The Design of Phase II Clinical Trials Testing Cancer Therapeutics: Consensus Recommendations from the Clinical Trial Design Task Force of the National Cancer Institute Investigational Drug Steering Committee

Lesley Seymour[1], **S. Percy Ivy**[2], **Daniel Sargent**[3], **David Spriggs**[4], **Laurence Baker**[5], **Larry Rubinstein**[2], **Mark J Ratain**[6], **Michael Le Blanc**[7], **David Stewart**[8], **John Crowley**[7], **Susan Groshen**[9], **Jeffrey S Humphrey**[10], **Pamela West**[11], and **Donald Berry**[8] **on behalf of the Clinical Trial Design Task Force**

[1] NCIC Clinical Trials Group, Kingston, Ontario, Canada [2] National Cancer Institute, Bethesda, Maryland [3] Mayo Clinic, Division of Biomedical Statistics and Informatics, Rochester, Minnesota [4] Memorial Sloan-Kettering Cancer Center, Department of Medicine, New York, New York [5] University of Michigan, Department of Medicine and Pharmacology, Ann Arbor, Michigan [6] University of Chicago, Department of Medicine, Section of Hematology/Oncology, Chicago, Illinois [7] Fred Hutchinson Cancer Research Center, Cancer Research and Biostatistics, Seattle, Washington [8] University of Texas M. D. Anderson Cancer [9] Keck School of Medicine, Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, California [10] Bristol-Myers Squibb, Plainsboro, New Jersey [11] The Emmes Corporation, Rockville, Maryland. Center, Houston, Texas

Abstract

The optimal design of phase II studies continues to be the subject of vigorous debate, especially with regards to studies of newer molecularly targeted agents. The observations that many new therapeutics 'fail' in definitive phase III studies, coupled with the numbers of new agents to be tested as well as the increasing costs and complexity of clinical trials further emphasizes the critical importance of robust and efficient phase II design.

The Clinical Trial Design Task Force(CTD-TF)of the NCI Investigational Drug Steering Committee (IDSC) has published a series of discussion papers on Phase II trial design in Clinical Cancer Research. The IDSC has developed formal recommendations regarding aspects of phase II trial design which are the subject of frequent debate such as endpoints(response vs. progression free survival), randomization(single arm designs vs. randomization), inclusion of biomarkers, biomarker based patient enrichment strategies, and statistical design(e.g. two stage designs vs. multiple-group adaptive designs).

While these recommendations in general encourage the use of progression-free survival as the primary endpoint, the use of randomization, the inclusion of biomarkers and the incorporation of newer designs, we acknowledge that objective response as an endpoint, and single arm designs, remain relevant in certain situations. The design of any clinical trial should always be carefully evaluated and justified based on the characteristic specific to the situation.

Corresponding author: Lesley Seymour MD FRCPC PhD, National Cancer Institute of Canada Clinical Trials Group, 10 Stuart Street, Kingston, ON, Canada K7L 3N6, Phone: (613) 533-6430, Fax: (613) 533-2941, lseymour@ctg.queensu.ca.

BACKGROUND

Many new drugs targeting molecular pathways are ready for clinical development, necessitating the use of efficient trial designs to quickly and accurately identify promising agents, while also identifying those for which all further development should be stopped. Although the development of some drugs is discontinued after phase I, the major drug development decision is generally made on the basis of phase II results. While traditional oncology trial designs using the endpoint of response and a single arm design appear to have performed this task reasonably well for cytotoxic agents, the same does not appear to be true for newer agents where high rates of tumor shrinkage may not be expected, nor for combinations of agents (such as a new drug combined with standard treatments). Certainly, success rates for phase III trials appear to be falling [i]. This has led to considerable scientific discussion, debating the advantages and disadvantages of using response vs. progression[ii] or other imaging endpoints[iii], single arm vs. randomized designs[iv], patient enrichment and biomarker endpoints[v] and optimal statistical designs, such as adaptive design or phase I/II designs.

