Proceedings of the 2021 Winter Simulation Conference
S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo, and M. Loper, eds.

SENSITIVITY ANALYSIS IN CLINICAL TRIAL SIMULATION AT SAS INSTITUTE

Wendy Xi Jiang
Bahar Biller
Jim Box

SAS Institute Inc.,
Cary, NC 27513, USA

Barry L. Nelson

Dept. of Industrial Engr. & Mgmt. Sci.
Northwestern University
Evanston, IL 60208-3119, USA

ABSTRACT

Clinical trial enrollment is expensive, important, and subject to many uncertainties. Simulation captures these uncertainties, so SAS® Institute created the Clinical Trial Enrollment Simulator (CTrES) as a tool specifically for enrollment planning. However, simulation provides no mathematical expression from which to extract sensitivity measures that are critical for problem diagnosis and management. This paper describes sensitivity analysis technology created for CTrES requiring only the output data obtained from simulation of the base scenario, and demonstrates it on a realistic enrollment planning problem for the United States.

1 INTRODUCTION

The design of any clinical trial includes the development of a plan to enroll a target number of patients while remaining within an available budget. Clinical trial enrollment planning can be a daunting task for clinical research organizations (CROs) and pharmaceutical companies considering the level of uncertainty under which the planning is done. Given the tight deadlines for creating the enrollment plan and the difficulty in capturing the sources of uncertainty, these plans often ignore the variability in the process and create inaccurate predictions of total cost and total time for enrollment. SAS® Institute has been partnering with the healthcare industry for 40 years and has developed an analytical tool known as the SAS Clinical Trial Enrollment Simulator (CTrES) for CROs and pharmaceutical companies. This tool equips its users with the power to develop high-reliability plans for enrolling patients in clinical trials. SAS is a founding member organization of the CEO Roundtable on Cancer, which is committed to the health and well-being of employees with the belief that cancer can be prevented, and lives can be prolonged (Goodnight 2007).

There are three sequential events that affect the enrollment timeline of a clinical trial: (i) starting clinical research efforts in a country; (ii) activating the clinical research sites in that country; and (iii) enrolling and tracking patients who arrive at each site. The timing of these events and their successful execution determine the performance of the clinical trial enrollment plan. The typical key performance indicators (KPIs) are the time it takes to enroll a target number of patients in the clinical trial, and the total cost of starting up the countries, activating the sites, enrolling patients and tracking the enrolled patients. Of these, the time to enroll patients is the most important. Obtaining accurate predictions of these KPIs is often challenging because the events of country start-up, site activation, and patient enrollment and tracking are connected through a sequence of subprocesses, each of which is subject to a high level of uncertainty.

The following are some representative subprocesses corresponding to the events (i)–(iii) enumerated above that can be the reasons why a clinical trial enrollment plan achieves low patient enrollment or high cost. After preparing the core regulatory package and completing the regulatory timeline under event (i), the pharmaceutical company could be unsuccessful in obtaining regulatory approval in a country while still incurring the country activation cost. After collecting information about a site, waiting for the availability of personnel, and spending the time needed to start up the site in event (ii), site activation may still fail. Even if a site is successfully activated, it may fail to enroll patients. Moreover, after the arrival of patients,

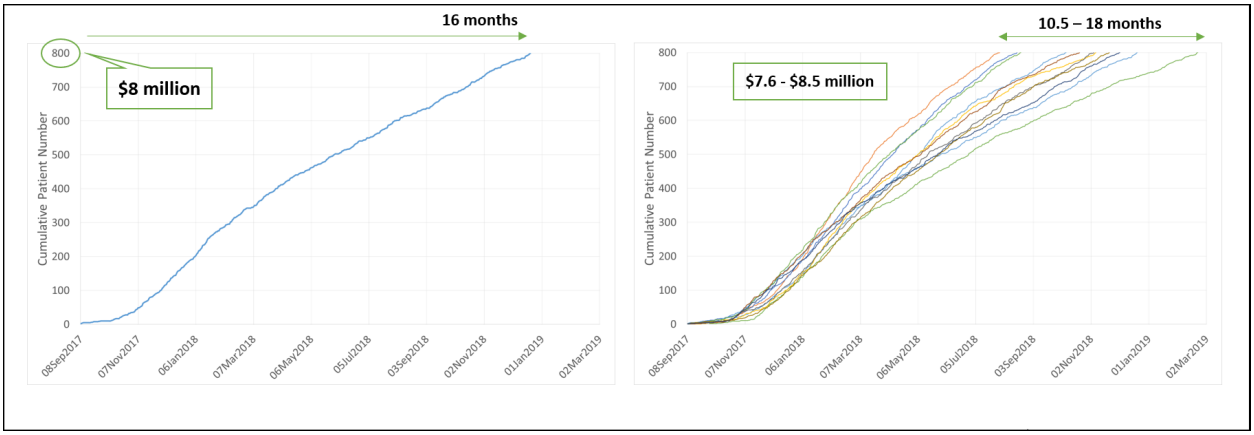

Figure 1: Illustrating time versus patient enrollment in deterministic (LHS) and stochastic (RHS) solutions.

only the successful completion of screening will result in the actual enrollment of patients in the clinical trial under event (iii).

Thus, there is a high degree of uncertainty at every step of clinical trial enrollment planning, from the probability that a single site will succeed to enroll patients to the random arrival of patients to a potential site. In their 2013 impact report, the Tufts Center for the Study of Drug Development noted that as many as 37% of sites missed their enrollment targets and 11% failed to enroll a single patient. This lack of certainty turns enrollment planning into a difficult task. In fact, 80% of clinical trials fail to meet enrollment timelines, and one third of Phase III clinical trial terminations stem from poor patient enrollment planning (Cognizant 2015). Failure to reach the target patient enrollment in time could lead to delays in getting medicine to the market and result in significant cost overruns.

