PhUSE EU Connect 2018

Paper TT08

Simulating Enrollment Plans

Jim Box, SAS Institute, Cary, NC, USA

ABSTRACT

Developing a plan to reach a Study Enrollment goal can be a tricky business. Choosing countries and sites with their associated rates and costs complicates things. Simulation is an excellent way to understand sources of variability and comparing options to meet the enrollment goals.

INTRODUCTION

One of the most important things to do once a clinical trial has been designed is to develop a plan for enrolling subjects. Decisions about what countries to use and what sites in those countries are driven by striking a balance between the anticipated enrollment rate and the costs involved in the regulatory process of initiation countries and sites and with the individual payments to the sites for the enrolled subjects. Many assumptions are made on these rates and costs, usually from combining previous experience and a feasibility study. In a feasibility study, a team will contact many potential sites and ask questions about the types of patients they typically see with the therapeutic area of interest. Those questions tease out the following information:

- How long will it take to get your site initiated (ready to enroll subjects)?
- How many subjects do you expect to enroll each month?
- How much will it cost to get initiated? How much will it cost to treat the subjects per the protocol?

For example, a site might say it will take six months to get initiated and they will enroll 10 subjects per month, so the site will contribute zero for the first six months and 10 per month after that. In general, when putting together an enrollment plan for a proposal, this sort of estimation is done for several sites and summed up over time until the required number of subjects are enrolled. This is a reasonable approach to get a ballpark estimate of how long enrollment would take, but it is missing a key component: an understanding of the expected variation. Simulation is a way to get an understanding on the width of possibilities for a given enrollment plan, and it provides a method of comparing multiple possible scenarios. It also allows for leveraging the simulated data to assist in budgetary and supply chain planning.

SIMULATION METHODS

There are multiple ways to approach simulation, but for a problem of this nature, discrete-event simulation (DES) is an excellent modeling approach. DES looks at how a system changes over small slices of time. A typical application is in modelling customers interacting with a queue – arrival times are sampled from a distribution, customers wait in line while tellers are seeing other people, and finally they are seen by the teller and leave the system. The simulation tracks the individuals as they go through the system, and the state of each individual changes over time. Subjects going through a clinical trial can be modeled in the same way. Figure 1 shows a very simple approach to enrolling subjects at a single site.

