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

Michael L. Maitland1,2,3, Christine Hudoba4, Kelly L. Snider5, and Mark J. Ratain1,2,3

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 with a low fraction resulting in 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 a standardized review of abstract and text. 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 these 363 published trials.

Results: Of 363 published phase II combination chemotherapy trials, 262 (72%) were declared to be positive. Among 3,760 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 showing 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 done by 2001-2002 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. Clin Cancer Res; 16(21); 5296–302. ©2010 AACR.
Yield of Phase II Combination Trials

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.

Abstracts for each of the remaining phase II trials were screened by one author (C. Hudoba) 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. Maitland) 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. Hudoba and M.L. Maitland) came to a consensus decision. Studies were rated as positive if the author(s) concluded that a combination showed 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.

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 preliminary reports, or were obviously underpowered or unlikely to be relevant to any subsequent study (Supplementary Table S1).

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. Hudoba) compared each remaining citing article with the original source paper to categorize the relationship between the two. Any articles with unclear citing article with the original source paper to categorize the relationship between the two. Any articles with unclear relationships were reviewed by another author (M.L. Maitland) 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 pharmacokinetic-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. Maitland). 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.

**Results**

**Source and citing papers**

Our MEDLINE search identified a total of 575 phase II combination chemotherapy clinical trials published in 2001 and 2002 (Fig. 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 January 2003 and October 2007 that cited at least one of the original 363 published trials identified 3,760 unique citing papers. Exclusion of the noninvestigational publication types yielded 2,741 unique citing papers. After a review of each citing article with respect to the source papers that it 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 (Fig. 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%) were 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 nonrandomized (45% versus 19%, \( P = 0.004; \) Fig. 3).

Our search yielded 20 unique relevant randomized phase III trials. Ten of these were positive, seven were negative, two were equivocal, and one was a noninferiority 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 was 0.038.

**Design of phase II trials cited by subsequent phase III trials**

To describe the 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 of the table in alphabetical order) and the 7 negative phase III trials (bottom half of the table on grey field). To avoid complicating the comparisons, the two equivocal and one noninferiority trials 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, prespecified null hypothesis (7 of 10 versus 0 of 7; \( P = 0.01 \) Fisher’s exact test). Notably, three of the single-arm

---

**Table 1. Combination phase III trials leading to new acceptable standard of care treatment**

<table>
<thead>
<tr>
<th>First Author</th>
<th>Journal of publication</th>
<th>Year</th>
<th>Disease</th>
<th>Combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comella</td>
<td>Ann Oncol</td>
<td>2005</td>
<td>Colorectal</td>
<td>OXAFU</td>
</tr>
<tr>
<td>Eichhorst</td>
<td>Blood</td>
<td>2006</td>
<td>CLL</td>
<td>Fludarabine + cyclophosphamide</td>
</tr>
<tr>
<td>Falcone</td>
<td>J Clin Oncol</td>
<td>2007</td>
<td>Colorectal</td>
<td>FOLFOXIRI</td>
</tr>
<tr>
<td>Flinn</td>
<td>J Clin Oncol</td>
<td>2007</td>
<td>CLL</td>
<td>Fludarabine + cyclophosphamide</td>
</tr>
<tr>
<td>Habermann</td>
<td>J Clin Oncol</td>
<td>2006</td>
<td>Lymphoma (diffuse large B-cell)</td>
<td>R-CHOP</td>
</tr>
<tr>
<td>Hitt</td>
<td>J Clin Oncol</td>
<td>2005</td>
<td>Head and neck</td>
<td>Paclitaxel + cisplatin + fluorouracil</td>
</tr>
<tr>
<td>Long</td>
<td>J Clin Oncol</td>
<td>2005</td>
<td>Cervix</td>
<td>Cisplatin + topotecan</td>
</tr>
<tr>
<td>Petrylak</td>
<td>New Engl J Med</td>
<td>2004</td>
<td>Prostate</td>
<td>Docetaxel + estramustine</td>
</tr>
<tr>
<td>Rifkin</td>
<td>Cancer</td>
<td>2006</td>
<td>Multiple myeloma</td>
<td>Pegylated liposomal doxorubicin + vincristine + dexamethasone</td>
</tr>
<tr>
<td>Rothenberg</td>
<td>J Clin Oncol</td>
<td>2003</td>
<td>Colorectal</td>
<td>FOLFOX4</td>
</tr>
</tbody>
</table>

