

East6: Software for Designing, Simulating and Monitoring Group Sequential Trials

Hrishikesh Kulkarni, Cytel, Pune, India
Sheetal Solanki, Cytel, Pune, India

ABSTRACT

Cytel's **East6™** is powerful interactive software designed to plan, simulate and monitor clinical trials, with an emphasis on group sequential and adaptive trials. It comes with an easy-to-use, intuitive GUI with the ability to compute *Sample Size / Power* for trials with continuous, binary and time-to-event endpoints. One can incorporate rules for early termination using flexible efficacy and futility boundaries and compare multiple designs graphically and through tables. *Simulations* can be used to explore various operating characteristics of a trial wherein specific steps can be customized by supplying user defined R functions.

Interim Monitoring enables the user to monitor an ongoing trial and compute inference about the effect size. It can interface with external systems like SAS® for analyzing the accrued data. This article will demonstrate these features with the help of a case study starting with a Fixed Sample Design and then extending it to a Group Sequential Design.

INTRODUCTION

East is designed to meet the planning and monitoring needs of clinical trials. It has been developed over several years by accommodating the advanced and complex statistical concepts in the clinical trials domain.

This article focuses on the varied functionalities of East and the plethora of statistical features that it offers. Two case studies shall be considered which would cover the following features – *Study Design, Simulations, Interim Monitoring, Charts* and a couple of other functionalities like creating multiple scenarios and exploring them using the *Output Summary* feature.

DESIGN

Interactively designing a group sequential study is much more complex than designing a fixed sample study. The patient resources needed in a group sequential study depend not only on the desired power and significance level but also on other factors like the number of looks planned at the data, the stopping boundary family used, the type of hypothesis sought to be rejected etc. To take full advantage of the group sequential methodology, one must have highly interactive software available, both at the study design stage and at the interim monitoring stage.

East is especially being developed with this interactivity in mind. It allows user to design clinical trials using flexible efficacy and futility boundary families allowing for combinations of families. For instance, Efficacy boundaries can be formed using Spending Functions approach whereas futility boundaries can be formed using the p-value method. The boundaries at all the interim looks can be viewed on different scales like Z-scale, Score-scale, p-value scale etc. East has the ability to incorporate accrual and dropout processes at the design stage and explore different scenarios by observing the effect of dropouts on the overall accrual and study duration. Lastly, the Compare Design feature enables user to compare multiple designs in a tabular as well as graphical manner.

INTERIM MONITORING

This is a separate module of East having the ability to monitor a group sequential trial.

It performs computations of repeated p-values, confidence intervals, naïve confidence intervals at each interim look and adjusted p-value, confidence interval and the median unbiased estimate of the effect size at the end of the study. Based on the interim data, the module arrives at the conclusion such as the study could be terminated early establishing efficacy or futility

SIMULATIONS

This is a very useful feature which provides the user a lot of insight into the study design. It can simulate an on-going clinical trial and compute the empirical power of the study. These simulations can be performed to verify different operating characteristics of a trial. One can explore different combinations of accrual and dropout rates to observe their effect on overall operating characteristics like power, accrual duration and study duration.

All important steps in simulations like generating various types of data and computation of test statistic can be customized by supplying user defined R functions for executing these specific steps. This opens up the software for a lot of customization that may not otherwise be possible.

As mentioned above, East features shall be explored with the help of two case studies. In the first case study, discussion focuses on Design and Interim Monitoring of a normal endpoint clinical trial of Schizophrenia whereas the second case study shall focus on Design and Simulations of a time-to-event endpoint clinical trial of Breast cancer.

GUI of EAST

East has a very powerful yet very attractive and interactive User Interface. The complete software is grouped into four modules- Home, Data Editor, Design and Analysis. In this article, the focus is on Home and the Design menus.

Home menu is broadly divided into three panels **Library**, **Ribbon Bar** and **Help**. **Library** is the repository of data such as workbooks, datasets, plots and nodes for Design, Interim Monitoring and Simulations. This data appears in the tree-view. It also allows the user to perform basic tasks such as renaming and saving the node, writing user-defined notes for a particular node.

In **Figure 1**, an example of the view is shown. A dataset namely BloodparElements.cyd is opened and analysed. The analysis node can be seen linked to the dataset in the Library. Next, a group sequential study is designed. The Interim Monitoring dashboard, Simulation and the Design chart, if invoked are appropriately placed below the design.

