PhUSE 2016

Paper SD04

EnForeSys - an Advanced Patient Enrollment Forecaster with Monte Carlo Simulations

Ritika Yadav, Cytel, Pune, India
Swechhya Bista, Cytel, Pune, India

ABSTRACT

Planning patient enrollment is a major challenge for a successful clinical trial. One common reason for trial failure is inability to meet targeted patient enrollment. Pharmaceutical companies face unacceptable cost and time overrun in conducting clinical trials due to uncertainties and complexities involved in the enrollment process.

Dependable and efficient trial planning requires projection of many important aspects like number of sites, their start up time, enrollment rates in certain time period. It also needs decision rules driven by event triggers using appropriate methodology which addresses the operational complexities.

EnForeSys® is an enrollment forecasting tool. We will demonstrate how EnForeSys uses the power of Monte Carlo simulations to model the effects of various parameter perturbations and risks on the recruitment process. It helps in efficient planning of a trial to achieve expected enrollments within acceptable duration, thus fast tracking the approval process.

INTRODUCTION

EnForeSys is software developed for planning of patient enrollment in clinical trials. It is designed to overcome the challenges faced by pharmaceutical companies in meeting enrollment targets during a clinical trial. According to a paper, *Impact Report Tufts CSDD, 2013*, the most common cause for incomplete Phase 3 trials is related to enrollment. This is not surprising when as many as 37% of trial sites miss enrollment targets, and 11% fail to enrol even a single patient.

EnForeSys helps to estimate trial duration using site and recruitment information. It also helps in predicting the best and the worst case scenarios for achieving the desired target. In addition it can be used to compute accrual rates required to attain the target in specified time period. EnForeSys assists in planning and execution of clinical trials by predicting subject enrollment and site initiation based on known probability distributions, measuring and analysing site performance, estimating the probability of achieving milestones, and validating enrollment timelines.

GUI of EnForeSys

EnForeSys is designed on an architect platform; a platform designed by Cytel specifically for trial planning, design and execution. It has powerful, user friendly and attractive graphical user interface. The complete software is grouped into four modules- Home, Data Editor, Analysis, and EnForeSys. This paper focuses on Home and Plan Recruitment of the EnForeSys menu.

Home menu is broadly divided into three panels Library, Ribbon Bar and Help. All objects created within EnForeSys are saved in the Library in proper order using a tree-structure. Hence, library is a repository of data such as workbooks, plan, plots, tables and datasets. It also allows the user to perform tasks such as renaming and saving the workbooks, invoking plots, tables, detailed and summary outputs etc.

The Ribbon Bar has sub-menus which offer options like opening, creating and saving workbooks and datasets, importing external data files in excel data formats, saving functionalities. This panel also includes an important feature: Global Options. In Global options the defaults in terms of values as well as input/output precision can be set. A context sensitive Help panel is also provided which shows useful information about the control that is selected or is being edited.

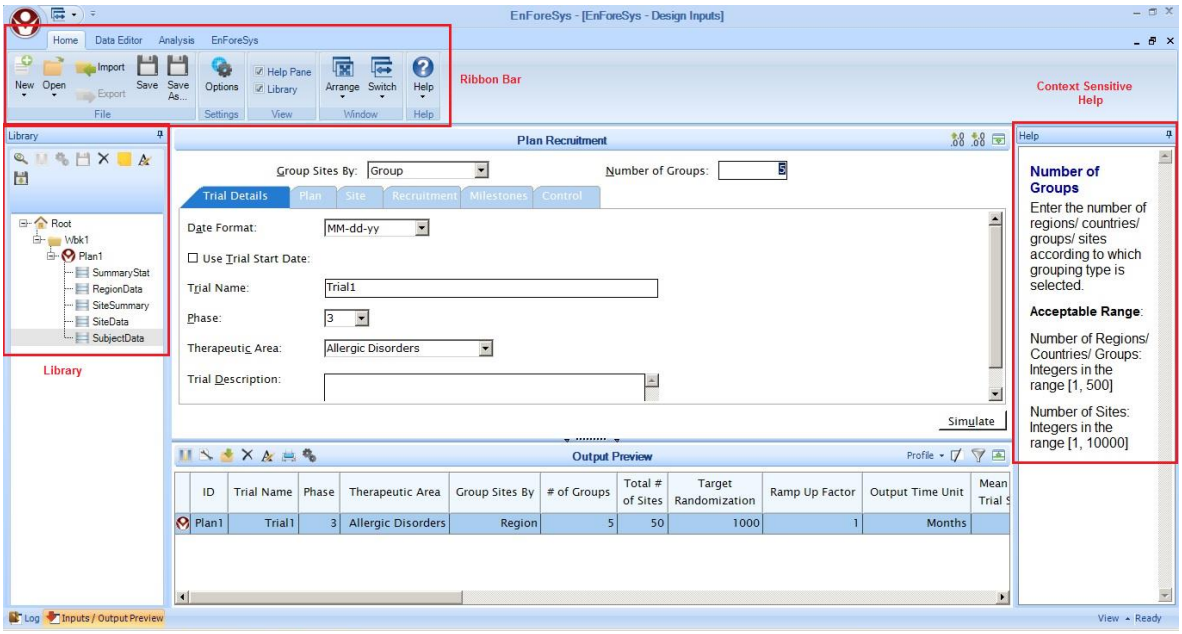


Figure 1: Home Menu

PLAN RECRUITMENT

Planning of patient enrollment is one of the crucial steps for successful completion of any clinical trial. There are various factors that play an important role in success of any enrollment process. Some of these factors are site initiation time, accrual rate, screen periods, site and screen failure probability etc. EnForeSys takes all these factors into account and uses Monte Carlo simulations to predict the duration of the trial. The various inputs that can be entered by the user are grouped logically in various tabs. Let us consider an example of a clinical trial, “Impova” to describe these inputs in various tabs. Impova is a drug aimed at alleviating the symptoms of seasonal allergies.

