A Model to Select Chemotherapy Regimens for Phase III Trials for Extensive-Stage Small-Cell Lung Cancer

T. Timothy Chen, John P. Chute, Ellen Feigal, Bruce E. Johnson, Richard Simon

Background: Many more phase II studies have favorable outcomes than the subsequent phase III trials. We used historical data from phase II and phase III studies for patients with extensive-stage small-cell lung cancer (SCLC) to generate a statistical model to provide assistance in selecting chemotherapy regimens from phase II studies for subsequent use in phase III randomized studies. Methods: Information from 21 phase III trials for patients with extensive-stage SCLC initiated during the period from 1972 through 1990 was reviewed to identify those that were preceded by phase II studies of the same regimen. We used data from all the trial pairs to develop a statistical model in which the number of patients, the median survival of patients, and the number of deaths observed in the phase II trial are used to estimate the statistical power of the subsequent phase III trial. All statistical tests were two-sided. Results: Nine phase II studies were identified that preceded phase III trials of the same regimen. The regimens from two phase II studies with the greatest expected power in the phase III trial (0.62 and 0.58) both demonstrated significantly prolonged survival when compared with standard treatment in subsequent phase III trials (P<.001 and P = .002, respectively). The regimens from six of the other phase II studies, for which the median power expected in the phase III trial was 0.28 (range, 0.19–0.52), showed no difference when compared with standard treatment in a phase III trial. Conclusions: Phase II studies for particular regimens that have an expected power of greater than 0.55 provide a reasonable basis for proceeding with a phase III trial. [J Natl Cancer Inst 2000;92:1601–7]

Phase III randomized clinical trials are the definitive method for determining whether a particular therapeutic regimen prolongs the survival of patients compared with existing standard regimens. Phase III trials, however, are expensive and time-consuming. Moreover, a recent review (1) of all North American phase III randomized trials for patients with extensive-stage small-cell lung cancer (SCLC) that were initiated during the period from 1972 through 1990 found that in only five (24%) of 21 such trials did the experimental arm show a statistically significant survival advantage relative to the control arm and that the survival differences in these five studies were small, 0.8–3.0 months (1). This review also showed that, in the majority of phase III randomized studies of SCLC, the experimental regimen did not prove to be superior to the control regimen despite the optimism of investigators after examining the results of a phase II trial.

Currently, phase II studies use response rate, not survival, as the major end point; however, in the case of phase II trials for patients with extensive-stage SCLC, survival data have usually been collected before a phase III trial is initiated. We believe that survival data from phase II studies of extensive-stage SCLC can be better utilized to help decide which regimens should be brought to phase III trial.

We have developed a statistical model based on our National Cancer Institute (NCI) intramural experience (2), on extramural phase II studies, and on data from the North American cooperative groups’ randomized phase III trials for patients with extensive-stage SCLC performed over a 22-year period (1). The time period from 1972 through 1990 was chosen because it includes the time of the introduction of combination chemotherapy and allows the passage of enough time so that the survival data are mature and nearly all of the studies have been published. This model was developed so that clinical researchers can use the survival data from a phase II study to help predict whether their new regimen is likely to increase the survival of patients with extensive-stage SCLC compared with standard chemotherapy regimens in a prospective trial. In this report, we provide evidence that this model offers a potential tool for clinicians to assess the likelihood that a phase II chemotherapy regimen will prove superior to standard regimens when applied in a phase III trial for extensive-stage SCLC.

Methods

Phase III Trials

The phase III trials initiated during the period from 1972 through 1990 for patients with extensive-stage SCLC were identified through a review of the database of the Cancer Therapy Evaluation Program, NCI, Bethesda, MD. A review of these studies was recently published (1).