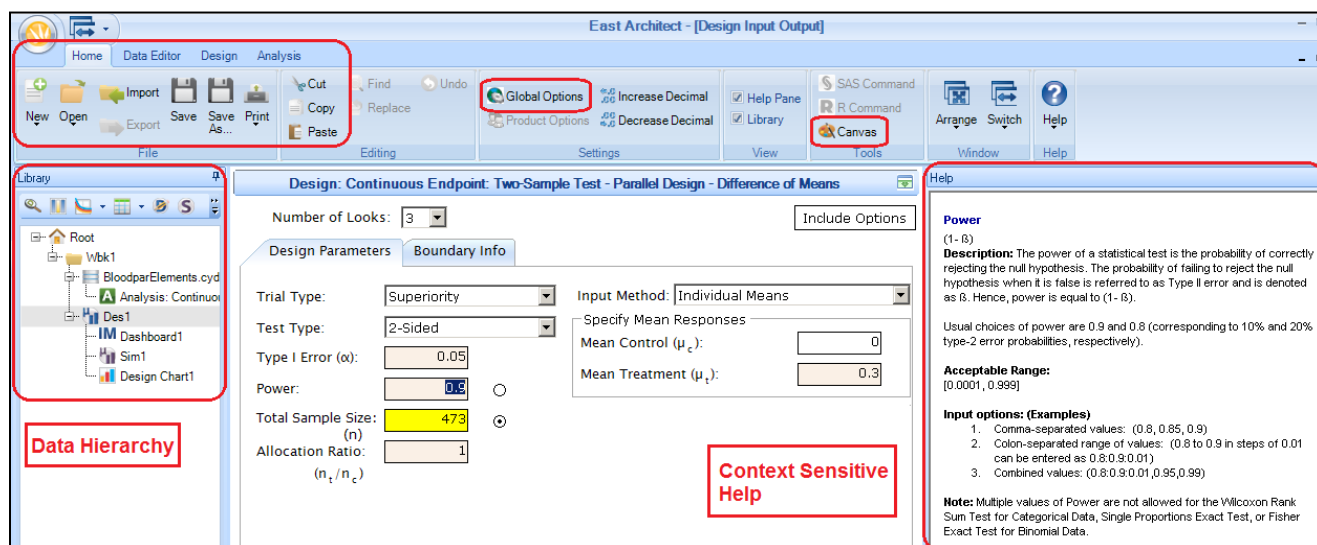


Figure 1: Home Menu

The **Ribbon Bar** has sub-menus which offer useful options like opening or creating new workbooks and datasets, saving and printing functionalities. One of the important features in this panel is **Global Options** where the defaults in terms of values as well as input/output precision can be set. The **Import** functionality enables importing external data files in Excel format as well as datasets from software packages like SAS and SPSS®. The **Canvas** functionality enables creating customized reports from the designs and analysis that have been carried out using East. Use of Canvas functionality is briefly explained at the end of this article.

The right hand side panel is of the context sensitive **Help**. It shows useful information about the control that is selected or is being edited. The help content consists of the definition of that quantity, suggested values, acceptable range if applicable and lastly the different ways of entering value(s) in East. Some of the ways are demonstrated in this article.

STUDY DESIGN

Designing a clinical trial is a very crucial step towards the development of any drug. While designing a trial, many factors are to be taken into account. Some of the important ones are the primary and/or secondary endpoint of the response variable, the significance level (α), desired power value ($1-\beta$) and the difference to be detected (δ). One can either design a Fixed Sample Study and analyze it after the collection of complete data or design a Group Sequential Study and monitor the data at the specified intervals of time. The Design feature of East provides well-recognized error spending approaches to control the Type I error rate (e.g., Lan-DeMets - O'Brien-Fleming or Pocock, Peto

methods, Wang-Tsiatis and Pampallona-Tsiatis methods and many other) that will enable early termination of a study when either no beneficial treatment effect is seen or a statistically robust demonstration of efficacy is observed.

CASE STUDY1 – NORMAL ENDPOINT SCHIZOPHRENIA TRIAL

Consider a two-arm trial to determine if a new drug is efficacious for treating negative symptoms in Schizophrenia as compared to the placebo.

Primary endpoint – Improvement from baseline to week 26 in the Negative Symptoms Assessment (NSA)

The effectiveness is defined by $\delta = \mu_t - \mu_c$ where μ_t and μ_c denote the mean difference in NSA at the baseline and NSA at week 26 for Treatment arm and Control arm respectively.

The trial is designed to test the null hypothesis $H_0: \delta = 0$ versus the one sided alternative hypothesis

$H_1: \delta > 0$. From the past data on related studies, 2-point improvement in NSA is expected which means $\delta \geq 2$. The between-subject standard deviation σ is believed to be 7.5. A Fixed Sample Design (FSD) for this trial will be created first and then it will be extended to a Group Sequential Design (GSD).

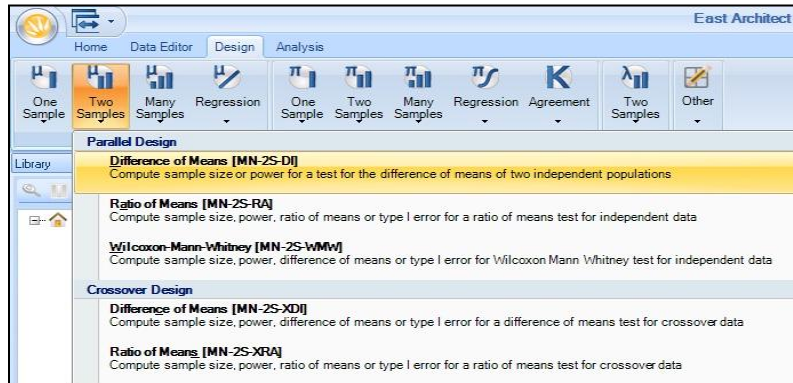


Figure 2: Design Menu

ID	Trial Type	No. of Looks	Test Type	Specified α	Specified Power	Attained Power	nt/nc	Sample Size	Completers	Input Method	δ1	μc	Mean Treatment (Alt.)	σ	Test Statistic
FSD	Superiority	1	1-Sided	0.025	0.9	0.9	1	592	592	Individual Means	2	0	2	7.5	Normal

Figure 3: Fixed Sample Design

With these input parameters, the design FSD is fully powered if a total of 592 subjects are enrolled (296/arm). However, there is considerable uncertainty about the true value of δ . If the true effect was not of 2 units and was greater than that, say 2.5 then the design won't require 592 subjects to detect this difference. In fact, a sample size of 380 would suffice to achieve the power of 90% and a sample size of 592 would overpower the trial. Generally, the sponsor is not in favour of designing an overpowered trial. Here is a need to consider a design that might be more flexible with respect to sample size. The Group sequential designs can help us in this situation.

