Introduction to Clinical Trials

Anne O’Neill

January 20-21, 2016
What is a Clinical Trial?

• A clinical trial is defined as an experiment on humans being carried out in order to evaluate one or more potential beneficial therapies.
Types of Clinical Trials

• The three types of clinical trials that we will talk about:
  • Phase I trials
  • Phase II trials
  • Phase III trials
TYPES OF CLINICAL TRIALS, CONT’D

• Typically, a Phase I clinical trial refers to a new treatment that is to be tried on humans for the first time.
• It is a study intended to estimate the so-called maximum tolerable dose (MTD) of a new regimen.
• In such studies, toxicity and pharmacologic information is obtained from dose escalation experiments, whereby volunteers are subjected to increasing doses of the regimen according to a predetermined schedule.
TYPES OF CLINICAL TRIALS, CONT’D

• Typically, a Phase I study has 20-40 patients.
• Typically, patients entered in Phase I trials constitute a very advanced disease population. Thus, they are those who are refractory to therapies believed to be beneficial.
• As a result, the evaluation of side effects on this advanced disease population may not necessarily be the same for patients who ultimately will receive the therapy for an evaluation of benefit.
Phase II cancer trials are initiated after the completion of Phase I trials.

A Phase II clinical trial is a small scale study investigating the efficacy and safety of a regimen.

Efficacy can be measured in terms of response to therapy (response could be clinical response or a change in a marker) and/or progression-free survival.

Safety is assessed in terms of toxicity.
A typical Phase II study has 30-50 patients. Phase II trials can enter newly diagnosed patients (who have received none or limited therapy). Phase II trials can enter patients who have failed 1 or more prior therapies. Phase II trials can investigate a single dose and schedule while others can test combinations of drugs.
After a regimen is shown to be reasonably effective, it is natural to compare it with the current standard treatment. This leads to a Phase III clinical trial. A Phase III clinical trial is a randomized large scale comparative study involving two or more treatment arms. Phase III trials could have between 250-5000, some 10,000 patients.
The Phase III trial is the most rigorous and extensive type of scientific clinical investigation of a new treatment.

This is perhaps the most important of the three types clinical trials since it involves a substantial number of patients and has the potential for changing physician practice and having a major impact on patients’ lives.
Randomization

• The fundamental scientific principle underlying the comparison of patient groups receiving different therapies is that the groups must be alike in all important aspects and only differ in the regimen that each group receives. Otherwise differences between the groups may not be due to the treatments under study.

• Most phase III studies are randomized studies.

• Some phase II studies are randomized.
Guidelines for Design of Phase I Clinical Trials

• The main endpoint of a phase I trial is generally the determination of the maximum tolerated dose (MTD). This is established by entering patients at a pre-planned dose, suspending accrual for a review of toxicity and escalating the dose for the next group of patients if an unacceptable rate of dose limiting toxicity (DLT) is not seen.

• Generally, the dose below that which produced an unacceptable rate of DLT is designated as the MTD, which will be used for further testing in phase II studies.
Since a small number of patients (typically 3-6) are treated at each dose level during the dose escalation, the design of a phase I study will often specify that an additional group of patients (usually n=10) be treated at the MTD once it has been determined.

Anticipated DLTs must be defined in the protocol. This is outlined by the PI.
For the primary endpoint, MTD determination, the Statistical Considerations section must explicitly describe the escalation scheme:
- the number of patients at each dose
- a minimum suspension time between doses
- a decision rule for dose escalation (further exploration of a dose given observed toxicity).

The Statistical Considerations section should also specify the probability of escalation at a given dose for various possible values of the true toxicity rate.

Actual total accrual will depend on observed toxicity and it may be difficult to project a time to completion because of the need to suspend after accrual at each dose.