TRIAL DETAILS

The trial detail consists of various input fields related to trial configuration such as trial name, trial start date, phase, therapeutic area, trial description. Let us consider that Impova will be tested on 700 subjects in a multi country multicenter, randomized, double blind, placebo-controlled trial which will begin on 1st January, 2017. The figure below shows all these input fields on the trial details tab.

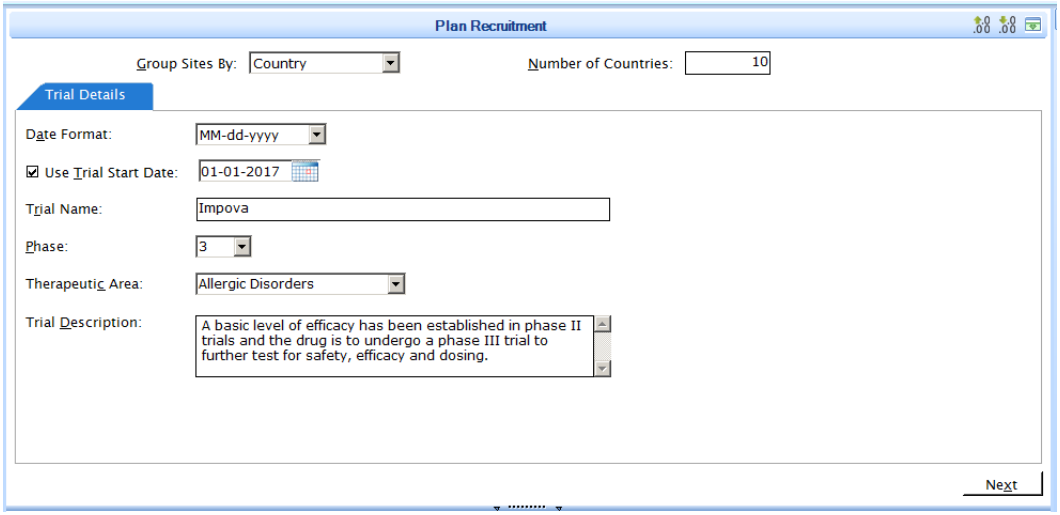


Figure 2: Trial Details

GROUP SITES BY

EnForeSys provides user the option of grouping sites by region, country or group. Using region and country option the sites can be grouped by their geographical location. Using the “Region-Country” option three levels of grouping of sites is also possible. The ‘Group’ option provides the user to group sites based on some other similar attributes. In Impova trial subjects will be enrolled in multiple sites across ten countries in Europe and North America.

DATE & DATE FORMATS

The default output option in EnForeSys is elapsed time. User has an option to display dates in output by checking "Use Trial Start Date" option. Use of this option gives an idea to the user about the dates at which the different events would occur for the planned trail. The software supports different date formats used worldwide. In this example we are considering 1st January, 2017 as the start date of the trial.

PLAN

The plan tab consists of some of the most important inputs like target randomization, randomization duration, ramp up factor, output time unit and interim analysis parameters.

TARGET RANDOMIZATION

Target randomization is the number of subjects that we aim to randomize at the end of the trial. Let us consider the target randomization as 700 for this trial.

OUTPUT TIME UNIT

It is time unit used for displaying the outputs. It can be days, weeks, months or years. The output time unit selected for this trial is months.

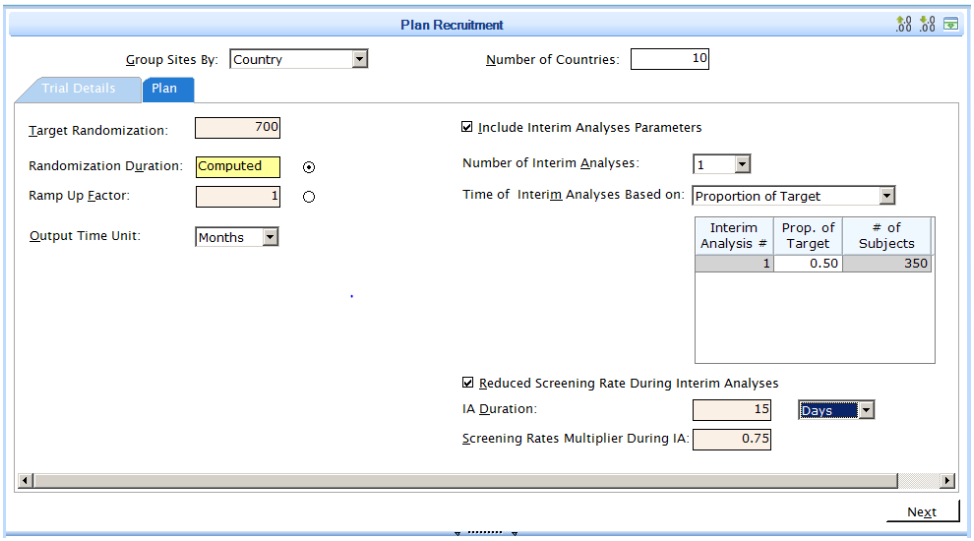


Figure 3: Plan

RANDOMIZATION DURATION AND RAMP-UP FACTOR

User can either select randomization duration or ramp-up factor. When the Ramp-up Factor is selected as the input EnForeSys multiplies the screening rate by this ramp up factor and calculates the trial duration based on the site and recruitment information. If the user selects the option of entering randomization duration, screening rates entered by user will be adjusted by EnForeSys by multiplying with an appropriately chosen ramp up factor so as to get the Target Randomization within user's desired randomization duration. Let us select the option of estimating randomization duration for ramp-up factor of value 1.