GROUP SEQUENTIAL TRIAL

Let us explore different scenarios in the Schizophrenia trial by designing a group sequential trial. A design is constructed with two interim looks and 90% power to detect $\delta=2$ such that if in fact, $\delta=2.5$, the trial will stop early. This can save a lot of resources in terms of number of subjects to be enrolled, the study duration and hence the capital investment. If the number of looks is changed to 3, keeping all other parameters constant, an additional tab called

Boundary Info appears. Here, one can select the boundary family and the alpha spending function to be used for the trial. East provides different boundary families like the Spending function, Haybittle-Peto, Wang-Tsiatis and many others. Combination of these families for obtaining boundaries for efficacy and futility is also possible as can be seen from the **Figure 4**. This kind of flexibility is unique to East and not commonly seen in other software packages which deal with group sequential trials.

Look #	Info. Fraction	Cum. α Spent	Efficacy Boundary
1	0.333	0.0001	3.7069
2	0.667	0.0060	2.5127
3	1.000	0.0250	1.9929

Figure 4: Boundary Information Tab

Another major design parameter to be specified on this tab is the timing of the interim analyses. In this case study, consider that the looks are taken at equal spacing. Generally, the type 1 error should be spent conservatively in the early stages of a trial so as to avoid premature termination. Thus, the early stopping boundary is obtained by the conservative O'Brien-Fleming spending function. The three designs **FSD**, **FSD_1** and **GSD_1** can be seen in the **Figure 6** below.

INCORPORATING ACCRUAL / DROP-OUTS

Since the response on primary endpoint for this trial will only be observed after 26 weeks, the actual saving in sample size will depend upon the accrual rate. In this case study, the accrual rate is anticipated to be 8 subjects per week and response lag of 26 weeks is to be taken into consideration. Thus, 208 subjects would have been already accrued and just a traditional group sequential trial may not lead to a lot of saving in this case. The accrual and dropout information needs to be incorporate in the study. East provides the facility to enter the Accrual/Dropout information for designs with normal and binomial endpoints along with time-to-event endpoint as against East@5.4 which allows it only for time-to-event endpoint. Consider that 10% of the subjects from both the arms are expected to dropout from the trial. To incorporate this information, edit **GSD_1** by clicking on tool and bring in **Accrual/Dropout Info Tab** from the **Include Options** button.

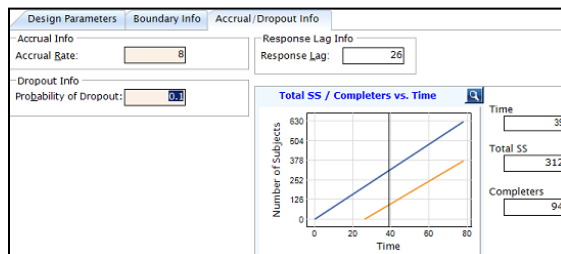


Figure 5: Accrual/Dropout Input Tab

Figure 5

Introducing response lag and dropout information

Figure 6

Comparison of sample sizes and completers across fixed sample design and group sequential designs

Output Summary				
	Wbk1:FSD	Wbk1:FSD_1	Wbk1:GSD_1	Wbk1:GSD_2
Test Parameters				
Trial Type	Superiority	Superiority	Superiority	Superiority
No. of Looks	1	1	3	3
Test Type	1-Sided	1-Sided	1-Sided	1-Sided
Specified α	0.025	0.025	0.025	0.025
Specified Power	0.9	0.9	0.9	0.9
Attained Power	0.9	0.901	0.9	0.9
Model Parameters				
Input Method	Individual Means	Individual Means	Individual Means	Individual Means
Difference in Means ($\mu_t - \mu_c$)	2	2.5	2	2
Mean Control (μ_c)	0	0	0	0
Mean Treatment (μ_t)	2	2.5	2	2
Std. Deviation (σ)	7.5	7.5	7.5	7.5
Test Statistic	Normal	Normal	Normal	Normal
Allocation Ratio (nt/nc)	1	1	1	1
Boundary Parameters				
Efficacy Boundary			LD (OF)	LD (OF)
Spacing of Looks			Equal	Equal
Accrual & Dropout Parameters				
Accrual Rate				8
Response Lag				26
Probability of Dropout				0.1
Sample Size				
Maximum	592	379	599	666
Expected Under H0			597.774	665.894
Expected Under H1			480.137	651.061
Completers				
Expected Under H0			597.774	597.774
Expected Under H1			480.137	480.137

Figure 6: Design Comparison

The designs **FSD** and **FSD_1** demonstrate the use of fixed sample design for varying values of δ . This design is extended to a group sequential design, **GSD_1** with two interim and one final look. Both **FSD** and **GSD_1** have 90% power to detect $\delta=2$ with a one-sided level of 0.025 test. Their sample size commitments too are almost the same. However, under **FSD** there is no possibility of early stopping whereas under **GSD_1**, it is possible to stop early and thereby save on sample size. If the drug is effective, the expected sample size required is 480 in case of **GSD_1** while it remains to be 592 for **FSD**. Under **GSD_1**, the upfront commitment of 599 subjects is required to detect $\delta=2$. On adding the dropouts in **GSD_2**, this commitment increases to 666 and the expected sample size is increased from 480 to 651 under H_1 .