The Statistical Consideration section also outlines the secondary endpoints.
Example of a Phase I Design

As an example, you may see something like the below text in a protocol defining DLT's:

“Toxicities of high dose chemotherapy can be divided into two groups. The first group consists of unavoidable complications in patients with aplasia (sepsis and hemorrhage) and thus are not considered dose-limiting. The second group, dose-limiting toxicities, consist of lethal or life threatening (grade 4-5) organ damage such as: cardiotoxicity, hemorrhagic cystitis, veno-occlusive disease, interstitial pneumonitis, renal failure, and CNS toxicity. Grade 3 peripheral neuropathies are also considered dose-limiting. Reversible grade 4 mucositis is not considered dose-limiting.”
Example of a Phase I Design, CONT’D

For this example, assume that the investigator is testing the following doses of Drug X:
“Doses of Drug X in mg's

-------------------
20
30
40
50
60

The following design is proposed to determine the maximum tolerated dose (MTD): Patients will enter the study in sets of three. Initially three patients are entered and receive Drug X, which lasts approximately five weeks.
Example of a Phase I Design, CONT’D

(cont’d) The following scheme will be used to escalate dosages.
The first dose of Drug X will be 20 mg and will be dose-escalated in increments of 10 mg.
Dosages of Drug X can be escalated if in the first three patients, none have experienced dose-limiting toxicity (DLT) within the first 5 weeks of treatment.”

So in general, within the first ‘Y’ weeks OR the first ‘Y’ cycles, escalations can occur.
Example of a Phase I Design, CONT’D

(you can summarize the following in the stat section for 3+3 design)

• Enter 3 patients.
• If zero out of the first 3 patients experience DLT, then the dose can be escalated.
• If 2 out of the first 3 patients experience DLT, then the MTD has been exceeded, and dose escalation will be stopped.
• If one out of the three patients has DLT at the current dose level, then the new set of three additional patients will be treated at that same dose level.
• If there is no DLT in the additional three patients, then the dose can be escalated.
• If one or more of these three additional patients experiences DLT, then patient entry at that dose level is stopped, the MTD has been exceeded, and dose escalation will be stopped. The MTD will be the next lower dose.

QUESTION: will 3 patients be entered at the same time/day?
QUESTION: what if see 2/3 with DLT in first dose, then what?
Example of a Phase I Design, CONT’D

“The following table gives the operating characteristics of this scheme. In this example, to ensure that the toxicity at the MTD is acceptable, an additional 10 patients will be accrued at the MTD. The estimation of toxicity rates will be based on these 10 additional patients. It is estimated that approximately 40 patients will be accrued over a period of 1.5 to 2 years. This estimate is roughly based on the expectation of accruing 2 patients per month and studying 5 levels with an average of 3-6 patients per level plus the additional 10 patients at the MTD.”
Table 1 - Probability of Escalation

<table>
<thead>
<tr>
<th>True Rate of DLT (%)</th>
<th>Probability of Escalation</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>.71</td>
</tr>
<tr>
<td>30</td>
<td>.49</td>
</tr>
<tr>
<td>40</td>
<td>.31</td>
</tr>
<tr>
<td>50</td>
<td>.17</td>
</tr>
<tr>
<td>60</td>
<td>.08</td>
</tr>
</tbody>
</table>
Interpretation of Table 1

• Therefore, if the true underlying proportion of toxic events is 30% at the current dose, there is a 49% chance of escalating to the next dose.

• QUESTION: What trend is found in Table 1?

• This is the most common phase I design (3+3 design), others do exist.

• Note: there are secondary endpoints in phase I trials. These can focus on preliminary efficacy endpoints.
Program for this phase I example

Programs:
- Table 1: binomial calculation.
Guidelines for Design of Phase II Clinical Trials

• Usually, response or some other ‘quickly’ discernible objective measure will be the primary endpoint of a phase II study, so that efficacious regimens can be moved to phase III testing ‘rapidly’ and agents with little effect can be dropped from consideration.

• The first stated objective of the study should reflect the endpoint of primary interest. The primary objective of a phase II study may be to “investigate”, “estimate”, or “describe” the endpoints of interest. The words “determine” and “establish” should be avoided, as such goals are not consistent with the exploratory nature of a phase II study.
Randomized phase II studies can be used to study several treatments simultaneously. Some are designed to be able to compare the arms, some are not.

