Oncology drug development has seen a shift from the development of cytotoxic chemotherapy to that of molecularly targeted agents. Different mechanisms of action and toxicity profiles may mean that traditional oncology trial designs are no longer optimal for the development of these agents. Furthermore, the wealth of agents that are being developed, coupled with a constrained research environment has increasingly highlighted the need for efficient clinical trial design, both to filter agents as well as to advance promising agents rapidly into clinical development. We review the adaptations to traditional Phase I and II clinical trial design that are hoped to address some of the current challenges of drug development for oncology therapeutics.

Keywords: clinical trial • drug development • oncology

Oncology drug development has seen an almost complete shift from the evaluation of cytotoxic chemotherapy to molecularly targeted agents in the last decade. The use of traditional trial designs to develop these agents that have novel efficacy and toxicity profiles has not always been successful. Less than 10% of new agents receive regulatory approval and at least 50% of oncology agents fail at the Phase III evaluation [1]. These statistics are likely to worsen further given the rash of negative trials reported in the last 2–3 years [2–5] and the large number of new agents entering clinical trials [6]. Clearly, there is an urgent need to optimize clinical trial design.

Cytotoxic chemotherapeutic agents generally affect DNA or cell division leading to nonspecific DNA damage or cell cycle arrest, and are not usually ‘targeted’, although some exceptions exist. These agents are usually administered parenterally, have a dose–response relationship with higher doses resulting in more toxicity, but also more efficacy, with dose escalation often limited by organ toxicity. The maximum administered dose (MAD) and associated recommended Phase II dose (RPTD; also referred to by some authors as maximum tolerated dose [MTD]) is the traditional end point in Phase I trials, with the quantification of tumor shrinkage (response rate [RR]) the most common end point to assess preliminary evidence of efficacy in Phase II trials.

By contrast, molecularly targeted agents inhibit specific proteins within pathways believed critical to the malignant phenotype including oncogenesis, metastases, cell signaling and angiogenesis. These agents often have a toxicity profile characterized by low grade, but chronic, adverse effects, especially when they are administered orally in a chronic schedule. Not uncommonly, some adverse effects are ‘off-target’ effects (i.e., related to the chemical structure rather than a sequela of the effect on the putative target). In addition, many of these agents do not have a linear dose–response relationship, resulting in the desired biologic effect occurring, at least hypothetically, at a dose substantially lower than the RPTD [7]. In some instances, toxicities result from
unanticipated or unknown on-target effects, such as those observed in the early development of sorafenib, initially thought to be a RAF inhibitor, but which in fact had vascular endothelial growth factor (VEGF) receptor inhibitory activity as well. Ethical issues, such as the desire to minimize exposure of patients to both too high doses (and toxicity) and too low doses (with lack of efficacy), and concerns regarding acquisition of tissue for biomarker studies, add complexity to design questions. Further complicating development, the traditional surrogate of efficacy, RR, may not be relevant, as the clinical benefit may result from stabilizing disease rather than tumor regression, or may be apparent only in select subsets of patients.

In this review we focus on recent challenges of early cancer therapeutic development and some possible solutions for studies in adult populations.

**Phase I trials**

The primary objective of a Phase I trial is to evaluate the toxicity profile, and determine the RPTD of a new agent or combination of agents. Despite this seemingly simple objective, numerous different Phase I trial designs have been proposed to address the challenges described above (Table 1 & Figure 1). A number of recommendations have been developed to assist researchers with choosing the optimal design and end points for their study [8–10].

- **Rule-based design**

  Rule-based designs, founded on those originally described by Fibonacci, are commonly used and are often referred to as the 3+3 designs [8]. Patients are enrolled into small cohorts (usually of three patients), with a starting dose initially defined by data from animal models. Provided that no predefined dose-limiting toxicities (DLTs) occur in the first three patients (i.e., ≤33% in the first cycle), the dose is escalated to the next dose level using a predefined algorithm. Dose escalation proceeds until a DLT occurs; at this point, the dose level is either expanded (if ≤1 DLT) or patients are accrued to a lower dose level (if ≥2 DLT). If a further DLT is observed in the expanded dose level this is declared the MAD and the dose level below is usually considered the RPTD. While this trial design continues to be the most commonly used [11,12], it has been criticized on the grounds of inefficiency, lack of statistical foundation, ethical considerations (the number of patients treated at low doses) and applicability to targeted agents where acute DLT in cycle one may not be relevant.