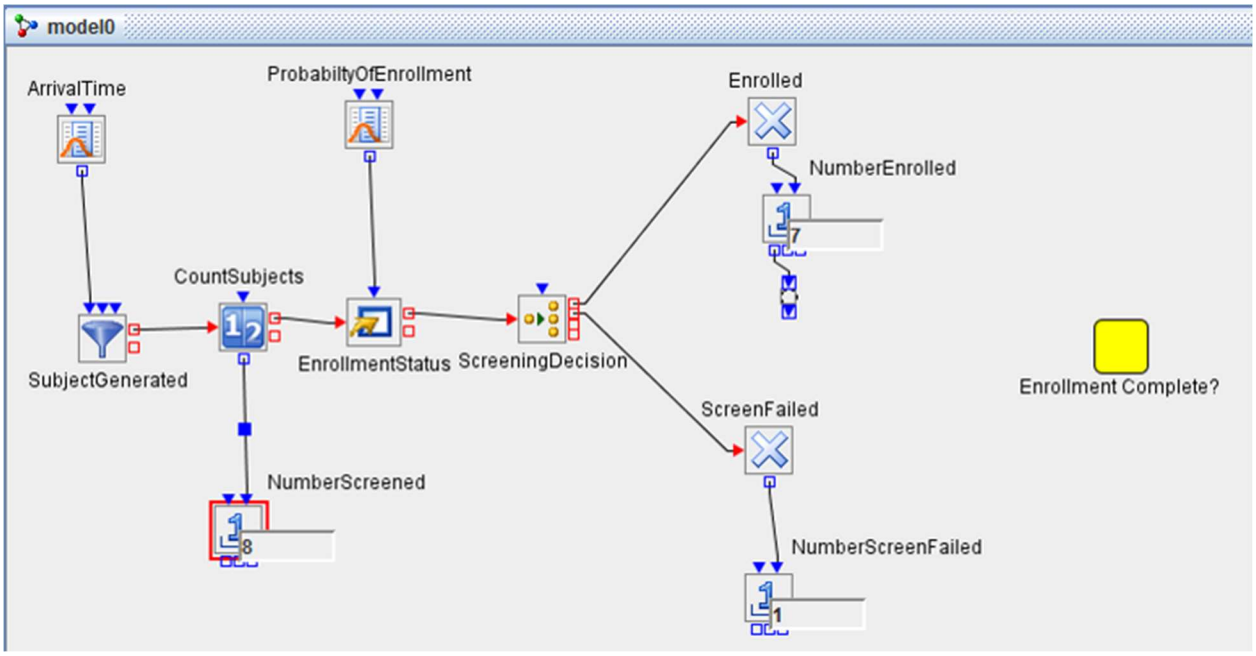

Figure 1. Discrete Event Simulation Flow for Single Site Enrollment

The first step in a DES model is to design the flow of the system. In this simple case, subjects arrive at the site, and then either follow the path to study enrollment or are screen failed out of the process. There is a stopping rule established – in this case, a count of the number of subjects enrolled – that determines when the system stops new subjects being generated. Finally, probability distributions are defined for the random events. In this system, the time between subject arrivals is sampled from a distribution that is characterized by how many subjects are expected to arrive per month. That is, if 15 subjects per month are expected, then on average a subject arrives every two days. The screening decision is simulated from a binomial distribution with a specified probability of success (say, 90%).

In a more complex model, the system would also account for countries being approved for the study, sites getting initiated and starting to enroll subjects, and the subjects themselves moving through the trial. The system would have many changes over time, which is a key aspect of discrete event simulation.

The most important thing about a simulation of this nature is repetition – a single run through this system would produce a single estimate of the time and cost of enrolling the study. Repeating the simulation multiple times (typically 100) will produce distributions of the time and cost values, which will lead to a deeper understanding of the likely outcomes.

SIMULATION PARAMETERS

One example of a more complex model for enrollment will have these key points:

1. A start time for the study is established (Day 0).
2. Countries are selected and get approval to start initiating sites. Countries will have different startup time periods.
3. Sites are initiated in the countries and can start enrolling subjects
4. Subjects start arriving at sites for screening.
5. Some subjects fail the screening process, others are enrolled into the study.
6. Subjects progress through the study. Some discontinue early, others reach the last scheduled visit.
7. The study reaches the enrollment target and no new subjects enter the screening process.
8. The study remains operational until all subjects finish the study or drop out.

As subjects are progressing through the study, the costs incurred for visits should be tracked, as well as the initiation costs for the sites and the countries.

When designing a discrete event simulation, it is vital to identify the key parameters that will be used to define the probability distributions and costs. These parameters will be very different for different therapeutic areas and will vary across the sites and countries. Some of the key parameters to consider are as follows:

COUNTRY LEVEL PARAMETERS

The country level parameters focus on the time between project start and site identification/initiation.

- **Country Startup Time** – the time between the decision to start a trial (Day 0) and the beginning of the site initiation process. This will typically vary due to the time to get regulatory approval in the different countries.
- **Startup Cost** – how much does this cost.
- **Startup Probability** – it's useful to consider the chance that approval will never be achieved in a country as you are working through the startup process – maybe there is a 99% chance of getting approval in the US vs an 85% chance in Brazil – some of the replications of the study might have only one country successfully started up.
- **Concurrent Site Startups Possible** – this is a constraint to the system. For example, you may plan on starting up 40 sites but only have the personnel for 30, so the 31st site to attempt to start up would sit in a queue until staff became available.