INCLUDE INTERIM ANALYSES

During enrollment process we may want to analyze the accrual data. This option allows user to analyze data based on the selected interim type. EnForeSys provides two options for selecting interims "Proportion of Target" and "Elapsed Time from Trial Start".

In the interim analysis based on proportion of target, the analysis of the accrual data is performed when the number of subjects randomized reaches user specified proportion of the target randomization. When the Elapsed time from trial start is selected as the option to perform the interim analysis, the analysis is done after the trial reaches the user entered time point. In certain cases user would like to slow down the recruitment process as he would be waiting for the results of interim analysis. This can be done by selecting the option of "Reduced Screening Rate During Interim Analyses".

In our example let us consider the Interim analysis based on "Proportion of Target" with a value of 0.5. Also let the duration of this interim analysis be 15 days with a reduced screening rate multiplier of 0.75.

SITE:

EnForeSys generates randomizations based on site level parameters entered by user like number of sites in a

group, sites opening schedule, maximum number of randomizations from that site and site failure probability. Each of these inputs can be represented by different distributions. Default values for the parameters of the distribution for these inputs can be changed through "Setup Advanced Parameters" button.

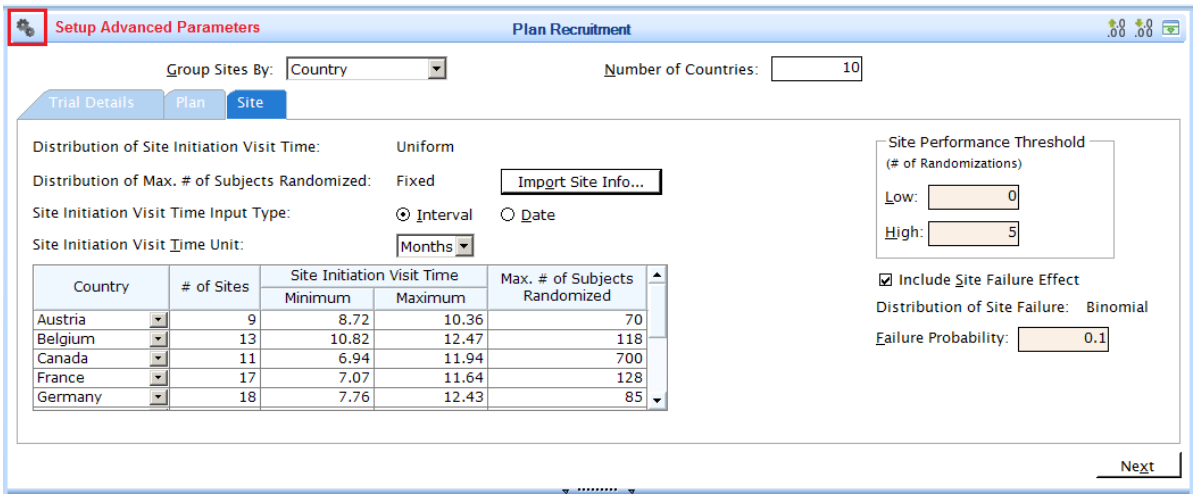


Figure 4: Site

SITE INITIATION VISIT TIME

It is the time at which a site is considered to be open for accrual of subjects.

RANDOMIZATION CAP

Many times regulatory authorities require certain percentage of patients recruited from particular group/country if the company want to market the drug in that country; Randomization Cap can help in this case. Also the extent of recruitment of subjects is driven by several factors which may lead to biased selection from the sites/group of sites dominated by these factors. Using an upper limit to the number of subjects randomized by these groups can eliminate this bias.

SITE PERFORMANCE THRESHOLD

Site Performance Threshold uses number of randomizations as a measure of performance of a site. A low performing site is one that randomizes subjects less than the lower threshold while a site is considered high performing if it randomizes subjects more than the higher threshold.

SITE FAILURE EFFECT

This input accounts for the possibility that a certain percentage of sites may not enroll any subject for screening during the trial. This percentage can be specified as Site Failure Probability.

For the Impova example, let us consider the site initiation visit time following a uniform distribution with minimum and maximum values as shown in figure 4, the low and high threshold values for the site performance to be 0 and 5 and the site failure probability to be generated from binomial distribution with probability 0.1.

RECRUITMENT

In this tab we specify the randomization information which forms the basis upon which EnForeSys generates subject recruitment through a user selected recruitment model, Poisson or Uniform.

INPUT RATE

User has an option to specify either the screening or randomization rate. Based on the selected option the rates are entered by the user.

SCREENING / RANDOMIZATION RATE

The screening/randomization rates for the accrual of subjects can be represented using different distributions. The distribution to represent the screening/randomization rates can be selected from “Setup Advanced Parameters button”.