Phase II Trials

For each published phase III trial of patients with extensive-stage SCLC, we attempted to identify a phase II study that preceded the randomized trial. Phase II studies were initially identified through a review of published references. The authors and chairpersons of the lung cancer committee of the cooperative groups that published these studies were then contacted if possible to confirm that the phase II study we identified was the one that gave rise to the randomized phase III trial. Information was obtained on the dates of the phase II studies, the number and sex of patients, the treatment regimens, the response rates, the median survival, and the number of patient deaths at the time of the phase II analysis. We included only phase II studies that were performed by the cooperative group that conducted the subsequent phase III trial, that were performed at an institution that was a participating member of the subsequent phase III trial, or that were performed at a single institution using a regimen analogous to that tested in a subsequent cooperative group trial. We classified the phase II and phase III trials on the basis of whether they were initiated before or up through 1981. We chose this division because it was an approximate midway point, and it coincides with the initiation of the NCI Cancer Therapy Evaluation Program.
with the transition from cyclophosphamide-based to cisplatin-based chemotherapy at our own institutions and by the cooperative groups (1,2).

We have plotted the phase II median survival (Fig. 1, A) and response rate (Fig. 1, B) versus the median survival of patients treated on the experimental arm of the subsequent phase III trials. The ordinary least-squares regression lines are calculated and drawn on both panels of Fig. 1.

The Model

The statistical power of a phase III clinical trial is the probability of obtaining a statistically significant result. Statistical power calculations are usually based on assumed distributions of survival for the control and experimental regimens, with the medians specified. For example, one might assume that the distributions of survival are exponential in form, with a 9-month median for the control group and a 12-month median for the experimental group. The specification of the median for the control group is usually based on data from previous phase III trials using the control or current regimen. The median for the experimental group, however, is usually hypothetical. The experimental median usually represents the smallest anticipated medically significant improvement in survival over current therapy. This anticipated median survival is used to help determine the number of patients needed for the required statistical power.

In our model, survival data from the phase II study of the experimental regimen are used to calculate the expected power of the phase III trial. The subscript $c$ represents the control values, and the subscript $e$ represents the experimental values throughout. $P(\lambda_c, \lambda_e, d)$ denotes the statistical power of a phase III trial comparing a control treatment with an exponential distribution of survival duration with a failure rate $\lambda_c$ versus an experimental treatment with an exponential distribution of survival with a failure rate $\lambda_e$ when a total of $d$ deaths will be observed in the phase III trial. The exponential distribution has a constant hazard rate by definition, and this distribution assumption is appropriate in our situation. We assume statistical significance when a two-tailed test results in $P < 0.05$.

Expected power is defined as the average value of $P(\lambda_c, \lambda_e, d)$ with regard to the distributions of $\lambda_c$ and $\lambda_e$ that are expected on the basis of available data.

At the time of planning a phase III trial, our model requires that information about $\lambda_e$ be specified as a gamma probability distribution, as is conventional for Bayesian analysis with exponential distributions. The model uses Bayesian analysis as a way to incorporate information about the effectiveness of an experimental treatment into the planning of a phase III trial (3). This information can be formally modeled and used in the computation of the expected power of a phase III trial. The gamma distribution has two parameters, $\alpha_c$ and $\beta_c$. The parameter $\alpha_c$ represents the amount of information (number of deaths) on which the prior distribution is based and is assumed to be a fixed number. The parameter $1/b_c$ represents the total patient-years of survival (until death or censoring) in previous experience with the treatment and is derived so that the median survival is at a specified value. (Equivalently, given $\alpha_c$ and median survival, we calculate $b_c$.)

The statistical analysis takes into account the evidence of increases in median survival for control treatments over time. The distribution of exponential failure rates for arms of phase III trials initiated from 1972 through 1981 well represented by a gamma distribution with parameters $\alpha = 136$ and $1/b_c = 1380$, corresponding to a median survival of approximately 7.0 months. For control arms of trials initiated after 1981, the distribution was well represented using $\alpha = 136$ and $1/b_c = 1741$, corresponding to a median survival of approximately 8.9 months.