SITE LEVEL PARAMETERS

Once the country successful starts up, the site initiation process can begin. There are two classifications of sites:

- Primary sites are specific investigators that have been identified and for which the parameters of interest are known
- Candidate sites are more akin to "batches" of sites that all have the same parameters. For example, a set of candidate sites might be classified as "fast enrolling with a central Institutional Review Board"

The site parameters of interest include:

- **Time-to-Identify (Candidate Sites Only)** – a distribution of the number of days it will take to find specific sites from this candidate group
- **Number Available (Candidate Sites Only)** – how many sites you can get from the specific candidate group
- **Startup Time** – distribution of the number of days between when the site is ready (identified and country is approved) and the site is ready to open for enrollment
- **Startup Cost** – the cost for getting the site through the initiation process
- **Startup Success Probability** – a value that represents the chance a site will successfully get through initiation
- **Probability of Zero Patients** – there is always a chance that the site will get initiated and can enroll, but for whatever reason never does – essentially the site gets lost to follow-up
- **Expected Recruitment Rate –** this is generally expressed as the number of patients per month to star the screening process. Often, sites are overly optimistic about how many patients they will see per month, and it is useful to have a way to address that. One approach is to break this rate into two periods – for example, the site will get 20 patients per month for the first six months, then see 12 per month after that.
- **Available Subjects –** it is a good practice to have a cap on the number of subjects that could possibly come from one site
- **Per Subject Cost (Screening) –** this is the cost incurred by the site to screen subjects
- **Per Subject Cost (Enrolled) –** this is the cost incurred by an enrolled subject who completes the study. Subjects who discontinue early would get a pro-rated value allocated to them.

STUDY LEVEL PARAMETERS

At the study level, it is important to identify how long the trial will last, and what sort of discontinuation rate is expected.

- **Discontinuation Rate** – the percent of subjects who discontinue early for any reason. It can be useful to split this up into cohorts – for example, the 60% in the active arm will see a 20% dropout rate, while the 40% in the control arm will see a 25% rate
- **Schedule of Assessments** – Identify the study days for visits. At each of the study visits, there would be a chance for each subject to discontinue, to get the overall discontinuation rate identified above

SCENARIO LEVEL PARAMETERS

Scenarios are a combination of countries, sites and a study that the simulation will be executed on. Generally multiple scenarios will be compared to determine the optimal approach. For example, once scenario may be a trial that will be run in 10 specific sites in the US, and the comparison scenario would look at adding up to 10 possible sites in Brazil, with the understanding that it will take longer and cost more to get Brazil up and running, but once it is operational, sites will enroll faster and the subjects will cost less. Will this have a positive impact on enrollment times and cost? The ability to address such a question is the key benefit of doing simulation.

- **Country Selection** – identify which countr(ies) to include in the simulation.
- **Country Screen Failure Rate** – each country will have its own probability of a subject failing screening.
- **Site Selection** – identify which sites to include in each country. If using candidate sites, provide the maximum number of sites.
- **Target Enrollment** – this is the total number of subjects you are trying to enroll. Once this target is reached, enrollment stops, but the simulation continues until every subject exits the trial.
- **Maximum/Minimum Subjects per Country** – consider putting caps on the total number of subjects who can enroll from any specific country.
- **Target Start Date** – this is Day 0, the start date of the simulation. All other dates are in refence to this.
- **Number of Replications** – decide how many times to re-run the simulation. 100 times gives a good estimate of variation without too much risk of producing extreme edge cases that may skew the results.

RESULTS & BENEFITS

The results from running multiple replications of a scenario will provide a good understanding of the possible variation. Figure 2 shows the results of the scenario with using just the 10 sites identified in the US.

