Published in final edited form as:

Clin Cancer Res. 2010 November 1; 16(21): 5296–5302. doi:10.1158/1078-0432.CCR-10-0669.

Analysis of the Yield of Phase II Combination Therapy Trials in Medical Oncology

Michael L. Maitland^{1,2,3}, Christine Hudoba⁴, Kelly L. Snider⁵, and Mark J. Ratain^{1,2,3}

¹Section of Hematology/Oncology - Department of Medicine, The University of Chicago, Chicago, IL

²Committee on Clinical Pharmacology and Pharmacogenomics, The University of Chicago, Chicago, IL

³Cancer Research Center, The University of Chicago, Chicago, IL

⁴Department of Surgery, The University of Chicago, Chicago, IL

⁵Pritzker School of Medicine, The University of Chicago, Chicago, IL

Abstract

Purpose—Phase II clinical studies screen for treatment regimens that improve patient care, but screening combination regimens is especially challenging. We hypothesized that recognized flaws of single arm trials could be magnified in combination treatment studies, leading to many reported positive phase II trials but a low fraction resulting practice-changing phase III trials.

Experimental Design—We searched MEDLINE and identified 363 combination chemotherapy clinical trials published in 2001 and 2002. Studies were rated as positive, negative, or inconclusive based on standardized review of abstract and text. The Web of Science Index (Thomson Reuters, NY, NY) was searched for all articles published between 2003 and October 2007 that cited at least one of these 363 published trials.

Results—Of 363 published phase II combination chemotherapy trials, 262 (0.72) were declared to be positive. Among 3760 unique subsequent citing papers, 20 reported randomized phase III trials of the same combination in the same disease as the source paper, and 10 of these resulted in improved standards of care. Estimating from these data, the likelihood that a published, positive phase II combination chemotherapy trial will result in a subsequent trial demonstrating an improvement in standard of care within five years was 0.038 [95% confidence interval- 0.016, 0.064].

Conclusions—The contributory value of combination chemotherapy phase II trials performed by 2001-02 standards is low despite the participation of more than 16,000 subjects. Future phase II studies of combination regimens require better methods to screen for treatments most likely to improve standards of care.

Keywords

combination chemotherapy; phase 2; cancer clinical trials

Corresponding Author: Michael Maitland, Section of Hematology/Oncology, MC 2115 Department of Medicine, University of Chicago, 5841 S Maryland Ave., Chicago, IL 60637, Tel: 773.834.8981, Fax: 773.702.9698, maitland@uchicago.edu.

These findings were presented, in part, at the ASCO Annual Meeting May 31, 2009 in Orlando, Florida.

INTRODUCTION

The first controlled studies to assess relative benefits of treatment in drug development are typically phase II trials. They serve as screening tests for whether or not to proceed with larger, more definitive phase III trials that establish improvements in standards of care (1). Although phase II trials in therapeutic areas other than oncology are frequently randomized (2), phase II trials in oncology have typically been single arm trials with the response rate as a common endpoint (2,3). In oncology, there has been a higher failure rate in phase III testing than other fields of medicine and the transition with the poorest performance has been from phase II to phase III (4,5). One explanation for this failure in the phase II to phase III transition is a systematic overestimate of treatment effects in phase II trials that is evident even in positive III studies (3).

Although combination therapy presents an opportunity to advance cancer care, how well typical phase II development methods screen for regimens that improve care is not known. On theoretical grounds, single arm phase II trials have fundamental flaws leading to greater uncertainty in their outcomes than randomized studies and this uncertainty could be magnified in combination treatment studies (6). Recent empiric evidence of relationships between phase II and phase III oncology trials have either focused on development of monotherapy (7) or been based on published phase III studies and the phase II trials they cited (3,8). These efforts have provided descriptive data, but have not fully evaluated all the preceding phase II publications. It is the set of all preceding phase II trials that is needed to determine a denominator for evaluating the yield of phase II trials in the context of a screening test (to determine the positive and negative predictive value).

To provide a benchmark for future analyses of the predictive value of phase II oncology combination therapy trials we began with all published phase II studies and linked these to subsequent published phase III trials by citation indexing. The purpose of this literature search and citation review was to evaluate the overall state, not drug or disease-specific state, of phase II combination anticancer agent development.¹ As would be an expected consequence of publication bias, previous investigations have demonstrated the vast majority of published phase II trials to have drawn positive conclusions (3,8). We sought to determine how many of these “positive” phase II trials led to subsequent positive phase III trials that improved standards of cancer care.