The Statistical Considerations section must include an explicit accrual goal, an estimate of accrual rate, and the expected time to complete accrual.

The standard in ECOG phase II studies is to adjust/increase the total sample size to allow for patients who never begin protocol therapy, ineligibility, unevaluability, and/or pathology exclusions. It may or may not be the standard for non ECOG phase II studies.

The accrual design must be specified, i.e., whether patients will be accrued in one stage, two or more stages, or some other formal sequential design. If a multistage design is used, criteria for moving onto the second (and subsequent) stages of accrual must be given. Any early stopping rules for adverse events (AEs) for example, must also be explicitly stated.
Guidelines for Design of Phase II Clinical Trials, CONT’D

- Phase I/II, II/III, and I/II/III studies were generally discouraged. Instead a sequence of studies at the appropriate exploratory or confirmatory level was recommended. Phase I trials may study a group of patients at the apparent MTD without being categorized as phase II, and clear evidence of efficacy from a phase II trial should be available at the outset of a phase III trial.
- Both the references given in the study and prior large trials should be checked to determine if there is an “industry standard" against which this trial would be judged. Comparisons with historical control rates can only be used as a guide for further phase III testing, and never to support claims of efficacy.
Example of a Phase II Trial with a Randomized Two Arm One Stage Design: E4101

• “A randomized phase II study is proposed where the primary endpoint of the study is the proportion experiencing clinical benefit, i.e. CR, PR, or Stable disease with the patients classified as Stable continuing in that state for 6 months, in patients receiving Arimidex and ZD1839 or Faslodex and ZD1839.

Sixty-eight patients will be randomized to each arm and the combination will be considered worthy of further testing in a phase III trial if 22 or more of the 68 eligible patients in either arm experience clinical benefit as defined above. If the true rate of clinical benefit for this combination is 40%, the probability of observing 22 or more patients in either arm experiencing clinical benefit is 92%. If the true rate of clinical benefit is 25%, then the probability of observing 22 or more patients in either arm experiencing clinical benefit is 11%. Allowing 10% of the patients to be ineligible, up to 74 patients could be randomized to each arm.

Another endpoint of the study is to document toxicities…For another secondary objective, tissue blocks…”

• Note: Next example will review similar secondary endpoints in details.
• Note: Full stat section of E4101 is available for practice-see Anne.
E4101 (cont’d)Primary endpoint was clinical benefit, so calculate sample size around that endpoint.

Power calculations are made for secondary endpoints based on the sample size calculated for the primary endpoint.

QUESTIONS: What type I and type II error rates are typically used/taught? What type I and type II error rates are generally used in phase II trial design and why?

Where do the null and alternative hypotheses come from?

Discuss and demonstrate sample size calculation. Also see handouts.

Sample size can be calculated via:

• NQUERY (pc)
• TWOSTG (unix) (using only the first stage part) or STPLAN (unix)
• BIN1SAMP (desmon in R)

Trial designed in 2001. Could a smaller sample size have been used? Should it have been?
E4101 questions, continued

QUESTIONS

• Even though this is a one stage design, could one or both arms be closed prior to reaching its respective accrual goal? If yes, why and mention at least 2 reasons.

• When accrual goal is met, will the data be immediately ready for analysis?

• When the time comes for analysis, what denominators will be used for primary and secondary analyses? ECOG? DFHCC studies?

• What happens at end of trial if both arms meet criteria for further study?

• What happens at end of trial if neither arm meets criteria for further study?