  These designs are based on the assumption of a linear dose–response relationship, such that an increase in dose is associated with increasing efficacy and toxicity. This has generally performed well when cytotoxic agents are considered; however, this model may not be robust for molecularly targeted agents, especially when given by mouth where even chronic low grade toxicity may be intolerable, where doses are limited for other reasons (volume of a drug, or number of pills that can be taken orally, cost constraints for production of monoclonal antibodies) or where there is a disparate dose–response relationship for toxicity and efficacy.

  Rule-based designs, such as the 3+3 design, have been criticized for the proportion of patients treated at low, likely nonefficacious, doses. Newer designs, such as accelerated titration (AT), attempts to address this by enrolling a single patient to each early dose level until a prespecified

<table>
<thead>
<tr>
<th>Table 1. Advantages and disadvantages of major Phase I trial designs.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phase I study design</strong></td>
</tr>
<tr>
<td>Traditional 3+3 design</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Accelerated design</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Continual reassessment model</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Escalation with overdose control</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

DLT: Dose-limiting toxicity; MAD: Maximum administered dose.
level of toxicity (e.g., ≥ grade 2) is observed. Thereafter, three patients per dose level are accrued until the MAD/RPTD is reached. Simon and colleagues demonstrated that AT designs have the potential to reduce the number of patients necessary to determine the RPTD and increase the number of patients receiving a potentially therapeutic dose. Unfortunately, these theoretical advantages may not translate in clinical practice. In a review of the traditional versus AT design, the proportion of patients that were treated below the MTD was lower for the AT design compared with the 3+3 design (58 vs 71%); however, the total number of dose levels was higher for the AT design, while the length of study was similar for both designs.

Other described designs include those where dose escalation decisions are based on pharmacokinetic data (challenging due to logistics and interpatient variability), designs including intrapatient dose escalation (considered problematic and biased if intrapatient dose escalation contributes to dose decisions) and even patient-directed dose selection. However, model-based designs are of particular interest for use with targeted agents, especially those that include response and biomarker data in the RPTD decision.

### Model-based designs

Model-based designs adapt to new data that become available during the clinical trial. Proponents consider them to have the potential to increase efficiency, treat more patients at or near optimal doses and address several questions within the context of a single trial. Although there are no direct comparisons of efficiency between model-based designs and traditional designs, reviews suggest that the new trial designs might result in more efficient and effective treatment selection.
in fewer cohorts, or fewer patients treated at lower dose levels [5,16]. Unfortunately, despite the potential advantages of model-based designs, they have been utilized in only a minority of Phase I clinical trials [12]. Potential reasons for this include the added complexity of the trial and the need for biostatistical support during the conduct of the study [17]. Commonly discussed examples are the continual reassessment method (including trivariate continual reassessment model [Tri-CRM]) and escalation with overdose control (EWOC).

The continual reassessment method selects the first dose near the predicted RPTD based on statistical modeling [18]. Toxicity data obtained from the first patient enrolled onto the trial are then used to reassess the probability of a DLT occurring at a specific dose level; this information is then used to select the next appropriate dose. Proponents of the model believe that it offers a more accurate and precise measurement of the MTD with fewer patients experiencing DLT [19]. However, there are concerns that patients would be exposed to toxic doses of the experimental agent due to the rapid increase in dosing that the model proposes [20]. In order to address this, modified continual reassessment methods have been proposed that decrease the rate of dose escalation, thus reducing the risk that a trial participant will be exposed to a toxic dose of experimental therapy [16,21,22].

The EWOC design minimizes the risk of overdose by specifically setting the probability of a dose above that desirable to a preset low level. The trial design then proceeds in a similar manner in order to approach the RPTD efficiently. Simulation studies have demonstrated that a greater proportion of patients are treated at optimal doses compared with rule-based design [23], and has the potential additional safeguard over the continual reassessment method.

