go to the people who participated in the workshop on "Practical challenges in the conduct of biomarker-guided clinical trials" and the MRC Hubs for Trials Methodology Research Network's Stratified Medicine Working Group (SMWG).

References

1. Landeck L, Kneip C, Reischl J, Asadullah K. Biomarkers and personalized medicine: current status and further perspectives with special focus on dermatology. Exp Dermatol. 2016;25(5):333-9.
2. Bailey AM, Mao Y, Zeng J, Holla V, Johnson A, Brusco L, et al. Implementation of biomarker-driven cancer therapy: existing tools and remaining gaps. Discov medicin. 2014;17(92):101-14.
3. Antoniou M, Jorgensen AL, Kolamunnage-Dona R. Biomarker-Guided Adaptive Trial Designs in Phase II and Phase III: A Methodological Review. PLoS ONE. 2016;11(2):e0149803.

4. Antoniou M, Kolamunnage-Dona R, Jorgensen AL. Biomarker-Guided Non-Adaptive Trial Designs in Phase II and Phase III: A Methodological Review. J Pers Med. 2017;7(1).
5. Freidlin B, Korn EL. Biomarker-adaptive clinical trial designs. Pharmacogenomics. 2010;11(12):1679-82.
6. Lin JA, He P. Reinventing clinical trials: a review of innovative biomarker trial designs in cancer therapies. Br Med Bull. 2015;114(1):17-27.
7. Antoniou M, Jorgensen AL, Kolamunnage-Dona R. Fixed and Adaptive Parallel Subgroup-Specific Design for Survival Outcomes: Power and Sample Size. J Pers Med. 2017;7(4).
8. Middleton G, Crack LR, Popat S, Swanton C, Hollingsworth SJ, Buller R, et al. The National Lung Matrix Trial: translating the biology of stratification in advanced non-small-cell lung cancer. Ann Oncol. 2015;26(12):2464-9.
9. Phase II trial of olaparib in patients with advanced castriation resistant prostate cancer (TOPARP). [Available from: http://www.icr.ac.uk/our-research/our-research-centres/clinical-trials-and-statistics-unit/clinical-trials/toparp.
10. ATLANTIS: An adaptive multi-arm phase II trial of maintenance targeted therapy after chemotherapy in metastatic urothelial cancer. [Available from: http://www.crukctuglasgow.org/eng.php?pid=atlantis.
11. PRIMUS 001 [Available from: http://www.crukctuglasgow.org/eng.php?pid=primus001.
12. Trial of perioperative endocrine therapy - individualising care (POETIC). [Available from: https://www.icr.ac.uk/our-research/centres-and-collaborations/centres-at-the-icr/clinical-trials-and-statistics-unit/clinical-trials/poetic_trial.
13. Robertson J, Dowsett M, Bliss J, Morden J, Wilcox M, Evans A, et al. Abstract GS1-03: Peri-operative aromatase inhibitor treatment in determining or predicting longterm outcome in early breast cancer – The POETIC* Trial (CRUK/07/015). Cancer Research. 2018;78(4 Supplement):GS1-03-GS1-.
14. FOCUS4 [Available from: http://www.focus4trial.org/.
15. Pirmohamed M, Burnside G, Eriksson N, Jorgensen AL, Toh CH, Nicholson T, et al. A randomized trial of genotype-guided dosing of warfarin. N Engl J Med. 2013;369(24):2294-303.
16. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet Med. 2013;15(7):565-74.
17. Roche MI, Berg JS. Incidental Findings with Genomic Testing: Implications for Genetic Counseling Practice. Curr Genet Med Rep. 2015;3(4):166-76.
18. Jorgensen AL, Prince C, Fitzgerald G, Hanson A, Downing J, Reynolds J, et al. Implementation of genotype-guided dosing of warfarin with point-of-care genetic testing in three UK clinics: a matched cohort study. BMC Med. 2019;17(1):76.