• What happens at end of trial if one of the arms meets criteria for further study?
Example of a Phase II Trial with One Arm and a Two Stage Design: E4293

“The primary goal of this Phase II study is to determine (1) the response rate of prolonged, low-dose infusional topotecan in the treatment of colorectal cancer. The primary outcome to be analyzed is objective tumor response. A true response rate of 25% would be considered active in this patient population. A two-stage design will be used for this study. Twenty-four patients will be initially entered, assuming that at least 21 will be eligible. If there are at least three responses among the 21 eligible patients, an additional 32 patients will be entered, of whom 29 will be assumed to be eligible. If at least 8 responses are observed among the 50 eligible patients, then the treatment will be considered promising. Given current accrual rates for advanced colorectal cancer at the institutions chosen for this study, the initial accrual stage should be completed within 11 months. If the study continues to the second stage, a little over one additional year will be required to finish accrual. Table 1 gives the characteristics of this design as a function of the true response rate.”
QUESTIONS
Why use two stages in a trial?
• Pro’s:
• Con’s:

What n is the power based on?
### Table 1. Operating Characteristics of the Early Stopping Procedure

<table>
<thead>
<tr>
<th>True Response Rate</th>
<th>0.10</th>
<th>0.15</th>
<th>0.20</th>
<th>0.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability of Stopping Early (&lt;3/21 responses)</td>
<td>0.65</td>
<td>0.37</td>
<td>0.18</td>
<td>0.07</td>
</tr>
<tr>
<td>Probability of Eventually Rejecting Drug</td>
<td>0.90</td>
<td>0.59</td>
<td>0.27</td>
<td>0.10</td>
</tr>
</tbody>
</table>
E4293 questions, continued

Comments/QUESTIONS

• Alternative hypothesis is 25%, null is 10%.
• What trends are apparent in the table?
• What are the ‘take home’ messages/thoughts?, i.e what about the absolute difference between the null and alternative hypothesis and implications… What if targeted a 20% RR instead of 25% (in terms of power and clinical significance)? What if targeted a 35% RR?
E4293 Response

“Table 2 below gives 90% confidence intervals for the true but unknown response rate of this regimen, given possible observed response rates and assuming that the trial continues to 50 eligible patients. For example, if 6 responses are observed, the 90% confidence interval for the true response rate is 6.2% - 27%.”

QUESTION: Can a ‘standard’ confidence interval (CI) be calculated?
Table 2. 90% Confidence Interval for the True Response Rate

<table>
<thead>
<tr>
<th>Observed Responses</th>
<th>Observed Response Rate</th>
<th>90% Confidence Interval for the True but Unknown Response Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>6%</td>
<td>(4.0%, 27%)</td>
</tr>
<tr>
<td>4</td>
<td>8%</td>
<td>(4.4%, 27%)</td>
</tr>
<tr>
<td>5</td>
<td>10%</td>
<td>(5.2%, 27%)</td>
</tr>
<tr>
<td>6</td>
<td>12%</td>
<td>(6.2%, 27%)</td>
</tr>
<tr>
<td>7</td>
<td>14%</td>
<td>(7.4%, 28%)</td>
</tr>
<tr>
<td>8</td>
<td>16%</td>
<td>(8.6%, 29%)</td>
</tr>
<tr>
<td>9</td>
<td>18%</td>
<td>(10%, 31%)</td>
</tr>
<tr>
<td>10</td>
<td>20%</td>
<td>(12%, 32%)</td>
</tr>
<tr>
<td>11</td>
<td>22%</td>
<td>(13%, 34%)</td>
</tr>
<tr>
<td>12</td>
<td>24%</td>
<td>(15%, 36%)</td>
</tr>
<tr>
<td>13</td>
<td>26%</td>
<td>(16%, 38%)</td>
</tr>
<tr>
<td>14</td>
<td>28%</td>
<td>(18%, 40%)</td>
</tr>
</tbody>
</table>
“Table 3 below provides 90% confidence intervals for the true complication rate among the 56 patients registered (assuming all 56 receive protocol treatment (some or all?)). For example, if 4 patients experience a particular (severe or worse) complication, then the true rate for this complication is likely to be in the interval 2.5% - 15.6%.”

