Efficient Designs for Testing New Agents and Regimens*  

Andrea Ardizzoni, MD

Current clinical chemotherapy trial procedures in non-small cell lung cancer (NSCLC) require large numbers of patients and last up to 10 years before providing results that are often disappointing. Different designs are available that offer the advantages of smaller numbers of patients, shorter duration, and earlier identification of ineffective or highly active agents and regimens. An alternative method is proposed, combining phase II screening, randomized phase II studies, and multicenter, comparative, randomized phase III trials. Issues under discussion include the combined study in phase II of different disease stages and the need for recruitment to larger trials.  

(CHEST 1996; 109:835-868)

A new active agent that has shown promise in preclinical studies is subject to a long clinical trial process. Initially, three to five small (20 to 30 patients per study), single-center, phase II studies are commonly carried out, taking 1 to 2 years, to screen for antitumor activity and distinguish between active and inactive new drugs. A good response rate (>20%) permits progress toward registration, which requires a broader multicenter, usually international, phase II study of more than 100 patients taking another 1 to 2 years. A somewhat lower response rate is usually achieved in these trials, because radiographs showing the response are reviewed centrally and because, as these studies are large and multicentered, patients are less carefully selected.

Phase II studies of combinations with known active agents (eg, cisplatin) can then begin. These studies are usually single center, small (30 to 40 patients) and last 1 to 2 years. The primary end point is again to assess the toxicity and response rate of the combination. In these studies, a drug combination is considered promising when response rate exceeds 30 to 40%.

Finally, randomized, multicenter phase III trials begin comparing promising drug combinations, including the new agent, with a standard chemotherapy regimen. The primary end point is usually survival. As 400 to 500 patients are required in each study to obtain meaningful results in terms of survival differences, this is likely to take another 4 to 5 years. Response rates will probably now be 20 to 40%, with improvements in survival uncommon.

The whole process takes 8 to 10 years, involves 1,000 patients, and sometimes gives only very poor conclusions. For these reasons, alternative methods are required that are faster, use fewer patients and reject ineffective agents earlier in the process.

ALTERNATIVE TECHNIQUES

Screening of New Agents

The usefulness of the phase II process for the screening of activity is based on the assumption that the objective response is a surrogate marker of clinical efficacy; the higher the response rate, the higher the chance of obtaining a clinical benefit in terms of survival and symptom control. As most of the new drugs that have become available for non-small cell lung cancer (NSCLC) have been shown to be inactive, it is important for an efficient drug-clinical screening process to reject ineffective agents quickly and to identify effective agents without overestimating their activity.

Simon1 developed a phase II design that meets these requirements. This is a two-stage trial design with n1 patients in the first stage and n patients in total. Two proportions of response, p0 and p1, are set; p0 identifies the level of activity of no clinical interest and p1 the target activity level of clinical interest. For a single agent in NSCLC, p0 could be, for example, 5 to 10% and p1 20 to 25%. If the observed response rate at the end of the first stage is r1/n1 or less (r=number of responding patients), then the trial is terminated and the drug is rejected as being of little interest. Otherwise, accrual continues to a total of n patients. At the end of the second stage, the drug is rejected if the observed response rate is r/n or less (Table 1). If the target level of activity is reached, combination studies are warranted.

This process offers the advantages of short duration and relatively small numbers of patients. A maximum of 20 patients in the first stage and of 40 patients overall are generally enough. To obtain even more reliable

*From the Istituto Nazionale per la Ricerca sul Cancro, Genova, Italy. Reprint requests: Dr. Ardizzoni, Department of Medical Oncology I, Istituto Nazionale per la Ricerca sul Cancro, Viale Benedetto XV, 1016132 Genova, Italy
results, the study can be multicenter, should have central registration of patient enrolment, central review of responses, and eligibility criteria as strict as possible (i.e., previously untreated, stage IV, performance status less than 2, measurable disease).