The Investigational Drug Steering Committee (IDSC) of the National Cancer Institute Cancer Therapy and Evaluation Program (NCI CTEP) appointed a Clinical Trial Design Task Force (TF)to advise on the design of early (phase I and II) clinical trials (Table 1). In keeping with its broad mandate, TF members include IDSC members, as well as external representation from academia and the pharmaceutical industry(Table 2). Members have expertise in early clinical trial design, conduct and analyses and include statisticians, clinicians, imaging specialists, pharmacologists, biomarker experts, radiation and systems biologists, as well as patient advocates. One of the first initiatives of the TF was to coordinate a Phase II Workshop attended by TF members, other IDSC members and a number of invited experts in the field. This Phase II Workshop formed the basis of a series of discussion documents on all aspects of phase II design and conduct, published in this journal[i,ii,iii,iv,v].. Subsequently, at the request of the IDSC, the TF formulated specific recommendations for the design of phase II clinical trials (Figures 1 and 2). These recommendations, while based on the original Phase II Workshop and subsequent publications, underwent extensive discussion and revision to ensure broad applicability and acceptance and were formally approved by the IDSC.

We report here on the IDSC's recommendations regarding phase II trial design.

REVIEW AND DISCUSSION

Types of Phase II Trials

While the most common grouping of phase II trials is by design, i.e. single arm and randomized trials (Figure 3), another conceptual grouping is the ultimate aim of the trial. Is the trial to be used to 'screen' for any evidence of activity of a new drug which has recently completed phase I testing, to look for preliminary hints of activity and guide selection of tumor types for further study? Or, is the trial designed to provide a go/no-go answer to allow the conduct of a definitive registrational trial in a specific disease?. The TF considered a number of ways to categorize phase II trials, including the concept of 'screening' (e.g. IIa) versus 'decision-making' (i.e. IIb - sufficiently robust to support progression to phase III). However, the consensus opinion was that for most drugs, conducting two phase II trials in sequence would be inefficient. We recognized that there are circumstances where such a "proof of concept" approach might be reasonable, such as seeking a signal regarding the selection of tumor types for further study (for e.g. when not readily apparent from preclinical or phase I studies)or for biomarker based studies to validate a proposed mechanism of action. Ideally, however, these concepts would be embedded, possibly adaptively, in a single phase II trial. Thus, we used single-arm versus randomized studies as our primary categorization of phase II trials (Figure 3).

Selection of the Appropriate Primary Endpoint

The advantages and limitations of objective tumor response as the primary endpoint in phase II trials, and alternative endpoints such as multinomial endpoints[vi], response as a continuous variable, progression free survival, imaging, patient reported outcomes[vii] and biomarker endpoints were reviewed by Dhan[iii], Shankar[iii] and McShane[v], and will not be further addressed here. We believe that the decision on the most appropriate endpoint is critical to inform the appropriate design. While the TF accepted that response-based endpoints are still relevant for some agents (when tumor shrinkage and clinically relevant response rates are expected) and some trials, the recommendations emphasize the need to consider the inclusion of a progression-free survival primary endpoint as more informative[viii], [ix]. Overall survival is not recommended as an endpoint, as subsequent therapy may confound conclusions, and progression is usually substantially earlier, thus shortening the duration of the trial and follow up.

Randomization, Blinding and Crossover

The TF agreed that randomization was generally required to evaluate the efficacy of combinations of agents (for e.g. approved drugs and investigational agents). Randomization is usually essential for a phase II trial where progression-free survival is the most appropriate endpoint. Nonetheless, single arm designs are still appropriate for the evaluation of a monotherapy or when a well-defined historical control database is available [x]. As for any trial the design and the selected null and alternative hypotheses must be carefully justified. If a randomized design is selected, blinding of the agents (against placebo, other doses of the same agent, or other active agents) should be considered. When the primary endpoint is progression-free survival or response based, designs that allow crossover after progression maintain the integrity of the study and can provide additional data that could inform the future development of the agent