EXPLORING SCENARIOS USING MULTIPLE VALUES FOR INPUT PARAMETERS

East can create multiple designs at a time and compare them graphically. Suppose, the user is interested in assessing the operating characteristics of one group sequential design over a range of δ and Power. It can be done by entering these multiple values as either comma separated values or in the form **From: To: StepSize**

Design Parameters

Boundary Info

Trial Type: Superiority
 Test Type: 1-Sided
 Type I Error (α): 0.025
 Power: 0.8, 0.9
 Sample Size (n): Computed
 Allocation Ratio: 1
 (n_t/n_c)

Input Method: Individual Means
 Specify Mean Responses
 Mean Control (μ_c): 0
 Mean Treatment (μ_t): 2:2.5:0.1
 Dist. of Test Stat.: Normal
 Std. Deviation (σ): 7.5

Figure 7: Multiple Values Input

Output Preview														
ID	Trial Type	No. of Looks	Test Type	Specified α	Specified Power	Attained Power	nt/nc	Spacing of Looks	Efficacy Boundary	Input Method	μ_c	Mean Treatment...	δ_1	Sample Size
Des1	Superiority	3	1-Sided	0.025	0.8	0.801	1	Equal	LD (OF)	Individual Means	0	2	2	448
Des2	Superiority	3	1-Sided	0.025	0.8	0.8	1	Equal	LD (OF)	Individual Means	0	2.1	2.1	406
Des3	Superiority	3	1-Sided	0.025	0.8	0.8	1	Equal	LD (OF)	Individual Means	0	2.2	2.2	370
Des4	Superiority	3	1-Sided	0.025	0.8	0.801	1	Equal	LD (OF)	Individual Means	0	2.3	2.3	339
Des5	Superiority	3	1-Sided	0.025	0.8	0.801	1	Equal	LD (OF)	Individual Means	0	2.4	2.4	311
Des6	Superiority	3	1-Sided	0.025	0.8	0.801	1	Equal	LD (OF)	Individual Means	0	2.5	2.5	287
Des7	Superiority	3	1-Sided	0.025	0.9	0.9	1	Equal	LD (OF)	Individual Means	0	2	2	599
Des8	Superiority	3	1-Sided	0.025	0.9	0.9	1	Equal	LD (OF)	Individual Means	0	2.1	2.1	543
Des9	Superiority	3	1-Sided	0.025	0.9	0.9	1	Equal	LD (OF)	Individual Means	0	2.2	2.2	495
Des10	Superiority	3	1-Sided	0.025	0.9	0.9	1	Equal	LD (OF)	Individual Means	0	2.3	2.3	453
Des11	Superiority	3	1-Sided	0.025	0.9	0.9	1	Equal	LD (OF)	Individual Means	0	2.4	2.4	416
Des12	Superiority	3	1-Sided	0.025	0.9	0.9	1	Equal	LD (OF)	Individual Means	0	2.5	2.5	383

Figure 8: Created Design Scenarios

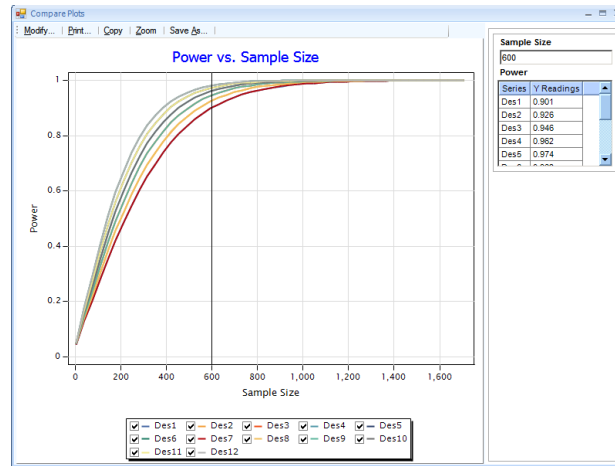


Figure 9: Comparison of Multiple Designs

To explore different combinations of Power and δ values, two values of Power are entered as comma separated values: "0.8, 0.9" and six values of δ from 2 to 2.5 in steps of 0.1: "2:2.5:0.1". This results into twelve design scenarios which can then be compared using the compare plots option. In the **Figure 8**, one can see trend in the sample size required to achieve desired power over the range of μ_t . From the plot in **Figure 9**, one can observe the behaviour of Power function over a range of δ for a constant sample size 600.

INTERIM MONITORING

This is another major feature of East which has the ability to monitor the trial on the Fisher information scale. It also computes the repeated as well as stage-wise adjusted inference like p-values, confidence intervals and unbiased estimates of parameters.

Let us monitor the Schizophrenia trial by creating the IM dashboard for **GSD_1**. This can be done by selecting the design and clicking on **IM** button. Refer to **Figure 10** for a blank IM dashboard. The central panel allows us to enter the look-by-look data in the **Test Statistic Calculator (TSC)** which can be invoked by clicking the **TS** button. An alternative input strip is also made available to enter the look data. The bottom half of the interim monitoring worksheet contains four charts, each with a corresponding data table to its right. These charts summarize the progress of the study and are useful tools for decision making by the data monitoring committee.