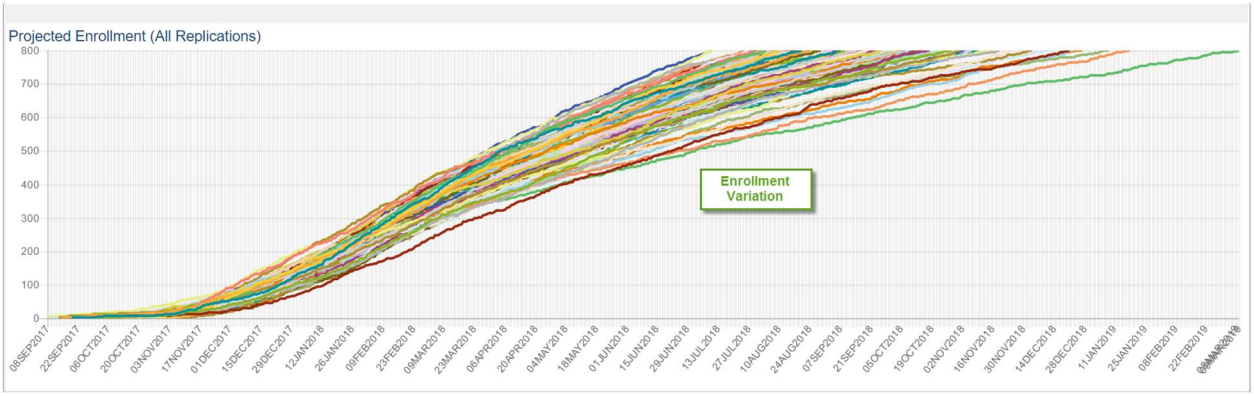

Figure 2. Simulation Results for a Single Country Scenario

The observed variability is due to randomness in startup times, enrollment rates, and the number of sites that successfully enroll any subjects. Having a greater appreciation for what might reasonably happen allows for better planning.

Figure 3 shows a scenario using the same 10 US sites, but also including up to 10 sites in Brazil. The Brazil sites have a higher risk of failure to start (as does the country startup), but once started will contribute subjects faster.