Biomarkers

Biomarkers are of considerable interest in the setting of the phase II study, but present significant challenges in their incorporation, measurement and interpretation. In most instances, the biomarker is not clinically validated as a predictive marker(of efficacy) early in the development of a new agent (i.e. at the time of the phase II trial). The IDSC's recommendations regarding biomarker utilization in early clinical trials are detailed in the current Focus Series [xi][xii][xiii][xiv][xv][xvi]. Because of these limitations, these recommendations encourage the prospective inclusion of molecular markers in phase II trials to evaluate predictive markers, but discourage prospective patient selection based on a biomarker (unless already clinically validated) except in the setting of an appropriate(and explicit)adaptive design. Phase II trials including patients with a specific biomarker but with multiple histological subtypes were considered of particular interest and may be a more efficient screening tool, especially when combined with an adaptive design.

Statistical Designs

Improved efficiencies in clinical trial design with associated shortening of development times for effective agents are highly desirable. Numerous designs have been proposed, including randomized selection designs (pick-the-winner), adaptive designs[xvii], randomized discontinuation designs[xviii], and other randomized designs[xix]. Prospectively specified adaptive designs are of particular interest in the context of phase II studies of molecularly targeted agents where biomarker identification and validation maybe emergent during the conduct of the trial, limiting the ability to select patients or identify optimal doses/schedules at the trial outset. Such adaptive designs are also particularly useful for trials including patients with a range of histologic subtypes but with biomarkers of interest. Adaptive designs in such

settings should be efficient and may result in improved precision. Despite the multiplicity of new designs that have been proposed, their inclusion in new trials has in general been modest at best[xx],[xxi]. Reasons postulated include requirements for statistical support as well as concerns about robustness, accrual and cost.

The TF is strongly supportive of designs that improve efficiency and shorten development time, such as adaptive designs, but recognized the need to continue to formally evaluate these designs to encourage wider acceptance and implementation. An ongoing initiative is the creation of a database to allow the formal testing, '*in silico*' of newer designs, in order to validate their use in future trials.

Interestingly, although formulated prior to the publication of the editorial, these recommendations are congruent with a review of phase II trials published in the Journal of Clinical Oncology[xxii], as well as with other reviews and recommendations[xxiii].

CONSENSUS RECOMMENDATIONS

A. Choosing the Appropriate Primary Endpoint

The first and critical decision point for the design of a phase II trial is based on the choice of the most appropriate primary endpoint, which should be tailored to the disease and drug(s) under investigation.

- Response-based endpoints such as those defined by RECIST, are standard, especially in early phase II trials. Other qualified biomarkers, such as molecular imaging or tumor markers, may be appropriate in select circumstances. Response based endpoints are appropriate primary endpoints if unambiguous and clinically relevant direct anti-tumor activity (such as tumor shrinkage) is hypothesized.

- If a response-based endpoint is not appropriate, especially in later phase II trials, progression-free survival is recommended as the primary endpoint. Other biomarker endpoints (such as tumor burden, tumor markers, novel imaging, tumor response, molecular biomarkers) and patient reported outcomes are always encouraged as secondary endpoints, especially in the context of studies that aim to qualify such endpoints. It is acknowledged that once qualified, these biomarker endpoints will become appropriate primary endpoints.

B(1). Study Design: Primary Endpoint is Tumor Response

Monotherapy trials—Single arm designs are acceptable. However, randomization should be encouraged to optimize dose and schedule or to benchmark activity against known active therapies.

Combination trials—With some exceptions (e.g. availability of a well validated robust control database), randomization is usually required for trials testing combinations of agents to establish efficacy. An example is standard therapy ± novel agent or combinations of novel agents.

B (2): Study Design: Primary Endpoint is Progression-Free Survival

Monotherapy or combination trials

1. With some exceptions (e.g. availability of a robust control database), randomization is required

2. For randomized trials, blinded designs are encouraged where feasible. While placebo controlled trials are challenging, they are encouraged whenever possible. Alternatives

include dose ranging, randomization vs. active controls or other novel agents, and randomized discontinuation and other crossover designs.

3. It may be informative to prospectively incorporate crossover to the standard therapy + novel agent for those patients initially assigned to the standard therapy alone although careful consideration should be given to the timing of crossover (for e.g., only after the primary endpoint has been observed). Such cross-over designs increase the access of patients to investigational agents, and also provide additional information about the activity of the study arms.