The industry practice in clinical trial enrollment design is to make many assumptions about enrollment rates and various components of cost, motivated by experience and learning from feasibility studies (Box 2018). In a feasibility study, a team contacts potential sites and asks questions about the types of patients that they typically treat in the therapeutic area of interest. The team also gathers answers to the following questions: (a) How long would it take to get your site ready to enroll patients? (b) How many patients would you expect to enroll each month? (c) How much would it cost to get ready for enrollment and how much would it cost to treat the patients according to the protocol? The answers to these questions are used to obtain a rough estimate of how long it would take to enroll a target number of patients.

An example of this rough estimate is provided on the left-hand side (LHS) of Figure 1. This figure plots the cumulative number of patients enrolled (y-axis) against time (x-axis) and presents the associated total cost to enroll, say, 800 patients. As implied by the construction of a single path on the LHS, the rough estimate based on the data from a feasibility study lacks any formal quantification of risk. This is an example of a deterministic but incomplete solution to the problem of KPI prediction in clinical trial enrollment planning. However, accounting for the uncertainty in the inputs provides a range of between 10.5 to 18 months for the time it takes to enroll 800 patients on the right-hand side (RHS) of Figure 1, which is generated by a CTrES simulation. Similar statements can be made for the total cost. The two prediction intervals for the total cost and the time it takes to enroll 800 patients clearly demonstrate the significant impact of input risk on KPI variability. The capability to quantify this risk for CROs and pharmaceutical companies has two noteworthy benefits: First, it informs them about the level of risk in their cost and enrollment predictions; second, it guides them towards the identification of enrollment plans to reduce uncertainty.

Stochastic simulation is a natural choice to capture the risk arising in different stages of a clinical trial enrollment plan. The use of simulation to mimic the clinical trial enrollment process can help overcome the three primary challenges of clinical trial enrollment planning (Handelsman 2012): 1) The patient enrollment

process consists of a long sequence of dynamic random events; 2) the hierarchical relationship among country startups, site activations, and patient screening and enrollment complicates the process of design and analysis of patient enrollment; and 3) enrollment planning must be driven by country, site, and patient data sets, and the solution must be robust to data uncertainty and scalable to any number of countries and sites under consideration.

In addition to the classical problem of KPI prediction, examples of the what-if questions that planners want to ask are the following: If mean site activation delay increased by 1 week, how would the mean KPI change? If mean screening failure probability increased by 1%, how would the mean KPI change? If the standard deviation of site activation delay increased by one week, how would the mean KPI change? Obtaining answers to these what-if questions helps CROs and pharmaceutical companies diagnose the current setup sensitivities and decide where to put management effort towards the design of better clinical trial enrollment plans. Answering these questions is the topic of this paper.

Each of these questions can be answered by creating a new scenario in the SAS CTrES User Interface. Specifically, the first question can be addressed by creating a second scenario where the mean site activation delay is increased by one week, and the simulation output data obtained from these two scenarios are compared. Unfortunately, a typical enrollment planning exercise may involve multiple countries and hundreds of sites. A study of the SAS CTrES simulation engine for a single-country, 10-site setting reveals 51 different stochastic inputs to support enrollment planning (Biller et al. 2019). Thus, at least 52 computationally intensive simulations (initial simulation + varying each input) would be needed just to evaluate the sensitivity to changes in the means for one possible scenario of countries and sites to activate. Thus, CTrES currently lacks the capability to quickly answer what-if questions in a way that scales with the number of countries and sites involved in a clinical trial design. *Our work reported here enables CTrES to overcome this limitation and equips CTrES with the power to answer what-if questions for any number of stochastic inputs using the output data obtained from simulation of the base scenario only.*

Answering the types of what-if questions posed above for the stochastic inputs of the simulation is a type of *local sensitivity analysis*, which focuses on the influence of the inputs on the output near a nominal setting. While SAS already has global sensitivity analysis capabilities, it does not support the type of local sensitivity analysis CTrES requires. *The focus of this paper is creation of local sensitivity analysis technology for CTrES.* Although the methods presented here were created for CTrES, they are broadly applicable to many simulation contexts. In this paper we use the term "local" to refer to small changes in inputs, "global" to refer to varying an input across its entire range, and "nominal" to refer to the baseline simulation and its parameters.

2 LITERATURE

Clinical trial enrollment planning has been studied from different perspectives for different purposes. However, most published research makes significant simplifying assumptions to formulate the problem as a mathematical model that is tractable.

From the perspective of production planning and supply chain design, the key is to position the right inventory of drugs at the right time at the right trial site considering both the cost of production, shipping, inventory carrying, enrollment, and duration of the clinical trial; e.g., see Zhao et al. (2018), Zhao et al. (2019). The problem is formulated as a multi-stage stochastic program and the only uncertainty considered is the number of patients, which is modeled as a countable number of scenarios where each scenario represents a possible realization based on previous trial data. Furthermore, the enrollment cost is either not considered or assumed to be independent of patient arrivals, which seems unrealistic in the scenarios modeled by CTrES.

Kouvelis et al. (2017) study the problem of maximizing the expected net present value of a drug considering the costs of clinical trial, the drug's likelihood of approval, and its subsequent expected revenue if approved given the maximum duration of the study. The problem is modeled as a discrete-time, discounted dynamic program determining when and how many test sites should be opened and the rate at which patients