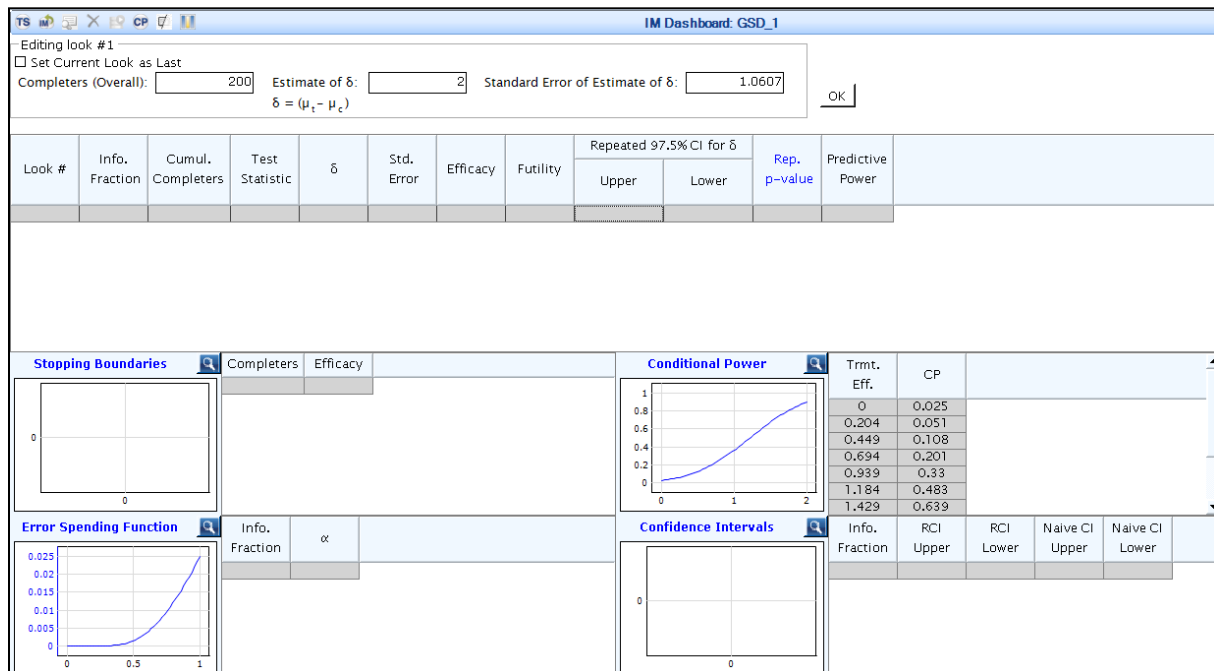


Figure 10: Interim Monitoring Dashboard

As can be seen from **Figure 10**, one needs to enter the interim data in terms of the “Overall Completers” at that look and the pair $(\hat{\delta}, SE(\hat{\delta}))$. The direct value of “Test Statistic” can also be entered in the TSC. One can need the design

level information while entering the data on IM. It can be easily accessed by clicking the button . This information can help the user to decide the sample size/completers at the interim looks. Although the study has been designed assuming three equally spaced analyses (suggested look positions by East design - 200, 400 and 600), departures from this strategy are permissible using the spending function methodology of LAN and DeMets (1983). At each interim analysis time point, based on the current look information, East will determine the amount of type-1 and type-2 error that can be spent and this is done on the basis of the spending functions specified at the design level. It will then re-compute the corresponding stopping boundaries. This strategy ensures that the overall type -1 error will not exceed the nominal significance level α .

The default value of $\hat{\delta}$ in TSC is the same as design δ and the default value of standard error at each look is computed as $\sqrt{\sigma^2 / (n_j * r * (1 - r))}$ where n_j is number of completers at j^{th} look; r is the Assigned Fraction at the design level. Although, the trial is designed with three looks, one can enter the data for more than three looks in IM. Suppose the user decides to take five looks and enters the number of completers at each of the hypothetical analyses as 100, 200, 300, 400, and 600. The $\hat{\delta}$ values and corresponding SE values at these looks were (1.878, 1.5), (2.146, 1.06066), (1.766, 0.86602), (1.756, 0.75), (1.7558, 0.61237).

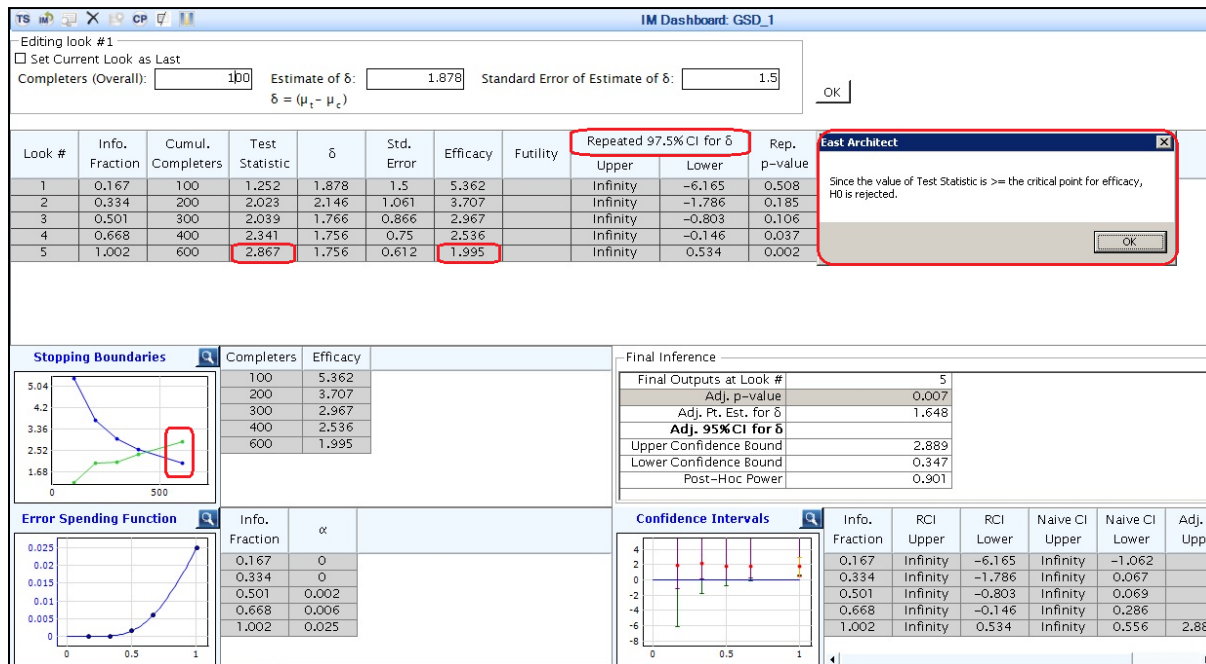


Figure 11: Monitoring Schizophrenia Trial

As the summary data from each interim look is entered into IM dashboard, East computes the 97.5% repeated confidence interval for the effect size. It can be observed here that as the test statistic approaches the upper stopping boundary, the lower bound of the repeated confidence interval approaches zero. If the test statistic crosses the upper stopping boundary then the null hypothesis is rejected and the confidence interval at that look is expected to exclude zero. At the 5th look, the upper boundary is crossed and the Conditional Power chart is replaced by the **Final Inference** table. This table displays the achieved post-hoc power (0.901) of the adopted sequential procedure. The adjusted p-value is 0.007, consistent with the rejection of H_0 .

