PEER REVIEWED

Gen Li, PhD, MBA | Robert Gray, MBA

Performance-Based Site Selection Reduces Costs and Shortens Enrollment Time

This article is a part of the authors' integrated effort to build a conceptual structure for managing study operations in clinical development by focusing on performance-based site selection.

Investigator site performance is one of the most important perennial challenges facing clinical trial execution teams across the industry. The inability to enroll the needed number of subjects in a targeted timeframe strains study budgets and resources, and prolongs cycle times.

Our analysis of 5,000 terminated clinical trials showed that sponsors attributed 30% or more of Phase III study terminations to enrollment difficulties. From a variety of perspectives, industry leaders and experts have invested tremendous time and energy trying to identify root causes and find solutions to this multibillion-dollar problem.

Although numerous vendors offer site selection services to clinical trial sponsors, little practical guidance exists in the public domain to select sites with better potential to enroll more patients. This article is a part of the authors' integrated effort to build a conceptual structure for managing study operations in clinical development[1–5] by focusing on performance-based site selection.

We are convinced the fundamental solution to better enrollment lies in better planning, or more specifically, in better planning information. With this in mind, we use a neuropathic pain clinical trial as an example to describe how performance-based site selection can be managed to improve operational deliverables, and how the same method can generate predictable patient enrollment outcomes for a planned clinical trial.

However, there are inherent uncertainties associated with clinical trial execution, as it is a part of scientific and medical exploration. We have no expectations of inventing a "magic bullet" to solve every research challenge; therefore the limitations to our approach are described as well.

What is Site Performance?

For a clinical trial that uses 200 sites to enroll 1,000 subjects, the resulting five subjects enrolled per site is often referred to as the study's "site enrollment performance." Though mathematically correct, this definition is of little practical value to people in clinical operations. There are numerous time-consuming tasks, such as contract negotiation and institutional review board (IRB) approval, that need to be completed before a site can be activated for enrollment. It would be impossible to have all 200 sites come online at the same time, since it is not possible for 200 sites to complete all the multiple-party-driven tasks at the same time. If the sites were activated in several batches months apart, those joining later would be at a disadvan-

tage, simply because they had less time to recruit their five subjects.[4]

The effect is the same in a slightly different trial recruitment scenario (sometimes called a "rescue mission"), in which too few sites are deployed initially and more sites have to be added later. Then there also is the less common case of a study team activating too many sites, where time and resources were used to activate the excess number of sites, resulting in poor collective performance of the sites and longer enrollment cycle time from "first-subject first-visit" (FSFV) to "last-subject first-visit" (LSFV).[3] Regardless of how any study unfolds, the mere number of subjects each site enrolls does not give a clear view of how successful those sites actually performed at enrollment.[2]

Site performance is a complicated concept, with many dimensions. In our definition, the performance of an investigative site demonstrates its ability to provide usable data to validate or invalidate defined medical assumptions in a planned clinical trial, in comparison to other sites in the same clinical trial. Enrollment performance is a leading indicator of overall site performance.

Two key concepts can help to delineate site performance in more meaningful ways:[5]

- "Site effectiveness index" (SEI), defined as the percentage of enrollment capacity, which means the capability of a site to enroll patients once activated, being calculated at all the deployed sites for the clinical trial, based on how quickly sites are activated and ready to enroll. SEI measures the success of the site activation process, which is independent of site performance, as being measured by the average site enrollment rate (see next).
- "Average site enrollment rate" (ASER), defined as the sites in a study and the average to enroll patients in a defined timeframe, such as the number of subjects per site per month. ASER measures collective site performance in a clinical trial, which is independent of site activation process, as being measured by SEI.

Successful patient recruitment in a clinical trial requires the selection of sites that are able to enroll the required number of subjects in the required timeframe, as well as the establishment of an effective site activation process. Failure on the part of the clinical trial operations staff to fulfill either of these two requirements is enough to make the study an operational debacle. Improvement in site activation and in enrollment-per-site-per-month can have the same quantitative impact on clinical trial deliverables; that is, a 10% increase in either SEI or ASER will reduce overall enrollment cycle time, from FSFV to LSFV, by the same 10%.[1]

A fundamental problem that underpins poor site selection is the simple lack of effective tools and information to help people select sites with solid potential to enroll the needed patient population.

Aside from improving the site activation process, which is always essential, clinical operations staff should tackle site performance challenges both by selecting sites with evidence-based top potential to perform, and by supporting moderately successful sites in ways that might help them perform better during the course of the study.[6] Unfortunately, the emphasis is too often placed on dealing with an enrollment crisis after it is in full swing, rather than on strategic planning that could have prevented the crisis in many cases. As a result, more money and resources are expended on attempting to improve site performance during the recruitment period—even for sites with little potential for improvement—instead of on selecting and supporting sites that are likely to succeed. A fundamental problem that underpins poor site selection is the simple lack of effective tools and information to help people select sites with solid potential to enroll the needed patient population.