QUESTION:
What is the denominator for these CI calculations? Can a ‘standard’ CI be calculated and why?
Table 3. 90% Confidence Interval for the Complication Rate

<table>
<thead>
<tr>
<th>Observed Number with Complications</th>
<th>Observed Complication Rate</th>
<th>90% Confidence Interval for the True but Unknown Complication Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.8%</td>
<td>(0.09%, 8.2%)</td>
</tr>
<tr>
<td>2</td>
<td>3.6%</td>
<td>(0.6%, 11%)</td>
</tr>
<tr>
<td>3</td>
<td>5.4%</td>
<td>(1.5%, 13.3%)</td>
</tr>
<tr>
<td>4</td>
<td>7.1%</td>
<td>(2.5%, 15.6%)</td>
</tr>
<tr>
<td>5</td>
<td>9%</td>
<td>(3.6%, 18%)</td>
</tr>
<tr>
<td>6</td>
<td>11%</td>
<td>(4.8%, 20%)</td>
</tr>
<tr>
<td>7</td>
<td>12.5%</td>
<td>(6%, 22.2%)</td>
</tr>
</tbody>
</table>
“Table 4 below provides power calculations for detecting differences in response rates for high versus low PK levels using Fisher's exact test. We assume that pharmacokinetic levels are dichotomized at the median (50% high/50% low), that the study will accrue 50 eligible patients, and that a 5% one-sided significance level is used. “

QUESTION
From Table 4 it is clear that this study will only have high power to detect rather large differences in PK levels. What else did these calculations assume about the denominator and pk level data?
Table 4. Power to Detect Difference in Response Rates Between High and Low PK Levels

<table>
<thead>
<tr>
<th>High PK Level Response Rate</th>
<th>Low PK Level Response Rate</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>60%</td>
<td>10%</td>
<td>0.98</td>
</tr>
<tr>
<td>50%</td>
<td>10%</td>
<td>0.91</td>
</tr>
<tr>
<td>40%</td>
<td>10%</td>
<td>0.73</td>
</tr>
<tr>
<td>40%</td>
<td>15%</td>
<td>0.53</td>
</tr>
<tr>
<td>33%</td>
<td>10%</td>
<td>0.54</td>
</tr>
<tr>
<td>33%</td>
<td>20%</td>
<td>0.19</td>
</tr>
<tr>
<td>25%</td>
<td>10%</td>
<td>0.29</td>
</tr>
</tbody>
</table>
Programs from E4293 example

- Table 1:
  twolook(unix)
  simon (desmon)
- Table 2:
  twocon(unix)
  twocon(desmon)
- Table 3:
  confin(unix)
  binci(desmon)
- Table 4:
  b2p(unix)
  b2p(desmon)
- Note that desmon programs are available in bob gray’s desmon R library. We encourage you to use those programs as they are more frequently updated than the unix programs.
Example of a Randomized Two Arm one stage phase II with AE primary endpoint and AE stopping rule E2198

- Discussed in DB2SAS training.
- See Anne if want full stat section for another example.
Phase II Trials

• Randomized or not
• One, two, multi stage ± AE stopping rule
• Pick the Winner
• Other
Guidelines for Design of Phase III Clinical Trials

- In general, phase III clinical trials are conducted using therapies (experimental treatments) that have shown promising results in adequate phase II testing, and are likely to be put into widespread use if found substantially better than standard therapies currently in use.
- As in phase II trials, the sample size calculation is based upon the primary endpoint with power calculations for secondary endpoints.
- Phase III clinical trials with the major goal of showing a difference between treatments (superiority) should have a type I error probability of 0.05. Trials with the major goal of showing equivalence may opt for larger type I error and smaller type II error.

**QUESTION:** why is type I error 5%?
Phase III clinical trials must be designed with a large enough sample size to have at least an 80% chance of detecting a clinically meaningful alternative to no treatment difference in a reasonably timely fashion.

The actual power used must be specified.

The alternative used for sample size calculations must be specified in the statistical considerations section.

This should represent the minimum treatment difference necessary for a change in clinical practice.

This difference tends to differ from one disease site to another.