EXPERIMENTAL DESIGN

IDENTIFICATION AND EVALUATION OF PHASE II COMBINATION CHEMOTHERAPY TRIALS PUBLISHED IN 2001 – 2002 (SOURCE PAPERS)

We searched SilverPlatter ^{MEDLINE} with WinSPIRS® for all phase II combination chemotherapy clinical trials published in 2001 and 2002. The terms “Antineoplastic- Combined-Chemotherapy-Protocols-therapeutic-use” were searched as Major MESH Descriptors. That set was combined with the set of publications identified by searching Publication Type = “Clinical-Trial-Phase-II,” and the resulting set limited to the publication years 2001 and 2002 and the English language literature.

To focus the analysis on development of combination regimens within medical oncology, the authors manually excluded publications that were not phase II or combination cancer therapy studies, entailed any radiation or surgery (mostly adjuvant and neoadjuvant studies), were

¹The regimens evaluated in the phase II setting and the corresponding disease sites in which they were tested are outlined in Supplementary Table 1, available on *Clinical Cancer Research* online.

preliminary reports, or were obviously underpowered or unlikely to be relevant to any subsequent study.²

Abstracts for each of the remaining phase II trials were screened by one author (C.H.) for an explicit statement by the authors for a positive, negative, or inconclusive outcome. Any abstracts without such an explicit statement were reviewed by another author (M.L.M.) for further interpretation. If the authors' conclusions were not obvious, the original manuscript was reviewed for a statement in the introduction and discussion sections. If no explicit statement was found in the manuscript text, two authors (C.H. and M.L.M.) came to a consensus decision. Studies were rated as positive if the author(s) concluded that a combination demonstrated activity with acceptable toxicity or that the regimen warranted further study without significant modifications. Studies were documented as negative in instances where the conclusion made a clear recommendation to not pursue the combination being tested. Studies were documented as inconclusive if the conclusion stated that further evaluation was necessary to make a determination before proceeding with randomized trials.

SUBSEQUENT PAPERS CITING THE ORIGINAL PHASE II COMBINATION CHEMOTHERAPY TRIALS PUBLISHED IN 2001 – 2002 (CITING PAPERS)

The Web of Science Index (Thomson Reuters, NY, NY) was searched for all articles published between January 2003 and October 2007 that cited at least one of the source papers in the final set described above. Citing articles of specific publication types (review, editorial, letter, meeting abstract, or correction) were automatically excluded from further review. One author (C.H.) compared each remaining citing article to the original source paper to categorize the relationship between the two. Any articles with unclear relationships were reviewed by another author (M.L.M.) to determine the ultimate categorization. Citing publications were categorized as (a) a subsequent trial of the same combination or (b) not a subsequent trial of the same combination (if it was either not a clinical trial or if it was a clinical trial testing anything other than the same drug combination that was tested in the original source paper).

Subsequent trials of the same combination were further subcategorized: (a) relevant randomized phase III trial (if it was a phase III trial testing the same combination in the same disease as the source paper); (b) subsequent randomized phase II, or biomarker or PK-based trial of the same combination in the same disease as the source paper; (c) altered dose, schedule, or setting of the same combination in the same disease as the source paper; (d) same combination as the source paper but in a different disease; (e) phase I trial of the same combination as the source paper in any disease setting; (f) miscellaneous (not relevant phase III trials including those testing the same combination as the source paper but in a different disease, and trials for which the primary difference between the citing and source paper was the number of subjects enrolled, the country in which the trial was conducted, or a specific focus on a subpopulation such as elderly subjects).

For citing papers determined to be relevant phase III trials, the manuscripts were reviewed by one author (M.L.M.). A phase III study was rated positive if the results led to a recognized change in, or acceptable addition to, standard of care (Table 1). To avoid bias against successful phase III trials, this could include not only a change in the label for one of the study agents or use of the regimen as a reference regimen in subsequent clinical trials, but also consideration of the regimen as acceptable in any developed health care system. The remaining studies were determined to be negative or equivocal.

---

²Supplementary Table 1 lists all excluded articles under each category, available on *Clinical Cancer Research* online

RESULTS

SOURCE AND CITING PAPERS

Our MEDLINE search identified a total of 575 phase II combination chemotherapy clinical trials published in 2001 and 2002 (Figure 1). We excluded 212 publications that were not consistent with the focus of our analysis, leaving a total of 363 trials for the main analysis; 179 (49%) were published in 2001 and 184 (51%) were published in 2002. Only 22 (6%) of these phase II trials were randomized. The Web of Science Index search for all articles published between 2003 and October 2007 that cited at least one of the original 363 published trials identified 3,760 unique citing papers. Exclusion of the non-investigational publication types yielded 2,741 unique citing papers. After review of each citing article with respect to the source papers that they matched, a total of 3,801 citing papers were categorized, as there were instances in which a citing paper cited more than one of the source papers from our final set.