Adaptive designs may also simultaneously address other questions within the trial. For example, Thall et al. described a method of adjusting the dose based on the probability of both toxicity and efficacy at different dose levels [24]. Dose selection is selected based on the probability of a specific dose fitting the toxicity and efficacy boundaries. It was reported that this method enabled the majority of patients to be treated at a dose close to the optimum for efficacy versus toxicity. Tri-CRM also allows the inclusion of efficacy end points into the model [25].

- Biomarkers & the Phase I trial

While traditional ‘biomarkers’ such as response and pharmacokinetics are commonly included in all Phase I design, less traditional biomarkers (functional imaging end points, pharmacodynamic changes in tissue or surrogate tissue biomarkers) have the potential to add value to Phase I trials, especially where it is anticipated that toxicity alone will not prove an appropriate surrogate to define the RPTD. A comprehensive review of functional imaging in early clinical trials is beyond the scope of this review, but excellent reviews are available [26].

Biomarkers, especially tissue-based ones, may be incorporated into Phase I clinical trials for a number of reasons. They may be used to confirm the agent is achieving the desired molecular effect (proof of principle) or penetrating into the tumor. They may also be used to help define the RPTD, either by demonstrating an effect estimate, based on data obtained from preclinical studies, likely to be associated with ‘efficacy’, or by demonstrating a dose response (or lack thereof). This is especially useful for drugs that are associated with little toxicity. Finally, their use in the early clinical trials setting may allow early identification of subsets of patients most likely to benefit. This can be further explored in the Phase III setting.

Ideally, biologic effects would be assessed by direct measurement in the tumor or, failing that, in qualified surrogate tissues. Direct tumor assessment is difficult to achieve as repeated tumor sampling from patients may not be feasible, although many trialists now design their studies to include this only at the RPTD and expansion. Alternatives include assessing biologic effect in tissue that is easier to acquire, such as peripheral blood or functional imaging. Although these ‘surrogates’ may confirm proof of principle, they require a clear understanding of the mechanism of action of the drug to be evaluated and robust qualified and/or validated assays [9,27]. While increasing efforts are being made to identify relevant biomarkers and develop appropriate assays prior to the institution of clinical trials, this has not always been possible in the past. Clearly, attempting to identify biomarkers and develop assays only during the conduct of Phase I and II clinical trials has not been very successful.

A meta-analysis of 2458 Phase I clinical trials reported that biomarkers only aided the determination of dose in 13% of studies and that, despite increasing inclusion of biomarker studies, the role in these studies is predominantly supportive [28]. These disappointing results may reflect the challenges of the inclusion of biomarkers into clinical trials: selection of the appropriate biomarker, availability of qualified or validated assay; ability to collect the required sample; appropriate interpretation of the results; ethical issues; cost and complexity and the generalizability of biomarkers across different tumor types. However, these results may also not be reflective of more recent clinical trials of this class of agents.
On the other hand, there are examples of the successful incorporation of biomarkers into early-phase clinical trials, demonstrating their potential impact in the future. In the development of the DNA repair protein polyADP ribose polymerase (PARP) inhibitor, Fong et al. not only utilized biomarkers in tumor and surrogate tissues to demonstrate proof of principle of the PARP inhibitor, but, based on preclinical data, enriched the trial with patients known to have germ-line BRCA mutations demonstrating an impressive RR in this subset of patients. These data have informed the further development of this agent. Biomarkers were also of use in the development of other agents such as bortezomib [30].

Most recommendations have advocated the continued exploration of toxicity across a dose range in Phase I clinical trials, the judicious inclusion of (preclinically) qualified and validated biomarkers, and have recommended against the selection of patients for inclusion into Phase I trials, unless there is a very strong hypothesis. Biomarkers at the current time are generally used for hypothesis generation and rarely impact on key clinical trial decisions. Biomarker discovery should commence early in the development of a new agent and recent guidelines from the National Cancer Institute Biomarker Task Force have recommended initiatives to enhance the early development of biomarkers.

### Phase II clinical trial design

The primary goal of the Phase II clinical trial is to collect preliminary data on the efficacy of a new therapy. If the design is appropriate, drugs that have poor or uninteresting activity can be identified and further development discontinued, and resulting data used to direct the future development of promising agents. The Phase II trial can also provide an opportunity to explore different doses and schedules, in order to select the optimal regimen in a larger and more homogeneous population than is possible during Phase I evaluation; the goal with these changes is to improve efficacy, not to acquire further primary toxicity data. In addition, Phase II trials can further evaluate biomarkers of interest to allow appropriate patient enrichment for later trials. On rare occasions, Phase II data has been the sole source of data leading to regulatory approval of an agent, but usually only in situations such as rare tumors or instances of drugs with overwhelming efficacy.