The study chair, study statistician, and disease committee chair must all agree on the appropriateness of the alternative chosen in the experimental design.
The primary analysis of a phase III study may be based upon all eligible patients only. If this is the case, the expected rate of ineligibility and exclusion on central pathology review should be estimated, and the total sample size adjusted accordingly.

That was the former standard in ECOG phase III designs. Newer designs in ECOG base the primary analysis on all patients, whether eligible or ineligible.

Check with your supervisor for standard in non ECOG phase III trials. (in house design of phase III trials is more rare in DFHCC relative to ECOG)
• Patients who refuse treatment after randomization, or who otherwise fail to start their assigned treatment, should be included in the primary outcome analyses as they were randomized: Intent-to-Treat analysis.

**QUESTION: WHY?**

• If for example, the primary analysis is an Intent-to-Treat analysis based upon eligible patients only, patients who do not start protocol treatment are included provided these patients were eligible. If more than a very small proportion of patients are expected not to start on their assigned therapy, then the effect on the power from including these patients should be estimated, and the total sample size adjusted to still provide adequate power. Again, newer designs in ECOG base the primary analysis on all patients, whether eligible or ineligible.

• **Be sure to read (or write!) the statistical section carefully for information on the population (denominator) for the primary endpoint.**
Phase III studies often have multiple endpoints, such as time to disease progression and survival. The power for each endpoint to be analyzed should be discussed. Usually this includes survival, even when it is not the primary endpoint.

Projected accrual rates, total sample size, projected length of accrual, and follow-up after termination of accrual must be specified.

The accrual rate for a particular study should be estimated from other studies previously conducted with the same disease (sub)types.

For intergroup studies (most often the case for ECOG studies), the estimated accrual rate should account for patient resources in the participating groups.

When in doubt, a conservative estimate of the accrual rate should be used. (CTEP for example, has a policy for early closure of slow accruing studies that ECOG uses.)
• The type of test to be used (e.g. log-rank), whether the test will be stratified, and if so how the strata will be formed, the size of the test, and whether the test will be one- or two-sided, must all be stated.

• Any assumptions such as a proportional hazards model or a cure rate model made in calculating power should also be clearly stated.

• The statistical section of all phase III studies must discuss the planned timing of the main (‘final’) analysis of the major endpoint(s). Usually this should be expressed in terms of an appropriate information time, such as the number of observed events, required to give the specified power.
Procedures for interim monitoring should also be discussed.

**QUESTION: do you know what an ‘interim analysis’ is?**

If no interim analyses are planned, a justification for this should be given.

If interim analyses are planned, then the method for performing interim analyses to preserve the overall type I error rate should also be specified. Often an O'Brien-Fleming use function boundary would be used to determine appropriate significance levels at the interim analyses.

Also, the method for determining the timing of the analyses must be discussed. Often, these would be scheduled to take place at certain fractions of the total planned information for the final analysis OR at regularly scheduled intervals once a certain percentage of the total information is observed.
In March 2002, following the review of Freidlin, Korn and George, (Controlled Clinical Trials 20:295-407, 1999) the following recommendations on timing of analyses were made for ECOG studies:

– The study design should specify the minimum information time at which interim analyses will begin, and the total information. For failure time endpoints, these should be expressed in terms of the number of events. Typically, the first analysis should occur at the first Data Monitoring Committee (DMC) meeting after the study reaches 25% of the total information. Other values could be considered, but should be justified.

– The O'Brien Fleming boundary will be truncated at a nominal one-sided level of $\alpha/50$.

– After the first interim analysis, interim analyses should be performed for every regularly scheduled DMC meeting (approximately every 6 months) until the stopping criteria are met or full information is required. The exception to this is if the number of events by the time of the next DMC meeting has increased by only a small fraction, i.e. 5%.

– The expected number of analyses to reach full information and the expected information times at the interim analyses should be given in the statistical section of the protocol.
It is also appropriate to include boundaries to allow early stopping in favor of the null hypothesis (usually repeated confidence intervals).
• It is not possible to provide uniform guidelines for phase III studies with more than two arms. Examples of such designs include comparing two or more experimental therapies with a control, comparing three or more experimental therapies with each other, four arm factorial designs, and comparing two arms in one stratum and two different arms in another stratum. The Statistical Considerations Sections for these studies should clearly state all hypotheses to be tested.