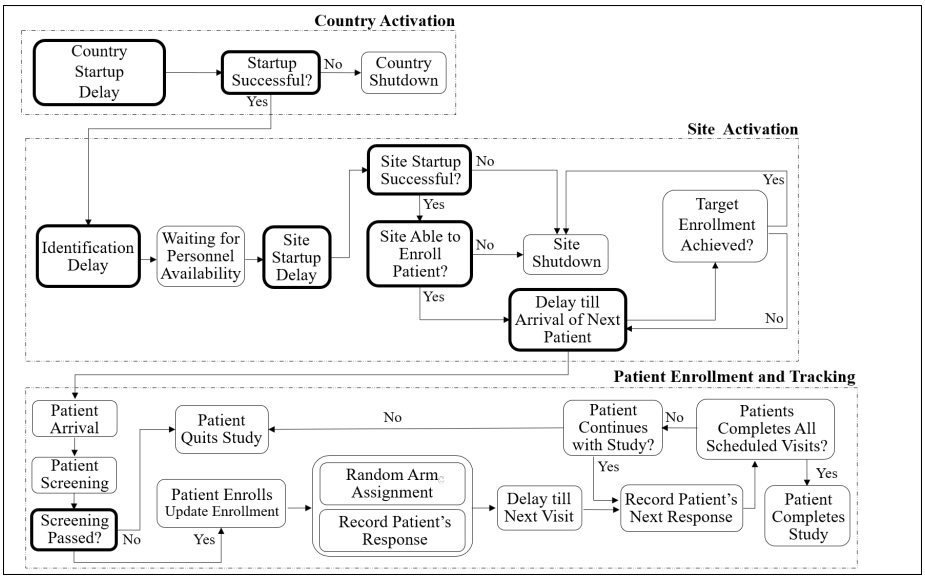

Figure 2: High-level view of clinical trial enrollment process flow.

should be recruited to achieve the optimum. To simplify the analysis, the paper assumes that the sites will be opened in a given order, which is restrictive unless all sites have identical capacity and zero startup cost. Moreover, under most cases, the recruitment rate is not controllable but rather a site-specific characteristic with uncertainty.

There are also many studies focusing on modeling of patient recruitment, e.g., Monte Carlo simulation models in Abbas et al. (2007) and the Pareto-Poisson statistical model in Mijoule et al. (2012). The most widely used is the Gamma-Poisson model in the empirical Bayesian framework proposed by Anisimov. This purely statistical model not only enables the prediction of recruitment with confidence bounds, but also evaluates various site performance measures and approximates the minimal number of sites needed with confidence (Anisimov 2008; Anisimov 2009; Anisimov 2016). The model accounts for the natural variation in recruitment over time, in recruitment rates among different sites, and in site startup delays (Anisimov 2008). However, the real-life clinical trial enrollment process is far more complex because of the uncertainty associated with site startup and enrollment success, and the patient screening success. Mijoule et al. (2012) further study to what extent estimation error of the arrival rate generates an error in the prediction of the trial duration, which is known as "input uncertainty" in the simulation literature.

3 THE CLINICAL TRIAL ENROLLMENT MODEL

Figure 2 presents a high-level illustration of the CTrES process flow, which is implemented in SAS Simulation Studio, a Java-based discrete-event simulation tool (Hughes et al. 2018). Thus, SAS Simulation Studio serves as the engine for SAS CTrES to address clinical trial enrollment planning questions for SAS customers; it is made available through a web interface as software as a service.

The simulation model is composed of three modules consistent with the three events introduced in Section 1: (i) country activation, (ii) site activation, and (iii) patient enrollment and tracking. Each module introduces a specific entity flowing through the corresponding portion of the logic illustrated in Figure 2: (i) Country entities in the Country Activation module, (ii) Site entities in the Site Activation module, and (iii) Patient entities in the Patient Enrollment and Tracking module. Each entity has attributes that are subject to uncertainty characterized by probability distributions based on expert opinion and historical data. Within each replication, the realized value of each uncertain attribute of each entity is updated after the corresponding subprocess and recorded right before leaving the corresponding module.

Table 1: Distributions relevant for local sensitivity analysis.

Level	Uncertainty	Input Distribution	Input Parameters
Country	Startup Delay	Triangular Distribution	Min, Mode, Max
	Startup Success	Bernoulli Distribution	Probability
	Screening Failure	Bernoulli Distribution	Probability
Site	Startup Delay	Triangular Distribution	Min, Mode, Max
	Startup Success	Bernoulli Distribution	Probability
	Enrollment Success	Bernoulli Distribution	Probability
	Identification Delay	Triangular Distribution	Min, Mode, Max
	Patient Arrival Rate	Nonstationary Poisson	High Rate, Low Rate
	Duration of High Rate	Triangular Distribution	Min, Mode, Max

Box (2018) outlines the following key points for the CTrES process flow: (a) A start time is established as Day 0 for the clinical trial study. (b) Countries are selected and may receive approval after a certain duration of delay. Countries may have different values for the startup delay during country activation. Once a country successfully starts up, site initiation begins. (c) Sites are initiated in the countries that start up successfully and can start enrolling patients. (d) Patients start arriving at sites that are activated successfully and able to enroll for screening. (e) Some of the patients fail the screening process while those passing the screening test are enrolled in the study. (f) Patients progress through the study. Some of the patients quit the study early while others reach the last scheduled visit. (g) As soon as total patient enrollment reaches the target enrollment, the patient arrival process is terminated. (h) The study remains operational until all the patients that are still flowing through the system either finish the study or drop out.

The two primary KPIs of interest for a CTrES user are the time it takes to enroll a given target number of patients, say 800, denoted as "TimeToEnrollTarget," and the associated total cost of the clinical trial, denoted as "TotalCost," which is the sum of country and site activation costs and the costs of screening and enrolling 800 patients. There are two other timeline KPIs: "FirstTimeEnrolls," which is the time the first patient enrolls, and "EnrollmentDuration," which is the time between "FirstTimeEnrolls" and "TimeToEnrollTarget." Figure 2 shows that the enrollment of a patient is recorded after passing the screening process and before the execution of the follow-up response model. Thus, the two primary KPIs of interest and the other two timeline KPIs are updated and recorded in the module for patient enrollment and tracking.

For these KPIs, only the stochastic inputs associated with countries and sites are relevant for the development of the local sensitivity analyzer for SAS CTrES. Table 1 lists those uncertain inputs and their corresponding probability distributions, which are the sources of uncertainty in the process flow illustrated in Figure 2. The use of the three-parameter triangular distribution to capture the uncertainty associated with the length of subprocesses is common practice so that expert users can provide the corresponding input parameters, i.e., minimum, most likely, and maximum values. The Bernoulli distributions are used to capture the uncertainty associated with a subprocess happening or not. Notice that although the enrollment of each patient is subject to the probability of passing screening, the screening failure distribution is specified at the country level.