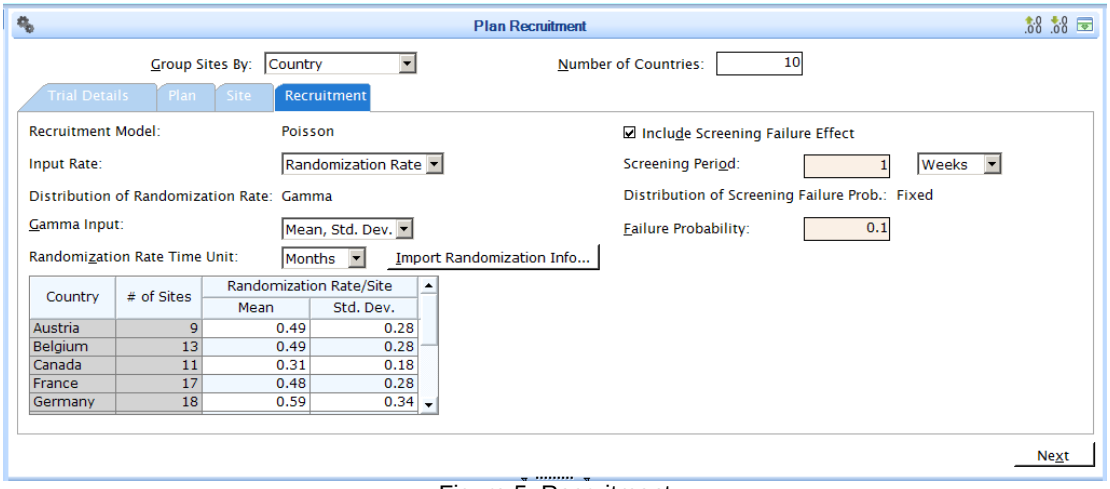


Figure 5: Recruitment

SCREENING FAILURE EFFECT

It is likely that a considerable number of subjects who are screened will not be randomized into the clinical trial, for instance because they do not meet certain criteria as per the trial protocol. EnForeSys provides the option to account for these screening failures in the enrollment process by generating a screen failure probability. This option can be activated by checking the option of "Include Screening Failure Effect". Furthermore, we can even input a screening period, which is the time required by the subject to get the screening results.

For Impova trial, let the recruitment model follow a Poisson process with gamma distribution for randomization rates. The parameters of gamma distribution for randomization rates of different countries are shown in figure 5. The trial also includes screening failure effect with a failure probability of 0.1 and screening period of 1 week.

MILESTONES:

Aside from exploring how long the overall enrollment is likely to take, user may also be interested in knowing how long it might take to randomize a certain number of subjects or how many subjects have been randomized by a particular time point. EnForeSys allows user to set a variety of milestones based on the following criteria:

- Number of subjects screened
- Number of subjects randomized
- Date
- Elapsed Time From Trial Start (in our chosen time unit)

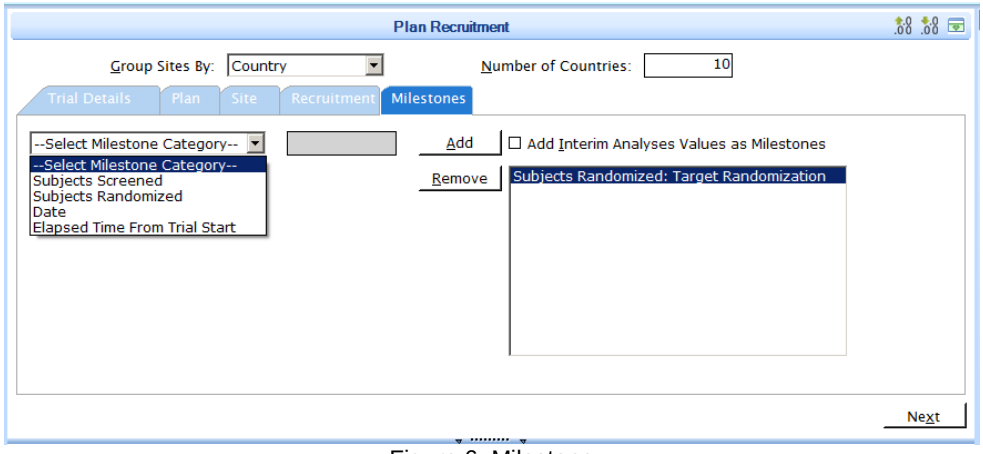


Figure 6: Milestone

CONTROL (SIMULATION PARAMETERS):

The Control tab is the final stage of the input process, in which we can specify different simulation parameters and output options. We can select the Random Number Seed to either be 'Clock' or 'Fixed'. The fixed seed ensure that your results match every time you perform the simulations using the same seed and input values. While if the Clock seed option is selected, EnForeSys computes simulation outputs depending upon the random seed generated based on the system time.

EnForeSys produces a number of output files summarizing the simulation results for each run. In the Output Options panel we can choose to save these output files in the .cydx format or as CSV files. We can select the former option by setting Output Type as 'Case Data' by means of the drop-down menu. These saved data can be used for further analyses of the simulation output beyond the charts and tables and reports produced by EnForeSys.

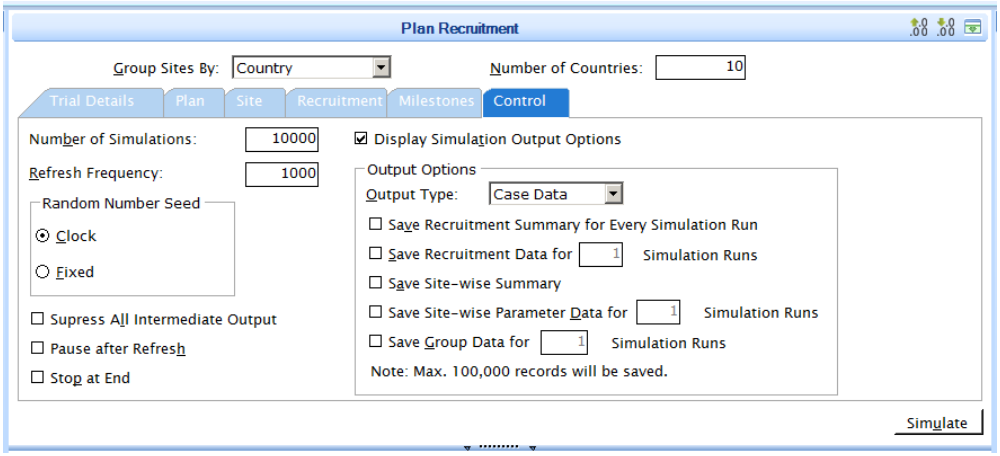


Figure 7: Control

While simulating the trial, EnForeSys generates intermediate outputs in the form of plots and table (refer Figure 8), which provide summary of the progress of simulations. These intermediate outputs can be suppressed by using the option "Suppress All Intermediate Output".