Information about the failure rate for the experimental treatment $\lambda_e$ available at the time of planning the phase III trial can also be specified as a gamma probability distribution. Before the phase II study of the experimental regimen, we specified our prior distribution by determining $\alpha_c$ and $\beta_c$ to give an expected median ($m_c$) approximately equal to that expected for the control treatment but with the probability 0.20 that $m_c$ is greater than 12 months. We used $\alpha_c = 10$ and $1/b_c = 125$. These values can be considered realistic approximations because fewer than 20% of past experimental regimens have yielded medically important improvements.

The parameters specified above as representing expectations for the experimental regimen prior to conducting a phase III trial are updated in the model based on phase II results in the following way: At the time of planning a phase III trial, $\alpha_c = 10 + d_c$, where $d_c$ is the number of deaths observed in the phase II trial, and $1/b_c = 125 + T_c$, where $T_c$ is the sum of the survival times in months (until death or

---

**Fig. 1.** Plots of phase II median survival versus the corresponding phase III experimental median survival (A) and phase II response rates versus the corresponding phase III experimental median survival (B). The x-axes represent the median survival (in months) of patients with extensive-stage small-cell lung cancer (SCLC) treated in nine phase II studies (A) and the overall response rates of patients with extensive-stage SCLC treated in eight phase II studies for which response rates were available (B). The y-axes represent the phase II median survival (A) and response rate (B) versus the phase III experimental median survival. The lines represent the least-squares regression lines. The Pearson correlation coefficients are 0.38 ($P = 0.57$) (A) and 0.13 ($P = 0.67$) (B).
last follow-up observed for patients in the phase II trial. These simple updating rules are a result of the complementarity of the exponential distribution of survival and the gamma prior distribution of the hazard parameter.

These specifications permit us to compute the expected power of a phase III trial following a particular phase II study. For the log-rank test with exponential survivals of moderate-sized treatment effects, the power for specific values of \( \lambda_c \) and \( \lambda_e \) can be approximated by

\[
F^{-1} \left( \frac{\log(\lambda_c/\lambda_e) - z_{1-\alpha/2}}{\sqrt{d}} \right)
\]

where \( F^{-1} \) is the inverse of the standard normal distribution function, \( z_{1-\alpha/2} \) is the upper 100(1 –\( \alpha/2 \)) percentile of the normal distribution, and \( d \) is the total number of deaths expected in the phase III trial at the time of analysis. To compute expected power, we averaged this quantity with regard to the gamma distributions of \( \lambda_c \) and \( \lambda_e \) described above. We used \( d = 256 \) as the total number of deaths to be observed in a phase III trial, corresponding to a 90% statistical power for detecting a 50% increase in median survival with a two-sided 5% significance level. The averages were approximated by Monte Carlo integration with 10,000 samples.

**RESULTS**

**Phase III Trials**

Twenty-one North American cooperative group phase III trials of chemotherapy regimens for extensive-stage SCLC were identified that were initiated during the period from 1972 through 1990 and for which the results have been published (1,4–24). Only five (24%) of the 21 trials demonstrated a statistically significant difference between the experimental treatment arm and the standard treatment arm. These five studies revealed a statistically significant trend toward an increase in overall patient survival. The median survival of patients treated on the control arm was 7.0 months for studies started during the period from 1972 through 1981; it was 8.9 months for studies started during the period from 1982 through 1990 (\( P = .001 \)).

**Phase II Studies**

We identified nine phase II studies (25–33) that tested a regimen subsequently studied in a phase III trial (1). The information from these phase II studies is summarized in Table 1. We excluded the other 12 phase III trials (1) because we found differences in either the chemotherapeutic drugs used or the chemotherapy schedule between the phase II regimens cited as the basis for the phase II trials and the phase III trials performed (4–7,9,12,13,16,18,20,23).