The input that is quite different and worth more explanation is the patient arrival process, specified at the site level. The arrival process of patients is characterized by a piecewise-constant non-stationary Poisson process (NSPP) with two pieces because sites tend to have patients arriving at a higher rate at the beginning of the clinical trial. This is followed by a lower arrival rate after a certain duration of high enrollment. Moreover, there is uncertainty about this length of time that the arrival rate is high; a triangular distribution is used to capture this uncertainty.

4 SENSITIVITY MEASURES AND NEW CHALLENGES

The what-if questions described in the Introduction can be summarized as the quantification of the expected change in the mean KPI per unit change in the mean or standard deviation of each uncertain subprocess, as characterized by probability distributions specified in Table 1. Since the KPI is the simulation output and we are interested in its mean, where no confusion will arise, we will redefine KPI as the expectation of the simulation output from now on, i.e., KPI $\equiv \mathrm{E}(\text{output})$. Therefore, the goal is to measure the sensitivity of each KPI to the mean or standard deviation of each input distribution, near a nominal setting. This goal fits in the framework of output-property-with-respect-to-input-property sensitivity measures proposed in Jiang et al. (2019) and Jiang et al. (2021). The sensitivity measures of interest in the context of CTrES are two special cases of the general family: *mean sensitivity to mean* (MSM) and *mean sensitivity to standard deviation* (MSSD). The MSSD measure is built upon the *mean sensitivity to variance* (MSV) measure described in Jiang et al. (2021) through replacing the variance with the standard deviation.

For ease of explanation, we focus on a single output and a single input distribution. Let Y be the simulation output, $\mathrm{E}(Y)$ be one of the KPIs, and $X \sim F(\cdot|\boldsymbol{\theta})$ be one of the uncertain inputs that are listed in Table 1 with distribution parameter $\boldsymbol{\theta}$. Without loss of generality, let $\boldsymbol{\theta} \in \Re^p$ where $p \geq 1$. Further, let $\mu = \mu(\boldsymbol{\theta})$ and $\sigma = \sigma(\boldsymbol{\theta})$ be the mean and standard deviation of input X, both of which are differentiable with respect to $\boldsymbol{\theta}$ around the nominal setting $\boldsymbol{\theta} = \boldsymbol{\theta}^0$.

Recapping the definition introduced in Jiang et al. (2021), the MSM measure is defined as the directional derivative of $\mathrm{E}(Y)$ with respect to μ along a normed direction $\vec{\mathbf{d}}$ from the nominal parameter setting $\boldsymbol{\theta}^0$:

$$\mathrm{MSM}_{\vec{\mathbf{d}}} = \frac{\partial \mathrm{E}(Y)}{\partial \mu_{\vec{\mathbf{d}}}} = \frac{\vec{\mathbf{d}}^T \nabla_{\boldsymbol{\theta}^0} \mathrm{E}(Y)}{\vec{\mathbf{d}}^T \nabla_{\boldsymbol{\theta}^0} \mu}.$$

A practically meaningful direction is the steepest-ascent direction of the mean, $\vec{\mathbf{d}} = \nabla_{\boldsymbol{\theta}^0}\mu / \|\nabla_{\boldsymbol{\theta}^0}\mu\|$, which is a defensive (aggressive) choice assuming the goal is to identify the maximal sensitivity. Similarly, MSSD is defined as

$$\mathrm{MSSD}_{\vec{\mathbf{d}}} = \frac{\partial \mathrm{E}(Y)}{\partial \sigma_{\vec{\mathbf{d}}}} = \frac{\vec{\mathbf{d}}^T \nabla_{\boldsymbol{\theta}^0} \mathrm{E}(Y)}{\vec{\mathbf{d}}^T \nabla_{\boldsymbol{\theta}^0} \sigma}.$$

For sensitivity with respect to the standard deviation, meaningful directions are the steepest ascent direction along which σ increases the fastest: $\vec{\mathbf{d}} = \nabla_{\boldsymbol{\theta}^0}\sigma / \|\nabla_{\boldsymbol{\theta}^0}\sigma\|$; and the minimum-mean-change direction, which minimizes the rate of change in the mean of the input while increasing its standard deviation. The minimum-mean-change direction can be determined through solving an optimization problem similar to the one in Section 2.1 of Jiang et al. (2021) after replacing σ^2 with σ.

In the context of CTrES, the gradient of the mean or standard deviation of the inputs with respect to the input parameter, $\nabla_{\boldsymbol{\theta}^0}\mu$ or $\nabla_{\boldsymbol{\theta}^0}\sigma$, are known and the key is estimating $\nabla_{\boldsymbol{\theta}^0}\mathrm{E}(Y)$, the stochastic gradient. For stochastic gradient estimation in CTrES, we used the regression-based method of Wieland and Schmeiser (2006) as extended by Lin et al. (2015).

However, the framework in Jiang et al. (2019) and Jiang et al. (2021) is not sufficient for conducting local sensitivity analysis for all of the CTrES inputs. In Table 1, only sensitivity to Screening Failure fits perfectly within the previous work. *The technical contribution of this paper is to derive sensitivity measures for the others.* For instance, the sensitivity to inputs following a triangular distribution needs a problem-specific direction. The sensitivities to the remaining inputs require new methods. We describe these new challenges and our solutions to them in four subsections below, and then demonstrate their use in an illustrative case study.

4.1 Direction $\vec{\mathbf{d}}$ for Triangular Distribution

The challenge presented by the triangular distribution is that its support depends on the distribution parameters and that makes the gradient of the mean or standard deviation with respect to the input parameters hard to

interpret. In this case, the meaningful directions described above are not appropriate. This is an example of a problem-specific direction that we need to determine in the context of CTrES.