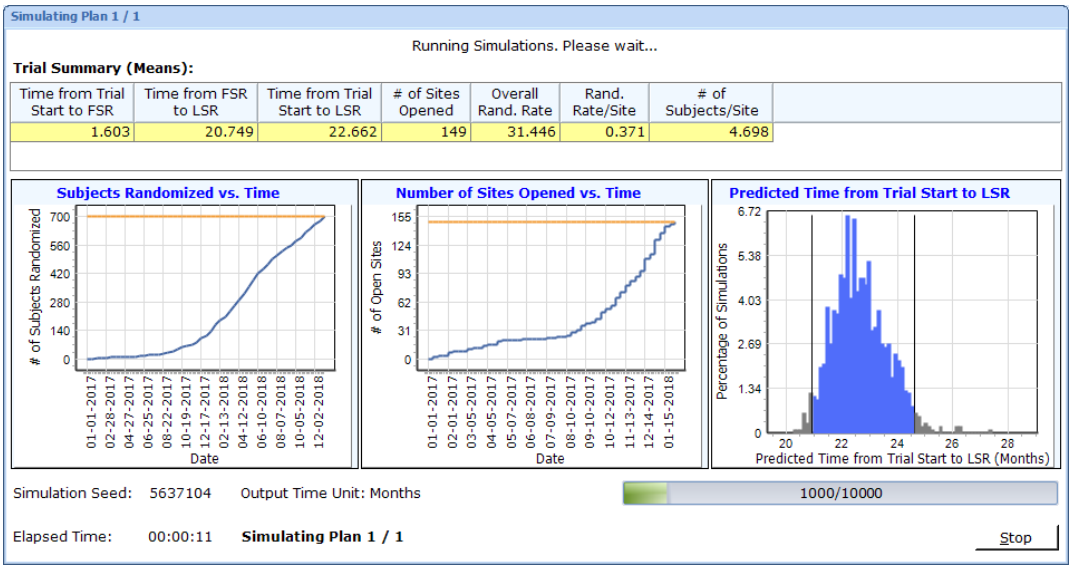
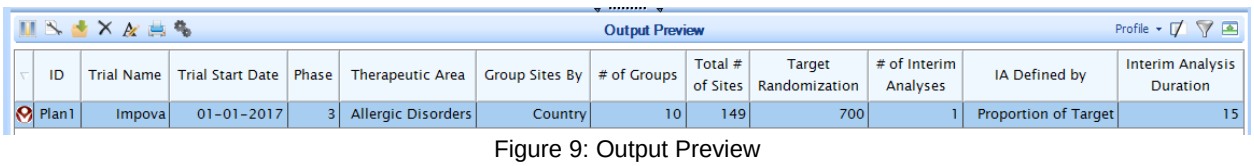


Figure 8: Intermediate Outputs

OUTPUT PREVIEW

After simulations have been performed, the simulating window closes automatically and a one line summary of the results is displayed in the Output Preview window. This plan can be saved into the library.



ID	Trial Name	Trial Start Date	Phase	Therapeutic Area	Group Sites By	# of Groups	Total # of Sites	Target Randomization	# of Interim Analyses	IA Defined by	Interim Analysis Duration
Plan1	Impova	01-01-2017	3	Allergic Disorders	Country	10	149	700	1	Proportion of Target	15

Figure 9: Output Preview

OUTPUT DETAILS:

Select the design in the Library pane and click the Details button, or alternatively simply double-click on the design

itself. Both actions open a page containing a set of summary tables providing statistics such as minimum, maximum, mean, median and 95% credible intervals for a variety of key parameters.

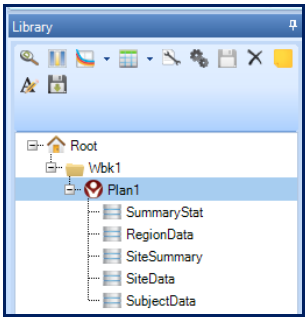


Figure 10: Plan Saved in Library

Different tables are displayed in the output details window which are explained below.

TRIAL RECRUITMENT SUMMARY

Some of the important outputs of the trial are displayed in this table. For example:

1) Time from Trial Start to First Site Initiation Visit
2) Time from Trial Start to First Screening
3) Time from Trial Start to First Subject Randomization
4) Time from First Subject Randomization to Last Subject Randomization
5) Time from Trial Start to Last Subject Randomization
6) Screening and Randomization Rate per Site

Trial Recruitment Summary

Parameter	Minimum	Maximum	Mean	Median	95% Credible Interval	
					LL	UL
Time from Trial Start to FSIV (Months)	5.9E-5	3.2	0.4	0.2	9E-3	1.3
Time from Trial Start to First Screening (Months)	1.4E-2	7.4	1.6	1.5	0.2	3.8
Time from Trial Start to FSR. (Months)	0.2	7.7	1.9	1.8	0.5	4.2
Time from FSR. To LSR. (Months)	16.1	25.6	20.8	20.8	18.2	23.2
Time from Trial Start to LSR. (Months)	19.5	26.6	22.7	22.6	20.9	24.7
Overall Randomization Rate (Subjects/Month)	26.59	37.457	31.415	31.402	28.694	34.203
Screening Rate/Site (Subjects/Site/Month)	0.316	0.534	0.412	0.411	0.361	0.467
Randomization Rate/Site (Subjects/Site/Month)	0.281	0.476	0.37	0.37	0.324	0.42
# of Subjects Screened/Site	5.007	5.463	5.221	5.221	5.101	5.349
# of Subjects Randomized/Site	4.698	4.698	4.698	4.698	4.698	4.698
# of Screen Failures	46	114	77.936	78	60	97
# of Sites Opened	149	149	149	149	149	149
# of Sites with at least 1 Subject Screened	112	141	128.222	128	120	136
# of Sites with no Subject Screened	8	37	20.778	21	13	29