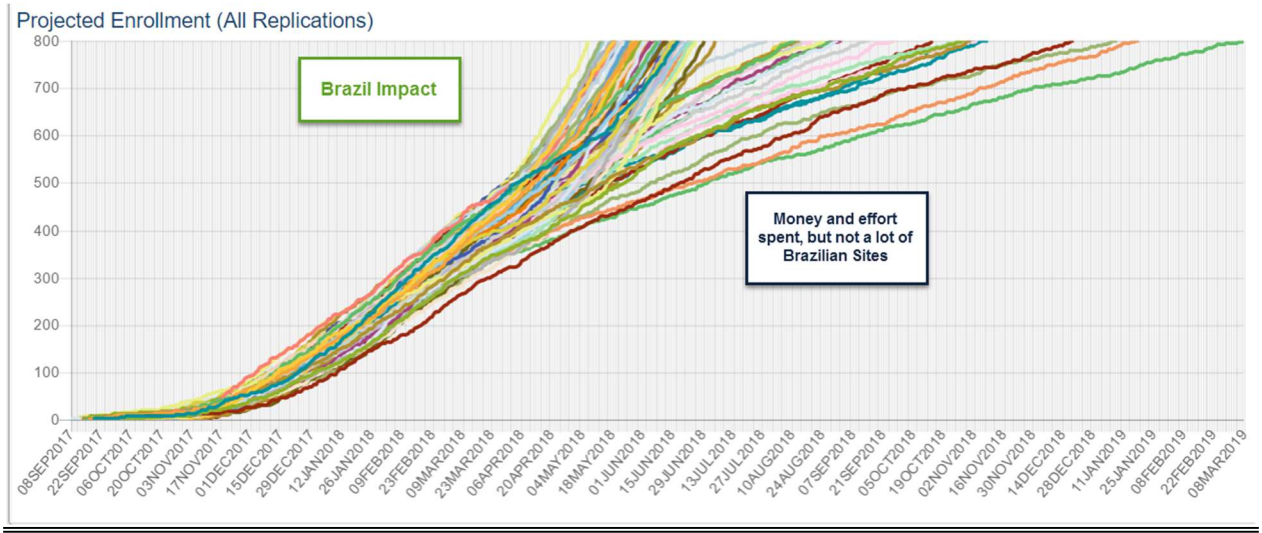

Figure 3. Simulation Results for a Two Country Scenario

Once a scenario is chosen and an enrollment plan is created, one of the replications from the scenario is chosen to anchor the expectations. Generally, the median replication is the best choice, but it might make sense to choose a different one, depending on the amount of risk involved for the team.

One of the key benefits of doing simulation is the creation of subject-level datasets for each replication. These datasets allow for further analysis and planning opportunities. Figure 4 shows the expected trial visits for a chosen replication. This can be used for managing supply chain and for developing expectations around monitoring visits.

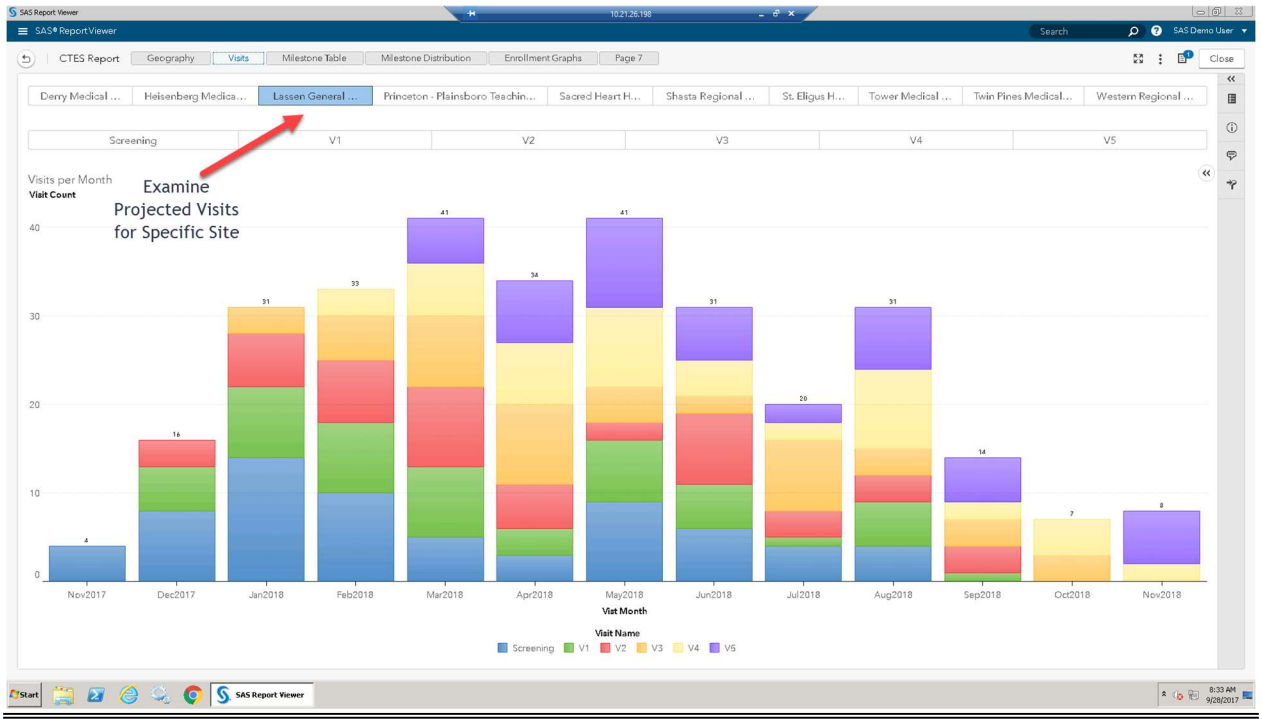

Figure 4. Projected Site Visits for a Chosen Replication

Figure 5 shows a distribution of the last patient last visit dates – this could be useful to evaluate the overall acceptability of the scenario. For example, we might learn that there is a 5% chance the LPLV will not happen by the desired end date, and other scenarios must be explored.