DISEASE DISTRIBUTION AND OUTCOMES

Lung, breast, and colorectal cancers were the three most studied disease sites (Figure 2). A total of 16,008 subjects participated in the published phase II trials. Of the 363 phase II trials, according to explicit statements by the authors for a positive, negative, or inconclusive outcome, 262 (72%) were designated positive, 74 (20%) as negative, and 27 (7%) were inconclusive (percentages do not sum to 100% due to rounding). Randomized phase II studies were more likely to draw a negative conclusion than those that were non-randomized (45% vs. 19%, $p = 0.004$) (Figure 3).

Our search yielded 20 unique relevant randomized phase III trials. Ten of these were positive, 7 were negative, 2 were equivocal, and 1 was a non-inferiority trial (Table 1). Given that 10 positive unique relevant randomized phase III trials resulted from the collection of phase II trials published in 2001–2002 and 262 were declared to be positive, the estimated positive predictive contributory value (PPCV) for phase II trials of combination anticancer agents is 0.038.

DESIGN OF PHASE II TRIALS CITED BY SUBSEQUENT PHASE III TRIALS

To describe features of “true positive” phase II trials (the declared positive studies that resulted in practice-improving phase III trials) that should be replicated in future studies, we evaluated in detail the study characteristics of the unique source trials, cited by the 20 randomized phase III trials described above (Table 2). The table displays the 10 positive phase III trials (top half table in alphabetical order) and the 7 negative phase III trials (bottom-half table on grey field). To avoid complicating the comparisons the 2 equivocal and 1 non-inferiority trial are not displayed or included in the analysis. Positive phase III trials were more likely than negative phase III trials to be based on phase II trials with a stated, pre-specified null hypothesis (7/10 vs. 0/7; $p = 0.01$ Fisher’s exact test). Notably, 3 of the single arm source trials also reported a pre-specified alternative hypothesis and the measured endpoint was consistent with that alternative hypothesis. These standard elements of a rigorous single arm trial design were infrequently reported or achieved in studies that led to phase III trials.

DISCUSSION

The challenges facing cancer therapeutics development in the phase II setting have been well recognized, but only recently has there been growing consensus that these trials could be improved by consistent use of comparator arms (9,10). Testing of combination regimens in single arm studies is particularly problematic. Although single arm studies (with comparison to inferred historical controls) have been conventional for phase II cancer therapeutics development they are uncommon in other fields of medicine (2). In the case of testing single

agents, the comparison is with the natural history of the disease, but in combination therapy including the addition of a second or third agent, the comparator is the other agents in the regimen for which the response rate is typically non-zero. Testing of combinations introduces a set of additional concerns about dosing and safety of each of the components of the new combination and these concerns have been ignored even in recently published studies (11). Consequently, we hypothesized that although single arm trials would be commonly used in combination therapy development, they would be particularly ineffective in screening for treatments that would actually change standards of care. Randomized phase II trials of combination therapy were 15-fold less common than single-arm trials during this study period, so we could not perform a rigorous analysis of their value relative to the single-arm trials.

Previous analyses of phase II cancer therapy trials have attempted to identify predictive factors for success in phase III (3,7,8), and either focused on development of new single agents or began with published phase III trials and then identified phase II trials cited by the phase III trials. To describe fully the spectrum of combination therapy phase II trials in medical oncology and the subsequent results of these studies we took an alternative approach. We performed a cross-sectional analysis of all published phase II trials in a two-year period and then during a 5 year follow-up period identified all publications that cited these “source” trials. In these analyses we have approached phase II trials as screening diagnostic tests for regimens that would warrant more rigorous and expensive testing in confirmatory phase III trials. As single arm designs were overwhelmingly favored during the 2001–2002 publication period, it is not surprising that some led to positive phase III trials. In the sample of phase II trials that were associated with subsequent phase III trials, the true positive single arm studies were designed and interpreted more rigorously than the false positive studies. Through this approach we confirmed the vast majority of published phase II combination therapy studies to have single arm design, to have interpreted their data to warrant further investigation, and to have a low positive predictive contributory value. Despite the participation of over 16,000 cancer patients, few new standards of care were achieved.