Table 1 shows that a median of 20 patients were treated in each of the nine phase II studies. The number of patients treated in different phase II studies varied widely, ranging from five to 106 patients. The median survival of patients treated in the phase II studies ranged from 7.4 to 13.5 months. The median survival of patients treated in the phase II studies is plotted versus the median survival of patients treated on the experimental arm of the corresponding subsequent phase III trial in Fig. 1, A. The least-squares regression line is given in the figure, and the Pearson correlation coefficient is .38 (\( P = .57 \)). The median survival of the patients treated on randomized phase III trials did not appear to increase compared with the phase II studies because of the passage of time. The median survival times of patients were modestly longer in six of the nine phase II studies than in the experimental arms of the subsequent phase III trials that tested the same regimen (range of differences, 0.6–4.1 months). However, a nonparametric paired Wilcoxon signed rank test comparing the survival of patients in the nine phase II trials with the survival of patients treated on the experimental arms of the subsequent phase III studies indicated that the median survival times of these groups were not significantly different (\( P = .11 \)).

Response rates are frequently used to assess the likelihood that a phase II regimen will increase survival over standard treatment in a phase III trial. For eight of the nine phase II trials, response rates were available and these were compared with the survival observed in the subsequent phase III trials. Fig. 1, B, shows that the response rates in the phase II studies did not correlate with the median survival of patients treated with a particular regimen in the subsequent phase III trial. The least-squares regression line is given in the figure, and the Pearson correlation coefficient is .13 (\( P = .67 \)). To give one example of this poor correlation, 95% of the patients treated on the phase II study by Natale et al. (30) showed a partial or complete response (Table 1); however, in the subsequent phase III study, patients receiving the same experimental regimen had a median survival duration of only 8.1 months.

Since response rates in phase II trials have a poor correlation with the phase III survival results, our predictive model utilizes the survival information instead of response rates. To retrospectively test how well the statistical model estimates the outcome of the phase III trials from phase II survival data, we analyzed the nine phase II studies that gave rise to subsequent phase III trials of the same regimen. In Table 2, the results (median survival and number of deaths) of the nine phase II studies and the expected power of each corresponding phase III study as obtained from the model are shown. The expected power is the usual statistical power averaged with regard to the size of

<table>
<thead>
<tr>
<th>Authors (reference No.) of phase II studies</th>
<th>No. of patients (male/female) in phase II studies</th>
<th>Phase II regimen</th>
<th>Response rate, CR/PR, %</th>
<th>Median survival, mo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Williams et al. (25)</td>
<td>12 (5/7)</td>
<td>POCC,</td>
<td>17/58</td>
<td>11.0</td>
</tr>
<tr>
<td>Messeih et al. (26)</td>
<td>20 (NA)</td>
<td>E-VAC</td>
<td>35/20</td>
<td>7.4</td>
</tr>
<tr>
<td>Zacharski et al. (27)</td>
<td>13 (130)</td>
<td>CVM plus warfarin</td>
<td>NA</td>
<td>9.1</td>
</tr>
<tr>
<td>Lowenbraun et al. (28)</td>
<td>106 (80/26)</td>
<td>HD-CAV</td>
<td>12/59</td>
<td>9.7</td>
</tr>
<tr>
<td>Markman et al. (29)</td>
<td>27 (NA)</td>
<td>CAV/HEM</td>
<td>52/44</td>
<td>13.5</td>
</tr>
<tr>
<td>Natale et al. (30)</td>
<td>20 (NA)</td>
<td>CAV/EP</td>
<td>65/30</td>
<td>12.2</td>
</tr>
<tr>
<td>Krook et al. (31)</td>
<td>5 (4/1)</td>
<td>EP schedule</td>
<td>0/80</td>
<td>10.5</td>
</tr>
<tr>
<td>Lohrert et al. (32)</td>
<td>40 (26/14)</td>
<td>EIP</td>
<td>37/34</td>
<td>9.7</td>
</tr>
<tr>
<td>Murphy et al. (33)</td>
<td>22 (15/7)</td>
<td>Flora E</td>
<td>9/73</td>
<td>9.9</td>
</tr>
</tbody>
</table>

*CR = complete response; PR = partial response; C = cyclophosphamide; A = Adriamycin (doxorubicin); V or O = vincristine; P = procarbazine; C1 = CCNU (lomustine); M = methotrexate; H = hexamethylmelamine (altretamine); E = etoposide; P = cisplatin; HD = high dose; I = ifosfamide; NA = not available.
the treatment difference anticipated on the basis of the median survival observed in the phase II trial, the number of deaths observed in the phase II trial, and the distribution of median survivals expected in the phase III trial for the control treatment.