The Phase II trial design is not intended to be the definitive assessment of efficacy and should thus be economical in size, yet also be adequately powered to yield sufficiently robust data to avoid false-positive results and a subsequent negative Phase III trial. These competing demands require a carefully designed Phase II trial, tailored to the agent, disease and aim of the trial.

#### Phase II trial end points

Logically, the first decision in the design of the Phase II clinical trial should be the choice of end point. Ideally, the end point should be one that can be rapidly reached, but should be a robust surrogate for the ‘gold standard’ – overall survival (OS) and improvement in quality of life. The end point most commonly used remains RR, but time-dependent end points (e.g., progression-free survival (PFS)) are increasingly being used. Novel end points such as biomarkers, functional imaging or alternate ways to measure response are increasingly included.

#### Response-based end point

The objective RR is defined as the proportion of patients that achieve a complete or partial response (PR) defined by a set of specific criteria. For solid tumors, the response evaluation criteria in solid tumor (RECIST) is a commonly accepted standard. The change in the size of the tumor is categorized as complete response (CR; disappearance of all evidence of disease), a PR (a 30% reduction in the longest diameter of the sum of the measured lesions), progressive disease (PD; increase in size of the sum of measured lesions by 20%) and stable disease (SD; all measurements between PR and PD). RR is not a direct measure of clinical benefit to the patient, although it has been demonstrated to be a surrogate of OS in some but not all malignancies, at least for cytotoxic agents.

Targeted agents have a number of challenges with regards to the selection of the primary efficacy end point. Numerous studies have demonstrated very modest RR but OS benefit, including erlotinib in non-small-cell lung cancer (RR: 8.9%) [41], sorafenib in renal cell carcinoma (RR: 2%) [42] and cetuximab for colorectal cancer (RR: 8%) [43]. Traditional single-arm Phase II studies using 5% as the null hypothesis and 20% for the alternate hypothesis RR, respectively (commonly used in Phase II clinical trials), would have declared these useful agents as ‘not promising’. Despite justifiable criticisms regarding the use of RR in Phase II trials of targeted agents, a review of 89 trials evaluating 19 targeted agents found that RR was the primary or co-primary end point in 69% of trials, with responses noted in 43% of them. Higher RRs were associated with a higher likelihood of receiving regulatory approval and no targeted agent without evidence of RR received regulatory approval. Unfortunately, designing Phase II trials with a lower RR will increase sample sizes and may increase the risk of a negative Phase III trial. Adaptations to the standard RECIST have been proposed to address these issues, including the use of waterfall plots and use of change in tumor size as a continuous variable, but as yet require prospective testing and validation.
- **Novel adaptations to the assessment of response**

  Stable disease as defined by RECIST is a default category that encompasses both patients that have definite tumor shrinkage but not to the degree to meet the PR bar, as well as patients who have enlarging tumors but not to the level of PD. These two scenarios are likely to be biologically distinct but are assigned the same response category (Figure 2).

  Trial designs have been proposed where response is assessed as a continuous variable, whereby the mean of the difference in tumor size is assessed at a specific time point [45]. We reviewed performance of this methodology retrospectively for the Phase III trial NCIC CTG PA1, which evaluated a novel metalloproteinase inhibitor, BAY 12-9566, versus gemcitabine chemotherapy [46].

  The first interim analysis was based on the absence of progression at 8 weeks, and the trial continued as the proportion of patients with PD was sufficiently low. The trial was ultimately negative, stopping at the second interim analysis, demonstrating a favorable outcome for patients randomized to gemcitabine (control arm). However, in the retrospective modeling, if the response assessment used response as a continuous variable, the trial would have been appropriately halted at the first interim analysis [47].