Denoting the parameters of a triangular distribution as (a,b,c), where a is the minimum, b is the mode, and c is the maximum, the mean and standard deviation of the distribution are given by

$$\mu = \frac{a+b+c}{3} \qquad \sigma = \sqrt{\frac{a^2+b^2+c^2-ab-ac-bc}{18}}.$$

For sensitivity with respect to the mean (i.e., MSM), the unit-norm steepest ascent direction of the mean, where the probability density function shifts to the right by $\sqrt{3}/3$ units (i.e., $\vec{\mathbf{d}} = (\sqrt{3}/3, \sqrt{3}/3, \sqrt{3}/3)^\top$), still makes sense for CTrES. Along this direction, the mean increases at the fastest rate while the standard deviation is kept constant. Thus, the effect of input-distribution location is isolated with minimal change to its spread. For this reason that we propose parameterizing the triangular distribution by the mean μ directly when estimating MSM. It can be shown that this reparameterization leads to a stochastic gradient estimator that is the same as the original parameterization using the method of Wieland and Schmeiser (2006), at least approximately.

For sensitivity with respect to the standard deviation (i.e., MSSD), we chose a meaningful direction to be the direction where the end points of the probability density function move in opposite directions by the same amount, i.e, $\vec{\mathbf{d}} = (-\sqrt{2}/2, 0, \sqrt{2}/2)^\top$. The triangular distribution has no unique min-mean-change direction because of having more than two parameters. However, this particular min-mean-change direction is practically meaningful for CTrES because the expert users who provide the parameters are often confident about the mode but not the support of the distribution. Thus, the sensitivity measure that tells users the impact of adjusting the minimum and the maximum of a triangular distribution without affecting the mean or mode is the most useful. It is for this reason that we propose parameterizing the triangular distribution by the minimum and maximum parameters when estimating MSSD.

4.2 Sensitivity with Respect to Piecewise-constant NSPP

The piecewise-constant NSPP in CTrES consists of two distinct arrival rates, λ_{high} and λ_{low}, over two intervals $[0, L_{\text{high}})$ and $[L_{\text{high}}, T)$, where L_{high} is the duration of the time when the arrival rate is high. The duration L_{high} has a triangular distribution, and T is the time necessary to enroll the required number of patients. Because this piecewise-constant NSPP has two intervals with uncertain length, it is particularly challenging to directly measure the sensitivity with respect to its mean or standard deviation.

As suggested in Morgan et al. (2016), each interval in a piecewise-constant NSPP can be regarded as a single input distribution to the simulation with the observation interval matching the simulation interval. Therefore, the sensitivity with respect to this NSPP can be decomposed into sensitivities with respect to two independent Poisson processes. We describe the Poisson process as interarrival times following exponential distribution so that the corresponding stochastic gradient can be estimated using the method of Wieland and Schmeiser (2006). Since $\mu = \sigma$ for exponential distribution, we only measure sensitivity to the mean of the interarrival time. With this formulation, the sensitivity falls within the framework of Jiang et al. (2021).

The stochastic input L_{high} is problematic because it affects the number of arrivals under the high and low rates. Therefore, we reformulated the sensitivity question to be "How sensitive are the KPIs to the *actual* duration of the time when the arrival rate is high?" To obtain this, we simply do a regression of the simulation output on the observed value of L_{high} of all sites and the sensitivities are the corresponding coefficients.

4.3 Sensitivity with respect to Bernoulli Inputs

For the inputs following a Bernoulli distribution, only sensitivity with respect to the mean of the input, i.e., $\mu \equiv \mathrm{E}(X) = p$, makes sense. Thus, MSM requires estimating the stochastic gradient $\partial \mathrm{E}(Y)/\partial p$. For

Screening Failure, the screening test results of at least 800 patients are recorded within each replication, so the stochastic gradient we need can be estimated by using the method of Wieland and Schmeiser (2006). For Startup Success and Enrollment Success, on the other hand, only a single outcome (0 or 1) is observed in each replication. Therefore, the method of Wieland and Schmeiser (2006) applies only through batching the observed input variates across replications. However, the method of batching sacrifices samples size for the subsequent linear regression to estimate $\nabla_{\boldsymbol{\theta}^0}\mathrm{E}(Y)$. This could be a serious problem for CTrES which often involves tens or hundreds of stochastic inputs.

We overcome this challenge as follows. Notice that when there is a single Bernoulli random variable, $X \sim \text{Bernoulli}(p)$, the $\partial \mathrm{E}(Y)/\partial p$ can be derived by conditioning, i.e.,

$$\mathrm{E}(Y) = \mathrm{E}(Y|X=1)p + \mathrm{E}(Y|X=0)(1-p) \quad \Rightarrow \frac{\partial \mathrm{E}(Y)}{\partial p} = \mathrm{E}(Y|X=1) - \mathrm{E}(Y|X=0). \tag{1}$$

Expression (1) can be estimated directly from the output data by

$$\hat{\eta} = \frac{\partial \widehat{\mathrm{E}}(Y)}{\partial p} = \frac{\sum_{i=1}^{n} Y_i \cdot \mathrm{I}\{X_i = 1\}}{\sum_{i=1}^{n} \mathrm{I}\{X_i = 1\}} - \frac{\sum_{i=1}^{n} Y_i \cdot \mathrm{I}\{X_i = 0\}}{\sum_{i=1}^{n} \mathrm{I}\{X_i = 0\}}.$$

4.4 Interacting Inputs

In the context of CTrES, there are inputs that interact with each other. For example, as shown in Figure 2, the impact of a site's startup delay and interarrival time only matter if that site starts up successfully and is able to enroll patients. Similarly, the site-specific inputs of a country have impact on the KPIs only when that country starts up successfully. With such interacting inputs, it is tricky to find an appropriate sensitivity measure.

Specifically, if the result of the country startup is failure, then no variates will be observed from uncertain inputs at the site level for all sites in the country. Similarly, if the startup or enrollment of a site fails, no variates will be observed from the inputs Patient Arrival Rate or Duration of High Rate for that site. One solution is measuring the sensitivities *conditional* on the successful startup of the country and all sites, and successful enrollment at all sites. However, this is a situation that rarely happens, so such a conditional sensitivity does not answer the what-if questions that help with plan management. What CTrES users want is an unconditional sensitivity measure.