C. Patient Selection and Enrichment Strategies

Monotherapy or combination trials

1. A goal of Phase (I and) II development should be to define biomarkers predictive of efficacy and/or toxicity. Where feasible and appropriate, molecular biomarkers should be explored in order to identify subsets of patients of interest for future study.

2. Enrollment should in general not be limited by biomarker status unless there are strong confirmatory and supportive *clinical* data justifying the enrichment strategy. Adaptive statistical designs may be used to allow modification of enrollment if data suggest a biomarker is predictive.

3. In an un-selected trial (i.e. patients not defined by a biomarker), the patient population of primary interest (i.e. a cohort defined by a biomarker) should be predefined and the study powered accordingly to detect an effect in that subset.

4. Multi-disease phase II designs should be considered, especially if the objective is to test a biomarker-focused hypothesis.

D. Statistical Designs

Prospective designs that adapt to what is learned during the trial can improve the efficiency of drug development and provide greater precision. Available adaptations include stopping early, continuing longer than anticipated, dropping arms (or doses), adding arms, focusing on patient subsets, assignment of better performing treatment arms with greater probability, and seamlessly moving from phase I to II or phase II to III during a single trial.

CONCLUSIONS

The TF formulated recommendations (Figure 2) for the design of phase II trials of anticancer agents, based on consensus gained during a Workshop and extensive discussions with members of the IDSC, the TF and external experts. These recommendations were subsequently approved by the IDSC. While these recommendations in general encouraged the use of progression-free survival as the primary endpoint, the use of randomization, the inclusion of biomarkers and the use of newer designs, they acknowledge that objective response and single arm design remain relevant in appropriate circumstances. The design of any clinical trial should always be carefully evaluated and justified.

Acknowledgments

References

i. Adjei A, Christian M, Ivy P. Novel Designs and End Points for Phase II Clinical Trials. Clin Cancer Res March 15 2009;15:1866–1872.

ii. Dhani N, Tu D, Sargent DJ, Seymour L, Moore MJ. Alternate Endpoints for Screening Phase II Studies. Clin Cancer Res 2009;15:1873–1882. [PubMed: 19276273]

iii. Shankar LK, Van den Abbeele A, Yap J, Benjamin R, Scheutze S, FitzGerald TJ. Considerations for the Use of Imaging Tools for Phase II Treatment Trials in Oncology. Clin Cancer Res 15:1891–1897. [PubMed: 19276276]

iv. Rubinstein, Larry; Crowley, John; Ivy, Percy; LeBlanc, Michael; Sargent, Dan. Randomized Phase II Designs. Clin Cancer Res 2009;15:1883–1890. [PubMed: 19276275]

v. McShane LM, Hunsberger S, Adjei AA. Effective Incorporation of Biomarkers into Phase II Trials. Clin Cancer Res 2009;15:1898–1905. [PubMed: 19276274]

vi. Goffin J, Tu D. Phase II Stopping Rules That Employ Response Rates and Early Progression. J Clin Onc 2008;26(22):3715–3720.

vii. Wagner LI, Wenzel L, Shaw E, Cella D. Patient-Reported Outcomes in Phase II Cancer Clinical Trials: Lessons Learned and Future Directions 2007. J Clin Oncol :5058–5062.

viii. Fleming TR, Rothmann MD, Lu HL. Issues in Using Progression-Free Survival When Evaluating Oncology Products. J Clin Oncol 27:2874–2880. [PubMed: 19414672]

ix. Freidlin B, Korn EL, Hunsberger S, et al. Proposal for the Use of Progression-Free Survival in Unblinded Randomized Trials. J Clin Oncol 2007;25:2122–2126. [PubMed: 17513819]

x. Korn E, Liu P-Y, Lee SJ, et al. Meta-Analysis of Phase II Cooperative Group Trials in Metastatic Stage IV Melanoma to Determine Progression-Free and Overall Survival Benchmarks for Future Phase II Trials. JCO Feb 1;2008 :527–534.