This investigation establishes a benchmark by which efforts to improve the process of combination cancer therapeutics could be measured. If studies are well-designed, the positive predictive value of the studies will only be adversely affected if trials fail to determine treatments to be unlikely to improve standards of care. These data set a very low bar- a positive predictive contributory value of less than 4%. Although this is admittedly an estimate, a number this low raises several questions about the limitations of this study. We have focused solely on the published literature and a cross section of only two years: 2001 and 2002. This was an arbitrary selection based on when we began this study, in order to ensure nearly 5 year follow-up. Using the 5 year cut-off might have eliminated a few subsequent trials, but this seemed an appropriate timeframe to capture most of the relevant subsequent phase III trials. Few molecularly targeted agents were tested in the source phase II trials, but during this time some of these agents skipped the phase II combination development process entirely and then failed in phase III,(12,13) so these phase III failures were not counted against the phase II screening process. Our calculation of the positive predictive contributory value includes both “positive” phase II trials that did not proceed to phase III and those that led to negative phase III trials. We considered both of these outcomes to constitute a “false positive” phase II screening test. We provided no weighting in our analysis for publications in higher impact journals or those enrolling more patients than others. As this is an initial benchmarking estimation study, we thought it would be fair to consider each patient’s election to participate in any phase II combination therapy trial to be of equal value and analyzed in the units of academic productivity- the published study. Arguably the most important subset of phase II trials in this analysis were those that were directly associated with subsequent phase III trials. As expected, the data suggest that true positive phase II trials are conducted and interpreted with greater rigor and discipline than false positive trials.

We do not conclude that if a phase II trial does not lead to a positive phase III trial that it is a waste of resources. To the contrary, the important issue is that the total of all phase II trial activity leads to more rapid progress in standards of care. These data highlight the importance of improving the positive predictive contributory value of phase II trials. A clear implication is that there is much to gain by increasing the threshold for declaring a trial to be positive. Although there were few randomized phase II trials (6%), these trials were clearly more likely to conclude that a new combination regimen was insufficiently better than the comparator arm to warrant further study. Notably rather than subsequent phase III trials, 46 unique subsequent trials were randomized phase II or biomarker development studies and these efforts could be a very sensible approach to developing new, effective combination regimens.

These data verify the woeful state of combination cancer therapy development in the recent past. Empirically, this problem has been increasingly recognized, and some obvious solutions have begun testing to improve the productivity of the entire phase II development enterprise in oncology (9,10,14). There is new evidence that one suggestion, increasing the size of single arm studies, will not improve the predictive value of phase II studies as screening tests (15). So, the elimination of single arm combination therapy studies as a convention is an important first step. However, the increasing number of novel cancer therapeutics available for testing means the potential combinations increase by a permutation function. To exploit maximally this opportunity in cancer therapeutics will require serial innovations in the selection, phase I, and phase II development of these treatments to help the greatest number of patients in the shortest period of time.

Statement of translational relevance

Combining drugs can be an effective approach to improving the therapeutic index for treating disease. Phase II clinical trials are the typical setting in which new approaches are first tested for evidence of clinical effects. In oncology these studies have tended to enroll fewer patients and more frequently not to use control arms than in other fields of medicine. Because testing combinations entails additional variables on dose, toxicity, and therapeutic activity, the design of phase II trials is important to subsequent success. This investigation estimated the yield of previously completed phase II clinical trials of combination chemotherapy for advances in cancer care. Although thousands of patients enrolled in this cross-section of phase II trials, few advances affecting routine care of cancer patients were made.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

KLS was supported by a Calvin Fentress Research Fellowship from the University of Chicago. MLM was supported by mentored career development award K23CA124802.

REFERENCES

1. Temple R. Current definitions of phases of investigation and the role of the FDA in the conduct of clinical trials. *Am Heart J* 2000;139:S133–S135. [PubMed: 10740119]
2. Michaelis LC, Ratain MJ. Phase II trials published in 2002: a cross-specialty comparison showing significant design differences between oncology trials and other medical specialties. *Clin Cancer Res* 2007;13:2400–2405. [PubMed: 17438099]