The median expected power for the nine phase III trials based on the corresponding phase II trials was 0.30 (range, 0.19–0.62). The two phase II trials that gave the highest expected powers (0.62 and 0.58) were followed by phase III studies that showed a statistically significant difference between the experimental regimen and the standard treatment arm (Table 2) (28,29). In particular, the phase III trial by Markman et al. (29) showed that patients treated with the regimen developed by Markman et al. (29) (expected power = 0.58) lived statistically significantly longer than those treated on the control arm (P = .002). Likewise, the phase III trial by Daniels et al. (8) showed that patients treated with the phase II trial experimental regimen developed by Daniels et al. (8) (expected power = 0.62) lived statistically significantly longer than patients given standard therapy (P<.001).

The other seven phase II trials had expected powers ranging from 0.19 to 0.52, and six (86%) of the seven subsequent phase III trials showed no difference between the survival of patients on the experimental arm and that of patients on the standard arm (1). The exception was the phase II regimen developed by Loehrer et al. (32), which gave an expected power of 0.21, but the subsequent phase III trial showed a statistically significant difference between the survival of patients treated on the experimental arm and that of patients treated on the standard arm (P = .044; median survival = 9.1 months versus 7.3 months). However, this phase III trial by Loehrer et al. (22) may be a poor example to compare with the power estimate of the model because the median survival of patients on the control arm (7.3 months) is lower than what has been observed in most other studies conducted after 1981.

To illustrate how expected power can be determined from future phase II trials, Fig. 2 reflects the decade of treatment up through 1981 (Fig. 2, A) and the decade after 1981 (Fig. 2, B). The two panels of this figure show expected power as a function of median survival observed in the phase II trial. The different lines correspond to different numbers of deaths (events) observed in the phase II trial. The computations assume that the phase III trial will be sized to include 256 deaths (300 patients followed for 3 years; 150 patients on each arm). To apply to the nine phase II trials retrospectively, three phase II studies up through 1981 are represented in Fig. 2, A, and six phase II studies identified after 1981 are represented in Fig. 2, B.

**DISCUSSION**

We recently reviewed randomized phase III trials for patients with extensive-stage SCLC that were initiated during the period from 1972 through 1990 and found that promising experimental regimens in phase II studies rarely resulted in prolonged survival when compared with standard treatment in a phase III trial. Our analysis showed that a high expected power generated by the model (>0.55) was associated with a higher likelihood that an experimental regimen would increase survival relative to the standard regimen. However, promising results from a phase II trial are rarely sufficient to establish the superiority of a new regimen in the absence of a confirmatory phase III trial. Conversely, a low expected power (≤0.55) scored for a particular phase II regimen does not necessarily predict that the regimen will not prolong survival compared with standard treatment in a phase III trial.