Can Performance-Based Site Selection Work in the Real World?

Site selection is one of the most challenging tasks for clinical trial teams. Although many teams try very hard to use available data and information to help them pick the best sites, existing site selection methods remain in need of improvement if they are to consistently identify sites with greater potential to enroll patients. A better way to identify top candidate sites is to look at their historic performance by using relevant and well-structured data. An ideal database would collect details on operations and other parameters from several thousand multicenter, interventional, mostly industry-sponsored clinical trials that have already been completed. Sources of the data would be a combination of governments, not-for-profit organizations, commercial organizations, and pharmaceutical companies. Such a database would be updated daily and would mirror the industry's clinical development effort in disease areas, indications, and geographic distribution of sites globally. Examples of these parameters include milestones at site level, clinical trial level, and clinical project level; number of patients enrolled by site, by country, and by trial; inclusion and exclusion criteria, along with when and how inclusion and exclusion criteria are modified in the course of trial execution; and many, many more.

Although a few of the largest pharmaceutical companies hold significant

amounts of detailed site performance data, any single company likely still accounts for less than 20% of the data available for research activity pertaining to the disease indication in which it concentrates. Besides, rarely has any company been able to dominate a disease condition for very long, and with so many pharmaceutical companies moving into new therapeutic areas and new disease conditions as they try to find new markets, they often have little historical data from which to benefit.

At any rate, based on a large and still-growing database that includes more than 15 million points of site performance data in a single category relating to site activation, a series of integrated statistical models have been developed by the lead author's company to identify candidate investigative sites with high potential to outperform sites selected using an existing method. In the remainder of this article, we use the example of a neuropathic pain clinical trial to describe how a performance-based site selection can improve operational deliverables in clinical trials, and how potentially the same method can be used to help predict patient enrollment outcome for a planned clinical trial.

In this particular example, since all the sites were activated by the same business process, SEI, which measures the success of site activation process, is a constant. Performance in any single site would have an impact on ASER. The more high-performing sites there are in the pool of sites, the more improvement of ASER a sponsor will be able to achieve.

Table 1 shows performance in a recently completed neuropathic pain study for which the lead author's company assisted a project team in identifying site candidates. Briefly, raw data were extracted from the database. The raw data included several thousand investigative sites that participated in clinical trials related to neuropathic pain that had completed enrollment. Multiple dimensions of parameters (evaluators) measuring the historical performance of these sites in the trials in which they participated were fed into a series of statistical models, which transformed the parameters into a single composite performance score for each site. For its new study, the project team selected 15 sites from a list of those recommended via this modeling.

Overall, the 15 sites recommended enrolled 81 patients and completed 55; the 20 sites selected by an existing method enrolled 79 patients and completed 44. In this trial, the "existing method" involved the project team making the best effort possible to gather available information on its own in order to choose promising sites.

The enrollment rates were 5.4 patients per site for the sites being recommended, compared to 4.0 patients per site for sites selected by existing method (a 35% difference). The completion rates were 3.7 patients per site for the recommended sites, compared to 2.2 patients per site for those selected by existing method (a 68% difference).

Performance-based site selection not only reduced nonperforming sites from 45% (nine of 20) to 20% (three of 15), it also improved the desired deliverables significantly for enrolled patients. For example, there was a 15% difference between the two groups of sites that enrolled patients as measured by the number of patients who completed the trial (or by the patient dropout rate, as it is more often expressed).

In addition to enrollment numbers, we have always intuitively believed that higher rated sites would also deliver better quality study data—the ultimate desired deliverable in conducting any clinical trial. Interestingly, there was also a significant difference in data quality between the two groups of sites. As shown in the Figure 1, the data query rate, as measured by the number of complexity-adjusted queries per completed subject was 17% lower in the recommended sites (complexity-adjusted query is the total number of query transactions between sites and the project team as documented in the Query Detail Report, which approximately reflects time and resources used to solve outstanding queries during the course of trial execution).