3. Zia MI, Siu LL, Pond GR, Chen EX. Comparison of outcomes of phase II studies and subsequent randomized control studies using identical chemotherapeutic regimens. *J Clin Oncol* 2005;23:6982–6991. [PubMed: 16192585]
4. Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? *Nat Rev Drug Discov* 2004;3:711–715. [PubMed: 15286737]
5. DiMasi JA, Grabowski HG. Economics of new oncology drug development. *J Clin Oncol* 2007;25:209–216. [PubMed: 17210942]
6. Ratain MJ, Sargent DJ. Optimising the design of phase II oncology trials: the importance of randomization. *Eur J Cancer* 2009;45:275–280. [PubMed: 19059773]
7. 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;26:1346–1354. [PubMed: 18285606]
8. Chan JK, Ueda SM, Sugiyama VE, et al. Analysis of phase II studies on targeted agents and subsequent phase III trials: what are the predictors for success? *J Clin Oncol* 2008;26:1511–1518. [PubMed: 18285603]
9. Adjei AA, Christian M, Ivy P. Novel designs and end points for phase II clinical trials. *Clin Cancer Res* 2009;15:1866–1872. [PubMed: 19276272]
10. Cannistra SA. Phase II trials in journal of clinical oncology. *J Clin Oncol* 2009;27:3073–3076. [PubMed: 19451415]
11. Hamberg P, Verweij J. Phase I Drug Combination Trial Design: Walking the Tightrope. *J Clin Oncol* 2009;27:4441–4443. [PubMed: 19704054]
12. Herbst RS, Giaccone G, Schiller JH, et al. Gefitinib in combination with paclitaxel and carboplatin in advanced non-small-cell lung cancer: a phase III trial--INTACT 2. *J Clin Oncol* 2004;22:785–794. [PubMed: 14990633]
13. Herbst RS, Prager D, Hermann R, et al. TRIBUTE: a phase III trial of erlotinib hydrochloride (OSI-774) combined with carboplatin and paclitaxel chemotherapy in advanced non-small-cell lung cancer. *J Clin Oncol* 2005;23:5892–5899. [PubMed: 16043829]
14. Seymour L, Ivy SP, Sargent D, et al. 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. *Clin Cancer Res* 2010;16:1764–1769. [PubMed: 20215557]
15. Tang H, Foster NR, Grothey A, Ansell SM, Goldberg RM, Sargent DJ. Comparison of Error Rates in Single-Arm Versus Randomized Phase II Cancer Clinical Trials. *J Clin Oncol* 2010;28:1936–1941. [PubMed: 20212253]

Figure 1.
Flow of literature search and curation

Figure 2.
Distribution of phase II combination studies by disease site

Figure 3.
Distribution of investigators' conclusions for phase II combination studies

Table 1

Combination phase III trials leading to new acceptable standard of care treatment

First Author	Journal of Publication	Year	Disease	Combination
Comella	Ann Oncol	2005	Colorectal	OXAFUFU
Eichhorst	Blood	2006	CLL	Fludarabine + Cyclophosphamide
Falcone	J Clin Oncol	2007	Colorectal	FOLFOXIRI
Flinn	J Clin Oncol	2007	CLL	Fludarabine + Cyclophosphamide
Habermann	J Clin Oncol	2006	Lymphoma (diffuse large B-cell)	R-CHOP
Hitt	J Clin Oncol	2005	Head and Neck	Paclitaxel + Cisplatin + Fluorouracil
Long	J Clin Oncol	2005	Cervix	Cisplatin + Topotecan
Petrylak	New Engl J Med	2004	Prostate	Docetaxel + Estramustine
Rifkin	Cancer	2006	Multiple Myeloma	Pegylated Liposomal Doxorubicin + Vincristine + Dexamethasone
Rothenberg	J Clin Oncol	2003	Colorectal	FOLFOX4

Table 2

Design of phase II trials cited by subsequent phase III trials

Phase III Trial (author/PMID)	Source Phase II Trial (author/PMID)	Reported Null Hypothesis	Reported Alternative Hypothesis	Endpoint c/w Alternative Hypothesis
Comella 15837702	Ravaioli 12011134	•	•	•
Eichhorst 16219797	Hallek 11529853	•		
Falcone 17470860	Souglakos 12039926	•		
Flinn 17283364	Hallek 11529853	•		
Habermann 16754935	Vose 11208830	•		
Hitt 16275937	Hitt 12377658	•	•	•
Long 15911865	Fiorica 11925125			
Petrylak 15470214	Savarese 11331330	•	•	•
Rifkin 16404741	Hussein 12412170			
Rothenberg 12775730	Maindrault-Goebel 11334725			
<i>Alberola 12947054</i>	<i>Laack 11916547</i>			
<i>Gibson 15908667</i>	<i>Basaran 12017379</i>			
<i>Gridelli 12837810</i>	<i>Palmeri 11557121</i>		•	•
<i>Laack 15197195</i>	<i>Laack 11290433</i>			
<i>Okamoto 17579629</i>	<i>Quoix 11521802</i>			
<i>Rocha Lima 15365074</i>	<i>Rocha Lima 11870159</i>			
<i>Zielinski 15735116</i>	<i>Conte 11688520</i>			