xi. LoRusso PM, Boerner SA, Seymour LK. An Overview of the Optimal Planning, Design and Conduct of Phase I Studies of New Therapeutics. Clin Cancer Res 2010;16 in press.

xii. Dancey JE, Dobbin KK, Groshen S, Jessup JM, Hruszkewycz A, Koehler M, et al. Guidelines for the development and incorporation of biomarker studies in early clinical trials of novel agents. Clin Cancer Res 2010;16 in press.

xiii. Senderowicz A. What information is needed to conduct first-in-human oncology trials in the U.S. A view from a former FDA medical reviewer. Clin Cancer Res 2010;16 in press.

xiv. Rock EP, Molloy VJ, Humphrey JS. Data quality for academic investigators. Clin Cancer Res 2010;16 in press.

xv. Forster MD, Calvert AH. Performing Phase I Clinical Trials within the European Union Regulations and Guidelines. Clin Cancer Res 2010;16 in press.

xvi. Ivy SP, Siu L, Garrett-Mayer E, Rubinstein L. Approaches to phase 1 clinical trial design focused on safety, efficiency and selected patient populations: A report from the Clinical Trial Design Task Force of the National Cancer Institute Investigational Drug Steering Committee. Clin Cancer Res 2010;16 in press.

xvii. Coffey CS, Kairalla JA. Adaptive clinical trials: progress and challenges. Drugs R D 2008;9(4):229–42. [PubMed: 18588354]

xviii. Rosner GL, Stadler W, Ratain M. Randomized discontinuation design: application to cytostatic antineoplastic agents. J Clin Oncol 2002;20:4478–84. [PubMed: 12431972]

xix. Rubinstein LV, Jorn EL, Freidlin B, Hunsberger S, Ivy SP, Smith NA. Design issues of randomized phase II trials and a proposal for phase II screening trials. J Clin Oncol 2005;23:7199–206. [PubMed: 16192604]

xx. El-Maraghi RH, Eisenhauer EA. Review of phase II trial designs used in studies of molecular targeted agents: outcomes and predictors of success in phase III. J Clin Oncol 2008;26(8):1346–54. [PubMed: 18285606]

xxi. Lee JJ, Feng L. Randomized Phase II Designs in Cancer Clinical Trials: Current Status and Future Directions. J Clin Oncol 2005:4450–445. [PubMed: 15994154]

xxii. Cannistra S. Phase II Trials in Journal of Clinical Oncology. J Clin Oncol Jul 1;2009 :3073–3076. [PubMed: 19451415]