**Randomized Phase II Trials of New Regimens**

When a new active agent has been identified, the next step is to incorporate the drug into combination regimens. These new regimens are usually tested in phase II studies. If the response rate is found to be higher than that which would usually be achieved with standard regimens, phase III randomized comparative studies are planned to verify the superiority of the new regimen over the standard regimen. Unfortunately, almost none of the phase III studies conducted in the last 20 years have shown a clear-cut superiority of a specific regimen over another. This is thought to be due to the fact that results of phase II studies tend to overestimate the efficacy of the regimens tested. In fact, the response rate in phase III studies is usually lower than that of phase II trials. Most phase III trials would not even have been started if the actual response rate of a certain regimen had been more precisely identified during phase II studies. Randomized phase II trials might be the solution to this problem.

Randomized, noncomparative phase II trials can test the activity of several combinations or single agents simultaneously. Randomization is usually central, offering two important advantages. Patients’ eligibility is checked and the results for all registered patients are reported. This contrasts with normal phase II studies in which patients are often unreported. Furthermore, patient selection is reduced, which helps eliminate the bias caused by the unconscious selection of the most suitable cases. Moreover, similar to phase III trials, the results are assessed using an “intention to treat analysis,” therefore avoiding “exclusion biases.” Other advantages include the possibility of stopping inactive regimens early. Finally, response is reviewed centrally, which is of paramount importance in reducing interobserver variability (Table 2). Phase II randomized trials can be conducted consecutively without interruption. As soon as an arm has reached the target accrual, the regimen under test in that arm can be replaced by a new one without stopping the study. The results of the closed arm can be analyzed separately.

Suggested eligibility criteria for randomized phase II trials include measurable disease, performance status less than 2 according to the World Health Organization score system, absence of prior chemotherapy, and a clinical cancer (extension) above stage II.

For example, the Italian Task Force for Lung Cancer (FONICAP) has completed two sequential randomized phase II studies testing new chemotherapy programs with a standard reference regimen in metastatic NSCLC (Table 3). The first examined the mitomycin-ifosfamide-cisplatin (MIP) regimen vs MVP plus interferon vs cisplatin plus carboplatin. The response rate for MIP was only 14%, compared with the 40 to 70% obtained in the phase II studies and the 30 to 40% in phase III.

In the second trial, mitomycin-vindesine-cisplatin (MVP) was used instead of MIP as a reference regimen. Again the response rate with MVP was substantially lower than the one achieved in nonrandomized phase II studies and closer to the response rates reported in randomized phase III.

These results indicate that randomized phase II studies give results closely similar to those from phase III studies, and therefore represent a more reliable tool on which to base the decision to embark on a comparative phase III study (i.e., response rate above 30 to 40% in the setting of randomized phase II trial). Randomized phase II trials can make some phase III studies redundant, removing the need for thousands of

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Advantages of Randomized Phase II Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central registration</td>
<td>Check eligibility</td>
</tr>
<tr>
<td>All patients enrolled are reported on</td>
<td>Reduced patient selection</td>
</tr>
<tr>
<td>Central review of results</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Response Rate According to Study Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regimen</td>
<td>Phase II, %</td>
</tr>
<tr>
<td>MIP</td>
<td>40-70</td>
</tr>
<tr>
<td>MVP</td>
<td>40-60</td>
</tr>
</tbody>
</table>

*FONICAP*—Italian Task Force for Lung Cancer.
patients to participate in studies providing little additional information. This approach helps avoiding disappointing results from phase III trials following apparently promising phase II studies.

It was suggested that intermediate end points other than response rate (eg, complete resectability and pathologically complete responses in the neoadjuvant setting) might accelerate the process, allowing new agents and combinations to make a clinical contribution even sooner.

Some participants questioned the choice of the reference regimen. It was pointed out that the reference regimen is used to monitor the conduct of the study, not as the basis for comparisons with the new regimen. For example, unusual response rates with the reference regimen may suggest irregularities in the study.

