TTEDesigner User's Manual

John D. Cook
Department of Biostatistics, Box 447
The University of Texas, M. D. Anderson Cancer Center
1515 Holcombe Blvd., Houston, Texas 77030, USA
cook@mdanderson.org

September 3, 2007

1 Introduction

The natural end point of clinical trials is often the time to an event. The event under consideration may be undesirable, such as disease progression or death, in which case the purpose of a trial would be to maximize the time to this event. Alternatively, the event being studied may be desirable, such as time to transplant engraftment, in which case the trial would seek to minimize the time to the event.

Statisticians often dichotomize time-to-event outcomes by picking some time window and monitoring only whether the event occurred during the window. For example, rather than monitoring the safety of a trial based on survival time, one might consider only whether patients survived six months. Clearly there are several drawbacks to this approach. First, the choice of observation window is arbitrary. Why pick six months rather than, say, eight months? Second, information is thrown away. The dichotomized outcome does not distinguish a patient who died immediate after treatment from a patient who survived five months. Finally, partial information is not used. In a dichotomized model, if we know a patient has survived five months, we have no useable information. With a time-to-event model, every day the patient survives provides a little more information.

This software will guide the user in designing a safety monitoring rule for a single-arm trial with a time-to-event endpoint using a Bayesian statistical model.

Until this software was written, such designs required trial and error simulation to achieve the operating characteristics desired by researchers. This software inverts the design process by asking the user for desired operating characteristics and searching for the parameters that best satisfy these desires.

In general, the desired operating characteristics are not possible to achieve: as in all of statistics, there is the inevitable trade-off between sample size and error probabilities. However, when the desired operating characteristics cannot be achieved, the software informs the user what *can* be achieved and offers alternatives between meeting conflicting requirements.

2 Probability model

Let T_E represent the time-to-event for an experimental treatment. We assume that $T_E \,|\, m_E$ follows an exponential distribution with median m_E. (We model the median rather than the mean because in our experience physicians are more likely to think of event times in terms of medians. One could model the mean by using the fact that the median of an exponential random variable is $\ln 2$ times the mean.) We assume *a priori* that m_E has an inverse gamma distribution. See Section 5 for more on the inverse gamma distribution.

Similarly, let T_S represent the time-to-event for the standard treatment to which the experimental treatment is being compared. We assume that $T_S \,|\, m_S$ is exponentially distributed with median m_S, and that m_S has an inverse gamma distribution. Since the standard treatment is better understood, the distribution on m_S will be more concentrated than the prior distribution on m_E.

Throughout the trial, we update the distribution on m_E and will stop the trial if it appears sufficiently unlikely given the data that the experimental treatment is superior. If the event being monitored is undesirable, we will stop the trial if at any point

$$P(m_E > m_S + \delta \,|\, \text{data}) < p_L \tag{1}$$

for some small cutoff probability p_L. Here δ is a requirement improvement and may be set to zero. Similarly, if we are monitoring the time to a desirable event, we stop the trial if

$$P(m_E < m_S - \delta \,|\, \text{data}) < p_L. \tag{2}$$

For more information regarding this model, including its robustness, see [1].

3 Software installation

The TTEDesigner software is available from the M. D. Anderson Biostatistics Software Download site at

`http://biostatistics.mdanderson.org/SoftwareDownload/`

The software depends on Microsoft's .NET framework version 2.0. If the necessary version of the .NET framework has not been installed on your computer, the installation software will download the framework from Microsoft and prompt you to install it. *You must have a connection to the internet while installing the software*, unless the framework has already been installed.

4 Software use

The TTEDesigner software guides the user through the design of a time-to-event safety monitoring trial in four steps.

1. Basic design parameters.
2. Desired operating characteristics.
3. Simulation to solve for possible values of p_L.
4. Stopping boundaries.

Note that throughout the software, clicking on underlined text provides context-sensitive help to be displayed in the panel to the right. Also, if an input is invalid, a red icon will appear next to the offending value. Mouse over the icon to see an explanation of the error.

4.1 Step 1

The first step in using the software is specifying the basic design parameters.

The distribution on median time-time-to-event for the standard treatment has shape α_S and scale β_S. If historical data are available, α_S is the number of events and β_S is the total time-on-test for the data set. If you have summary statistics rather actual data, you could use the Parameter Solver application, available from the software download site referenced above, to solve for the parameters α_S and β_S.

The distribution on median time-time-to-event for the experimental treatment has shape α_E and scale β_E. If the parameters for the standard treatment, α_S and β_S, are filled in first, the software will create default values for the corresponding parameters for the experimental treatment.

The method of determining these default values and its justification are as follows. We assume *a priori* that the distributions on m_S and m_E have the same mean. Otherwise we would have to believe one treatment is better. If we truly believe that time-to-event is better on the standard treatment, the experiment

is unethical. But if we choose a prior that reflects a belief that time-to-event is better on the experimental treatment, we are starting the trial by giving the new treatment an unearned advantage. Although we assume m_S and m_E have the same mean, we assume that the variance is greater on m_E since we have less experience with it. In order to preserve the mean and increase the variance, we set

$$\alpha_E = 3$$

and

$$\beta_E = 2\beta_S/(\alpha_S - 1).$$

Setting $\alpha_E = 3$ means that in a sense the prior on m_E contains as much information as three observations, thus creating a diffuse prior. The equation for β_E gives the value necessary for m_E to have the same mean as m_S.