xxiii. Booth CM, Calvert HA, Giaccone G, et al. Design and Conduct of phase II studies of targeted anticancer therapy: recommendations from the task force on methodology for the development of innovative cancer therapies (MDICT). Eur J Cancer 2008;44:25–9. [PubMed: 17845846]
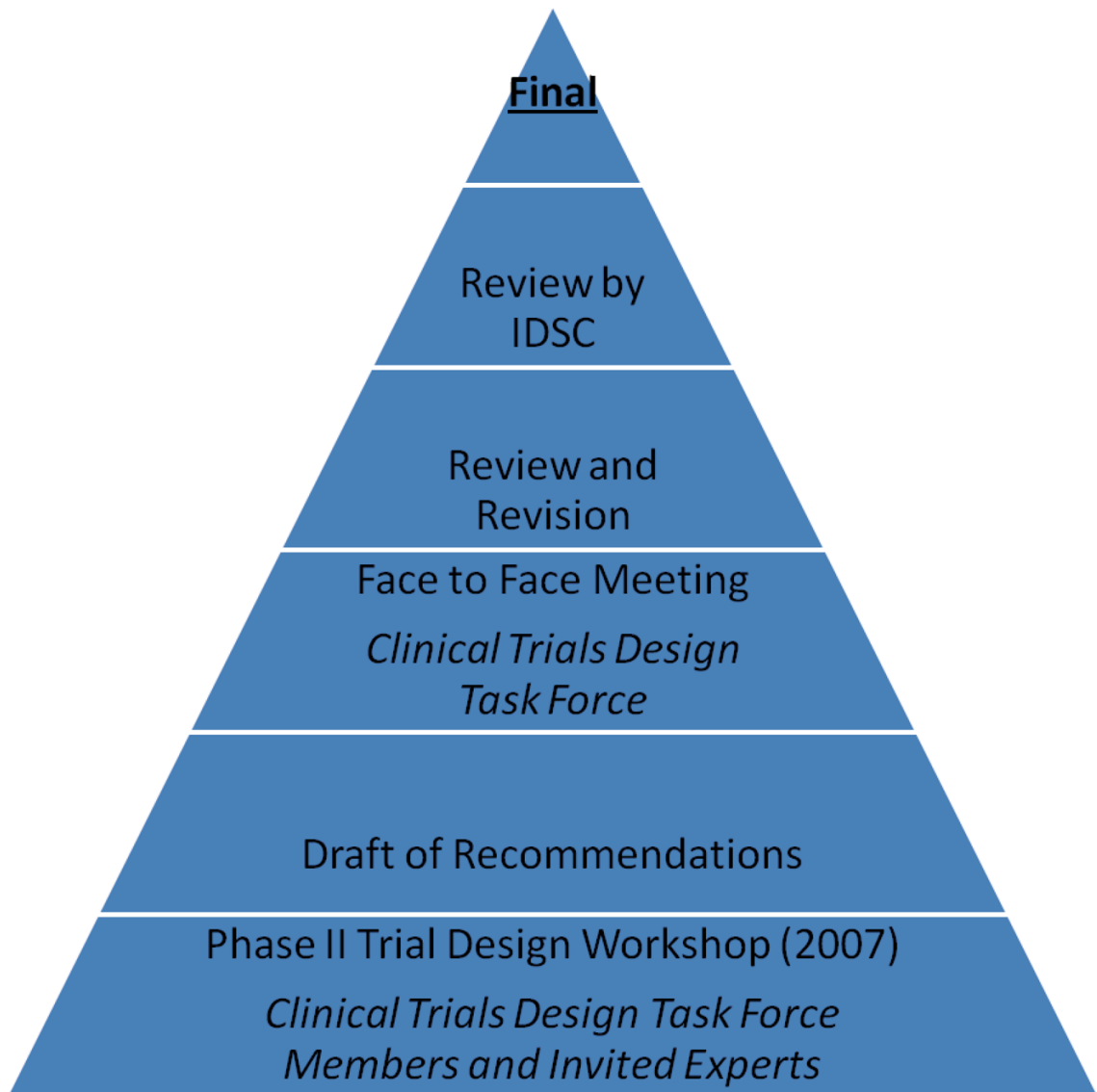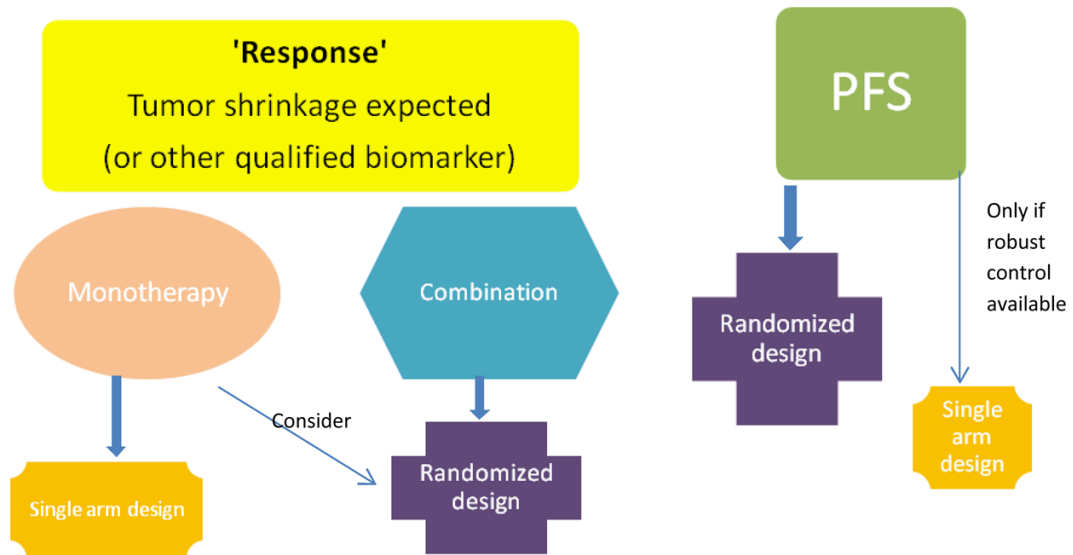
Final