Figure11: Trial Recruitment Summary

TRIAL START TO SIV SUMMARY

This table displays the summary of the amount of time taken to initiate a certain percentage of sites.

Time from Trial Start to SIV Summary (Months)

Parameter	Minimum	Maximum	Mean	Median	95% Credible Interval	
					LL	UL
10% SIV	2.1	7.8	5.3	5.3	3.4	7.1
20% SIV	6.7	9.2	8	8	7.5	8.5
25% SIV	7.6	9.4	8.4	8.4	8	8.9
30% SIV	8	9.8	8.9	8.9	8.4	9.3
40% SIV	8.8	10.5	9.6	9.6	9.2	10
50% SIV	9.4	11	10.2	10.2	9.9	10.6
60% SIV	10.2	11.4	10.9	10.9	10.6	11.2
70% SIV	10.9	11.6	11.3	11.4	11.1	11.4
75% SIV	11.1	11.8	11.4	11.4	11.4	11.6
80% SIV	11.3	12	11.6	11.6	11.4	11.8
90% SIV	11.5	12.3	12	12	11.8	12.2

Figure12: Trial Start to SIV Summary

SITE PERFORMANCE TABLE

This table gives the summary of low, medium and high performing sites based on the low and high threshold values entered by the user on the Site Tab.

Site Performance Summary

# of Rand.	# of Sites					
	Mean	Median	25th %ile	75th %ile	95% Credible Interval	
					LL	UL
0	22.049	22	19	25	15	30
1-4	61.29	61	58	65	50	72
>=5	65.661	66	63	68	58	73

Figure 13: Site Performance Summary

COUNTRY WISE RECRUITMENT SUMMARY

This table provides a country-wise breakdown of several key parameters we encountered earlier in table "Trial Recruitment Summary", for instance the mean values for the number of active sites, the first site initiation visit (FSIV) in output time unit, the number of subjects screen and randomized, and so on.

Country Wise Recruitment Summary (Mean)

Country	# of Sites Opened	FSIV (Months)	# of Subjects		Screening Time (Months)		Screening Rate/Site (Subjects/Site/Month)	Rand. Rate/Site (Subjects/Site/Month)
			Screened	Randomized	First Subject	Last Subject		
Austria	9	8.9	55.123	49.644	9.5	22.1	0.474	0.427
Belgium	13	10.9	67.945	61.151	11.5	22.3	0.477	0.429
Canada	11	7.4	43.87	39.492	8.6	22.2	0.302	0.272
France	17	7.3	105.055	94.495	8.1	22.3	0.468	0.421
Germany	18	8	93.666	84.294	8.7	19.3	0.571	0.515
Netherlands	8	11.4	42.382	38.167	11.7	21.6	0.517	0.466
Russia	17	10.9	114.9	103.461	11.4	22.3	0.62	0.558
Sweden	8	7.9	41.447	37.307	8.9	22.1	0.38	0.342
UK	13	10.3	57.067	51.363	11.1	22.3	0.398	0.358
United States	35	0.4	156.242	140.626	1.6	22.4	0.274	0.247

Figure 14: Country Wise Recruitment Summary

MILESTONE SUMMARY

In the tables listed under Milestones Summary user can find out how long it takes for user chosen milestones to be achieved. As discussed earlier, different milestones can be selected as per the user's preference. The summary of these milestones are displayed under milestones summary as shown in figure 15.

Milestones Summary

Monitor Interim Analyses

Interim #	Proportion of Target Randomization	IA Target Rand.	Time from Trial Start to IA (Months)				Time from IA to LSR (Months)			
			Mean	Median	95% Credible Interval		Mean	Median	95% Credible Interval	
					LL	UL			LL	UL
1	0.5	350	16	16	15.1	17	6.6	6.6	5.5	8

Monitor by # of Subjects Randomized

# of Subjects Randomized	Time to Achieve Milestone (Months)				# of Subjects Screened			
	Mean	Median	95% Credible Interval		Mean	Median	95% Credible Interval	
			LL	UL			LL	UL
700	22.7	22.6	20.9	24.6	777.961	778	761	797

Monitor by Dates

Date	# of Subjects Screened				# of Subjects Randomized			
	Mean	Median	95% Credible Interval		Mean	Median	95% Credible Interval	
			LL	UL			LL	UL
06-05-2017	10.127	10	3	20	8.318	8	2	18

Figure 15: Milestone Summary

OUTPUT SUMMARY:

The main aspects of the trial such as the expected duration of the randomization, the randomization rates, and the number of screen failures are given in the Output Summary. This is a single table containing key input parameters and output values. Output summary table is displayed below:

	Wbk1:Plan1
Trial Details	
Group Sites By	Country
# of Groups	10
Total # of Sites	149
Plan	
Target Randomization	700
# of Interim Analyses	1
Interim Analysis Duration	15
Screening Rates Multiplier During IA	0.75
Ramp Up Factor	1
Site Information	
Distribution of Site Initiation Time	Uniform
Max # of Subjects Distribution	Fixed
Distribution of Site Failure	Binomial (0.1)
Recruitment Information	
Recruitment Model	Poisson
Distribution of Randomization Rate	Gamma
Randomization Rate Time Unit	Months
Distribution of Screening Failure Rate	Fixed (0.1)
Screening Period	1
Simulation Results (Overall)	
Mean Time from Trial Start to FSR	1.9
Mean Time from Trial Start to FSIV	0.4
Mean Time from Trial Start to LSR	22.7
Mean # of Sites Opened	149
Mean # of Subjects Screened per Site	5.22
Mean # of Screen Failures	77.936
Mean # of Sites with at least 1 Subject Screened	128.222
Mean Randomization Rate/Site	0.37
Output Time Unit	Months

Figure 16: Output Summary

PLOTS AND TABLES:

Graphical outputs can convey information in more meaningful manner. Hence, EnForeSys provides options to view these outputs in plots and tables, which makes it easier for the user to interpret. Different histograms and line plots are used to display these outputs graphically. Some of the important plots and tables are described below:

SUBJECTS RANDOMIZED VS. TIME

This plot displays the timeline of subject randomization. On the y-axis are the number of subjects randomized and on the x-axis we find the trial duration expressed in terms of elapsed time or calendar dates. The mean, median and 95% credible intervals are plotted in this plot. User has the flexibility to view this plot for the overall trial or a particular region or a particular site.

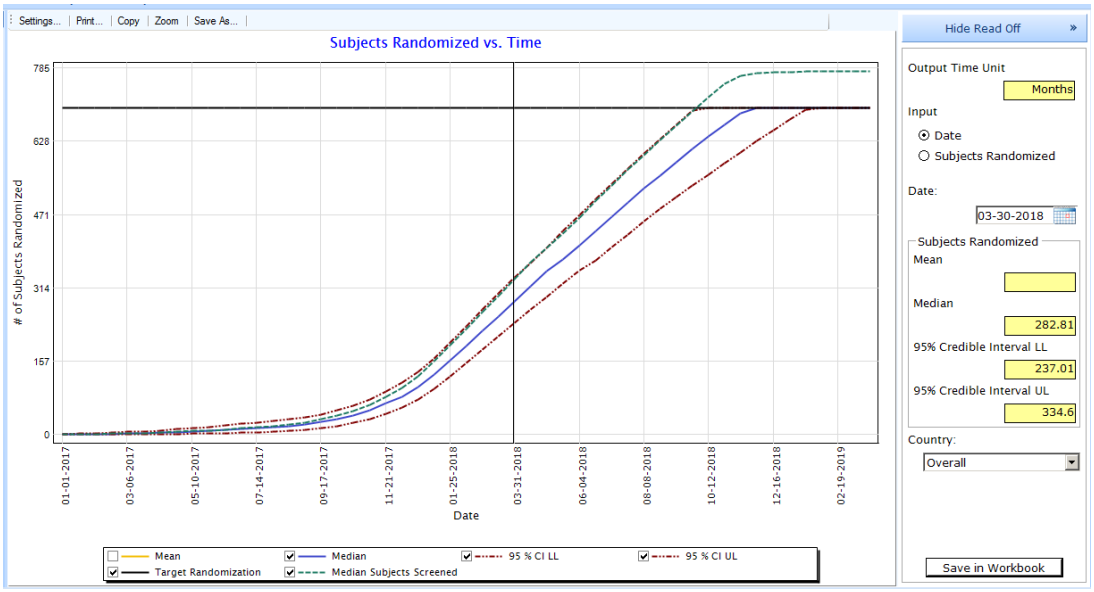


Figure 17: Subject Randomized Vs. Time

SITE PERFORMANCE PLOT

The Site Performance plot displays the number of sites for low, medium and high performance using three different colors, with the x-axis representing the number of subjects randomized against the y-axis of the number of sites.

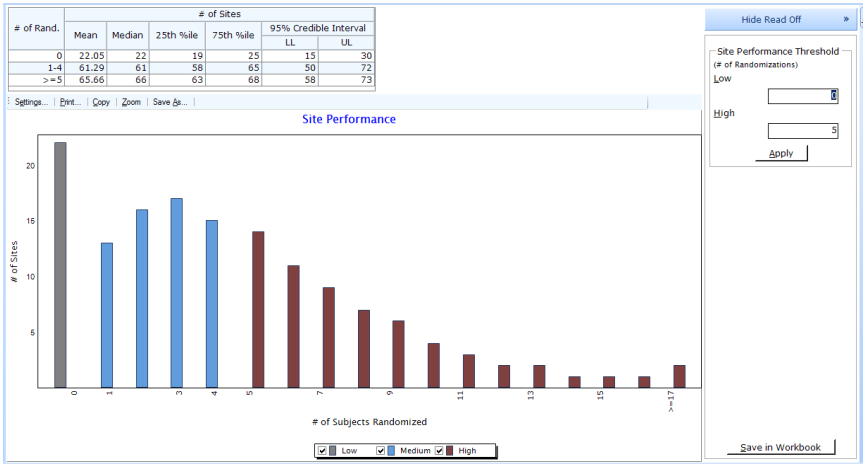


Figure 18: Site Performance Plot

MILESTONE PROBABILITY PLOT

The Milestone Probability plot displays the probability of crossing the milestones defined in the input stage at each date or elapsed time. The user can edit the milestone value and view the plot updated accordingly.