Abbreviations: CLL, chronic lymphocytic leukemia; OXAFU, oxaliplatin, folinic acid, fluorouracil; FOLFOXIRI, fluorouracil, folinic acid, oxaliplatin, irinotecan; R-CHOP, rituximab, cyclophosphamide, doxorubicin, vincristine, prednisone; FOLFOX4, fluorouracil, folinic acid, oxaliplatin.
source trials also reported a prespecified alternative hypothesis and the measured end point 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). The testing of combination regimens in single-arm studies is particularly problematic. Although single-arm studies (with comparison with 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 consists of the other agents in the regimen for which the response rate is typically nonzero. The testing of combinations introduces a set of additional concerns about the dosing and safety of each of the components of the new combination and these concerns have been ignored even in recently published studies.

---

**Fig. 1. Flow of literature search and curation.**

Original MEDLINE search

Phase II combination chemotherapy trials

Published 2001 - 2002

(n = 575 published "source" trials)

Source trials for main analysis

n = 363

Total unique citing papers

n = 3, 760

Minus type of work:
correction, editorial, letter, meeting abstract, review

n = 2, 741

Corresponds to total matched citing papers

n = 3, 801

Unique relevant randomized phase III trials

(n = 20)

Unique relevant positive randomized phase III trials

(n = 10)

Not a subsequent trial of the same combination

n = 3, 274

Subsequent trial of the same combination

n = 504

Subsequent randomized phase II, or biomarker or Pharmacokinetic-based trial

n = 63

Altered dose/schedule/setting

n = 300

Different disease

n = 68

Phase I

n = 50

Miscellaneous
(larger n, trial conducted in different country, subpopulation focus, not relevant phase III)

n = 23
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 do 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 did a cross-sectional analysis of all published phase II trials in a two-year period, and then during a five-year follow-up period identified all publications that cited these “source” trials. In these analyses we 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 <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, to ensure nearly five-year follow-up. Using the five-year cutoff 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, i.e., the published study. Arguably the most important subset of phase II trials in this analysis consisted of 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 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). Thus, 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 and in phase I and phase II development of these treatments to help the greatest number of patients in the shortest period of time.

### Table 2. Design of phase II trials cited by subsequent phase III trials

<table>
<thead>
<tr>
<th>Phase III trial (author/PMID)</th>
<th>Source phase II trial (author/PMID)</th>
<th>Reported null hypothesis</th>
<th>Reported alternative hypothesis</th>
<th>End point c/w alternative hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comella 15837702</td>
<td>Ravaiali 12011134</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Eichhorst 16219797</td>
<td>Hallek 11529853</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Falcone 17470860</td>
<td>Souglakos 12039926</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Flinn 17283364</td>
<td>Hallek 11529853</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Habermann 16754935</td>
<td>Vose 11208830</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Hitt 16275937</td>
<td>Hitt 12377658</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Long 15911865</td>
<td>Fiorica 11925125</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Petrylak 15470214</td>
<td>Savaresse 11331330</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Rifkin 16404741</td>
<td>Hussein 12412170</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Rothenberg 12775730</td>
<td>Mainauidt-Goebel 11334725</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Alberola 12947054</td>
<td>Laack 11916547</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Gibson 15908667</td>
<td>Basarar 12017379</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Grideli 12837810</td>
<td>Palmeri 11557121</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Laack 15197195</td>
<td>Laack 11290433</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Okamoto 15799629</td>
<td>Quoix 11521802</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Rocha Lima 15365074</td>
<td>Rocha Lima 11870159</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Zielinski 15735116</td>
<td>Conte 11688520</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
</tbody>
</table>
Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Grant Support

K.L. Snider was supported by a Calvin Fentress Research Fellowship from the University of Chicago. M.L. Maitland was supported by mentored career development award K23CA124802.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 03/16/2010; revised 08/20/2010; accepted 08/24/2010; published OnlineFirst 09/13/2010.

References

Clinical Cancer Research

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

Michael L. Maitland, Christine Hudoba, Kelly L. Snider, et al.

Clin Cancer Res 2010;16:5296-5302. Published OnlineFirst September 13, 2010.

Updated version
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-10-0669

Cited articles
This article cites 15 articles, 12 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/16/21/5296.full.html#ref-list-1

Citing articles
This article has been cited by 15 HighWire-hosted articles. Access the articles at:
/content/16/21/5296.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.