In this case study, illustration of Design and Interim Monitoring was done for a Normal endpoint Schizophrenia trial. Couple of other features like creating and comparing multiple design scenarios were also covered. In the next case study, the Simulations feature will be demonstrated to explore real life scenarios in a clinical trial.

CASE STUDY2 – SURVIVAL ENDPOINT BREAST CANCER TRIAL

Breast cancer these days is prevalent among women worldwide and accounts for nearly one in four cases of cancer among women. Novartis launched the Femara versus Tamoxifen Adjuvant (FEMTA) trial in March 1998, a two-arm, phase-III, randomized, double-blind trial to compare 5 years of treatment with either Femara or Tamoxifen

PhUSE 2012

in postmenopausal women with operable, invasive breast cancer. For details on this trial, one should visit <http://www.ncbi.nlm.nih.gov/pmc/articles/pmc2893024/>

Primary Endpoint - Disease-Free Survival (DFS), which was defined as the time from random assignment to the earliest time of invasive recurrence of cancer in local, regional, or distant sites or death from any cause.

Input Details - For the comparison of Femara and Tamoxifen, a sufficient number of subjects were enrolled to provide 80% power using a one-sided, 0.025-level test to detect a 20% reduction in the hazard rates. The Survival Percentage for Tamoxifen (control arm) at end of 5 years was known to be 75% and Hazard Ratio (λ_t/λ_c) of 0.8065 was to be detected. The Hazard rates on the two arms can be computed as $\lambda_c = -\ln(0.75)/60$ which is 0.004794 and $\lambda_t = 0.003867$. This was a 5-year study; 3 year enrollment and 2 additional years of follow-up. A group

sequential design incorporating the above inputs was constructed with two interim looks and one final look. These three looks were spaced at the 30%, 70% and 100% information. Additionally, a futility boundary with the spending function Gamma (-4) was also introduced. Refer to the design **GSD_1** in **Figure 13** on next page for this scenario.

Observe that because of introducing efficacy and futility boundaries, expected number of events reduced from 679 (**FSD**) to 477 (under H_0) or to 571 (Under H_1), but the upfront commitment in terms of events is 702.

A clinical trial and particularly an oncology trial is a very complex trial as a lot of factors are to be considered while designing a study for such a trial. The Accrual/ Dropouts/Hazard information may not be constant throughout the duration of study. Secondly, the subjects need to be followed either till the end of the study or for a specified period of time after the drug administration. Designing a clinical trial with such inputs may not be possible without a software support. East has all the required options to handle such situations. The information on non-proportional hazard rates can be provided as inputs in different ways like simple hazard rates or % Survival Rates or Median Survival Times. Non-constant accrual/dropout information can be entered as well in the form of tables available on the

Accrual/Dropout Info tab.

Consider the Breast Cancer study to explore **Design** and **Simulations** features for the scenarios stated below-

- Probability of subjects dropping out from both the arms by month 60 is 10% (**GSD_2**)
- Non-constant accrual (20% in year1; 80% in years 2 and 3) (**GSD_3**)
- Fixed follow-up of 24 months for each subject (**GSD_4**)

Observe the impact of introducing dropouts in **GSD_2** of **Figure 13**. The Maximum Sample Size under H_0 and H_1 increases. Suppose this design is simulated in East. **Figure 12** shows the table of average Accruals, Dropouts and Study Duration from the simulation results. On an average, 231 subjects drop out over the course of the trial having an average duration of 52 months.

Average Accrual, Dropouts and Study Duration								
Look #	Average Look Time	Average Events	Average Sample Size	Average Events		Average Dropouts		Average Follow up
				Control	Treatment	Control	Treatment	
1	29.045	211	3557.268	116.349	94.6511	42.664	42.9355	13.7034
2	46.329	491	4409	269.379	221.6208	99.005	100.254	25.7467
3	59.961	702	4409	379.747	322.2527	141.402	143.1335	36.782
Total	51.697	573.006	4343.348	315.074	257.9315	115.411	116.9876	30.2425

Figure 12: Simulating Design GSD_2

Incorporating Non-constant Accrual Rates

Refer to the design output for **GSD_3** in **Figure 135**. Due to non-constant accrual rates, the sample size is further increased to 4640 (4441 under H_0 and 4583 under H_1).

Introducing Follow-up Time

In some time-to-event trials, subjects are only followed for a fixed period of time. In the current example, suppose the subjects were treated for 24 months and were no longer followed. If event was not observed, the subject was censored. Refer to the design output for **GSD_4** in **Figure 13**. A huge impact of the 24 month follow-up can be seen on the sample size of the study. The sample size is increased from 4640 to 7264.