Therefore, we propose a new random variable that considers the interaction among inputs. We demonstrate it for the case when $X \sim F(\cdot|\theta)$, $B \sim \text{Bernoulli}(p)$, and $\theta = \mathrm{E}(X)$. Because X is observable on each replication, the realized value of the parameter, $\widehat{\theta}$, can be obtained as a function of the observed X's. Define a new variable $\widehat{\theta}' = B\widehat{\theta}$ which has $\mathrm{E}(\widehat{\theta}') = \theta' = p\theta$ and B is the input that interacts with X. We can apply the method of Wieland and Schmeiser (2006) to estimate the stochastic gradient of $\mathrm{E}(Y)$ with respect to θ' using OLS by regressing Y_j on the observed new random variate $\widehat{\theta}'_j$, $j = 1, 2, \ldots, n$, i.e.,

$$Y = \beta_0 + \beta_1 \widehat{\theta}' + \varepsilon. \tag{2}$$

However, if we use the model in (2) we have $\nabla_{\theta^0}\mathrm{E}(Y) = p\beta_1$ where β_1 can be estimated via OLS. Thus, for unconditional sensitivity we use $p\widehat{\beta}_{1,\mathrm{OLS}}$ as the estimator of the gradient of $\mathrm{E}(Y)$ with respect to θ.

4.5 Dependence Because Total Enrolled Patients is Fixed

A CTrES simulation stops when a fixed number of patients, say 800, is enrolled. This forces a constraint on the number of patients recruited at each open site because they have to total to 800. Therefore, there is functional dependence among the observed arrival processes of open sites. The goal here is to decide how to parameterize the interarrival time such that the dependence works in our favor for local sensitivity analysis. For the purpose of explaining our solution, suppose each site has only as single arrival rate, instead of both high and low.

Table 2: Mean values of selected site-specific stochastic inputs.

Site	Startup Success	Startup Delay	Interarrival Time	High Duration
1	100%	15 weeks	2.03 days	17 weeks
2	95%	6 weeks	3.04 days	9 weeks
3	95%	15 weeks	2.34 days	13 weeks
4	100%	19 weeks	2.03 days	17 weeks
5	99%	25 weeks	1.52 days	22 weeks
6	95%	17 weeks	3.04 days	13 weeks
7	99%	17 weeks	1.52 days	13 weeks
8	99%	13 weeks	2.03 days	17 weeks
9	99%	26 weeks	1.22 days	26 weeks
10	99%	13 weeks	1.52 days	13 weeks

Let $\widehat{\theta}^{(i)}$ be the observed parameter of the exponential distribution of the interarrival time for site i. Suppose there are S sites where each is affected by its Startup Success $B^{(i)} \sim \text{Bernoulli}(p^{(i)})$. The regression model for estimating the gradient of $\text{E}(Y)$ with respect to $\theta^{(i)}$ at the nominal setting is given by $Y = \beta_0 + \sum_{i=1}^{S} \beta_i B^{(i)} \widehat{\theta}^{(i)} + \varepsilon$. Analysis of the model is straightforward if $B^{(i)}$ is independent of $B^{(j)}$ and $B^{(i)}$ is independent of $\widehat{\theta}^{(j)}$ for $i \neq j$, $\forall i, j$. However, the latter assumption does not hold because the simulation terminates when 800 patients are enrolled.

Specifically, when θ is the rate parameter λ, the arrival counting process of site i, $N^{(i)}(t)$, is $\text{Poisson}(\lambda^{(i)} t)$, and the time it takes to enroll 800 patients can be represented as:

$$T = \inf\left\{ t \geq 0 : \sum_{i=1}^{S} B^{(i)} N^{(i)}(t) = 800 \right\} \quad \Rightarrow \sum_{i=1}^{S} B^{(i)} N^{(i)}(T) = 800. \tag{3}$$

Therefore, the observed arrival rate of site i is $\widehat{\lambda}^{(i)} = B^{(i)} N^{(i)}(T)/T$ and Equation (3) is equivalent to $\sum_{i=1}^{S} \widehat{\lambda}^{(i)} = 800/T$, which shows that $\widehat{\lambda}^{(i)}$'s are not independent. If $\widehat{\lambda}^{(i)}$ is larger than expected, the observed rates of other sites must be smaller to compensate. Such dependence among predictors of the regression makes sense from a local sensitivity point of view. Thus, we propose parameterizing the interarrival time by the rate parameter and using $p^{(i)}\widehat{\beta}_i$ as the change in $\text{E}(Y)$ per unit increase in the observed rate at site i.

If, on the other hand, we let $\theta^{(i)}$ be the mean interarrival time $\mu^{(i)}$, then the observed mean interarrival time at site i is given by $\widehat{\mu}^{(i)} = T/\left(B^{(i)} N^{(i)}(T)\right)$ when $B^{(i)} = 1$, and is undefined otherwise. In this case, we no longer have the sum of $B^{(i)}\widehat{\mu}^{(i)}$ to be some constant and the relationship among the $\widehat{\mu}^{(i)}$'s depends on the observed $B^{(i)}$'s. Therefore, the resulting regression coefficients are hard to interpret.

5 AN ILLUSTRATIVE CASE: ONE COUNTRY WITH TEN SITES

In this section, we illustrate the results discovered via local sensitivity analysis on a CTrES simulation with one country and ten sites. This is a realistic case for a clinical trial in the U.S.; however, specific parameter values are chosen only for demonstration purposes. The country and all sites are subject to the uncertainties specified in Table 1 and there are 52 stochastic inputs with 104 parameters in total.

The country startup delay is assumed to last for an average of 14 days without any possibility of failure and without any identification delay for any of its sites. The country startup cost is assumed to be zero and there is at least one person that is available to initiate each site as soon as country activation is completed. The enrollment probability is 95% for each of the sites except Site 5 where the enrollment probability is 90%. Table 2 provides mean values of selected site-specific stochastic inputs to present insights into what may be expected prior to running CTrES and performing local sensitivity analysis. For each site, the mean interarrival time is provided in days for the duration of high enrollment that is given in weeks. We also

Table 3: Cost assumptions.