- **Multinomial stopping rule**

  The multinomial stopping rule incorporates both RR and the early progression rate (defined as the proportion of patients with progressive disease at the first evaluation time point) [48]. Potentially, this design may reduce the false-negative rate of trials evaluating cytostatic agents, which may have a high stable disease rate but a low RR leading to their identification as ‘uninteresting’ agents. However, incorporation of the proportion of early progression rate, which will be inversely proportional to the proportion of patients with CR, PR or SD, will overcome this challenge. Dent et al. evaluated the performance of this rule by retrospectively applying the multinomial rule to 39 Phase II trials that utilized the Gehan [49] or Fleming rule [50]. Of these 39 trials there was disagreement between the standard and multinomial rule in nine trials, all of which would have been appropriately discontinued using the multinomial rule.

  A subsequent review, applied retrospectively to 15 trials of single-agent targeted therapies, demonstrated that the multinomial rule stopped trials more frequently in the first stage compared with the Fleming rule [51]. The task force Methodology for the Development of Innovative Cancer Therapies (MDICT) has recommended consideration of the use of multinomial end points for agents with low RRs [52].

  The use of response as an end point is particularly challenging in some malignancies (mesothelioma, brain tumors and those with a predilection for bone-only disease) and in patients with no measurable disease. PET and functional and molecular imaging (the assessment of physiological and molecular change, respectively) have the potential to enhance conventional anatomical imaging. This may enable a more accurate evaluation of response and biologic activity of a novel agent [53]. Although promising, these have not yet been validated as surrogates of the old standard of OS and increased quality of life. The next major revision of RECIST should include these imaging modalities. Other initiatives include the collection and enumeration of circulating tumor cells [54].

- **Time-dependent end points**

  Time dependent end points (e.g., PFS and OS) overcome the challenge of evaluating cytostatic agents that may induce stable disease rather than tumor regressions. PFS is more frequently utilized compared with OS as OS requires longer follow-up and there is potential bias from patients who receive subsequent lines of therapy. A recent review suggests that PFS is

![Figure 2. Simulated waterfall plot of response to two drugs (drug A and drug B). Both agents have the same stable disease rate. Both drug A and drug B have the same response rate; however, drug B appears to have a greater efficacy than drug A.](image-url)
Early-phase oncology clinical trial design

**Phase II trial designs**

Although nononcology Phase II trials generally include a prospective control arm [56], this has not been the case in oncology [44]. Of the numerous different oncology Phase II trial designs (Figure 3), the most hotly debated issue is the role of randomization. A single-arm Phase II trial is a simple and efficient trial design that has been used extensively in oncology. The trial is frequently designed to accrue patients in two stages; for example, the Simon two-stage design [49,50,57]. Patients are accrued being utilized more commonly and may predict for greater success in the Phase II setting [55]. PFS does not overcome all the challenges of assessing response as described above, as the time point of progression still has to be robustly defined, which is usually defined by RECIST. Furthermore, a randomized design is almost certainly required as time-dependent end points are more sensitive to differences in patient characteristics than RR, especially when historical controls are unlikely to be robust.

---

**Figure 3. Examples of Phase II clinical trial designs.**

- **Single arm**
  - Stage 1
  - Stage 2
  - Historical control
  - Regimen A
  - Regimen B (or noncomparative concurrent control)
  - Efficient design, short follow-up
  - Allow several regimens/schedules to be evaluated in a single trial
  - Requires use of historic controls

- **Randomized noncomparative**
  - Historical control
  - Regimen A
  - Regimen B (or noncomparative concurrent control)
  - Efficient design to evaluate more than one regimen/schedule
  - Best regimen may not be superior to standard therapy
  - Prospective control arm
  - Larger sample size necessary; underpowered trial may increase false-positive results

- **Randomized comparative**
  - Regimen A
  - Regimen B
  - Regimen C
  - Effiecient design to evaluate more than one regimen/schedule
  - Best regimen may not be superior to standard therapy
  - Prospective control arm
  - Larger sample size necessary; underpowered trial may increase false-positive results

---

**Pros/cons**

- Efficient design, short follow-up
- May underestimate efficacy of cytostatic agents; some diseases difficult to measure; use of historic controls
- Allow several regimens/schedules to be evaluated in a single trial
- Requires use of historic controls
- Efficient design to evaluate more than one regimen/schedule
- Best regimen may not be superior to standard therapy
- Prospective control arm
- Larger sample size necessary; underpowered trial may increase false-positive results