<table>
<thead>
<tr>
<th>Authors (reference No.), phase II/phase III</th>
<th>Median survival, mo, phase II studies</th>
<th>No. of deaths</th>
<th>Expected power</th>
<th>Phase III study outcome</th>
<th>Median survival, mo, phase III studies</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors (reference No.), phase II/phase III</td>
<td>Median survival, mo, phase II studies</td>
<td>No. of deaths</td>
<td>Expected power</td>
<td>Phase III study outcome</td>
<td>Median survival, mo, phase III studies</td>
<td>P*</td>
</tr>
<tr>
<td>Trials (pre-1981)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Williams et al. (25)/Daniels et al. (8)</td>
<td>11.0</td>
<td>7</td>
<td>0.62</td>
<td>Positive</td>
<td>10.0 vs. 7.0</td>
<td>.001</td>
</tr>
<tr>
<td>Messleh et al. (26)/Jackson et al. (11)</td>
<td>7.4</td>
<td>20</td>
<td>0.30</td>
<td>Negative</td>
<td>9.4 vs. 7.8</td>
<td></td>
</tr>
<tr>
<td>Zacharski et al. (27)/Chainin et al. (14)</td>
<td>9.1</td>
<td>13</td>
<td>0.52</td>
<td>Negative</td>
<td>9.3 vs. 7.9</td>
<td></td>
</tr>
<tr>
<td>Trials (post-1981)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lowenbraun et al. (28)/Johnson et al. (15)</td>
<td>9.7</td>
<td>67</td>
<td>0.19</td>
<td>Negative</td>
<td>6.7 vs. 8.0</td>
<td></td>
</tr>
<tr>
<td>Markman et al. (29)/Ettinger et al. (17)</td>
<td>13.5</td>
<td>18</td>
<td>0.58</td>
<td>Positive</td>
<td>10.6 vs. 9.8</td>
<td>.002</td>
</tr>
<tr>
<td>Natale et al. (30)/Roth et al. (19)</td>
<td>12.2</td>
<td>19</td>
<td>0.47</td>
<td>Negative</td>
<td>8.1 vs. 8.6</td>
<td></td>
</tr>
<tr>
<td>Krook et al. (31)/Maksymiuk et al. (21)</td>
<td>10.5</td>
<td>5</td>
<td>0.28</td>
<td>Negative</td>
<td>9.5 vs. 10.5</td>
<td></td>
</tr>
<tr>
<td>Loehrer et al. (32)/Loehrer et al. (22)</td>
<td>9.7</td>
<td>38</td>
<td>0.21</td>
<td>Positive</td>
<td>9.1 vs. 7.3</td>
<td>.044</td>
</tr>
<tr>
<td>Murphy et al. (33)/Miller et al. (24)</td>
<td>9.9</td>
<td>13</td>
<td>0.25</td>
<td>Negative</td>
<td>9.9 vs. 9.5</td>
<td></td>
</tr>
</tbody>
</table>

*Two-sided.
†Numbers of patients who died conservatively estimated as two thirds of patients enrolled.

Table 2. Expected power of nine phase III trials for extensive-stage small-cell lung cancer based on results of phase II studies

Median survival, mo, phase II studies and No. of expected Phase III Median survival, mo, P* trials. The exception was the phase II regimen developed by Loehrer et al. (32), which gave an expected power of 0.21, but the subsequent phase III trial showed a statistically significant difference between the survival of patients treated on the experimental arm and that of patients treated on the standard arm (P = .044; median survival = 9.1 months versus 7.3 months). However, this phase III trial by Loehrer et al. (22) may be a poor example to compare with the power estimate of the model because the median survival of patients on the control arm (7.3 months) is lower than what has been observed in most other studies conducted after 1981.

To illustrate how expected power can be determined from future phase II trials, Fig. 2 reflects the decade of treatment up through 1981 (Fig. 2, A) and the decade after 1981 (Fig. 2, B). The two panels of this figure show expected power as a function of median survival observed in the phase II trial. The different lines correspond to different numbers of deaths (events) observed in the phase II trial. The computations assume that the phase III trial will be sized to include 256 deaths (300 patients followed for 3 years; 150 patients on each arm). To apply to the nine phase II trials retrospectively, three phase II studies up through 1981 are represented in Fig. 2, A, and six phase II studies identified after 1981 are represented in Fig. 2, B.