Site	Startup Cost	Screening Cost	Enrollment Cost
1	\$5,000	\$2,000	\$8,000
2	\$5,000	\$2,000	\$8,000
3	\$5,000	\$2,000	\$8,000
4	\$6,000	\$2,500	\$8,500
5	\$10,000	\$2,000	\$10,000
6	\$5,000	\$2,500	\$7,500
7	\$10,000	\$3,000	\$9,000
8	\$3,000	\$2,000	\$7,000
9	\$10,000	\$2,500	\$7,500
10	\$5,000	\$1,000	\$6,500

present the site-specific cost assumptions in Table 3. It can be inferred from Table 2 that Site 9 is the fastest enrolling site, followed by Site 5, Site 7, and Site 10. It can also be inferred from Table 3 that Site 8 and Site 10 are the least costly sites while Site 5 and Site 7 have the largest costs. Finally, the mean screening failure is assumed to be 15% for each of the ten sites.

The two primary KPIs are the mean time it takes to enroll 800 patients (denoted as "TimeToEnrollTarget") and the mean of the implied total cost (denoted as "TotalCost"). The simulation is run for 6000 replications and mean TimeToEnrollTarget and mean TotalCost are estimated as 61 weeks and 8 million dollars. Using the 6000 replications, we measure the sensitivity of each KPI to the mean and standard deviation of each stochastic input and screen out the inputs without statistically significant effects on the means of the KPIs. We find that the changes in the standard deviations of stochastic inputs have negligible impact on the mean KPIs in this particular illustration. Therefore, we only report the MSM measures. When we interpret the sensitivity measure, the change in the mean of an input is in its actual units, i.e., in weeks for Startup Delay and Duration of High Enrollment, in days for Patient Interarrival Time and in percentages for Startup Success, Enrollment Success, and Screening Failure. For ease of representation, we express the unit of cost in thousands of dollars (K).

First, we examine sensitivity with respect to the mean patient screening failure probability. We find that the mean TotalCost would increase by \$23K if the mean failure probability increased by 1%. A protocol design causing an average of 1% increase in the mean of this stochastic input would also take an average of six days longer to reach the 800-patient enrollment target. This sensitivity applied existing concepts from Jiang et al. (2019) and Jiang et al. (2021); the remainder require the new developments in this paper.

Consider increasing the mean duration of high enrollment by one week. We identified the mean TotalCost to increase by \$5K and \$4K, respectively, when the mean duration of high enrollment increased by one week at Site 5 and Site 7. We also identified a reduction of 1.19 days and 1.15 days in the mean TimeToEnrollTarget. When the mean duration of high enrollment increased by one week at Site 10, the mean TotalCost decreased by \$6K while the mean TimeToEnrollTarget decreased by 1.28 days. Notice that the pattern of effect on mean TotalCost is driven by the sites' attributes. Site 10 is one of the fastest enrolling sites as shown in Table 2. Site 10 also incurs the lowest patient screening and enrollment costs (Table 3). Furthermore, the mean startup delay at Site 10 is significantly shorter than those at Site 5 and Site 7. Therefore, if Site 10 were enrolling a larger number of patients for a longer duration, then the contribution of the two most costly Sites 5 and 7 to the overall patient enrollment target could be reduced. Consequently, the total cost would decrease. For this reason if there would be any investment into increasing the duration of high enrollment, the local sensitivity analysis recommends consideration of Site 10.

Next, we investigate the sensitivity of mean KPIs to an increase of one day in the mean patient interarrival time and report our findings in Table 4. Notice that Site 2 is not included in the table: Increasing the mean interarrival time of patients to Site 2 by one day has no statistically significant impact on mean TimeToEnrollTarget and mean TotalCost. Similarly, we identify such a change at Sites 1, 3, 6 and 9 to

Table 4: Impact of one-day increase in mean interarrival time on mean KPIs.

Site	1	3	4	5	6	7	8	9	10
TimeToEnrollTarget (week)	3.15	1.22	2.58	4.66	1.28	3.46	1.92	10.07	2.97
TotalCost ($)			-29K	-97K		-75K	38K		103K

Table 5: Impact of 1% increase in mean startup success and in mean enrollment success.

Site	5	7	8	10
Startup Success	$\$1.9K$	$\$2.1K$	$-\$1.0K$	$-\$3.3K$
Enrollment Success	$\$1.8K$	$\$2.4K$	$\$1.0K$	$-\$3.6K$

have no significant impact on mean TotalCost. However, an increase of one day in mean patient interarrival time at the fastest enrolling site (Site 9) is expected to increase the mean TimeToEnrolllTarget by 10 weeks. The sensitivity of mean TimeToEnrollTarget to an increase of one day in patient interarrival time is smaller but still significant at each of the remaining sites. In particular, it would be advisable to avoid any protocol design that would reduce the patient arrival rate to Site 10 because an increase of one day in mean interarrival time at this site would increase the mean TotalCost by $\$103K$. This would be due to the higher enrollment contribution from Sites 5 and 7 to reach the 800-patient target. We also observe the one-day increase in mean interarrival time at Sites 4, 5 and 7 increases mean TimeToEnrollTarget by approximately three weeks, five weeks and four weeks, respectively, while decreasing mean TotalCost by $\$29K$, $\$97K$, and $\$75K$. Therefore, the results of the local sensitivity analysis allow the user to decide whether any protocol design at these sites, despite reducing the mean TotalCost, would be worthwhile due to an increase of at least three weeks in mean TimeToEnrollTarget.