PhUSE 2012

Output Summary					
BreastCancer.FSDBreastCancer.GSD_1BreastCancer.GSD_2BreastCancer.GSD_3BreastCancer.GSD_4					
Test Parameters					
Trial Type	Superiority	Superiority	Superiority	Superiority	Superiority
No. of Looks	1	3	3	3	3
Sidedness	1-Sided	1-Sided	1-Sided	1-Sided	1-Sided
Specified α	0.025	0.025	0.025	0.025	0.025
Attained α	0.025	0.024	0.024	0.024	0.024
Specified Power	0.8	0.8	0.8	0.8	0.8
Attained Power	0.8	0.8	0.8	0.8	0.8
Model Parameters					
Hazard Ratio (Alt.)	0.807	0.807	0.807	0.807	0.807
Var (Log HR)	Null	Null	Null	Null	Null
Allocation Ratio (nt/nc)	1	1	1	1	1
Boundary Parameters					
Efficacy Boundary		Gm (-4)	Gm (-4)	Gm (-4)	Gm (-4)
Futility Boundary		Gm (-4) (NB)	Gm (-4) (NB)	Gm (-4) (NB)	Gm (-4) (NB)
Spacing of Looks		Unequal	Unequal	Unequal	Unequal
Sample Size					
Maximum	4108	4247	4409	4640	7264
Expected Under H0	4108	4049	4195	4441	6689
Expected Under H1	4108	4188	4345	4583	7076
Events					
Maximum	679	702	702	702	702
Expected Under H0	679	476.774	476.774	476.774	476.774
Expected Under H1	679	571.006	571.006	571.006	571.006
Accrual Duration					
Maximum	36	36	36	36	36
Expected Under H0	36	34.322	34.253	34.713	33.625
Expected Under H1	36	35.5	35.477	35.631	35.224
Study Duration					
Maximum	60	60	60	60	60
Expected Under H0	55.94	43.078	42.791	43.875	36.592
Expected Under H1	59.996	51.731	51.567	52.099	47.814

Figure 13: Breast Cancer Trial Designs

SIMULATING NON-PROPORTIONAL HAZARD RATES

In a clinical trial, there is always a possibility that the drug takes a while to become effective and the response might be observed with some delay. For example- there can be a situation of no difference in survival curves for Treatment and Control for first six months (HR =1). After six months, the drug starts showing its effect and thus HR= 0.8065 is to be detected henceforth. Consider that the hazard rate for Control arm is same throughout the study. Let it be 75% survival or in terms of rate, $\lambda_c = -\ln(0.75)/60 = 0.0048$. East allows the user to assess the impact of non-proportional hazard rates through simulations. Design **GSD_3** is simulated to explore this scenario.

Simulation Parameters		Response Generation Info		Accrual/Dropout Info		Simulation Control Info	
Survival Information							
☒ Using Hazard Rates ☐ Using % Survival Rates							
# of Hazard Pieces 2							
Piece	Starting At	Hazard Rates		Hazard Ratio			
		Control	Treatment				
1	0.000	$-\ln(0.75)/60$	0.005	1.000			
2	6.000	0.005	0.004	0.807			

Figure 14: Simulating non-proportional hazards

Note that the mathematical formulae can also be provided as inputs. On simulating with these parameters, one can observe that the power goes down from 80% to 64%. Refer to the column named **GSD_3_Delayed_Resp** in the **Figure 15**. This is because of the delay in response on Treatment arm. Certain adjustments like increase in number of events from 702 to 1000 can be made. On simulating with modified number of events, the lost power is recovered. Refer to the column named **GSD_3_Increased_Events**. The desired power was achieved by increasing the number of events but this was done at the expense of Study Duration. As the events were increased to 1000, Study Duration was also increased from 53 to 67. To avoid this, total SS is increased from 4640 to 6000. Avg. Study duration changes from 67 to 56. Refer to column **GSD_3_IncreasedSS**.

PhUSE 2012

Output Summary			
	BreastCancer.GSD_3:Delayed_Resp	BreastCancer.GSD_3:Increased_Events	BreastCancer.GSD_3:Increased_SS
Test Parameters			
Trial Type	Superiority	Superiority	Superiority
Sidedness	1-Sided	1-Sided	1-Sided
Attained Power	0.635	0.843	0.807
No. of Looks	3	3	3
Model Parameters			
No. of Hazard Pieces	2	2	2
Allocation Ratio (nt/nc)	1	1	1
Accrual & Dropout Parameters			
Followup Duration	Until End of Study	Until End of Study	Until End of Study
Sample Size			
Maximum	4640	4640	6000
Events			
Maximum	702	1000	1000
Simulation Control Parameters			
Fix at Each Look	Total No. of Events	Total No. of Events	Total No. of Events
No. of Simulations	10000	10000	10000
Simulation Results at Study Termination			
Average Study Duration	53.156	67.481	56.049
Average Sample Size	4589.127	4636.445	5959.338
Average Events	598.689	819.323	833.513
Average Power	0.635	0.843	0.807

Figure 15: Simulating Scenarios

TAKING ADMINISTRATIVE LOOKS

At the initial stages of a group sequential trial, we might be more interested in assessing the drug for futility than for efficacy. This could be because we don't want to declare efficacy with a smaller sample size which will be the case at the start of the trial. In this situation, we would prefer to take an administrative look just to check for futility. This is a new concept and is being introduced in East. East allows the user to take such administrative looks for both efficacy and futility boundaries. User can mark any of these boundaries as absent for a particular look. Some conditions like at least one of the two boundaries must be present for each look are to be satisfied. Let us edit the design **GSD_4** and create **GSD_5** by keeping all other parameters constant and marking the first look efficacy boundary as absent. This can be done on **Boundary Info** Tab.

Spacing of Looks		☐ Equal ☒ Unequal		Boundary Scale: Z Scale		Recalc	
Look #	Info. Fraction	Stop for Efficacy	Stop for Futility	Cum. α Spent	Cum. β Spent	Efficacy Boundary	Futility Boundary
1	0.300	☐	☒		0.0086	NA	0.8223
2	0.700	☒	☒	0.0072	0.0576	-2.4469	-0.7725
3	1.000	☒	☒	0.0250	0.1997	-1.9982	-1.9982

Figure 16: Missing Boundary Scenario

Refer to **Figure 17** for the design output of this missing boundary scenario. For **GSD_5**, comparatively less sample size than **GSD_4** is required. From the design output, the effect of marking the first efficacy boundary as absent can be seen. Cumulative alpha spent at the first look is 0 and hence the corresponding efficacy boundary is also not computed. The impact of a missing boundary can be clearly observed in features like Interim Monitoring and Simulations. The concept of missing boundary is yet to be incorporated for both IM and Simulations.