• When more than two arms are used, the protocol should acknowledge the effect of multiple comparisons on the overall type I error rate for the study. In such studies, consideration should be given to adjusting the levels of the individual tests to control the overall type I error rate.
Example of a Phase III Design: E1292

9. Statistical Considerations
The primary objective of this study is to compare adjuvant therapy with 5-FU and Levamisole to adjuvant therapy enhanced with peri-operative therapy with continuous infusion 5-FU in Dukes' Stage B3 and C patients with regard to the disease free interval and survival. This study is designed to be able to detect a clinically meaningful improvement in survival with 82% power while maintaining an overall significance level of 5% in a one-sided test, assuming patients are followed for 3 years after closure of the study to accrual. If 5 years of follow-up time were available, the power will be 91%. For Dukes' C patients, a clinically meaningful improvement in survival was taken to be an improvement in 5-year survival from 55% to 65%. For Dukes' B3 patients, this is an improvement in 5-year rates from 75% to 81%. To estimate the required sample sizes we assume exponential survival and a group sequential design with an O'Brien-Fleming upper bound for early rejection of the null hypothesis of no treatment difference.
Assuming a 5-year survival rate of 75% for Stage B3 patients and 55% for Stage C patients in the adjuvant therapy alone arm, this study will require 800 analyzable patients (400 in each arm) to obtain the stated power using a stratified log-rank test. Calculations assume that approximately 70% of the patients will be Dukes' C and 30% will be Dukes' B3. Since pathological staging is determined only after randomization, the accrual goal needs to be expressed in terms of Dukes' A, B, and C patients. Using data on staging it is projected that 40% of all patients will be Dukes' B3 or C thus requiring a total accrual goal of 2000 patients to obtain the required number of B3 and C patients. As these projections are very approximate, accrual will be closely monitored during the first year and the accrual goals will be revised as necessary. For Dukes' A, B1, and B2 patients there will be no formal treatment comparisons. However, these patients will be followed to gather more information about disease progression and survival.
Peri-operative therapy will require the full participation of the group's surgeons and will be feasible only in member institutions. Thus, this study will not benefit from referrals from outside institutions, which normally comprise 60-75% of the patient population in adjuvant trials. Assuming that the annual rate will be 750-875 total patients/year (300-350 B3 and C patients/year), this study will need about 2-3 years to complete accrual.

**QUESTION:**
Will accrual start the first day that the trial opens?
If yes, why
If no, why
We will assume that the patients will be followed for 3 years after the closure of the study to accrual and that 3 interim analyses will be conducted at 2.5 years (one-half year before the end of accrual), 3.5 years, and 5 years and if the study is not terminated at this time, a final analysis will be conducted at 6 years (3 years of follow-up after closure to accrual). The critical values corresponding to these times will be 3.4, 2.4, 1.9, and 1.74 corresponding to nominal significance levels of .0003, .007, .03, and .04. The chosen interim analysis times closely approximate equal increments in the total number of expected deaths (25%, 50%, and 75% respectively).
Summary of Phase III E1292 example

- Show this design with seqoppr6
- Discuss different parts of the output from seqoppr6

- You can also use seqoppr in bob's R library to practice more with the program and output (we used seqoppr6 on unix today)
Summary of Phase III E1292 example

• Reminder:
This was a phase III trial with interims scheduled at certain time intervals different from interims scheduled at semi regular intervals different from interims scheduled at certain fractions of the total information. YOUR study may or may not have similar timing of interim analyses. Statistical sections among other things also need to include definitions of endpoints and power calculations for the secondary endpoints. This is discussed more in the phase III training notes.
Phase III Trials

Can have:

• Superiority
• Noninferiority
• Equivalence

See Bob Gray’s phase III notes.