Review by
IDSC

Review and
Revision

Face to Face Meeting

*Clinical Trials Design
Task Force*

Draft of Recommendations

Phase II Trial Design Workshop (2007)

*Clinical Trials Design Task Force
Members and Invited Experts*

Figure 1. Process for Development of Recommendations
IDSC – Investigational Drug Steering Committee

Figure 2.
Process/Flow or Approaches for determination of Phase 2 trial design recommendations. RCT – randomized controlled trial, 2^0 – secondary, PRO – patient related outcomes, PFS = progression-free survival.
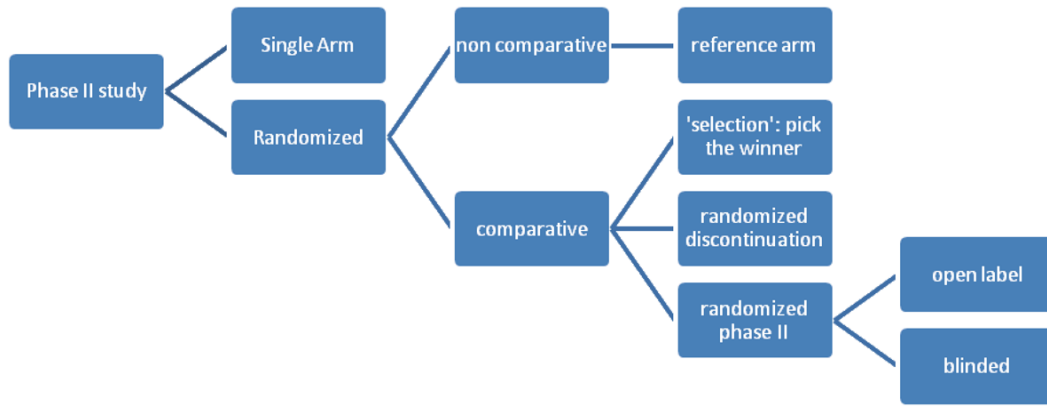
Figure 3.
Types of Phase II Studies

Table 1
Objectives of the Clinical Trial Design Task Force

IDSC – Investigational Drug Steering Committee; NCI – National Cancer Institute.

Provide guidance to the IDSC on current best practice for all aspects of early clinical trial design
Identify areas where further research and investigation are needed to improve the quality and or efficiency of early clinical trial design
Assist NCI and the IDSC or as needed with the implementation of recommendations for further research and investigation
Collaborate with other IDSC taskforces to minimize duplication of efforts
Develop and implement a plan to disseminate and publish the recommendations and guidance formulated by the taskforce and approved by the IDSC

Table 2

Past and Present Members of the Clinical Trial Design Taskforce

Chair	Lesley Seymour
Co-Chair	Donald Berry
CTEP	S Percy Ivy
Clinical/Pharmacology	A Adjei, S Yao, L Baker, S Lutzker, J Humphrey, D Stewart, A Dowlati, P Keegan, P LoRusso, M Ratain, D Spriggs, J Collins, M Grever, C Erlichmann
Statistics	J Crowley, S Groshen, M Le Blanc, L Rubinstein, D Sargent,
Imaging	L Shankar, A Shields
Advocate	D Collyar
Non voting	R Agarwal, L Minasian, P Ujhazy, L Jensen, P West
Past members	G Eckhart, S. Arbuck, M Christian, G Fyfe, R Humphrey, M Sznol, M Villalona-Calero, M Weinblatt