Design: Survival Endpoint: Logrank Test Given Accrual Duration and Study Duration

Test Parameters:

Design ID: GSD_5
 Trial Type: Superiority
 No. of Looks: 3
 Sidedness: 1-Sided
 Specified α : 0.025
 Attained α : 0.024
 Specified Power: 0.8
 Attained Power: 0.8

Stopping Boundaries: Look by Look

Look #	Info. Fraction (s/s_max)	Events (s)	Sample Size (n)	Cumulative α Spent	Cumulative β Spent	Boundaries		Cum. Boundary Crossing Prob.			
						Efficacy Z	Futility Z	Under H0		Under H1	
								Efficacy	Futility	Efficacy	Futility
1	0.3	210	4744	0	0.0086	NA	0.8223	0	0.2055	0	0.0086
2	0.7	490	7243	0.0072	0.0576	-2.4469	-0.7725	0.0072	0.7839	0.4733	0.0576
3	1	700	7243	0.025	0.1997	-1.9982	-1.9982	0.0244	0.9756	0.8003	0.1997

Model Parameters:

HR = λ_1/λ_0
 Under H0: 1
 Under H1: 0.807
 Ratio of % Surv. at Period #1: 1.057
 Var (Log HR): Null
 Allocation Ratio (n_1/n_0): 1

Events, Sample Size, Dropouts, Pipeline and Analysis Times: Look by Look (Under H1)

Look #	Info. Fraction (s/s_max)	Sample Size (n)	Events (s)	Dropouts (d)	Pipeline (n-s-d)	Analysis Time	Boundary Crossing Prob.	
							Under H1	
							Efficacy	Futility
1	0.3	4744	210	85	4449	25.646	0	0.0086
2	0.7	7243	490	199	6554	38.891	0.4733	0.0489
3	1	7243	700	284	6259	59.722	0.327	0.1422

Boundary Parameters:

Spacing of Looks: Unequal
 Efficacy Boundary: Gm (-4)
 Futility Boundary: Gm (-4) (NB)

Events, Sample Size, Dropouts, Pipeline and Analysis Times: Look by Look (Under H0)

Look #	Info. Fraction (s/s_max)	Sample Size (n)	Events (s)	Dropouts (d)	Pipeline (n-s-d)	Analysis Time	Boundary Crossing Prob.	
							Under H0	
							Efficacy	Futility
1	0.3	4483	210	85	4188	24.564	0	0.2055
2	0.7	7243	490	197	6556	36.758	0.0072	0.5784
3	1	7243	700	282	6261	48.152	0.0172	0.1917

Figure 17: Design output with Efficacy boundary missing

In this case study, a fixed sample study was designed which was then extended to a group sequential trial. Several other scenarios involving non-constant accrual rates, non-proportional hazard rates, dropouts were explored. It can be observed that Simulations is a good tool to deal with such situations. Lastly, a feature allowing us to take administrative looks at the data was demonstrated. Once a study is designed, simulated and monitored, an inference is to be drawn based the observations. These observations unless tabulated will not be truly meaningful and hence the last section of this article talks about an interesting feature called Canvas which is a feature meant for reporting.

CANVAS

Software with advanced features for building alternative scenarios is very much required for conducting a clinical trial. Next step is to compile a report with all the salient observations from the design and analysis of the trial. A clinical study report is a document that gives the complete, verifiable, and accurate picture of a clinical portion of a study. The Canvas functionality of East facilitates generation of completely customized reports with absolutely no programming effort. Many times, a report containing the details from Design and IM stage as well as some exploratory observations obtained from Simulations is required. Canvas is the place to create such reports where one can extract the desired output quantities from different output nodes and then prepare a consolidated report out of them. It is a very simple task as one can actually drag and drop the quantities from the HTML output to the Canvas window. These quantities can either be a single parameter like alpha or a table like stopping boundaries or a chart as well. Once all the desired quantities are pasted to the Canvas, one can either view the resulting report in HTML form or export it to well-established tools like MS word or PDF. The finalized report format can be saved as template.

CONCLUSION

The article elaborates key features of East and their usage at different stages of a clinical trial. East provides software support in planning and execution of a clinical trial. There are several other functionalities like *Compare Design*, Calculators like *TSC* and *Conditional Power* etc. which add to the efficient usage of these broad level features. Finally, it also offers a useful tool for generating customized reports in widely used formats like Word or PDF. Some more interesting features are under development. One of them is a link between IM dashboard and external software like R and SAS. This link will allow us to import the raw data from external sources, run the user-defined R or SAS codes on that data and again display the output on IM dashboard. Some more features under development are Adaptive IM and Conditional Simulations which will enrich the software further.

PhUSE 2012

REFERENCES

Breast Cancer Trial: <http://www.ncbi.nlm.nih.gov/pmc/articles/pmc2893024/>
Webinars on Cytel products: <http://www.cytel.com/science-technology/webinars>
East5 User Manual

ACKNOWLEDGMENTS

The authors would like to thank all the colleagues at Cytel who gave their valuable inputs as and when required.

CONTACT INFORMATION

Hrishikesh Kulkarni hrishikesh.kulkarni@cytel.com +91-20-67090206

Sheetal Solanki sheetal.solanki@cytel.com +91-20-67090209

Cytel Statistical Software & Services Pvt. Ltd.
Survey Number 150, Lohia Jain IT Park
"A" Wing, 6th Floor, Paud Road, Kothrud, Pune 411 038, INDIA
Web: www.cytel.com

Brand and product names are trademarks of their respective companies.