When assessing the relationship between site performance scores (more details on performance score are explained later) and data query rates,

Table 1 Measures Differentiating the Results Between the Sites Selected Based on Their Past Performance and the Sites Selected Based on Existing Method*

	Number of Sites	Number of Sites with Enrolled Patients	Enrolled Number of Patients	Enrolled Patients Per Site	Completed Number of Patients	Completed Patients Per Site	Complexity-Adjusted Queries	Number of Queries Per Completed Patient
Performance-Based Site Selection	15	12	81	5.4	55	3.7	4,001	73
Existing Site Selection	20	11	79	4.0	44	2.2	3,747	85
Overall Clinical Trial	35	23	160	4.6	99	2.8	7,748	78
Percentage Difference				35%		68%		16%

*Since all sites were activated in a single batch, the time factor is the same for both performance-based site selection and existing site selection.

we were able to chart the result as shown in Figure 1. The correlation is clear: The higher the assigned performance score, the lower the data query rate.

Can Patient Enrollment Outcomes be Predicted?

To recap, as part of the performance-based site selection process, site performance scores were assigned to assist clinical trial teams in identifying sites. The performance score is a composite score derived from several variables (evaluators) related to the performance potential of the sites in patient enrollment. Examples of the evaluators include, but are not limited to, site enrollment duration, number of patients enrolled, patient data quality (as measured by query rate), clinical research coordinator experience, principal investigator experience, time to site activation, site experience in specific disease indication, site workload, etc.

Not all of these variables are independent. Enrollment drives many performance-related variables. The higher the performance score, the better the potential that the site will be a high-performing site in terms of patient enrollment. These scores have quantitative implications for the ability of the sites to enroll and retain subjects. Only the sites with scores of 4 (out of 5) or higher were recommended for selection by clinical trial teams.

The two trajectories in Figure 2 represent subject enrollment per site for two unrelated sets of sites. The blue line was drawn from a clinical trial that completed patient enrollment when we were at the planning stage for the clinical trial represented by the green line, which is the same neuropathic pain study we have been discussing in this article. Since these two clinical trials are similar in design, and for the same disease indication, we predicted the sites with similar performance scores that were selected for the planned study would yield similar enrollment results.

Figure 1 Negative Correlation Between Performance Score and Number of Complexity-Adjusted Queries per Completed Patient

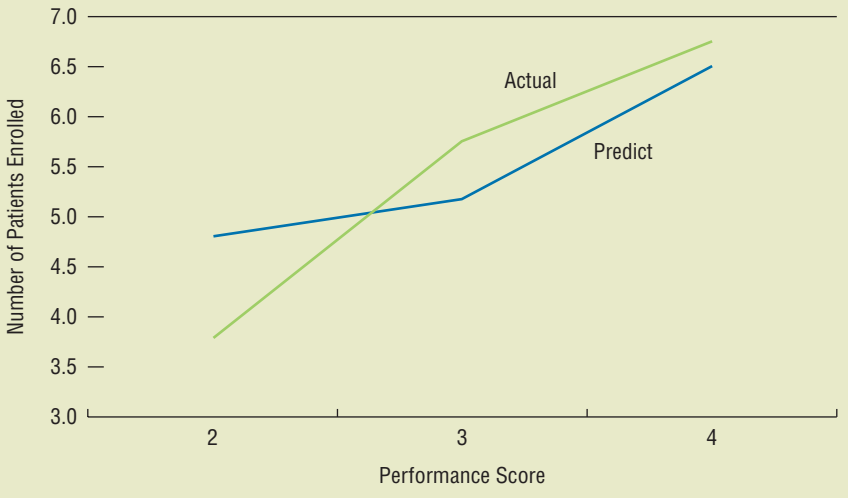
The similarity of the predicted and actual performance results illustrates how closely our assigned investigator performance scores predicted actual site performance. Therefore, it is possible to predict patient enrollment outcome for a planned clinical trial when the performance-based site selection method is being used.

There is No Magic Bullet

Although higher scores were assigned to sites with demonstrated high potential to do well in a planned clinical trial, sometimes lower scores for sites merely indicate that we do not have enough data to substantiate their performance potential. Given the dynamic nature of our business, inevitably some high-scoring sites will receive lower scores in the future, while some low-scoring sites will merit higher scores as more data emerge about their performance histories.

From the more than 30 clinical trials for which performance-based site selection scores were used in site identification, a close correlation between predicted and actual enrollment numbers has been established in more than 20 trials for which actual enrollment results were evaluated against the predictions. The rest of the clinical trials are actively recruiting patients (with one exception due to unsuccessful attempts to follow up with the client).

Figure 2 Correlation Between Site Performance Scores and Number of Patients Enrolled



In other words, although we may be less certain whether or not a single site may enroll more patients than other sites in the same trial, as affected by site personnel turnover and other reasons, we are much more confident about predicting the higher enrollment performance of one group of sites versus another group of sites, aided by extensive historic performance data and a robust modeling process.