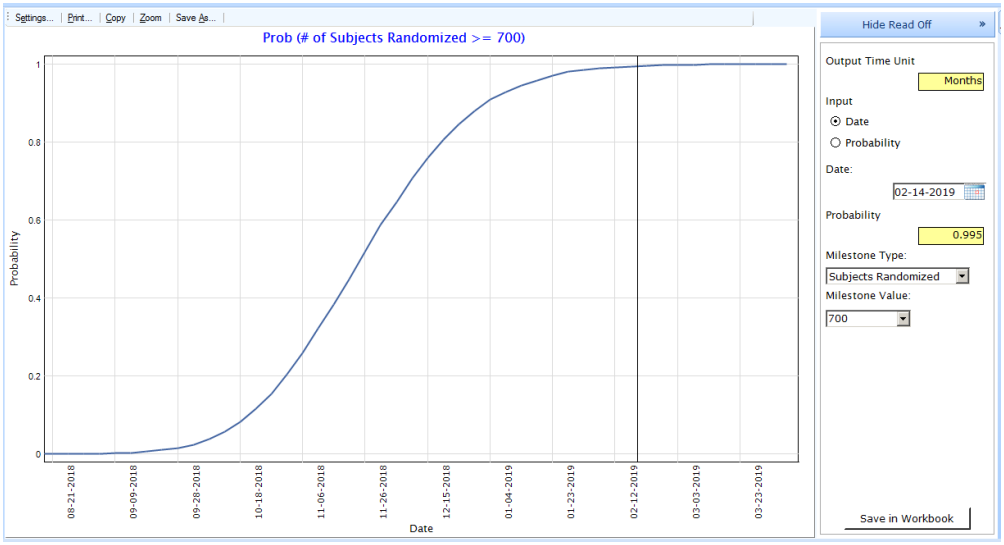


Figure 19: Milestone Probability Plot

USABILITY FEATURES:

IMPORT FUNCTIONALITY

EnForeSys provides an easier way to enter the Site and Recruitment information, by importing data through a CSV file. This is especially useful in the case of large number of regions when the number of parameters for each variable may change depending on the distribution. The parameters for the input variables can be mapped to headers of the CSV file.

EDIT PLAN

There may be certain scenarios where user might want to change a few parameters of a plan created earlier. In such cases instead of entering all the inputs again user can use the edit functionality provided by EnForeSys. When user edit's a plan all the inputs from that plan are inherited into a new plan and user can easily change the parameters as desired. To edit a plan, select the plan in the Library pane and click the Edit Plan button, or alternatively simply right-click on the plan itself to get the edit option. Both these actions open the input window with inputs inherited from the plan.

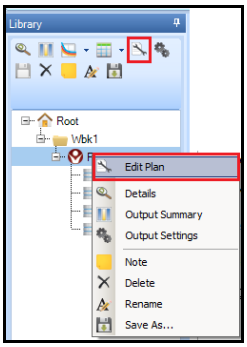


Figure 20: Edit Functionality

COMPARE TWO PLANS

Users can compare the outputs, plots and tables of two or more plans by using the Compare option. User simply needs to select the plans to be compared and invoke the desired outputs. This will display the side by side comparison of the outputs from the selected plans.

	Wbk1:Plan1	Wbk1:Plan4
Trial Details		
Group Sites By	Country	Region
# of Groups	10	7
Total # of Sites	149	50
Plan		
Target Randomization	700	1000
# of Interim Analyses	1	
Interim Analysis Duration	15	
Screening Rates Multiplier During IA	0.75	
Ramp Up Factor	1	1
Site Information		
Distribution of Site Initiation Time	Uniform	Uniform
Max # of Subjects Distribution	Fixed	Fixed
Distribution of Site Failure	Binomial (0.1)	Binomial (0.1)
Recruitment Information		
Recruitment Model	Poisson	Poisson
Distribution of Randomization Rate	Gamma	Gamma
Randomization Rate Time Unit	Months	Months
Distribution of Screening Failure Rate	Fixed (0.1)	Fixed (0.1)
Screening Period	1	0
Simulation Results (Overall)		
Mean Time from Trial Start to FSR	1.9	0.6
Mean Time from Trial Start to FSIV	0.4	0.3
Mean Time from Trial Start to LSR	22.7	15.3
Mean # of Sites Opened	149	50
Mean # of Subjects Screened per Site	5.22	22.22
Mean # of Screen Failures	77.936	111.105
Mean # of Sites with at least 1 Subject Screened	128.222	44.274
Mean Randomization Rate/Site	0.37	2.465
Output Time Unit	Months	Months
Other Parameters		
Trial Name	Impova	Trial1
Trial Start Date	01-01-2017	08-27-16
Phase	3	3
Therapeutic Area	Allergic Disorders	Allergic Disorders

Figure 21: Compare Plans

DISPLAY PRECISION

User can change the display precision of the outptus, either globally for all the plans to be created or locally for a particular plan. To change the display precision globally, we can use the global option from the ribbon bar. And for changing this locally we can use Output Settings option available in the library.

CONCLUSION

The paper elaborates key features of EnForeSys, a software which uses Monte Carlo Method for patient enrollment, and provides a more realistic forecasting technique which captures the possible risks and outcomes for any given scenario. It accounts for the nonlinearity and randomness of real-life enrollment processes. It identifies how different inputs like accrual rates, screening failures and site-specific parameters can affect the patient recruitment strategy.

Across many functions within the drug development industry, modeling and simulation methods are already gaining popularity, and will soon become indispensable tools for clinical operations. Trial planners who take advantage of such methods will be able to significantly reduce their study costs and timelines, while maximizing their chances of study success.

REFERENCES

EnForeSys User Manual
Impact Report Tufts CSDD, 2013

ACKNOWLEDGMENTS

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

CONTACT INFORMATION

Ritika Yadav ritika.yadav@cytel.com +91-20-67090212
Swechhya Bista swechhya.bista@cytel.com +91-20-67090248

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