---

**Future science group**

**Clin. Invest.** (2011) 1(1)

39
Review: Clinical Trial Methodology

Bradbury, Hilton & Seymour

Randomized Phase II trial designs
Randomization has a number of advantages, including the reduction of bias introduced from differences within patient cohorts, which could influence the interpretation of the trial results, by the introduction of a contemporaneous control cohort. In addition, different regimens, schedules or doses can be evaluated. Designs may be comparative or noncomparative. Randomized designs are recommended to evaluate combinations of novel agents with standard therapies, and when time-based end points are appropriate, as they are particularly sensitive to the bias that is introduced from improvements in the outcome of a disease over time, suggesting that the use of historic controls is not appropriate in such circumstances [36].

Noncomparative randomized designs
A number of noncomparative randomized trial designs may have been described. For example, patients may be randomized to one of a number of different regimens or schedules, but each arm is compared with historic controls with different stopping rules and not across the randomized arms. An alternative design is to randomize patients to an experimental treatment or a control arm. The trial is not powered for a formal statistical comparison across arms, but the incorporation of a formal control arm may be more robust than the use of historic data.

Comparative randomized designs
There are a number of comparative Phase II designs. The randomized selection design (pick the winner) described by Simon et al. randomizes patients to one of a series of experimental arms, from which the arm that appears most efficacious is chosen for further development over the others [61]. This design does not include a formal control arm or specifically utilize historic data. It has the disadvantage that an arm may appear superior to the other within a Phase II trial, but is actually inferior to standard therapy [61].

A second design is a randomized comparative Phase II design with a prospective control arm. The trial may be open label or blinded. Randomization is particularly applicable to the evaluation of combination therapies, where the addition of a novel agent to standard therapy is difficult to evaluate without a contemporaneous control. Importantly, randomized trials are not intended to be definitive efficacy trials and attempts to maintain efficiency by keeping patient numbers down may result in a small and underpowered trial with the risk of false-positive results. Other designs include the randomized discontinuation design. All patients receive the active agent and at a prespecified time point those patients with stable disease are randomized to either continue therapy or discontinue treatment [62].

Adaptive Phase II trial designs
In the Phase II setting, adaptive designs based on Bayesian principles can be used to adjust patient allocation to arms proving more efficacious (and away from an arm that is inferior), adding or discontinuing treatment arms or adjusting sample size. For example, Giles et al. utilized an adaptive design to evaluate three chemotherapy regimens for the treatment of acute myeloid leukemia [63]. Initially patients were randomized equally to the three arms (troxacetabine and araC; troxacetabine and idarubicin; idarubicin and araC), however, further randomizations were weighted toward treatment arms that had a higher probability of demonstrating improved efficacy. If the probability of assignment to an arm fell below a specific level the arm was dropped. After accruing five of 24 patients in total, the troxacetabine and idarubicin arm was discontinued. Subsequently the troxacetabine and araC arm was discontinued after accruing 11 patients and the trial was stopped [63]. A number of
innovative trials using adaptive designs and in some cases including biomarkers are ongoing, and some have been reported (Figure 4) [64].

■ Incorporation of biomarkers into the Phase II trial

The considerations discussed with regards to the inclusion of biomarkers in Phase I trials are equally applicable to Phase II trials. Importantly, molecularly targeted agents are likely to have efficacy in a select population of patients and it is therefore desirable to be able to evaluate the agent in the population most likely to benefit, making enrichment (for a defined biomarker) attractive in the Phase II setting. However, in the absence of a validated biomarker and robust biomarker assay, such patient selection may limit the information gleaned from the trial (e.g., the assumption that responses will be seen in patients with marker of interest and subsequent limitation of accrual may be incorrect). Most recommendations suggest that, in the absence of robust data supporting enrichment, that a strategy of unselected enrollment, evaluation of outcomes with biomarker data and, if necessary, the inclusion of additional patients with the putative biomarker, if appropriate, is more informative [53,65]. Results can then be incorporated into the design of definitive Phase III trials. There have been a number of very notable examples of the successful inclusion of a biomarker to select the population of interest in the Phase II trial, for example trastuzumab for patients with HER2-positive breast cancer [66], imatinib mesylate for patients with Philadelphia ph-chromosome-positive chronic myeloid leukemia [67] and imatinib mesylate for patients with KIT-positive gastrointestinal stromal tumor [68].