Trial	Disease	Primary endpoint(s)	Number of arms	Trial design	Type of biomarker(s)	Role of biomarker(s)	Responsibility for overall management	Primary funding source(s)	Challenges
NLMT	Advanced non-small cell lung cancer	Best objective response; Durable clinical benefit; Progression-free survival time	8	Bayesian adaptive umbrella design	Genetic markers	To determine arm/treatment allocation	Early Drug Development (EDD) Trial Management Team based within the Cancer Research UK Clinical Trials Unit (CRCTU), University of Birmingham	Cancer Research UK	Uncertainty regarding total costs of trial – resolved by submitting estimated future costs and CRUK providing agreement in principle; additional costs of biomarker analysis – resolved by funding molecular screening platform as separate entity; significant

									dropout due to recruiting patients with advanced disease; CTU personnel required to interpret biomarker reports themselves to determine relevant treatment arm
TOPARP	Metastatic castration resistant prostate cancer	Treatment response according to pre-specified criteria	TOPARP-A: single arm TOPARP-B: two-arm randomised	TOPARP- A: Open-label, single arm, two part adaptive design phase II trial. TOPARP-B: Open-label, two-arm randomised, each arm with a	Genetic markers	TOPARP-A: Biomarker development - to identify predictive biomarkers of response to olaparib TOPARP B: Biomarker validation - biomarker guided patient	Institute of Cancer Research, UK	Trial run under the NCRN-AZ initiative (CRUK and AZ funded)	Complex sampling collection and processing requirements outside standard pathway at sites. QA sample failures at central labs which lead to

				single stage phase II design.		selection for eligibility to confirm sub-group identified in A			delays in biomarker results being available. Greater CTU and lab activity/resource required to manage challenges and ensure collaborators' expectations were met.
ATLANTIS	Metastatic urothelial cancer	Progression-free survival	3	Adaptive multi-arm design	Homologous recombination deficiency and genetic markers	To determine arm/randomisation treatment	Clinical Trials Unit, University of Glasgow	Cancer Research UK	
PRIMUS001	Metastatic pancreatic cancer	Progression-free survival	2	Adaptive design	Genetic markers	For subgroup analysis of primary outcome, and to determine eligibility for recruitment following interim	Clinical Trials Unit, University of Glasgow	Cancer Research UK	

						analyses			
SALONICA	Ovarian cancer	Progression-free survival	1	Sequence of single arm trials, but plans to progress to Bayesian adaptive randomised design	Genetic markers	Initially for subgroup analysis, and then to determine randomisation ratio	Clinical Trials Unit, University of Glasgow	N/A – planning stage	
TASTER	Chronic Myeloid Leukaemia	Progression-free survival	1	Series of single arm trials	Biomarkers contributing to molecular signatures	To determine eligibility for which single arm trial	Clinical Trials Unit, University of Glasgow	N/A – planning stage	
POETIC	Breast cancer	Relapse free survival	2	Two-arm parallel randomised controlled trial	Genetic marker and Gene expression profile	To determine eligibility and for subgroup analyses	Institute of Cancer Research, UK	Cancer Research UK	

FOCUS 4	Metastatic colorectal cancer	Progression-free survival	I 3 molecularly stratified trials and 1 non-stratified trial	Multi-arm, multi-stage umbrella design	Genetic markers	To determine arm/randomisation treatment	MRC Clinical Trials Unit at UCL	NIHR/MRC EME Programme and Cancer Research UK	Intensive CTU resource requirements for the multi-tasking aspects of the adding and dropping arms design; High costs of running trial – resolved by securing joint funding between MRC/NIHR EME and Cancer Research UK and having trial conducted in a CTU with separate core funding; Delays in biomarker results turnaround

									and failed samples; Pathology number discrepancies; Failure to send GCP compliant documents from pathology lab to CTU; Needing comparison-specific CRFs therefore sites having to deal with several separate databases;
EU-PACT	Atrial fibrillation and venous thromboembolism	Time in therapeutic INR range during first three months of treatment	2	Two-arm parallel randomised controlled trial	Genetic markers	Predict therapeutic dose	Wolfson Centre for Personalised Medicine, University of Liverpool	European Commission Seventh Framework Programme	Need for rapid turnaround of genotyping results to allow same-day treatment initiation at

									predicted dose – resolved by working with industrial collaborator to develop efficient point of care test