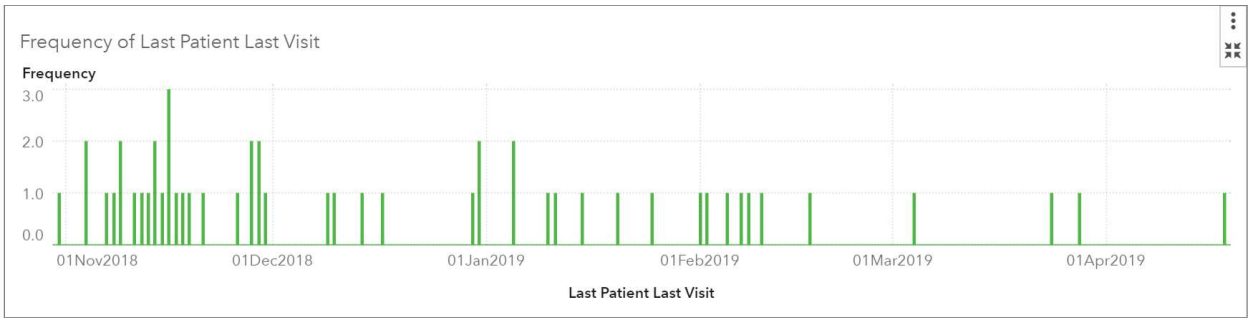

Figure 5. Distribution of Last Patient Last Visit dates

Figure 6 shows a cost breakdown by all subjects for a specific replication. This can be used for budget forecasting and can easily be broken down by site and by time, to further aid in budget creation.

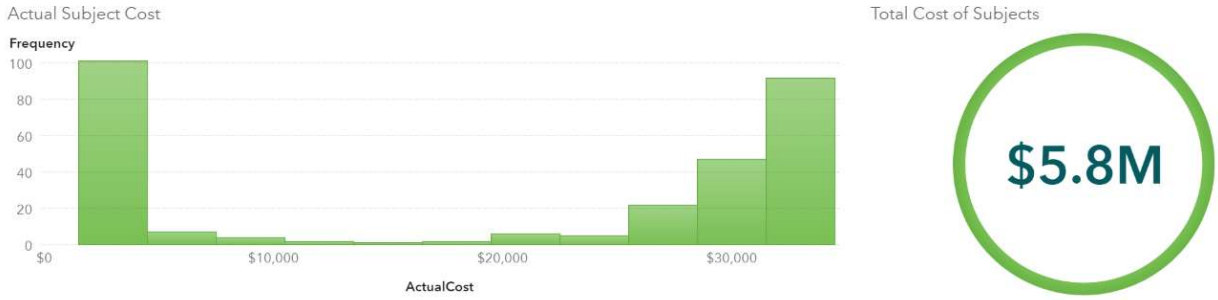

Figure 6. Distribution of Subject Costs for a Chosen Replication

CONCLUSION

Creating enrollment plans for clinical trials is a complicated endeavor. Static methods can give a reasonable ballpark estimation, but simulation methods can give a much deeper understanding of possible outcomes. Discrete-event simulation allows for the comparison of many potential scenarios and is a source of rich data that can be used to project key milestones and costs that allows for more comprehensive clinical trial planning.

CONTACT INFORMATION

(In case a reader wants to get in touch with you, please put your contact information at the end of the paper.)
Your comments and questions are valued and encouraged. Contact the author at:

Jim Box
SAS Institute
Cary, NC, USA

Email: jim.box@sas.com