**DISCUSSION**

We recently reviewed randomized phase III trials for patients with extensive-stage SCLC that were initiated during the period from 1972 through 1990 and found that promising experimental regimens in phase II studies rarely resulted in prolonged survival when compared with standard treatment in a phase III trial. Our analysis showed that a high expected power generated by the model (>0.55) was associated with a higher likelihood that an experimental regimen would increase survival relative to the standard regimen. However, promising results from a phase II trial are rarely sufficient to establish the superiority of a new regimen in the absence of a confirmatory phase III trial. Conversely, a low expected power (≤0.55) scored for a particular phase II regimen does not necessarily predict that the regimen will not prolong survival compared with standard treatment in a phase III trial.
Indeed, one phase II regimen with an expected power of less than 0.55 increased survival statistically significantly in the subsequent phase III trial.

Although the initial application of our model toward estimating phase III results from phase II studies of extensive-stage SCLC appears promising, the use of historical data to predict the efficacy of a particular treatment poses several problems. Differences in patient selection, treatment regimens, and supportive care are unaccounted for when a current treatment regimen is compared with historical data based on previous regimens. Indeed, for patients with extensive-stage SCLC, it has been demonstrated that the number of women enrolled in a clinical trial, the use of cisplatin-based regimens, and improved supportive care may each have an impact on patient outcomes over time (1,2,35–37). Therefore, a historical database that includes 22 years of different clinical trials may include artifacts that render the database invalid for the generation of a model. However, neither sex nor cisplatin-containing regimens had an independent, statistically significant impact on patient survival in our historical database. Therefore, we believe that the database we have used to generate our model is a valid comparison group against which to compare future phase II regimens for extensive-stage SCLC. Nevertheless, individual cooperative groups could use values of $a_c$ and $b_c$ representative of their own phase III control group experiences for use with this model. To model a particular trial in which the median survival for the control group was $m_c$ with $d$ deaths, we set $a_c = d$ and determine the value $b_c$ numerically. We have also applied our approach to obtain a predictive model for limited-stage SCLC (38).

The use of the model described herein has implications for the future design of phase II trials. As we have illustrated in Fig. 2, panels A and B, the expected power to predict whether a regimen that appears promising in a phase II trial will be successful in a phase III trial depends on the number of deaths observed in the phase II trial and the median survival observed. For example, consider a phase II study with 50 deaths observed with a median survival of 14 months. The expected power of the subsequent phase III trial is 0.80 (taken from Fig. 2, B). In contrast, if 10 deaths are observed, with a median survival of 14 months, the expected power is 0.57. Equation 1 or Fig. 2, panels A and B, can be used to obtain the number of deaths required in the phase II trial. However, it can be seen from Fig. 2, A and B, very small phase II trials (fewer than 25 patients) provide an inadequate basis for evaluating the applicability to a phase III trial of survival results that may appear promising in a phase II trial. In our study, we found that the total number of patients in all nine phase II studies is only 265. This finding indicates that, up to 1990, the phase II trials were designed with inadequate size. The situation is different today because of the utilization of an optimal two-stage or three-stage design to differentiate the specified desirable and undesirable response rates (39,40).

The validity of our statistical model and its usefulness in assessing experimental treatment regimens for patients with extensive-stage SCLC can be verified.
only through future prospective studies. Our results thus far suggest that a particular phase II regimen with an expected power of greater than 0.55 provides a reasonable basis for a subsequent phase III evaluation of the regimen. The use of this model may expedite the randomized study of regimens that show promise in phase II studies and would give pause to researchers prior to conducting a phase III trial if the model results suggest that the experimental regimen will be unlikely to prolong survival compared with standard therapy.

REFERENCES


(33) Murphy PB, Hainsworth JD, Greco FA, Hande KR, DeVere RF, Johnson DH. A phase II trial of cisplatin and prolonged administration of


**NOTES**

The views and opinions expressed herein are those of the authors and are not to be construed as the official opinion of the U.S. Navy or the Department of Defense.

*Present address*: B. E. Johnson, Lowe Center for Thoracic Oncology, Department of Adult Oncology, Dana Farber Cancer Institute, Boston, MA.

Manuscript received December 22, 1999; revised July 19, 2000; accepted July 26, 2000.