**Summary of the Alternative Method of Drug Testing in NSCLC**

This proposal for an alternative method aims to require fewer patients and less time than the current approach. The target response rates are made rather demanding because the objective is the identification of distinctly better agents rather than agents that have merely similar performance to those already available.

Step 1 is single-agent screening using a two-stage Simon’s design, involving about 20 to 40 patients and lasting 1 year. The agent is accepted if the response rate is higher than 20%. The use of a tissue bank is recommended so that biologic markers can eventually be correlated with response.

Step 2 is a multicenter, randomized phase II combination study incorporating a standard reference regimen. This requires 100 to 150 patients and lasts 1 to 2 years. The regimen is accepted if the response rate is higher than 30 to 40%.

Step 3 is a multicenter comparative randomized phase III trial. This requires at least 500 patients and takes 4 to 5 years.

This approach is faster (5 to 7 vs 8 to 10 years) and requires fewer patients (<700 vs >800) than current approaches. There is also the advantage that more agents and combinations are likely to be rejected early in the process and the start of almost certainly negative phase III studies can be prevented.

**Should Different Disease Stages Be Studied Separately or Together in Phase II Studies?**

Participants at the workshop offered a range of views on the desirability of studying different stages of disease (eg, III and IV) in phase II trials. There was a view that eligibility for phase II trials should be open so that patients can reflect real world experience and phase III eligibility. This would help increase accrual rates. However, it is well known that response rate decreases with the increase of stage and therefore response rates should be analyzed separately by stage and the number of patients increased if more than one stage is included in the study. Comparisons may be difficult if one study has, for example, 70% stage III patients and another has 70% stage IV patients. Other participants argued that tumor size is a better predictor of likely response than the stage of the disease. It was believed that the question of stage III vs stage IV or bulky vs nonbulky within studies was perhaps a numeric or design problem: given sufficient number of patients or adequate stratification, the problem could be avoided. There was some disquiet that the definitions of the various stages have not yet been standardized, which further complicates comparisons. It is possible that this approach could be assisted by using different end points for different stages (eg, complete resection rate and pathologically complete response rate for patients with stage IIIa disease undergoing surgery after chemotherapy, quality of life, and symptoms control for patients with stage IV disease). There is increasing interest in quality-of-life data for registration procedures, and ample anecdotal evidence of patients whose conditions improve visibly, or simply by their own account, despite having no measurable formal response. Conversely, there are responding patients who report no subjective improvement.

**The Need for Larger Trials**

The proposed method could eliminate many of the less worthwhile trials that consume valuable time for the participants, leaving them free to concentrate on accruals to better opportunities, particularly larger studies.

As the expected survival improvement with novel chemotherapeutic regimens is expected to be low (3 to 5% increase of 2 to 3 years’ survival rates or 20 to 30% increase in median survival), participants were concerned that useful trials required large numbers of patients, which are difficult to achieve, yet only a tiny minority of patients seem to be recruited in clinical trials. It was believed that trials could be studies faster if the number of patients apparently available was exploited more fully, perhaps by relaxing the many excluding factors (eg, age, performance status, stage). The point was made that trials recruited on the basis of so many exclusions do not appear to represent real life. It was suggested that patient power should be enlisted to overcome the reluctance of the medical community to put forward candidates for these trials.

It was noted that the European Organization for Research and Treatment of Cancer (EORTC) board has decided to require standards of quality and recruitment from its contributing centers, and that many centers have been discarded on that basis. It was hoped...
that the new policy would enhance the quality and quantity of patients entered for trials.

Finally, the view was expressed that large trials might be hampered because individual researchers were reluctant to join a larger program in which their contribution might be diluted, with a consequent loss of professional recognition. It was felt forcefully that patients' health was more important than personal career advancement.

ACKNOWLEDGMENTS: The author is indebted to P. Bruzzi for his useful comments and suggestions.

REFERENCES
1 Simon R. Optimal two stage design for phase II clinical trials. Control Clin Trials 1989; 10:1-10