Next, we switch our focus to the stochastic inputs Site Startup Success and Site Enrollment Success. Table 5 shows that each of Sites 5, 7 and 10 continues to have a significant effect on mean TotalCost. For example, if startup success probability increased by 1% at Site 10, the mean TotalCost would decrease by $\$3K$. If enrollment success probability increased by 1% at the same site, the mean TotalCost would decrease by $\$4K$. Therefore, protocol redesign and incentives at Site 10 to increase startup and enrollment probabilities may be beneficial. This takeaway leads to a broader decision as to whether application of incentives to accelerate country activation may have any benefit. We find that an increase of one week in mean country startup time has no significant effect on mean TimeToEnrollTarget but on FirstTimeEnrolls only; it increases the mean time to first enrollment by almost a week (6.47 days).

6 CONCLUSION

SAS CTrES is a powerful tool for CROs and pharmaceutical companies for clinical trial enrollment planning because it is capable of capturing all the uncertainties throughout the process and quantifying the risk in the cost and enrollment prediction beyond traditional deterministic solutions. However, CTrES lacks the capability to quickly answer the what-if questions that are important for problem diagnosis and management of a clinical trial. We extend the framework in Jiang et al. (2019) and Jiang et al. (2021) to enable CTrES to conduct local sensitivity analysis to answer the what-if questions for any number of stochastic inputs without running addition simulations beyond the basic scenario. Instead of directly opening more sites to improve only the most important KPI—the time it takes to enroll a given target number of patients—the sensitivity measures suggest smart resource management and effort allocation strategies that are both time efficient and cost efficient.

ACKNOWLEDGMENTS

This work was partially supported by National Science Foundation Grant Number 1634982.

REFERENCES

Abbas, I., J. Rovira, and J. Casanovas. 2007. "Clinical Trial Optimization: Monte Carlo Simulation Markov Model for Planning Clinical Trials Recruitment". *Contemporary Clinical Trials* 28(3):220–231.

Anisimov, V. V. 2008. "Using Mixed Poisson Models in Patient Recruitment in Multicentre Clinical Trials". In *Proceedings of the World Congress on Engineering*, Volume 2, 1046–1049. London, U.K.

Anisimov, V. V. 2009. "Predictive Modelling of Recruitment and Drug Supply in Multicenter Clinical Trials". In *Proceedings of Joint Statistical Meeting*, 1248–1259. Washington, D.C.

Anisimov, V. V. 2016. "Predictive Hierarchic Modeling of Operational Characteristics in Clinical Trials". *Communications in Statistics-Simulation and Computation* 45(5):1477–1488.

Biller, B., A. Mokashi, I. Oliveira, S. Pathan, J. Yi, and J. Box. 2019. "Time Travel into the Future of Clinical Trial Enrollment Design". In *Proceedings of the SAS Global Forum 2019 Conference*. Cary, NC: SAS Institute Inc.

Box, J. 2018. "Simulating Enrollment Plans". In *Proceedings of the 2018 PhUSE EU Connect Paper TT08*. Cary, NC: SAS Institute Inc.

Cognizant 2015. "Patient Recruitment Forecast in Clinical Trials". https://www.cognizant.com/whitepapers/patients-recruitment-forecast-in-clinical-trials-codex1382.pdf. Accessed: 2020-05-18.

Goodnight, James 2007. "SAS". http:ceoroundtableoncancer.org/members/sas. Accessed: 2020-05-18.

Handelsman, D. 2012. "Applying Business Analytics to Optimize Clinical Research Operations". In *Proceedings of the 2012 SAS Global Forum*, 1–8. Cary, NC: SAS Institute.

Hughes, E., R. Pratt, and B. Biller. 2018. "Solving Business Problems with SAS Analytics and OPTMODEL". Technology workshop presented at INFORMS Analytics Conference, Baltimore, MD.

Jiang, W. X., B. L. Nelson, and L. J. Hong. 2019. "Estimating Sensitivity to Input Model Variance". In *Proceedings of the 2019 Winter Simulation Conference*, edited by N. Mustafee, K.-H. G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, 3705–3716. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Jiang, X., B. L. Nelson, and L. J. Hong. 2021. "Meaningful Sensitivities: A New Family of Simulation Sensitivity Measures". *IISE Transactions*:in press.

Kouvelis, P., J. Milner, and Z. Tian. 2017. "Clinical Trials for New Drug Development: Optimal Investment and Application". *Manufacturing & Service Operations Management* 19(3):437–452.

Lin, Y., E. Song, and B. L. Nelson. 2015. "Single-experiment Input Uncertainty". *Journal of Simulation* 9(3):249–259.

Mijoule, G., S. Savy, and N. Savy. 2012. "Models for Patients' Recruitment in Clinical Trials and Sensitivity Analysis". *Statistics in Medicine* 31(16):1655–1674.

Morgan, L. E., A. C. Titman, D. J. Worthington, and B. L. Nelson. 2016. "Input Uncertainty Quantification for Simulation Models with Piecewise-constant Non-stationary Poisson Arrival Processes". In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, 370–381. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Wieland, J. R., and B. W. Schmeiser. 2006. "Stochastic Gradient Estimation using a Single Design Point". In *Proceedings of the 2006 Winter Simulation Conference*, edited by L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, 390–397. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Zhao, H., E. Huang, R. Dou, and K. Wu. 2019. "A Multi-objective Production Planning Problem with the Consideration of Time and Cost in Clinical Trials". *Expert Systems with Applications* 124:25–38.

Zhao, H., K. Wu, and E. Huang. 2018. "Clinical Trial Supply Chain Design Based on the Pareto-optimal Trade-off Between Time and Cost". *IISE Transactions* 50(6):512–524.

AUTHOR BIOGRAPHIES

WENDY XI JIANG is an Operations Research Specialist at SAS Institute. Her e-mail address is Xi.Jiang@sas.com.

BAHAR BILLER is a member of SAS Advanced Analytics Center of Excellence. Her e-mail address is Bahar.Biller@sas.com.

JIM BOX is a Principal Data Scientist at SAS Institute. His e-mail address is Jim.Box@sas.com.

BARRY L. NELSON is the Walter P. Murphy Professor in the Department of Industrial Engineering and Management Sciences at Northwestern University. His e-mail and web addresses are nelsonb@northwestern.edu and www.iems.northwestern.edu/~nelsonb/