Performance-based site selection is far from a perfect method; no perfect method exists. We cannot eliminate nonperforming sites, but by using the right tools, we can substantially reduce the number we select for our studies. We cannot guarantee that a site can retain all the patients it enrolls, but sites with high percentages of completing patients in the past probably know something about keeping subjects engaged to complete a study. We also cannot expect sites to provide flawless data, but it is reasonable to think that sites with great data quality histories may continue the trend.

We cannot eliminate nonperforming sites, but by using the right tools, we can substantially reduce the number we select for our studies.

Further, not all clinical trials have trouble enrolling patients in time; the popularity of influenza vaccine clinical trials points to this fact quite well. Meanwhile, some academic clinical trials appear to have "unlimited" time to enroll patients, and in such situations, performance-based site selection may not apply.

Conclusion

The number of sites deployed in a clinical trial is a key driver in financial and resource costs in clinical development programs. For a pharmaceutical company that activates 1,000 sites annually, using performance-based site selection yielding a 10% improvement in site enrollment would reduce by 100 the total number of sites needed (other factors being equal). Assuming that activating a site costs $30,000, the sponsor would save $3 million a year in site activation costs alone.

Alternately, that same sponsor might opt to continue using 1,000 sites, but select them based on historical enrollment performance. In that case, if enrollment time were reduced by 10% for a clinical trial with a planned enrollment period of 300 days, the savings would be 30 days. Multiplied across a drug development program, such savings would be substantial.[1]

Performance-based site selection has been able to provide sites that outperform other sites selected by existing methods by 20% to 40%, as measured by number of patients enrolled per site; so the 10% improvement discussed above is a conservative estimate.

To summarize, our industry can use past clinical trial performance of investigators and sites to predict future performance. There is an abundance of data available today, but the art and science of better site identification involves more than just access to those data. What sets apart the outcomes of the performance predictions are the comprehensiveness and objectivity with which, and accuracy of how, the past performance data are interpreted and applied.

Even so, relatively modest improvements in site performance can reduce costs, or shorten enrollment cycle times, or both. One does not need a massive database and the perfect method to achieve at least some improvement. Making the extra effort to learn more about sites before engaging them, and then applying the information objectively can be all it takes to realize marginal improvement. After all, a 1% improvement in average enrollment could mean $60 million in cost savings annually for our industry.

Acknowledgments

Beth Harper, chief clinical officer for CenterPhase Solutions, provided structural input and ongoing support throughout the writing process. Ellen E. Wilson provided comprehensive editorial support. Jim Cucinotta from Sanofi-Aventis also reviewed the article.

References

1. Li G. 2009. Finding the sweet spot. October 2, 2009. Available at: http://pharmexec.findpharma.com/pharmexec/R%26D/Finding-the-Sweet-Spot/ArticleStandard/Article/detail/631734.
2. Li G. 2008. Site activation: the key to more efficient clinical trials. December 12, 2008. Available at: http://pharmexec.findpharma.com/pharmexec/R%26D/Site-Activation-The-Key-to-More-Efficient-Clinical/ArticleStandard/Article/detail/571374.
3. Li G, Sirabella L. 2010. Planning the right number of investigative sites for a clinical trial. 24(4): 54–8.
4. Li G, Harper B, Sirabella L. 2010. Activating investigative sites in a staggered process: problems and solutions.
5. Li G. 2009. Site Effectiveness Index and Methods to Measure and Improve Operational Effectiveness in Clinical Trial Execution. U.S. patent publication No. 2010-0250273.
6. Eriksson L-O, Harper BD. 2002. Successful subject recruitment: one size does not fit all. November 1, 2002. Available at: http://appliedclinicaltrialsonline.findpharma.com/appliedclinicaltrials/Patient%2FSubject+Recruitment/Successful-Subject-Recruitment-One-Size-Does-Not-F/ArticleStandard/Article/detail/82492. **ACRP**

Gen Li, PhD, MBA, is the founder and president of PhESi, LLC, which provides services to clinical development organizations around the world in the areas of investigator site selection, site activation process support, clinical trial planning, and business process improvement support. Previously, he was head of productivity for Pfizer Development Operations, director of performance management for medical development at Pharmacia, and senior resource manager for Bristol-Myers Squibb. He can be reached at gen.li@phesi.com.

Robert Gray, MBA, is director of global enrollment strategies for the Global Operations Services COE for Johnson & Johnson Pharmaceutical Research and Development, LLC. He has more than 17 years of experience in strategic planning, management, and monitoring of globally executed clinical trials. Most recently, he developed and led a clinical trial services department that specializes in clinical trial intelligence, protocol feasibility, and patient recruitment and compliance solutions. He is one of the earliest clients to have utilized the database and the concepts to improve clinical trial deliverables noted in this article. He can be reached at rgray5@its.jnj.com.