Recent consensus statements have been published from the Clinical Trial Design Task Force of the National Cancer Institute Investigational Drug Steering Committee and the MDICT [53,65]. The conclusions formulated from both initiatives were similar and emphasize the need for careful design for each agent, rather than a cookie-cutter approach where all trials are randomized (or, not). Response rate (used as a RR, continuous variable or in a multinomial rule) was felt to remain a standard end point in many instances, with PFS being advantageous especially when RR is difficult to evaluate. For combination regimens, randomized trial designs are superior to that of single-arm trials, with single-arm trials a reasonable design for single-agent studies, especially early ‘screening’ trials or where robust historic data are available. Both initiatives recommended the continued development and evaluation of novel designs and end points, including adaptive designs and seamless designs (Phase I/II or Phase II/III).

Future perspective

The last 10 years have seen an almost complete shift toward the development of drugs that target specific molecular aberrations, without, at least initially, a parallel shift in the design of early clinical trials. For some diseases, development has nonetheless been highly successful; for example, the development of imatinib mesylate for chronic myeloid leukemia and gastrointestinal stromal tumors, and trastuzumab for the treatment of HER2-positive breast cancer. However, in other diseases targeted agents have had a more modest benefit in the unselected population [41,43,69], suggesting that traditional trial designs may not be ideal or efficient and emphasizing the need for successful biomarker discovery and validation conducted in parallel with the clinical evaluation of a new agent. Enormous efforts are being made by drug developers and early clinical trialists to improve the drug development process, by developing biomarkers prior to clinical development, and developing and testing novel and efficient trial designs, tailored to the drug and the putative target tumor, to improve both efficiency and success rates. Emerging technologies evaluating multiple biomarkers in small samples, functional imaging and collection and characterization of circulating tumor cells will assist these efforts. In parallel with these advances new trial designs are being incorporated more easily into clinical development plans [70]. Such initiatives are critical to ensure that the myriad of new agents already on the horizon are evaluated efficiently.

![Figure 4. Summary of the design of a prospective biomarker-driven Phase II trial (The Biomarker-Integrated Approaches of Targeted Therapy for Lung Cancer Elimination [BATTLE] clinical trial program [64]) incorporating an adaptive design.](Image 222x135 to 582x282)
**Executive summary**

- The shift from evaluating cytotoxic chemotherapy to targeted agents, coupled with the desire to improve the efficiency of the drug development process, has led to a debate regarding the relevance of traditional oncology trial designs to these novel therapies.
- There is significant interest in model-based trial designs, the evaluation of novel end points and the effective incorporation of biomarkers in the early development of a new agent. However, this has not yet been translated into practice, with the majority of trials continue to employ traditional designs.
- Strategies including the publication of recommendations regarding the design of clinical trials and initiatives to improve early biomarker development may aid the adoption of novel designs to improve the evaluation of molecularly targeted agents.

**Financial & competing interests disclosure**

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

**Bibliography**

Papers of special note have been highlighted as:
- of interest
- Key article with regard to the challenges facing Phase I and II trials with the introduction of targeted agents.
- Key article with regard to the challenges facing Phase I and II trials with the introduction of novel agents.
- Key article with regard to the challenges facing Phase I and II trials with the introduction of targeted agents.
- Key article with regard to the challenges facing Phase I and II trials with the introduction of targeted agents.
- Key article with regard to the challenges facing Phase I and II trials with the introduction of targeted agents.

10. LoRusso PM, Boerner SA, Seymour L. Key article with regard to the challenges facing Phase I and II trials with the introduction of targeted agents.
Early-phase oncology clinical trial design

Review: Clinical Trial Methodology


Key article with regard to the challenges facing Phase I and II trials with the introduction of targeted agents.


Key article with regard to the challenges facing Phase I and II trials with the introduction of targeted agents.


66 Key article with regard to the challenges facing Phase I and II trials with the introduction of targeted agents.
67 Key article with regard to the challenges facing Phase I and II trials with the introduction of targeted agents.