The default values for α_E and β_E are only suggestions; you may enter any other positive values you wish. Once values have been entered for α_E and β_E, these values will not change if the corresponding standard parameters are changed so as to preserve your input. However, if you want to have the software recalculate default values, clear the values in α_E and β_E before changing α_S and β_S.

Note that β_S and β_E are both entered in units of months.

It is important to correctly specify the kind of event being monitored. If this event is classified as "bad," an undesirable event, then the method will use the inequality in (1) to monitor the trial, maximizing the time-to-event. If the event is classified as "good," a desirable event, then the method will use the inequality in (2) to monitor the trial, minimizing the time-to-event. The default value is "bad" because in our experience the event monitored is often death.

The δ parameter is the amount of improvement the experimental treatment is required to demonstrate. Note that the meaning of "improvement" depends on whether the event being monitored is good or bad. Delta is the required increase in the time to a bad event or the required decrease in the time to a good event. When in doubt, set $\delta = 0$.

The accrual rate is necessary for the simulations which determine parameters. Enter the number of patients expected to be treated per month in the trial. Note that this value should reflect the rate you *expect*, not the rate you *hope* for. It is quite common for people to greatly over-estimate the accrual rate in clinical trials.

The monitoring rate specifies how often the stopping rule is evaluated. If the rule is evaluated every two months, enter 2 for this value. If the stopping rule will be evaluated each time an event occurs, enter 0, corresponding to continuous monitoring.

The design is not terribly sensitive to accrual rate or monitoring rate and so both of these may be rough guesses.

When you are finished entering the values for Step 1, click the Next button.

4.2 Step 2

Most of the parameters in the time-to-event model considered here are fairly easy to specify. They have intuitive interpretations or rules of thumb for specifying their values. The exception is p_L. The main benefit of the TTEDesigner is its guidance in choosing a value of p_L.

The parameter p_L controls how often the trial will stop. Smaller values of p_L make trials harder to stop: a trial with $p_L = 0$ will never stop, and a trial with $p_L = 1$ will stop immediately. But it is not obvious how p_L should be set between these extremes for a given trial.

In order to determine p_L, the software asks for two scenarios: one in which the experimental treatment is inferior so the trial should stop, and one in which the experimental treatment is superior and one would hope that the trial would not stop. Specify a scenario by giving a hypothetical value for the time-to-event, and then specify how often you would like the trial to stop under that scenario.

Enter the maximum sample size for the trial and click the Next button.

4.3 Step 3

In general, the desired operating characteristics expressed in Step 2 are contradictory and one must pick two out of three criteria to satisfy. With a large enough sample size, one can pick a value of p_L so that the trial stops as often as you'd like under the scenario where it should stop, and stop as little as you like under the scenario where it should not stop. However, the required sample size may be prohibitive, and so we tolerate larger error probabilities.

For a given sample size, we can find a value of p_L to exactly satisfy the stopping characteristics under one scenario or the other, though in general not both.

In this step, the software will calculate the values of p_L to satisfy each of the stopping characteristics separately, and provide three compromise values in between, trading off worse performance according one criterion for better performance according to another.

Click the Calculate button to begin the calculation that will fill the table at the top of the form. This calculation may take a few minutes, depending on the speed of your computer.

Once the calculations are done, the table will give five values of p_L along with the probability of early trial termination and the average number of patients

enrolled under both scenarios.

It may be a difficult decision to choose one of the five designs, though these designs represent the best that can be done with the specified sample size. If none of these designs are acceptable, click the Previous button and must go back to Step 2 to increase the maximum number of patients. Otherwise select a design by clicking the check box corresponding to that value of p_L and click the Next button.

4.4 Step 4

This step gives a table of stopping boundaries for the design selected in the previous step. The stopping boundary table could be used to conduct the trial, or could be used as another view of the operating characteristics of the design.

Note that the stopping boundaries are given in units of **days** while all other times in the software have been specified in **months**. This is because months are the most convenient unit for eliciting physician input (at least in the context of trials we have seen at our institution) while data collected in the conduct of a trial is most conveniently recorded in units of days.

For more information on using this table to conduct a clinical trial, see [2].

5 Appendix: inverse gamma distribution

Let $\alpha > 0$ be the shape parameter and $\beta > 0$ the scale parameter. Then the inverse gamma distribution has PDF

$$\left(\frac{\beta^\alpha}{x^{\alpha+1}\Gamma(\alpha)}\right) e^{-\beta/x}$$

for all positive x.

If $\alpha > 1$ the mean is

$$\frac{\beta}{\alpha - 1}.$$

If $\alpha > 2$ the variance is

$$\frac{\beta^2}{(\alpha-1)^2(\alpha-2)}.$$

If X is distributed as a gamma distribution with parameters (α, β) then $1/X$ is distributed as an inverse gamma with parameters $(\alpha, 1/\beta)$.

NB: The β in our parameterization of the inverse gamma corresponds to $1/\beta$ in another common convention.

References

[1] Peter F. Thall, Leiko H. Wooten, and Nizar M. Tannir. Monitoring Event Times in Early Phase Clinical Trials: Some Practical Issues, *Clinical Trials* 2, 467-478 (2005).

[2] John D. Cook. Continuous safety monitoring in single-arm time-to-event trials without software. MDACC technical report UTMDABTR-006-05. `http://www.mdanderson.org/pdf/biostats_utmdabtr_001_06.pdf`